

Fagos CrAssBcn: los nuevos virus crAssphage indicadores de contaminación fecal humana

Clara Gómez Gómez

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Fagos CrAssBcn

Los nuevos virus crAssphage indicadores de contaminación fecal humana

Clara Gómez Gómez / Julio 2025



Departamento de Genética, Microbiología y Estadística

Facultad de Biología

Fagos CrAssBcn: los nuevos virus crAssphage indicadores de contaminación fecal humana

Fags CrAssBcn: els nous virus crAssphage indicadors de contaminació fecal humana

Memoria presentada por CLARA GÓMEZ GÓMEZ para optar al grado de **Doctora por la Universitat de Barcelona**

Programa de Doctorado EEES: Biotecnología

Tesis realizada bajo la dirección de la Dra. Mª Teresa Muniesa Pérez y la Dra. Cristina García Aljaro en el Departamento de Genética, Microbiología y Estadística de la Facultad de Biología de la Universitat de Barcelona

La doctoranda,

Clara Gómez Gómez

Barcelona, 2025

Directora y tutora,

Mª Teresa Muniesa Pérez

Directora,

Cristina García Aljaro

A mi familia

Como dijo Michael J. Fox (más conocido en casa como Marty McFly)

“*La familia no es algo importante.*

¡Lo es todo!”

“Sé el cambio que quieres ver en el mundo”

Mahatma Gandhi

Tesis financiada con la ayuda para la contratación de investigadores predoctorales de la Universitat de Barcelona (PREDOCS-UB 2020)

Diseño de la portada: Juan Manuel Gálvez Castilla

AGRADECIMIENTOS

Pues nada, ya han pasado 4 años volando y aquí estoy, escribiendo los agradecimientos justo al final. Mi intención era ir escribiéndolos durante la redacción de la tesis, pero fui posponiéndolo porque para mí es la parte más difícil y, quizás, la más importante. Expresar con palabras el agradecimiento a todas las personas que han sido parte fundamental de este proceso no es tarea fácil. Es precisamente por todas estas personas que, si pudiera regresar al momento en que dudaba sobre si hacer la tesis o no (por lo difícil que fue conseguir la beca y siendo honesta, casi renuncio), me diría sin pensarlo que siguiera insistiendo hasta lograrlo, porque gracias a esa decisión he vivido experiencias maravillosas y he conocido a personas increíbles que han hecho que todo el esfuerzo mereciera la pena. En este apartado voy a intentar explicaros por qué, siendo consciente de que es la parte que todos vais a leer.

Primero de todo quiero agradecer a las dos personas que han hecho posible esta tesis. Mis dos directoras. **Maite**, des del primer dia que em vas fer classe de Seguretat Alimentària a quart de grau, vaig tenir clar que volia formar part del teu grup. Vaig intentar apropar-m'hi demanant unes pràctiques d'estiu, però em vas proposar directament fer el TFM amb vosaltres. No puc estar-te més agraïda, perquè aquell gest va marcar l'inici d'una etapa molt important per mi. Amb el temps he arribat a la conclusió que ets una persona realment admirable. Sempre has estat molt present: cada matí passant pel laboratori saludant, preguntant pel cap de setmana, les vacances o comentant l'experiment del dia anterior. Allò que sovint m'expliquen altres amics que estan fent la tesi, que el director només apareix un cop al mes i dona gràcies, aquí no passa. I t'asseguro que això es valora moltíssim. Gràcies pels teus consells, per la teva eficiència, per saber trobar la part positiva quan jo m'angoixava amb uns resultats que no acabaven de sortir, sobretot amb la petita-gran crisi que he tingut a una setmana de dipositar. Gràcies per tot el coneixement que m'has transmès, per confiar en mi des del primer moment i per acompanyar-me amb tanta dedicació al llarg d'aquests anys.

Cristina, gràcies per estar sempre disposada a resoldre qualsevol dubte amb claredat i bona predisposició. El teu esforç constant i la passió que poses en la recerca són una gran font d'inspiració, perquè treballar amb una persona tan implicada i dedicada anima a donar el millor d'un mateix. Et vull agrair especialment per ser un exemple i un suport constant, aportant sempre molt, tant en l'àmbit acadèmic com en l'humà. Un clar exemple d'això van ser els ànims que em vas donar durant el congrés d'Amersfoort, que em van donar molta força per superar un moment difícil.

También quiero agradecer a toda la gente de los despachos que me han dado su apoyo. Gràcies **Anicet** per sempre confiar en les meves capacitats i animar-me en els congressos. **Eli**, quan vaig veure la presentació del teu recorregut professional durant la defensa de la teva

plaça, em vaig quedar realment impressionada. Vaig sentir una profunda admiració per la teva passió per la ciència, i tant de bo algun dia pugui arribar a aquest punt. **Lorena**, gracias por haber estado presente en las distintas etapas de este recorrido que hizo posible esta tesis. Te agradezco tus buenos consejos, tu soporte constante a lo largo de estos años y la disposición con la que siempre has compartido tus conocimientos. Gracias también a **Susana,** por echar siempre una mano cuando lo necesitaba.

Quiero agradecer y felicitar especialmente a esas tres personas que han dado un paso importante en su trayectoria profesional. Un logro más que merecido, fruto tanto de su valía personal como de su profesionalidad. **Loles,** gracias por tu alegría contagiosa y ese salero andaluz con el que siempre vienes a saludarnos. Eres una grandísima profesional, y tu ayuda ha hecho que todo sea mucho más llevadero. No sé qué haríamos sin ti. **Miriam**, gràcies per estar sempre disposada a ajudar quan ho necessito, ja sigui en temes de laboratori o burocràtics. Has estat una clienta fidel de la nostra petita empresa familiar, i aquests trofeus de fusta que demanes cada any signifiquen molt per mi. També vull agrair a la **Laura**. Sento com si fossis la nostra germana gran, sempre disposada a donar un cop de mà en tot i a cuidar-me en els congressos. Realment ets una persona molt propera i atenta, sempre vens a veure'ns al laboratori ja sigui per saludar, per feina, per xerrar i explicar-nos les teves històries o bé per menjar-te una taronja en companyia. Ara, comences com a professora a la universitat i de debò crec que seràs una directora de tesi meravellosa i estic molt contenta de que siguis part del **laboratori 10**. Un lugar donde no solo he pasado largas horas de trabajo, sino que también he compartido risas y momentos inolvidables con las personas que forman parte de él. Muchos se preguntarán, ¿es posible ser feliz en el trabajo? A lo que yo digo: sí, siempre y cuando tengas una familia tan especial como la que hemos formado en este laboratorio. Para los que no estéis familiarizados con esto, en el 10 siempre hemos bromeado· con que somos como una familia: un matrimonio, sus dos hijas y el abuelo. Aunque entre broma y broma, al final, realmente os siento como tal. Así que quiero contaros esta historia y mi paso por este lugar tan especial, que nunca voy a olvidar.

Llegué a este lugar en 2017 para hacer el Trabajo de Fin de Máster. Siempre he sido una persona bastante tímida, a la que le cuesta abrirse en entornos nuevos, por lo que llegar a un sitio desconocido siempre me resulta abrumador. Sin embargo, ahí estabas, **Dani**, la persona que desde el primer momento me hizo sentir cómoda y con la que pude ser yo misma. Gracias por tanto. Nunca habría imaginado que, después de nuestras competiciones de hipos y estornudos, las charlas frikis, el interés por Japón y tantas otras cosas, acabaríamos formando ese “matrimonio” con dos hijas a las que cuidar. Con el tiempo me di cuenta de que te habías convertido en el corazón del laboratorio, y aunque ahora estés trabajando en otro sitio, tu presencia sigue notándose cada día en tu silla. Todos sentimos muchísimo tu marcha, la

depresión en el lab 10 fue real… PD: ¡Cuanta falta me haces para comentar semanalmente One Piece! Ahora que estoy al día y la cosa está tan emocionante, cada episodio me deja totalmente impactada…

A punto de terminar el TFM, conocí a mi querido amigo, sensei, comodín y viejito prematuro: **Pedro** (alias el abuelo). Al principio no interactuamos mucho, pero cuando volví en 2020, ahí estabas y pudimos conocernos mejor. Gracias por todas las risas, los consejos, y los momentos compartidos durante este camino. Me sigue preocupando tu salud muscular y articular. Ya lo dices tú que a partir de los 25 todo va cuesta abajo y sin freno, y eso queda retratado con total fidelidad en tu video de tesis (el mejor vídeo te diré, que bien me lo pasé haciéndolo). Eres una de esas personas de las que todos deberíamos aprender, ya sea por tu sabiduría (gracias por enseñarme el noble arte del "robatori"), por tu forma de ser, tan generosa y auténtica, o por ese corazón de romántico empedernido y de padre increíble.

Estuvimos una época, Dani, Pedro y yo solos. La verdad que esos días los recuerdo con mucho cariño, íbamos los tres juntos a todas partes. Pero entonces llegó la chispa. **Sara.** ¡Ay niños, no saben lo que me pasó! Encontré a esa compañera o, mejor dicho, amiga (que con el tiempo llegó a ser como nuestra hija mayor, de quien debíamos cuidar para que no perdiera ni olvidara nada), con la que compartir los 4 años de tesis. Antes de que llegaras, me comentaron que eras unos años mayor que yo y que tenías un hijo, y pensé "quizás no tenga demasiado tiempo de hacer vida fuera del lab y no acabemos de conocernos del todo". ¡Pero que equivocada estaba! Desde la primera semana ya empezamos con nuestras tonterías, a quedarnos sin aire de la risa, a probar clases de todo tipo (Krav Maga, gimnasio, piscina) y la de anécdotas que tenemos. Te deseo mucha suerte porque eres la próxima en pasar por esto, y recuerda que siempre estaré orgullosa de mi marrano.

Y luego llegó la pequeña, mi querida **Gloria**. Esa personita que amolda su estado de ánimo al tiempo meteorológico. Contigo, nuestro vocabulario mejoró con todas esas palabras que nadie usa excepto tu y el laboratorio 10 se convirtió en un consultorio ya fuera para tratar nuestros dramas del día a día o bien para inventarnos cotilleos para mantenernos despiertas las tardes de trabajo. Entre tus frecuentes heridas en el trabajo, que me toca curar, tus despistes y tu cara de cachorrito cuando me pides probar mi tapper, realmente me he sentido como una madre, así ya tengo experiencia para el futuro. Eres una persona muy positiva (aunque no siempre lo parezca), y gracias a todo lo que me cuentas me has hecho ver que no todo es negro, y que, aunque sea poquito, cada uno puede contribuir a mejorar las cosas. Esa idea fue la inspiración para la frase dedicatoria de mi tesis.

Ahora, también quiero agradecer a todas las personas del laboratorio 9, nuestros vecinos. Empezaré por los dos pilares de este laboratorio: Raquel y Julia. **Raquel,** cuantas veces me

habré levantado de la silla para ir corriendo al 9 y enterarme porque te ríes o chillas emocionada. Eres una bellísima persona que personifica la alegría y la energía. Agradezco mucho el tiempo que hemos pasado charlando de animalitos, cosas frikis o de los sucesos raros que te ocurren, sobre todo en el metro. Y que sepas que me has generado una obsesión por los loros con tantos videos que me pasas. **Julia**, no sé si te lo había dicho nunca pero tu risa se me contagia mucho, cuando te escucho reír de lejos me rio yo sola. Gracias por haber sido mi gran compañera de prácticas y por tener un corazón tan grande. Recuerdo cuando me dijiste que Japón no te interesaba, y luego fuiste y volviste fascinada. Me alegró mucho verte descubrir algo tan maravilloso, que ahora podemos compartir. Luego están las incorporaciones más recientes. **Helga,** tengo que decir que me encanta tu estilo de vestir; siempre pienso “que mona va”. Eres una persona muy agradable y des del primer momento fue fácil conectar contigo. Te incorporaste al grupo, como yo, mientras hacías el máster y ya eres una más de esta familia. Algo curioso es que ahora cada vez que vea sacarina, no podré evitar pensar en ti. **Gabi**, a veces me acuerdo de que quizás tu foto está colgada en el salón de alguna familia del sudeste asiático y me meo de la risa. No coincidimos cuando hiciste el máster, pero me hablaron muy bien de ti y ahora, he podido comprobar por mí misma que eres una persona maravillosa. **Laura** (italiana), me encantó volver a verte por aquí, fue una gran sorpresa. Me alegro muchísimo de que hayas podido encontrar la manera de quedarte con nosotros porque eres un gran fichaje, una gran persona que le va a dar una nueva vida al laboratorio. Por cierto, me fascina lo rápido que has aprendido a hablar español y lo bien que lo haces. **Irene,** desde el primer momento fuiste una persona muy cercana, y siempre me alegra el día cuando te asomas al 10 para saludarnos. Incluso antes de que formaras parte del 9, cuando nos cruzábamos por los pasillos pensaba “que maja parece esta chica”, y sin duda, mi intuición no me falló.

En cuanto a gente de otros laboratorios, quiero agradecer también a **Andrés**. He sido muy afortunada de haber conocido a alguien tan alegre, abierto y que le pone ese toque colombiano que siempre logra alegrarme el día. Desde el primer momento me llamaste “Dra. Clara/Gómez” lo cual me hacía mucha gracia entonces, y aún más ahora que ese título está por hacerse realidad, tanto para mí como para ti. El día que tu presentes estaré en primera fila orgullosa de mi amigo. Espero algún día ir a Colombia y que seas mi guía por tu isla, que ya tiene un lugar en mi lista de viajes pendientes, de las maravillas que me cuentas de allá.

También quiero agradecer a **Maite Polo**, quien me ha facilitado muchísimo las cosas a lo largo de este camino. Aquellas clonaciones habrían sido imposibles sin tu ayuda, y qué decir de lo pesada que he sido con los pedidos. Me has salvado en más de una ocasión, siempre con una sonrisa y una disposición admirable. He disfrutado mucho trabajar contigo, y sé que, cada

vez que escuche una caja arrastrándose por el suelo, no podré evitar recordar que ese es el sonido que anuncia tu llegada al laboratorio 10.

También quiero agradecer al resto de personas estupendas del departamento. **Sandra**, gràcies per compartir la passió pel món coreà. **Ignasi** (el millor taxista en bici) gràcies pels "bon dia" amb un somriure que regales cada matí i que em fan millorar el dia al moment. **Cristina,** molts ànims amb el que queda de tesi. **Ayalke,** gracias por tu simpatía. I també al **David** i a l **'Albert**, gràcies per tenir sempre mostres disponibles.

No me olvido tampoco de la gente que ya no está en el departamento pero que han formado parte de este camino. Primero quiero agradecer a alguien que solo formó parte del laboratorio 10 durante un año. **Geno**, esta vez puedo dedicarte unas palabras con mucho más sentido que aquella primera dedicatoria que te escribí sin conocerte apenas y pensando que te caía mal. Después de tantas risas, anécdotas, bailes de TikTok, el viaje a Valencia y tantos momentos compartidos, puedo decir con toda certeza que te considero una gran amiga. Aunque ahora estemos lejos y no coincidamos tanto como antes, estoy convencida de que cuando volvamos a vernos será como si no hubiera pasado el tiempo, y creo que eso es precisamente lo que define una verdadera amistad. **Ari,** també et vull agrair a tu per haver format part d'aquesta etapa. Has sigut una integrant molt important del grup, sent la cosina de la família i amb tu hem compartit moltíssims bons moments i molts riures. La veritat és que se't troba molt a faltar al laboratori. **Anna** gràcies per la teva companyia als diversos congressos i per ser la heroïna apaga focs del departament. Em fa infinita il·lusió que formis part del tribunal de la meva tesi. **Robert**, et queda camí per aprendre d'en Dani a donar ensurts. **Aiora**, el laboratori sense tu no és el mateix, però al menys em conformo en veure't i escoltar el teu riure pels passadissos o durant els dinars. También agradecer a **Maryury, Pablo**, **Elena, Paula, Hong Xia, Marta, Eva**, a todos los TFG/TFM que han pasado por el laboratorio: **Pablo** (em faria molt feliç tornar-te a veure pel laboratori)**, Andreu, Laura**, **Vero, Sergio** y todos aquellos a los que no he mencionado por culpa de mi memoria.

También dedicaré una línea a **Don Geraldo**, esa locura que comenzó con un juguete que recogió Gloria en una cabalgata de reyes y acabo siendo el dios de nuestro altar, generando procesiones, ligas de música o pequeños infartos por su horrible aspecto. La verdad que esta figura nos unió mucho más a todos. Agradecer en general al **club de lectura** por animarme a volver a leer, algo que había dejado abandonado y que he retomado con mucha motivación, lo cual me permite aprender algo nuevo cada día. También al **club de las series**, que actualmente está bastante olvidado y que creo que deberíamos recuperar, porque siempre nos echamos unas buenas risas. Desde Seven seeds (no tiene desperdicio), Sharknado (aún me pregunto cómo puede haber 6 películas), series que al terminarlas no sabemos muy bien

que acabamos de ver (The Leftovers), Attack on Titan (orgullosa de introducir a gente “ignorante” al mundo Ochinokokyu) hasta Severance (vaya ida de olla de serie). Y, ¿cómo olvidarme de la **liga de música?** Eso que sacó nuestra máxima competitividad, pero que también me permitió descubrir un montón de canciones nuevas, y que cada vez que las escucho, me trae a la menta a quien las compartió. Y también agradecer a las **tardes de juegos** y **de roll plays**, que han sido pocas, pero muy divertidas. Intento recordar ese roll play que hicimos tan macabro y me vienen a la mente bebés vegetativos, Max chuches, Maria Eugenias, traiciones y aventuras románticas por doquier.

Ahora le toca el turno a toda esa gente de fuera del ámbito laboral, pero que también han sido personas que me han apoyado y animado en este camino, siendo parte de mi vida. A mis amigas de la universidad, **Judit, Elena, Jessica, Luce, Amparo y Carles**. A mi queridísima amiga del Máster, **Nathy**. A mis amigos del cole, **Nati, Claudia, Alberto, Joan** y **Martos.** Y también a mis amigos de Japón, **Yuko, Jeff, Shuya** y **Misaki**, la estancia allí no habría sido lo mismo sin vosotros. 日本のこんな素敵な国で、温かく迎えてくれて本当にありがとう。またすぐに会えるのを楽しみにしてるね。

Y por último a mi familia de sangre. **A mis padres** (**Natalia y Roberto**), creo que ni ahora ni en otras vidas os podría agradecer todo lo que habéis hecho por mí. A menudo recordaré la frase que más he oído en relación con mi tesis: “¿y el Nobel para cuándo?”. Que inocentes sois, aunque agradezco esa gran confianza en mí y en mi trabajo. Siempre habéis estado dispuestos a escuchar mis presentaciones, incluso cuando me mirabais con esos ojitos de “no estamos entendiendo ni papa”, pero orgullosos de contar a todos que vuestra hija habla de cosas muy difíciles. Gracias por haber hecho posible todo esto. Por supuesto tampoco pueden faltar el agradecimiento a mis hermanas, **Maria** y **Elena**, mi cuñado **Albert** y a mi adorable sobrinita **Mila,** a quien tiene mucho por delante y apoyaré del mismo modo que han hecho mis tíos y primos, **Yulia** y **Kolia** (hice la tesis en castellano para que la pudierais entender), **Fernando, Dolores, Lola**, **Leo, Polina** y **Esteban**. Gracias a todos por estar siempre presentes. También agradecer a mi abuelita de Rusia. **Бабуля**, хоть нас всегда разделяло расстояние, твоя любовь и твой голос в каждом звонке всегда помогали мне чувствовать тебя рядом. Спасибо, что всегда меня поддерживала. Мне бы очень хотелось, чтобы ты могла присутствовать на защите, но то, что ты приехала отпраздновать мои 30 лет, уже многое для меня значит. Когда я наконец закончу диссертацию, смогу спокойно “постоять на спине”, как ты говоришь. Y por supuesto también agradecer a mis abuelos que ya no están, **Bibi** y **Dina**, a quienes recuerdo con muchísimo cariño ahora y siempre. Así como también a mi preciosa gatita **Eve**.

Y, por último, quiero darte las gracias a ti, **Uri**. Gracias por estar siempre cuando te necesito, por acompañarme, por entenderme incluso en los días en los que ni yo misma me entendía. Gracias por animarme (especialmente con esos stickers tuyos llenos de adjetivos positivos), por tu infinita paciencia y por vivir esta tesis conmigo como si fuera tuya. Sé que podrías dar tú mismo esta presentación de tantas veces que te la voy a repetir… y, aun así, sin una sola queja. No hay mejor persona con la que podría compartir todo esto y mi vida entera. Tengo muchísima suerte de haberte conocido.

Sin darme cuenta he llegado hasta aquí, y sé que no lo habría conseguido sin cada uno de vosotros. Gracias de corazón.

Julio 2025

RESUMEN

El estudio de la microbiota intestinal es esencial para comprender su impacto en la salud humana. Este ecosistema dinámico se estructura mediante complejas interacciones entre bacterias, bacteriófagos y otros microorganismos presentes en la microbiota. Los procesos evolutivos de las diferentes poblaciones y la acción de los fagos sobre las poblaciones bacterianas regulan su estabilidad y diversidad. En este contexto, el descubrimiento de crAssphage, posiblemente el fago más abundante en el intestino humano, ha revolucionado el conocimiento del viroma intestinal. El aislamiento de crAssphage mediante técnicas convencionales presenta desafíos significativos relacionados con su dificultad de propagación *in vitro* y la escasa visibilidad de las calvas de lisis que producen. A pesar de ello, disponer de crAssphage aislados es esencial para profundizar en su conocimiento.

En la presente tesis doctoral, se ha logrado aislar 25 nuevos fagos tipo crAss en aguas residuales de Cataluña. Los fagos crAss aislados (crAssBcn) presentan una notable diversidad genética, posiblemente impulsada por la presión selectiva y la interacción con los sistemas de defensa de su hospedador bacteriano, *Bacteroides intestinalis*. La disponibilidad de viriones aislados ha permitido realizar ensayos experimentales más precisos sobre su ciclo de vida, superando así las limitaciones inherentes a los estudios realizados sólo con datos metagenómicos. Se observó la coexistencia de estos fagos virulentos con su hospedador bacteriano, cuyas poblaciones presentan fenotipos sensibles y resistentes a la infección. Los resultados sugieren que dicha resistencia, independiente del receptor del fago, está vinculada a una región genética en el hospedador *B. intestinalis* que está sujeta a variación de fase, la cual modula diferencialmente la expresión de polisacáridos capsulares (CPS) que, según se expresen o no, permiten la infección de los fagos.

Los fagos del tipo crAss destacan como potenciales marcadores universales de contaminación fecal humana debido a su abundancia en heces, su ubicuidad y su resistencia a los tratamientos de desinfección. Sin embargo, buscando la selección de un candidato universal representativo de este grupo tan abundante y heterogéneo, los análisis de genomas completos, previamente anotados en bases de datos (tanto aislados como no), revelan la necesidad de considerar fagos tipo crAss específicos para ser aplicados en cada región geográfica. La ausencia de una relación filogenética entre ellos, su posible presencia en heces de origen animal y la inexistencia de un candidato con distribución global uniforme refuerzan la hipótesis de considerar variantes locales para su aplicación. Siguiendo esta premisa, se llevaron a cabo estudios experimentales más exhaustivos sobre la persistencia de los viriones crAssBcn autóctonos, siendo el primer estudio que ha evaluado la inactivación natural de

fagos tipo crAss aislados. Los resultados confirman su mayor persistencia en comparación con otros indicadores microbiológicos e, incluso, con ciertos virus entéricos.

En conjunto, los hallazgos obtenidos en esta tesis subrayan la necesidad de profundizar en la comprensión del papel ecológico de estos virus, con el propósito de desarrollar nuevas estrategias para el control ambiental, contribuyendo a mejorar la gestión de la calidad del agua, al tiempo que permitirá explorar nuevas aproximaciones diagnósticas y terapéuticas para abordar desequilibrios en la microbiota intestinal.

Palabras clave: crAssphage, fagos tipo crAss, Crassvirales, bacteriófagos, *Bacteroides,* microbioma intestinal, polisacáridos capsulares, región con variación de fase, desinfección de aguas, microbiota fecal, *Microbial Source Tracking*, persistencia, inactivación natural, indicador de contaminación fecal.

RESUM

L'estudi de la microbiota intestinal és essencial per comprendre el seu impacte en la salut humana. Aquest ecosistema dinàmic s'estructura mitjançant complexes interaccions entre bacteris, bacteriòfags i altres microorganismes presents a la microbiota. Els processos evolutius de les diferents poblacions i l'acció dels fags sobre les poblacions bacterianes regulen la seva estabilitat i diversitat. En aquest context, el descobriment del crAssphage, possiblement el fag més abundant de l'intestí humà, ha revolucionat el coneixement del viroma intestinal. L'aïllament de crAssphage mitjançant tècniques convencionals presenta desafiaments significatius relacionats amb la dificultat de la seva propagació *in vitro* i la poca visibilitat de les calbes de lisi que produeixen. Tot i això, disposar de crAssphage aïllats és essencial per aprofundir en el seu coneixement.

En la present tesi doctoral s'han aconseguit aïllar fins a 25 nous fags tipus crAss d'aigües residuals de Catalunya. Els fags crAss aïllats (crAssBcn) presenten una notable diversitat genètica, possiblement impulsada per la pressió selectiva i la interacció amb els sistemes de defensa del seu hoste bacterià, *Bacteroides intestinalis*. La disponibilitat de virions aïllats ha permès realitzar assaigs experimentals més precisos sobre el seu cicle de vida, superant així les limitacions inherents als estudis realitzats només amb dades metagenòmiques. S'ha observat la coexistència d'aquests fags virulents amb el seu hoste bacterià, les poblacions del qual presenten fenotips sensibles i resistents a la infecció. Els resultats suggereixen que aquesta resistència, independent del receptor del fag, està vinculada a una regió genètica de l'hoste *B. intestinalis* sotmesa a variació de fase, la qual modula de manera diferent l'expressió de polisacàrids capsulars (CPS), que, segons s'expressin o no, permeten la infecció dels fags.

Els fags de tipus crAss destaquen com a potencials marcadors universals de contaminació fecal humana degut a la seva abundància en femtes, la seva ubiqüitat i la seva resistència als tractaments de desinfecció. No obstant això, cercant un candidat representatiu universal d'aquest grup tan abundant i heterogeni, els anàlisis de genomes complets, prèviament anotats en bases de dades (tant aïllats com no), revelen la necessitat de considerar fags tipus crAss específics per a ser aplicats a cada regió geogràfica. L'absència d'una relació filogenètica entre ells, la seva possible presència en femtes d'origen animal i la inexistència d'un candidat amb distribució global uniforme reforcen la hipòtesi de considerar variants locals per a la seva aplicació. Seguint aquesta premissa, es van dur a terme estudis experimentals més exhaustius sobre la persistència del virions crAssBcn autòctons, essent el primer estudi que avalua la inactivació natural de fags tipus crAss aïllats. Els resultats confirmen la seva major persistència en comparació amb altres indicadors microbiològics i, fins i tot, amb certs virus entèrics.

En conjunt, els resultats obtinguts en aquesta tesi subratllen la necessitat d'aprofundir en la comprensió del paper ecològic d'aquests virus, amb el propòsit de desenvolupar noves estratègies per al control ambiental, contribuint a millorar la gestió de la qualitat de l'aigua, alhora que permetrà explorar noves aproximacions diagnòstiques i terapèutiques per tractar desequilibris en la microbiota intestinal.

Paraules clau: crAssphage, fags tipus crAss, Crassvirales, bacteriòfags, *Bacteroides*, microbioma intestinal, polisacàrids capsulars, regió amb variació de fase, desinfecció d'aigües, microbiota fecal, *Microbial Source Tracking*, persistència, inactivació natural, indicador de contaminació fecal.

ABSTRACT

The study of the gut microbiota is essential for understanding its impact on human health. This dynamic ecosystem is structured through complex interactions among bacteria, bacteriophages, and other microorganisms present in the microbiota. Evolutionary processes among the different populations and the action of phages on bacterial populations regulate their stability and diversity. In this context, the discovery of crAssphage, possibly the most abundant phage in the human gut, has revolutionized our understanding of the gut virome. Isolating crAssphage using conventional techniques presents significant challenges related to its difficulty in propagating *in vitro* and the poor visibility of the lysis plaques it produces. Despite this, having isolated crAssphage is essential to further our understanding.

In this doctoral thesis, 25 new crAss-like phages were isolated from wastewater in Catalonia. The isolated crAss phages (crAssBcn) exhibit notable genetic diversity, likely driven by selective pressure and interactions with the defense systems of their bacterial host, *Bacteroides intestinalis*. The availability of isolated virions has allowed for more precise experimental assays on their life cycle, thus overcoming the limitations inherent to studies based solely on metagenomic data. The coexistence of these virulent phages with their bacterial host was observed, with bacterial populations displaying both sensitive and resistant phenotypes to infection. The results suggest that this resistance, independent of the phage receptor, is linked to a genetic region in the host *B. intestinalis* subject to phase variation, which differentially modulates the expression of capsular polysaccharides (CPS), enabling or preventing phage infection depending on their expression.

These crAss-like phages stand out as potential universal markers of human fecal contamination due to their abundance in feces, ubiquity, and resistance to disinfection treatments. However, in the search for a universal representative candidate of this abundant and heterogeneous group, analyses of complete genomes, both annotated in databases (isolated and non-isolated), reveal the need to consider region-specific crAss-like phages for application in each geographic area. The absence of a phylogenetic relationship among them, their possible presence in animal feces, and the lack of a globally distributed uniform candidate reinforce the hypothesis of considering local variants for practical use. Following this premise, more exhaustive experimental studies were conducted on the persistence of the native crAssBcn phages, making this the first study to evaluate the natural inactivation of isolated crAss-like phages. The results confirm their greater persistence compared to other microbiological indicators and even certain enteric viruses.

Altogether, the findings of this thesis highlight the need to deepen our understanding of the ecological role of these viruses, with the aim of developing new strategies for environmental

control, contributing to improved water quality management, while also enabling the exploration of new diagnostic and therapeutic approaches to address imbalances in the gut microbiota.

Keywords: crAssphage, crAss-like phages, Crassvirales, bacteriophages, *Bacteroides*, gut microbiome, capsular polysaccharides, phase-variable region, water disinfection, fecal microbiota, Microbial Source Tracking, persistence, natural inactivation, fecal contamination indicator.

ÍNDICE

AGRADECIMIENTOS ..... **i**

RESUMEN ..... **ix**

RESUM ..... **xi**

ABSTRACT ..... **xiii**

ÍNDICE ..... **xv**

ABREVIATURAS **1**

1. INTRODUCCIÓN **5**

1.1 Microbiota intestinal **7**

1.1.1. Bacterias intestinales: un componente fundamental de la microbiota .. 9

1.1.2. Viroma intestinal 13

1.2. Bacteriófagos **14**

1.2.1. Ciclos de replicación de los bacteriófagos 16

1.2.2. Dinámica de control de poblaciones bacterianas mediada por fagos . 18

1.3. CrAssphage **25**

1.3.1. Taxonomía 25

1.3.2. Hospedador bacteriano 26

1.3.3. Características genéticas 27

1.3.4. Fagos tipo crAss aislados 28

1.3.5. Prevalencia de crAssphage en el microbioma intestinal humano 29

1.3.6. Posibles usos de los fagos tipo crAss 33

1.4. La contaminación fecal del agua **35**

1.4.1. Indicadores de contaminación fecal 39

1.4.2. Bacteriófagos como microorganismos indicadores 41

1.4.3. Microbial Source Tracking 43

1.4.4. Fagos tipo crAss como indicadores de contaminación fecal humana . 46

2. OBJETIVOS **49**

3. INFORMES **53**

3.1. Informe sobre el factor de impacto **55**

3.2. Informe de coautoría y de participación en las publicaciones **56**

4. RESULTADOS **59**

4.1. Capítulo 1: 61

La caracterización de los fagos tipo crAss aislados destaca la heterogeneidad genética de los Crassvirales y su distribución mundial

4.2. Capítulo 2: 99

La regulación de la biosíntesis de polisacáridos por la región de fase variable 7 (PVR7) en *Bacteroides intestinalis* es responsable de la resistencia a los fagos crAssBcn

4.3. Capítulo 3: 129

El reclutamiento de genomas completos de fagos tipo crAss revela su presencia en viromas de pollo, pocos fagos específicos de humanos y la ausencia de detección universal

4.4. Capítulo 4: 157

Persistencia de los fagos crAssBcn en condiciones de inactivación natural y procesos de desinfección, y su potencial como marcadores de trazabilidad de muestras de origen humano

5. DISCUSIÓN GENERAL 185

6. CONCLUSIONES 203

7. BIBLIOGRAFÍA 207

8. ANEXOS 247

8.1. Anexo 1: Otras publicaciones de la doctoranda no incluidas en la tesis .. 249

8.2. Anexo 2: Medios de cultivo y tampones 263

8.3. Anexo 3: Protocolos realizados 275

8.4. Anexo 4: Reacciones de PCR, qPCR, RT- qPCR y cebadores utilizados .. 301

8.5. Anexo 5: Plásmidos utilizados 307

ABREVIATURAS

AAI	Identidad de aminoácidos promedio (en inglés, *Average Amino acid Identity*)
ABB	Caldo basal anaerobio (en inglés, *Anaerobe Basal Broth*)
ADN	Ácido desoxirribonucleico
ARD	Dinámica de carrera armamentística (en inglés, *Arms Race Dynamics*)
ARN	Ácido ribonucleico
BBE	Agar *Bacteroides* bilis esculina (en inglés, *Bacteroides Bile Esculin Agar*)
bd	Bidestilada
BIF	Bacterias Indicadoras Fecales
BPRM	Medio de recuperación de fagos de *Bacteroides* (en inglés, *Bacteroides Phage Recovery Medium*)
BSA	Albúmina sérica bovina (en inglés, *Bovine Serum Albumin*)
°C	Grados centígrados
CrAssphage	En inglés, *Cross-Assembly phage*
CPS	Polisacáridos capsulares (en inglés, *Capsular Polysaccharides*)
CRISPR	Repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (en inglés, *Clustered Regularly Interspaced Short Palindromic Repeats*)
CsCl	Cloruro de cesio
Da	Dalton
DIG	Digoxigenina
DO	Densidad Óptica
EDAR	Estación de Depuración de Aguas Residuales
EDTA	Ácido etilendiaminotetraacético

FSD	Dinámica de selección fluctuante (en inglés, *Fluctuating Selection Dynamics*)
g	Gramos
h	Horas
ICTV	Comité Internacional de Taxonomía de Virus (en inglés, *International Committee on Taxonomy of Viruses*)
IPTG	Isopropil-β-D-tiogalactopiranósido
kb	Kilobase
KtW	Dinámica matar al ganador (en inglés, *Kill the Winner*)
L	Litro
LB	Luria-Bertani
LPS	Lipopolisacáridos
M	Molar
MAG	Genomas derivados de ensamblado metagenómico (en inglés, *Metagenome-Assembled Genomes*)
min	Minutos
ml	Mililitros
MST	Trazabillidad de la fuente microbiana (en inglés, *Microbial Source Tracking*)
NGS	Secuenciación de nueva generación (en inglés, *Next Generation Sequencing*)
nm	Nanómetros
ORF	Pauta de lectura abierta (en inglés, *Open Reading Frame)*
PBS	Tampón fosfato salino (en inglés, *Phosphate Buffered Saline*)
PCR	Reacción en cadena de la polimerasa
***p*-crAssphage**	crAssphage prototípico (en inglés, *prototypical cross-Assembly phage)*
PES	Poliéter sulfona

ppm	Partes por millón
PS	Polisacáridos
PtW	Dinámica de cargar con el ganador (en inglés, *Piggyback the Winner*)
p/v	Peso/volumen
PVR	Región con variación de fase (en inglés, *Phase Variable Region*)
qPCR	PCR cuantitativa a tiempo real
R-M	Sistema de restricción-modificación
rpm	Revoluciones por minuto
RT-qPCR	PCR cuantitativa con transcriptasa inversa
s	Segundos
SDS	Dodecil sulfato de sodio (en inglés, *Sodium Dodecyl Sulfate*)
spp	Especies
ss	Semisólido
UFP	Unidades Formadoras de calvas de lisis (del inglés, *Plaque Forming Units*)
UV	Ultravioleta
VLP	Partículas similares a virus (en inglés, *Virus-Like Particles*)
WT	Salvaje, silvestre, natural (en inglés, *Wild Type*)
x*g*	Fuerza centrífuga (unidades de gravedad)
X-gal	5-bromo-4-cloro-3-indolil-β-D-galactopiranósido
µg	Microgramo
µl	Microlitro
µm	Micrómetro
µM	Micromolar

1. INTRODUCCIÓN

1.1 Microbiota intestinal

Las superficies de la piel y las mucosas de los animales superiores están colonizadas por una gran variedad de microorganismos, conocidos colectivamente como la microbiota. Esta comunidad está compuesta principalmente por bacterias, arqueas y virus, además de microorganismos eucariotas, como protozoos, levaduras y hongos (Hoffmann et al., 2013; Huttenhower et al., 2012; Parfrey et al., 2011) y en algunos casos nemátodos intestinales (Hewitson and Maizels, 2014). En los seres humanos la microbiota alcanza su mayor densidad en el tracto gastrointestinal, con más de 100 billones de microbios concentrados principalmente en el intestino distal (Ley et al., 2006; Kamada et al., 2013; Sommer and Bäckhed, 2013). En conjunto, estos microorganismos poseen un repertorio genético aproximadamente 150 veces mayor que el del genoma humano (Karlsson et al., 2011; Wexler and Goodman, 2017), lo que ha llevado a considerar al microbioma intestinal como un “segundo genoma” (Bubier et al., 2021).

Este ecosistema microbiano ha evolucionado junto a los vertebrados durante millones de años, dando lugar a una relación mutualista en la que el hospedador proporciona nutrientes y un ambiente adecuado, mientras que la microbiota contribuye a múltiples funciones fisiológicas esenciales para la salud (Hooper and MacPherson, 2010). A pesar de su importancia, el ecosistema microbiano intestinal aún no se comprende completamente, y su diversidad y dinámica funcional no están bien definidas (Eckburg et al., 2005).

Durante mucho tiempo, la microbiota se consideró principalmente una posible fuente de patógenos (Sommer and Bäckhed, 2013). Sin embargo, los avances en la investigación han transformado esta perspectiva, subrayando la importancia de comprender la microbiota intestinal no como un simple reservorio de microorganismos potencialmente dañinos, sino como una comunidad ecológica compleja y dinámica, que puede influir en la fisiología y la salud de su hospedador (Lozupone et al., 2012).

Entre las funciones más destacadas de la microbiota intestinal se encuentran la digestión y fermentación de carbohidratos complejos, la producción de compuestos esenciales, como vitaminas y ácidos grasos de cadena corta, el desarrollo del epitelio intestinal y el sistema inmunológico, así como la protección contra patógenos oportunistas y el desarrollo del sistema linfoide asociado al intestino (Hooper and MacPherson, 2010; Stecher and Hardt, 2011; Renz et al., 2012; Sommer and Bäckhed, 2013; Horvath et al., 2022; Quan et al., 2023). Asimismo, el papel de la microbiota en el desarrollo y la regeneración tisular, la proliferación celular, la homeostasis ósea, el metabolismo y la modulación del comportamiento ha ampliado la comprensión de su influencia en la fisiología del hospedador, desafiando la noción de que

estos procesos dependen únicamente del programa genético del individuo (Sommer and Bäck, 2013).

A pesar de estos beneficios, el equilibrio entre el microbioma y el hospedador es delicado. Desequilibrios en el microbioma intestinal (disbiosis) podrían contribuir al desarrollo de diversas enfermedades (Kamada et al., 2013; Bubier et al., 2021), como la obesidad, enfermedades inflamatorias y metabólicas, trastornos neurológicos y cáncer (Lozupone et al., 2012; Sommer and Bäckhed, 2013; Quan et al., 2023). Estas disrupciones podrían estar asociadas a cambios en la composición y función de la microbiota, lo que subraya la importancia de mantener su equilibrio (eubiosis) que, asimismo, contribuye en el equilibrio del individuo (homeostasis) para preservar la salud.

El modelo representado en la Figura 1 muestra los mecanismos a través de los cuales el microbioma intestinal y el genoma del hospedador influyen conjuntamente en el fenotipo de una enfermedad.

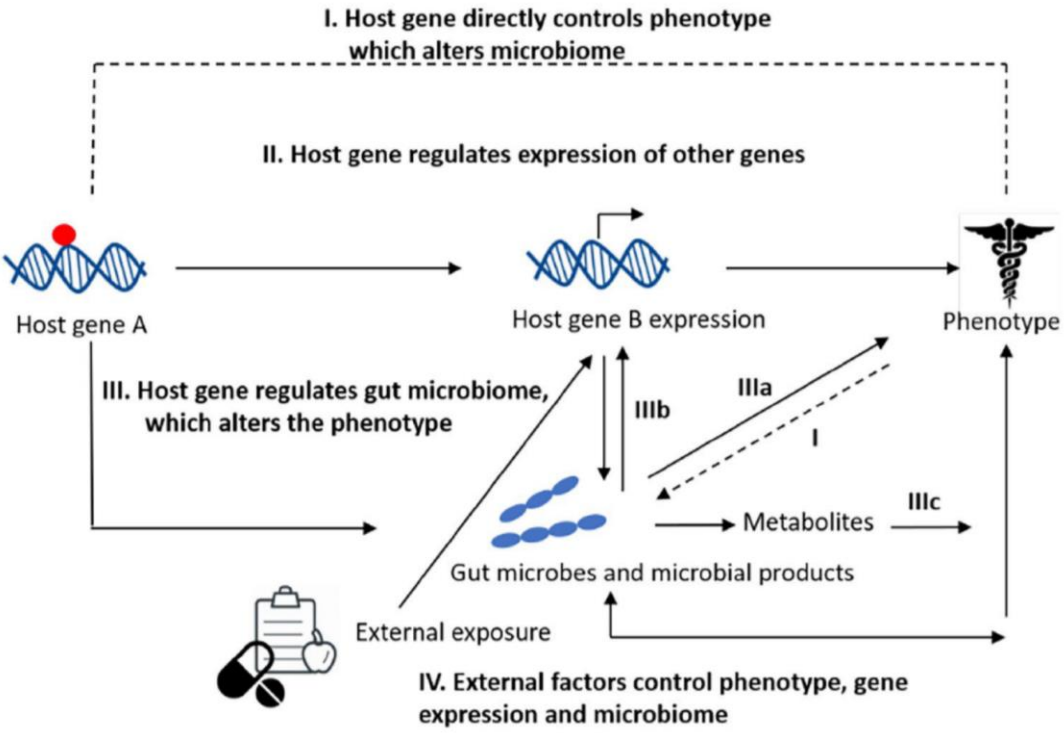


Figura 1. Influencia del genoma del hospedador y su microbioma intestinal en el fenotipo del individuo. Figura extraída de Bubier et al., (2021). Bajo licencia CC BY 4.0. No se realizaron modificaciones.

En este contexto, ciertas variantes genéticas en el hospedador pueden modular el fenotipo de manera directa (vía I) o indirecta, ya sea modulando la composición de la microbiota de manera directa (vía III) o regulando la expresión de otros genes (vía II), lo que también puede

provocar alteraciones en la microbiota. Ésta, a su vez, puede influir en el fenotipo del hospedador de diferentes maneras: directamente a través de moléculas en su superficie celular (vía IIIa), mediante metabolitos que afectan procesos fisiológicos (vía IIIc) o indirectamente al modular la expresión de genes del hospedador (vía IIIb). Finalmente, factores externos, como la dieta o los fármacos (vía IV) pueden alterar la microbiota intestinal, lo que repercute en la interacción entre la genética del hospedador y su fenotipo.

Estudios basados en el análisis de microbiomas intestinales humanos obtenidos en diversas regiones geográficas han evidenciado que cada individuo alberga una comunidad microbiana única. Esta comunidad se establece al nacer, cuando el recién nacido es colonizado por microorganismos provenientes de la madre y del entorno. A medida que el lactante va creciendo, su microbiota evoluciona y se estabiliza progresivamente hasta adquirir una composición característica de la edad adulta, alterándose de nuevo en la vejez. Este proceso de colonización está influenciado tanto por factores genéticos como ambientales, lo que explica la variabilidad observada entre individuos, e incluso dentro de una misma familia, y que existan diferencias importantes en distintos grupos poblacionales de distintas regiones geográficas (Sommer and Bäckhed, 2013). El análisis del microbioma mediante herramientas "ómicas" ha revelado que, aunque la composición general de la microbiota intestinal puede mantenerse estable durante largos períodos de tiempo, su estructura a nivel de especies es altamente dinámica y fluctúa en respuesta a estímulos externos, como la dieta o el uso de antibióticos (Ley et al., 2008; Lozupone et al., 2012; Sommer and Bäckhed, 2013; David et al., 2014; Mirzaei and Maurice, 2017).

1.1.1. Bacterias intestinales: un componente fundamental de la microbiota

Debido a las limitaciones inherentes a los métodos de cultivo convencionales, que no permiten detectar los microorganismos no cultivables, el estudio de la microbiota intestinal ha recurrido a enfoques moleculares que han permitido una caracterización más precisa y completa de su composición y diversidad. Entre estos métodos se encuentra la secuenciación del gen del ARN ribosomal 16S (16S ARNr) y las técnicas de secuenciación de nueva generación (*Next Generation Sequencing* o NGS). Estas técnicas han revelado que el intestino humano alberga entre 500 y 1000 especies bacterianas distribuidas mayoritariamente en unos pocos filos, incluidos *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Actinobacteria* y *Verrucomicrobia*, con un claro predominio de microorganismos anaerobios. En particular, aproximadamente el 90% de esta comunidad está compuesta por los filos *Firmicutes* (géneros *Lactobacillus, Bacillus, Clostridium, Enterococcus* y *Ruminococcus*) y *Bacteroidetes* (géneros *Bacteroides* y *Prevotella*) (Mirzaei and Maurice, 2017; Pinart et al., 2022).

La densidad microbiana aumenta progresivamente desde el intestino proximal hacia el distal (Figura 2), reflejando las importantes variaciones fisiológicas a lo largo del tracto intestinal. Entre los factores clave que influyen en esta distribución se encuentran:

- Un gradiente de pH que se vuelve menos ácido a medida que se avanza hacia el colon.
- Una mayor secreción de péptidos antimicrobianos en las regiones proximales.
- Un gradiente de oxígeno, con concentraciones más altas en la parte inicial del intestino y condiciones predominantemente anóxicas en la porción distal, lo que favorece el crecimiento de comunidades bacterianas anaerobias.

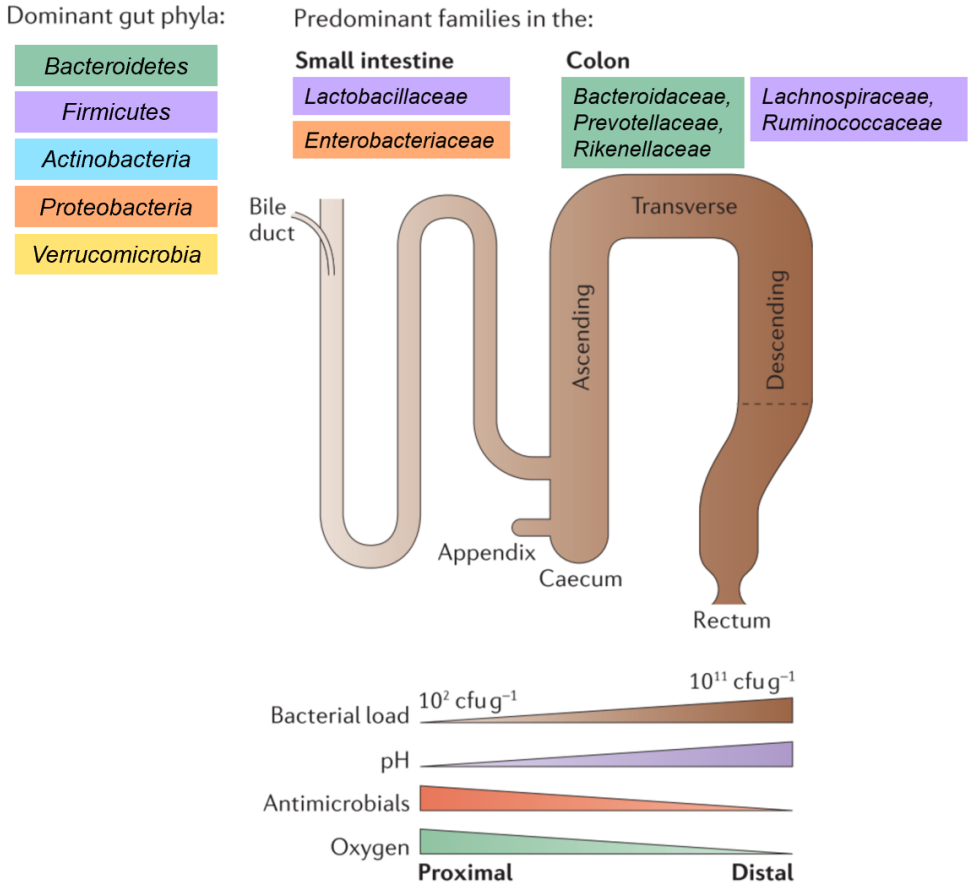


Figura 2. Distribución microbiana en el tracto intestinal humano. Figura adaptada de Donaldson et al., (2015).

Bacteroides

Bacteroides es uno de los géneros más abundantes del intestino. En particular, las especies del género *Bacteroides* son bacterias anaerobias obligadas, Gram negativas y resistentes a los ácidos biliares (Salyers, 1984), que viven exclusivamente en el tracto gastrointestinal de los mamíferos, demostrando una fuerte adaptación a este entorno (Ley et al., 2008). Pueden

alcanzar densidades de hasta 10^{11} células por gramo de contenido intestinal y desempeñan un papel crucial en la eubiosis del microbioma intestinal, actuando como especies fundamentales en su estructura (Shin et al., 2024).

Para adaptarse y sobrevivir en el intestino, *Bacteroides* muestra una notable plasticidad génica gracias a su habilidad de incorporar genes que provienen de otras bacterias mediante transferencia horizontal de genes, así como su capacidad para activar o desactivar genes específicos según sus necesidades (Risse et al., 2015; Yan et al., 2022; Shin et al., 2024). Ejemplos de esta adaptabilidad incluyen su capacidad para crecer en presencia de concentraciones bajas de oxígeno, a pesar de considerarse anaerobios (Baughn and Malamy, 2004), adaptar su fuente de alimento según la disponibilidad de nutrientes (Bjursell et al., 2006; Wexler and Goodman, 2017), exportar compuestos antibacterianos dirigidos a bacterias competidoras del tracto digestivo (Russell et al., 2014), y modular la respuesta inmune del hospedador para protegerse (Wexler, 2007). Sin embargo, su diseminación a otros tejidos puede desencadenar infecciones graves debido a su carácter oportunista, agravadas por su creciente resistencia a múltiples antibióticos (Shoemaker et al., 2001; Yan et al., 2022), lo que los convierte en los patógenos anaerobios con las tasas de resistencia más elevadas (Wexler, 2007). En consecuencia, dependiendo de su ubicación, de la composición de la microbiota circundante y de la disponibilidad de nutrientes, y posiblemente del estado de salud del hospedador, el género *Bacteroides* demuestra su versatilidad ya que puede actuar como mutualista, comensal o incluso patobionte (Wexler and Goodman, 2017).

Más en detalle, una de las principales funciones de *Bacteroides* es la degradación de polisacáridos (PS) complejos de cadena larga que no han podido ser absorbidos directamente por el hospedador (Figura 3). Esta capacidad se debe a la presencia de los *loci* de utilización de PS (*polysaccharide utilization loci* o PULs), un conjunto de genes que determinan los nichos metabólicos que estas bacterias pueden ocupar en el intestino (Lee et al., 2013; Donaldson et al., 2015). La expresión de los PULs no ocurre de manera simultánea, sino que está estrictamente regulada por la presencia de intermediarios de degradación de PS (Raghavan et al., 2014). Este mecanismo de control transcripcional permite a *Bacteroides* responder dinámicamente a cambios en la dieta, modulando su metabolismo en función de las variaciones en la disponibilidad de nutrientes. Por ejemplo, en períodos de ayuno o dietas bajas en fibra, comienzan a consumir glicanos de la mucosa del hospedador, asegurando así su persistencia en el ecosistema intestinal (Sonnenburg et al., 2005).

Además de su capacidad para degradar PS, también son capaces de sintetizarlos, formando cápsulas y glicoproteínas que modulan las interacciones, tanto microbianas como con el hospedador (Figura 3) (Pumbwe et al., 2006; Martens et al., 2009). *Bacteroides* emplea un

mecanismo de variación de fase para regular la expresión de sus *loci* de biosíntesis de polisacáridos capsulares *(capsular polysaccharide* o CPS) de manera alternante, como se comenta en detalle en el apartado (1.3.5). La variación de fase es un proceso aparentemente estocástico, controlado por promotores reversibles y también por proteínas inhibidoras *trans-locus* (Krinos et al., 2001). La expresión de uno o múltiples CPS en *Bacteroides*, que se expresan o inhiben de forma más o menos jerárquica (Chatzidaki-Livanis et al., 2010), genera una diversidad estructural en su superficie. Esto facilita su adaptación al entorno intestinal (Peterson et al., 2007), aunque esta misma cápsula puede representar una barrera para la difusión de nutrientes (Martens et al., 2009).

Otro mecanismo de variabilidad se encuentra en la capacidad de *Bacteroides* para modificar sus lipopolisacáridos (LPS) otorgándole resistencia frente a los péptidos antimicrobianos del hospedador (Wexler and Goodman, 2017).

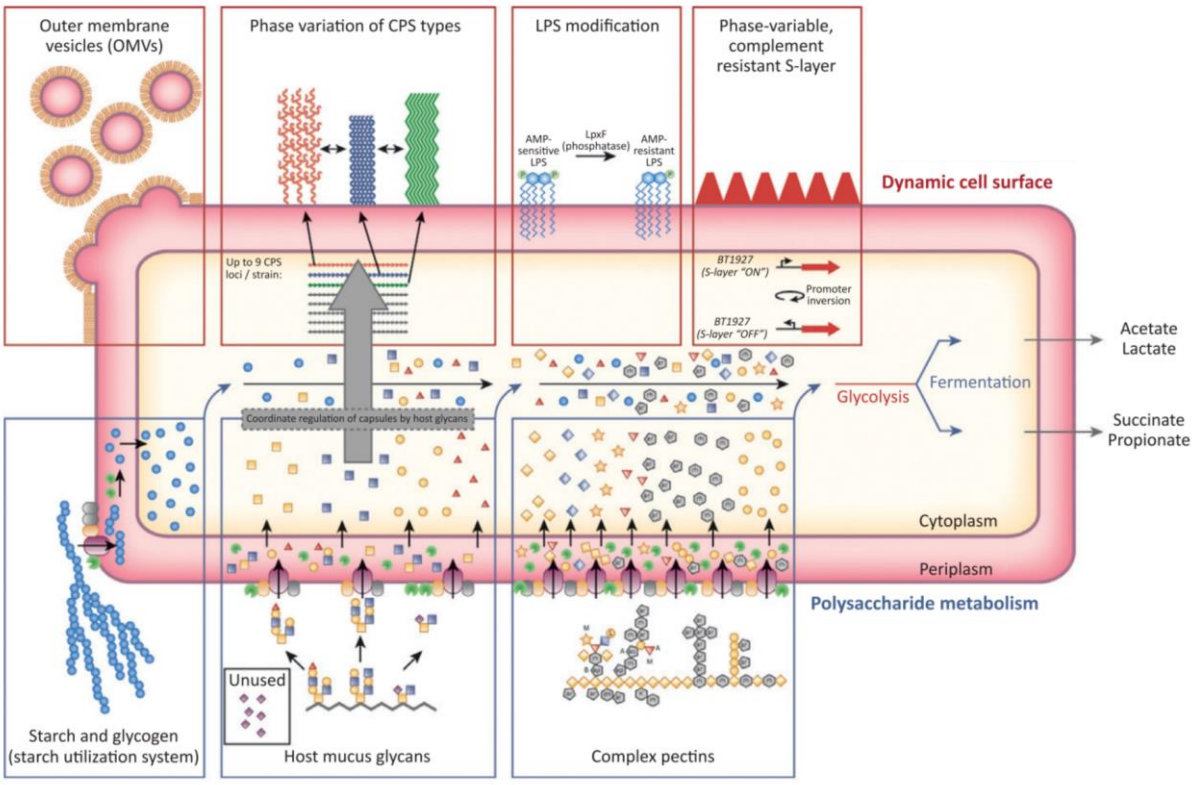


Figura 3. Infografía de *Bacteroides thetaiotaomicron* donde se muestra la dinámica de su superficie celular y el metabolismo de los PS. Figura adaptada de Porter et al., (2018).

1.1.2. Viroma intestinal

Tradicionalmente, los virus eran considerados únicamente agentes patógenos; sin embargo, hoy en día, esta visión se ha replanteado, identificándolos como componentes fisiológicos esenciales del microbioma (Scarpellini et al., 2015), tras el primer informe sobre la existencia de un viroma intestinal humano publicado en 2003 (Breitbart et al., 2003). Se denomina viroma al conjunto de todos los virus encontrados en un organismo.

Gracias a los continuos avances en las técnicas de metagenómica, que abarcan tanto tecnologías de secuenciación de alto rendimiento como metodologías asociadas al procesamiento de datos, se ha logrado un progreso significativo en el estudio de los viromas intestinales (Li et al., 2014). Estas innovaciones han permitido abordar uno de los principales desafíos de este campo: **"la materia oscura vírica** o ***viral dark matter*"**,** que hace referencia a aquellas secuencias virales que no se alinean con ninguna secuencia de referencia y, por ende, no pueden ser clasificadas taxonómicamente (suelen ser entre el 40 y el 90% de las secuencias obtenidas de un viroma) (Roux et al., 2015; Krishnamurthy and Wang, 2017). La carencia de marcadores filogenéticos universales en los genomas virales complica adicionalmente el estudio de su diversidad y dinámica en las comunidades microbianas (Reyes et al., 2012).

Diversos estudios han demostrado que la microbiota del intestino alberga innumerables poblaciones de virus (o partículas similares a virus, en inglés *virus-like particles* o VLPs), incluyendo a patógenos humanos, virus de plantas adquiridos por la dieta y bacteriófagos (virus que infectan bacterias), siendo éstos últimos los predominantes (Carding et al., 2017; Edwards et al., 2019; Minot et al., 2013; Reyes et al., 2012; Shkoporov and Hill, 2019). Se estima que en el intestino el número de fagos supera en 1 – 100 veces al de bacterias (Fortier and Sekulovic, 2013), siendo por tanto las entidades más numerosas de la microbiota intestinal. A pesar de que la comunidad vírica se presenta como relativamente más estable que el componente bacteriano (Broecker et al., 2017), también se evidencia una alta diversidad en su composición a nivel interpersonal (Minot et al., 2013; Ogilvie and Jones, 2015; Shkoporov et al., 2019). Esta variabilidad se explica, en parte, por la rápida evolución de algunos miembros residentes del viroma, lo que favorece la aparición de nuevas especies virales a lo largo del tiempo.

1.2. Bacteriófagos

Los bacteriófagos, también conocidos como fagos, son virus que infectan a bacterias, aunque algunos estudios también incluyen en este grupo a los virus que infectan arqueas, mientras que otros hablan de virus de arqueas (Abedon and Murray, 2013). Los fagos representan las entidades biológicas más numerosas, ubicuas y diversas de la biosfera, estimando que existen más de 10^{31} fagos en la Tierra (Suttle, 2005). Como parásitos intracelulares obligados, dependen exclusivamente de su hospedador para replicarse y propagar. Estos virus, en su forma extracelular de virión, exhiben una notable variabilidad en cuanto a su tamaño, que oscila entre 20 y 200 nm. Su material genético, ADN o ARN de cadena simple o doble, se encuentra protegido por una cápside proteica, que en algunos casos también puede estar rodeada por una envoltura lipídica. La cápside puede adoptar diversas morfologías, como poliédrica, filamentosa o pleomórfica, y en ciertos casos está conectada a una cola mediante una estructura especializada que facilita el reconocimiento de la bacteria hospedadora y la inyección del material genético en su interior (Dion et al., 2020; Jamal et al., 2019; Sharma et al., 2017; Zhu et al., 2022; Zinke et al., 2022). En particular, los fagos con cola y genoma de ADN bicatenario son los más representados en las bases de datos.

En cuanto a su clasificación taxonómica, hasta hace poco, se basaba en criterios físicos, tales como la estructura, tamaño y forma de la cápside, el tipo de material genético, el rango de hospedadores, el ciclo de replicación y la patogenicidad, entre otros (Sharma et al., 2017; Simmonds et al., 2017). Recientemente, el Comité Internacional de Taxonomía de Virus (ICTV) (https://ictv.global/taxonomy) ha implementado cambios drásticos en el sistema de clasificación. Estos ajustes incluyen la eliminación de grandes familias de fagos, como *Siphoviridae, Podoviridae* y *Myoviridae*, basadas en características morfológicas, así como del orden Caudovirales (Turner et al., 2023; Xihui et al., 2024; Zhu et al., 2022). En su lugar, el ICTV ha adoptado un enfoque fundamentado en las secuencias genómicas como criterio para la clasificación (Simmonds et al., 2017) y ha introducido el nombre en formato binomial para la designación de especies virales. Esta decisión responde al creciente número de secuencias genómicas identificadas gracias al uso de la metagenómica. Este nuevo sistema ha permitido la creación de nuevos taxones que reflejan de forma más precisa la diversidad de los fagos. Por ejemplo, se ha incorporado la nueva clase Caudoviricetes, que agrupa todos los fagos con cola, cápside icosaédrica y genoma de ADN de doble cadena, que constituyen la mayoría de los fagos intestinales descritos (Pargin et al., 2023).

No obstante, a pesar de las ventajas que ofrece la metagenómica, existen ciertas limitaciones. Entre ellas se encuentran la carencia de genomas de referencia anotados, la insuficiente resolución para reconstruir genomas con secuencias homólogas, la elevada diversidad

derivada del mosaico genético de los genomas fágicos y la dificultad de ensamblar secuencias que corresponden a un único genoma viral presente dentro de una mezcla metagenómica (Dion et al., 2020; Dutilh et al., 2014; Shang et al., 2021; Simmonds et al., 2017). En ocasiones, en las bases de datos se encuentran genomas completos de fagos, pero aún es más frecuente encontrar genomas derivados de ensamblado metagenómico (*Metagenome-Assembled Genomes* o MAGs) que son genomas usualmente incompletos que han sido reconstruidos a partir de los datos metagenómicos (Setubal, 2021). Un ensamblaje incorrecto, o realizado a partir de datos insuficientes puede llevar a secuencias de genomas que no existan en la realidad. De este modo, las técnicas de metagenómica se deben complementar con nuevas técnicas de cultivo que permitan tener colecciones de fagos aislados, determinar su genoma con exactitud y con los que a su vez se pueda determinar el rango de hospedadores para estudios más en profundidad de las relaciones entre virus y hospedador y sus dinámicas de infección (Fitzgerald et al., 2021).

Después de numerosos años de investigación en diferentes matrices ambientales y, mediante la aplicación de un enfoque integral que combina tanto técnicas tradicionales de estudio como métodos moleculares, se ha demostrado de manera concluyente que los fagos están ampliamente distribuidos en una gran variedad de entornos. Entre las metodologías tradicionales empleadas se incluyen la citometría de flujo, la microscopia de epifluorescencia y la microscopia electrónica de transmisión, así como métodos de aislamiento en cultivo (Ackermann, 2009; Brussaard, 2009; Ortmann and Suttle, 2009). Complementariamente, entre las técnicas moleculares se encuentran la qPCR y la PCR dirigidas a diversos genes dentro del genoma fágico (Adriaenssens and Cowan, 2014), el uso de marcadores basados en polimorfismos de longitud de fragmentos de restricción (RFLP), la electroforesis en gel con gradiente de desnaturalización (DGGE) (Clokie et al., 2011) y, más recientemente, técnicas de secuenciación masiva. Los fagos se han detectado en ecosistemas acuáticos (Chow and Suttle, 2015; Paul et al., 2002), en la microbiota de organismos superiores (Manrique et al., 2017; Sutton and Hill, 2019), en suelos (Williamson et al., 2017), e incluso, en condiciones extremas de salinidad o temperatura (Pietilä et al., 2010; Säwström et al., 2008). Estos hallazgos evidencian su ubicuidad y adaptabilidad en la biosfera, remarcando su relevancia ecológica y evolutiva en diversos nichos ambientales. Esta amplia distribución se atribuye en gran medida a la abundancia y dispersión de sus organismos hospedadores (Clokie et al., 2011), hecho que también influye en la variabilidad interpersonal de los fagos dada su alta especificidad por dichas cepas bacterianas (Minot et al., 2013; Rodriguez-Valera et al., 2009). En general, se encuentran fagos en todos los nichos donde están presentes sus bacterias hospedadoras, e incluso se detectan fagos en ambientes donde sus bacterias hospedadoras se han inactivado y ya no están presentes.

El estudio de la presencia de fagos en diversos entornos, además de contribuir a la comprensión de la biología viral, permite el desarrollo de una amplia variedad de aplicaciones en el campo de la biotecnología. Estas aplicaciones incluyen, entre otras, el uso de fagos para el biocontrol de bacterias en la industria alimentaria que pueden ocasionar enfermedades o el deterioro del producto (Gutiérrez et al., 2017), así como el biocontrol de patógenos en sectores como la agricultura y la acuicultura (Dy et al., 2018). En el ámbito de la salud, los fagos se pueden emplear en la terapia fágica, tanto en humanos como en animales (Barrow, 2001; Loc-Carrillo and Abedon, 2011) ofreciendo ventajas como un menor impacto sobre la composición de la microbiota intestinal en comparación con los antibióticos. Además, los fagos son herramientas valiosas para el descubrimiento de nuevas moléculas antimicrobianas (Liu et al., 2004) y para explorar interacciones proteína-ligando (Azzazy and Highsmith Jr, 2002), lo que abre posibilidades para el desarrollo de terapias inmunológicas (Bazan et al., 2012). Adicionalmente, también son ampliamente usados como indicadores de contaminación fecal (Toribio-Avedillo et al., 2021).

1.2.1. Ciclos de replicación de los bacteriófagos

Dada su estrecha interacción con diferentes microorganismos, los fagos tienen la capacidad de modular tanto la composición de las poblaciones bacterianas circundantes como sus funciones ecológicas en los ecosistemas (Koskella and Brockhurst, 2014; Weitz and Wilhelm, 2012). Esta interacción con su bacteria hospedadora se manifiesta a través de dos ciclos de replicación distintos (Morimoto et al., 2022; Reyes et al., 2012): el ciclo lítico (fagos virulentos) y el ciclo lisogénico (fagos atemperados).

En el **ciclo lítico** los fagos identifican a su hospedador al reconocer receptores específicos presentes en la superficie bacteriana, como el LPS, proteínas de membrana externa o ácidos teicoicos (Sharma et al., 2017). Tras esta identificación, se adhieren a la célula e inyectan su material genético en el citoplasma bacteriano. Posteriormente, se secuestra la maquinaria biosintética del hospedador para replicar sus ácidos nucleicos y para sintetizar las nuevas proteínas virales. Las proteinas se ensamblan en cápsides dentro de las que se empaqueta el ácido nucleico, formando viriones completos, que finalmente son liberados del interior de la célula mediante la lisis de la bacteria, provocando su muerte e iniciando un nuevo ciclo infeccioso (Jamal et al., 2019).

En algunos casos, parte del genoma del hospedador puede ser empaquetado en la cápside del fago, lo que da lugar a un mecanismo de transferencia horizontal de genes conocido como **transducción**. Este proceso puede conferir nuevas características al hospedador, como

mayor patogenicidad (Muniesa et al., 2004; O'Shea and Boyd, 2002), resistencia a antibióticos (Colomer-Lluch et al., 2011) e incluso nuevas capacidades metabólicas (Borodovich et al., 2022; Hehemann et al., 2010; Hurwitz and U'Ren, 2016). Además, este ciclo impulsa la rápida evolución de la población viral, favoreciendo la aparición de nuevas especies de virus que incorporan nuevo material genético (Minot et al., 2013). Por otro lado, los hospedadores intentan defenderse mediante la modificación y diversificación de los receptores bacterianos, promoviendo en los hospedadores una mayor tasa de mutación (Pal et al., 2007; Reyes et al., 2012).

En el **ciclo lisogénico** el ADN del fago inyectado se integra en el ADN cromosómico del hospedador (lisógeno), mediante recombinación homóloga (Casjens and Hendrix, 2015) o no homóloga (Harshey, 2015). El ADN fágico incorporado en el genoma bacteriano, denominado profago, permanece en estado latente en el genoma bacteriano y se replica de manera concomitante con el hospedador, sin provocar consecuencias negativas a la bacteria infectada. También se puede dar el caso de la **pseudolisogenia**, donde el genoma del fago persiste en el interior de la célula hospedadora, sin que haya multiplicación del genoma fágico, ni replicación sincronizada con el ciclo celular. Este fenómeno generalmente se da bajo condiciones de limitación nutricional, que restringen la replicación del ADN bacteriano o la síntesis de proteínas (Mirzaei and Maurice, 2017; Shkoporov and Hill, 2019).

Bajo ciertas condiciones, como cambios en la densidad celular, factores ambientales o alteraciones metabólicas (Howard-Varona et al., 2017; Paul, 2008), los genes del profago que permanecían reprimidos pueden reactivarse, provocando la transición del ciclo lisogénico al lítico (inducción del ciclo lítico). Este cambio permite la replicación y liberación de nuevos viriones, asegurando la supervivencia del fago y la continuidad de su ciclo de vida. Cuando la inducción del ciclo lítico ocurre en respuesta al estrés ambiental, se generan restos celulares que actúan como fuentes de nutrientes para la población bacteriana superviviente o para otras poblaciones del mismo ecosistema. Además, este proceso contribuye a una dinámica constante de diversidad (Rodriguez-Valera et al., 2009) y puede favorecer la eliminación de competidores, lo que amplía el nicho ecológico de las bacterias que albergan el fago en su genoma (Figura 4).

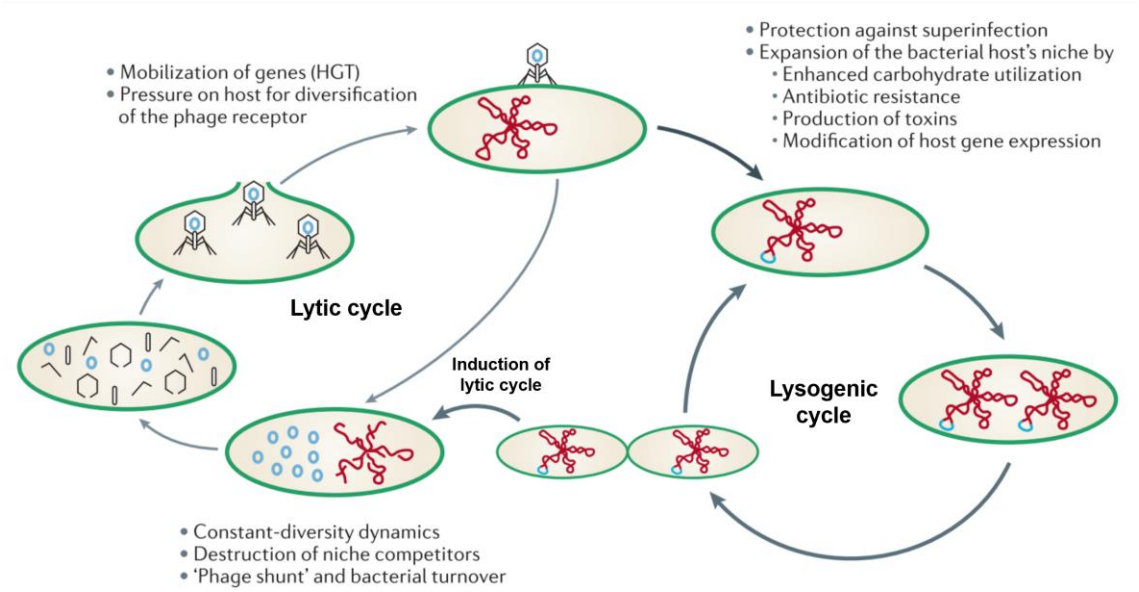


Figura 4. Impacto de los ciclos de vida de un fago en el microbioma intestinal humano. Figura adaptada de Reyes et al., (2012).

En ocasiones, los profagos pueden llegar a perder los genes necesarios para la producción de partículas virales infecciosas (profagos crípticos), por lo que acaban siendo elementos genómicos derivados de profagos que se pueden mantener selectivamente en caso de suponer ventajas al hospedador (Howard-Varona et al., 2017).

En términos generales, se estima que diariamente, entre el 20% y el 50% de las bacterias presentes en la biosfera son infectadas y lisadas por fagos. Además, aproximadamente la mitad de las bacterias cuyo genoma ha sido secuenciado contienen al menos una secuencia fágica integrada en su cromosoma (Shkoporov et al., 2022).

1.2.2. Dinámica de control de poblaciones bacterianas mediada por fagos

El estudio de las interacciones entre fagos y sus hospedadores ha revelado que la dinámica de estos virus no se limita únicamente a la eliminación total de las poblaciones bacterianas, sino que se han identificado escenarios de coexistencia en los que la relación fago-hospedador resulta en beneficios mutuos, lo cual sugiere una compleja interdependencia evolutiva y funcional entre ambos microorganismos (Rodríguez-Rubio et al., 2020). Si bien estas relaciones son claramente evidentes a nivel comunitario, se vuelven mucho menos perceptibles cuando se analizan a nivel de célula bacteriana individual y un único virión. De hecho, en este contexto, la infección por el fago suele tener efectos negativos para el hospedador, ya que puede enfrentarse constantemente a la amenaza de lisis o verse obligado

a producir viriones de forma continua, lo que implica un significativo coste metabólico (Shkoporov et al., 2022).

En los ambientes marinos, predominan las interacciones depredador-presa (ciclo lítico), más específicamente descritas bajo la dinámica "matar al ganador" **(*Kill the winner* o KtW)** (Winter et al., 2010). Este modelo ecológico postula que los fagos desempeñan un papel crucial en la regulación de las poblaciones bacterianas, permitiendo la coexistencia de especies bacterianas de crecimiento lento con aquellas de crecimiento más rápido. Las bacterias de rápido crecimiento alcanzan elevadas densidades poblacionales, lo que las hace más susceptibles a infecciones por virus. Esta presión viral reduce sus poblaciones, favoreciendo así la diversidad microbiana al evitar la dominancia de especies altamente competitivas (Ogilvie and Jones, 2015), además de maximizar la eficiencia en el uso de recursos por parte de la comunidad (Manrique et al., 2017).

En el entorno intestinal, se manifiesta un patrón distintivo en las interacciones virus-hospedador, caracterizado por una mayor complejidad debido a la influencia de una gran cantidad de factores bióticos y abióticos, entre los que se incluyen la estructura anatómica del intestino, la presencia de un sistema inmunológico, la constante entrada desde el ambiente de nuevas bacterias y fagos, así como factores como el pH ácido, presencia de sales biliares o bien la composición de la dieta, entre otros (Jończyk et al., 2011; Shkoporov and Hill, 2019).

Para empezar, también se encuentran fagos que siguen el ciclo lítico. En individuos sanos, se observa un equilibrio entre dos dinámicas evolutivas distintas: la "carrera armamentística" (**A*rms-Race Dynamics*** o **ARD**) y la "dinámica de selección fluctuante" (***Fluctuating Selection Dynamics"*** o **FSD**), también conocida como **teoría de la Reina Roja.**

En el modelo **ARD**, la interacción depredador-presa se caracteriza por una constante carrera evolutiva en la que la bacteria desarrolla mecanismos de defensa para evadir la infección, lo cual conlleva un elevado coste tanto a nivel metabólico (consumo de ATP) como a nivel competitivo (por ejemplo, pérdida del pilus que actúa como receptor, necesario para la movilidad o la formación de biofilms) (van Houte et al., 2016). Al mismo tiempo, los fagos evolucionan continuamente para superar dichas defensas, ampliando sus rangos de infectividad. Este proceso culmina en la disminución de la diversidad, ya que solo las cepas que han acumulado mecanismos de resistencia amplios prevalecen, mientras que las variantes sensibles son eliminadas.

Por otor lado, el modelo **FSD** permite la coexistencia prolongada de fagos y bacterias, manteniendo fluctuaciones dinámicas en sus poblaciones. Esto fomenta una mayor diversidad microbiana al mantener polimorfismos de defensa especializados en las comunidades

bacterianas. Como resultado, los fagos desarrollan rangos de hospedadores más estrechos y específicos (Mirzaei and Maurice, 2017).

De este modo, a nivel poblacional, la constante presión ejercida por los fagos impulsa a las bacterias a generar nuevas variantes genotípicas y fenotípicas, fomentando así la diversidad. Este proceso puede ocurrir a través de varios mecanismos (Figura 5): (1) las poblaciones sensibles a los fagos son susceptibles a la infección, pero simultáneamente generan variantes fenotípicas reversibles que se vuelven temporalmente insensibles, así como mutantes con resistencia permanente; (2) estrategias como la lisogenia permiten a las células infectadas resistir a nuevas infecciones, prolongar su supervivencia y posponer la lisis; (3) la lisis de células infectadas reduce la subpoblación sensible, al mismo tiempo que impulsa el crecimiento de variantes insensibles o mutantes resistentes; (4) la disminución de la densidad de fagos, junto con los costes metabólicos asociados a la resistencia y la constante reversión de las variantes fenotípicas, contribuyen a la persistencia de subpoblaciones sensibles y promueven la diversificación global de la población bacteriana (Shkoporov et al., 2022).

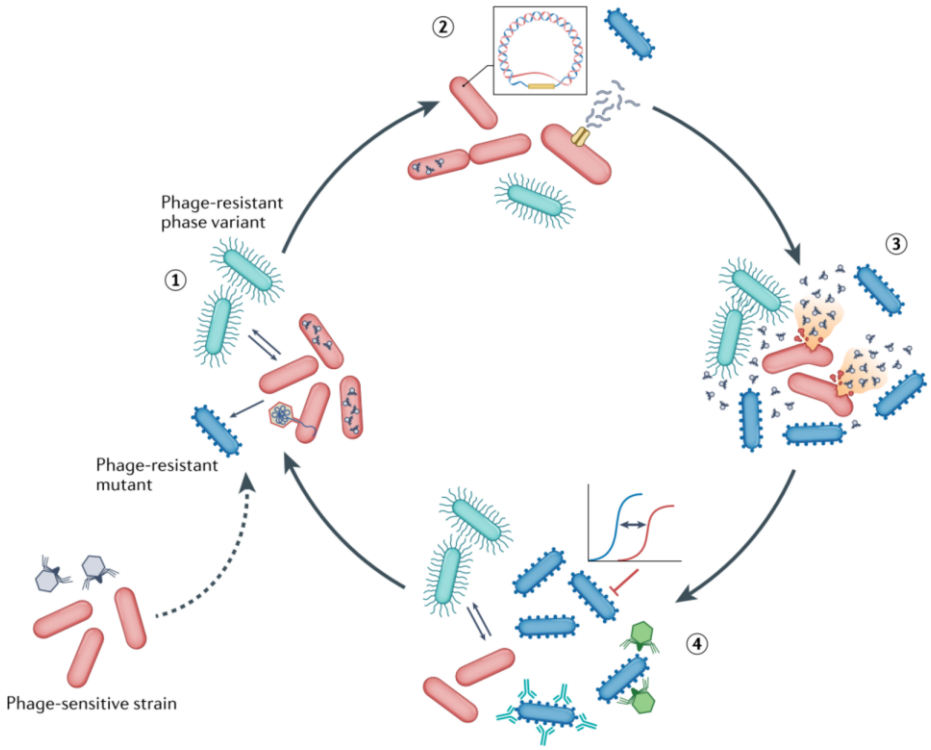


Figura 5. Interacciones que permiten la coexistencia de fagos y bacterias a nivel de población. Figura adaptada de Shkoporov et al., (2022).

En relación con los mecanismos de defensa bacteriana frente a los fagos, se han identificado diversas estrategias, las cuales abarcan tanto modificaciones estructurales como mecanismos de respuesta adaptativa. Entre estas estrategias (Figura 6) (Mirzaei and Maurice, 2017) se encuentran:

- La modificación de los receptores de superficie o la producción de CPS que enmascaran estos receptores, dificultando así la adsorción del fago. Además, el estado metabólico de las bacterias desempeña un papel clave en la regulación de estas estructuras superficiales, confiriendo una resistencia transitoria a la infección sin necesidad de recurrir a mutaciones irreversibles (De Sordi et al., 2019). Esta plasticidad fisiológica y estructural permite que las bacterias ajusten su susceptibilidad a los fagos en función de las condiciones ambientales, asegurando su supervivencia en distintos contextos ecológicos (explicado más en detalle en el apartado 1.3.5).
- Mecanismos basados en los sistemas de restricción-modificación (R-M) en las bacterias para degradar el genoma fágico cuando entra en el citoplasma mediante restricción en dianas específicas (Mirzaei and Maurice, 2017).
- Mecanismos de inmunidad CRISPR-Cas (Garneau et al., 2010) que permiten un reconocimiento específico de secuencias virales previamente encontradas en infecciones anteriores, otorgando así una defensa adquirida altamente eficaz contra infecciones recurrentes.
- Mecanismos tipo BREX (BacteRiophage EXclusion) (Goldfarb et al., 2015), un sistema compuesto de diversas proteinas que metilan el ADN bacteriano, lo que le permite detectar el ADN fágico y prevenir su replicación ya que éste no presenta metilación en la secuencia de reconocimiento del sistema BREX.
- Algunas bacterias implementan el suicidio celular altruista mediante el cual una célula infectada se autodestruye (por ejemplo, induciendo una toxina) con el propósito de proteger a la población circundante, evitando así la propagación del fago hacia células sanas.

Por otro lado, los fagos han desarrollado mecanismos de contra-defensa para superar estas barreras. Estos incluyen (Samson et al., 2013):

- El reconocimiento de motivos variables o nuevos receptores bacterianos.
- La producción de enzimas hidrolasas capaces de degradar el CPS.
- Modificar sus propios sitios de restricción o interferir con el sistema CRISPR-Cas para evitar la degradación de su genoma.
- También puede sintetizar antitoxinas que neutralizan el suicidio bacteriano, permitiendo así la continuidad del ciclo infeccioso.

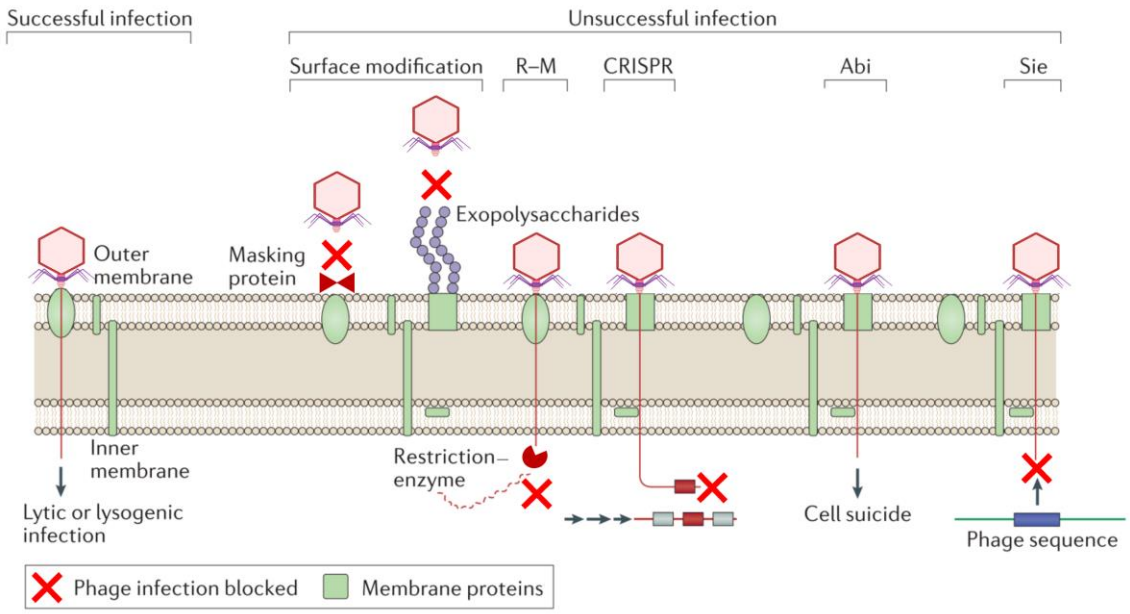


Figura 6. Mecanismos de defensa de las bacterias frente la infección de fagos. Figura adaptada de Mirzaei and Maurice (2017).

Sin embargo, existe una marcada predominancia de fagos que siguen un ciclo de vida lisogénico (Minot et al., 2013). Este hecho se sustenta por la predominancia observada de genes como integrasas y recombinasas, indispensables para la lisogenia (Mirzaei and Maurice, 2017), y en los recuentos de VLPs en muestras fecales. Estos recuentos oscilan entre 10^8 y 10^9 VLPs por gramo de muestra, suponiendo una proporción aproximada de 1:1 entre virotipos y filotipos bacterianos en el intestino, en marcado contraste con lo que se ha reportado en ambientes oceánicos y en suelos, donde las proporciones alcanzan hasta 10:1 (Carding et al., 2017; Reyes et al., 2012; Williamson et al., 2005).

Este tipo de relación simbiótica puede beneficiar tanto al fago como a su hospedador, con importantes implicaciones en el intercambio genético y la alteración de fenotipos bacterianos (Ogilvie and Jones, 2015). Esta interacción favorece a las bacterias intestinales en su adaptación al entorno y así mantener su estabilidad en la microbiota (Duerkop et al., 2012; Minot et al., 2011). Las funciones que aportan los profagos incluyen la codificación de genes involucrados en la obtención de energía, como los relacionados con el metabolismo de carbohidratos y aminoácidos, así como genes que confieren virulencia, como las toxinas, y los que otorgan resistencias a antibióticos. De acuerdo con Owen et al., (2020), estas funciones se encuentran codificadas en *loci* “accesorios”, los cuales suelen estar altamente expresados durante la lisogenia, mientras que los genes asociados a la producción de viriones están reprimidos.

Además, pueden desempeñar un papel en la prevención de infecciones por el mismo fago o de uno estrechamente relacionado, mediante el mecanismo de exclusión de sobreinfección (Hunter and Fusco, 2022). Esta capacidad también se respalda por la distribución inversa observada entre la presencia de profagos y los sistemas CRISPR, lo que sugiere que los lisógenos comensales pueden desarrollar resistencia a otras infecciones mediante la protección conferida por sus profagos (Bondy-Denomy et al., 2016). Por otro lado, existen casos en los que se liberan profagos activados sólo por una fracción de la población lisogénica, que posteriormente puede infectar y lisar especies competidoras (Imamovic et al., 2016; Manrique et al., 2017).

La predicción de secuencias correspondientes a profagos continúa siendo un reto en la metagenómica del viroma intestinal, ya que los métodos bioinformáticos no siempre logran diferenciarlas con certeza del resto del genoma bacteriano (Nayfach et al., 2021).

En el entorno intestinal (Figura 7), el ciclo lisogénico no se selecciona únicamente cuando la densidad de bacterias es baja, sino que también puede activarse en condiciones de alto título de hospedadores y de nutrientes suficientes. Esto permite a los fagos aprovechar la rápida proliferación bacteriana para replicarse simultáneamente, un fenómeno conocido como el modelo “cargar con el ganador” (***piggyback the winner*** o **PtW)** (Knowles et al., 2016; Silveira and Rohwer, 2016).

En el lumen intestinal, donde la concentración de bacterias es especialmente alta, prevalece una dinámica PtW, lo que resulta en una baja proporción de virus libres respecto a bacterias (VMR – *viral to microbe ratio*). Sin embargo, a medida que se avanza hacia el colon terminal, la disminución de nutrientes y el aumento del estrés oxidativo promueven la inducción de los profagos. En contraste, en la capa de mucina, los niveles de bacterias se mantienen relativamente bajos, lo que provoca un cambio hacia la dinámica KtW. Además, se ha observado que ciertas partículas fágicas se adhieren a la mucina, proporcionando una inmunidad mediada por bacteriófagos (Shkoporov and Hill, 2019).

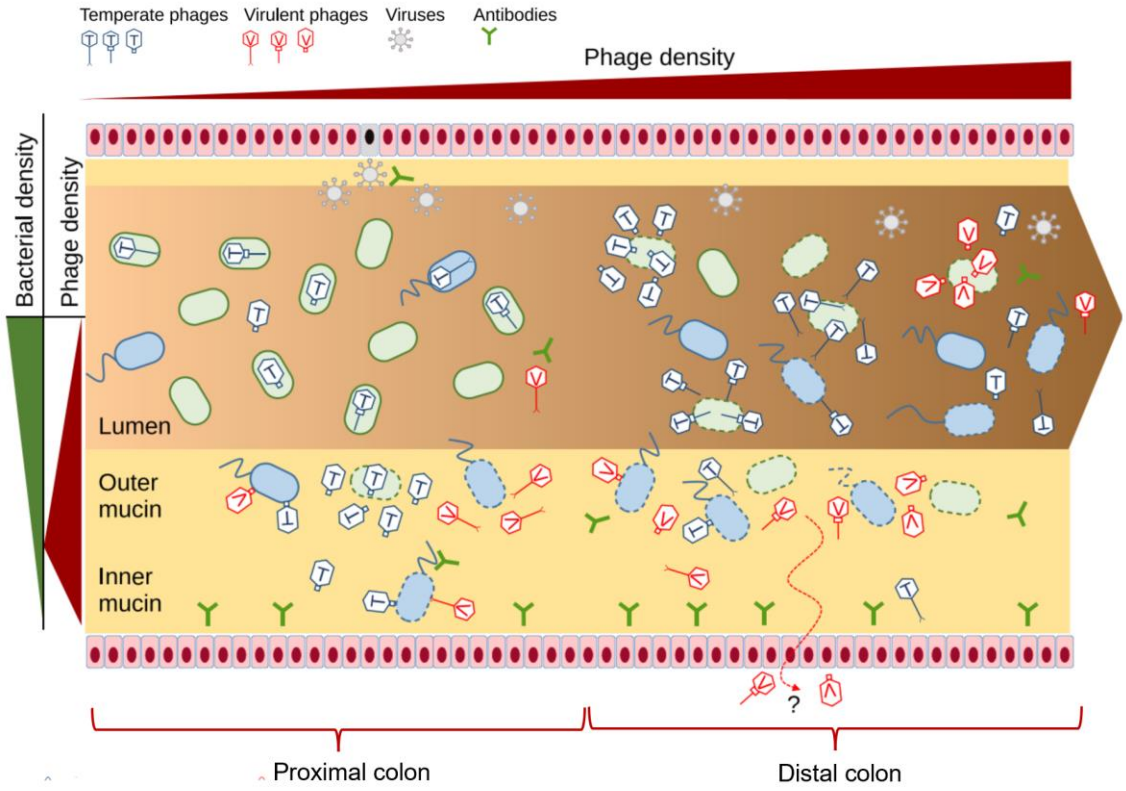


Figura 7. Dinámica de los fagos en el intestino. Figura adaptada de Shkoporov and Hill (2019).

Hay que destacar también, que pese a la elevada especificidad por el hospedador que presentan los fagos en ensayos *in vitro*, pueden expandir su rango de hospedadores permitiendo una amplia infectividad entre cepas de la misma especie, cuando se encuentran en el ambiente intestinal en ratones (Hsu et al., 2019; Modi et al., 2013).

Sin embargo, queda mucho por descubrir sobre las relaciones que se establecen entre los fagos y sus hospedadores dentro del intestino. Esta falta de conocimiento se debe, en parte, a diversas limitaciones metodológicas, como la imposibilidad de cultivar aproximadamente el 60% de las bacterias intestinales (Browne et al., 2016). Además, aunque el uso de técnicas "ómicas" ha permitido avances significativos en la comprensión de la biología de los fagos, como el hallazgo de dominios adhesivos que facilitan la fijación al mucus intestinal (Barr et al., 2013), estos métodos dan una visión muy restringida y no capturan la complejidad de las dinámicas entre bacterias y fagos en su totalidad.

En este contexto, dado que la composición bacteriana intestinal está estrechamente relacionada con la eubiosis del microbioma, se ha planteado la posibilidad de que ciertos grupos de fagos desempeñen un papel clave en la regulación y mantenimiento de una microbiota intestinal saludable.

1.3. CrAssphage

En el año 2014, la aplicación de técnicas avanzadas de metagenómica por Dutilh et al., (2014) permitió el ensamblaje cruzado de múltiples lecturas provenientes de varios viromas fecales humanos. Este enfoque condujo a la identificación de una nueva secuencia genómica circular de ADN de cadena doble, con una longitud aproximada de 100 kb, perteneciente a un grupo de fagos hasta entonces desconocido. La mayoría de los genes de esta secuencia no mostraban homología con secuencias previamente registradas en bases de datos, lo que evidenciaba su carácter novedoso. Este virus, descrito como "el más abundante del cuerpo humano", está presente en al menos el 50% de la población, con una mayor prevalencia en comunidades industrializadas (Honap et al., 2020), aunque también se ha identificado en poblaciones remotas en menor medida, lo que sugiere su amplia distribución. Representando hasta el 90% de la carga total de ADN vírico intestinal en personas de culturas occidentales, y más del 20% del ADN fecal total, este nuevo fago fue denominado ***p*-crAssphage** (fago prototipo de ensamblaje cruzado, en inglés *prototypical cross-Assembly phage*) en referencia a la metodología de ensamblaje cruzado utilizada para su identificación.

En realidad, este nuevo crAssphage asociado al intestino humano no constituye una entidad única, sino un grupo diverso de virus, como demostraron Yutin et al., (2018), quienes identificaron homologías distantes y sintenias genómicas en secuencias extraídas de bases de datos. Los fagos identificados posteriormente, que comparten similitudes genéticas con el *p*-crAssphage, así como un ancestro común, han sido clasificados bajo la denominación de "fagos tipo crAss" (***crAss-like phages***).

A raíz del desarrollo y optimización de nuevas herramientas avanzadas tras el descubrimiento de crAssphage, se llevaron a cabo análisis metagenómicos más exhaustivos de la materia oscura viral. Estos estudios permitieron identificar numerosas familias virales previamente desconocidos, entre los que se incluyen "*Flandersviridae*", "*Quimbyviridae*", "*Gratiaviridae*" (Benler et al., 2021), así como los Gubaphage (Camarillo-Guerrero et al., 2021), los fagos gigantes Lak (Devoto et al., 2019), y el fago atemperado de *Bacteroides dorei*, el Hankyphage (Benler et al., 2018).

1.3.1. Taxonomía

Antes de la implementación de los cambios en la clasificación taxonómica por parte del ICTV, la categorización de los fagos tipo crAss se basaba en características morfológicas. En este contexto, dichos fagos fueron asignados al orden Caudovirales y a la familia *Podoviridae*, sustentado en análisis predictivos de proteínas de cola (Yutin et al., 2018). Sin embargo, con

la eliminación del orden Caudovirales, la adopción de un enfoque basado en taxonomía genómica y mediante análisis filogenéticos de genes estructurales conservados (subunidad grande de la terminasa, proteína principal de la cápside y la primasa de ADN), se estableció el orden Crassvirales. Esta nueva clasificación agrupa a los fagos tipo crAss en un sistema más detallado, compuesto en el momento de la redacción de esta tesis por 4 familias, 11 subfamilias, 42 géneros y 73 especies (Turner et al., 2023).

De manera anecdótica, el origen del nombre propuesto para las 4 familias se debe a que estos virus son muy abundantes en el intestino de los mamíferos. Por esta razón, los nombres de las familias provienen de la palabra “intestino” en 4 idiomas: *Intestiviridae* (latín), *Crevaviridae* (bosnio: creva), *Suoliviridae* (finlandés: suolet) y *Steigviridae* (irlandés: stéig). En cuanto a la nomenclatura de las subfamilias, esta se basa en sinónimos del término “crAss”, que en inglés significa “grosero”. Los nombres propuestos incluyen: *Crudevirinae* (*crude*, crudo), *Obtuvirinae* (*obtuse*, obtuso), *Churivirinae* (*churish*, tosco), *Coarsevirinae* (*coarse*, áspero), *Doltivirinae* (*doltish*, torpe), *Loutivirinae* (*loutish*, grosero), *Uncouvirinae* (*uncouth*, descortés), *Bearivirinae* (*bearish*, bruto), *Boorivirinae* (*boorish*, rudo) y *Oafivirinae* (*oafish*, patoso). Para los géneros, se crearon nombres modificados a partir de razas de perro, añadiendo el sufijo “virus”, dado que estos virus, habiendo evolucionado junto con los humanos durante tanto tiempo, pueden considerarse como “el otro compañero de la especie humana”. Concretamente, *p*-crAssphage ha sido recientemente reclasificado dentro de la familia *Intestiviridae* y en la especie *Carjivirus communis.*

1.3.2. Hospedador bacteriano

Para determinar el hospedador bacteriano de los fagos tipo crAss, ya que no se había aislado el virión, se emplearon varios enfoques metodológicos independientes. En primer lugar, se realizó un análisis de espaciadores CRISPR, dado que estos contienen secuencias complementarias a fragmentos del genoma viral. Aunque no se identificaron coincidencias perfectas, las secuencias más similares se encontraron en dos especies intestinales pertenecientes al filo *Bacteroidetes*: *Prevotella intermedia* y *Bacteroides*. Adicionalmente, se aplicaron perfiles de co-ocurrencia para predecir posibles hospedadores, lo que permitió inferir que las especies del género *Bacteroides* eran los hospedadores más probables de estos fagos. Esta predicción fue respaldada por la identificación de una proteína viral que contenía dominios BACON (del inglés *Bacteroidetes*-associated carbohydrate-binding often N-terminal domain), los cuales están involucrado en la adherencia a glucoproteínas y, hasta la fecha, solo se habían encontrado en genomas de *Bacteroidetes* y en metagenomas del tracto intestinal. La presencia de estos dominios BACON, expuestos en la cápside del fago, sugiere un posible

mecanismo de interacción con su hospedador, en concordancia con la hipótesis de la adherencia de los fagos a la mucina intestinal. Según esta teoría, los fagos pueden unirse a las glucoproteínas de la mucina, lo que facilita encuentros más frecuentes con sus bacterias hospedadoras y potencia su capacidad infecciosa dentro del ecosistema intestinal (Dutilh et al., 2014; Koonin and Yutin, 2020). Antes del descubrimiento de crAssphage, (Reyes et al., 2013), en un experimento de trasplante de viroma fecal humano a ratones libres de gérmenes previamente inoculados con una comunidad bacteriana que incluía varias especies de *Bacteroides*, identificaron un *contig* (secuencia ensamblada a partir de fragmentos más cortos) que análisis genómicos posteriores mostraron una identidad del 91 % con *p*-crAssphage, lo que apunta a *Bacteroides* como el hospedador más probable para este fago. Otros estudios han sugerido posibles hospedadores bacterianos adicionales, además de los previamente propuestos. Entre estos, se han identificado mediante datos de CRISPR, especies pertenecientes a los géneros *Porphyromonas* y *Parabacteroides*, entre otros (Gulyaeva et al., 2022; Tomofuji et al., 2022).

1.3.3. Características genéticas

La caracterización de los fagos tipo crAss ha revelado una alta heterogeneidad de sus genomas. No obstante, el avance de las bases de datos de secuenciación masiva y el desarrollo de herramientas bioinformáticas especializadas, junto con el uso de criomicroscopía electrónica para la reconstrucción estructural de viriones aislados, han permitido la asignación funcional de la mayoría de las proteínas virales. Estos avances han contribuido significativamente a la comprensión del genoma de estos virus, identificando características compartidas que establecen su relación filogenética, al tiempo que los diferencian de otros grupos virales (Bayfield et al., 2023; Yutin et al., 2018).

Desde una perspectiva estructural y funcional, los fagos tipo crAss presentan dos regiones genómicas con orientaciones opuestas (Rossi et al., 2020). La primera región codifica proteínas asociadas a la replicación del ADN, mientras que la segunda, de mayor tamaño, codifica proteínas involucradas en la transcripción y la formación de nuevos viriones. Entre los genes más conservados dentro de las distintas subfamilias se encuentran aquellos que codifican proteínas estructurales de la cápside, la terminasa (que empaqueta el ADN), la proteína portal (que conecta la cápside con la cola) y la primasa (que inicia la transcripción del ADN a ARN para la posterior síntesis de proteínas) (Smith et al., 2023). Entre los hallazgos más recientes en el estudio de estos fagos, se ha descrito la proteína “muzzle” localizada en el extremo de la cola del fago, la cual presenta un plegamiento estructural previamente desconocido, denominado “crAss fold”. Se ha postulado que esta estructura actúa como un

mecanismo regulador en la expulsión de componentes virales además de facilitar la interacción con la superficie de la célula hospedadora. Adicionalmente, se ha observado que los viriones almacenan proteínas virales tanto en la cápside como dentro de la cola, una característica inusual entre los bacteriófagos conocidos (Bayfield et al., 2023). Además, se ha identificado que las espículas de la cola de estos fagos contienen dominios de unión a carbohidratos, así como posibles dominios de glucosilhidrolasas. Esto sugiere que dichas espículas podrían desempeñar un doble papel: por un lado, en el reconocimiento de los PS presentes en la superficie del hospedador y, por otro, en su despolimerización para facilitar el acceso a la membrana externa y permitir la adsorción del fago (Knecht et al., 2020).

Desde una perspectiva evolutiva, los fagos tipo crAss presentan un origen monofilético, lo que sugiere una descendencia común dentro de este grupo (Yutin et al., 2021). Además, han desarrollado un mecanismo de replicación distintivo caracterizado por la alternancia recurrente entre polimerasas de ADN tipo A y tipo B, que se cree que está mediada por recombinación homóloga (Smith et al., 2023). Asimismo, se ha identificado en su genoma una nueva ARN polimerasa (RNAP) cuya estructura guarda similitudes con la ARN polimerasa dependiente de ARN presente en eucariotas. Otra característica que los distingue de otros virus es la presencia de al menos una nucleasa de la familia PDDEXK, presuntamente involucrada en la replicación.

1.3.4. Fagos tipo crAss aislados

Para poder comprender mejor el papel biológico de estos fagos en la regulación del microbioma intestinal humano, es crucial poder propagarlos *in vitro* y aislarlos. A pesar de que más de 600 MAGs de fagos tipo crAss han sido identificados dentro del orden Crassvirales, y que son detectados en elevadas cantidades en diversos ecosistemas mediante ensayos de qPCR (García-Aljaro et al., 2017; Li et al., 2024; Liang et al., 2018; Morrison et al., 2020; Park et al., 2020; Sabar et al., 2022; Sala-Comorera et al., 2021; Stachler et al., 2017) hasta las aportaciones de esta tesis doctoral, solo se habían aislado 7 fagos tipo crAss, lo que resulta sorprendente si realmente son tan abundantes. Entre las posibles explicaciones para esta discrepancia se incluyen el uso de cepas bacterianas inadecuadas como hospedadores, la presencia de fagos más virulentos de *Bacteroides* en las muestras que podrían lisar las células bacterianas antes de que éstas permitan aislar a los fagos tipo crAss, o la aplicación de métodos convencionales de aislamiento basados en calvas de lisis, lo que resulta problemático en aquellos virus incapaces de generar calvas visibles.

El primer fago fue aislado 4 años después del descubrimiento de este grupo de fagos, el fago ΦCrAss001 (familia *Steigviridae*, especie *Kehishuvirus primarius*), obtenido a partir de muestras de heces humanas y utilizando *Bacteroides intestinalis* como hospedador (Shkoporov et al., 2018). Este fago, forma calvas de lisis visibles lo que permitió su aislamiento y apoyó la naturaleza virulenta de los fagos tipo crAss. Un par de años después, partiendo de muestras de aguas residuales, fueron aislados los fagos DAC15 y DAC17 (ambos pertenecientes a la familia *Steigviridae*, especie *Wulfhauvirus bangladeshii*) a partir de *Bacteroides thetaiotaomicron*. Remarcar que su propagación en este hospedador fue limitada, creciendo mejor en cepas mutantes con cápsulas polisacarídicas específicas (Hryckowian et al., 2020). En 2021, utilizando muestras fecales humanas, se aisló ΦCrAss002 (familia *Intestiviridae*, especie *Jahgtovirus secundus*) con *Bacteroides xylanisolvens*. Debido a su incapacidad para formar calvas de lisis, fue necesario realizar enriquecimientos consecutivos con el hospedador durante al menos 3-5 días para poder aislarlo (Guerin et al., 2021). Más recientemente, en el mismo mes en que se publicó nuestro estudio (capítulo 1 de esta tesis), Papudeshi et al., (2023) reportaron el aislamiento de 3 nuevos fagos tipo crAss de aguas residuales utilizando la técnica de la doble capa de agar y empleando *Bacteroides cellulosilyticus WH2* como hospedador. Éstos fueron asignados a las especies *Kehishuvirus winsdale* (Bc01), *Kolpuevirus frurule* (Bc03) y *Rudgehvirus redwords* (Bc11). Adicionalmente, el fago φ14:2, originalmente aislado de la bacteria acuática *Cellulophaga* (perteneciente al filo *Bacteroidetes*), que inicialmente no fue clasificado, fue posteriormente incluido como miembro del orden Crassvirales (familia *Steigviridae*, especie *Akihdevirus balticus*) (Holmfeldt et al., 2013).

En relación con la morfología de estos fagos, los análisis predictivos realizados por Yutin et al., (2018) sobre la región de la cola, sugirieron que presentan una cola corta, similar a las de los podovirus. Esta hipótesis fue confirmada mediante observaciones realizadas con microscopía electrónica de transmisión y criomicroscopía electrónica de los fagos aislados (Guerin et al., 2021; Holmfeldt et al., 2013; Hryckowian et al., 2020; Papudeshi et al., 2023; Shkoporov et al., 2018). Concretamente, los resultados mostraron que estos fagos presentan una cápside icosaédrica de entre 77 y 88 nm de diámetro y una cola corta no contráctil de entre 36 y 44 nm con espículas de diferentes tamaños.

1.3.5. Prevalencia de crAssphage en el microbioma intestinal humano

Se ha observado que las cepas de crAssphage dentro de un mismo individuo tienden a agruparse filogenéticamente, lo que sugiere la coexistencia de múltiples poblaciones de estos virus estrechamente relacionados (Edwards et al., 2019; Guerin et al., 2018). Estas

poblaciones pueden variar en abundancia y además muestran un patrón de agrupación geográfica (Edwards et al., 2019). Los mismos autores también han propuesto que crAssphage ha mantenido una estructura genómica estable a lo largo de millones de años. Este hallazgo podría generar cierta controversia respecto a la idea ampliamente aceptada de que el viroma intestinal es exclusivo de cada individuo, ya que sugiere que algunos fagos podrían estar muy conservados en el entorno del intestino humano. Esta hipótesis es respaldada por diversos estudios longitudinales sobre el viroma intestinal humano, ensayos de colonización fago-hospedador en animales libres de gérmenes y co-cultivos *in vitro*, los cuales han evidenciado una persistencia prolongada de diferentes fagos tipo crAss, incluso de hasta 4 años de persistencia de la misma cepa viral (Gulyaeva et al., 2022; Hedžet et al., 2022; Reyes et al., 2013; Shkoporov et al., 2019; Siranosian et al., 2020). Durante los ensayos en cultivos bacterianos líquidos con el fago ΦcrAss001, y a lo largo de múltiples pases seriados, no se observó la clarificación de los cultivos (Shkoporov et al., 2018). Esto sugiere una coexistencia en la que la densidad de la población hospedadora se mantiene estable, permitiendo la propagación del fago sin causar una lisis masiva del hospedador. Estos hallazgos se alinean con el modelo PtW, el cual postula que, en presencia de altas concentraciones de hospedador, los fagos pueden adoptar estrategias no destructivas o modos de replicación alternativos como la lisogenia. No obstante, en este caso se descarta la presencia de un ciclo lisogénico tradicional debido a la ausencia de: 1) profagos en bases de datos públicas de genomas bacterianos, 2) genes asociados a la lisogenia en el genoma de ΦcrAss001 y 3) lisógenos obtenidos experimentalmente (Shkoporov et al., 2021). Además, la lisogenia tiende a favorecerse en comunidades bacterianas de rápido crecimiento, estableciendo una correlación entre la prevalencia de profagos y el tiempo mínimo de duplicación celular. Este patrón es particularmente evidente en cepas de *Firmicutes* y *Proteobacteria*; en cambio, especies del filo *Bacteroidetes*, el porcentaje de lisógenos es significativamente menor (Kim and Bae, 2018), aunque se han detectado fagos atemperados de *Bacteroides*, como es el caso de Hankyphage (Shareefdeen and Hynes, 2021).

Diversos mecanismos han sido propuestos para explicar la prevalencia y la permanencia a largo plazo de los fagos tipo crAss en el intestino humano, a pesar de su carácter virulento. Uno de los sistemas podría ser la heterogeneidad espacial del hábitat microbiano intestinal, donde ciertas regiones de la mucosa resultan inaccesibles para los fagos, funcionando como refugios para las bacterias hospedadoras (Heilmann et al., 2012; Lourenço et al., 2020). Otro factor clave es la resistencia fenotípica de las bacterias, que da lugar a una población heterogénea en términos de susceptibilidad al fago sin requerir mutaciones genéticas específicas para su adquisición (Lourenço et al., 2018), como podría ser el caso de la variación de fase en la expresión de los CPS. Esta variabilidad es regulada por inversiones de ADN

catalizadas por invertasas de la familia recombinasas de serina y tirosina, lo que modula la expresión de receptores de superficie bacterianos y, en consecuencia, su sensibilidad a la infección por fagos (Coyne et al., 2003; Hedžet et al., 2022; Porter et al., 2020; Shkoporov et al., 2021). También se ha propuesto la existencia de un ciclo de vida atípico o un proceso evolutivo de ARD impulsado por la constante aparición de nuevas mutaciones tanto en los fagos como en sus hospedadores (Scanlan, 2017). En apoyo a esta hipótesis, estudios recientes sugieren un mecanismo de coevolución caracterizada por una intensa presión selectiva sobre los genes que codifican la ARN polimerasa y las proteínas de la cola del fago (Brown et al., 2021). En particular, los genes que codifican estas últimas están asociados con elementos retroactivos generadores de diversidad (Morozova et al., 2020), los cuales emplean transcriptasas inversas para introducir mutaciones selectivas en regiones específicas del ADN, dando lugar a proteínas de cola hipervariables (Guo et al., 2014).

Más en concreto, Shkoporov et al., (2021) han propuesto dos de los mecanismos mencionados que podrían operar simultáneamente (Figura 8):

- **Ciclo de infección ralentizado que retrasa la liberación de la nueva progenie viral**, que podría explicarse de tres maneras: 1) una infección en estado portador (presencia persistente de partículas virales, formadas o no, en el citoplasma del hospedador); 2) una pseudolisogenia inestable (genoma fágico en forma episomal), o bien 3) la inhibición de la propagación del fago a través de mecanismos celulares.
- **Adquisición dinámica y reversible de resistencia al fago dependiente de la variación de fase de los CPS**, lo que permitiría una replicación fágica restringida. Como resultado, el fago podría persistir a largo plazo gracias a un equilibrio dinámico y espontáneo entre diferentes tipos de CPS, los cuales determinan la sensibilidad o resistencia de las células de *Bacteroides.* Este equilibrio podría estar impulsado por la presión selectiva ejercida por múltiples fagos. En este sentido, la expresión de ciertos CPS que confieren protección contra la infección de un fago podría, a su vez, hacer que las células sean sensibles a otro fago presente en el mismo entorno (Porter et al., 2020).

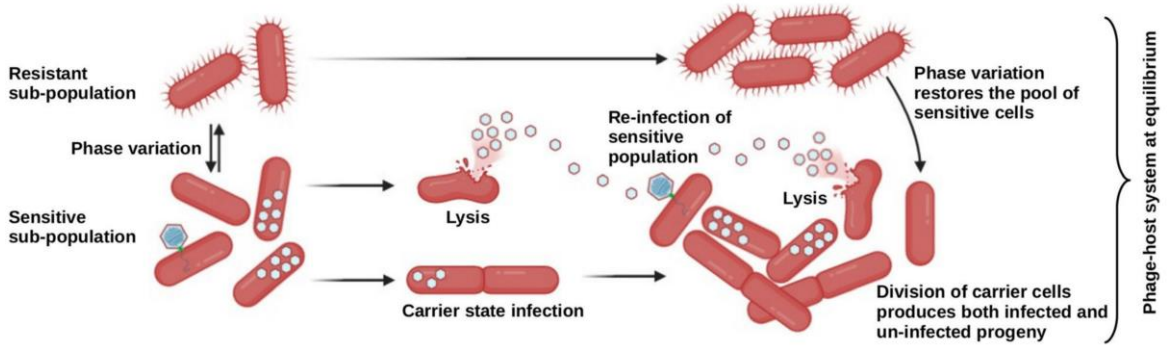


Figura 8. Modelo de co-exsistencia de fagos tipo crAss en su hospedador *Bacteroides.* Figura extraída de Shkoporov et al., (2021), bajo licencia CC BY 4.0. No se realizaron modificaciones.

Esta segunda hipótesis resulta altamente plausible, dado que el género *Bacteroides* presenta una notable abundancia de sistemas de variación de fase (Ben-Assa et al., 2020). En este sentido, el estudio de Krinos et al., (2001) sobre los mecanismos que usa *Bacteroides fragilis* para evadir el sistema inmunológico del hospedador, demostró que estos microorganismos son capaces de modular la expresión de ciertos genes mediante un sistema de encendido y apagado (ON/OFF). Este proceso se basa en inversiones reversibles (como giros de 180º) de fragmentos de ADN, en las cuales los promotores asociados, ubicados junto o dentro de la **región con variación de fase** (en inglés, *Phase Variation Región*, PVR) se orientan alineados con los genes, promoviendo su expresión (ON) o, por el contrario, giran en el sentido opuesto, no alineándose con sus genes, evitando su expresión (OFF). Entre las zonas reguladas por fases variables se incluyen aquellas involucradas en la biosíntesis de CPS, así como en la producción de un PS extracelular, fimbrias, glicoproteínas de la superficie externa, vesículas de membrana externa o también, en la síntesis de proteínas involucradas en sistemas de R-M (Ben-Assa et al., 2020; Chatzidaki-Livanis et al., 2008; Coyne and Comstock, 2016; Fletcher et al., 2007; Nakayama-Imaohji et al., 2016; Porter et al., 2020).

En particular, se ha demostrado que las especies del género *Bacteroides* contienen un total de ocho *loci* responsables de la síntesis de CPS, que en *B. fragilis* se han denominado alfabéticamente de PSA a PSH (polisacárido A-H). Cada CPS presenta una organización genética similar, lo que implica que cada uno representa un operón compuesto entre 11 y 22 genes. De estos ocho *loci*, siete (con excepción del *locus* PSC) presentan regiones PVR que contienen el promotor y que están flanqueadas por repeticiones invertidas de alrededor de 20 pb que permiten la variación de fase, lo que regula la orientación del promotor y, por consiguiente, la expresión de los genes involucrados en la biosíntesis de los PS (Chatzidaki-Livanis et al., 2010; Krinos et al., 2001). La alternancia entre los fenotipos “ON” y “OFF” está

influenciada por las condiciones ambientales, pero en general se cree que son de naturaleza estocástica.

Los genes de cada operón están situados aguas arriba de la PVR. Los dos primeros genes codifican para proteínas pertenecientes a las familias UpxY y UpxZ (donde la "x" varía dependiendo del PS específico, desde Upa a Uph). UpxY, miembro de la familia de proteínas tipo NusG, actúa como un factor de antiterminación transcripcional al unirse a la ARN polimerasa, facilitando así la transcripción de los siguientes genes de cada operón. A continuación, se encuentra el gen que codifica para la proteína UpxZ, una proteína sin ortólogos conocidos, que tiene como función interferir con la acción de las UpxY de otros CPS, inhibiendo así la síntesis de esos otros CPS. Esta proteína exhibe diversas capacidades inhibitorias, lo que permite regular de manera jerárquica la biosíntesis de los diferentes CPS. Como resultado de este mecanismo, cada célula prioriza la expresión de un solo CPS, lo que contribuye a la plasticidad fenotípica de las poblaciones bacterianas. Esto se traduce en la existencia de células con distintas combinaciones de CPS en su superficie dentro de una misma población, lo que se cree que les otorga una ventaja adaptativa al entorno (Chatzidaki-Livanis et al., 2010; Saba et al., 2024).

1.3.6. Posibles usos de los fagos tipo crAss

- **Uso como biomarcador de patologías:** en situaciones de disbiosis, como las observadas en patologías intestinales, tales como la colitis ulcerosa y la enfermedad de Crohn, el conocimiento sobre la dinámica de los fagos es limitado. Actualmente, no está claro si los fagos siguen un modelo generalizado o si cada enfermedad presenta una dinámica específica y diferenciada. Sin embargo, estudios han evidenciado una mayor prevalencia de fagos pertenecientes a la antigua clasificación Caudovirales en adultos con colitis ulcerosa, en contraste con lo observado en pacientes con enfermedad de Crohn. Este hallazgo sugiere que los fagos podrían funcionar como potenciales biomarcadores para la distinción entre diversas patologías intestinales (Norman et al., 2015).

  En lo que se refiere a los fagos tipo crAss, estudios de prevalencia no han identificado asociaciones perjudiciales entre estos virus y la salud humana, lo que sugiere que pueden considerarse entidades benignas dentro del microbioma intestinal. De hecho, se ha observado que una reducción en la abundancia y diversidad de estos fagos podría estar relacionada con diversas condiciones, como por ejemplo la obesidad, el síndrome metabólico o enfermedades inflamatorias, entre otros (Cervantes-Echeverría et al., 2023; Chu et al., 2023; Clooney et al., 2019; Draper et al., 2018; Tomofuji et al., 2022). Sin

embargo, el análisis de la abundancia de los hospedadores bacterianos de los fagos tipo crAss revela que su concentración también disminuye en estos trastornos, lo que podría estar vinculado a una menor estabilidad de crAssphage. Por otro lado, en algunos casos, los niveles de fagos tipo crAss se ven aumentados, como ocurre en pacientes con colitis ulcerosa, en comparación con individuos sanos, mientras que la población de *Bacteroides* disminuye (Yang et al., 2020). O bien, su presencia no se correlaciona con pacientes que presentan un cuadro de diarrea (Liang et al., 2016).

- **Uso como marcador de contaminación fecal**: se ha establecido que los fagos tipo crAss poseen una distribución global y se encuentran en alta abundancia en heces humanas, mientras que su presencia en muestras contaminadas con heces de origen animal, desde primates hasta felinos, es significativamente menor (Edwards et al., 2019; García-Aljaro et al., 2017; Li et al., 2021; Malla et al., 2019), lo cual permite la capacidad de distinguir una contaminación fecal de origen humano respecto a una contaminación fecal de origen animal. Estas evidencias han llevado a la propuesta de estos virus como indicadores universales de contaminación fecal humana (García-Aljaro et al., 2017; Sala-Comorera et al., 2021; Stachler et al., 2017), aunque la variabilidad genética sea notable en el orden Crassvirales (Edwards et al., 2019; Guerin et al., 2018; Koonin and Yutin, 2020).

1.4. La contaminación fecal del agua

El agua constituye un recurso esencial para la vida, y su calidad es un factor determinante en la salud pública. Las actividades humanas han generado un impacto significativo en los ecosistemas acuáticos, debido a la liberación de contaminantes orgánicos e inorgánicos, microorganismos, xenobióticos, metales pesados, residuos de petróleo y otras sustancias nocivas. Estos contaminantes provienen, en gran medida, de la descarga directa de desechos fecales o aguas residuales no tratadas, así como de la escorrentía urbana y agrícola, los desbordamientos de alcantarillado, los vertidos industriales y las actividades ganaderas (Figura 9), lo que contribuye al deterioro de la calidad del agua y a la alteración de los sistemas hídricos (Caruso, 2013).

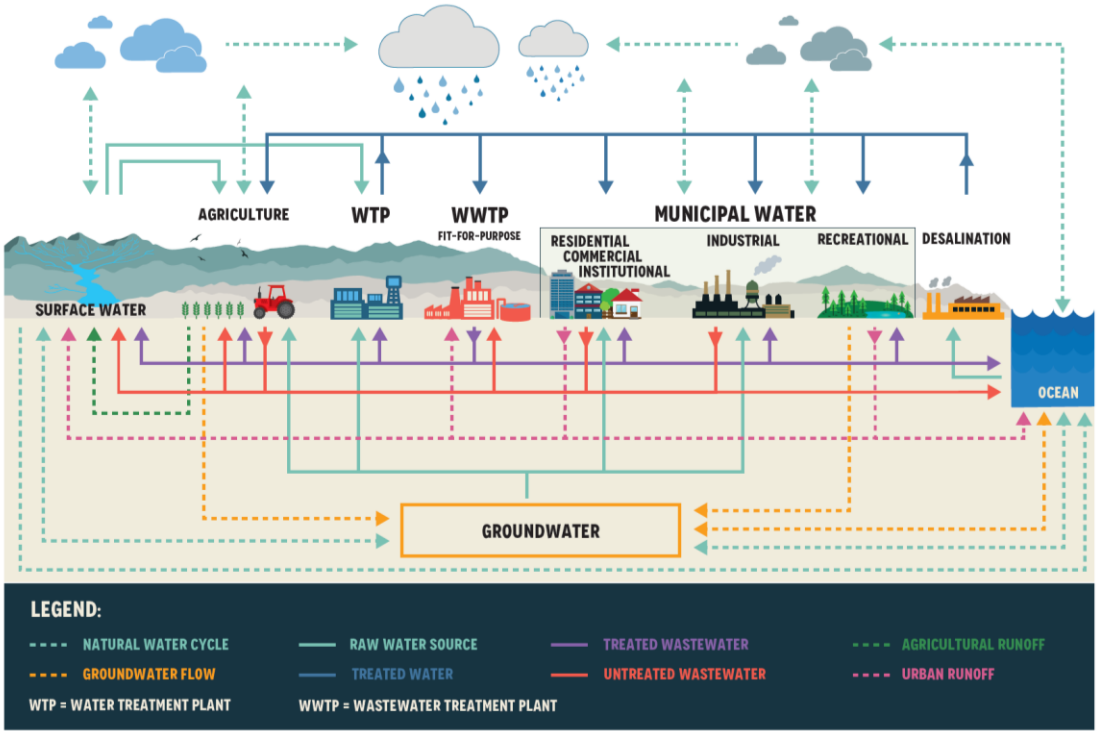


Figura 9. Ciclo del agua: usos y tratamientos. Figura extraída de UNESCO (2017), bajo licencia CC BY-SA 3.0 IGO. No se realizaron modificaciones.

En particular, la creciente población, la industrialización, la globalización y el cambio climático han incrementado el riesgo de transmisión de múltiples enfermedades (Hernando-Amado et al., 2019; Rodó et al., 2013; Sharma et al., 2003) debido principalmente a la contaminación fecal. Se estima que el agua en condiciones insalubres es responsable del 50% de las enfermedades a nivel mundial y contribuye a la muerte de millones de personas cada año (Wen et al., 2020; World Health Organization, 2019). Por lo tanto, garantizar el acceso a agua potable segura es una prioridad fundamental para la salud pública a nivel global. En este

contexto, el Servicio de Salud Pública de los Estados Unidos implementó en 1914 las primeras regulaciones formales para la calidad microbiológica del agua de consumo, estableciendo así un precedente en la protección sanitaria (Teaf et al., 2018). En la actualidad, en el marco de la Directiva 2000/60/CE de la Unión Europea, conocida como la Directiva Marco del Agua (DMA), que estableció un marco comunitario de actuación para la política de aguas, muchos países, entre ellos España, mediante el Real Decreto Legislativo 1/2001, donde se incorporan las bases para su aplicación, han priorizado la mejora del saneamiento del agua y la reducción del vertido de aguas residuales en el entorno natural. Este enfoque se desarrolla bajo el paradigma de **One Health**, promoviendo estrategias de gestión integrados que reconocen la interconexión entre la salud humana, animal y ambiental (World Health Organization, 2019).

Las aguas residuales, definidas como todos aquellos efluentes líquidos y sólidos relacionados con la actividad humana (aguas domésticas, industriales, etc) que son vertidos al sistema de saneamiento y conducidos a través de colectores residuales para su tratamiento en estaciones de depuración de aguas residuales (EDAR), han sido ampliamente reconocidas como un vector significativo en la propagación de diversas enfermedades infecciosas, incluyendo el cólera y la fiebre tifoidea, entre otras (Kilaru et al., 2023; Mari et al., 2012; McCall et al., 2020; Shuval, 1993).

La implementación de medidas eficaces de tratamiento y vigilancia de aguas residuales se ha vuelto indispensable para mitigar la diseminación de patógenos. Además, la epidemiología basada en el análisis de esta matriz ha emergido como una herramienta clave para la identificación y monitoreo de microorganismos dentro de la compleja comunidad microbiológica presente en las aguas residuales, permitiendo una mejor comprensión de los riesgos asociados y el desarrollo de estrategias de control más efectivas (Kilaru et al., 2023; McCall et al., 2020; Sims and Kasprzyk-Hordern, 2020). En este sentido, resulta esencial disponer de una solución efectiva que incremente la sensibilidad en la detección de contaminación fecal, complementando a los indicadores tradicionales (apartado 1.4.1), a lo largo de todo el ciclo del agua, desde su captación, potabilización y distribución, hasta la recolección de los efluentes residuales, su depuración, regeneración y posterior retorno al medio ambiente.

No obstante, con frecuencia, la gestión de aguas residuales recibe poca atención social en comparación con los desafíos asociados al abastecimiento de agua, aunque ambos aspectos están muy relacionados (UNESCO, 2017). Una descarga inadecuada de esta matriz puede constituir una fuente importante de contaminación no solo del agua potable, sino también vías ambientales adicionales, como las recreacionales (Marsalek and Rochfort, 2004), alimentarias o incluso superficies como las manos u otros utensilios (Kotay et al., 2017). Por ejemplo,

utilizar aguas para la irrigación de campos de cultivo con una calidad insuficiente debido a un tratamiento de depuración poco efectivo, puede resultar en la contaminación de alimentos (Katzenelson et al., 1976; Yadav et al., 2023). Asimismo, la liberación de aguas residuales crudas o parcialmente tratadas al mar, puede comprometer la seguridad alimentaria de productos marinos como el marisco (Bryan, 1977; Oliveira et al., 2011; World Health Organization, 2003). Por lo tanto, el vertido de residuos fecales representa una de las principales causas de contaminación de los recursos hídricos, debido a la presencia de grandes cantidades de microorganismos, entre los cuales también se incluyen patógenos excretados por individuos infectados. Una de las rutas de infección más ocurrentes es la fecal-oral mediante la ingestión de agua o alimentos contaminados, pero también existen otras posibles vías de transmisión hídrica de patógenos, como la generación de aerosoles, el contacto directo (por ejemplo, a través de una herida) o bien mediante vectores de transmisión de enfermedades, como ocurre con la malaria (World Health Organization, 2017) (Figura 10).

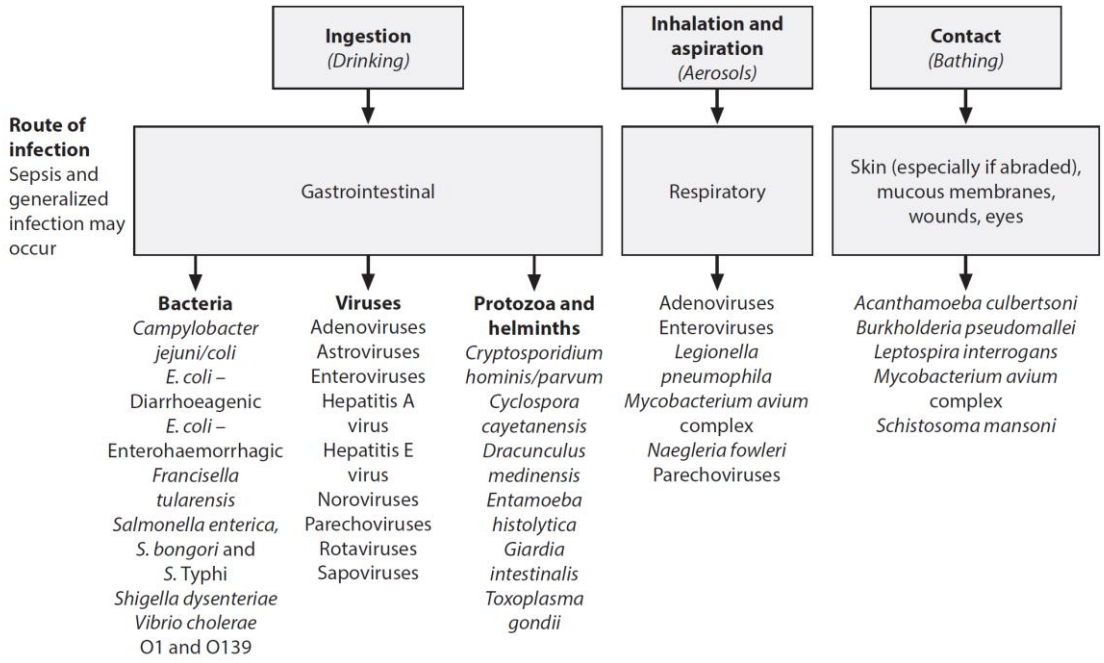


Figura 10. Rutas de transmisión hídrica que siguen varios patógenos. Figura extraída de World Health Organización (2017), bajo licencia CC BY-NC-SA 3.0 IGO. No se realizaron modificaciones.

El tratamiento de aguas residuales tiene como objetivo la eliminación de contaminantes físicos, químicos y microbiológicos con la finalidad de generar un efluente suficientemente limpio para la posterior reutilización. Sin embargo, desde la perspectiva de la preservación de la salud humana y la prevención de enfermedades de transmisión hídrica, la principal prioridad

radica en la eliminación de microorganismos patógenos, asegurando el cumplimiento de los estándares nacionales y/o internacionales de calidad de los efluentes (Momba et al., 2019).

La presencia de patógenos en la población es relativamente baja, con menos del 1% de las personas infectadas en un momento concreto, salvo en caso de brotes o en zonas donde los patógenos son endémicos. No obstante, un individuo infectado puede excretar, durante un período de tiempo limitado, una cantidad significativa de microorganismos patógenos en sus heces, dependiendo del tipo de patógeno y de la condición del individuo afectado (Payment and Locas, 2011). Por ello, su detección en muestras depende de su circulación en la comunidad en el momento del muestreo. En este contexto, es necesario detectar y cuantificar los patógenos que siguen la vía fecal-oral, la cual consta de 3 fases principales (Figura 11). En primer lugar, el patógeno es excretado al ambiente a través de las heces de humanos y animales, siendo la cantidad excretada y la ocurrencia en la población un factor determinante para su concentración final en el medio. En segundo lugar, el patógeno alcanza una fuente de agua, un entorno significativamente diferente al del tracto intestinal, diluyéndose en ella. Allí debe resistir diversas condiciones ambientales que pueden afectar su viabilidad, como la exposición a la luz ultravioleta, fluctuaciones de temperatura, salinidad e interacciones bióticas con la microbiota autóctona, además de los posibles tratamientos implementados al agua. La capacidad del microorganismo para persistir en este ambiente es crucial, ya que cuanto mayor sea su tiempo de supervivencia, mayores serán las probabilidades de encontrar un nuevo hospedador susceptible a la infección. Durante esta fase, los patógenos pueden permanecer en estado planctónico, depositarse en sedimentos o adherirse a la vegetación acuática, entre otros mecanismos de supervivencia. En la última fase, el microorganismo es ingerido nuevamente a través del consumo de agua contaminada o de alimentos expuestos a aguas fecalmente contaminadas, facilitando así su propagación (Atwill, 2005; Korajkic et al., 2019; Naughton and Mihelcic, 2017).

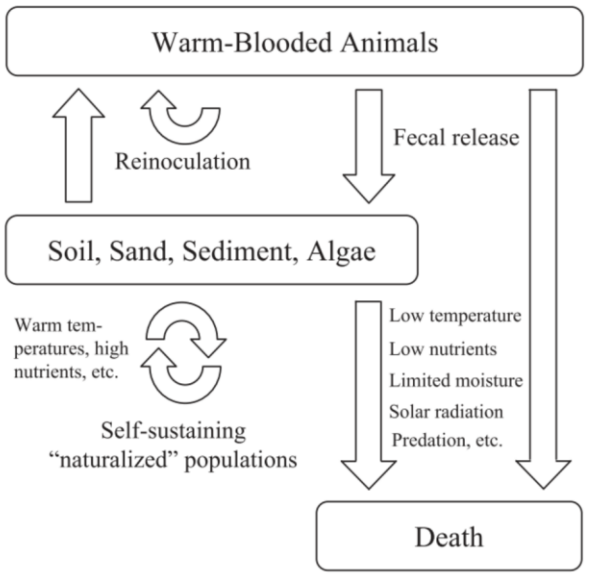


Figura 11. Fases del ciclo fecal-oral de *E. coli*, similar al de los microorganimos patógenos fecales (Ishii and Sadowsky, 2008).

La persistencia de los patógenos en distintos ambientes acuáticos es altamente variable entre especies microbianas, lo que complica su monitoreo y cuantificación. Los métodos empleados para la detección de estos microorganismos incluyen tanto técnicas de cultivo como enfoques moleculares, siendo ambas complementarias (Pinar-Méndez et al., 2022). Sin embargo, debido a la amplia diversidad de patógenos que existen, estos procedimientos pueden variar según el patógeno en cuestión y resultar costosos, prolongados y, en algunos casos, pueden sobreestimar la cantidad de patógenos viables o presentar limitaciones en la detección dadas sus distribuciones irregulares y bajas concentraciones en aguas (Field and Samadpour, 2007), incluso cuando se utilizan muestras de heces frescas (Girones et al., 2010; Korajkic et al., 2019; Liu et al., 2014; Ramírez-Castillo et al., 2015). Ante estas dificultades, el uso de microorganismos indicadores se ha establecido como una estrategia válida para la identificación de contaminación fecal y la evaluación de la seguridad microbiológica. A pesar de ciertas limitaciones, su aplicación está ampliamente reconocida y aceptada en el ámbito científico como en el regulatorio, facilitando un monitoreo más eficiente y acorde con los estándares establecidos por las normativas vigentes en cada país, cuyos límites varían según el destino del agua, siendo más estrictos en caso del agua de consumo.

1.4.1. Indicadores de contaminación fecal

El uso de los microorganismos indicadores de contaminación fecal permite, a través de un único análisis, detectar la presencia o ausencia de un determinado tipo de contaminación

fecal, así como evaluar la eficacia de los procesos de desinfección y depuración, siendo más viable y económico que el análisis de todos los microorganismos potencialmente peligrosos (Meays et al., 2004). Esto, a su vez, permite inferir la posible presencia de patógenos de origen fecal en el medio analizado.

Según la Organización Mundial de la Salud (OMS), un microorganismo indicador adecuado debe cumplir con los siguientes criterios fundamentales (World Health Organization, 2017):

- Ser detectable mediante técnicas simples, rápidas y de bajo coste.
- Estar siempre presente en altas concentraciones en heces de humanos y animales y, por ende, ser excretados en grandes cantidades a las aguas residuales.
- Presentar recuentos superiores a los de los patógenos fecales.
- Ser incapaces de multiplicar en el ambiente.
- Exhibir una persistencia e inactivación similares a los patógenos fecales.
- No ser patogénico, garantizando así una manipulación segura por parte de los usuarios.
- Mantener una presencia constante en heces independientemente del estado de salud del hospedador.

El uso de estos indicadores se implementó inicialmente como una herramienta para alertar sobre un mayor riesgo de presencia de una contaminación fecal, comenzando con el monitoreo del agua potable. Con el tiempo, su aplicación se expandió a efluentes de aguas residuales, cuerpos de agua ambientales y el control de la seguridad alimentaria. Entre los indicadores más utilizados y mejor comprendidos se encuentran los bacterianos (bacterias indicadoras fecales, BIF), como los coliformes totales, coliformes termotolerantes o fecales, *Enterococcus* spp. y *Clostridium perfringens.* Estos microorganismos, que habitan en grandes cantidades en el tracto gastrointestinal de animales de sangre caliente (Murray and Baron, 2007), se excretan de manera recurrente en las heces (Rochelle-Newall et al., 2015).

En términos generales, la detección de patógenos no se asocia con el concepto de indicador, ya que cada patógeno es representativo únicamente de su propia presencia y, su ausencia no implica necesariamente la ausencia de otros microorganismos patógenos (Payment and Locas, 2011). En particular, es importante destacar que la medición de indicadores bacterianos no ha demostrado una correlación consistente con la presencia de patógenos, principalmente debido a las diferencias significativas en su fisiología y filogenia (Harwood et al., 2014).

Un agua contaminada fecalmente no solo puede albergar bacterias, sino también patógenos virales (principales causantes de gastroenteritis a nivel mundial) (Liu et al., 2016; Platts-Mills et al., 2015; Troeger et al., 2018) y protozoos, como *Giardia* y *Cryptosporidium*, entre otros

(Harwood et al., 2005; Jurzik et al., 2010). Estos microorganismos presentan una mayor persistencia, sobreviviendo más tiempo en el ambiente, incluso bajo condiciones ambientales adversas o durante procesos de desinfección, como los utilizados en algunas EDARs o en las estaciones de reutilización de agua (ERA). Esto ocurre incluso cuando las bacterias ya han sido inactivadas, y las aguas cumplen con los estándares de calidad para bacterias coliformes (Anderson et al., 2005; Grabow, 2001; IAWPRC Study group on health related water microbiology, 1991; Jofre, 2007; Lin and Ganesh, 2013). Además, establecer una relación directa entre la concentración de indicadores bacterianos de contaminación fecal y el riesgo real para la salud humana resulta complicado, ya que la mayoría de los animales también excretan estas bacterias en sus heces, lo que genera una amplia distribución entre diferentes especies hospedadoras (Harwood et al., 2014). Por ejemplo, se ha identificado una posible relación en fuentes como el ganado bovino, debido a la alta prevalencia de patógenos zoonóticos (Soller et al., 2010). Además, las bacterias pueden presentar otras limitaciones, como el hecho de que no siempre provienen exclusivamente de fuentes fecales y, bajo ciertas condiciones, pueden encontrarse o multiplicarse en el ambiente (Byappanahalli et al., 2012; Ferguson and Signoretto, 2011; Winfield and Groisman, 2003). Por esta razón, ha sido necesario identificar nuevos indicadores más apropiados para poder correlacionar la presencia de contaminación fecal de origen vírico. Entre las alternativas propuestas se encuentran (1) los virus entéricos humanos, aunque al ser patógenos no cumplen todos los requisitos de los indicadores, y (2) los bacteriófagos.

1.4.2. Bacteriófagos como microorganismos indicadores

Los bacteriófagos presentan varias ventajas frente a los indicadores bacterianos. Son más abundantes, poseen características morfológicas similares a las de muchos virus entéricos, y por tanto una persistencia más parecida a éstos, y se excretan al medio ambiente de manera comparable a los patógenos virales. Por estas razones su uso permite una detección más precisa de la contaminación vírica de origen fecal (Ballesté et al., 2021; Grabow, 2001; Jofre et al., 1986, 2014; Toribio-Avedillo et al., 2021). Como consecuencia, en los últimos 20 años algunos grupos de bacteriófagos han sido incluidos en las directrices sobre la calidad del agua en diversos países (North Carolina Administration, 2011; República de Colombia, 2014; The European Parliament and the Council of the European Union Regulation (EU), 2020; U.S. EPA, 2016; Western Australian Government, 2012).

En la actualidad, los fagos se emplean para diversas finalidades. Una de ellas es su uso como indicadores de proceso (Amarasiri et al., 2017), ya que permiten evaluar la eficacia de los tratamientos de desinfección. Un ejemplo es su aplicación para evaluar la eficacia de los

tratamientos aplicados en las EDAR, que pueden verse afectados por fallos operativos, variaciones en la cantidad y calidad de agua, así como por fluctuaciones estacionales que influyen en el rendimiento de la planta (Momba et al., 2019). Otra finalidad es su uso como indicadores fecales, ya que su aplicación contribuye a estimar el riesgo para la salud asociado a la contaminación fecal y a la posible presencia de virus entéricos patógenos, tanto en matrices hídricas como agua potable, aguas residuales, aguas recreativas y aguas subterráneas, así como matrices sólidas como suelos, sedimentos y alimentos (Guelin, 1948; Hyman and Abedon, 2009; IAWPRC Study group on health related water microbiology, 1991).

Los bacteriófagos, capaces de infectar bacterias entéricas, más comúnmente utilizados como indicadores de contaminación fecal son los siguientes:

- **Colifagos somáticos**: grupo heterogéneo de virus capaces de infectar a *E. coli* y otras bacterias coliformes a través de la pared celular, tras unirse a receptores específicos en la membrana externa (Muniesa et al., 2003). Constituyen el grupo de bacteriófagos indicadores mayoritario en casi todas las muestras ambientales y presentan una resistencia superior a la de indicadores bacterianos (Jofre et al., 2016).
- **Colifagos ARN F-específicos:** infectan *E. coli* y otras bacterias coliformes mediante la adhesión a los pili sexuales de la célula hospedadora. Representan el segundo grupo más común en el entorno ambiental (Jofre et al., 2016). Ciertos genotipos de este grupo de fago son específicos de humanos (genotipos II y III) mientras otros genotipos son más frecuentes en animales (genotipos I y IV, predominantes, pero no exclusivos) (Schaper et al., 2002). Por lo tanto, el genotipo de estos fagos puede detectarse mediante hibridación o por qPCR.
- **Fagos que infectan especies de *Bacteroides***. Suelen estar presentes en concentraciones más bajas en comparación con los colifagos. La mayoría exhiben una morfología tipo siphovirus, caracterizadas por una cápside icosaédrica, una cola larga y flexible, y ADN de doble cadena. Tienen un rango de infección limitado, ya que se adhieren a receptores específicos en la pared celular de su hospedador (Puig et al., 2001).

Algunos de estos fagos, debido a la alta especificidad hacia hospedadores bacterianos de origen exclusivamente humano o animal, se utilizan como método microbiano para trazar el origen de una contaminación fecal (*Microbial source tracking* o MST) (Harwood et al., 2014; Jebri et al., 2017; Jofre et al., 2014).

1.4.3. Microbial Source Tracking

El MST engloba un conjunto de métodos que surgieron a finales del siglo XX (Wiggins, 1996) y que fueron diseñados para diferenciar entre fuentes humanas y no humanas de contaminación fecal. Una vez determinado el origen de la contaminación, es posible estimar los riesgos sanitarios asociados y sobre todo aplicar medidas de remediación dirigidas a abordar el problema desde su origen (Farnleitner and Blanch, 2017). Localizar y rastrear la fuente de contaminación fecal no es una tarea simple, pero contribuye a una gestión más eficiente y ayuda a resguardar la salud pública (Blanch et al., 2004; Jagals and Grabow, 1996). Estos métodos se han aplicado para determinar el origen de restos fecales humanos, pero también animales, como diversos grupos de ganado y animales de granja (vacas, cerdos, pollos) (Gómez-Doñate et al., 2011; Weidhaas et al., 2010), mascotas (Kildare et al., 2007) o animales salvajes (Jiang et al., 2007; Nguyen et al., 2018; Sinigalliano et al., 2010). Dentro de los métodos de MST, existen varios basados en bacteriófagos, como la distribución de serotipos/genotipos de fagos F-específicos de ARN, fagos que infectan especies de *Bacteroides* y fagos que infectan cepas específicas de *Enterococcus* (Jofre et al., 2014).

Otros indicadores tradicionales, como los colifagos o las BIF, son ampliamente utilizados en el monitoreo de la calidad del agua, pero no permiten, en general, identificar el origen de la contaminación, ya que están presentes en el intestino tanto de humanos como de muchos otros animales. En contraste, los fagos de *Bacteroides,* empleando cepas hospedadoras específicas, pueden proporcionar una mayor precisión en la determinación del origen de la contaminación fecal (Holcomb and Stewart, 2020). En concreto, debido a la estrecha relación entre ciertas cepas de *Bacteroides* y los hospedadores humanos o animales, los fagos que infectan cepas exclusivamente humanas o animales también reflejan esta especificidad, lo que los convierte en herramientas valiosas para ser empleados en MST (Grabow et al., 1995; Jofre et al., 2014; Kator and Rhodes, 1992; Muniesa et al., 2012; Payan et al., 2005; Tartera et al., 1988).

Además, los fagos de *Bacteroides* presentan una notable persistencia ambiental, incluso en condiciones extremas y ante tratamientos de desinfección con cloro, temperaturas elevadas, radiación UV, desinfección química y tratamientos aplicados a lodos, entre otros (Costán-Longares et al., 2008; Durán et al., 2003; Guzmán et al., 2007; Mocé-Llivina et al., 2003; Tartera et al., 1988). También muestran una lenta inactivación en matrices hídricas como agua dulce (Durán et al., 2002), agua de mar (Mclaughlin and Rose, 2006; Mocé-Llivina et al., 2005) o sedimentos (Chung and Sobsey, 1993), lo que les permite representar de forma más precisa el comportamiento de microorganismos fecales más persistentes.

Hace años se sugirió que los colifagos somáticos podrían replicarse en ambientes acuáticos (Grabow, 2001), comprometiendo su uso como indicadores. Posteriormente se demostró que esto es improbable y solo podría suceder en circunstancias excepcionales (Jofre, 2009). En cualquier caso, no se cuestiona que los fagos que infectan a *Bacteroides* tengan capacidad de multiplicarse fuera del tracto intestinal, debido a los requisitos de anaerobiosis de su hospedador. Por lo tanto, tal como indican diversos estudios, los fagos de *Bacteroides* que infectan a cepas específicas muestran un buen potencial como posibles marcadores de contaminación fecal humana (Ebdon et al., 2007; Jofre et al., 1986; Mcminn et al., 2014; Payan et al., 2005).

Sin embargo, la aplicación de fagos que infectan a *Bacteroides* presenta ciertas limitaciones. Entre ellas, la necesidad de realizar cultivos anaerobios para el crecimiento de sus cepas hospedadoras, lo que dificulta su implementación en análisis rutinarios. Asimismo, la composición de la microbiota intestinal varía considerablemente según factores geográficos (con variaciones en la dieta), los factores genéticos del hospedador y los factores ambientales. Esto implica que se requiere el aislamiento de una cepa de *Bacteroides* adecuada y específica de humanos o animales distinta para ser usada en cada región geográfica (Ballesté et al., 2019; Stewart et al., 2013), lo que complica aún más su aplicación. Se han detectado fagos de la cepa *B. fragilis* RYC2056 en aguas residuales de diferentes partes del mundo; sin embargo, estos fagos están asociados tanto a contaminación humana como animal, lo que los convierte en un candidato poco adecuado para MST (Jofre et al., 2014). Además, su baja concentración en aguas residuales (Tabla 1) (Jofre et al., 2014) contrasta con los criterios ideales para un indicador eficiente, aunque existen métodos para la concentración de estos fagos en caso de niveles bajos de contaminación (Méndez et al., 2004). En cambio, el análisis de fagos infecciosos para cepas específicas humanas de *B. fragilis* HSP40 (Tartera et al., 1989), *B. thetaiotaomicron* GA17 (Payan et al., 2005), *B. fragilis* HB13 (Payan et al., 2005) o *B. ovatus* GB-124 (Ebdon et al., 2007) ha ofrecido buenos resultados en la detección específica de la contaminación fecal humana de la zona geográfica donde se aislaron (Blanch et al., 2006, 2004; Harwood et al., 2013; Tartera and Jofre, 1987). Sin embargo, estas cepas no son útiles de forma universal, siendo unas cepas más adecuadas para ciertos países que otras (Blanch et al., 2006).

Tabla 1. Concentraciones de diversos fagos que infectan especies de *Bacteroides*, colifagos y fagos ARN F-específicos en aguas residuales, según diferentes áreas geográficas. Adaptada de Lucena et al., (2004), Jofre et al., (2014) y Remesh and Viswanathan (2024).

Fago	Área geográfica	Fagos de *Bacteroides* en aguas residuales (UFP/100 ml)*	Referencias
Fago de *B. fragilis* HSP40	España, Francia, Israel, Sud África	$5 \cdot 10^3$ - 10^4	(Armon and Kott, 1995; Gantzer et al., 1998; Grabow et al., 1993; Puig et al., 1999; Tartera et al., 1989; Tartera and Jofre, 1987)
	USA, Gran Bretaña, Suecia, Tailandia	$<10^2$	(Kator and Rhodes, 1992; Puig et al., 1999; Sirikanchana, 2013)
Fago de *B. tethaiotaomicron* GA17	España, Francia, Chipre, Colombia, Túnez	$10^4 – 5 \cdot 10^5$	(Blanch et al., 2006; Jebri and Yahya, 2013; Payan et al., 2005)
	Reino Unido y Suecia	$\sim 3 \cdot 10^{3A}$	(Payan et al., 2005)
Fago de *B. ovatus*. GB124	Reino Unido	$10^2 – 5 \cdot 10^{4B}$	(Ebdon et al., 2007; Ebdon et al., 2012; Payan et al., 2005)
	España	$\sim 10^{4A}$	(Payan et al., 2005)
Fago de *B. fragilis* HB13	España	$5 \cdot 10^4 – 5 \cdot 10^5$	(Payan, 2006; Payan et al., 2005)
	Colombia	$5 \cdot 10^3 – 1 \cdot 10^4$	(Campos, 2013)
Fago de *Bacteroides* sp. ARABA 84	Suiza	$1 – 5x10^3$	(Wicki et al., 2011)
Fago de *B. fragilis* RYC2056	Argentina Colombia Francia España	$\sim 1x10^4$ $\sim 6x10^3$ $\sim 4x10^4$ $\sim 5x10^4$	(Lucena et al., 2004)
p-crAssphage [C]	España	$\sim 3 \cdot 10^6 – 10^8$	(Cuevas-Ferrando et al., 2022; García-Aljaro et al., 2017)
	Italia	$\sim 5 \cdot 10^3 – 2 \cdot 10^7$	(Crank et al., 2020)
	USA	$\sim 1 \cdot 10^6 – 2 \cdot 10^9$	(Korajkic et al., 2019; Papp et al., 2020; Tandukar et al., 2020; Wu et al., 2023; Wu et al., 2020)
	Chile	$\sim 2 \cdot 10^8$	(Jennings et al., 2020)
	Irlanda	$\sim 7 \cdot 10^4 – 5 \cdot 10^7$	(Sala-Comorera et al., 2021)
	Australia	$\sim 9 \cdot 10^5 – 2 \cdot 10^{10}$	(Ahmed et al., 2022; Holm et al., 2022)
	Tailandia	$\sim 6 \cdot 10^6 – 2 \cdot 10^7$	(Kongprajug et al., 2019; Sangkaew et al., 2021)
	Reino Unido	$\sim 6 \cdot 10^6 – 2 \cdot 10^7$	(Farkas et al., 2019)
	Japón	$\sim 9 \cdot 10^9 – 1 \cdot 10^{11}$	(Malla et al., 2019)
	Egipto	$\sim 1 \cdot 10^3 – 8 \cdot 10^5$	(Hamza and Abd-Elmaksoud, 2023)

Fago	Área geográfica	Fagos de *Bacteroides* en aguas residuales (UFP/100 ml)*	Referencias
Otros fagos indicadores			
Colifagos somáticos	Argentina Colombia Francia España	~ $6 \cdot 10^5$ ~ $6 \cdot 10^5$ ~ $1 \cdot 10^6$ ~ $1 \cdot 10^7$	(Lucena et al., 2004)
Fagos ARN F-específicos	Argentina Colombia Francia España	~ $7 \cdot 10^4$ ~ $2 \cdot 10^5$ ~ $5 \cdot 10^5$ ~ $7 \cdot 10^5$	(Lucena et al., 2004)

[A] Difícil cuantificación debido a la turbidez de las placas
[B] Concentraciones en efluentes secundarios
[C] Valores en copias genómicas (CG)/100mL, únicamente detectado por métodos moleculares.
*UFP: unidades formadoras de placas

Es cierto que cada vez se emplean técnicas más avanzadas, como la NGS, para identificar nuevos marcadores de MST, logrando incluso distinguir entre contaminación fecal de niños y adultos (Bauza et al., 2019). Aun así, la falta de métodos estandarizados para la detección rutinaria de los fagos de *Bacteroides* mediante qPCR ha impulsado la búsqueda de marcadores alternativos que cumplan con los criterios esenciales de un buen indicador, dado que hasta la fecha no se ha identificado uno que sea universalmente aplicable. En la actualidad, los fagos tipo crAss han despertado un gran interés, ya que se perfilan como un candidato prometedor.

1.4.4. Fagos tipo crAss como indicadores de contaminación fecal humana

Gracias a sus características particulares, estos bacteriófagos han sido propuestos como un marcador MST más fiable para el monitoreo de la calidad y seguridad del agua, permitiendo detectar contaminación fecal humana en diversos tipos de aguas (Stachler and Bibby, 2014). En comparación con otros indicadores bacterianos y virales, así como con virus humanos (Ahmed et al., 2021; Ballesté et al., 2019; Crank et al., 2020; Greaves et al., 2020; Hamza and Abd-Elmaksoud, 2023; Stachler and Bibby, 2014; Wu et al., 2020; Zhang et al., 2021), presentan ventajas claves, como su alta prevalencia y abundancia en heces humanas, lo que facilita su excreción en grandes cantidades a las aguas residuales (Tabla 1). Esto deriva en altas tasas de detección molecular en aguas contaminadas con heces humanas, tales como aguas residuales, pluviales, ríos, subterráneas y arroyos (Farkas et al., 2019; Stachler et al., 2018), donde no se observan variaciones estacionales. Además, estos fagos también se

pueden considerar indicadores útiles para las cadenas de producción alimentarias (Farkas et al., 2019; Lee et al., 2023; Nam et al., 2023).

Estudios basados en los genomas de los fagos tipo crAss indican que éstos exhiben una notable persistencia tanto ante procesos de desinfección como a los mecanismos naturales de inactivación tras su liberación en el entorno, hecho que se evidencia con su amplia distribución. Otro aspecto fundamental es su correlación con diversos virus responsables de enfermedades de transmisión hídrica, como norovirus, poliomavirus o adenovirus (Crank et al., 2019; Farkas et al., 2019; Hamza and Abd-Elmaksoud, 2023; Wu et al., 2020), lo que le permite ser un indicador más confiable para el monitoreo a largo plazo de virus patógenos (Sabar et al., 2022). De todos modos, los fagos de *Bacteroides* son indicadores de contaminación fecal, y correlacionan muy bien con la presencia de la misma, pero no son indicadores de virus patógenos, por lo que en ocasiones no correlacionan con estos virus, del mismo modo que los indicadores bacterianos de contaminación fecal no siempre correlacionan con los patógenos bacterianos presentes en las aguas.

Evaluar la persistencia de los indicadores es importante ya que la buena persistencia ambiental es una característica esencial en un buen indicador de contaminación microbiana. Muchos estudios que evalúan persistencia ante diferentes tratamientos de desinfección e inactivación ambiental se han basado en métodos moleculares (Ahmed et al., 2019, 2018a; Ballesté et al., 2021; Greaves et al., 2020; Hamza and Abd-Elmaksoud, 2023; Kongprajug et al., 2019; Sala-Comorera et al., 2021; Tandukar et al., 2020; Wu et al., 2020). Sin embargo, los métodos moleculares detectan únicamente un fragmento de ADN o ARN en lugar del microorganismo completo, y no siempre reflejan con precisión la infectividad de las partículas virales detectadas, y muchas pueden haber perdido su capacidad infecciosa (Aw and Rose, 2012; Greaves et al., 2020; Wu et al., 2020). Otro desafío es la susceptibilidad de las técnicas moleculares a la inhibición causada por diversas sustancias presentes en el ambiente. Esta inhibición limita la sensibilidad de los ensayos ya sea por la inhibición directa de determinados compuestos o por las estrategias aplicadas para minimizar posibles inhibiciones, como la dilución o la purificación del ácido nucleico (McKee et al., 2015).

La detección de los fagos por métodos moleculares permite detectarlos de forma estable incluso después de los procesos de desinfección (Botes et al., 2013), posiblemente debido a que los tratamientos reducen la infectividad de los fagos, pero no degradan su cápside, por lo que el ADN permanece estable y por lo tanto sigue siendo detectable. Esto representa una limitación para la evaluación del riesgo para la salud porque se detectan partículas a pesar de que ya no son infecciosas. Por ello, para poder evaluar con mayor precisión la persistencia de los fagos tipo crAss, es crucial realizar un estudio conjunto, aislando tanto el virión como a su

hospedador, para poder evaluar las partículas infecciosas, además de realizar un seguimiento por métodos moleculares. Aunque hasta el momento de este estudio se habían identificado varios viriones de fagos de tipo crAss (como se menciona en el apartado 1.3.4), el número sigue siendo insuficiente para obtener resultados robustos. El aislamiento de más fagos es, por tanto, el objetivo primario de esta tesis doctoral, ya que permitirá realizar ensayos más fiables sobre su persistencia y profundizar en el estudio de la interacción entre el fago y su hospedador.

2. OBJETIVOS

El objetivo general de la presente tesis es aislar y caracterizar nuevos fagos tipo crAss con la intención de comprender su rol ecológico y funcional en la regulación de la microbiota intestinal, evaluando su interacción con su huésped y su dinámica de infección. Además, se busca explorar su potencial como indicador de contaminación fecal de origen humano en nuestra región geográfica. Para alcanzar el objetivo general, se han planteado varios objetivos específicos que se detallan a continuación:

- Identificar y aislar nuevos fagos tipo crAss a partir de muestras con contaminación fecal humana confirmada en nuestra área geográfica, mediante el diseño y optimización de métodos moleculares y de cultivo. (Capítulo 1).
- Caracterizar los viriones asilados y llevar a cabo la secuenciación y el análisis comparativo de sus genomas completos. Esta caracterización contribuirá a ampliar el repertorio de fagos disponibles en las bases de datos públicas, facilitando su estudio filogenético y permitiendo una mejor comprensión de su diversidad, así como el mapeo de su distribución geográfica (Capítulo 1 y 3).
- Evaluar la interacción fago-hospedador y la dinámica de infectividad de los fagos crAss mediante la generación de mutantes de *Bacteroides* con fenotipos de resistencia o sensibilidad estables. Esto permitirá esclarecer los mecanismos que explican la variabilidad en la susceptibilidad y su impacto en la prevalencia del hospedador y la abundancia de los fagos en el intestino (Capítulo 2).
- Determinar, mediante herramientas bioinformáticas, el fago tipo crAss más adecuado para ser utilizado como indicador de contaminación fecal de origen humano universal. Esto se logrará utilizando una base de datos personalizada con genomas completos de fagos para evaluar su abundancia y prevalencia en diversos metagenomas virales de todo el mundo, lo que permitirá identificar aquellos fagos exclusivos de humanos, analizar su relación filogenética y delimitar su prevalencia (Capítulo 3).
- Evaluar la persistencia de los fagos crAssBcn frente a diferentes tratamientos de desinfección y a condiciones naturales, con el fin de comprender los efectos individuales de cada parámetro fisicoquímico sobre estos virus. Estos ensayos mediante técnicas moleculares y de infectividad contribuirán a validar su aplicación como indicador de contaminación fecal humana en evaluaciones de la calidad del agua (Capítulo 4).

3. INFORMES

3.1. Informe sobre el factor de impacto

Los artículos que conforman la memoria de esta tesis doctoral han sido publicados en revistas científicas internacionales registradas en el ***Journal Citation Report*** o se encuentran en proceso de preparación y son relevantes en la línea de investigación en la que ha estado trabajando la doctoranda durante los años de su tesis doctoral.

- El artículo "*Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution*" fue publicado el 18 de julio de 2023 en la revista **Nature Communications** (DOI: 10.1038/s41467-023-40098-z). Dicha revista se encuentra en el primer decil (D1) y en el primer cuartil (Q1) de su categoría (*Multidisciplinary Sciences*) con un factor de impacto en 2023 de 14,7.

- El artículo "*Regulation of Polysaccharide Biosynthesis by Phase Variable Region 7 (PVR7) in Bacteroides intestinalis Enhances Resistance to CrAssBcn Phages*" se encuentra actualmente en preparación. Se pretende finalizarlo y someterlo a consideración en la revista **Nature Communications**.

- El artículo "*Recruitment of complete crAss-like phage genomes reveals their presence in chicken viromes, few human-specific phages, and lack of universal detection*" fue publicado el 3 de octubre del 2024 en la revista **The ISME Journal** (DOI: 10.1093/ismejo/wrae192). Dicha revista se encuentra en el D1 y en el Q1 de todas sus categorías (*Microbiology*) con un factor de impacto en 2024 de 10.

- El artículo "*Persistence of crAssBcn phages in conditions of natural inactivation and disinfection process and their potential role as human source tracking markers*" fue publicado el 20 de diciembre del 2024 en la revista **Science of the Total Environment (STOTEN)** (DOI: 10.1016/j.scitotenv.2024.177450). Dicha revista se encuentra en el D1 y en el Q1 de en todas sus categorías (*Environmental Science*) con un factor de impacto en 2024 de 8.





Dra. Mª Teresa Muniesa Pérez

Dra. Cristina García Aljaro

Barcelona, julio de 2025

3.2. Informe de coautoría y de participación en las publicaciones

La doctoranda Clara Gómez Gómez ha participado en todas las publicaciones que forman parte de su tesis de la manera que se detalla a continuación. Ninguno de los coautores de los artículos ha utilizado los datos y las publicaciones para la elaboración de su tesis doctoral.

- Ramos-Barbero, M. D., **Gómez-Gómez, C.**, Sala-Comorera, L., Rodríguez-Rubio, L., Morales-Cortes, S., Mendoza-Barberá, E., Vique, G., Toribio-Avedillo, D., Blanch, A. R., Ballesté, E., Garcia-Aljaro, C., & Muniesa, M. (2023). Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution. *Nature Communications*, *14*(1), 4295. DOI: 10.1038/s41467-023-40098-z

  La doctoranda es co-primera autora de esta publicación. Ha desempeñado un papel activo en la investigación, llevando a cabo los experimentos que incluyen el aislamiento de todos los fagos tipo crAss, su posterior caracterización y la implementación de los ensayos de infectividad. Asimismo, ha colaborado en el análisis de resultados y la elaboración de tablas y figuras, además de contribuir a la redacción y revisión del articulo bajo la supervisión de sus directoras de tesis.

- **Gómez-Gómez, C.**, Ramos-Barbero, M. D., Sala-Comorera, L., Blanco-Picazo, P., Rodríguez-Rubio, L., Garcia-Aljaro, C., & Muniesa, M. (2025). Regulation of polysaccharide biosynthesis by phase variable region 7 (PVR7) in *Bacteroides intestinalis* enhances resistance to crAssBcn phages. Artículo en preparación.

  La doctoranda ha realizado todos los experimentos para la obtención de mutantes sensibles y resistentes, así como los procedimientos de clonación, los ensayos de PCR, qPCR y RT-qPCR para la identificación de mutantes y el análisis de la expresión genética relativa y ensayos de inactivación de fagos mediante el uso de fracciones de membrana. Asimismo, ha colaborado en el análisis de los resultados obtenidos, la confección de tablas y figuras y en la escritura y revisión del artículo bajo la supervisión de sus directoras.

- Ramos-Barbero, M. D., **Gómez-Gómez, C.**, Vique, G., Sala-Comorera, L., Rodríguez-Rubio, L., & Muniesa, M. (2024). Recruitment of complete crAss-like phage genomes reveals their presence in chicken viromes, few human-specific phages, and lack of universal detection. *The ISME Journal*, 18(1), wrae192.
  DOI: 10.1093/ismejo/wrae192

  La doctoranda ha realizado gran parte del procesamiento y preparación de las muestras, la purificación de partículas fágicas, la extracción de ADN vírico para los

análisis metagenómicos, con el objetivo de conseguir los viromas necesarios para el estudio. También ha contribuido activamente a la revisión del artículo, bajo la supervisión de sus directoras de tesis.

- **Gómez-Gómez, C.**, Ramos-Barbero, M. D., Sala-Comorera, L., Morales-Cortes, S., Vique, G., García-Aljaro, C., & Muniesa, M. (2024). Persistence of crAssBcn phages in conditions of natural inactivation and disinfection process and their potential role as human source tracking markers. *Science of the Total Environment*, 957, 177450. DOI: 10.1016/j.scitotenv.2024.177450

  La doctoranda ha participado activamente en el diseño experimental del estudio y ha llevado a cabo la mayor parte de los ensayos. También ha realizado los análisis de los resultados obtenidos, su organización y discusión, liderando la redacción de la primera versión del artículo, así como sus posteriores revisiones bajo la supervisión de sus directoras de tesis.





Dra. Mª Teresa Muniesa Pérez

Dra. Cristina García Aljaro

Barcelona, julio de 2025

4. RESULTADOS

4.1. Capítulo 1:

La caracterización de los fagos tipo crAss aislados destaca la heterogeneidad genética de los Crassvirales y su distribución mundial

Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution

María Dolores Ramos-Barbero, **Clara Gómez-Gómez**, Laura Sala-Comorera, Lorena Rodríguez-Rubio, Sara Morales-Cortes, Elena Mendoza-Barberá, Gloria Vique, Daniel Toribio-Avedillo, Anicet R. Blanch, Elisenda Ballesté, Cristina Garcia-Aljaro y Maite Muniesa

Nature Communications (2023); 14:4295.

DOI: 10.1038/s41467-023-40098-z

Introducción

El análisis *in silico* de metagenomas fecales humanos ha permitido identificar numerosas secuencias genómicas pertenecientes a un nuevo grupo de fagos, denominados crAssphages (Dutilh et al., 2014), clasificados dentro de la familia *Intestiviridae* (especie *Carjivirus communis*) (Turner et al., 2023) y que infectan al género *Bacteroides*. Tras su identificación, se destacó su distribución global y su alta especificidad por el intestino humano, lo que explica su abundancia en heces humanas, aunque también se han detectado en niveles significativamente inferiores en muestras con contaminación fecal de origen animal (García-Aljaro et al., 2017). Estos fagos, particularmente el *p-crAssphage* (el primero descrito y considerado el prototipo de los crAssphages), han sido propuestos como indicadores universales de contaminación fecal humana (Dutilh et al., 2014; Guerin et al., 2018; Stachler et al., 2018; Sabar et al., 2022), pese a la diversidad genética existente dentro del orden Crassvirales. Asimismo, se ha sugerido que estos fagos presentan una persistencia ambiental superior a la de *Escherichia coli* y comparable a la de los colifagos somáticos, una característica clave para un indicador eficaz (Ballesté et al., 2019).

A pesar de su abundancia en los metagenomas fecales humanos (Edwards et al., 2019; Yutin et al., 2021), hasta el inicio de esta tesis solo se habían aislado siete fagos del tipo crAss, evidenciando las dificultades inherentes a su aislamiento mediante técnicas de cultivo. Esta limitación representa un desafío significativo para el estudio de la relación fago-hospedador,

así como para la comprensión de su persistencia e inactivación en el medio ambiente, aspectos cruciales para evaluar su potencial como indicador de contaminación fecal. Además, los métodos moleculares actuales no permiten distinguir entre virus infecciosos e inactivados, lo que añade una dificultad adicional para su caracterización.

Objetivo

En este capítulo se busca aislar y caracterizar nuevos fagos de tipo crAss virulentos (los cuales infectan especies de *Bacteroides*) a partir de muestras de aguas residuales de nuestra zona geográfica lo que representa un avance importante en la descripción de este grupo de virus específicos del ser humano. El objetivo de este trabajo es aumentar la cantidad de viriones aislados para, mediante la combinación de técnicas de cultivo y moleculares, profundizar en la comprensión de la biología de los fagos tipo crAss y evaluar su potencial como indicadores universales de contaminación fecal humana.

Resultados y discusión

Tras detectar la presencia de fagos crAss mediante varios ensayos de qPCR en 12 muestras de aguas residuales de cinco plantas de tratamiento del área urbana de Barcelona, y después de evaluar el hospedador más adecuado para su propagación, se lograron aislar 25 nuevos fagos tipo crAss (ΦCrAssBcn1 - ΦCrAssBcn25). Este aislamiento se consiguió tras varios pasos de propagación consecutivos en cultivos de *Bacteroides intestinalis*, lo que resultó en un aumento significativo del número de partículas virales y una mayor señal de hibridación positiva (20-30% de las placas). Los fagos positivos fueron aislados, purificados y confirmados como del tipo crAss utilizando el ensayo de qPCR diseñado a partir del genoma del primer fago crAss aislado, el ΦCrAss001.

Una vez aislados los fagos crAssBcn, se caracterizó su morfología identificándose como del tipo podovirus mediante microscopía electrónica, sin observarse diferencias entre ellos a pesar de su origen diverso. Los fagos crAssBcn presentaron un diámetro de cápside de 77 ± 0,8 nm y una cola corta de 40 ± 0,9 nm. Durante la propagación en cultivo líquido, se alcanzaron títulos máximos de hasta 2×10^9 ufp/ml a las 24 horas, con aumentos de 2,5 a 4 unidades logarítmicas entre las 0 y 24 horas, siendo la curva de propagación similar a la de ΦCrAss001. El valor promedio del tamaño de explosión (*burst size*) fue de 64,6 ± 16 fagos por célula infectada. Sin embargo, a pesar de ser capaces de propagar, los fagos no afectaron el crecimiento de la cepa hospedadora, ya que los cultivos no se lisaron (o levemente) y el número de células hospedadoras aumentó en los cultivos de propagación, alcanzando una densidad óptica (DO) cercana a 1 a las 24 horas. Esta falta de lisis de la bacteria hospedadora sugiere una posible interacción dinámica entre el fago y el hospedador que permite la persistencia de la población bacteriana a la vez que propagan los fagos, posiblemente debido

a: 1) una infección del fago que permanece en estado de portador, 2) una pseudolisogenia o 3) una resistencia parcial de la población bacteriana basada en la variación de fase, implicando plausiblemente a los receptores de los fagos (Porter et al., 2020).

La secuenciación de los 25 fagos crAssBcn permitió identificar 24 genomas completos de ADN de doble cadena y un genoma incompleto, con tamaños que oscilaban entre 97,685 y 103,497 pb, que contenían entre 101 y 114 pautas de lectura abierta (O*pen Reading Frame* o ORF). La comparación intergenómica de estos fagos evidenció que ninguno compartía una secuencia idéntica, ni a nivel genómico ni proteico, lo que confirma su notable diversidad genética. A pesar de su diversidad, todos los fagos aislados se agruparon taxonómicamente dentro del orden Crassvirales, familia *Steigviridae*, subfamilia *Asinvirinae* y género *Kehishuviru*s. Según el criterio de demarcación de especies para virus basado en identidad genética, los fagos crAssBcn se clasificaron en seis especies distintas, siendo la especie I la más predominante. Sin embargo, al compararlos con otros fagos tipo crAss del género *Kehishuvirus*, ningún crAssBcn coincide con estos (a excepción del fago ΦCrAssBcn20 clasificado en la misma especie que el primer aislado ΦcrAss001). Esta heterogeneidad de los fagos de tipo crAss remarca la necesidad de utilizar genomas de fagos tipo crAss con una identidad suficiente y de longitud representativa, o genomas obtenidos de fagos aislados, para una correcta clasificación (Yutin et al., 2021). La anotación funcional indicó que se trata de fagos virulentos, dado el alto número de genes asociados con el metabolismo y la replicación, y la ausencia de genes relacionados con la lisogenia. Finalmente, el análisis filogenético reveló que los fagos crAssBcn forman un clúster junto con ΦCrAss001, pero presentan diferencias significativas en comparación con otros fagos de la familia *Steigviridae*, destacando su singularidad evolutiva.

Los 24 genomas completos de los fagos crAssBcn mostraron una variabilidad significativa en ciertos genes potenciales, con una identidad de aminoácidos promedio (*Average Amino acid Identity* o AAI) cercana al 50%. Los ORFs más variables codificaban para 1) espículas de la cola, que presentan alta variabilidad debido a la coevolución con los receptores bacterianos (Brown et al., 2021), generando potencialmente nuevas vías de infección; 2) proteínas de la cola y 3) de cápside, que derivan de un ancestro común pero divergieron en dos tipos distintos, aparentemente como resultado de un evento de integración; 4) holinas; 5) proteínas hipotéticas, en las que se identificaron variantes de estas proteínas y de las holinas en ramas filogenéticas distintas, apoyando la hipótesis de especiación por intercambio de genes ortólogos (Gabaldón and Koonin, 2013); 6) endonucleasas HNH, que desempeñan funciones clave en la recombinación, reparación y empaquetado del ADN; y 7) ADN polimerasas. Adicionalmente, aunque en menor medida, también se detectó variabilidad en timidilato sintasas, lectinas tipo C y ARN polimerasas.

En particular, las ADN polimerasas I de tipo A se dividieron en dos variantes principales: una de 706 aminoácidos presente en las especies V y VI, y otra de 775 aminoácidos en las especies I a IV. Estas variantes presentan un 45% de identidad entre sí, lo que refleja eventos evolutivos o de recombinación. A pesar de sus diferencias estructurales, los sitios activos de estas polimerasas permanecieron conservados, destacando la preservación de sus funciones esenciales. Esta diversidad observada podría ser consecuencia de la adaptación al entorno bacteriano y de la reconfiguración modular de los genomas fágicos. Mantener la diversidad en los módulos replicativos podría ser beneficioso para los fagos, ya que les permitiría adaptarse a los cambios en la coevolución con las bacterias o evadir sus sistemas de defensa (Smug et al., 2023).

Además, los genes adyacentes a las polimerasas también mostraron alta variabilidad. Sin embargo, no fue posible identificar algunas de las proteínas codificadas, clasificadas como proteínas hipotéticas. Finalmente, la variabilidad observada en la ADN polimerasa tipo I plantea dudas sobre su idoneidad como marcador universal de fagos, como había sido propuesta (Liang et al., 2016), mientras que genes más conservados, como la terminasa (Guerin et al., 2018), podrían ser alternativas más adecuadas.

El análisis de la distribución geográfica mundial de los fagos crAssBcn reveló su presencia en el 51% de los metagenomas analizados, provenientes de países como Irlanda, EE. UU., España, Uganda, Malaui, Camerún, Etiopía y Finlandia. Este hallazgo resultó sorprendente, ya que el análisis filogeográfico de los fagos tipo crAss ha identificado clústeres locales de fagos genéticamente similares, aunque también se ha documentado una cepa con distribución global (Edwards et al., 2019). Este fenómeno podría explicarse por los movimientos de migración humana o por la estabilidad y adaptabilidad de los fagos al medio ambiente. Además, los fagos crAssBcn mostraron una representación mayor que ΦCrAss001, destacando España, seguida de Etiopía y Malaui, como los metagenomas con la mayor abundancia relativa de fagos crAssBcn.

Conclusiones

Gracias a la metodología de aislamiento utilizada, se han obtenido 25 nuevos fagos crAssBcn, los cuales, tras su caracterización, mediante la secuenciación de sus genomas, mostraron una notable diversidad genética, evidenciada por la variabilidad en genes clave como las espículas de la cola y las polimerasas. A pesar de esta diversidad, los fagos crAssBcn comparten una morfología tipo podovirus y presentan una propagación eficiente en cultivos de *B. intestinalis*, sin detrimento del número de células bacterianas, lo que sugiere una interacción dinámica con el hospedador. Los crAssBcn se clasificaron dentro del orden Crassvirales, aunque con diferencias significativas en comparación con otros fagos tipo crAss

previamente descritos, lo que resalta la necesidad de utilizar genomas representativos para una clasificación precisa. Además, el análisis filogenético mostró que los fagos crAssBcn forman un clúster cercano a ΦCrAss001, pero presentan particularidades evolutivas que subrayan su singularidad. La distribución geográfica de estos fagos sugiere una dispersión global, posiblemente influenciada por la migración humana, lo que refuerza su potencial como indicadores universales de contaminación fecal humana. No obstante, la variabilidad observada en ciertas regiones del genoma resalta la necesidad de explorar otros marcadores moleculares más precisos para un seguimiento microbiológico adecuado, además de determinar cuál es el miembro más representativo de los Crassvirales a nivel global.

nature communications

Article

https://doi.org/10.1038/s41467-023-40098-z

Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution

Received: 2 February 2023

Accepted: 7 July 2023

Published online: 18 July 2023

Check for updates

María Dolores Ramos-Barbero[1,2], Clara Gómez-Gómez[1,2], Laura Sala-Comorera[1,2], Lorena Rodríguez-Rubio[1], Sara Morales-Cortes[1], Elena Mendoza-Barberá[1], Gloria Vique[1], Daniel Toribio-Avedillo[1], Anicet R. Blanch[1], Elisenda Ballesté[1], Cristina Garcia-Aljaro[1] & Maite Muniesa[1]

Crassvirales (crAss-like phages) are an abundant group of human gut-specific bacteriophages discovered in silico. The use of crAss-like phages as human fecal indicators is proposed but the isolation of only seven cultured strains of crAss-like phages to date has greatly hindered their study. Here, we report the isolation and genetic characterization of 25 new crAss-like phages (termed crAssBcn) infecting *Bacteroides intestinalis*, belonging to the order Crassvirales, genus *Kehishuvirus* and, based on their genomic variability, classified into six species. CrAssBcn phage genomes are similar to ΦCrAss001 but show genomic and aminoacidic differences when compared to other crAss-like phages of the same family. CrAssBcn phages are detected in fecal metagenomes around the world at a higher frequency than ΦCrAss001. This study increases the known crAss-like phage isolates and their abundance and heterogeneity open the question of what member of the Crassvirales group should be selected as human fecal marker.

The discovery of crAssphage, a highly human-specific resident of the gut, has been of paramount importance, as it is the only universal marker of human fecal pollution described so far[1–4]. The first crAssphage (cross-assembly phage) is a bacteriophage or group of phages first identified by in silico studies of human fecal metagenomes[4]. This first p-crAssphage (prototypical-crAssphage) has recently been classified as the family *Intestiviridae*, species *Carjivirus communis*[5]. After the discovery of p-crAssphage, other phages (crAss-like phages), similar in their genomic architecture and sharing the same ancestor with p-crAssphage, mostly uncultured, and highly abundant in the mammalian gut and other habitats, have been discovered, all conforming a new order of Crassvirales[3,6,7]. Currently, the order Crassvirales comprises four families, 11 subfamilies, 42 genera and 73 new species (https://ictv.global/taxonomy). All of them seem to infect bacteria in the phylum Bacteroidetes and, based on the few virions isolated, they specifically infect the *Bacteroides* genus.

Given the abundance of crAss-like phage sequences identified in different metagenomes[7,8] and the increasing number of studies detecting crAssphages (or crAss-like phages) in different ecosystems by qPCR[2,9–12], it is intriguing why only seven crAss-like phages have been isolated to date. Possible explanations are the use of unsuitable host bacteria, the relatively low proportion of virulent crAss-like phages in comparison with other lytic phages infecting *Bacteroides*, or the inability of infectious crAss-like phages to generate visible plaques of lysis that allow their isolation.

Efforts to isolate crAss-like phages were not successful until four years after their discovery, when ΦCrAss001 (family *Steigviridae*, species *Kehishuvirus primarius*[5]) was isolated in *Bacteroides intestinalis*[5,13,14]. The difficulties in isolating crAss-like phages in pure culture persist, with only six other crAss-like phages isolated to date: DAC15 and DAC17 (family *Steigviridae*, species *Wulfhawirus bangladeshii*)[5], were isolated on *Bacteroides tethaiotaomicron*[15] in

[1]Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Diagonal 643. Annex. Floor 0, E-08028 Barcelona, Spain. [2]These authors contributed equally: María Dolores Ramos-Barbero, Clara Gómez-Gómez, Laura Sala-Comorera. e-mail: mmuniesa@ub.edu

2020, ΦCrAss002 (family *Intestiviridae*, species *Jahgtovirus secundus*)[5] on *Bacteroides xylanisolvens* in 2021[16], and recently, three more crAss-like phages have been isolated infecting *Bacteroides cellulosilyticus* and assigned to three new species (*Kehishuvirus winsdale* (Bc01), *Kolpuevirus frurule* (Bc03), and *Rudgehvirus redwords* (Bc11)[17]. All other reported crAss-like genomes are the result of composite assemblies and have never been propagated in pure culture in a laboratory. Consequently, there is limited information about the biology and replicative cycle of crAss-like phages. It has been established that they are virulent phages of podovirus-like morphology (short-tailed) with a relatively large (100 Kb) double-stranded circular DNA genome. Studies on the replication of crAss-like phages suggest they do not follow a common lytic cycle. Plaques of lysis of ΦCrAss001 are visible in agar overlays, but the infected liquid bacterial cultures are not cleared, suggesting that phage and host co-exist. This has been attributed to phase-variation of bacterial capsular polysaccharides, which maintains a dynamic equilibrium between phage sensitivity and resistance of *Bacteroides* cells and allows the phage and host cell to multiply in parallel[18]. ΦCrAss002 does not form plaques or spots on lawns of sensitive cells, nor does it lyse liquid cultures, even at high titers, and the phage only propagates if co-cultured with the host for a minimum of 3–5 days[16]. Phages DAC15 and DAC17 might form plaques of lysis but they propagate poorly in the *B. thetaiotaomicron* wild type strain and more readily in mutant strains with single capsular polysaccharides[15].

Metagenomic analysis has enabled the identification of many crAss-like phage genome sequences, and multicenter investigations have concluded that crAss-like phages, especially p-crAssphage, are highly abundant in human feces and have a global distribution[8]. As a result, p-crAssphage, and in general crAss-like phages, have been proposed as universal indicators of human fecal pollution[10,12,19,20], despite the genetic variability within the Crassvirales order[8]. qPCR analysis has also revealed crAss-like phages in samples with animal fecal contamination, although at much lower levels than in human feces[12]. It has also been suggested that crAss-like phages persist in the environment for longer than *Escherichia coli* and comparably with somatic coliphages[21], an essential attribute of an effective fecal indicator.

Nevertheless, the molecular methods used so far for crAss-like phages detection are limited in that they cannot distinguish between the target and non-infecting or inactivated viruses. Moreover, the abundance and persistence of crAss-like phages in the human gut has not yet been explained mechanistically and, as suggested previously, the phage-host relationship can only be properly studied with isolated phage-host pairs[16].

In this work, we present the isolation and characterization of 25 new virulent crAss-like phages infecting *B. intestinalis* that represents a step forward in the analysis of this ubiquitous human-specific phage group.

Results

Isolation of crAss-like phages

CrAss-like phages were searched in 12 sewage samples from five wastewater treatment plants using three different qPCR assays, the one designed from ΦCrAss001 genome in this study and two previously described[12,19]. Depending on the qPCR assay employed and the samples, values in wastewater ranged from 10^3–10^6 gene copies (GC)/ml of crAss-like phages. If considered the qPCR assay designed from ΦCrAss001, the one finally used, 10^3–10^5 gene copies (GC)/ml of crAss-like phages were detected (Supplementary Table 1). Assuming that one GC corresponds to one phage particle, the analyzed waters contained 10^3 to 10^5 viral particles/ml.

Different qPCR assays revealed the presence of crAss-like phages but were unable to determine whether they were infectious phages that could be isolated or propagated, a task that has proven challenging to date[16,18]. The three different qPCR assays were tested in enrichment cultures of different strains of *Bacteroides* (Supplementary Table 2), to identify the most suitable host for the propagation of crAss-like phages present in the wastewater samples and thus facilitate their isolation. Three consecutive propagation steps were performed, and the abundance of crAss-like phages was monitored with the different qPCR assays after each step. *B. intestinalis* was the only host that generated a notable increase in the number of crAss-like phage particles detected with the qPCR designed for ΦCrAss001, after the second or third propagation step in nine of the 12 samples analyzed (Supplementary Table 1). Phage suspensions obtained directly from wastewater samples or after propagation in these nine samples were diluted and plated. Plaques of lysis (plaque-forming units: pfu) generated by crAss-like phages were detected by plaque blot hybridization using the CrAss1-ORF46 probe. A clear hybridization signal was obtained only from the propagated suspensions, which facilitated the enumeration of the tiny plaques (Supplementary Fig. 1). On average, 20–30% of the plaques visualized on the soft agar overlay showed a positive signal for the CrAss1-ORF46 probe, although the proportion may have been higher, given the difficulty in visualizing the smallest plaques (Supplementary Fig. 1). Finally, 1–5 positive plaques from each wastewater sample, well separated from the other plaques and showing a clear hybridization signal, were isolated, purified, and confirmed to be crAss-like phages by PCR. Twenty-five phages (named crAssBcn phages, from ΦCrAssBcn1 to ΦCrAssBcn25) from samples of the different wastewater plants and collected in different dates were randomly selected for further characterization (Supplementary Table 1).

Characterization of the new crAss-like phages

All crAssBcn phages showed a podovirus-like morphology with head diameters of 77 ± 0.8 nm and a short tail of 40 ± 0.9 nm with a characteristic trident shape (Fig. 1). Despite the different origins of the crAss-like phages analyzed, no morphological differences between them were observed.

When propagating in liquid culture, maximum titers of crAssBcn phages were obtained after 24 h of propagation reaching values of up to 2×10^9 pfu/ml, with increases of 2.5–4.0 log units between 0 and 24 h (Supplementary Fig. 2a). An initial explosion in the infectious process at 2 h of incubation, with an increase of 1 log unit, was followed by a lower increase at 2–4 h. A second substantial increase of 1 log unit was observed at 4–6 h of incubation, followed by a slight increase at 6–9 h. From 9 to 24 h, the number of infectious crAssBcn phages remained constant, all reaching values of 4×10^7 -2×10^9 pfu/ml. Although differences were observed, the mean propagation curve of the 25 crAssBcn phages (dotted blue line in Supplementary Fig. 2a) did not differ significantly from that of ΦCrAss001 (red line) (Wilcoxon matched pairs test, $p = 0.125$). Likewise, no differences were observed between individual crAssBcn phages (grey lines) ($p > 0.05$).

The average burst size of the crAssBcn phages was calculated as 64.6 ± 16 phages per infected cell. However, the crAssBcn phages did not affect the host strain growth. The infected cultures were not cleared during propagation, and instead the number of host cells increased throughout, reaching an optical density (OD) at 600 nm close to 1.0 at 24 h (Supplementary Fig. 2b). A similar pattern was revealed by measuring the number of culturable cells, which fluctuated until 9 h after infection and reached the highest concentration at 24 h (on average 4×10^8 cfu/ml). The concentration of the uninfected *B. intestinalis* control was only slightly higher (on average 8.8×10^8 cfu/ml) than the infected *B. intestinalis* cultures after 24 h (Supplementary Fig. 2b).

Intergenomic comparison of the 25 crAssBcn phages and other crAss-like phages

Sequencing of the 25 isolates allowed the identification of 24 dsDNA complete genomes and one draft genome (ΦCrAssBcn25). The 24 complete phage genomes ranged from 97,685 to 103,497 bp in size

https://doi.org/10.1038/s41467-023-40098-z

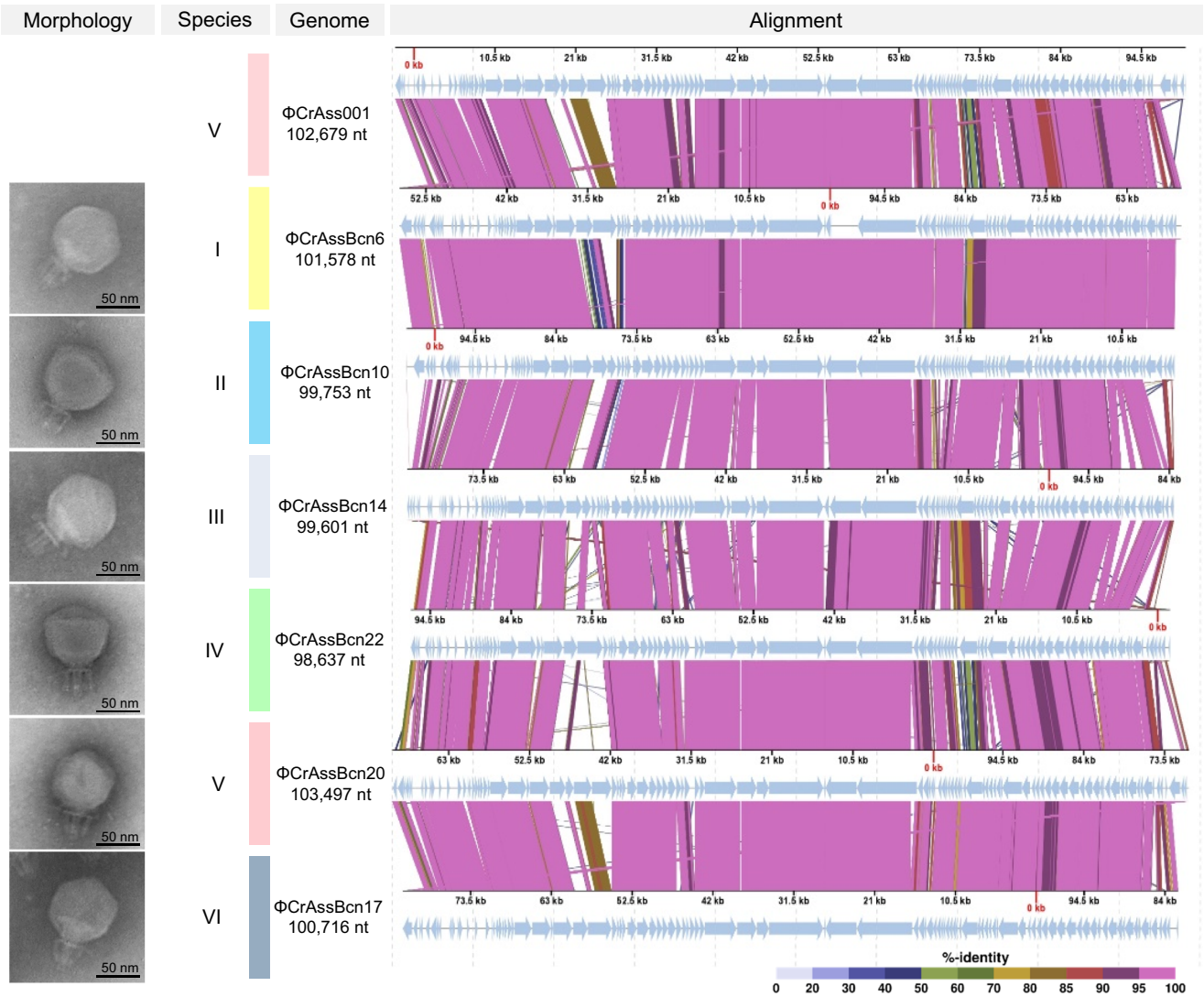


Fig. 1 | Comparison of the six representative crAssBcn phages of each species and ΦCrAss001. Electron micrographs and CDS alignment of the genomes of six crAssBcn phages representative of the six species (groups I-VI). Light blue arrows indicate the ORFs. The color of the bands between the seven genetic maps shows the percentage of identity between each sequence as indicated in the legend at the bottom of the figure.

and contained 101 to 114 ORFs (Supplementary Table 3, Supplementary Fig. 3).

Comparison of the 24 crAssBcn phages showing complete genomes (incomplete phage ΦCrAssBcn25 genome was excluded from the following analyses) revealed that none shared an identical nucleotide sequence (Fig. 2, Supplementary Fig. 3) and they all also differed at the amino acid level confirming that technically none had identical genomes. The phage with the largest complete genome within each species, according to the grouping suggested by VIRIDIC (in nt), was selected as representative of that species(Figs. 1, and 2) Based on the recently revised taxonomy[5], all the crAssBcn phages were classified within the order Crassvirales, family *Steigviridae*, subfamily *Asinivirinae*, and genus *Kehishuvirus*.

According to the International Committee for the Taxonomy of Viruses[22], the main species demarcation criterion for bacterial and archaeal viruses is currently set at a genome sequence identity of 95% over 85% of the complete genome. Using this criterion, the group of 24 crAssBcn phages could be divided in six different species, (I to VI) (Fig. 2, Supplementary Table 3), each comprising from 1 to 11 phage genotypes. Species I was the most numerous (11 phages), representing 44% of the isolated crAssBcn phages. Only ΦCrAssBcn20 (species group V) showed a genome sequence identity of > 95% with ΦCrAss001, being a member of the same species (*Kehishuvirus primarius)* (Fig. 2).

Functional annotation of all crAss-like phages is presented in Supplementary Data 1. Exemplifying the crAssBcn phages, the genetic map of the isolate ΦCrAssBcn6, representative of species I, shows the distribution of 105 ORFs (Supplementary Fig. 4). All the crAssBcn phages were probably virulent, as genes related to lysogeny (i.e., integrase, excisionase, and lysogenic module genes) were not detected. Additionally, the high number of ORF genes related to metabolism and replication (Supplementary Fig. 4) suggests the phages had a high replicative potential, which is characteristic of virulent phages.

To explore the phylogenetic relationship between crAssBcn phages with their closest relatives, we performed a phylogenetic analysis including as query the 25 crAssBcn phages and the 15 uncultured crAss-like phages assigned within the 15 species of the family *Steigviridae*[6] (in Supplementary Table 4). The proteomic tree showed that the 25 crAssBcn phages clustered with different levels of similarity on the same branch, where ΦCrAss001 was also clustered. In contrast, crAss-like phage genomes of species assigned to *Steigviridae* family were close but in different branches (Fig. 3).

Alignment of the ΦCrAssBcn6 genome, the representative of species I, with the genomes of the 15 phages of the *Steigviridae* family

https://doi.org/10.1038/s41467-023-40098-z

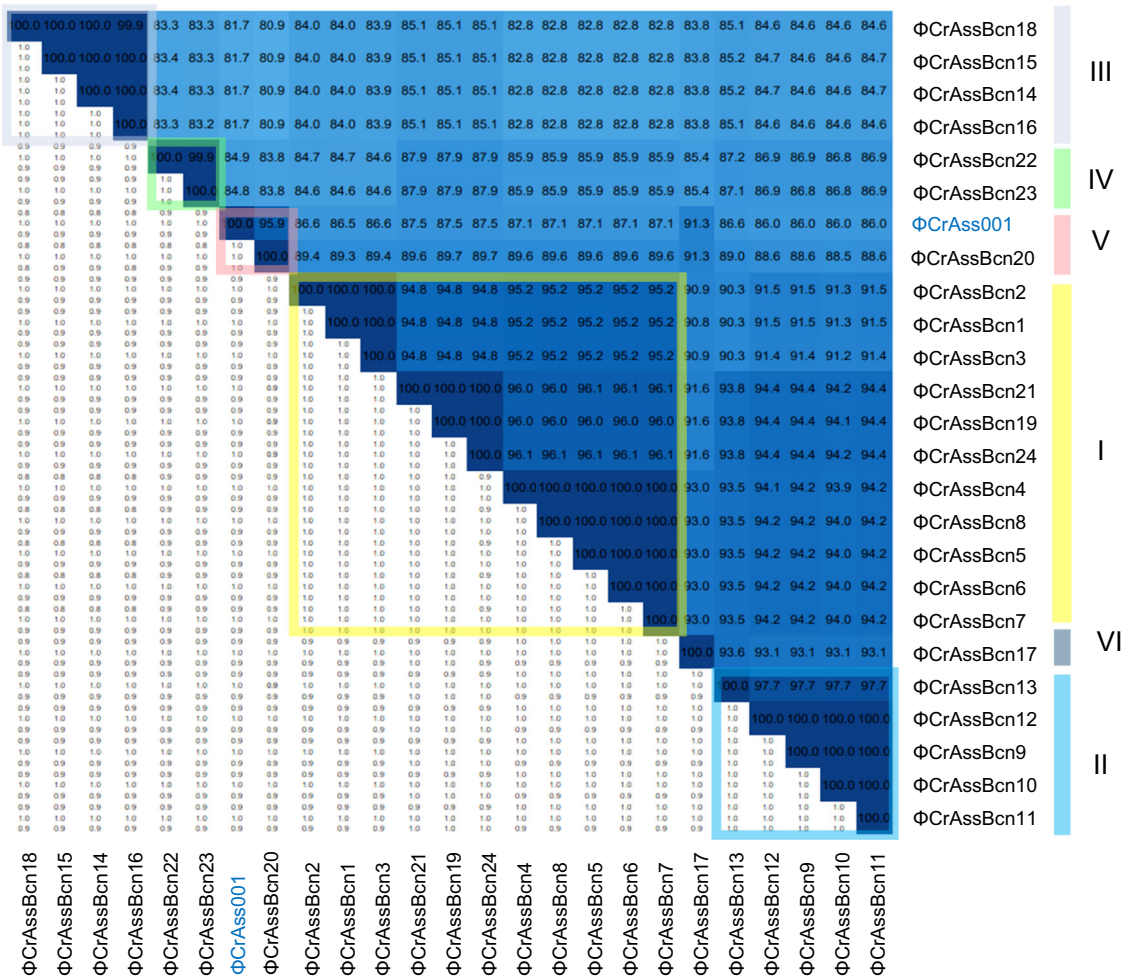


Fig. 2 | Intergenomic comparison of the 24 complete crAssBcn phages genomes and their classification into species. Heatmap generated by VIRIDIC shows the intergenomic similarity values (right half) and alignment indicators (left half). The percent identity between two genomes was determined by BLASTn, integrating intergenomic similarity values with data on genome lengths and aligned genome fractions. In the right half, the more closely related the genomes, the darker the color. The numbers inside the map represent the identity values for each genome pair, rounded to the first decimal. In the left half, the aligned genome length is indicated, with darker colors corresponding to smaller aligned genome fractions. On the right side of the heatmap, the six species (I to VI) in which each virus is classified are indicated in colored squares.

allowed to visualize the low identity at the genome level (Supplementary Fig. 5), therefore crAssBcn phages cannot be assigned to any of the previously described species, except ΦCrAssBcn20, as indicated above.

Variability within the crAssBcn phages

When comparing the genomes of the 24 crAssBcn phages (ΦCrAssBcn25 was excluded), the first observation was that some ORFs were detected in some, but not in all, of them (Supplementary Table 5). Some ORFs were absent in 100 to 8% of the phage genomes. Among them, many encoded hypothetical proteins, endonucleases or tail fibers (Supplementary Table 5).

Considering those ORFs present in all or most of the genomes, we aimed to identify those showing the highest variability among the 24 crAssBcn phage genomes. To avoid generating an excessively large data matrix, which would have been difficult to interpret, only the genomes of the six representative phages were used to ascertain the differences between the *ca* 100 ORFs of each phage. Using a reciprocal best-match analysis, the comparison revealed the ORFs with the lowest % of amino acid identity (AAI) and hence with the highest variability in the phage genomes (Supplementary Fig. 6). Seven ORF with the lowest % AAI were selected and aligned (Figs. 4 and 5) to illustrate the degree of variability observed. According to the last annotation of ΦCrAss001 (released in January 2023) and a recent cryoEM identification of ΦCrAss001 structural proteins[14], the most variable ORFs (AAI close to 50%) encode a tail spike, a tail protein, a head protein, a holin, a HNH-endonuclease, a hypothetical protein (Fig. 4), and a DNA polymerase I (Fig. 5). To a lower extent, a certain variability in thymidylate synthases, C-type lectin, and RNA polymerases was also observed (See Supplementary Data 2 for detailed information). The phylogenetic trees generated with these highly variable proteins were constructed with the sequences of crAssBcn phages and the closest hits showing a genome coverage larger than 50% and an identity over 50% or, in case that no sequences accomplished these criteria, the ten best hits in the databases were used (datasets are included in Supplementary Data 3). The tail proteins and the HNH-endonucleases of all crAssBcn phages derive from a single common ancestor, although some are not present in all the crAssBcn phages (below the detection limit) (Fig. 4). The crAssBcn phages head proteins derive from a single cluster that later shows a dichotomy in two different variants, a dichotomy that includes the head protein of phages ΦCrAssBcn20 and ΦCrAss001, that differ even though both phages belong to the same species V (Fig. 4). In contrast, the tail spike proteins, holins and the hypothetical protein of crAssBcn phages are located in different clusters, similarly to other crAss-like phages of the databases (Fig. 4).

https://doi.org/10.1038/s41467-023-40098-z

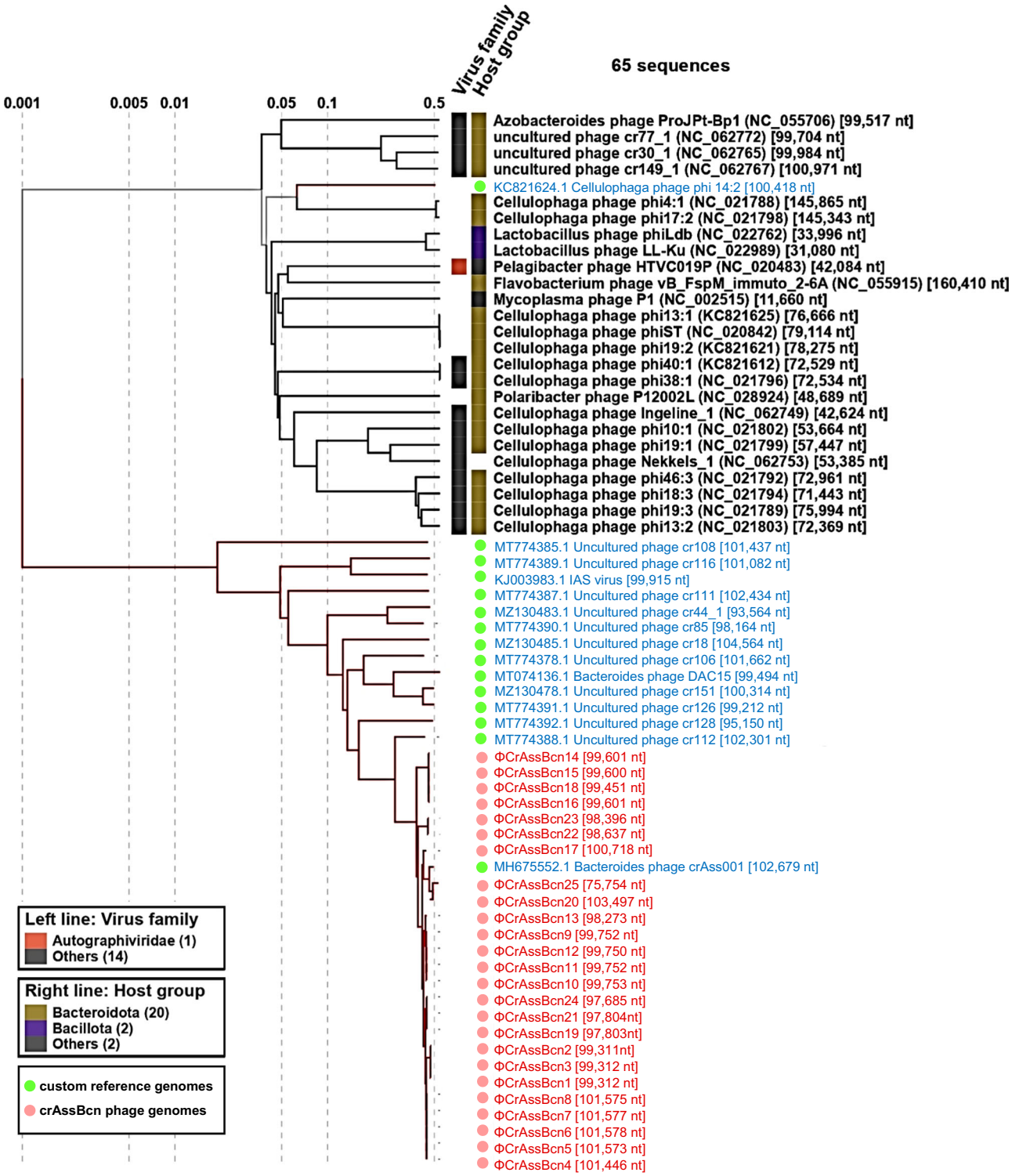


Fig. 3 | Phylogenetic relationship of the crAssBcn phages. Viral proteomic tree constructed with Viptree, where the 25 crAssBcn phages (in red) and 15 crAss-like phages of the *Steigviridae* family (in blue) were included as a query, showing the phylogenetic relationship of each phage with its closest relatives.

Analysis of the DNA polymerase A of the representative crAssBcn phages species I-VI (Fig. 5) showed that all six species encoded a family A DNA polymerase I. However, the alignment and % AAI for the polymerases confirmed two different types with a 45% of AAI between them; one of 706 aa, identified in species V and VI (phagesΦCrAssBcn20, ΦCrAssBcn17) coincident with the ΦCrAss001 DNA polymerase, and another one of 775 aa in 22 phages in species I to IV (Fig. 5). Both types of polymerases show hits in the databases (Fig. 5), as the polymerase type of species V-VI is present in other phage genomes assigned to *Steigviridae* family (for example GenBank accession number YP_010111476.1), or to polymerases detected in metagenomes or in sequences previously reported by Yutin et al. [7] (Fig. 5). The polymerase variant of species I-IV showed the closest hit (97.4%) in a metagenome without species assigned (MAG sequences in Fig. 5), and one of the closest phage genomes was that of the uncultured phage cr106_1, assigned to species *Mahstovirus faecalis* (accession N°

https://doi.org/10.1038/s41467-023-40098-z

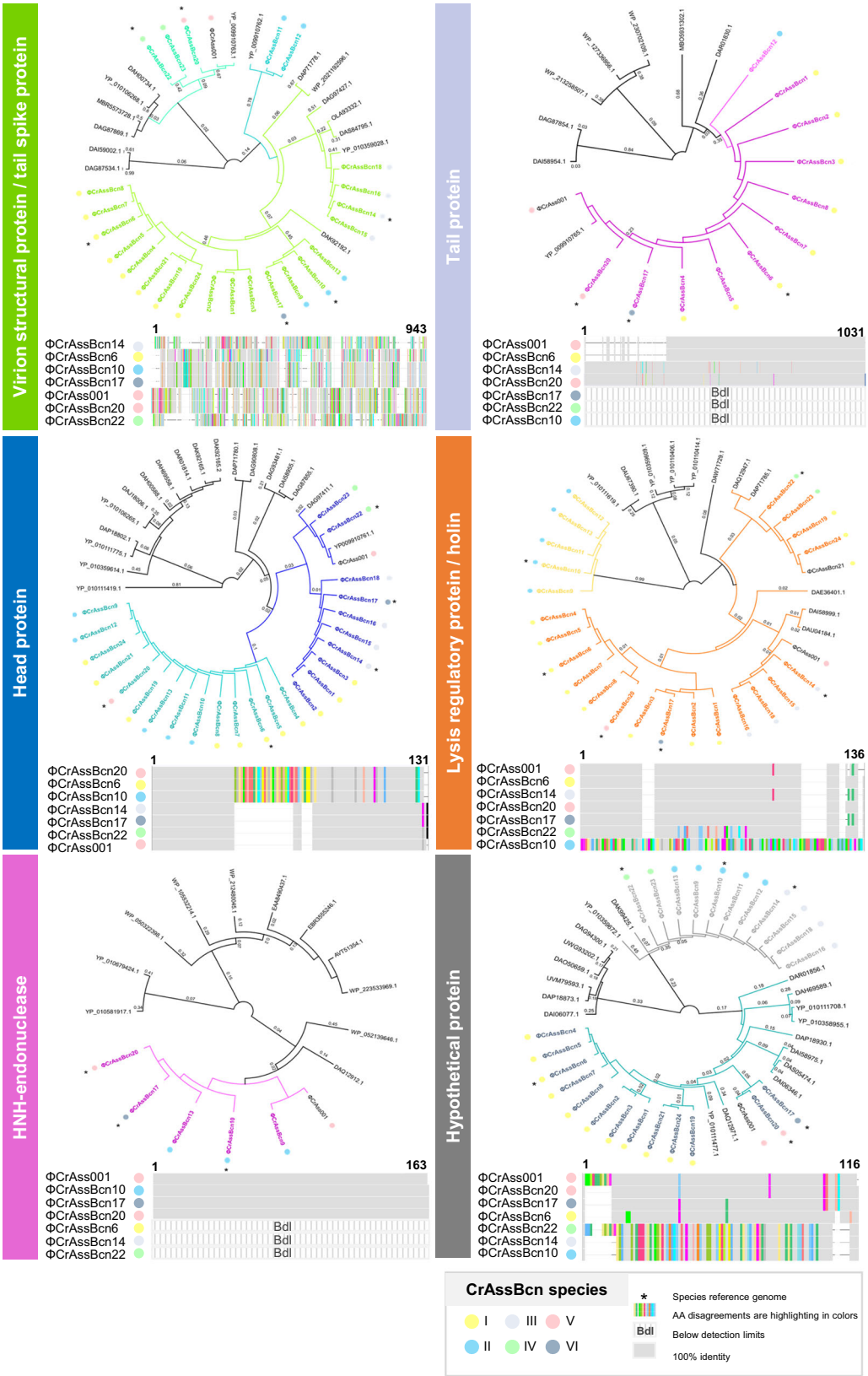
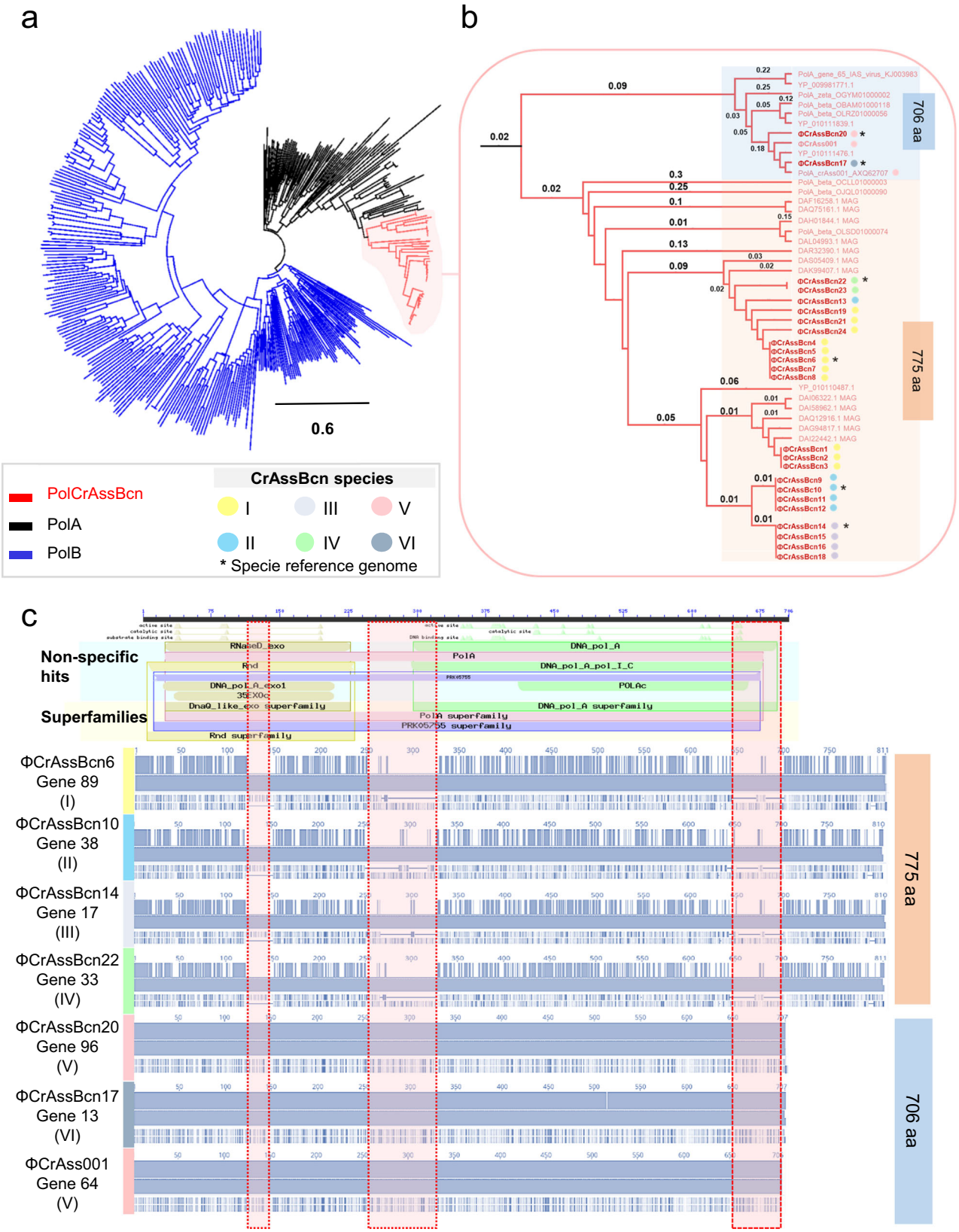


Fig. 4 | Comparison of the most variable ORFs of crAssBcn phages. Phylogenetic tree and multialignment of the amino acid sequences of the different variable ORFs of the CrAssBcn phages and ΦCrAss001. In the phylogenetic tree, the 24 crAssBcn phage sequences are displayed in coloured lines, while sequences from the databases are displayed in black lines. Only branch labels over 0.001 are shown. The multialignment of the different ORFs was constructed with the representative genomes of the six species where 100% identity is displayed in grey bars and aminoacid disagreements are shown in coloured bars. The color code for the crAssBcn species is indicated in the legend. Those proteins not detected in a given species are indicated a Bdl (below detection limit) and consequently are not present in the phylogenetic tree. The asterisk indicates the species reference genome.

https://doi.org/10.1038/s41467-023-40098-z



YP_010110487.1) (Fig. 5 and Supplementary Data 3), but with only a 79.3% of identity. Phylogenetic tree (Fig. 5) shows that PolA evolved from PolB, that both variants of the PolA of crAssBcn phages seem to have evolved from a common PolA ancestor and that from this point, they have diverged in the two different types observed. The type with 775 aa appears to be also the most abundant in the crAssBcn phages and is also the most represented among the crAss-like phage sequences available in the databases.

The archetypical structure of the polymerase A should comprise three domains to sustain its activity. The polymerase active site with three subdomains: fingers (which bind an incoming nucleotide and interact with the single-stranded template), palm (which harbors the

https://doi.org/10.1038/s41467-023-40098-z

Fig. 5 | Comparison of the polymerase of the six species of crAssBcn phages. a. hylogenetic tree constructed with the polA of crAssBcn phage, other PolA and PolB proteins available in databases and reported in crass-like phages by Yutin et al., (2021). **b** Detail of the branch of the phylogenetic tree were crAssBcn phages are located showing the two separated branches containing the 706aa and the 775aa family A polymerases. **c** Multiallignment of PolA of the six species of crAssBcn phages. Upper part shows the polymerase A domains using the polymerase A gene of phage ΦCrAss001 as reference. From top to bottom this chart shows first the length of the protein. Small triangles indicate the aminoacids involved in conserved active, catalytic and DNA binding sites of each domain. Colored bars show the closest hits found in the conserved domain database (CCD) for each domain, these can be specific hits (with a high confident association) or non-specific and the superfamily to which the highest-ranking hit belongs. In the bottom part, the alignment of the polymerases of the six species (groups I-VI) and ΦCrAss001. Vertical red shadows bands highlight the gaps observed between the 775aa and the 706aa polA.

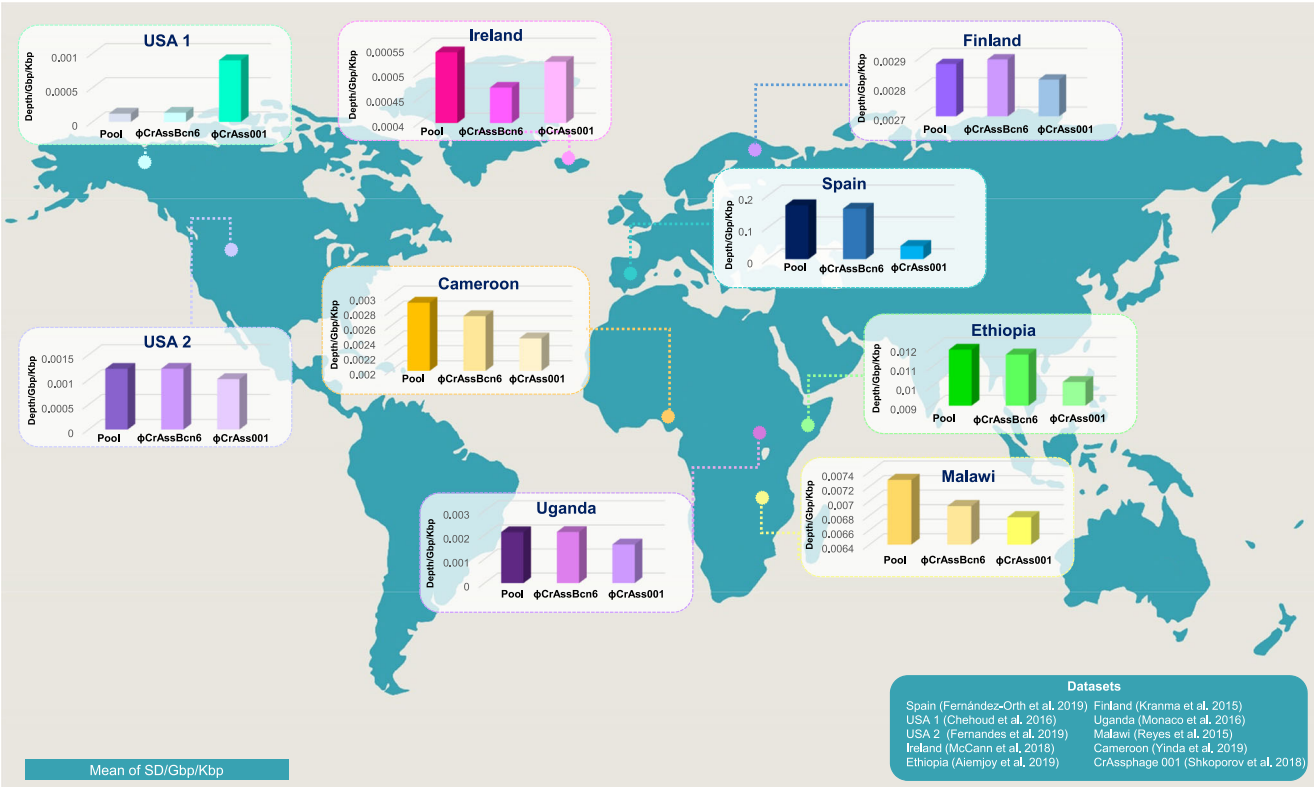


Fig. 6 | Geographical distribution of the new crAssBcn phages. The presence and abundance of the crAssBcn phages in different gut metagenomes from around the globe was analyzed by fragment recruitment of the pool of 25 crAssBcn phages (Pool), the representative ΦcrAssBcn6 and the model phage ΦCrAss001. Only metagenomes with positive recruitments are depicted, which are from the USA (USA 1 and 2), Finland, Spain, Ireland, Uganda, Malawi, Ethiopia, and Cameroon. Only mapped sequences with coverage of at least 70 % and best-hit reads were considered. Additionally, relative abundances of studied crAssphages were normalized by metagenome size (Gbp) and phage length (Kbp). World map figure in the background has been obtained from Designspot/Freepik.

catalytic amino acid residues and also binds an incoming dNTP) and thumb (which binds double-stranded DNA), and the additional associated domains for 3′–5′ exonucleolytic proofreading.

Both polymerases showed a hit for 3′–5′ exonuclease as well as a hit for the polymerase A domain (Supplementary Fig 7). Despite the observed differences in the 3′–5′ exonuclease and polymerase A domains between the two types (Fig. 5) the active sites of both domains appeared conserved.

The genes immediately upstream and downstream the polymerase, annotated as muconolactone isomerase and hypothetical protein, respectively, were also analyzed. While the ORF upstream the polymerase is conserved among the crAssBcn phages (Supplementary Data 2), the ORF immediate downstream the polymerase shows a high degree of variability among them, similar to the variability found in the ORF encoding the polymerase. This highly variable ORF was selected in our previous analysis (Hypothetical protein in Fig. 4), and encodes two types of proteins, one in species II, III and IV and another one in species I, V and VI. The nucleotide and aminoacidic sequences and the domain analysis did not, however, allow the identification of these proteins.

Biogeographical distribution of crAssBcn phages

In the assessment of the geographical distribution of crAssBcn phages, a positive hit was obtained in 51% of the metagenomes analyzed (Supplementary Data 3), including datasets from Ireland, USA (two datasets), Spain, Uganda, Malawi, Cameroon, Ethiopia and Finland (Fig. 6 and Supplementary Fig. 8). The pool of crAssBcn phages was more represented than ΦCrAss001, except in the USA-1 metagenome, as was ΦcrAssBcn6, with the exception of USA-1 and Ireland (Fig. 6). The largest recruitment and therefore the highest relative abundance corresponded to metagenomes from Spain, followed by Ethiopia and Malawi (Supplementary Fig. 8).

Discussion

The propagation strategy resulted in the successful isolation of 25 new crAss-like phages from wastewater. The abundance of crAss-like phages in wastewater using the different qPCR assays[12] differed because the assays target different crAss-like phage genomes. In fact, two of them[12,19] did not detect crAssBcn phages.

In this study, using different environmental and clinical isolates of *Bacteroides* and three different qPCR assays, only *B. intestinalis* allowed

efficient propagation of crAss-like phages detected by the qPCR assay targeting ΦCrAss001. Although other species of *Bacteroides* have been described as crAss-like phage hosts[15,16,23], *B. intestinalis* is also the host of ΦCrAss001, suggesting that this family comprises phages infecting this *Bacteroides* species. In their natural habitat, it is possible that these phages may infect other hosts[13] and it is also possible that other crAss-like phages infecting *B. intestinalis* would have been detected using a different set of primers and probes.

Attempts to isolate crAss-like phages from plaques of lysis generated directly from wastewater samples (without propagation) were unsuccessful, probably because the proportion of crAss-like phages among the entire phage pool infecting *Bacteroides* strains was too low and felt below the detection limit of the double agar layer technique after the dilutions performed to obtain separated plaques. Another factor hindering the isolation of crAss-like phages was the poor visibility of the generated plaques. The use of this *B. intestinalis* strain allowed the preferential propagation of crAss-like phages among other phages in the original sample, increasing their proportion and allowing their isolation. However, the use of a specific qPCR targeting ΦCrAss001 and this particular host promoted the isolation of phages related to ΦCrAss001. While other strategies based on the sequencing of the viruses present in bacterial cultures supernatants would have allowed the identification of other crAss-like phage genomes, they would not have guaranteed their isolation.

The average burst size of the crAssBcn phages was of 64.6 ± 16 phages per infected cell. In contrast, ΦCrAss001 shows a burst size of 2.5 phages per infected cell[13], although, microscopy observations indicated progeny counts of >50 phages per cell[13], closer to the burst size of crAssBcn phages. In addition, the propagation curves of crAssBcn phages consistently show a sequence of smaller bursts during the 9 h of incubation, starting at 2 h after infection and after 4 h and 6 h. Phage propagation did not seem to be detrimental for the growth of the strain, suggesting that the host population was maintained through a dynamic phage-host interplay, involving at least a fraction of the bacteria. These observations, in agreement with previous studies[13,18] indicate that crAssBcn phages are not temperate, and as lysogeny is ruled out, the persistence of the host bacteria during phage propagation may be due to a carrier-state infection, pseudolysogeny, or phase variation-based resistance involving receptor variation[24], alone or in combination[18]. Phase-variable systems that promote resistance, show a dynamic of spontaneous reversion of some resistant cells to a susceptible state. The sensible fraction of the bacterial population could generate enough susceptible cells to propagate the phage progeny while the resistant fraction maintains the bacterial population[24].

Samples of wastewater were obtained from different origins to increase the possibility of variation in any isolated phages. Although all belonged to the same genus and morphological group, the crAssBcn phages differed at the genomic and the protein level between them and when compared with crAss-like genomes of the *Kehishuvirus* genus.

The ORFs with the highest variability AAI (close to 50%) encoded variable proteins such as C-type lectin, which facilitates phage adaptation to the host[25], or thymidylate synthase genes, used by phages to acquire metabolites for DNA replication independently of the host. The variability of phage-hypermodified thymidylate synthases would expand the range of nucleotide modifications used to counteract bacterial defense systems[26].

A high level of variability was found in endonucleases, which are involved in DNA recombination, repair, and packaging, and have been proposed as a source of phage genetic diversity through positive selection pressure[4,27]. In addition, tail spikes were among the most variable genes, in accordance with the strong selective pressure these structures are under[14,28,29]. However, it was not expected to be so high in phages infecting the same host strain and presumably using the same receptor. A possible explanation for this observation is that capsular polysaccharides or other cell surface proteins regulated by phase variation act as highly dynamic crAss-like phage receptors[17,18]. In such a *Red Queen* scenario, the phage tail protein is adapting to constantly evolving target cell surface proteins[29]. This would suggest an evolution from a single ancestor, like the one observed for the tail protein. It is also possible that different tail spikes can attach to the same receptor or that spike variability allows phages to infect the bacteria through different receptors. This might have been the case of the variable tail spikes, located in different branches of the tree, plausibly as a consequence of one or more speciation events. Similarly, different variants located in different branches were observed for the hypothetical protein and the holin, with these variants showing hits with other phage sequences in the databases, reinforcing the hypothesis of speciation through orthologous gene exchange[30]. In a lower degree, the head protein can be traced back to a common ancestor, but have diverged from there in two different types, apparently after an integration event affecting the middle part of the protein.

Another unexpected result was the difference in the AAI (less than 50%) of the DNA polymerase A gene of species V and VI compared to the other species, considering they all belong to the same genus and infect the same host. Evolutionary switches in DNA polymerase from type A to B have been described in different families of crAss-like phages[7], as well as the absence of both enzyme types[31], but this is not the case of crAssBcn phages, that encoded only a family A DNA polymerase I. Polymerase A is present in the 25% of the dsDNA phages described[32], including crAss-like phages of the *Steigviridae* family[7]. However, differences in the DNA polymerase I of species V and VI were observed compared to the other species, these differences affect 3′−5′ exonuclease and polymerase A motifs. DNA polymerases of family A have replicative and repair activities. A 3′−5′ exonuclease activity that provides proofreading and repairs the mispaired nucleotides and a 5′−3′ DNA polymerase activity for DNA synthesis. Like in phage T7 and other organisms, the DNA polymerases of crAssBcn phages have a bipartite architecture with a C-terminal polymerase and N-terminal 3′-‘5′ exonuclease domains that are encoded in a single gene[33]. This structure supports the dual replicative and repair activities of the polymerase A in crAssBcn phages[33].

According to the propagation curves, the replication of the different crAssBcn phages did not show differences regardless the type PolA encoded and, in addition, the different domains in the polymerase seem to be active. We concluded that the differences in the two DNA polymerases do not affect their activity and that differences among phages is the result of evolutive or recombination events, resulting in the co-existence of different polymerases with the same function. Considering that similar sequences of both polymerase types can be found in other crAss-like phages, the possibility of recombination events between phages, or between bacterial hosts and phages seem plausible[7,34]. Moreover, the highly variable ORF located immediately downstream the polymerase gene might have been mobilized in the same module as the polymerase gene during a previous recombination event. The idea, proposed by Botstein in 1980[35], is that phages evolve by shuffling interchangeable functional modules, and that selection acts on those modules facilitating the emergence of new mosaic genotypes that provide advantages in each niche. The phylogenetic analysis suggests a common PolA ancestor evolved in an early stage from PolB. From this ancestor, the two PolA variants of this study diverged showing a dichotomy. Albeit the two PolA variants are found in different crAss-like phages, the PolA of 775 aa is more represented than the smaller one, suggesting it could be more successfully spread. The divergence between the two PolA types might result in a non-orthologous gene displacement, as it was previously proposed to explain the evolution of viral DNA polymerases[34,36,37]. Maintaining diversity in the replicative modules might be beneficial for phages as it would give them the possibility of adapting to changes in the process

of coevolution between bacteria and phages, or of escaping bacterial defenses[38].

DNA polymerase I has been proposed as a signature gene for the identification of viral sequences in metagenomic samples, the reconstruction of phylogenetic trees, and the exploration of viral diversity and evolution in different habitats[39,40]. However, the variability of DNA polymerase observed in the present study suggests that it may not be a suitable target for the universal detection of crAss-like phages[41], and that other more conserved genes, such as terminase[3], should be used in preference. Conversely, the RNA polymerase, a hotspot of positive selection and thus a variable region[29], was found to be only slightly variable among the crAssBcn phages.

When compared to other crAss-like phages of the *Kehishuvirus* genus, the crAssBcn phages did not match any of them, with the exception of ΦcrAssBcn20 classified as belonging to the same species (*primarius*) as ΦCrAss001, although differences can be observed, for instance in the head protein (Fig. 4). In agreement with previous studies[7], this observation highlights the need to use crAss-like phage genomes with a sufficient identity at a representative length of their genomes or genomes obtained from isolated crAss-like phages for a correct classification. Many of the publicly available complete crAssphage genome sequences have been cross-assembled from multiple individuals or represent pooled samples[29], which may introduce significant sequence variation and confound variant analysis. In other cases, crAss-like phage identification is based only on a genome fragment (sometimes very short).

The high prevalence of crAssBcn phages in global gut metagenomes, showing a higher representation than their closest relative ΦCrAss001, was striking. Phylogeographical analysis of crAss-like phages has revealed local clusters of genetically similar phages although a geographically widespread phage strain has been reported[8]. According to this, crAssBcn phages are highly abundant in Spain but are also unexpectedly detected in other countries. The conservation of such ubiquitous strains, including the crAssBcn phages, can be attributed to recent human migration movements, high fitness, or environmental stability[8].

Due to its ubiquity in the global human population and abundance in fecally polluted water, p-crAssphage has been proposed as a human fecal indicator. Other Crassvirales are also highly abundant in the mammalian gut, particularly, in humans[3,8,31,42].However, before other Crassvirales can be employed as human fecal markers, several issues need to be addressed. To date, crAssphage has been largely targeted in many reports, and reported as if it were a single virus, even though there is a great heterogeneity within the Crassvirales order. The few virions isolated until now are not p-crAssphage, infect different *Bacteroides* strains, and have low synteny and very low homology[16]. Assuming the diversity of crAss-like phages, the most abundant and widespread should be the best candidate to be selected as a potential fecal indicator. However, to establish which member of the Crassvirales order is the most representative on a global level will require extensive research. The detection of crAssBcn phages in metagenomes of different countries is another indication that new crAss-like phages, as yet unidentified, could be more abundant and ubiquitous than those described so far. Moreover, the 25 crAssBcn isolates illustrate crAss-like phages diversity within the same genus, which may hinder comprehensive screenings of crAssphage in contaminated water samples. The diversity of crAss-like phages should be taken under consideration particularly when designing new qPCR assays other than those already described[19]; for instance, some qPCR assays being used for crAssphage detection[8,12,19,43] would fail to detect crAssBcn phages, despite their geographical distribution.

Therefore, there is a pressing need to determine which member of the order should be used as fecal human marker and if there are potential new candidates globally widespread. If this cosmopolitan group of phages are to be used for microbial source tracking and wastewater-based epidemiology[44], an accurate molecular marker should probably be better identified by the genetic characterization of isolated crAss-like phages and the definition of a core crAss-like phage genome. Otherwise, a large amount of data identified as crAssphage, but which actually belong to different crAss-like viruses, with no comparability between locations and no correlation with human fecal pollution and fecal microorganisms may be generated, with the risk that the potential of this new human-specific indicator remains unfulfilled.

The discovery of crAssphage in the human gut has been a showcase achievement of metaviromics[1,4]. By integrating culture-based and computational efforts, the isolation of new crAss-like phages allows research to go beyond metagenomics and the sequence variation hurdle and may provide new insights into the life cycle of these human-gut phages.

Methods

No ethical approval was necessary for sample collection. The research presented complies with all relevant ethical regulations from the bioethics commission of the University of Barcelona.

Strains, bacteriophages, media, and culture conditions

Bacteroides strains of different species (Supplementary Table 2) were used in propagation cultures as hosts to detect crAss-like phages from wastewater and to evaluate their sensitivity to crAssBcn phages. Exopolysaccharide-producing *B. intestinalis* strain APC919/174 (DSM 108646) was selected as the definitive host for the isolation of crAssBcn phages, and bacteriophage ΦCrAss001 (DSM 109066)[13] was used as the positive control.

Anaerobe basal broth (ABB) (Thermofisher) was used for the growth of the *Bacteroides* strains. For the plaque assays, ABB containing 0.7% (w/v) of agar-agar for soft agar or 1.4% (w/v) of agar-agar for the agar plates were used.

Manipulation was performed in aerobic conditions while incubation of cultures and agar plates was done at 37 °C for 24 or 48 h in anaerobic jars (GasPak; BBL) with CO_2 atmosphere generators (Anaerocult A; Merck).

Wastewater samples

Twelve raw influent samples were obtained from five urban wastewater treatment plants (WWTPs) in Catalonia (NE Spain) over a period of one year. One of the WWTPs served a population of almost 384,000 inhabitants (Gavà); two plants served populations of between 100,000 and 290,000 inhabitants (Manresa and Igualada); and two plants served a population of more than 2,000,000 inhabitants (Besòs and Prat). Samples were collected in sterile containers, transported to the laboratory within two hours of collection and processed immediately for bacteriophage isolation as described below. Ten ml samples were filtered through 0.22 μm pore size, low-protein-binding (PES) membranes (Millipore) to remove bacteria and other particulate material.

Bacteriophage enumeration

One ml of phage suspension from wastewater was used to infect one ml of cultures of *Bacteroides* cultures at the middle-exponential growth phase (OD_{600} of 0.3) in 8 ml of ABB medium and incubated anaerobically and statically at 37 °C for 24 h. For subsequent enrichment cultures, 1 ml of the phage suspension from the first culture was filtered and used to infect a new *Bacteroides* culture under the same conditions.

For plaque assays, ten-fold serial dilutions of each phage suspension prepared with SM buffer (200 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris-HCl, pH 7.5) were enumerated by the double agar layer method[45] with the respective *Bacteroides* host strains in ABB soft agar. Plates were incubated anaerobically at 37 °C for 24 h. Spot assays were performed as described for plaque assays but without the addition of

phage to the ABB soft agar. A 10 µl drop of phage suspension was directly applied to the solidified lawn of each host strain and dried prior to incubation. Negative controls were prepared without the addition of phage.

Plaque blot hybridization

CrAss1-ORF46 probe is a digoxigenin (DIG)-labeled 50-bp probe (5′-ACCTGCTTCTACACTTTCCTTAGATGAACTAA-TATCTAACCCAGCTCTAT-3′) located in the ΦCrAss001 genome and commercially available. Plaques observed in the soft agar layer were transferred to a nylon membrane (Hybond N + , Amersham Pharmacia Biotech) and hybridized with the probe. Hybridization was performed at 53 °C, according to the standard procedure[46]. Stringent hybridization was achieved with the DIG-DNA Labeling and Detection Kit (Roche Diagnostics) according to the manufacturer's instructions.

Plaques showing a positive signal were recovered from the soft agar overlayer with a sterile loop, resuspended in 200 µl of SM buffer and submitted to a chloroform treatment to eliminate bacterial cells. The isolated plaques confirmed to be crAss-like phages by qPCR were further propagated in larger volumes of *B. intestinalis* culture for subsequent analysis.

Purification of phage particles

Suspensions containing each individual phage were further purified by cesium chloride (CsCl) density gradients using ultra clear thin wall tubes (Beckman), 1 ml of 20% (w/v) sucrose and three densities of CsCl (1.3, 1.5, and 1.7 g/ml)[46]. Samples were ultracentrifuged at 22,000 x *g* for 2 h at 4 °C in a Swinging-Bucket SW-41 Rotor in a Beckman ultracentrifuge.

The visible grey bands corresponding to bacteriophages[46] were collected by puncturing the tube, obtaining a 0.5 ml volume that was dialyzed using prepared dialysis membranes (MWC 12–14 kDa) (Thermofisher) in dialysis buffer (Tris 0.1 M, EDTA 0.2 mM, pH 8) for 2 h. The dialysis buffer was replaced with fresh buffer and further dialyzed for 18 h with magnetic stirring.

Infectivity assays

To evaluate the dynamics of infection of the 25 crAssBcn phages, each phage suspension was used to infect 1 ml of a middle-exponential growth phase culture of *B. intestinalis* grown in ABB at a multiplicity of infection (MOI) of 0.001. Tubes were incubated anaerobically at 37 °C, and the growth was monitored by measuring the OD of the cultures at 600 nm. In parallel, the number of colony-forming units grown in ABB agar, and the number of plaques of lysis plaques obtained by the double agar layer method (pfu/ml) were evaluated at intervals for 24 h. As a control, a *B. intestinalis* culture without phage was grown under the same conditions. All phage counts were normalized by the increase in concentration at each time point with respect to the initial value. Data were analyzed using GraphPad Prism 9 (GraphPad Software, www.graphpad.com). Comparisons between the average increase in the concentration of all crAssBcn phages *vs* ΦCrAss001 and among crAssBcn phages were evaluated with a Wilcoxon matched-pairs test and Friedman test with Dunn's multiple comparison test.

To calculate the burst size, adsorption assays were performed after the first burst (approximately at 150 min) at a MOI of 0.001 with *B. intestinalis* grown anaerobically to middle-exponential growth phase, in accordance with Kropinski[47], with modifications to adjust to the anaerobic conditions and the growth rate of *B. intestinalis*. Free phages were quantified by the double agar layer method as described above.

Electron microscopy observations

Ten µl of the concentrated CsCl phage suspensions were dropped onto copper grids with carbon-coated Formvar films and negatively stained with 2% ammonium molybdate (pH 6.8) for 2 min. Phages were visualized using a Jeol 1010 TEM (JEOL Inc.) operating at 80 kV.

Phage DNA isolation

DNA isolation was performed with the QIAamp DNA blood mini kit (Qiagen GmbH,), following the manufacturer's instructions. The DNA was suspended in a final volume of 200 µl of sterile bidistilled water. The DNA concentration of each pooled sample was evaluated using a Qubit® Fluorometer (Life Technologies) and the DNA quality was further confirmed by the 2100 Bioanalyzer system (Agilent Technologies).

PCR and qPCR assays

A PCR assay (UP: 5′-ACCTGCTTCTACACTTTCCTT-3′/LP: 5′-AGTGCTCCAGAATAGGATTGT-3′) and a qPCR assay using TaqMan hydrolysis probe (UP: 5′-ACCTGCTTCTACACTTTCCTT-3′/LP: 5′-AGTGCTCCAGAATAGGATTGT-3′/Probe: 6FAM- ATATCTAACC-CAGCTC-MGBNFQ) was designed from the ΦCrAss001 genome (NC_049977.1) targeting a gene coding for a hypothetical protein (ORF 46) and previously described qPCR assays for the detection of crAss-like phages[12,19] were used. Primers and probes were confirmed as specific for the detection crAss-like phages available in genomic databases.

PCR amplification was performed using DreamTaq Green DNA Polymerase (Fermentas) in a GeneAmp PCR system 2700 (Applied Biosystems). qPCR amplifications were carried out using the standard run in the StepOne™ Real Time PCR System (Applied Biosystems) in a 20 µl reaction mixture with TaqMan® Environmental Master Mix 2.0 (Applied Biosystems). The reaction contained 9 µl of the sample DNA or standards with known DNA concentration prepared from gBlocks™ Gene Fragments used for quantification. The results were analyzed with the Applied Biosystems StepOne™ Instrument program. All samples were run in triplicate (including the standards and negative controls). The number of gene copies (GC) was defined as the mean of the triplicate data obtained.

Sequencing

Five µl of DNA at a concentration of 0.2 ng/µl was fragmented and used to prepare libraries for whole genome sequencing with the Kapa Hyper Plus Kit (Roche) according to the manufacturer's protocol. Libraries were purified using AmPure beads (Beckman Coulter Inc.), checked for fragment distribution and size and quantified in a TapeStation 4200 and the Agilent High Sensitivity D1000 ScreenTape system (Agilent Technologies) in a Quantus™ Fluorometer (Promega). An equimolar pool of the individual 25 phage genomes was separately sequenced by NextSeq System (Illumina)with a high output run of 300 cycles.

Sequence trimming, genome recovery and functional annotation

Raw reads were trimmed by Trimmomatic (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36)[48]. The quality of trimmed reads was checked by FastQC[49]. Paired-end reads were joined using fq2fa from the idba package v1.1.3[50]. Additionally, pair-end filtered reads were assembled individually by SPAdes v3.13.0 (-k 21,33,55,77,99,127)[51]. In order to recover additional complete viral genomes from the assembly sequences, genomes were assembled applying the Trusted contigs strategy[52], available in SPADEs, that used the ΦCrAss001 genome[13] as a reference to provide a guidance during de novo assembly. ORF prediction was carried out using Prodigal[53]. Genomic maps were generated using Geneious Prime version 20231.1.

Clustering and intragenome comparison of crAssBcn phages

Phage genomes were clustered at 90% by cd-hit-est[54], the phage genome similarities were calculated using VIRIDIC[55] (BLASTn-based). VIRIDIC[55] calculates the intergenomic similarity of two viruses using BLASTn and checks each comparison for genomic synteny[56]. The reciprocal best match strategy was chosen for the ORF comparison in aminoacids, (rbm.rb script from enveomics package[57], selected default parameters). Taxonomic assignation of recovered phage genomes was

done by VIPtree (genome-wide similarity-based)[58]. Multiple alignments of selected protein sequences were performed by MUSCLE[59]. Phylogenetic tree reconstruction was drawn using IQ-TREE[60] and Geneious Prime 2023.1.1 (tree build method neighbor-joining). The polymerase phylogenetic analysis includes a database of polymerases A and B previously described[7]. Additionally, non-redundant database (updated in April 2023)[61] (BLASTp) best hit matches were also included in the analyses, considering those hits showing a genome coverage larger than 50% and an identity over 50%. In case that no sequences accomplished these criteria, then the ten best hits were selected.

Functional protein annotation and structure prediction

Protein sequences were searched using the InterProScan software v5.47-82.0[62] to identify signatures from the InterPro member databases; Pfam[63], SMART[64], TIGRFAMs[65] and CDD[61]. Additionally, all crAssBcn phages ORFs were mapped against all ΦCrAss001 genes (annotation update January 2023) and taking in consideration recent identification of previously mis-annotated proteins encoded by the ΦCrAss001 genome[14] in order to obtain a more accurate functional gene annotation. Prediction of the structure and conserved active domains of the polymerase encoded in the crAssBcn phages was performed using Alphafold[66], and the NCBI's Conserved Domain Database and SPARCLE[67].

Presence and abundance of crAssBcn phages around the world

To evaluate the geographical distribution of the crAssBcn phages, we performed a fragment recruitment of the crAssBcn phages against a viral metagenome collection that we created consisting of 1,255 human gut viromes from children and adults and in 14 countries (Supplementary Data 3). These metagenomes were mapped against 1) the genomes of the 25 crAssBcn phages pooled together (Pool), 2) the ΦCrAssBcn6 genome representing species 1, which formed the larger group of crAssBcn phages, and 3) the ΦCrAss001 genome, using standalone BLASTn with a cutoff of 70% query coverage, e-value $\geq 10^{-1}$ and filtered by the 'best hit' option. Next, crAssBcn phage abundances were calculated using sequencing depth values normalized by dataset size (Gbp) and genome length (Kbp) (sequencing depth/Gbp/Kbp). Fragment recruitment data were plotted by the enveomics.R package in the R statistical tool[57]. Additionally, the graphs (bars, boxplots and heatmaps) were drawn with Plotly by R[68] and heatmapper[69].

Statistics and Reproducibility

Statistical analyses were performed using the GraphPad Prism 9 (GraphPad Software, San Diego, CA, US). Unpaired t-tests were conducted to identify differences between treatments. Significant differences were set at $p < 0.05$. Experimental data presented are the average of three independent experiments. Bioinformatic analysis were performed in duplicate and are reproducible. At least five electron micrographs were taken for each phage, and a selected one for each phage is presented in Fig. 1.

No statistical method was used to predetermine sample size, no data were excluded from the analyses; the experiments were not randomized, and the Investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The crAssBcn phages genomes generated in this study have been deposited in GenBank" with the GenBank accession codes available in Supplementary Table 3 and are publicly available. Other supporting data generated in this study (tree gene sequences, tree alignments and values) can be found in https://data.cyverse.org/dav-anon/iplant/home/lolesramosub/Ramos-Barbero%2C%20MD%2C%20G%C3%B3mez-G%C3%B3mez%2C%20C.%20%2C%20Sala%2C%20L%20%28...%29%26%20Muniesa%2C%20M.%202023%20Supporting%20information/Ramos-Barbero%202023%20Supporting%20information.rar. The following databases were used: Pfam (http://pfam.xfam.org/), SMART (http://smart.embl-heidelberg.de/) TIGRFAMs (https://www.jcvi.org/research/tigrfams) and CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Database of metagenomes in Fig. 6 is available in Supplementary Data 3. Source data for supplementary figure 2 are provided in the Source Data file. Source data for supplementary figure 6 are provided in Supplementary Data 2. The authors declare that all other data supporting the findings of this study are available within the paper and its supplementary files. Source data are provided with this paper.

References

1. Stachler, E., Akyon, B., De Carvalho, N. A., Ference, C. & Bibby, K. Correlation of crAssphage qPCR markers with culturable and molecular indicators of human fecal pollution in an impacted urban watershed. *Environ. Sci. Technol.* **52**, 7505–7512 (2018).
2. Sabar, M. A., Honda, R. & Haramoto, E. CrAssphage as an indicator of human-fecal contamination in water environment and virus reduction in wastewater treatment. *Water Res* **221**, 118827 (2022).
3. Guerin, E. et al. Biology and taxonomy of crAss-like bacteriophages, the most abundant virus in the human gut. *Cell Host Microbe* **24**, 653–664 (2018).
4. Dutilh, B. E. et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* **5**, 4498 (2014).
5. Turner, D. et al. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch. Virol.* **168**, 74 (2023).
6. Current ICTV Taxonomy Release | ICTV. https://ictv.global/taxonomy.
7. Yutin, N. et al. Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nat. Commun.* **12**, 1044 (2021).
8. Edwards, R. A. et al. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat. Microbiol* **4**, 1727–1736 (2019).
9. Park, G. W. et al. CrAssphage as a novel tool to detect human fecal contamination on environmental surfaces and hands. *Emerg. Infect. Dis.* **26**, 1731–1739 (2020).
10. Sala-Comorera, L. et al. crAssphage as a human molecular marker to evaluate temporal and spatial variability in faecal contamination of urban marine bathing waters. *Sci. Total Environ.* **789**, 147828 (2021).
11. Morrison, C. M. et al. Potential indicators of virus transport and removal during soil aquifer treatment of treated wastewater effluent. *Water Res* **177**, 115812 (2020).
12. García-Aljaro, C., Ballesté, E., Muniesa, M. & Jofre, J. Determination of crAssphage in water samples and applicability for tracking human faecal pollution. *Micro. Biotechnol.* **10**, 1775–1780 (2017).
13. Shkoporov, A. N. et al. ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides intestinalis*. *Nat. Commun.* **9**, 4781 (2018).
14. Bayfield, O. W. et al. Structural atlas of a human gut crassvirus. *Nature* **617**, 409–416 (2023).
15. Hryckowian, A. J. et al. *Bacteroides thetaiotaomicron*-infecting bacteriophage isolates inform sequence-based host range predictions. *Cell Host Microbe* **28**, 371 (2020).
16. Guerin, E. et al. Isolation and characterisation of ΦcrAss002, a crAss-like phage from the human gut that infects *Bacteroides xylanisolvens*. *Microbiome* **9**, 89 (2021).

17. Papudeshi, B. et al. Novel crAssphage isolates exhibit conserved gene order and purifying selection of the host specificity protein. *bioRxiv* 2023.03.05.531146 https://doi.org/10.1101/2023.03.05.531146 (2023).
18. Shkoporov, A. N. et al. Long-term persistence of crAss-like phage crAss001 is associated with phase variation in *Bacteroides intestinalis*. *BMC Biol.* **19**, 163 (2021).
19. Stachler, E. et al. Quantitative CrAssphage PCR assays for human fecal pollution measurement. *Environ. Sci. Technol.* **51**, 9146–9154 (2017).
20. Ahmed, W., Payyappat, S., Cassidy, M., Besley, C. & Power, K. Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. *Water Res* **145**, 769–778 (2018).
21. Ballesté, E. et al. Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. *Water Res* **155**, 233–244 (2019).
22. Lefkowitz, E. J. et al. Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res* **46**, D708–D717 (2018).
23. Hedžet, S., Rupnik, M. & Accetto, T. Broad host range may be a key to long-term persistence of bacteriophages infecting intestinal *Bacteroidaceae* species. *Sci. Rep.* **12**, 1–11 (2022).
24. Porter, N. T. et al. Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in *Bacteroides thetaiotaomicron*. *Nat. Microbiol* **5**, 1170 (2020).
25. McMahon, S. A. et al. The C-type lectin fold as an evolutionary solution for massive sequence variation. *Nat. Struct. Mol. Biol.* **12**, 886–892 (2005). *2005 12:10*.
26. Lee, Y. J. et al. Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses. *Proc. Natl Acad. Sci. USA* **115**, E3116–E3125 (2018).
27. Wilson, G. W. & Edgell, D. R. Phage T4 mobE promotes trans homing of the defunct homing endonuclease I-TevIII. *Nucleic Acids Res* **37**, 7110 (2009).
28. Siranosian, B. A., Tamburini, F. B., Sherlock, G. & Bhatt, A. S. Acquisition, transmission and strain diversity of human gut-colonizing crAss-like phages. *Nat. Commun.* **11**, 280 (2020).
29. Brown, B. P. et al. crAssphage genomes identified in fecal samples of an adult and infants with evidence of positive genomic selective pressure within tail protein genes. *Virus Res* **292**, 198219 (2021).
30. Gabaldón, T. & Koonin, E. V. Functional and evolutionary implications of gene orthology. *Nat. Rev. Genet* **14**, 360–366 (2013).
31. Koonin, E. V. & Yutin, N. The crAss-like phage group: how metagenomics reshaped the human virome. *Trends Microbiol* **28**, 349–359 (2020).
32. Wommack, K. E., Nasko, D. J., Chopyk, J. & Sakowski, E. G. Counts and sequences, observations that continue to change our understanding of viruses in nature. *J. Microbiol* **53**, 181–192 (2015).
33. Morcinek-Orłowska, J., Zdrojewska, K. & Węgrzyn, A. Bacteriophage-encoded DNA polymerases-beyond the traditional view of polymerase activities. *Int J. Mol. Sci.* **23**, 635 (2022).
34. Filée, J., Forterre, P., Sen-Lin, T. & Laurent, J. Evolution of DNA polymerase families: Evidences for multiple gene exchange between cellular and viral proteins. *J. Mol. Evol.* **54**, 763–773 (2002).
35. Botstein, D. A theory of modular evolution for bacteriophages. *Ann. N. Y Acad. Sci.* **354**, 484–491 (1980).
36. Koonin, E. V. Non-orthologous gene displacement. *Trends Genet.* **12**, 334–336 (1996).
37. Koonin, E. V. Temporal order of evolution of DNA replication systems inferred by comparison of cellular and viral DNA polymerases. *Biol. Direct* **1**, 39–39 (2006).
38. Smug, B. et al. Protein modularity in phages is extensive and associated with functions linked to core replication machinery and host tropism 2 determinants. *bioRxiv 12.27.521992* https://doi.org/10.1101/2022.12.27.521992 (2022).
39. Nasko, D. J. et al. Family A DNA polymerase phylogeny uncovers diversity and replication gene organization in the virioplankton. *Front Microbiol* **9**, 3053 (2018).
40. Adriaenssens, E. M. & Cowan, D. A. Using signature genes as tools to assess environmental viral ecology and diversity. *Appl Environ. Microbiol* **80**, 4470–4480 (2014).
41. Liang, Y., Zhang, W., Tong, Y. & Chen, S. CrAssphage is not associated with diarrhoea and has high genetic diversity. *Epidemiol. Infect.* **144**, 3549–3553 (2016).
42. Yutin, N. et al. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat. Microbiol* **3**, 38–46 (2018).
43. Cinek, O. et al. Quantitative CrAssphage real-time PCR assay derived from data of multiple geographically distant populations. *J. Med Virol.* **90**, 767–771 (2018).
44. Sims, N. & Kasprzyk-Hordern, B. Future perspectives of wastewater-based epidemiology: Monitoring infectious disease spread and resistance to the community level. *Environ. Int* **139**, 105689 (2020).
45. Adams, M. H. *Bacteriophages*. (New York, Interscience Publishers, 1959).
46. Sambrook, J. & Russell, D. Molecular cloning: A laboratory manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY* 999 (2001).
47. Kropinski, A. M. Practical advice on the one-step growth curve. *Methods Mol. Biol.* **1681**, 41–47 (2018).
48. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
49. Wingett, S. & Andrews, A. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* **7**, 1338 (2018).
50. Peng, Y., Leung, H. C. M., Yiu, S. M. & Chin, F. Y. L. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420–1428 (2012).
51. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput Biol.* **19**, 455–477 (2012).
52. Çabuk, U. & Ünlü, E. S. A combined de novo assembly approach increases the quality of prokaryotic draft genomes. *Folia Microbiol (Praha)* **67**, 801–810 (2022).
53. Hyatt, D. et al. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma.* **11**, 119 (2010).
54. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
55. Moraru, C., Varsani, A. & Kropinski, A. M. VIRIDIC-A novel tool to valculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* **12**, 1268 (2020).
56. Adriaenssens, E. M. & Rodney Brister, J. How to name and classify your phage: an informal guide. *Viruses* **9**, 70 (2017).
57. Rodriguez-R, L. M. & Konstantinidis, K. T. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Prepr* https://doi.org/10.7287/peerj.preprints.1900v1 (2016).
58. Nishimura, Y. et al. ViPTree: The viral proteomic tree server. *Bioinformatics* **33**, 2379–2380 (2017).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797 (2004).
60. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

61. Agarwala, R. et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7–D19 (2016).
62. Mulder, N. & Apweiler, R. InterPro and InterProScan: tools for protein sequence classification and comparison. *Methods Mol. Biol.* **396**, 59–70 (2007).
63. Finn, R. D. et al. Pfam: the protein families database. *Nucleic Acids Res* **42**, D222 (2014).
64. Letunic, I. & Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res* **46**, D493–D496 (2018).
65. Haft, D. H., Selengut, J. D. & White, O. The TIGRFAMs database of protein families. *Nucleic Acids Res* **31**, 371–373 (2003).
66. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). *2021 596:7873*.
67. Marchler-Bauer, A. et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* **45**, D200 (2017).
68. Plotly – Collaborative Data Science – 2015 Data Storytelling Studio @ MIT. https://datastudio2015.datatherapy.org/index.html%3Fp=745.html.
69. Babicki, S. et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* **44**, 147–153 (2016).

Acknowledgements

This research was funded by the Spanish Ministerio de Innovación y Ciencia (PID2020-113355GB-I00) the Agencia Estatal de Investigación (AEI) and the European regional fund (ERF). C. G-G has a fellowship from the University of Barcelona. MD.R.B has a Margarita Salas fellowship from the Universitat d' Alacant. L. S-C has a Maria Zambrano fellowship. G. V has a FPI grant from the Spanish Ministerio de Universidades. S. M-C has a grant from Colciencias (Republic of Colombia) and L. R-R. is lecturer of the Serra-Hunter program, Generalitat de Catalunya. M. M. is a researcher of the ICREA Academia 2022 program. Authors want to acknowledge Prof. J. Anton and Dr. M. Martinez-Garcia from University of Alicante for their valuable help and support with the use of their servers for the bioinformatic analysis of this study, and Prof. F. Navarro and Prof. L. Comstock for some of the strains used in this study.

Author contributions

M.D.R-B. conducted all bioinformatics analyses, prepared figures, and contributed to writing and revising the paper. C.G-G. isolated all phages, performed characterization and infectivity experiments and prepared figures. L.S.-C. conducted infection experiments, calculated all statistical analysis, prepared figures, and contributed to writing the paper. M.D-.R-B., C.G-G., and L.S-C. contributed equally to this work. L.R-R. wrote and revised the paper and discussed results. S.M-C. evaluated phages by qPCR and wrote sections of the paper. E.dM. and D.T-A. contributed to microscopy studies and phage characterization. G.V. performed the sampling, prepared media and performed some infection assays. A.R.B. gave technical support and conceptual advice and revised the paper. E.B. gave technical support and conceptual advice and commented on the manuscript at all stages. C.G-A. funded the study, discussed results, and commented on the manuscript at all stages. M.M. is the corresponding author, supervised and funded the study, designed experiments, analyzed data, and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-40098-z.

Correspondence and requests for materials should be addressed to Maite Muniesa.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

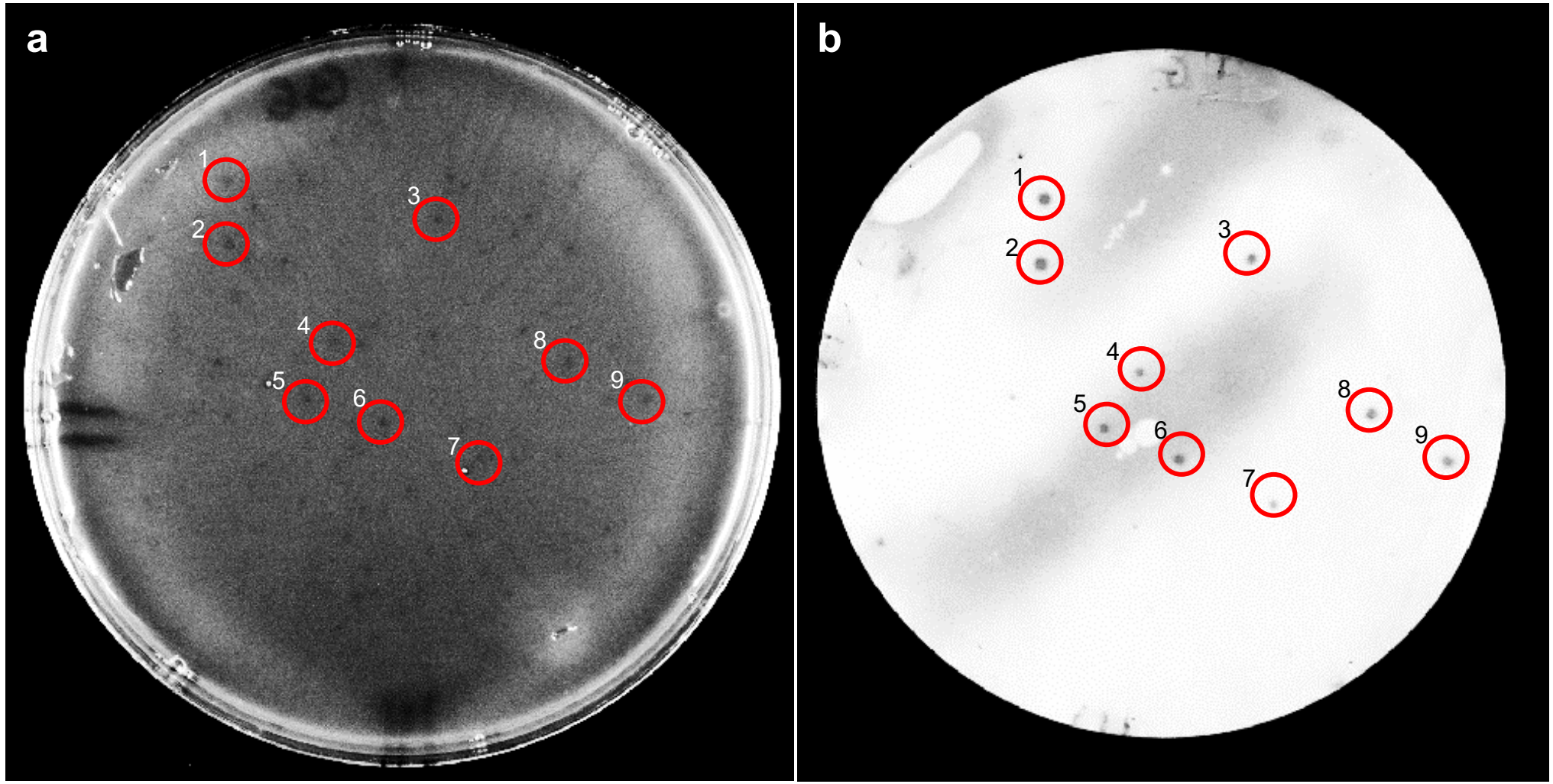
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

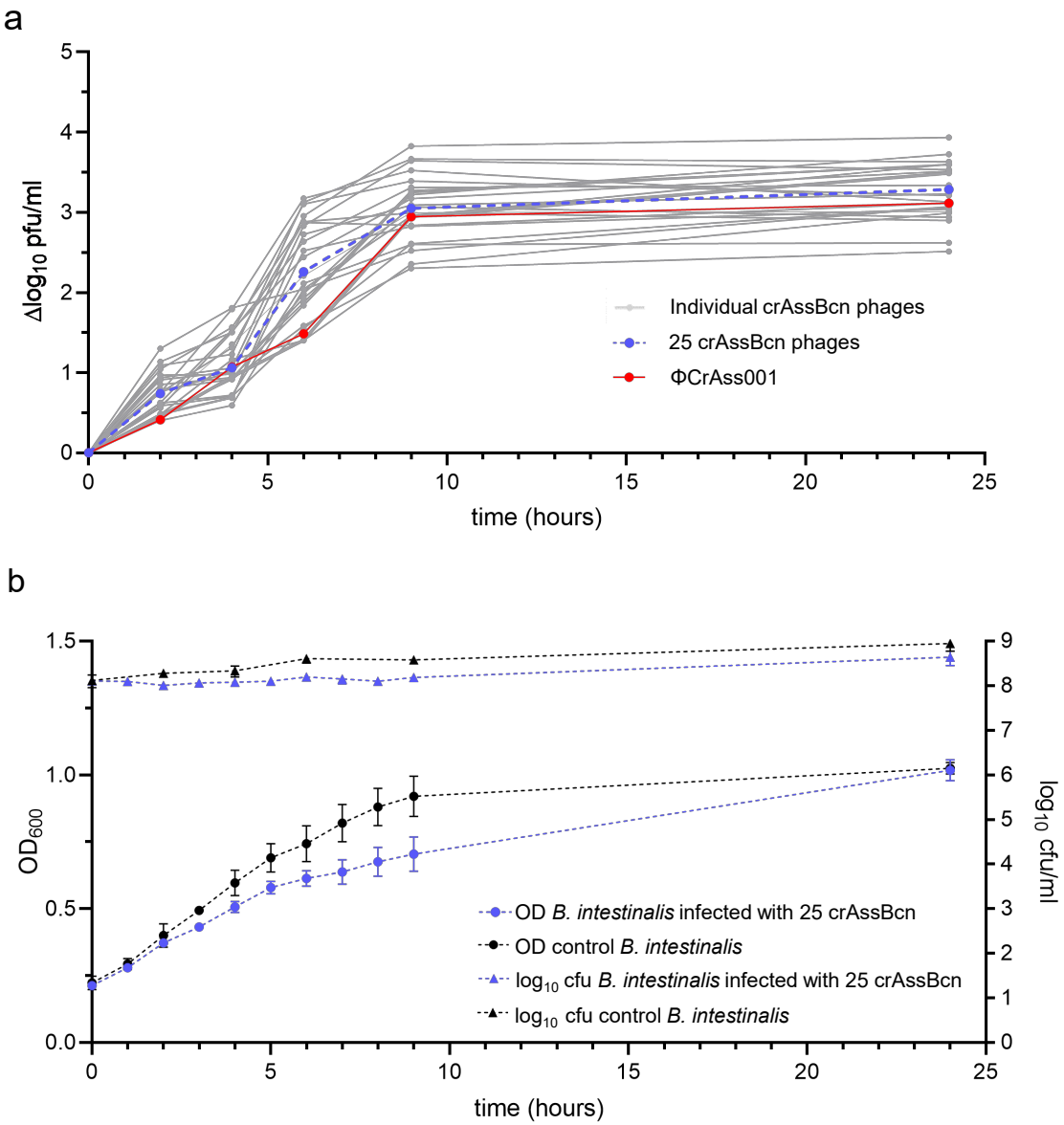
© The Author(s) 2023, corrected publication 2023

SUPPLEMENTARY INFORMATION

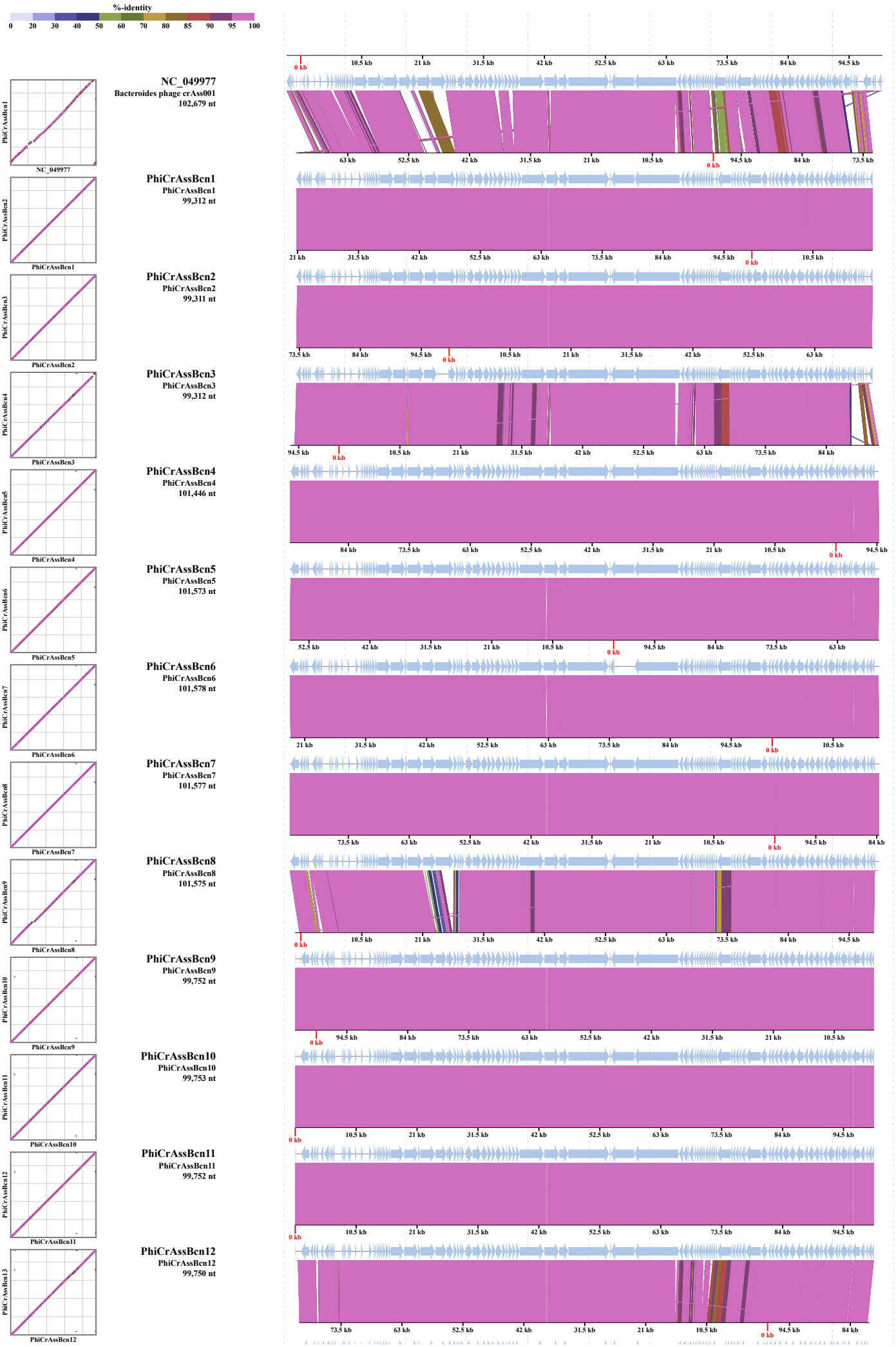
Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution

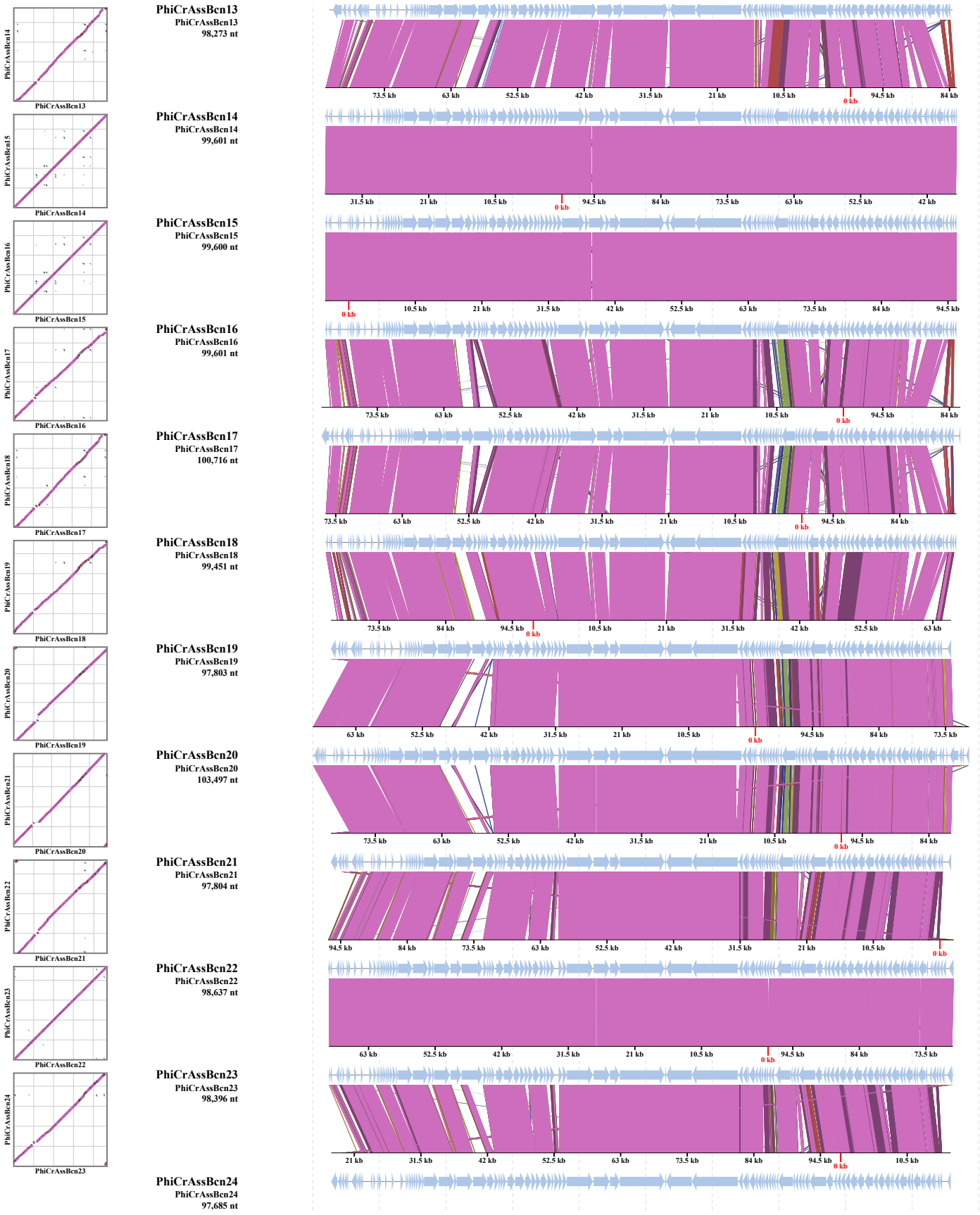


Supplementary Fig. 1. Detection of crAssphage lysis plaques. Representative morphology of the plaques produced by the crAssBcn phages in *B. intestinalis* and plaque blot hybridization. (a) Plaque enumeration by the double agar layer method of phage ΦCrAssBcn6 infecting *B. intestinalis* from a wastewater sample (WWTP Gava 1). (b) Plaque blot hybridization of phages transferred from (a) using the CrAss1-ORF46 probe.

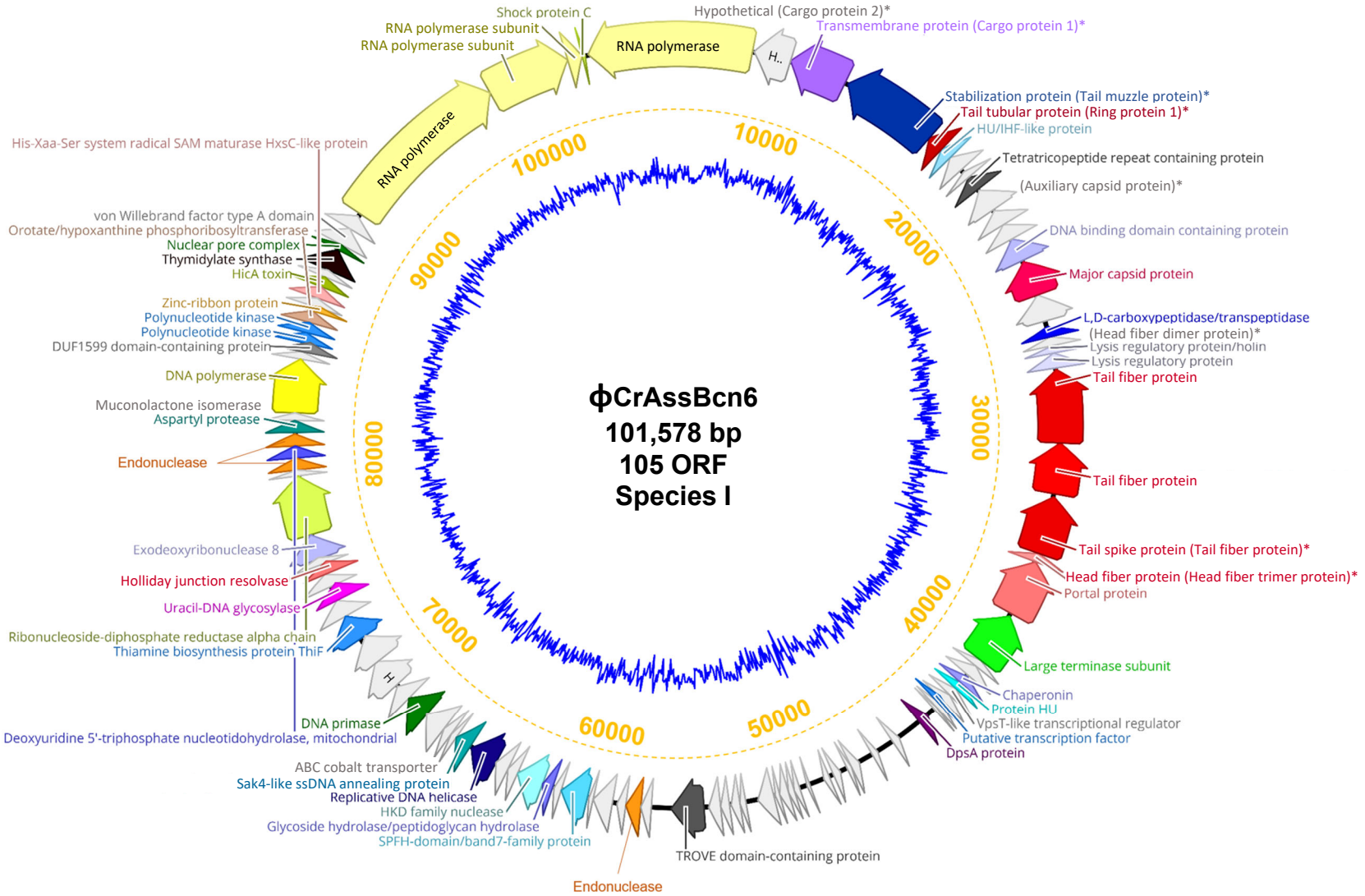


Supplementary Fig. 2. Infectivity dynamics of crAssBcn phages. (a) Increase in the number of infectious phage particles measured during 24 h by the double agar layer method and expressed in $\log_{10}$ pfu/ml. Values in the grey lines (crAssBcn phages) and red (ΦCrAss001) are the average of three replicates performed with each individual phage. Blue dotted line corresponds to the median of the 25 crAssBcn phages (b) Growth of the host strain *B. intestinalis* plotted on the left Y-axis was measured by the increase in optical density (OD) at 600 nm and values are presented with circles. Growth of the host strain *B. intestinalis* plotted on the right Y-axis was measured by colony counts (cfu/ml) and values are presented with triangles. Each value corresponds to the median of seven bacterial cultures uninfected (black) or infected with the 25 crAssBcn phages (blue). Source data are provided as a Source Data file.

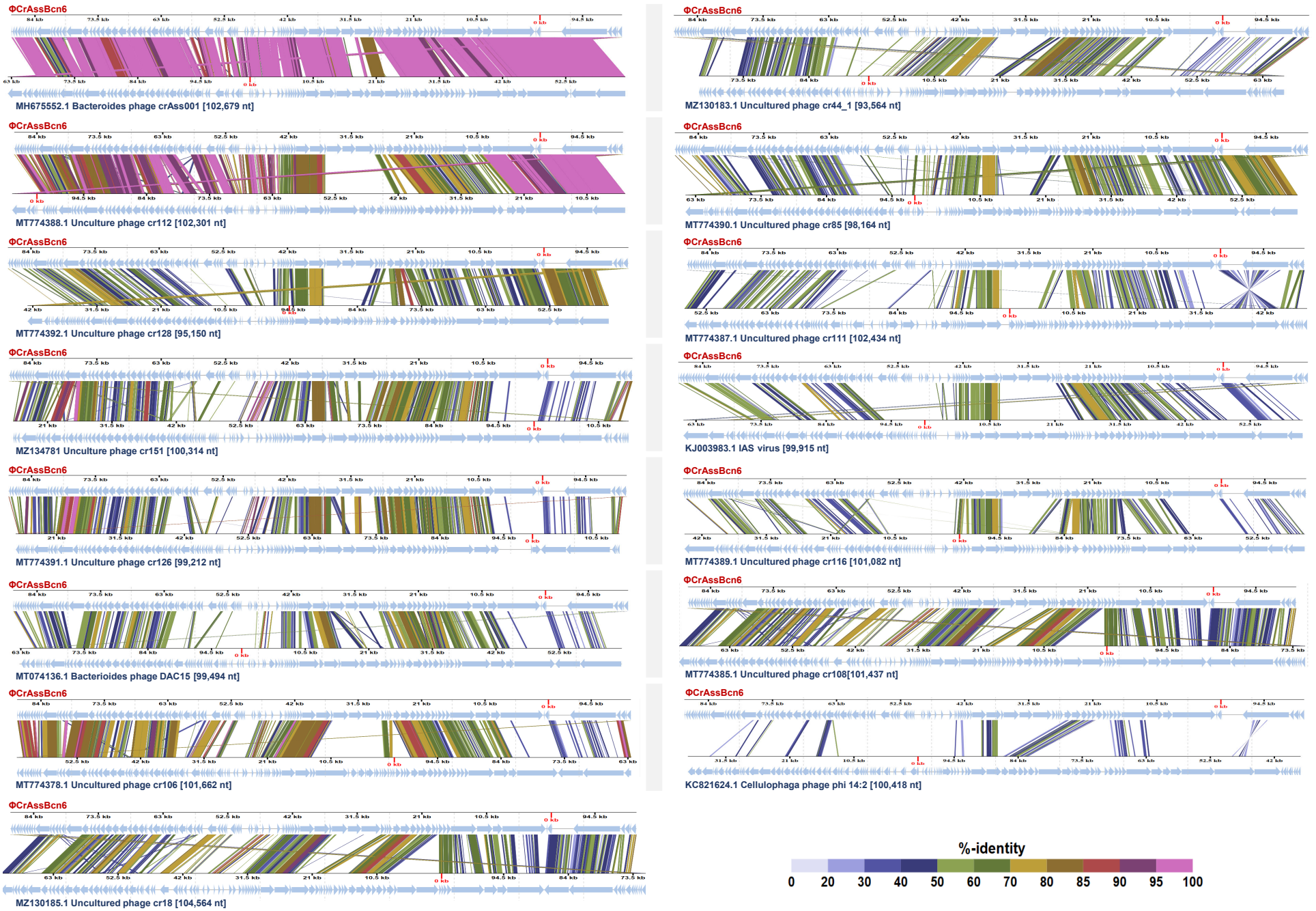

%-identity
0 20 30 40 50 60 70 80 85 90 95 100
NC_049977
Bacteroides phage crAss001
102,679 nt
PhiCrAssBcn1
PhiCrAssBcn1
99,312 nt
PhiCrAssBcn2
PhiCrAssBcn2
99,311 nt
PhiCrAssBcn3
PhiCrAssBcn3
99,312 nt
PhiCrAssBcn4
PhiCrAssBcn4
101,446 nt
PhiCrAssBcn5
PhiCrAssBcn5
101,573 nt
PhiCrAssBcn6
PhiCrAssBcn6
101,578 nt
PhiCrAssBcn7
PhiCrAssBcn7
101,577 nt
PhiCrAssBcn8
PhiCrAssBcn8
101,575 nt
PhiCrAssBcn9
PhiCrAssBcn9
99,752 nt
PhiCrAssBcn10
PhiCrAssBcn10
99,753 nt
PhiCrAssBcn11
PhiCrAssBcn11
99,752 nt
PhiCrAssBcn12
PhiCrAssBcn12
99,750 nt



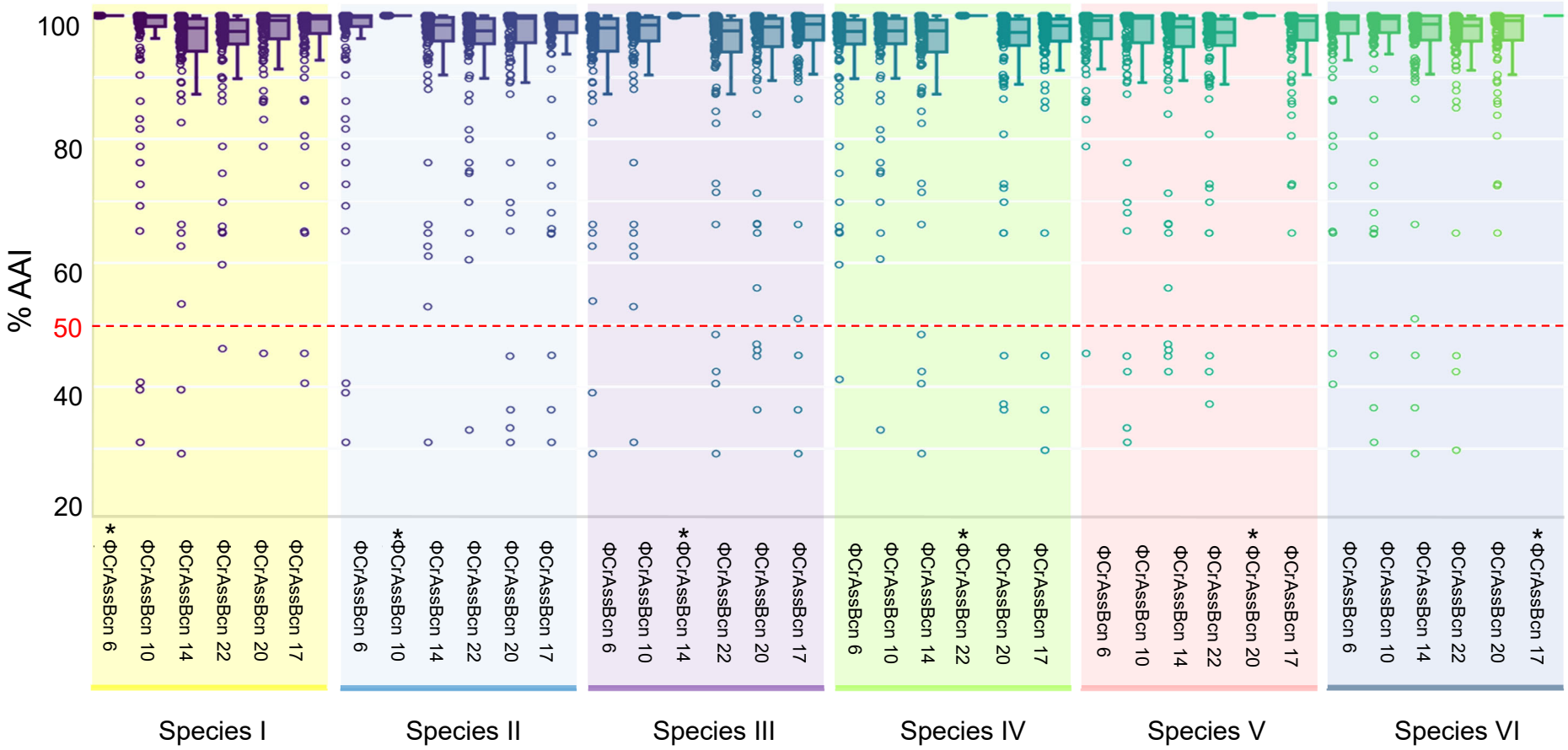
Supplementary Figure 3. Multiple genome alignment of the 24 CrAssBcn phages .CDS alignment of the 24 complete genomes and ΦCrAss001. Left charts show pairwise alignment of two consecutive genomes. Light blue arrows in the genetic maps (right side) indicate the ORFs in each genome. The color of the bands between the genetic maps shows the percentage of identity between each sequence as indicated in the legend at the top of the figure.



Supplementary Fig. 4. Circular genomic map of ΦCrAssBcn6. Genomic map of the representative phage of the most abundant species (group I) among the crAssBcn phages. The map was generated using Geneious Prime 2022.0.2. Different colored arrows show the predicted ORFs: yellow, with replicative functions; red, structural genes; orange, endonucleases; and grey, hypothetical proteins. The orange dashed ring indicates the genome position, and the inner blue ring displays the variations in the % Guanine/Cytosine. (*) In brackets the ORFs with new annotation reported in Bayfield *et al*., 2022.

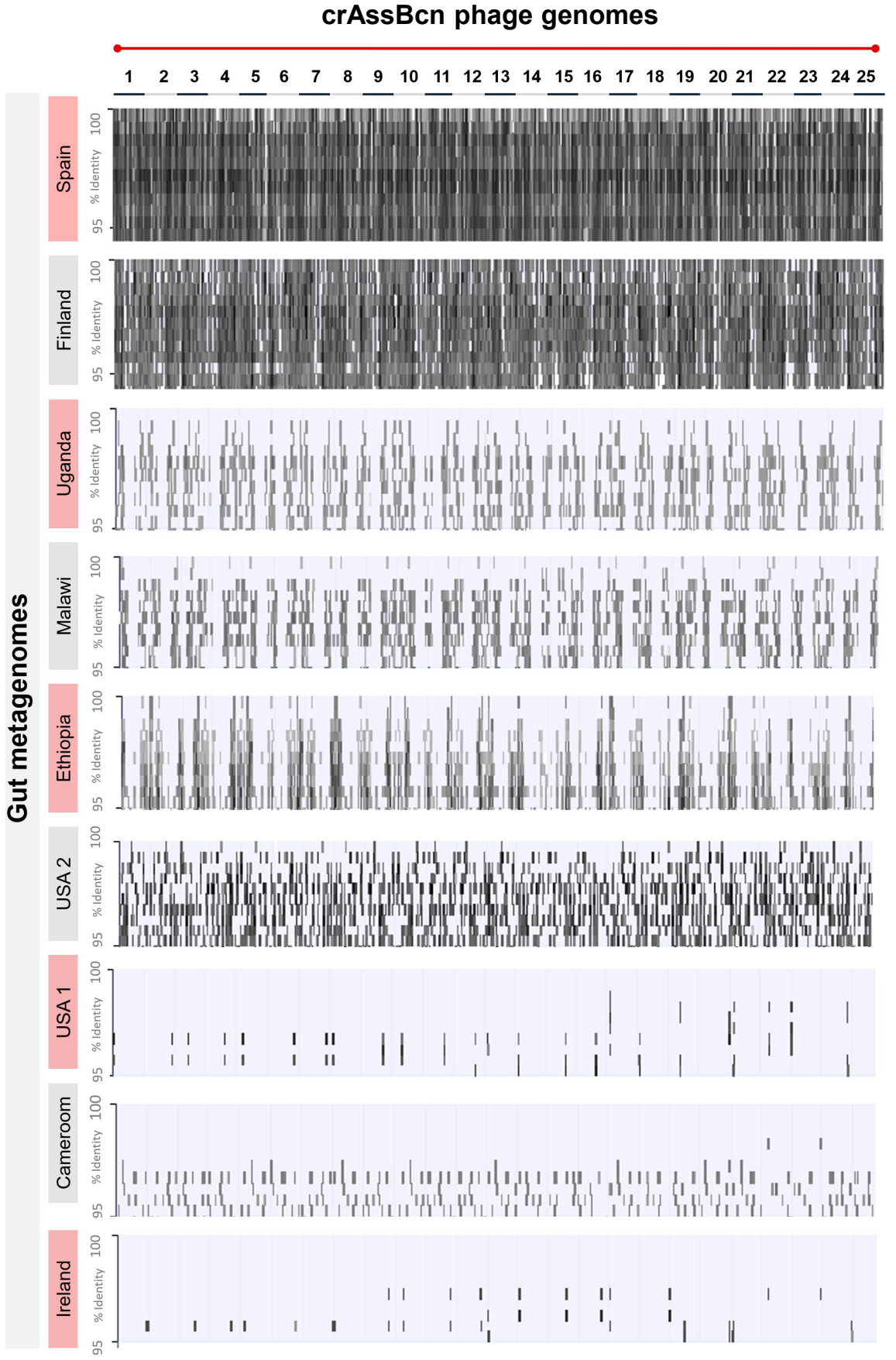


Supplementary Figure 5. Genetic identity of ΦCrAssBcn6 with phages of *Steigviridae* family. CDS alignment of the genome of phage ΦCrAssBcn6 with each one of the 15 crAss-like phage genomes assigned as species of the *Steigviridae* family. Light blue arrows indicate the ORFs in each genome. The color of the bands between the seven genetic maps shows the percentage of identity between each sequence as indicated in the legend at the bottom of the figure.



Supplementary Fig. 6. Comparison of the ORFs performed by reciprocal best match of the crAssBcn phages. The boxplot shows the percentage of amino acid identity (% AAI) of all ORFs comparing the representative crAssBcn phage of each species (marked with an *) with those of the other seven species (I-VI). (Nº ORFs for species I = 106, N for species II = 104, N for species III = 114 and N for species IV, V and VI = 105). The ORFs in the boxplots and upper circles have the highest % AAI and therefore the least variability among the phages. Central line of each boxplot represents the average value of the positive samples, the upper box include samples whose values are within the 75th percentile, and those in the lower white boxes are within the 25th percentile. The ORFs with higher variability correspond to the circles below 50 % (dotted red line). Source data are provided in Supplementary Data 2.

Supplementary Fig. 7. Conserved domains of the polymerase A of crAssBcn phages. Conserved domains of the product encoded in the ORF corresponding to polymerase A of each of the six crAssBcn phages representative of species I-VI. From top to bottom each chart shows first the length of the protein. Small triangles indicate the aminoacids involved in conserved active, catalytic and DNA binding sites of each domain. Colored bars show the closest hits found in the CDD for each domain are presented, these can be specific hits (with a high confident association) or non-specific and the superfamily to which the highest-ranking hit belongs.



Supplementary Fig. 8. Recruitment plots of crAssBcn phages mapping in gut viral communities from different countries. This figure showcases the results of a BLASTn search of metagenomic short reads against crAssBcn phage genomes (X-axis). The upper panel shows the crAssBcn phage genomes. Bars lower than the average represent regions with fewer mapped reads. The panel for the metagenome of each country represents the reads recruited for each phage, where the X-axis shows mapped reads and the Y-axis the percent identity (from 95 to 100 % of identity).

Supplementary Table 1- Evaluation of crAss-like phages in wastewater samples and after three consecutive propagations in *B. intestinalis* enrichment cultures. Values are expressed in Ct and in brackets are the estimated number of GC/ml, that corresponds to the number of CrAss-like phage viral particles.

WWTP	Ct (GC/ml)				CrAss-like phage*	Isolate
	Initial sample	1rst propagation	2nd propagation	3rd propagation	pfu/ml	
Besos 1	28.96	32.42	17.95	n.a	$6.0 \cdot 10^7$	ΦCrAssBcn1
	(10^4)	(10^3)	(10^8)			
Besos 2	31.95	24.96	13.22	n.a	$2.0\ 10^8$	ΦCrAssBcn2
	(10^4)	(10^6)	(10^9)			
Besos 3	30.76	26.74	14.09	n.a	$1.0\ 10^9$	ΦCrAssBcn3
	(10^4)	(10^5)	(10^9)			
Gava 1	29.18	25.16	14.78	13.28	$3.9 \cdot 10^8$	ΦCrAssBcn4
	(10^4)	(10^6)	(10^8)	(10^9)		ΦCrAssBcn5
						ΦCrAssBcn6
						ΦCrAssBcn7
						ΦCrAssBcn8
Gava 2	28.37	30.84	33.81	33.22	n.a	-
	(10^5)	(10^4)	(10^3)	(10^3)		
Igualada 1	30.44	24.98	12.03	13.60	$4.4 \cdot 10^9$	ΦCrAssBcn9
	(10^4)	(10^6)	(10^9)	(10^9)		ΦCrAssBcn10
						ΦCrAssBcn11
						ΦCrAssBcn12
						ΦCrAssBcn13
Igualada 2	30.03	34.16	34.09	34.59	n.a	-
	(10^4)	(10^3)	(10^3)	(10^3)		
Manresa 1	31.24	25.81	10.70	14.21	$1.4 \cdot 10^9$	ΦCrAssBcn14
	(10^4)	(10^5)	(10^{10})	(10^9)		ΦCrAssBcn15
						ΦCrAssBcn16
						ΦCrAssBcn17
						ΦCrAssBcn18
Manresa 2	30.32	33.40	33.45	31.62	n.a	-
	(10^4)	(10^3)	(10^3)	(10^4)		
Prat de	31.18	30.02	13.92	14.28	$2.8 \cdot 10^9$	ΦCrAssBcn19
Llobregat 1	(10^4)	(10^4)	(10^9)	(10^9)		ΦCrAssBcn20
						ΦCrAssBcn21
Prat de	32.98	21.02	13.89	13.54	$8.0 \cdot 10^8$	ΦCrAssBcn22
Llobregat 2	(10^3)	(10^7)	(10^9)	(10^9)		ΦCrAssBcn23
Prat de	32.98	23.31	14.72	n.a	$1.8 \cdot 10^8$	ΦCrAssBcn24
Llobregat 3	(10^3)	(10^6)	(10^8)			ΦCrAssBcn25

*Detected by plaque blot hybridization on plates from the last propagation step.

n.a: not analyzed

WWTP: wastewater treatment plant

Supplementary Table 2.- Bacterial strains used in this study.

Bacteria	Strain	Characteristics	Reference/source
Bacteroides fragilis	RYC2056	Clinical isolate	[1]
Bacteroides fragilis	HSP40	Clinical isolate	[2]
Bacteroides sp.	GB105	Municipal wastewater isolate. UK	Our collection
Bacteroides fragilis	GB124	Municipal wastewater isolate. UK	[3]
Bacteroides fragilis	GB307	Municipal wastewater isolate. UK	[4]
Bacteroides fragilis	GB331	Municipal wastewater isolate. UK	[4]
Bacteroides fragilis	GB135	Municipal wastewater isolate. UK	[4]
Bacteroides fragilis	NCTC 9343	Appendix abscess isolate	[5]
Bacteroides fragilis	3156	Abscess isolate	[6]
Bacteroides fragilis	DSM 2079	Human faeces isolate	[7]
Bacteroides fragilis	C13525	Clinical isolate	[2]
Bacteroides fragilis	ATCC 23745	Pleural fluid isolate	[8]
Bacteroides fragilis	C17898	Clinical isolate	[2]
Bacteroides fragilis	C350	Clinical isolate	Our collection
Bacteroides fragilis	C13182	Clinical isolate	[2]
Bacteroides fragilis	D1416	Clinical isolate	[2]
Bacteroides fragilis	638R C17898	Plasmid-free spontaneous rif^R mutant	[9]
Bacteroides fragilis	SA208	Municipal wastewater isolate. Barcelona	Our collection
Bacteroides fragilis	GA18	Municipal wastewater isolate. Barcelona	Our collection
Bacteroides fragilis	Ø7	Municipal wastewater isolate. Tarragona	Our collection
Bacteroides fragilis	PL122	Municipal wastewater isolate. Tarragona	[10]
Bacteroides fragilis	01-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	02-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	06-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	07-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	08-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	09-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	11-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides fragilis	*ΔmpiM44*	Constitutively expresses PSA	Provided by L.E. Comstock
Bacteroides fragilis	*Δtsr19M5*	Expresses extracellular polysaccharide EPS	Provided by L.E. Comstock
Bacteroides fragilis	*Δtsr15M8*	Constitutively expresses an external protein operon	Provided by L.E. Comstock

Bacteroides fragilis	*Δtsr15M4*	Does not express surface proteins that are constitutively expressed in the mutant *Δtsr15M8* Mutant of strain NCTC 9343	Provided by L.E. Comstock
Bacteroides fragilis	*ΔmpiM44ΔupaZΔupeZ*	Constitutively expresses PSA, PSE and PSC. Mutant of strain NCTC 9343	Provided by L.E. Comstock
Bacteroides fragilis	*ΔPSA*	Overexpresses PSA. Mutant of strain NCTC 9343	[11]
Bacteroides fragilis	*ΔungD1ΔungD2*	Expresses PSH. Mutant of strain NCTC 9343	Provided by L.E. Comstock
Bacteroides fragilis	*ΔPSC*	Overexpresses PSC. Mutant of strain NCTC 9343	Provided by L.E. Comstock
Bacteroides fragilis	*Δtsr19M3*	Does not express EPS. Mutant of strain NCTC 9343	Provided by L.E. Comstock
Bacteroides thetaiotaomicron	GA17	Municipal wastewater isolate. Barcelona	[12]
Bacteroides thetaiotaomicron	ATCC 29741	Perforated appendix isolate	[13]
Bacteroides thetaiotaomicron	HB13	Municipal wastewater isolate. Colombia	[12]
Bacteroides thetaiotaomicron	SA1610	Municipal wastewater isolate. Barcelona	[12]
Bacteroides thetaiotaomicron	04-HSP	Clinical isolate	Provided by F. Navarro
Parabacteroides distasonis	ATCC 8503	Human faeces isolate	[14]
Bacteroides ovatus	03-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides ovatus	05-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides ovatus	10-HSP	Clinical isolate	Provided by F. Navarro
Bacteroides intestinalis	DSM 108646	Human faeces isolate	[15]

Supplementary Table 3.- Genome characteristics of the 25 crAssBcn phages in this study, Species group and GenBank accession numbers

Phage	Length (bp)	Complete genome	% GC	Nº ORFs	Species group	GenBank Accession number	Hyperlink
ΦCrAssBcn1	99312	Yes	35.07	102	1	OQ221536	https://www.ncbi.nlm.nih.gov/nuccore/OQ221536
ΦCrAssBcn2	99311	Yes	35.07	102	1	OQ221537	https://www.ncbi.nlm.nih.gov/nuccore/OQ221537
ΦCrAssBcn3	99312	Yes	35.07	102	1	OQ221538	https://www.ncbi.nlm.nih.gov/nuccore/OQ221538
ΦCrAssBcn4	101446	Yes	35.18	103	1	OQ221539	https://www.ncbi.nlm.nih.gov/nuccore/OQ221539
ΦCrAssBcn5	101573	Yes	35.18	105	1	OQ221540	https://www.ncbi.nlm.nih.gov/nuccore/OQ221540
ΦCrAssBcn6*	101578	Yes	35.18	105	1	OQ221541	https://www.ncbi.nlm.nih.gov/nuccore/OQ221541
ΦCrAssBcn7	101577	Yes	35.18	105	1	OQ221542	https://www.ncbi.nlm.nih.gov/nuccore/OQ221542
ΦCrAssBcn8	101575	Yes	35.17	105	1	OQ221543	https://www.ncbi.nlm.nih.gov/nuccore/OQ221543
ΦCrAssBcn9	99752	Yes	35.08	104	2	OQ221544	https://www.ncbi.nlm.nih.gov/nuccore/OQ221544
ΦCrAssBcn10*	99753	Yes	35.08	104	2	OQ221545	https://www.ncbi.nlm.nih.gov/nuccore/OQ221545
ΦCrAssBcn11	99752	Yes	35.08	104	2	OQ221546	https://www.ncbi.nlm.nih.gov/nuccore/OQ221546
ΦCrAssBcn12	99750	Yes	35.08	103	2	OQ221547	https://www.ncbi.nlm.nih.gov/nuccore/OQ221547
ΦCrAssBcn13	98273	Yes	35.19	101	2	OQ221548	https://www.ncbi.nlm.nih.gov/nuccore/OQ221548
ΦCrAssBcn14*	99601	Yes	34.66	114	3	OQ221549	https://www.ncbi.nlm.nih.gov/nuccore/OQ221549
ΦCrAssBcn15	99600	Yes	34.67	114	3	OQ221550	https://www.ncbi.nlm.nih.gov/nuccore/OQ221550
ΦCrAssBcn16	99601	Yes	34.66	114	3	OQ221551	https://www.ncbi.nlm.nih.gov/nuccore/OQ221551
ΦCrAssBcn17*	100716	Yes	35.01	104	6	OQ221552	https://www.ncbi.nlm.nih.gov/nuccore/OQ221552
ΦCrAssBcn18	99451	Yes	34.66	110	3	OQ221553	https://www.ncbi.nlm.nih.gov/nuccore/OQ221553
ΦCrAssBcn19	97803	Yes	35.18	104	1	OQ221554	https://www.ncbi.nlm.nih.gov/nuccore/OQ221554
ΦCrAssBcn20*	103497	Yes	34.81	105	5	OQ221555	https://www.ncbi.nlm.nih.gov/nuccore/OQ221555
ΦCrAssBcn21	97804	Yes	35.18	104	1	OQ221556	https://www.ncbi.nlm.nih.gov/nuccore/OQ221556
ΦCrAssBcn22*	98637	Yes	35.09	105	4	OQ221557	https://www.ncbi.nlm.nih.gov/nuccore/OQ221557
ΦCrAssBcn23	98396	Yes	34.29	101	4	OQ221558	https://www.ncbi.nlm.nih.gov/nuccore/OQ221558
ΦCrAssBcn24	97685	Yes	35.08	103	1	OQ221559	https://www.ncbi.nlm.nih.gov/nuccore/OQ221559
ΦCrAssBcn25	75754	No	35.18	70	-	OQ221560	https://www.ncbi.nlm.nih.gov/nuccore/OQ221560

(*) Species reference genome.

Supplementary Table 4. Complete non-redundant crAss-like phage genomes of family *Steigviridae* used in this study (https://ictv.global/taxonomy).

Subfamily	Genus	Species	GenBank code	Description	
Asinivirinae	*Akihdevirus*	*Akihdevirus balticus*	KC821624	*Cellulophaga* phage phi14:2	dsDNA
Asinivirinae	*Kahnovirus*	*Kahnovirus copri*	MZ130483	uncultured phage cr44_1	dsDNA
Asinivirinae	*Kahnovirus*	*Kahnovirus oralis*	MT774390	uncultured phage cr85_1	dsDNA
Asinivirinae	*Kehishuvirus*	*Kehishuvirus splanchnicus*	MT774388	uncultured phage cr112_1	dsDNA
Asinivirinae	*Kehishuvirus*	*Kehishuvirus primarius*	MH675552	*Bacteroides* phage CrAss001	dsDNA
Asinivirinae	*Kolpuevirus*	*Kolpuevirus coli*	MZ130478	uncultured phage cr151_1	dsDNA
Asinivirinae	*Kolpuevirus*	*Kolpuevirus hominis*	MT774391	uncultured phage cr126_1	dsDNA
Asinivirinae	*Lahndsivirus*	*Lahndsivirus rarus*	MT774387	uncultured phage cr111_1	dsDNA
Asinivirinae	*Lebriduvirus*	*Lebriduvirus gastrointestinalis*	MZ130485	uncultured phage cr18_1	dsDNA
Asinivirinae	*Mahlunavirus*	*Mahlunavirus rarus*	MT774392	uncultured phage cr128_1	dsDNA
Asinivirinae	*Mahstovirus*	*Mahstovirus faecalis*	MT774378	uncultured phage cr106_1	dsDNA
Asinivirinae	*Pamirivirus*	*Pamirivirus faecium*	MT774389	uncultured phage cr116_1	dsDNA
Asinivirinae	*Paundivirus*	*Paundivirus hollandii*	KJ003983	IAS virus	dsDNA
Asinivirinae	*Pipoluvirus*	*Pipoluvirus rarus*	MT774385	uncultured phage cr108_1	dsDNA
Asinivirinae	*Wulfhauvirus*	*Wulfhauvirus bangladeshii*	MT074136	*Bacteroides* phage DAC15	dsDNA

Supplementary Table 5.- Proteins absent in some of the crAssBcn genomes. The presence (white) or absence (grey) of the ORFs in the 24 crAssBcn phages genomes are presented in comparison with the ORFs in the ΦCrAss001. The protein encoded by each ORF has been indicated according with the most recent annotation (January 2023) of the ΦCrAss001 genome.

CrAss001_ORF Nº protein	% ORF present	%ORF absent	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
CrAss001_18_endonuclease	0	100																								
CrAss001_62_hypothetical protein	8	92																								
CrAss001_26_tail fiber protein	8	92																								
CrAss001_2_hypothetical protein	12	88																								
CrAss001_41_endonuclease	12	88																								
CrAss001_92_hypothetical protein	24	76																								
CrAss001_37_endonuclease	28	72																								
CrAss001_99_HNH endonuclease	28	72																								
CrAss001_8_putative HNH endonuclease	28	72																								
CrAss001_102_CcmD family protein	36	64																								
CrAss001_104_transcriptional regulator	36	64																								
CrAss001_98_hypothetical protein	36	64																								
CrAss001_93_hypothetical protein	40	60																								
CrAss001_23_tail fiber protein	52	48																								
CrAss001_24_putative C-type lectin	60	40																								
CrAss001_25_tail fiber protein	64	36																								
CrAss001_58_zinc-ribbon protein	68	32																								
CrAss001_91_SPFH-domain	72	28																								
CrAss001_61_polynucleotide kinase	80	20																								
CrAss001_90_hypothetical protein	80	20																								
CrAss001_96_protein of unknown function	76	24																								
CrAss001_83_ABC-type cobalt transporter	84	16																								
CrAss001_56_His-Xaa-Ser system radical SAM maturase HxsC-like protein	84	16																								
CrAss001_48_shock protein C	84	16																								
CrAss001_101_hypothetical protein	92	8																								
CrAss001_103_hypothetical protein	92	8																								

Supplementary references

1. Puig, A., Queralt, N., Jofre, J., and Araujo, R. (1999). Diversity of Bacteroides fragilis strains in their capacity to recover phages from human and animal wastes and from fecally polluted wastewater. Appl Environ Microbiol *65*, 1772–1776.

2. Tartera, C., and Jofre, J. (1987). Bacteriophages active against Bacteroides fragilis in sewage-polluted waters. Appl Environ Microbiol *53*, 1632–1637.

3. Ebdon, J., Muniesa, M., and Taylor, H. (2007). The application of a recently isolated strain of Bacteroides (GB-124) to identify human sources of faecal pollution in a temperate river catchment. Water Res *41*, 3683–3690.

4. Ogilvie, L.A., Caplin, J., Dedi, C., Diston, D., Cheek, E., Bowler, L., Taylor, H., Ebdon, J., and Jones, B. v (2012). Comparative (meta)genomic analysis and ecological profiling of human gut-specific bacteriophage φB124-14. PLoS One *7*, e35053. 10.1371/journal.pone.0035053.

5. Sperry, J.F., Appleman, M.D., and Wilkins, T.D. (1977). Requirement of heme for growth of Bacteroides fragilis. Appl Environ Microbiol *34*, 386. 10.1128/AEM.34.4.386-390.1977.

6. Rashtchian, A., Dubes, G.R., and Booth, S.J. (1982). Tetracycline-inducible transfer of tetracycline resistance in Bacteroides fragilis in the absence of detectable plasmid DNA. J Bacteriol *150*, 141. 10.1128/JB.150.1.141-147.1982.

7. Wright, D.P., Knight, C.G., Parkar, S.G., Christie, D.L., and Roberton, A.M. (2000). Cloning of a Mucin-Desulfating Sulfatase Gene from Prevotella Strain RS2 and Its Expression Using a Bacteroides Recombinant System. J Bacteriol *182*, 3002. 10.1128/JB.182.11.3002-3007.2000.

8. Bergan, T., and Hovig, B. (1968). A new species, Sphaerophorus intermedius, isolated from empyema. Acta Pathol Microbiol Scand *74*, 421–430. 10.1111/J.1699-0463.1968.TB03496.X.

9. Privitera, F., Dublanchet, A., and Sebald, M. (1979). Transfer of multiple antibiotic resistance between subspecies of Bacteroides fragilis. J Infect Dis *139*, 97–101. 10.1093/INFDIS/139.1.97.

10. Gómez-Doñate, M., Payán, A., Cortés, I., Blanch, A.R., Lucena, F., Jofre, J., and Muniesa, M. (2011). Isolation of bacteriophage host strains of Bacteroides species suitable for tracking sources of animal faecal pollution in water. Environ Microbiol *13*. 10.1111/j.1462-2920.2011.02474.x.

11. Coyne, M.J., Tzianabos, A.O., Mallory, B.C., Carey, V.J., Kasper, D.L., and Comstock, L.E. (2001). Polysaccharide Biosynthesis Locus Required

for Virulence of Bacteroides fragilis. Infect Immun *69*, 4342. 10.1128/IAI.69.7.4342-4350.2001.

12. Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G.T., Blanch, A.R., Lucena, F., Jofre, J., and Muniesa, M. (2005). Method for isolation of Bacteroides bacteriophage host strains suitable for tracking sources of fecal pollution in water. Appl Environ Microbiol *71*, 5659–5662. 10.1128/AEM.71.9.5659-5662.2005.

13. Sutter, V.L., Emmerman, J., Randall, E., Zabransky, R.J., and Birk, R.J. (1985). Establishment of MICs of moxalactam for control and reference anaerobic organisms in agar dilution and microdilution techniques. Antimicrob Agents Chemother *27*, 424. 10.1128/AAC.27.3.424.

14. Eggerth, A.H., and Gagnon, B.H. (1933). The Bacteroides of Human Feces. J Bacteriol *25*, 389. 10.1128/JB.25.4.389-413.1933.

4.2. Capítulo 2:

La regulación de la biosíntesis de polisacáridos por la región de fase variable 7 (PVR7) en *Bacteroides intestinalis* es responsable de la resistencia a los fagos crAssBcn

Regulation of Polysaccharide Biosynthesis by Phase Variable Region 7 (PVR7) in Bacteroides intestinalis Enhances Resistance to CrAssBcn Phages

Clara Gómez-Gómez, María Dolores Ramos-Barbero, Laura Sala-Comorera, Pedro Blanco-Picazo, Lorena Rodríguez-Rubio, Cristina Garcia-Aljaro y Maite Muniesa.

Artículo en proceso de preparación.

Introducción

El intestino humano constituye un ecosistema complejo y dinámico que alberga una densa comunidad microbiana, compuesta por bacterias, arqueas, virus y hongos (Huttenhower et al., 2012). Entre los virus, los fagos desempeñan un papel clave en la estructuración y el equilibrio de las comunidades bacterianas, influyendo directamente en la composición microbiana intestinal (Manrique et al., 2016). Dentro del denominado “viroma saludable”, uno de los grupos de fagos más representativos y prevalentes es crAssphage, identificado inicialmente mediante ensamblaje metagenómico de muestras fecales humanas (Dutilh et al., 2014). El prototipo *p*-crAssphage marcó un hito en el estudio del viroma intestinal y, desde entonces se han aislado algunos fagos de tipo crAss, como ΦcrAss001 (Shkoporov et al., 2018) y los recientemente caracterizados crAssBcn (Ramos-Barbero et al., 2023), que infectan a *Bacteroides intestinalis.*

Estudios previos han demostrado que estos fagos son virulentos pero que no siguen un ciclo lítico típico ya que su activa replicación no causa una reducción de la población bacteriana (Cortés-Martín et al., 2024; Ramos-Barbero et al., 2023; Shkoporov et al., 2021, 2018). Esto sugiere una coexistencia basada en estados de infección como la pseudolisogenia o la resistencia mediada por variación de fase (Porter et al., 2020; Shkoporov et al., 2021). Este último mecanismo es frecuente en *Bacteroides* (Ben-Assa et al., 2020), ya que poseen PVR flanqueadas por repeticiones invertidas (Shkoporov et al., 2021). Estas regiones pueden cambiar su orientación por la acción de una invertasa específica (Coyne et al., 2003), lo que permite, según su orientación, activar o desactivar (ON/OFF) la expresión de ciertos genes,

en particular aquellos que codifican estructuras de superficie como los CPS (Krinos et al., 2001).

La expresión de cada *locus* CPS está controlada por un operón que incluye dos genes claves: *upxY* y *upxZ*. UpxY actúa como antiterminador de la transcripción, activando la expresión del correspondiente CPS y, en algunos casos, también la de otros *loci* en *trans*. Por otro lado, UpxZ inhibe la expresión de otros CPS, mediante la interferencia con la actividad de las proteínas UpxY de los otros CPS (Chatzidaki-Livanis et al., 2010; Saba et al., 2024a). Esta organización genética asegura que, a pesar del cambio de fase, aparentemente estocástico, que da lugar a la expresión de cada CPS, sea sólo un *locus* CPS el que finalmente se exprese activamente en cada célula (Chatzidaki-Livanis et al., 2010; Fletcher et al., 2007; Saba et al., 2024a). Así se establece un sistema de regulación excluyente que otorga gran plasticidad a la superficie bacteriana, facilitando tanto la evasión del sistema inmune como la resistencia frente a fagos (Carasso et al., 2024; Cortés-Martín et al., 2024; Porter et al., 2020, 2017a; Shkoporov et al., 2021).

Objetivo

El propósito principal de este capítulo es investigar la interacción entre los fagos tipo crAss y su hospedador bacteriano, con especial énfasis en el papel de la variación de fase en los genes involucrados en la biosíntesis de CPS. En este contexto, se pretende evaluar cómo estas modificaciones fenotípicas derivadas de dichas variaciones influyen en la susceptibilidad o resistencia de la bacteria frente a los fagos del grupo crAssBcn. Para ello, se seleccionarán y caracterizarán mutantes con perfiles de resistencia o sensibilidad a los fagos crAssBcn estables. El estudio de estos mutantes permitirá esclarecer los mecanismos moleculares implicados en la interacción el fago crAssBcn y *B. intestinalis*.

Resultados y discusión

Se estudió la dinámica de infección de los fagos crAssBcn en *Bacteroides intestinalis*. Todos los fagos replicaron eficazmente observándose dos fases de propagación: una rápida a las 2 h y otra entre las 4 y 6,5 h. A pesar de que los fagos propagaron hasta títulos cercanos a 10^9 ufp/ml, no se observó una disminución significativa en la cantidad de bacterias, que mantuvieron valores similares a los de los cultivos no infectados. Estos resultados, consistentes con estudios previos (Cortés-Martín et al., 2024; Shkoporov et al., 2021, 2018), sugieren que la propagación de los fagos no causa una lisis celular del cultivo en general, sino que un porcentaje significativo de células no son lisadas por los fagos. Entre las diversas posibilidades, descartadas la lisogenia o la pseudolisogenia, se proponen mecanismos de variación de fase como posible explicación para estos resultados.

Seguidamente, se seleccionaron mutantes de *B. intestinalis* con fenotipos estables de resistencia y sensibilidad a los fagos crAssBcn, evitando la reversión al fenotipo contrario asociada a la variación de fase. Tras infectar cultivos con tres fagos distintos y realizar sucesivas rondas de co-cultivos, se obtuvieron dos mutantes resistentes estables (R10 y R14), que no permitieron la propagación de los fagos. Además, se identificó un mutante espontáneo altamente sensible (S). Estos tres clones (R10, R14 y S) fueron seleccionados y se caracterizaron genómicamente junto con la cepa salvaje (WT). Aunque presentaban una identidad genómica del 99,99%, las diferencias genéticas observadas no se asociaron directamente con la sensibilidad a los fagos, sino con la variabilidad natural entre subcultivos. Se exploraron entonces mecanismos extracelulares (alteraciones del receptor) (Hesse et al., 2020) e intracelulares como sistemas CRISPR-Cas (Garneau et al., 2010) y BREX (Goldfarb et al., 2015; Yuan et al., 2024), sin hallarse diferencias relevantes entre cepas sensibles y resistentes. Sin embargo, se identificó una inversión específica en una PVR, asociada a la síntesis de CPS del operón CPS7, que presentaba una orientación distinta en las cepas resistentes en comparación con la WT y S, sensibles a la infección, lo que sugiere su implicación en el fenotipo resistente frente a los fagos crAssBcn (Carasso et al., 2024; Porter et al., 2020; Shkoporov et al., 2021). En los mutantes resistentes (R10 y R14), PVR7 se encontraba en orientación inversa (estado ON), activando la expresión del operón CPS7 aguas arriba. En cambio, en la cepa S y en la WT, la región estaba en orientación directa (estado OFF), no produciendo la expresión del operón. Estos resultados se confirmaron mediante PCR, secuenciación y análisis con PhaseFinder, vinculando la orientación de PVR7 con el fenotipo de resistencia o sensibilidad.

Una vez descubierto este aspecto, se investigó el papel de la orientación de la región PVR7 en la sensibilidad a la infección de los fagos mediante su clonación en vectores con las orientaciones ON y OFF. La introducción del fragmento PVR7-ON en la cepa S mediante conjugación triparental resultó en la adquisición de resistencia frente a los fagos crAssBcn. Por otro lado, la clonación del plásmido pG106 con el fragmento PVR7 en orientación OFF en las cepas R10 y R14, continúa en proceso, ya que hasta el momento no se han obtenido clones. Por ello, aún no ha sido posible confirmar si la introducción de dicho fragmento puede revertir el fenotipo resistente de las cepas R y conferirles susceptibilidad a la infección por los fagos crAssBcn, quedando este análisis pendiente de estuidos futuros. A pesar de que el resultado es preliminar y se pretende obtener nuevos clones con los fragmentos PVR7-ON y OFF, que deberán verificarse mediante secuenciación, los resultados obtenidos hasta el momento indican que la orientación ON del promotor en la región PVR7, actuando en *trans* sobre el antiterminador UpxY cromosómico, constituiría el principal determinante del cambio en el fenotipo de infectividad de los fagos crAssBcn.

No se observaron interferencias de otros operones CPS, lo que refuerza que la regulación de la sensibilidad a crAssBcn depende específicamente del estado ON/OFF de la región PVR7 que controla la expresión del operón CPS7.

Consistente con los resultados anteriores, el análisis proteómico y de expresión génica reveló una mayor expresión de varios genes del operón CPS7 en las cepas resistentes R10 y R14 en comparación con la WT y la S. Estos genes estaban asociados a funciones en la síntesis de PS y su sobreexpresión se vinculó con la orientación ON de la región PVR7 en las cepas resistentes. Empleando la RT-qPCR se evidenció esta modesta pero consistente sobreexpresión (*log_2-fold change* de 0,75-0,94).

Se demostró que la resistencia conferida por esta sobreexpresión del operón CPS7 no es específica frente a fagos crAssBcn ya que las cepas R10 y R14 y la cepa S conteniendo pG106:PVR7-ON también mostraron resistencia a 5 fagos aislados de aguas residuales no relacionados con los crAssBcn, ΦCrAss001 ni *p*-crAssphage. Esto sugiere que la resistencia no se debe a una modificación específica del receptor de los fagos crAssBcn, sino que más bien se trata de un mecanismo protector externo (sobreexpresión de CPS), formando una barrera que impide al fago alcanzar su receptor, sin necesidad de modificar o reducir su expresión (Shkoporov et al., 2021).

El receptor del fago crAssBcn aún no ha sido identificado, pero sabemos que se encuentra en la fracción de membrana de todas las cepas. Esto se confirmó mediante ensayos de inhibición de la infección con fracciones de membrana, en los que se observó una reducción en el número de fagos capaces de infectar las cepas al poner la suspensión de fagos en contacto con fracciones de membrana externa. Además, se confirmó que todas las cepas, resistentes o sensibles, expresan el receptor a niveles similares ya que la capacidad de inactivar fagos fue comparable entre las fracciones de membrana de todas las cepas, descartando que la resistencia se deba a una menor expresión del receptor en las cepas resistentes. Así, la regulación de CPS7 por inversión de fase es clave para la resistencia, independientemente del tipo de fago, y no involucra alteraciones del receptor.

Conclusiones

En *Bacteroides*, los CPS pueden influir en la interacción con fagos, otros microorganismos y el sistema inmune del hospedador. La expresión exclusiva de CPS7 podría también inhibir otros operones CPS, favoreciendo la estabilidad del fenotipo resistente. Aunque no se observaron otras inversiones en PVRs, no se descarta que otros sistemas puedan influir en la resistencia. Así, este estudio refuerza la idea de que la variación de fase de cápsula PVR7 es clave para la estrategia de defensa de *B. intestinalis* frente a la infección de fagos.

Regulation of Polysaccharide Biosynthesis by Phase Variable Region 7 (PVR7) in *Bacteroides intestinalis* Enhances Resistance to CrAssBcn Phages

Clara Gómez-Gómez[1], María Dolores Ramos-Barbero[1], Laura Sala-Comorera, Pedro Blanco-Picazo, Lorena Rodríguez-Rubio, Cristina Garcia-Aljaro, and Maite Muniesa*

Departament de Genètica, Microbiologia i Estadística. Universitat de Barcelona. Diagonal 643. Prevosti Building. Floor 0. E-08028 Barcelona. Spain.

[1] Both authors contributed equally to this study
***corresponding author.** Maite Muniesa. Phone: +34934039386. Fax: +34934039047
e-mail: mmuniesa@ub.edu

Running head: Resistance to crAssBcn phages relies on PVR7

ABSTRACT

Bacteroides intestinalis regulate surface structures like capsular polysaccharides (CPS) by phase-variable regions (PVRs). Up and down regulation of CPSs influences the interactions with bacteriophages. We investigated the factors conferring resistance to recently isolated crAssBcn phages. Through genomic and phenotypic analyses, we identified spontaneous phage-resistant mutants with the PVR region regulating CPS7 (PVR7) locked in the *ON* orientation, leading to upregulation of CPS7. In contrast, the WT and a phage-sensitive derivative exhibited PVR7 predominantly in the *OFF* state. By cloning PVR7 in *ON* orientation in sensitive strains, these become resistant to crAssBcn phages. The resistance conferred was a non-specific shielding mechanism since CPS7-mediated resistance extended to non-crAssBcn phages, suggesting a broad protective role rather than a specific modulation of the phage receptors. Phase variation is a dynamic mechanism that contributes to phage resistance and to the stability of gut microbial communities. This study enhances our understanding of crAss-like phage / host interactions and the evolutionary strategies shaping bacterial and phage survival in the human gut.

INTRODUCTION

The human gut is a complex ecosystem that contains a dense microbial community composed of bacteria, archaea, viruses, fungi and other microbial eukaryotes [1]. Ninety percent of the gut microbiota is represented by members of the Firmicutes and Bacteroidetes, along with phages targeting these groups in a ratio of approximately 1:1 [2]. Since phages can shape the structure of bacterial communities, and certain bacterial compositions are associated with homeostatic microbiota, studies have attempted to determine whether specific phage communities correlate with human health. In this context, groups of related phages that may contribute to maintaining a healthy gut microbiome have been identified [3] This 'healthy gut phageome' notably includes crAssphage.

The crAssphage (cross-assembly phage) group is one of the most abundant and widespread groups of human-associated viruses in the human gut. The prototypical p-crAssphage was the first crAss-like phage recovered by assembling the first complete crAssphage genome using contigs from several human fecal viral metagenomes [4].

So far, few crAss-like phages have been isolated. ΦCrAss001, described in 2018 [5] and

infecting *Bacteroides intestinalis,* was the first one. Later, other crAss-like phages (ΦCrAss002, DAC15, DAC17) [6–8] infecting different *Bacteroides* species were identified. More recently, we isolated 25 new crAss-like phages (crAssBcn phages), also infecting *B. intestinalis* [9]. CrAssBcn phages belong to the genus *Kehishuvirus,* like ΦCrAss001, but are classified into six different species [9].

Studies on the replicative cycle of isolated ΦCrAss001, ΦCrAss002 and crAssBcn phages indicate that crAss-like phages do not follow a typical lytic cycle. Phages propagation is not detrimental to the growth of *Bacteroides* strains [5,9,10] , which do not show a significant reduction in cell number even when the phages are actively replicating and reaching titers of up to 10^9 pfu/ml. This suggests that the *Bacteroides* population is maintained through a dynamic phage-host interplay involving at least a fraction of the bacterial cells [9,11] .

It has been confirmed that crAss-like phages are not temperate, and since lysogeny is ruled out, the persistence of the host bacteria during phage propagation may be due to a carrier-state infection (where the virions remain within the cell), pseudolysogeny, or phase variation-based resistance involving receptor switches [11,12.]. The latter is particularly promising considering the abundance of phase variation systems in *Bacteroides* [13]. Phase-variable systems that promote phage resistance may exhibit a dynamic of spontaneous reversion, converting some crAssphage-resistant cells to a phage-susceptible state. Through this mechanism, the susceptible fraction of the bacterial population would enable the propagation of phage progeny, while the resistant fraction continues to grow in the presence of the phage, thereby maintaining the overall bacterial population [9,11].

Bacteroides is one of the most abundant inhabitants of the human gut [14]. These Gram-negative anaerobic bacteria have the ability to invert an extensive number of DNA regions in their genomes [13] , similar to what has been described in other bacterial species, such as *Salmonella* [15]. These DNA inversions modulate *Bacteroides* protein expression in an *ON-OFF* manner through the reversible inversion of DNA segments containing the promoters for the ORFs located adjacent to the inversion site. These phase variable regions (PVR) include the promoter regions of at least eight distinct capsular polysaccharides (CPS) biosynthesis loci [16], the promoter for an extracellular polysaccharide (EPS) in *B. fragilis* [17], promoters for fimbriae regions [18] and promoters for other outer surface proteins [19,20] . Site-specific recombinases, from either the serine or tyrosine families, mediate these distinct DNA inversions [21].

The different CPS can surround the *Bacteroides* cell [22] , and the different capsular polysaccharide biosynthesis loci found in the genome of *Bacteroides* show a certain degree of synteny in their genomic organization. Particularly in *B. intestinalis*, all CPS loci except CPS-C include a phase-variable region of approximately 224 bp [11] flanked by two inverted repeats of 18 bp responsible for the inversion. The *ON-OFF* switch of the PVRs of the different CPS loci is mediated by the activity of a single DNA invertase located outside the CPS loci [21]. The ORF located directly upstream the PVR usually encodes for a protein designated UpxY, where “x” is replaced by the letter corresponding to the specific CPS (from UpaY to UphY). UpxY belongs to the family of NusG-like proteins and functions as a transcription antiterminator necessary for the expression of its respective CPS biosynthesis operon. Additionally, UpxY may act *in trans* to activate the expression of the other loci [23,24]. UpxY is followed by UpxZ, a protein that inhibits the synthesis of heterologous polysaccharides from other loci by interfering with UpxY proteins. In the different CPS loci of *Bacteroides*, UpxY and UpxZ are typically followed by 11 to 22 genes involved in CPS biosynthesis [11,16]. When the PVR is in the *ON* orientation, UpxY is transcribed, activating the expression of the genes in the corresponding CPS locus and simultaneously expressing UpxZ. UpxZ then blocks the synthesis of polysaccharides from the other CPS loci, regardless of their promoter orientations. This mechanism implies that, although promoter inversion occurs stochastically, *Bacteroides* prioritizes the expression of only one CPS locus at a time. The concomitant transcription of the other polysaccharide loci within a single cell is thereby inhibited, establishing a regulation network that determines which CPS is expressed [23,24].

Phase-variation of CPS is suggested as an important factor in the phage susceptibility phenotype in *Bacteroides* [11,12,25], although CPS expression is probably not the only mechanism involved [10] . The role of CPS in phage infection may be non-specific, be masking or exposing the phage receptor, or specific, by regulating receptor expression.

In this study, we used the recently isolated crAssBcn phages [9] to generate phage-resistant (R) and phage-sensitive (S) *B. intestinalis* spontaneous mutants, each displaying a stable phenotype towards phages, without the dynamic *ON-OFF* switching observed in the wild-type (WT) strain. Genome comparison of the mutants with the WT revealed that the PVR in the CPS region 7 (CPS7) was the only one affected and that R mutants showed the PVR7 permanently in the *ON* orientation, whereas the S mutant (and, to a lesser extent, the WT) displayed it predominantly in the *OFF* orientation.

By cloning the PVR in either the *OFF* or *ON* oritentation, we evaluated changes in phage resistance or sensitivity to determine whether CPS7 expression plays a major role in preventing crAssBcn phages infection, likely through a non-specific mechanism involving the physical masking of the phage receptor.

RESULTS

Infection dynamics of crAssBcn phages

CrAssBcn phages 6, 10 and 14, belonging to three different species of the genus *Kehishuvirus*[9] were used to infect *B. intestinalis*, and both bacteria and phages were enumerated at different time points over 24 h. All phages successfully propagated during the cultures reaching maximum titers of 6.2 x 10^8 pfu/ml, showing an increase of 3.0-3.6 log units after 24 h (Fig. 1A). The infectious process began with a rapid increase at 2 h, showing a ten-fold rise followed by a slower growth phase. After 4 h, a second burst of 1-2 log units occurred and lasted until 6.5 h, with a modest increase observed up to 9 h. After 9 h, the average titer of the three crAssBcn phages was 6.0 x 10^8 pfu/ml and remained stable up to 24h. Equivalent results were obtained by qPCR targeting crAssBcn phages with average values of 3.7 x 10^{10} gene copies (GC)/ml after 24 h (Fig. 1B). No significant differences were observed among the three individual crAssBcn phages ($p>0.05$). Although some variation was detected, the overall propagation curve of the 25 crAssBcn phages and of phage ΦCrAss001, which also infect *B. intestinalis*, followed the same trend (Supplementary Data 1) [9].

We confirmed previous results [9] showing that, during phage propagation, there was not significant reduction in the number of *B. intestinalis* cells, as evidenced by the absence of culture clearance measured by OD_{600}. Instead, host cell numbers continued to increase throughout the 24 h of incubation, reaching an OD_{600} of 1.02 and viable cell counts of 4.2 x 10^8 cfu/ml at 24 h. These values were similar to those obtained with the uninfected *B. intestinalis* cultures, which reached an optical density (OD_{600}) of 1.03 and 8.8 x 10^8 cfu/ml at the same time point (Fig. 1C and 1D). Throughout the experiment, the maximum difference in viable cell counts between infected and uninfected cultures was of 0.4 log units (Fig. 1D).

Infectivity experiments were performed at a multiplicity of infection (MOI) of 0.001. Previous studies have shown that differences in MOI do not to significantly impact the titers of phages and bacteria [10].

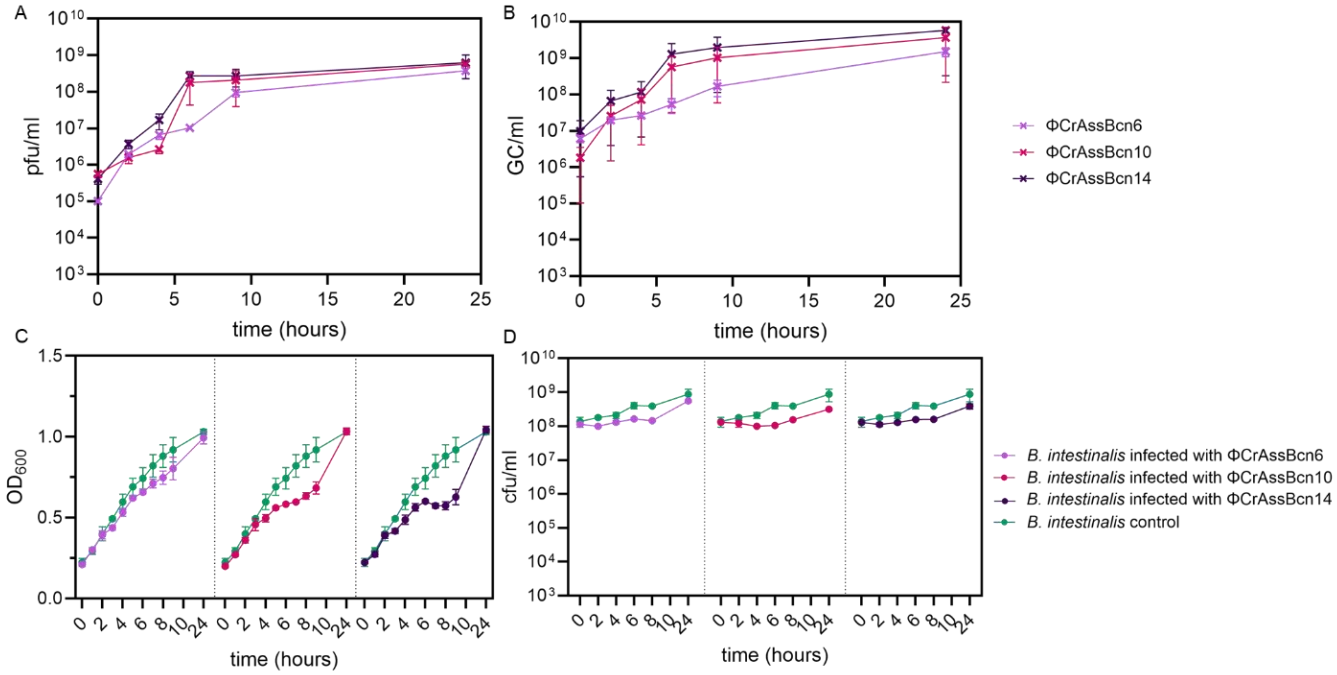


Figure 1. Infectivity dynamics of crAssBcn phages. A) Number of infectious phage particles measured by the double agar layer method (pfu/ml). B) Genome copies of phage particles measured by qPCR (GC/ml). C) Optical density at 600 nm of the host strain cultures of *B. intestinalis*. D) Growth of the host strain cultures *B. intestinalis* measured by colony counts (cfu/ml). Each value corresponds to the average of three independent experiments.

This lack of cell lysis, despite effective phage propagation, has been previously reported [5,10,11] and is an intriguing observation, raising questions about how the dynamics between crAss-like phages and their *Bacteroides* host are established. Phase variation or similar mechanisms have been suggested as relevant factors, although probably not the only ones, that may explain the phage-host dynamics of crAss-like phages with their bacterial host.

Selection of B. intestinalis mutants showing resistance and sensitivity to crAssBcn phages infection

Phase-variable systems are very common in Bacteroidetes and may promote phage resistance by spontaneously switching cells reverting resistant and susceptible states or vice-versa as has been shown for non-crAss-like phages [12,25]. The sensitive fraction of the bacterial population could provide enough susceptible cells to support cell propagation while the resistant fraction maintains the overall bacterial population [12].

Phage exposure has been shown to increase the proportion of cells displaying the phage resistant phenotype in *Bacteroides* bacterial cultures, although this phenotype might revert to a sensitive state in the absence of phage. Our aim was to obtain more stable resistant cells that do not revert, or do so at minimal levels, to the sensitive state. To this end, *B. intestinalis* cultures were infected with the three crAssBcn phages and, after 24h of co-culture, cells were plated and tested for resistance to the respective phages. From this first round of selection in the presence of the phage a 77 ± 8 % of colonies were resistant, as confirmed by spot test, phage enumeration and comparison to the WT strain. Ten spontaneous resistant colonies for each phage were selected and used as hosts in three consecutive co-cultures. After subcultivation, three spontaneous resistant mutants obtained with phages ΦCrAssBcn6, ΦCrAssBcn10 and ΦCrAssBcn14 (named R6, R10 and R14, respectively) were selected, as

they displayed an apparent stable resistant phenotype. To rule out lysogeny or pseudolysogeny in the resistant strains, the absence of crAssBcn phages in the cultures of R6, R10 and R14 was assessed by qPCR. All cultures were negative for the presence of crAssBcn phages.

After infectivity assays, R6 was discarded as it reverted to the sensitive phenotype when incubated in the absence of phage, showing plaques of lysis when infected with crAssBcn phages (Fig. 2A and B). The stability of R10 and R14 that maintained the resistant phenotype even after five successive co-cultures in the absence of phage was confirmed (Fig. 2). No plaques of lysis were observed with R10 and R14 and no propagation of the respective phages occurred (Fig. 2B). Consequently, these two clones were selected for further studies

During this procedure, controls consisted of cultures without phages. Ten colonies from these cultures were processed in the same way as those from the phage-bacteria co-cultures. In one of these controls, we detected a colony showing higher susceptibility to crAssBcn phages than the WT strain (Fig. 2). This sensitive derivative (S) showed a higher efficiency of plating than the WT reaching titers of up to 1.1 x 10^9 in only 6 h (Fig. 2A) and produced clearer lytic zones in spot tests (Fig. 2B). This sensitive (S) spontaneous mutant was used alongside R10, R14 and the WT in the subsequent experiments.

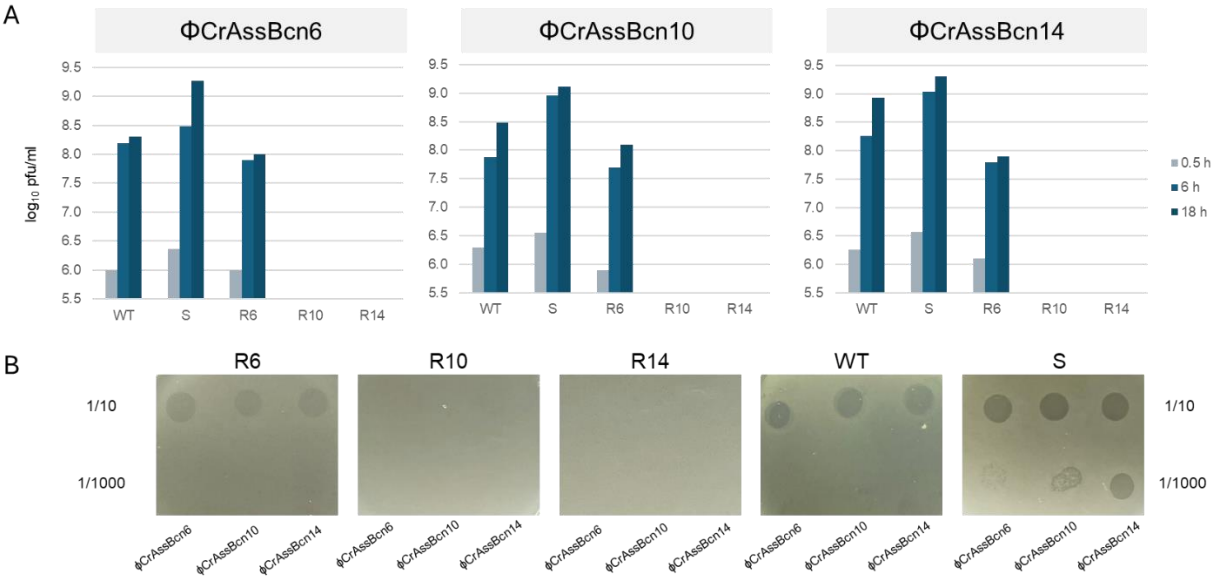


Figure 2- CrAssBcn phages infectivity on WT, R and S derivatives. Infection was conducted with phage suspensions of similar titer; ΦCrAssBcn6 (6.4 x 10^8 pfu/ml) ΦCrAssBcn10 (5.8 x 10^8 pfu/ml) and ΦCrAssBcn14 (4.0 x 10^8 pfu/ml). The hosts were the strains subcultured five successive times in the absence of phages. A) Phage propagation when inoculating ΦCrAssBcn6, ΦCrAssBcn10 and ΦCrAssBcn14 phages using WT, S, R6, R10 and R14 strains as hosts. B) Efficiency of plating of phages ΦCrAssBcn6, ΦCrAssBcn10 and ΦCrAssBcn14 onto the different strains.

Characterization of resistant and sensitive B. intestinalis mutants

To elucidate differences that might explain the phenotypic variation between the WT, S and R10 and R14 mutants, their genomes were sequenced (GenBank accession numbers will be available upon publication) and compared (annotations are provided in supplementary Data 2). Genome comparison (Fig. 3) showed an average nucleotide identity (ANI) of 99.99 % across shared genes of the four strains and only 55 ORFs were not identical among the four

strains. To further evaluate these differences, we performed reciprocal best match analysis of the four genomes at the nucleotide level. This analysis at nucleotide level revealed that 13 ORFs were not identical but still showed high ANIs (between 94-99 %), while the remaining differences were below the detection limit BdL (Supplementary Data 3).

We then compared the genome of the WT strain APC919/174 obtained in this study with the genome of the same strain deposited in the NCBI database (NCBI CP041379.1). The differences in gene content or sequence between R or S, and our WT strain were not greater than the differences between our sequenced WT strain and the NCBI reference genome of *B. intestinalis* APC919/174 (CP041379.1). We concluded that the variations observed were not related to phage susceptibility, but rather reflected the natural variability occurring in strains of the same species following subcultivation.

As genomic comparison could not explain differences between the strains, WT and R and S strains were analyzed for the presence of two known anti-phage defense systems, CRISPR-Cas and BREX, to assess whether they might be responsible of the resistant phenotype. Evidence of the presence of CRISPR spacers in either the WT, S or R10 and R14 strains was minimal (supplementary Data 4), and therefore it was considered that no CRISPR spacers matching ΦCrAssBcn10 and ΦCrAssBcn14 were present in the R strains. Moreover, no complete BREX cassette was present, and only an incomplete DNA methylation enzyme was observed across all strains. These results discarded both mechanisms as the cause of the resistant phenotype.

Beyond the genomic differences, phase variable systems have shown to influence phage sensitivity[11,12,25]. We therefore examined all five genomes for regions showing inversions with a special focus on those that displayed opposite orientations, particularly those affecting CPS operons. A total of 2,172 inverted regions were detected in the WT *B. intestinalis* genome, but only one region showed a clear distinct orientation in the R strains compared to the WT and S strains. This region corresponded to the phase variable region number 7 (PVR7, which affects polysaccharide capsular type 7 (CPS7) (Fig. 3).

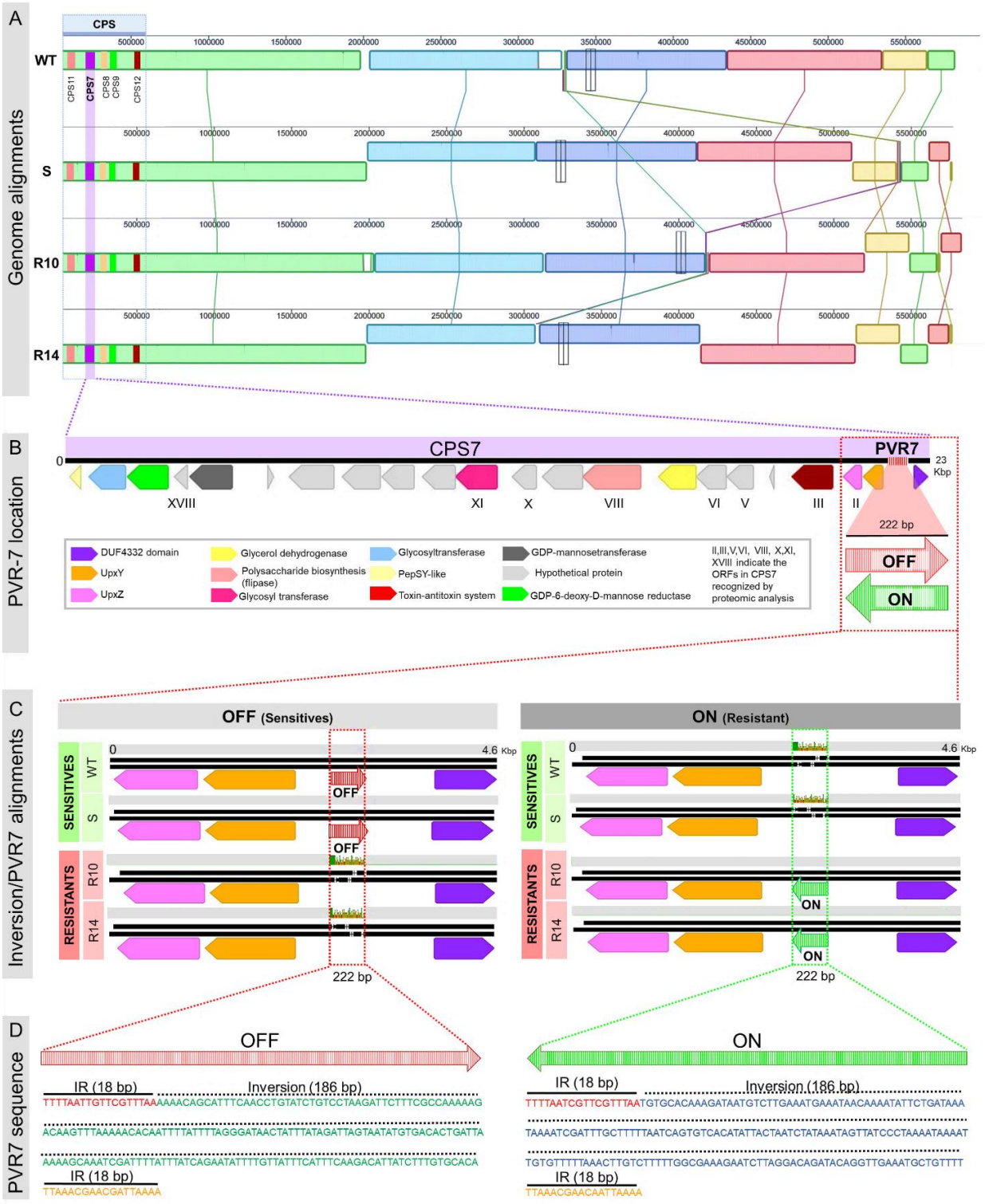


Figure 3.- Location of the CPS and the PVR7 in the resistant and sensitive strains of *B. intestinalis*. A) Multialignment of the genomes of the four strains. Mauve comparison of the four genomes shows minimal differences, particularly in the CPS operons. Location of the CPS is shown in colors in the left side. B) Detail of the CPS7 operon with the ORFs identified indicated in colors and those corresponding to hypothetical proteins labeled in grey. At the right side of the B pannel the PVR7 region is indicated. C) Detail of the PCR7 region showing the inversion and how it regulates the CPS expression in the WT and S strains compared to R10 and R14 strains. Resistant strains show the PVR7 in the opposite orientation than WT and S mutants. D) sequence of PVR7 (186 bp). *ON* orientation is the inverse and complementary sequence of the *OFF* orientation found by PhaseFinder (supplementary Data 5). The PVR7 is flanked by two inverted repeats (IR in yellow and red) of 18 b.

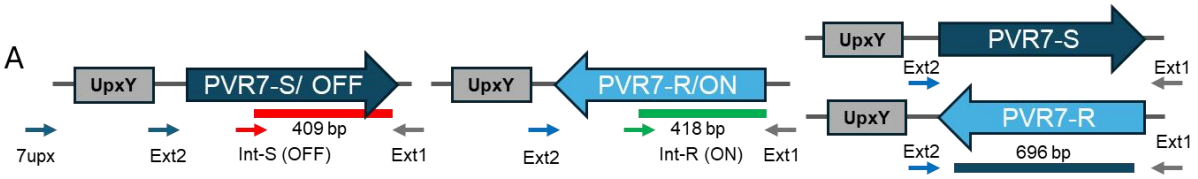
Identification of the phase variable region 7 involved in the resistance to crAssBcn phages

Twenty-one ORFs can be identified in CPS region 7, which displays the typical organization found in other CPS regions in *Bacteroides.* Briefly, the antiterminator UpxY and the transcriptional regulator UpxZ are located upstream the PVR7 and are followed by genes involved in capsular polysaccharide synthesis (Fig. 3).

The PVR7 orientation in the R and S mutants showed that in the R mutants the PVR7 is in the reverse orientation, promoting transcription of UpxY, which we identified as the *ON* phenotype. In contrast, in the S clone and the WT, PVR7 is orientated in the forward direction. This *OFF* state led to the lack of expression of the CPS7 operon. We confirmed the *ON* phenotype in the R mutants and the *OFF* state in WT and S clones by designing PCR primers targeting the 5' and 3' flanking regions of PVR7, along with an internal primer specific for the *ON* and *OFF* orientations (Fig. 4 and Supplementary Table 1). The R clones showed a strong amplification band in the *ON* orientation and only a faint band in *OFF* orientation, whereas the S mutants and the WT showed the opposite pattern.

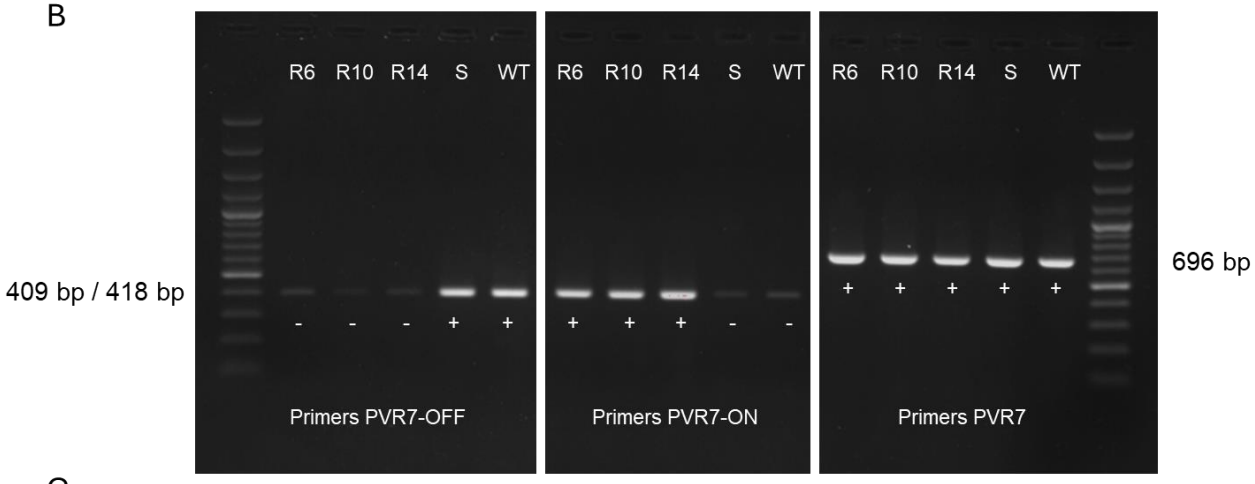
The number of reads corresponding to each PVR7 orientation were determined in the genome of the four strains using PhaseFinder. The data confirmed observations obtained with PCR amplification. R10 had no reads in the *OFF* (forward) orientation, and R14 had very few, with the majority of reads in the *ON* orientation, showing ON:TOTAL ratios of 1 and 0.93 respectively (Fig. 4C and Supplementary Data 5), correlating with the resistant phenotype. In contrast, the S mutant had only *OFF*-orientated reads, and WT had mostly OFF-orientated reads, with ON:TOTAL ratios of 0 and 0.02 respectively (Fig. 4C and Supplementary Data 5), correlating with sensitive phenotypes.

A



Sensitive phenotype | Resistant phenotype | Phase variable región 7

B



C

Strain	Span Forward (OFF)	Span Reverse (ON)	Span Ratio (ON/TOTAL)	Phenotype
WT	53	1	0.02	Sensitive
S	45	0	0	Sensitive
R10	0	40	1	Resistent
R14	1	13	0.93	Resistent

Figure 4. A) Schematic representation of the genomic organization of the PVR7. *OFF and ON* orientation and the primer pair combinations used to identify the PVR7 orientation as well as primers for amplification of the PVR7 area regardless its orientation. B) Gel electrophoresis of PCR products obtained with the different primer pairs for defining the PVR7 orientation of WT, S and R strains. C) Relative abundance of reads showing the *ON* and *OFF* orientations in the genomic analysis of the sensitive (WT and S) and resistant (R10 and R14) strains and the ratios of reads in *ON* orientation vs total number of reads being 1 all reads in the *ON* orientation.

Switch of phage sensitivity phenotypes in B. intestinalis by cloning the PVR7 oriented in ON

We wanted to explore the role of the PVR7 orientation in the response to phage infection. However, we were not certain whether the promoter PVR7 alone would be sufficient to lead to the CPS expression by acting *in trans* from a cloning vector on the antiterminator UpxY in the bacterial chromosome. To elucidate this, we amplified the whole PVR7 region in either the *ON* or *OFF* orientation with primers Ext1-bam/Ext2-bam (Supplementary Table 1). This fragment was first cloned into pGEM plasmid and subsequently transferred to plasmid pG106, generating constructs pG106::PVR7-ON and pG106::PVR7-OFF. These constructs were transformed into *E. coli* and, from there, conjugally transferred into *B. intestinalis* WT and its S mutant, and R10 and R14 strains. We were able to obtain one stable clone of the sensitive strain with the PVR7 in the ON orientation. The success of the cloning strategy was confirmed using primers M13f/Int-R or M13f/Int-S (Supplementary Table 1).

The S mutant harbouring pG106::PVR7-ON became resistant to infection by all three crAssBcn phages, confirming that the orientation *in ON* of the PVR7 region was the key determinant of phage resistance. Furthermore, the PVR7 promoter alone, rather than the presence of the antiterminator UpxY, was sufficient to induce a switch in phage sensitivity to the resistant phenotype compared to the parenteral strains.

Introduction of pG106::PVR7-OFF into R10 and R14 strains was unsuccessful and the expected effect of the R strains becoming susceptible to crAssBcn phages' infection could not be confirmed and remains pending of future analysis.

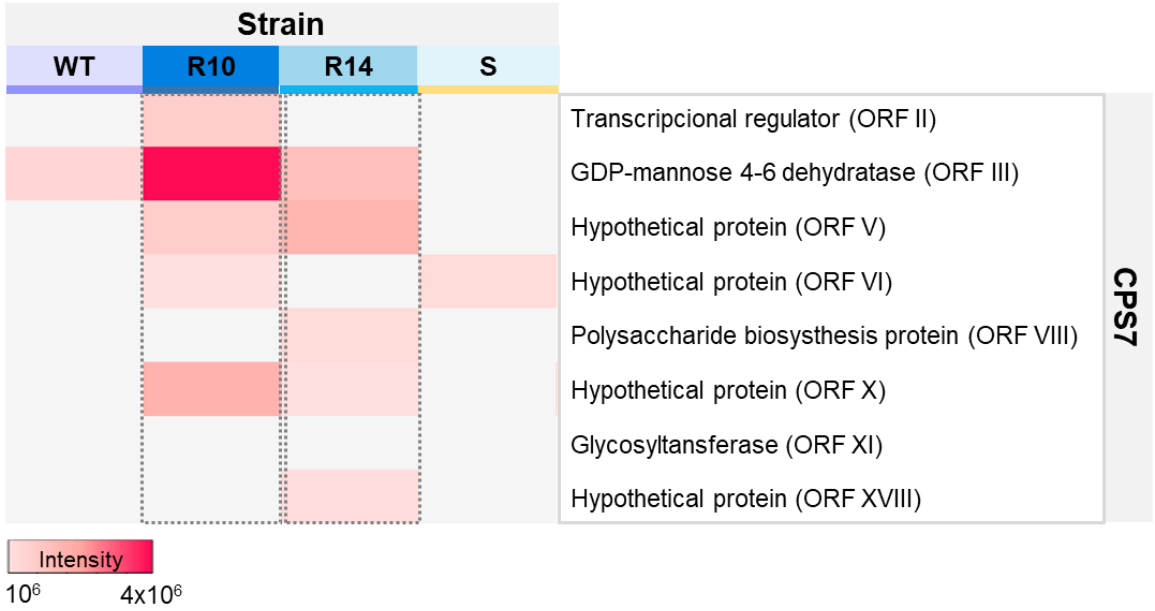
The control using the pG106 vector without insert did not result in any changes in phenotype in any of the strains compared to the original phage sensitivity of the parenteral strains.

These results are still preliminary and more efforts will be devoted to the generation of WT clones harbouring pG106::PVR7-ON as well as the generation of clones of R10 and R14 strains harbouring pG106::PVR7-OFF. Moreover, we will further confirm all the constructs generated, evaluate the cloned fragments by sequencing, and monitor the expression of ORFs in these constructs, as indicated in the following section. Despite the results are preliminary, they indicate that modification of PVR7 orientation alone was sufficient to confer resistance to crAssBcn phages, regardless of the potential role of other molecules or systems. This regulation was mediated by the promoter in PVR7, which we speculate may act *in trans* on the chromosomal UpxY. Although antiterminator UpxY has been described to also affect *in trans* other CPS operons [23], no other CPS systems appeared to interfere with the role of PVR7 in conferring resistance or sensitivity to crAssBcn phages. Despite the trans-acting role of the UpxY antiterminator on other CPS loci, the regulation between differenct loci is likely more influenced by the UpxZ proteins, encoded by the second gene in each of the CPS loci, which inhibit the synthesis of heterologous polysaccharides [23]. Supporting this, in our R mutants compared to the WT and S strains, we detected only differences in the inversion patterns of the PVR7 affecting CPS7 expression, with none of the other CPS loci apparently affected.

Expression of ORFs in CPS7 in R strains versus WT and S strains

Proteomic analysis was performed to identify differential protein expressions in the R strains compared to the WT and S strains. Proteomic results showed a wide range of different protein profiles. Focusing on proteins encoded in the ORFs of CPS7, whereas not all were detected in the proteomic analysis, we detected and identified those corresponding to ORFs II, III, V, VI, VIII, X, XI and XVIII in CPS7 (Fig. 3B). Among these, only UpxZ (ORF II), dehydratase (ORF III) a polysaccharide biosynthesis protein (ORF VIII) a glycosyltransferase (ORF XI) could be annotated, while the rest corresponded to hypothetical proteins. Relative expressions of these proteins indicated that most were upregulated in resistant strains R10 and R14, particularly in comparison to the WT strain and to a lesser extent in comparison to S strain (Fig. 5A).
The higher expression of CPS encoded by PVR7 locus in strains with the PVR7 in the *ON* orientation (R10 and R14 mutants was further evaluated through relative gene expression assays using RT-qPCR in 48 h cultures. We focused on the same ORFs identified in the proteomic analysis. The expression levels of the eight ORFs were quantified using 16SrRNA as the normalization control. All ORFs were upregulated in the R strains with log_2-fold changes ranging from 0.82 to 0.94 for R10 and from 0.75 to 0.86 for R14, compared to the WT strain (Fig. 5B).

A



B

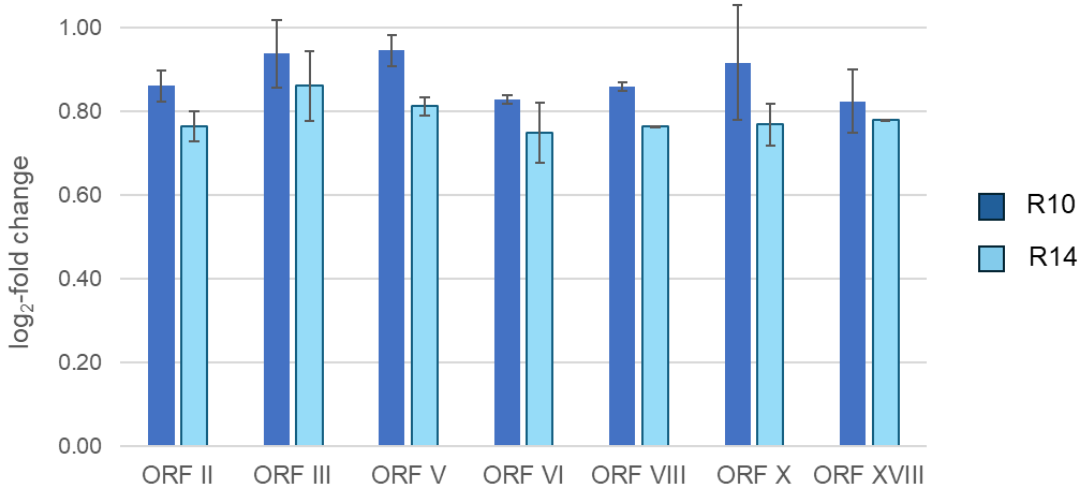


Figure 5.-. Upregulation of the ORFs of CPS7 in the resistant strains (R10 and R14) and the sensitive strains (WT and S). A) Heatmap constructed with the values of the relative abundances of each ORF in CPS7 identified by proteomic analysis. Higher relative abundances (supplementary data 6) are displayed by a stronger intensity. Right box shows the ORFs detected in the proteomic analysis and the corresponding protein encoded if identified. B) Relative gene expression (log_2-fold change) assays of the seven ORFs identified in the proteomic analysis in R10 and R14 gene expression in comparison to the WT strain using RT-qPCR.

Resistance phenotype is not specific of crAssBcn phages

It has been determined that resistance to crAss-like phage can be conferred either by reduced expression of the phage receptor or by a protective effect of capsular polysaccharides covering the phage receptor [11]. If resistance to crAssBcn phages mediated by PVR7 expression involves crAssBcn phage receptor, it should be more specific than simply masking the receptor through CPS overexpression. To explore this, we isolated five non-crAssBcn phages from wastewater using the WT *B. intestinalis* strain as host. Sequencing confirmed that these phages were not crAssBcn, ΦCrAss001, or p-crAssphage (data not shown).

Spot test with the non-crAssBcn phages on WT and S mutant strains showed clear lysis zones whereas the R10 and R14 mutants did not show lysis. Similarly, S strain harbouring pG106:PVR7-ON (S_R) was also resistant to the non-crAssBcn phages, just as they had become resistant to crAssBcn phages and ΦCrAss001. These results suggest that the resistance mechanism involving the CPS7 expression is not specific of crAssBcn phages and likely does not involve the crAssBcn phage receptor.

CrAssBcn phages receptor is not involved in the resistance phenotype

To determine whether differential expression of the phage receptor could also be responsible of the resistant phenotypes, we performed phage inactivation assays using the membrane fraction of the bacterial cells. When the membrane fraction of the WT strain was used, the number of lytic plaques in the phage suspension incubated with the membrane fraction was significantly lower than in the phage suspension control (Fig. 6). This reduction is attributed to the binding of phages to their receptor in the membrane, decreasing the number of free, viable phage particles. In contrast, the control using fragmented membrane fraction causes the largest decrease in plaques number, suggesting non-specific binding to the denaturalized membrane fragments in comparison to the specific receptor biding observed in the native membrane fractions. These results suggest that the receptor of crAssBcn phages is most likely located in the bacterial membrane fraction.

We observed that the inhibition with the membrane fractions of strains with resistant (R10, R14 and S_R (pG106::PVR7-ON)) or sensible (WT and S) phenotypes showed no significant differences ($p > 0.05$) in the number of ΦCrAssBcn10 plaques (Fig. 6). These results suggested that all strains expressed similar levels of the receptor regardless of their susceptibility to phages. This finding allowed us to exclude receptor downregulation as the cause of phage resistance. The same experiments conducted with phage ΦCrAssBcn14 showed similar trends (data not shown).

Although the receptor of crAssBcn phages remains unidentified, these results support the previous observation that phase variable systems and the up regulation of CPS drive resistance to crAssBcn phages in our study, without involving changes in receptor expression.

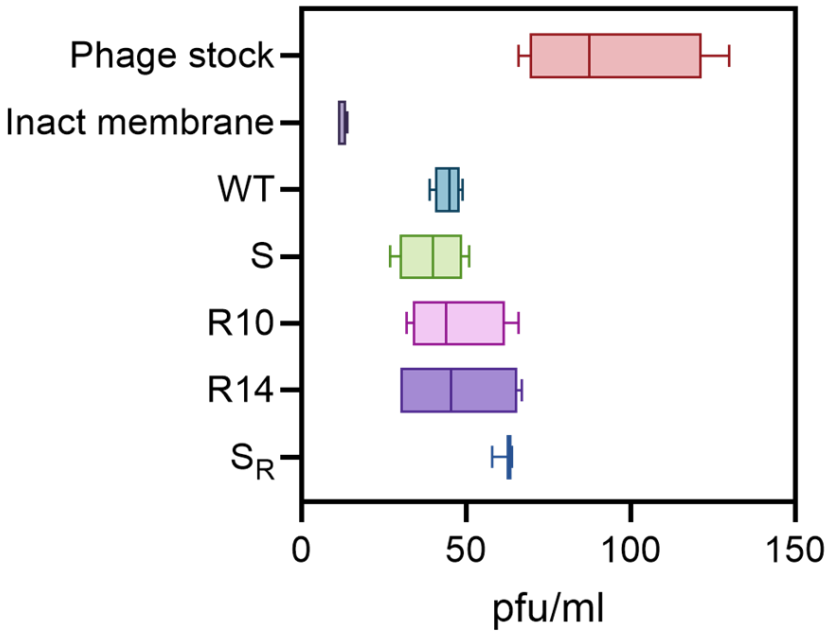


Figure 6.- Representative experiment of inhibition of crAssBcn phages infection. Phage enumeration (pfu/ml) of a suspension of phage ΦCrAssBcn10 against the membrane fraction of *B. intestinalis* strains (WT, S, R10, R14 and S_R (pG106::PVR7-ON). The phage stock not incubated with the membrane fraction was used as control. As non-specific binding control, an inactivated membrane fraction (inact membrane) of WT strain was used.

DISCUSSION

In 2014, a new virus named crAssphage (later designated as prototypical or p-crAssphage) was identified and characterized as highly human-specific [4]. Other members of the Crassvirales group, known as crAss-like phages, were subsequently described "in silico" [26,27] and a few were successfully isolated [5,6,8,9]

The isolation of virions enabled the discovery that crAss-like phages exhibit an unusual replication cycle that does not negatively affect the population of *Bacteroides* host strains. Despite active phage propagation, host cell numbers remain stable. Among several possible explanations, the presence of a heterogeneous population—comprising both phage-sensitive and -resistant cells—appears to be the most plausible [28]. The sensitive subpopulation allows for phage replication, while the resistant subset ensures host population stability.

This phenotypic heterogeneity could be attributed to phase-variable systems, which are common in the *Bacteroides* genus. These systems cause stochastic differences in gene expression between cells by turning operons related to surface structures *ON* or *OFF*, resulting in dynamic surface variation. This mechanism not only supports survival in the gut environment [19] but also leads to differential sensitivity to phages, as demonstrated for other *Bacteroides*-infecting phages [12,25].

Previous studies suggest that the differential expression of capsular polysaccharides is the key mechanism used by *Bacteroides* cells to either enhance or facilitate crAss-like phage infection [11,28]. Transcriptomic analysis has shown that changes in CPS expression correlate with susceptibility to phage infection.

Our recent study led to the isolation of a new set of 25 crAss-like phages, termed crAssBcn phages [9]. All belong to the genus *Kehishuvirus* but are distributed across six distinct species.

Three of these phages, each representing a different species, were selected in this study to investigate the mechanisms underlying differential sensitivity of *B. intestinalis* to crAssBcn phage infection. It was first confirmed that the sensitivity pattern mirrored those previously described for other crAss-like phages[11].

CrAssBcn phages form small, clear plaques and have a burst size of approximately 65 phages/cell [9]. Nevertheless, the host strain achieves densities of up to 10^9 cells/ml after 24 h and remains stable through 48 h of culture. All three crAssBcn phages displayed similar behavior, with no significant differences among them. Co-culture experiments yielded resistant cells, as previously reported [11], although these phenotypes were often unstable and reverted to sensitivity in the absence of phages in the culture. Through serial subculturing, two resistant mutants were isolated that retained phage resistance even after subsequent subcultures without phage exposure. This stable resistance allowed genomic comparison between the resistant mutants and the sensitive strains to identify possible genetic determinants of resistance.

Two main strategies, extracellular and intracellular, can confer phage resistance [29]. Extracellular resistance prevents phage attachment by modifying or masking receptor components, while intracellular mechanisms interfere post-adsorption by blocking genome injection, degrading phage DNA, disrupting assembly, or altering essential metabolites [30–32]. Well-characterized intracellular resistance systems in *Bacteroides* include restriction-modification (R-M), CRISPR-Cas, and BREX. R-M systems consist of a restriction endonuclease that cleaves foreign DNA and a methyltransferase that protects host DNA by methylation [13,33]. The CRISPR-Cas defense mechanism is composed of clustered regularly interspaced short palindromic repeats (CRISPR) that store phage-derived spacer sequences, guiding Cas proteins to cleave matching phage DNA upon reinfection [32.] Recently, the Bacteriophage exclusion (BREX) system, which protects against phages, has also been described. The BREX system comprises six to seven genes including a Lon-like protease, RNA-binding proteins, and methyltransferases [34,35]. More recently, research has uncovered a remarkable diversity of resistance systems [36–38], many of which remain poorly understood.

In our study, various strategies that could explain the resistant phenotype [39] were evaluated, but no significant differences were found between R and S strains that could account for resistance. However, no phage-matching spacers were detected, ruling out CRISPR-Cas activity, and no complete BREX cassette was present. Consequently, we focused on phase-variable regions (PVRs) as the most plausible source of resistance, consistent with previous findings [11,28] .

Particular attention was given to phase-variable regions influencing CPS expression [16] as well as restriction-modification systems, particularly Type I R-M system [13]. We observed only one differential inversion in the resistant mutants compared to sensitive strains, which corresponded to the PVR7, responsible for the expression of CPS7. The CPS7 region has previously been shown to be upregulated in the transcriptional response of the *B. intestinalis*-ΦCrAss001 pair, with mean $\log_2$-fold changes ranging between 1.3-2.3 [10]. In the same study, other genes in different phase-variable CPS biosynthesis clusters (PVR8, 9, 11 and 12) were also up- or down-regulated, consistent with previous observations of the same group [11]. In our study, we did not identify inversion changes in any PVRs besides PVR7. In contrast, our results showed moderate upregulation in CPS7 confirmed by two different techniques. First, proteomic analysis, that identified certain ORFs (though not all) of the CPS7 operon. These ORFs were upregulated in R strains compared to the WT and S strains. We then focused on these ORFs to evaluate gene expression via RT-qPCR, confirming increased expression of CPS7 genes. The observed upregulation was modest compared to prior transcriptomic studies [10]. The different techniques used, and the fact that our experiments were conducted in the absence of phage may explain the varying expression levels.

Although other CPS regions or systems might also contribute to phage resistance, inversion of PVR7, which leads to CPS7 upregulation, appears sufficient to confer resistance to

crAssBcn phages, as demonstrated after cloning the fragment in *ON* orientation. CrAssBcn phages belong to the same genus as ΦCrAss001, so a similar infection mechanism is expected, including the same or similar phage receptor, which remains unidentified in our study.

In *Bacteroides* the expression of diverse CPS is mediated by UpxY acting *in trans*, while inhibition of other CPS operons, leading to the expression of a single operon, is mediated by UpxZ. The absence of *upxY* in the *ON*-orientated fragment cloned into the WT and S strains did not influence in generating the resistant phenotype, suggesting that in these experiments UpxY was not essential for the CPS7 expression, and we hypothesized that the chromosomal UpxY may have been activated by the *ON* orientation in the plasmid, allowing the expression of the operon.

The expression by UpxY and inhibition by UpxZ of the other CPS in the cell occurs regardless of the inversion presented by their respective PVRs [23,24]. Therefore, the lack of inversion in other PVRs observed in our study does not preclude their potential involvement in phage resistance. Overexpression of CPS7 in resistant mutants may have inhibited other CPS operons, resulting in exclusive transcription of CPS7 operon in most cells [24]. The absence of detected inversions in other PVRs may also reflect the R strains' reduced capacity to generate diverse CPS-producing subpopulations, supporting the notion that CPS7 alone interferes with phage infection. However, no nucleotide differences between the R and the S and WT strains were detected in the PVR region or in invertases (data not shown), therefore the precise molecular mechanism behind resistance in the R mutants remains unclear.

The role of PVR7 and CPS7 expression in protecting resistant cells from phage infection confirmed through cloning the PVR7 region, which induced CPS7 upregulation, points towards an extracellular protection mechanism. Capsular biosynthesis is a non-specific phage defense mechanism shown to confer resistance against both crAss-like and non-crAss-like phages[11,12,25]. Resistance involves the formation of capsular polysaccharides that act as a shield, covering receptor structures regardless of their identity, thereby strains insensitive to infection. We confirmed that phages other than crAssBcn, which can infect *B. intestinalis* WT strain, do not infect the R and S_R mutants, supporting the extracellular protection mechanism as not specific to crAssBcn phages.

However, our experiments suggest the crAssBcn phages receptor is located in the membrane fraction and can bind phage particles in its native form, neutralizing a significant number of phages. Nonetheless, no differences in receptor expression between R, S_R and WT or S strains were observed, likely due to similar expression levels across all strains. This suggests receptor expression levels do not influence the infection phenotype further supporting the role of extracellular protection mechanisms. To date, specific bacterial receptors for crAssphages remain largely unidentified. Some studies suggest TonB-dependent receptors may bind to crAss-like phage tail fibers [7]. TonB-dependent systems in Gram-negative bacteria are outer membrane transport proteins that can function as phage receptors [40,41].

CPS in *Bacteroides* serves multiple roles, including protecting the bacterial cell surface. The coverage of the cell by different CPSs influences interactions with bacteriophages and other bacteria and can protect against the immune system or alter host physiology [42]. Variable expressions of different CPSs is also presented as an evolutionary advantage. For example, expression of certain CPSs, such as CPS7 and CPS8, is associated with enhanced biofilm formation, whereas the expression of all CPSs hinders biofilm development [43]. Switching between different capsules is also an adaptative feature that enhances cell fitness in response to various perturbations allowing flexible responses to threats and supporting population survival. Switching between different CPSs may also enable bacteria to better uptake certain nutrients, particularly sugars, which may be accessible through some CPSs but blocked by others.

Regarding their sensitivity to phages, *Bacteroides* liquid cultures of single CPS-expressing strains (e.g. CPS4, CPS7, CPS8) inoculated with phages regrew to intermediate densities, suggesting that a fraction of the population became resistant to phages. By generating heterogeneous subpopulations, bacteria enable the overall population to prevail while gaining specific advantages from interactions with different phages [44], as long as these interactions confer a competitive benefit.

Similar CPSs diversity and heterogeneity have been found in other species of gut Bacteroidetes, supporting the relevance of this plasticity for bacterial fitness. Concerning crAss-like phages, here we confirm the role of CPS7 in preventing CrAssBcn phage infection in a non-specific manner. As the mechanism only temporarily blocks phage infection without completely preventing bacteria-phage interactions, it is reasonable to hypothesize a yet unknown mutual benefit for both entities. This could lead, in future studies, to unveiling the role of crAss-like phages in the human microbiome.

METHODS

Strains, bacteriophages, vectors, media, and culture conditions

B. intestinalis strain APC919/174 (DSM 108646) was the host for the propagation of crAssBcn phages 6, 10 and 14 [9], and bacteriophage ΦCrAss001 (DSM 109066) [5] was used as the positive control. *E. coli* DH5α was used for preliminary cloning of *Bacteroides* genes.

Vector pGEM (Promega) was used for cloning of PCR fragments in *E. coli* DH5α. Vector pG106 (Addgene) a shuttle vector allowing cloning in *E. coli* and *B. intestinalis* [45] was used for definitive cloning in *B. intestinalis*. Vector pRK2073 [46] in *E. coli* HB101 was used for triparental mating.

Anaerobe basal broth (ABB) (Thermofisher) was used for the growth of the *Bacteroides*. For the plaque assays, ABB containing 0.7 % (w/v) of agar-agar for soft agar or 1.4 % (w/v) of agar-agar for the agar plates were used. For differential growth of *Bacteroides*, *Bacteroides* Bile Esculine (BBE) agar [47] was used. Luria broth (LB) and Luria Agar was used for growth and cloning procedures of *E. coli* [48].

When necessary, erythromycin (300µg/ml in agar plates and 50ug/ml in broth) [45], Ampicillin (100 µg/ml) for *E. coli* clones containing pGEM vector, or spectomycin (50µg/ml) for the helper HB101 strain [46] was added to the medium after sterilization.

Manipulation of *Bacteroides* was performed in aerobic conditions while incubation of cultures and agar plates was done at 37 ºC for 24 or 48 h in anaerobic jars (GasPak; BBL) with CO_2 atmosphere generators (Anaerocult A; Merck).

Two raw influent water samples were obtained from an urban wastewater treatment plant (WWTPs) in Catalonia (NE Spain). Samples were collected in sterile containers, transported to the laboratory within two h of collection and processed immediately for the isolation of non-crAss-like phages infecting the *B. intestinalis* strain APC919/174.

Bacteriophage enumeration

One ml of phage suspension was used to infect one ml of *B. intestinalis* culture at the middle-exponential growth phase (OD_{600} of 0.3) in 8 ml of ABB medium and incubated anaerobically and statically at 37 °C for 24 h. For subsequent enrichment cultures, one ml of the phage

suspension from the first culture was filtered through polyethersulfone (PES) low-protein-binding membranes of 0.22 µm of pore and used to infect a new *Bacteroides* culture under the same conditions.

For plaque assays, ten-fold serial dilutions of each phage suspension prepared with SM buffer were enumerated by the double agar layer method [49] with the respective *Bacteroides* host strains in ABB soft agar. Plates were incubated anaerobically at 37 °C for 24 h. Spot assays were performed as described for plaque assays but without the addition of phage to the ABB soft agar. Ten µl drops of 1/10 and 1/1000 dilutions of the phage suspension were directly applied to the solidified lawn of *B. intestinalis* host strain and dried prior to incubation. Negative controls were prepared without the addition of phage.

The dynamics of infection of the crAssBcn phages were evaluated by infecting a middle-exponential growth phase culture of *B. intestinalis* grown with each phage suspension at a multiplicity of infection (MOI) of 0.001. Cultures were inoculated in ABB medium and incubated anaerobically at 37 °C. Bacterial growth was monitored by measuring the OD of the cultures at 600 nm and by colony forming units (cfu/ml) grown in ABB agar. Phage propagation was enumerated by plaques assay of lysis plaques obtained by the double agar layer method (pfu/ml) and by qPCR [9] as described below at different time intervals during 24 h. As a control, a *B. intestinalis* culture without phage was grown under the same conditions.

Phage and bacterial nucleic acid isolation

DNA and RNA isolation from bacteria and phages was performed with QIAamp DNA Blood (Qiagen GmbH), following the manufacturer's instructions. The DNA was suspended in a final volume of 200 µl of sterile bidistilled water. The DNA and RNA concentration of each pooled sample was evaluated using a Qubit® Fluorometer (Life Technologies) and the DNA and RNA quality was further confirmed by the 2100 Bioanalyzer system (Agilent Technologies).

Generation of *B. intestinalis* mutants resistant and sensitive to crAssBcn phages infection

Three exponential growth-phase cultures of *B. intestinalis* (OD_{600}=0.3) were infected with three different CrAssBcn phages (6, 10 and 14) at a MOI= 1 and incubated for 24 h under anaerobic conditions at 37 °C. After incubation, bacterial cells in the cultures were decimal diluted and plated in AB agar plates. Isolated colonies were plated by double agar layer and infected with one drop of the respective phage suspension at 10^7 pfu/ml. A non-infected culture of the host strain was used as control. Those colonies showing positive spot assay, equal to the wild type (WT) strain were discarded and those showing lower or no infection compared to the WT were selected as possible resistant mutants (R). Colonies showing higher infectivity than the WT were considered as sensitive mutants (S). Suspected R mutants were subjected to reinfection with the respective phage as previously described in three consecutive steps and the resulting colonies were checked for phage resistance. Phage enumeration using standard double agar plaque assay and the R and S mutants as hosts was performed phage titer was compared with the WT strain.

RT-qPCR, qPCR and end-point PCR analysis

RNA qPCR amplifications were carried out using the standard run in the StepOne™ Real Time PCR System (Applied Biosystems) was performed in a 20 µl reaction mixture with Applied Biosystems™ Power SYBR™ Green RNA-to-C_T™ 1-Step Kit (Applied Biosystems) for RT-qPCR in one single step to convert RNA to cDNA and qPCR amplification with 2 µl of RNA.

DNA qPCR used TaqMan® Environmental Master Mix 2.0 (Applied Biosystems). The reactions contained 9 µl of the sample DNA or standards with known DNA concentration prepared from gBlocks™ Gene Fragments used for absolute quantification. All results were analyzed with the Applied Biosystems StepOne™ Instrument program. All samples were run in triplicate (including the standards and negative controls). The threshold cycle (Ct) values or the number of gene copies (GC) was defined as the mean of the triplicate data obtained.

End point PCR amplification was performed using DreamTaq Green DNA Polymerase (Fermentas) in a GeneAmp PCR system 2700 (Applied Biosystems). Amplimers were visualized by electrophoresis in agarose gel stained with RedSafe™ Nucleic Acid Staining Solution 20000x (INtRON). For relative quantification of the expression of ORFs upstream the inversion region was compared between the WT and the R and S mutants using the relative $\log_2$-fold expression. First, the RT-PCR Ct values of each ORF and strain were normalized with the Ct value obtained with the 16SrRNA of the same strain (ΔCt = Ct [Target]-Ct [16SrRNA]. $\Delta\Delta$Ct was the result of substracting the normalized value of each R strain to the normalized value of the WT strain, finally $\log_2$-fold change was calculated as $2^{-(\Delta\Delta Ct)}$.

Sequencing

Sanger sequencing was used for confirmation of the constructs pG106::PVR7-ON or OFF with an ABI PRISM Big Dye 3.1 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) in an ABI PRISM 3730 DNA Analyser (Applied Biosystems), following the manufacturer's instructions.

Whole genome sequencing of bacterial genomes was performed with five µl of bacterial DNA at a concentration of 0.2 ng/µl that was fragmented and used to prepare libraries using the Nextera XT Kit (Illumina. Inc.) according to the manufacturer's protocol. The libraries were purified with AmPure beads (Beckman Coulter Inc.), checked for fragment distribution and size, and quantified using a TapeStation 4200 and the Agilent High Sensitivity D1000 ScreenTape system (Agilent Technologies) in a Quantus™ Fluorometer (Promega). Wild type, resistant and sensitive bacterial strains were separately sequenced on the NextSeq System (Illumina) with a high output run of 300 cycles. Raw reads were trimmed by Trimmomatic (PE -phred33 ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36)[50] and the read quality was checked by fastQC[51]. Pair-end trimmed reads were assembled individually by SPAdes v3.13.0 [52] (-k 21,33,55,77,99,127). In order to recover complete or near to complete *Bacteroides intestinalis* genomes from the assembly sequences, genomes were assembled applying the Trusted contigs strategy, available in SPADEs, that used the *Bacteroides intestinalis* strain APC919/174 (CP041379.1) genome as a reference to provide a guidance during *de novo* genome reconstruction. Following, ORF prediction was performed by Prodigal [53] and functional annotation was done by Bakta [54] and Blastp [55]using NR database (update 2024) in parallel.

Evaluation of phage defense mechanisms, genome comparison, gene variation, CRISPR-Cas, and BREX system

B. intestinalis genomes were alignment by Mauve [56] and Genomic maps were generated using Geneious Prime® 2025.0.3. When comparing genomes, the identification of orthology gene between sensitive and resistant *Bacteroides* was performed by reciprocal best match and average sequence identity in nucleotides employing rbm.rb script from the Enveomics package [57] with its default parameters.

CRISPR arrays (Clustered Regularly Interspaced Short Palindromic Repeats) and their associated Cas proteins [58] were evaluated in each *B. intestinalis* genome by

CRISPRCasFinder 4.2.20 [58] The detected spacers were recluted against ΦCrAssBcn14 phage genome using map to reference option by Geneious Prime® 2025.0.3.

The putative BREX system[35] in *B. intestinalis* was investigated using as a reference the BREX_NZ_CP064940.1:3399567_3400007_I1224_RS12645 [GeneID=69507549], previously described in *B. intestinalis* NZ_CP064940 available in public database. Gene alignments were performed by Geneious Alignment tool by Geneious Prime® 2025.0.3.

Identification of phase variable regions in the resistant strains *vs* the sensitive strains

The phase-variable regions in both resistant strains and the sensitive and wild-type *Bacteroides* were analyzed using PhaseFinder (v1.0) [59] to identify invertible regions within raw genome reads, employing default parameters. Potential inversion sites were detected by locating inverted repeats within *Bacteroides* genomes using the 'locate' function. A reference database containing both forward and reverse orientations of these regions was subsequently generated via the 'create' function. Genomic reads were then mapped to this database using the 'ratio' function, which quantified the proportion of reverse-oriented reads relative to the total reads assigned to each region. Each invertible region underwent manual curation to characterize its coding sequences, gene annotations, and predicted functions as previously described [11] . Genomic regions were further examined and visualized using Mauve. Additionally, the identified PVR region was aligned in each genome in order to check the fragment direction in all cases (sensitives and resistants).

Cloning procedures

The PCR amplimer obtained using primers containing *Bam*HI restriction site is ligated in the pGEM vector (Promega) following the manufacturer instructions. The construct was electroporated in electroporation-competent *E. coli* DH5α cells that were prepared from 5 ml of cultures in LB medium and concentrated by centrifugation at 3000 *xg* for 15 min at 4 °C. Cells were three times washed in 4.5 ml of ice-cold double distilled water and suspended in 100 µl of ice-cold double-distilled water. The cells were mixed with the corresponding amount of plasmid DNA, transferred to a 0.2 cm electroporation cuvette (Bio-Rad) and electroporated at 2.5 kV with 25 F and 200 Ω resistance. Immediately after electroporation, the cells were transferred to 1 ml LB medium for 1 h at 37 °C with shaking. Cells were concentrated ten-fold from a 1 ml culture before plating on LB agar selective media, which includes X-Gal and IPTG at 40 µg/ml, and Ampicillin at 100 µg/ml in final concentration. The presence of the vector with the insertion in white ampicillin resistant colonies were confirmed by PCR analysis.

The construct was purified from the positive clones with the Plasmid DNA Extraction Midi Kit (Qiagen) and the insert was removed from pGEM vector by *Bam*HI digestion. The digested DNA was run in an agarose gel by electrophoresis, excised from the gel, purified with a QIAquick gel extraction kit (Qiagen) and ligated into *Bam*HI-digested pG106 vector and transformed into *E. coli* DH5α as described above using LB agar plates containing X-Gal and IPTG and erythromycin (300 µg/ml). The presence of pG106 with the insertion in white erythromycin-resistant colonies was confirmed by PCR analysis and several colonies with the PVR7 fragment in *ON* and *OFF* position were selected for being transferred to *B. intestinalis*.

For cloning in *Bacteroides*, triparental mating with the mobilizing strain HB101/pRK2073 was used to transfer the pG106 plasmid with the different inversions in *ON* and *OFF* from the donor strain *E. coli* DH5 to the receptor *B. intestinalis* strains. 50µl of the recipient *Bacteroides* strain, 10 µl of the helper strain HB101/pRK2073 and 10 µl of the *E. coli* donor strain carrying the plasmid pG106 were mixed and plated as a drop on a sterile ABB agar plate and incubated overnight at 37 °C under anaerobic conditions. After the required conjugation time, one ml of

ABB broth was used to collect the culture drop, and serial dilutions were made, followed by plating on *Bacteroides* selective BBE medium. Transconjugants were selected on BBE plates containing erythromycin and incubated anaerobically at 37 °C. The presence of the vector with the insertion in erythromycin resistant colonies with a black precipitate susceptible of being *B. intestinalis* were confirmed by PCR analysis.

Purification of bacterial membrane fraction

For screening purposes, LPS was obtained after proteinase K digestion of whole *Bacteroides* cells, and the LPS samples were separated by SDS-Tricine-PAGE as previously described and visualized by silver staining [60]. The LPS was extracted from the crude-cell envelope fraction according to the Westphal method, modified by Osborn [61].

Phage inactivation by membrane fractions

Phage suspensions (100 pfu/ml) were incubated with the bacterial membrane fractions of strains WT, S, R10, R14 and S (pG106::PVR7-ON) for 20 min without shaking. After incubation the number of viable phage particles was determined by double agar layer. In case the phage receptor is present in the membrane fraction, phages will bind to it, blocking and inhibiting their infective capacity, what should be confirmed by a decrease in the number of pfu in comparison to the phage stock. In order evaluate the specific receptor binding, an aliquot of the membrane fraction of *B. intestinalis* WT culture was inactivated by sonication (1 min), which denaturalized and fragmented the membrane structures.

Proteomic Analysis

A total proteomic analysis of membrane proteins from the R10, R14, S and WT strains was conducted to evaluate differences in the relative abundances of proteins corresponding to the ORFs located upstream and downstream of the switch. Proteins were extracted following a previously reported protocol [62] with some modifications: 1 L of bacterial culture was pelleted and resuspended in 10 ml of buffer (0.2 M Tris-HCl (pH 8), 1 M sucrose, and 1 mM EDTA). Lysozyme was added (1 mg/ml) and after 5 min incubation at room temperature, samples were placed in ice to allow spheroplast formation. Cells were centrifuged at 200,000 × g for 45 min at 4 °C and the pellet was resuspended in 10 mM Tris-HCl (pH 7.5), 5 mM EDTA, and 0.2 mM DTT, and DNase (1 mg/ml). Then, cells were broken using a French Press and the unbroken cells were pelleted. Supernatant was then centrifuged at 280,000–300,000 × g for 2–4 h at 4 °C. After centrifugation, the pellet, containing the crude membrane fraction, was resuspended in buffer containing 50 mM Tris-HCl (pH 8), 2 % (w/v) Triton X-100, and 10 mM $MgCl_2$ and incubated at 4 °C for 30 min to 1 h with agitation. After centrifugation at 105,000 × g for 1 h at 4 °C, the supernatant contained the membrane proteins.

Approximately 10 mg of each sample were loaded in SDS-PAGE gel to eliminate possible interferents. Then, samples were in-gel digested and resulting peptide mixtures were extracted from the gel matrix with 5 % formic acid solution (5 %FA/H2O) and 100 % ACN, and dried -down in a SpeedVac vacuum system. The dried-down peptide mixtures were analyzed in a nanoAcquity liquid chromatographer (Waters) coupled to a LTQ-Orbitrap Velos (Thermo Scientific) mass spectrometer. Peptides were trapped on a Symmetry C18™ trap column (5µm, 180µm x 20mm; Waters) and were separated using C18 reverse phase capillary column (biozenTM peptide XB-C18 column; 2.6 µm, 75 µm x 250 mm, Phenomenex). The gradient used for the elution of the peptides was 1 to 35 % B in 100 min, followed by a gradient from 35 % to 45 % in 20 min (A: 0.1 % FA; B: 100 % ACN, 0.1 %FA; flow rate: 350 nl/min). Eluted

peptides were subjected to electrospray ionization and peptide masses (m/z 300-1600) were analyzed. Generated. Raw data files were collected with ThermoXcalibur (v.2.2).

The raw data files obtained from the mass spectrometry analyses were used to search against the public Uniprot *Bacteroides* database using MaxQuant [63] in the default mode. This database also included a small database of laboratory contaminants, as well as the 10 proteins upstream and the 10 proteins downstream of the inversion region of the WT strain. Relative abundances (intensities) heatmap was generated using Heatmapper [64].

Statistics and Reproducibility

Statistical analyses were conducted using GraphPad Prism 10 (GraphPad Software). Friedman test with Dunn's multiple comparison was used to assess differences among the crAssBcn phages selected, with statistical significance set at $p<0.05$. D'Agostino-Pearson test was used to check normality. To assess the differences between resistant and sensitive phenotypes in infectivity assays, one-way ANOVA with Tuke's multiple comparison test was used, with statistical significance set at $p<0.05$. The experimental data presented represents the averages of at least three independent experiments. Bioinformatic analyses were performed in duplicate and yielded consistent, reproducible results.

Data availability

The crAssBcn phages genomes used in this study are deposited in GenBank with the GenBank accession codes from OQ221536 to OQ221560 and are publicly available. The *B. intestinalis* genomic data can be found in NCBI (NZ_CP041379.1). All data will be available upon publication.

The authors declare that all other data supporting the findings of this study are available within the paper and its supplementary files.

References

1. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
2. Mirzaei, M. K. & Maurice, C. F. Ménage à trois in the human gut: interactions between host, bacteria and phages. *Nature Reviews Microbiology 2017 15:7* **15**, 397–408 (2017).
3. Manrique, P. *et al.* Healthy human gut phageome. *Proceedings of the National Academy of Sciences* **113**, 10400–10405 (2016).
4. Dutilh, B. E. *et al.* A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* **5**, 4498 (2014).
5. Shkoporov, A. N. *et al.* ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects Bacteroides intestinalis. *Nat Commun* **9**, 4781 (2018).
6. Hryckowian, A. J. *et al.* Bacteroides thetaiotaomicron-infecting bacteriophage isolates inform sequence-based host range predictions. *Cell Host Microbe* **28**, 371–379 (2020).
7. Papudeshi, B. *et al.* Host interactions of novel Crassvirales species belonging to multiple families infecting bacterial host, Bacteroides cellulosilyticus WH2. *Microb Genom* **9**, 001100 (2023).

8. Guerin, E. *et al.* Isolation and characterisation of ΦcrAss002, a crAss-like phage from the human gut that infects Bacteroides xylanisolvens. *Microbiome* **9**, 89 (2021).

9. Ramos-Barbero, M. D. *et al.* Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution. *Nat Commun* **14**, 4295 (2023).

10. Cortés-Martín, A. *et al.* Adaptations in gut Bacteroidales facilitate stable co-existence with their lytic bacteriophages. *bioRxiv* 2024.11.17.624012 (2024) doi:10.1101/2024.11.17.624012.

11. Shkoporov, A. N. *et al.* Long-term persistence of crAss-like phage crAss001 is associated with phase variation in Bacteroides intestinalis. *BMC Biol* **19**, 163 (2021).

12. Porter, N. T. *et al.* Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. *Nat Microbiol* **5**, 1170 (2020).

13. Ben-Assa, N. *et al.* Analysis of a phase-variable restriction modification system of the human gut symbiont Bacteroides fragilis. *Nucleic Acids Res* **48**, 11040–11053 (2020).

14. Wexler, H. M. Bacteroides: the Good, the Bad, and the Nitty-Gritty. *Clin Microbiol Rev* **20**, 593 (2007).

15. Patrick, S. *et al.* Multiple inverted DNA repeats of Bacteroides fragilis that control polysaccharide antigenic variation are similar to the hin region inverted repeats of Salmonella typhimurium. *Microbiology (Reading)* **149**, 915–924 (2003).

16. Krinos, C. M. *et al.* Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature* **414**, 555–558 (2001).

17. Chatzidaki-Livanis, M., Coyne, M. J., Roche-Hakansson, H. & Comstock, L. E. Expression of a uniquely regulated extracellular polysaccharide confers a large-capsule phenotype to Bacteroides fragilis. *J Bacteriol* **190**, 1020–1026 (2008).

18. Coyne, M. J. & Comstock, L. E. A New Pillar in Pilus Assembly. *Cell* **165**, 520–521 (2016).

19. Fletcher, C. M., Coyne, M. J., Bentley, D. L., Villa, O. F. & Comstock, L. E. Phase-variable expression of a family of glycoproteins imparts a dynamic surface to a symbiont in its human intestinal ecosystem. *Proc Natl Acad Sci U S A* **104**, 2413–2418 (2007).

20. Nakayama-Imaohji, H. *et al.* DNA Inversion Regulates Outer Membrane Vesicle Production in Bacteroides fragilis. *PLoS One* **11**, (2016).

21. Coyne, M. J., Weinacht, K. G., Krinos, C. M. & Comstock, L. E. Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. *Proc Natl Acad Sci U S A* **100**, 10446 (2003).

22. Martens, E. C., Roth, R., Heuser, J. E. & Gordon, J. I. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. *J Biol Chem* **284**, 18445–18457 (2009).

23. Chatzidaki-Livanis, M., Weinacht, K. G. & Comstock, L. E. Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont Bacteroides fragilis. *Proc Natl Acad Sci U S A* **107**, 11976–11980 (2010).

24. Saba, J. *et al.* Bacteroides expand the functional versatility of a conserved transcription factor and transcribed DNA to program capsule diversity. *Nature Communications 2024 15:1* **15**, 1–16 (2024).

25. Carasso, S. *et al.* Inflammation and bacteriophages affect DNA inversion states and functionality of the gut microbiota. *Cell Host Microbe* **32**, 322-334.e9 (2024).

26. Yutin, N. *et al.* Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nat Commun* **12**, (2021).

27. Ramos-Barbero, M. D. *et al.* Recruitment of complete crAss-like phage genomes reveals their presence in chicken viromes, few human-specific phages, and lack of universal detection. *ISME J* **18**, wrae192 (2024).

28. Cortés-Martín, A. *et al.* Adaptations in gut Bacteroidales facilitate stable co-existence with their lytic bacteriophages. *bioRxiv* (2024) doi:10.1101/2024.11.17.624012.

29. Hesse, S. *et al.* Phage resistance in multidrug-resistant Klebsiella pneumoniae ST258 evolves via diverse mutations that culminate in impaired adsorption. *mBio* **11**, (2020).

30. Garb, J. *et al.* Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. *Nat Microbiol* **7**, 1849–1856 (2022).

31. Patel, P. H. *et al.* Anti-phage defence through inhibition of virion assembly. *Nat Commun* **15**, (2024).

32. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).

33. Tock, M. R. & Dryden, D. T. F. The biology of restriction and anti-restriction. *Curr Opin Microbiol* **8**, 466–472 (2005).

34. Yuan, Y. *et al.* Phosphorothioate DNA modification by BREX Type 4 systems in the human gut microbiome. *bioRxiv* 2024.06.03.597175 (2024) doi:10.1101/2024.06.03.597175.

35. Goldfarb, T. *et al.* BREX is a novel phage resistance system widespread in microbial genomes . *EMBO J* **34**, 169–183 (2015).

36. Doron, S. *et al.* Systematic discovery of antiphage defense systems in the microbial pangenome. *Science (1979)* **359**, (2018).

37. Tesson, F. *et al.* Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat Commun* **13**, (2022).

38. Millman, A. *et al.* An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* **30**, 1556-1569.e5 (2022).

39. Georjon, H. & Bernheim, A. The highly diverse antiphage defence systems of bacteria. *Nature Reviews Microbiology 2023 21:10* **21**, 686–700 (2023).

40. Lewis, J. M., Janda, K. E., Kotter, D. B., Grose, J. H. & McCleary, W. R. Characterization of the Attachment of Three New Coliphages onto the Ferrichrome Transporter FhuA. *J Virol* **97**, e00667-23 (2023).

41. Rabsch, W. *et al.* FepA- and TonB-dependent bacteriophage H8: Receptor binding and genomic sequence. *J Bacteriol* **189**, 5658–5674 (2007).

42. Porter, N. T., Canales, P., Peterson, D. A. & Martens, E. C. A subset of polysaccharide capsules in the human symbiont Bacteroides thetaiotaomicron promote increased competitive fitness in the mouse gut. *Cell Host Microbe* **22**, 494 (2017).

43. Béchon, N. *et al.* Capsular polysaccharide cross-regulation modulates Bacteroides thetaiotaomicron biofilm formation. *mBio* **11**, e00729-20 (2020).

44. Imamovic, L., Ballesté, E., Martínez-Castillo, A., García-Aljaro, C. & Muniesa, M. Heterogeneity in phage induction enables the survival of the lysogenic population. *Environ Microbiol* **18**, 957–69 (2016).

45. Jones, K. R., Belvin, B. R., Macrina, F. L. & Lewis, J. P. Sequence and characterization of shuttle vectors for molecular cloning in Porphyromonas, Bacteroides and related bacteria. *Mol Oral Microbiol* **35**, 181 (2020).

46. Jimenez, N. *et al.* Molecular Analysis of Three Aeromonas hydrophila AH-3 (Serotype O34) Lipopolysaccharide Core Biosynthesis Gene Clusters. *J Bacteriol* **190**, 3176 (2008).

47. Livingston, S. J., Kominos, S. D. & Yee, R. B. New medium for selection and presumptive identification of the Bacteroides fragilis group. *J Clin Microbiol* **7**, 448 (1978).

48. Sambrook, J. & Russell, D. Molecular cloning: A laboratory manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY* 999 (2001).

49. Adams, M. H. *Bacteriophages*. (New York, Interscience Publishers, New York, 1959).

50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–20 (2014).

51. Wingett, S. & Andrews, A. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* **7**, 1338 (2018).

52. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455–77 (2012).

53. Hyatt, D. *et al.* Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).

54. Schwengers, O. *et al.* Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* **7**, (2021).

55. Sayers, E. W. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **49**, D10–D17 (2021).

56. Darling, A. E., Mau, B. & Perna, N. T. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5**, (2010).

57. Rodriguez-R, L. M. & Konstantinidis, K. T. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Prepr* **4**, e1900v1 (2016).

58. Couvin, D. *et al.* CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res* **46**, W246–W251 (2018).

59. Jiang, X. *et al.* Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. *Science (1979)* **363**, 181–187 (2019).

60. Hitchcock, P. J. & Brown, T. M. Morphological heterogeneity among Salmonella lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J Bacteriol* **154**, 269 (1983).

61. Osborn, M. J. [21] Preparation of lipopolysaccharide from mutant strains of Salmonella. *Methods Enzymol* **8**, 161–164 (1966).

62. Thein, M., Sauer, G., Paramasivam, N., Grin, I. & Linke, D. Efficient subfractionation of gram-negative bacteria for proteomics studies. *J Proteome Res* **9**, 6135–6147 (2010).

63. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367–1372 (2008).

64. Babicki, S. *et al.* Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* **44**, W147–W153 (2016).

Acknowledgements

This research was funded by grant number PID2020-113355GB-I00 by MICIU/AEI/ 10.13039/501100011033 and by “ERDF A way of making Europe. L. S-C. is a Beatriu de Pinos fellow from the Catalan Government. M. M. is a researcher of the ICREA Academia 2022 program.

Authors' contributions

C. G-G generated the mutants and performed the laboratory assays and cloning. MD. R-B conducted bioinformatics analyses, prepared figures, and contributed to writing and revising

the paper. L. S-C performed infection experiments and calculated statistical analysis, prepared figures, and contributed to writing the paper. P. B.-P. isolated phages, performed infection experiments and conducted phage enumeration. L. R-R supervised and interpreted proteomic analysis, wrote and revised the paper and discussed results. C. G-A supervised infectivity experiments, analyzed data and revised the paper. M. M. is the corresponding author, supervised and funded the study, designed experiments, analyzed data, and wrote the paper. All authors revised the final manuscript.

Competing interests

The authors declare no competing interests.

Figure legends

Figure 1. Infectivity dynamics of crAssBcn phages. A) Number of infectious phage particles measured by the double agar layer method (pfu/ml). B) Genome copies of phage particles measure by qPCR (GC/ml). C) Optical density at 600 nm of the host strain cultures of *B. intestinalis*. D) Growth of the host strain cultures *B. intestinalis* measured by colony counts (cfu/ml). Each value corresponds to the average of three independent experiments.

Figure 2- CrAssBcn phages infectivity on WT, R and S derivatives. Infection was conducted with phage suspensions of similar titer; ΦCrAssBcn6 (6.4 x 10^8 pfu/ml) ΦCrAssBcn10 (5.8 x 10^8 pfu/ml) and ΦCrAssBcn14 (4.0 x 10^8 pfu/ml). The hosts were the strains subcultured five successive times in the absence of phages. A) Phage propagation when inoculating ΦCrAssBcn6, ΦCrAssBcn10 and ΦCrAssBcn14 phages using WT, S, R6, R10 and R14 strains as hosts. B) Efficiency of plating of phages ΦCrAssBcn6, ΦCrAssBcn10 and ΦCrAssBcn14 onto the different strains.

Figure 3.- Location of the CPS and the PVR7 in the resistant and sensitive strains of *B. intestinalis*. A) Multialignment of the genomes of the four strains. Mauve comparison of the four genomes shows minimal differences, particularly in the CPS operons. Location of the CPS is shown in colors in the left side. B) Detail of the CPS7 operon with the ORFs identified indicated in colors and those corresponding to hypothetical proteins labeled in grey. At the right side of the B pannel the PVR7 region is indicated. C) Detail of the PCR7 region showing the inversion and how it regulates the CPS expression in the WT and S strains compared to R10 and R14 strains. Resistant strains show the PVR7 in the opposite orientation than WT and S mutants. D) sequence of PVR7 (186 bp). *ON* orientation is the inverse and complementary sequence of the *OFF* orientation found by PhaseFinder (supplementary Data 5). The PVR7 is flanked by two inverted repeats (IR in yellow and red) of 18 b.

Figure 4. A) Schematic representation of the genomic organization of the PVR7. *OFF and ON* orientation and the primer pair combinations used to identify the PVR7 orientation as well as primers for amplification of the PVR7 area regardless its orientation. B) Gel electrophoresis of PCR products obtained with the different primer pairs for defining the PVR7 orientation of WT, S and R strains. C) Relative abundance of reads showing the *ON* and *OFF* orientations in the genomic analysis of the sensitive (WT and S) and resistant (R10 and R14) strains and the ratios of reads in *ON* orientation vs total number of reads being 1 all reads in the *ON* orientation.

Figure 5.-. Upregulation of the ORFs of CPS7 in the resistant strains (R10 and R14) and the sensitive strains (WT and S). A) Heatmap constructed with the values of the relative abundances of each ORF in CPS7 identified by proteomic analysis. Higher relative abundances (supplementary data 6) are displayed by a stronger intensity. Right box shows

the ORFs detected in the proteomic analysis and the corresponding protein encoded if identified. B) Relative gene expression (log_2-fold change) assays of the seven ORFs identified in the proteomic analysis in R10 and R14 gene expression in comparison to the WT strain using RT-qPCR.

Figure 6.- Representative experiment of inhibition of crAssBcn phages infection. Phage enumeration (pfu/ml) of a suspension of phage ΦCrAssBcn10 against the membrane fraction of *B. intestinalis* strains (WT, S, R10, R14 and S_R (pG106::PVR7-ON). The phage stock not incubated with the membrane fraction was used as control. As non-specific binding control, an inactivated membrane fraction (inact membrane) of WT strain was used.

Supplementary material

Supplementary Data 1. Infectivity of CrAssBcn phages. *Excel file*

Supplementary Data 2. *B. intestinalis* genome annotation. *Excel file*

Supplementary Data 3. Recriprocal best match of WT genes vs S, R10 and R14 genes (% nt identity between compared genes). The results were sorted by minimum % nucleotide identity. *Excel file*

Supplementary Data 4. Evaluation of CRISPR-Cas and BREX systems in the genomes of *B. intestinalis* strains. *Excel file*

Supplementary Data 5. Relative frequence of inversions in OFF (forward) and ON (reverse) in the genomic reads of the different strains of *B. intestinalis*. *Excel file*

Supplementary Data 6. Data on the relative abundance of the proteins encoded by the ORFs in the different B. intestinalis strains that were detected by proteomics used to generate the heatmap in Fig. 5A. *Excel file*

Supplementary Table 1. Oligonucleotides used in this study

Target	Name	Forward primer (5' – 3')	Name	Reverse primer (5' – 3')	Amplicon size
Confirmation of S phenotype (Internal S + external primers)	Int-S	TTCAACCTGTATCTGTCCTAAG	Ext1	CCACCTACCTATTCATATAAAG TGTGTGTA	409 bp
Confirmation of R phenotype (Internal R + external 2 primer)	Int-R	GTGCACAAAGATAATGTCTTGA	Ext1	CCACCTACCTATTCATATAAAG TGTGTGTA	418 bp
Confirmation of PVR7 (external 1 + external 2 primers)	Ext2	CGATAAGTGGCACGCATAGC	Ext1	CCACCTACCTATTCATATAAAG TGTGTGTA	696 bp
Cloning of insert PVR7 with *Bam*H1 cutting site.	Ext2-bam	CGCGGATCCCGATAAGTGGCACG CATAGC	Ext1-bam	CGCGGATCCCCACCTACCTAT TCATATAAAGTGTGTGTA	714 bp
pG106 primers	M13-f	CAGGAAACAGCTATGAC	M13-r	TGTAAAACGACGGCCAGT	800 bp – PVR7
Confirmation of ON orientation of PVR7 insert in pG106	M13-f	CAGGAAACAGCTATGAC	Int-R	GTAATATGTGACACTGA	425 bp – PVR7
Confirmation of OFF orientation of PVR7 inserts in pG106	M13-f	CAGGAAACAGCTATGAC	Int-S	GTGCACAAAGATAATGT	522 bp – PVR7
ORF II	ORFII-f	CTAATGTCTGTCTTGCCCTAC	ORFII-r	AGGGAGAGCATCAAGAAGA	119 bp
ORF III	ORFIII-f	GGGATATGCCAAGGATTATGT	ORFIII-r	GCAAACTCACGGACTGTAT	108 bp
ORF V	ORFV-f	GCTACCTCTATCGATGTTTACAA	ORFV-r	CCACCTCTGTCCAAGAATTAC	112 bp
ORF VI	ORFVI-f	GGAGACTCAAACCTGAAGATG	ORFVI-r	GTAGCTTCTCCTGTAAGAACAC	112 bp
ORF VIII	ORFVIII-f	TATGGATGTACGCCTGAAATG	ORFVIII-r	TCACGGATAGAAAGACCTGTA	110 bp
ORF X	ORFX-f	AATGTGTGGATTGCTGATAAAG	ORFX-r	CCCTGCGGCTAAACAATA	120 bp
ORF XVIII	ORFXVIII-f	TCGCAAGCTTGTCACTTT	ORFXVIII-r	CCTCGACCAATACATTACCTG	110 bp
Bacteroides 16S	16S-f	GTAGCGGTGAAATGCTTAGA	16S-r	CCTGTTTGATACCCACACTT	102 bp

4.3. Capítulo 3:

El reclutamiento de genomas completos de fagos tipo crAss revela su presencia en viromas de pollo, pocos fagos específicos de humanos y la ausencia de detección universal

Recruitment of complete crAss-like phage genomes reveals their presence in chicken viromes, few human-specific phages, and lack of universal detection

María Dolores Ramos-Barbero, **Clara Gómez-Gómez**, Gloria Vique, Laura Sala-Comorera, Lorena Rodríguez-Rubio y Maite Muniesa

The ISME Journal (2024); 18(1), wrae192

DOI: 10.1093/ismejo/wrae192

Introducción

Los fagos crAssphage, identificados por primera vez en el 2014 a través de un análisis metagenómico de heces humanas (Dutilh et al., 2014), pertenecen al orden Crassvirales, que comprende hasta el momento 4 familias, 11 subfamilias, 42 géneros y 73 especies (https://ictv.global/taxonomy), tratándose de una taxonomía en constante evolución gracias al descubrimientos de nuevos miembros (Yutin et al., 2021). Se trata de un grupo monofilético de fagos con genomas de ADN circular bicatenario de aproximadamente 100 kilobases. Aunque la mayoría de los fagos tipo crAss permanecen sin cultivar, los estudios sugieren que estos fagos son virulentos y tienen como hospedador a bacterias del género *Bacteroides* (Guerin et al., 2018; Yutin et al., 2021). Hasta la fecha, solo unos pocos fagos han sido aislados exitosamente por técnicas de cultivo, como por ejemplo ΦcrAss001, ΦcrAss002, DAC15, DAC17 o los nuevos 25 fagos (crAssBcn) aislados por nuestro grupo, mientras que el resto de los genomas conocidos provienen de ensamblajes compuestos, lo que limita su estudio.

El primer fago descrito del orden Crassvirales, *p*-crAssphage, se ha podido detectar en altas concentraciones en aguas residuales y biosólidos especialmente de países occidentales (Smith et al., 2023), mostrando una alta especificidad en las muestras humanas y una prevalencia global (Stachler and Bibby, 2014; Edwards et al., 2019; Smith et al., 2023). Sin embargo, su diversidad ha sido subestimada, confundiéndose con un único tipo de virus,

además de detectarse, aunque en menor escala, en muestras de animales (Stachler and Bibby, 2014; García-Aljaro et al., 2017; Li et al., 2021), posiblemente debido a una pérdida de la especificidad por los humanos o al uso de métodos moleculares para la detección de fragmentos de su genoma no específicos de humanos. Estas observaciones destacan la importancia de analizar genomas completos para identificar adecuadamente estos virus y confirmar su especificidad. Sin embargo, son pocos los estudios que han evaluado la especificidad humana o la distribución mundial de otros miembros de Crassvirales.

Objetivo

Este trabajo busca comprender la prevalencia y la especificidad de los fagos del orden Crassvirales por el intestino humano. Con este propósito, se evaluaron 91 genomas completos de fagos tipo crAss en un total de 1337 metagenomas virales, con el objetivo de determinar qué miembros del orden son exclusivos del ser humano. Además, se buscó caracterizar rasgos genómicos comunes, establecer relaciones filogenéticas, evaluar su distribución geográfica y prevalencia en metagenomas de diversas regiones del mundo, así como precisar diferencias entre viromas humanos y no humanos para delinear patrones específicos asociados a este grupo de fagos.

Resultados y discusión

Una vez creada la base de datos con los genomas fágicos completos, incluyendo los 25 fagos crAssBcn recientemente aislados en Barcelona, para evaluar la presencia y abundancia de estos fagos, se analizaron 60 viromas de origen humano (muestras fecales y orina humanas) y de muestras que pueden contener contaminación fecal humana (aguas residuales, suelo y alimentos). Los resultados mostraron que los fagos tipo crAss, especialmente los crAssBcn, tenían una alta representación en viromas de heces humanas y aguas residuales, pero una menor presencia en viromas de orina. En contraste, no se detectaron en viromas de alimentos o suelos con baja contaminación fecal, salvo en casos puntuales.

El análisis destacó que los fagos tipo crAss están estrechamente asociados con niveles altos de contaminación fecal humana, reflejado por las elevadas concentraciones de los indicadores bacterianos y virales como *E. coli* y colifagos somáticos en las muestras analizadas. Los resultados indican que estos fagos son más abundantes en muestras altamente contaminadas y escasos o ausentes en aquellas con niveles bajos de contaminación fecal, como alimentos y suelos agrícolas, al menos cuando son evaluados mediante un análisis metagenómico.

Se seleccionaron 28 fagos con una cobertura genómica horizontal superior al 40% en viromas de origen humano, incluyendo 27 fagos tipo crAss y todos los fagos crAssBcn que, dada su similitud, fueron representados en los siguientes experimentos únicamente por el

ΦCrAssBcn3. Estos 28 fagos también fueron detectados en viromas de aguas residuales, aunque con una cobertura generalmente menor, posiblemente atribuible a la mayor diversidad genética de estas muestras que reduce la recuperación de los genomas (Rodriguez-R and Konstantinidis, 2014; Newton et al., 2015; Zaheer et al., 2019). A pesar de esto, se priorizó la selección de fagos presentes en viromas fecales humanos para los análisis posteriores, independientemente de su abundancia relativa en aguas residuales.

Seguidamente, los 28 fagos humanos seleccionados en el ensayo anterior se buscaron en 24 viromas fecales de origen animal con alta calidad de secuencia, para determinar aquellos con menos del 40% de su genoma reclutado en este tipo de muestras, es decir ausentes. En general, los viromas de vacas, conejos, cerdos y primates no humanos carecían de fagos tipo crAss, lo que apoya su residencia preferente en el intestino humano. Sin embargo, se detectaron algunos fagos en los dos sets de viromas de pollos (China y España), aunque en éstos la contaminación fecal humana no pudo descartarse completamente. Concretamente, se identificaron 15 fagos altamente reclutados en ambos viromas y, por lo tanto, fueron excluidos por falta de especificidad humana. En total, se detectaron cuatro fagos exclusivamente presentes en muestras humanas, incluyendo dos fagos aislados (DAC15 y DAC17) y dos no cultivados (cr85 y cr17). Además, se seleccionaron también nueve fagos que presentaban menos del 40% de cobertura genómica en, al menos, uno de los dos viromas de pollo (especificidad parcial en humanos). Esto se debe a que, al no poder descartar una posible contaminación fecal humana en estos viromas de pollo, no se quiso excluir posibles candidatos basados en resultados de una única área geográfica. Por otro lado, se ha descrito la posible adaptación de especies de *Bacteroides* entre humanos y aves, lo cual aumenta la probabilidad de adaptación de los fagos que infectan dichas especies (Kollarcikova et al., 2020). Por lo tanto, tras este análisis se obtuvieron un total de 13 fagos tipo crAss humanos y ausentes, total o parcialmente, en viromas de animales.

Estos 13 fagos del tipo crAss fueron alineados para identificar alguna región común en su genoma, lo que podría ser de mucha utilidad a la hora de desarrollar herramientas moleculares para su detección. Sin embargo, el análisis reveló una alta heterogeneidad genética, distribuyéndolos en tres familias (*Steigviridae, Intestiviridae* y *Suoliviridae*), ocho géneros y 12 especies diferentes, sin poseer una región genómica central común entre ellos. En la comparación filogenómica, los fagos más específicos de humanos fueron agrupados en tres clústeres correspondientes a las tres familias distintas identificadas, lo que coincide con su clasificación taxonómica. Esto indica que la especificidad humana no deriva de un ancestro común, sino de adaptaciones evolutivas separadas, posiblemente a través de eventos simbióticos con su cepa bacteriana hospedadora dentro del reservorio humano.

El siguiente paso fue evaluar la distribución geográfica de los 13 fagos tipo crAss humanos, utilizando una colección de 1260 viromas intestinales humanos de 14 países diferentes. Se encontró que el 77% de los viromas reclutaron, al menos, uno de los 13 fagos. Sin embargo, los resultados mostraron una variabilidad en la distribución de los diferentes fagos en las diferentes regiones, con países como Tanzania que mostró la menor diversidad (solo 8 fagos), mientras que en otros países fueron detectados los 13 fagos a la vez, aunque ninguno de los fagos tipo crAss humanos predominó en todas las ubicaciones. Algunos fagos no aislados, como el cr53, cr12 y el cr112, presentaron abundancias relativas altas, mientras que ΦcrAss002 fue el más común a nivel global en términos de fagos aislados, aunque también fue detectado en pollos. Esto no sucedió con DAC15 y DAC17 que fueron los más específicos de humanos, pero en cambio fueron los menos abundantes globalmente. Además, se observó que la prevalencia de los fagos no solo dependía de su abundancia, sino también del reclutamiento completo de su genoma, lo que varió según el país. También influyó la profundidad de secuenciación de los viromas de cada país, la cual fue mayor en España, mientras que los metagenomas de China y Tanzania presentaron una menor profundidad, comprometiendo el reclutamiento de los fagos en estudio. Por ejemplo, el fago cr17 mostró una gran abundancia en USA 1, pero fue descartado por una baja cobertura genómica. El fago cr85 fue el más abundante globalmente, predominando en España, pero su presencia fue limitada en otros países. En cuanto a los fagos cr6 y cr12, destacaron por su amplia distribución y reclutamiento en múltiples países, pero tampoco alcanzaron la universalidad.

La especificidad de algunos fagos tipo crAss por el intestino humano se atribuye a su capacidad para infectar cepas humanas de *Bacteroides*, en lugar de ser una característica intrínseca de los fagos. La existencia de una distribución geográfica diversa de los fagos tipo crAss humanos se podría explicar debido a la sensibilidad de las cepas hospedadoras a distintos fagos según la región geográfica donde han sido aislados. Dado que *Bacteroides* presenta una gran capacidad de adaptación al entorno intestinal, quiere decir que estas diferencias geográficas podrías verse influenciadas por factores epigenéticos como la edad, la salud o la dieta. Por este motivo, la aparición de estos fagos pudo ocurrir de manera simultánea y aleatoria en diversas poblaciones humanas, lo que explicaría su variada distribución geográfica y la ausencia de una relación filogenética.

Por último, se exploró la distribución global del fago *p*-crAssphage, usando su genoma completo, ya que había sido propuesto como marcador de contaminación fecal humana, útil para el rastreo del origen de contaminación fecal o MST y de prevalencia universal (Edwards et al., 2019). Sin embargo, este fago no fue incluido en nuestro análisis debido a su alta prevalencia en ambos viromas de pollo. Además, se detectó de forma muy abundante en viromas de países como Irlanda, España y USA, pero, sin embargo, su presencia fue baja o

inexistente en el resto de las ubicaciones, comprometiendo su universalidad y contradiciendo estudios previos. Esto sugiere que, aunque algunos fagos de tipo crAss tienen una amplia distribución, no hay un fago universalmente presente en todos los países estudiados, además de que la mayoría de los Crassvirales no son exclusivos de viromas humanos.

Conclusiones

Este estudio se centró en determinar la especificidad humana de fagos tipo crAss ya descritos, destacando la dificultad de detectar secuencias virales en metagenomas y de reconstruir genomas completos debido a su diversidad genética. Los resultados sugieren que los fagos tipo crAss humanos no derivan de un ancestro común, sino que pertenecen a diferentes linajes dentro del orden Crassvirales, sin compartir rasgos genéticos comunes que expliquen su especificidad humana, lo cual dificulta el diseño de herramientas moleculares para su detección. Además, la distribución global de estos fagos no es uniforme, y su presencia parece depender de la relación simbiótica con las cepas bacterianas hospedadoras, que varían según la población humana y su localización geográfica. En ausencia de un candidato mejor, *p*-crAssphage continuará siendo el candidato utilizado dado que, cuando está presente, lo está de forma abundante, aunque es preciso ser conscientes de sus limitaciones en cuanto a especificidad y universalidad.

The ISME Journal, 2024, **18(1)**, wrae192
https://doi.org/10.1093/ismejo/wrae192
Advance access publication: 3 October 2024
Original Article

Recruitment of complete crAss-like phage genomes reveals their presence in chicken viromes, few human-specific phages, and lack of universal detection

María Dolores Ramos-Barbero, Clara Gómez-Gómez, Gloria Vique, Laura Sala-Comorera, Lorena Rodríguez-Rubio, Maite Muniesa*

Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Avinguda Diagonal, 643, Prevosti Building, Floor 0. Barcelona E-08028, Spain

*Corresponding author: Maite Muniesa, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Avinguda Diagonal, 643, Prevosti Building. Floor 0. Barcelona E-08028, Spain. Email: mmuniesa@ub.edu

Abstract

The order Crassvirales, which includes the prototypical crAssphage (p-crAssphage), is predominantly associated with humans, rendering it the most abundant and widely distributed group of DNA phages in the human gut. The reported human specificity and wide global distribution of p-crAssphage makes it a promising human fecal marker. However, the specificity for the human gut as well as the geographical distribution around the globe of other members of the order Crassvirales remains unknown. To determine this, a recruitment analysis using 91 complete, non-redundant genomes of crAss-like phages in human and animal viromes revealed that only 13 crAss-like phages among the 91 phages analyzed were highly specific to humans, and p-crAssphage was not in this group. Investigations to elucidate whether any characteristic of the phages was responsible for their prevalence in humans showed that the 13 human crAss-like phages do not share a core genome. Phylogenomic analysis placed them in three independent families, indicating that within the Crassvirales group, human specificity is likely not a feature of a common ancestor but rather was introduced on separate/independent occasions in their evolutionary history. The 13 human crAss-like phages showed variable geographical distribution across human metagenomes worldwide, with some being more prevalent in certain countries than in others, but none being universally identified. The varied geographical distribution and the absence of a phylogenetic relationship among the human crAss-like phages are attributed to the emergence and dissemination of their bacterial host, the symbiotic human strains of *Bacteroides*, across various human populations occupying diverse ecological niches worldwide.

Keywords: Crassvirales, crAss-like phage, bacteriophage, Bacteroides, Gut viromes, fecal microbiota, biogeography

Introduction

The crAssphage (cross-assembly phage) group of phages first came to the attention of the scientific community in 2014 after its *in silico* identification in a metagenomic analysis of human feces [1]. Initially considered as a single entity, the first prototypical crAssphage (or p-crAssphage) was later described as belonging to the new order Crassvirales, family *Intestiviridae*, species *Carjivirus communis*. The order Crassvirales currently comprises four families, 11 subfamilies, 42 genera, and 73 new species (https://ictv.global/taxonomy) [1–5]. All members of Crassvirales, also known as crAss-like phages, exhibit similar genomic architecture, featuring a relatively large double-stranded circular DNA genome of ∼100 kilobases. Crassvirales is a monophyletic group with a common evolutionary origin. Some of the phage proteins in this group are universally conserved across crAss-like phage subfamilies and share common ancestors [4, 5]. However, within the order, different branches can also be observed, and the taxonomy is continuously evolving with the discovery of new crAss-like phage members [6].

Although most crAss-like phages remain uncultured, analysis of the few isolated virions and the available genomes of crAss-like phages suggests they share a podovirus-like morphology and function as virulent phages, targeting bacteria within the phylum Bacteroidota, specifically the genus *Bacteroides* [4, 6].

Efforts to isolate crAss-like phages were unsuccessful until four years after the discovery of p-crAssphage, when ΦCrAss001 was isolated in *Bacteroides intestinalis* [7] and assigned to the family *Steigviridae*, species *Kehishuvirus primarius* [8]. The isolation of crAss-like phages in pure culture remains challenging, with only a few other crAss-like phages isolated to date: DAC15 and DAC17 (family *Steigviridae*, species *Wulfhauvirus bangladeshii*) [8], infecting *Bacteroides tethaiotaomicron* [9] in 2020; ΦCrAss002 (family *Intestiviridae*, species *Jahgtovirus secundus*) [8] infecting *Bacteroides xylanisolvens* in 2021 [10]; three new species in 2023 (family *Steigviridae*, species *Kehishuvirus tikkala* (Bc01), and *Kolpuevirus frurule* (Bc03), and family *Intestiviridae*, species *Rudgehvirus jaberico* (Bc11)) infecting *Bacteroides cellulosilyticus* [11]; although initially it was not classified, a *Cellullophaga* phage ϕ14:2, a marine isolate [12] was later included as a member of the Crassvirales order (family *Steigviridae*, species *Akihdevirus balticus*) [6]; and recently 25 new phages (crAssBcn phages 1–25) of the genus *Kehishuvirus*, belonging to at least six different species, infecting *B. intestinalis*

Received: 2 May 2024. **Revised:** 25 July 2024. **Accepted:** 1 October 2024
© The Author(s) 2024. Published by Oxford University Press on behalf of the International Society for Microbial Ecology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

[5]. All other crAss-like phages reported so far correspond to genomes generated as a result of composite assemblies and have never been propagated in pure culture in a laboratory, which limits their study.

p-crAssphage has been found to exhibit a high level of host specificity and high abundance in sewage and biosolids in Europe and the United States. Initially, it was thought to be less abundant in sewage samples from Asia and Africa [13, 14], but subsequent research using DNA sequences of p-crAssphage demonstrated its global prevalence [15]. Although crAss-like phages are thought to exhibit high human specificity, qPCR analysis has also revealed the presence of p-crAssphage in samples containing only animal feces [14, 16, 17], albeit at lower levels than in human feces [17]. In addition to p-crAssphage, previous taxonomical analysis described non-human, environmental phages in the Epsilon and Zeta groups of Crassvirales [6]. However, to date no more studies have assessed the level of human specificity of other members of the order Crassvirales and little information is available about the worldwide distribution of other crAss-like phages [4, 5].

To fully evaluate the Crassvirales group and to determine how this group of phages has spread across the human populations around the globe, a comprehensive understanding of their specificity and prevalence in the human gut is required. However, identifying human-specific crAss-like viruses within the vast diversity of Crassvirales remains an ongoing task and warrants further investigation.

In this study we used a custom database of 91 crAss-like phages. The database included only complete phage genomes to evaluate their abundance and prevalence in a collection of 1337 viral metagenomes with different levels of human pollution, from human feces, environmental samples (wastewater and soil), and food, as well as their absence in a collection of animals fecal viromes (feces of chickens, rabbits, cows, pigs, and non-human primates). The phylogenomic relationship and the geographical distribution of the human crAss-like phages were studied across metagenomes from a wide range of countries. This comprehensive metagenomic recruitment analysis aimed to determine which members of the order Crassvirales are exclusively found in humans, to ascertain whether the human-specific phages share common genomic traits and are phylogenetically related, and finally to delineate their biogeographical prevalence.

Materials and methods

Samples

This study evaluated the presence of 91 complete, non-redundant genomes of crAss-like phages selected from public databases (Supplementary Data 1) in 84 viromes obtained from human, food, environmental, and animal samples (Supplementary Data 2). Among these, environmental viromes including soil and wastewater, and viromes of dairy products (cheese, milk, and kefir), were obtained in this study. Viromes from human feces and urine, food viromes (vegetables, meat, and fish), and two viromes of chicken feces, were all generated by our group in previous studies [18–21]. A brief summary of the preparation methods for those viromes appears below. Other animal viromes (pig, rabbit, cattle, and non-human primates were generated by other authors and are available in the databases (Supplementary Data 2).

Human samples

Seven viromes from human feces and 16 viromes from human urine were obtained as described in previous studies [19, 20]. All samples were collected from healthy adults, were anonymized, and all individuals were volunteers. The original study that generated the sequences used here was approved by the Clinical Ethics Committee of the Hospital de la Santa Creu i Sant Pau (12/065/1350) [20].

Environmental samples

Six raw sewage samples were collected over one year period from the influent of two urban wastewater treatment facilities (Besòs and Prat) in Catalonia, northeastern Spain. Each facility serves a population of over 2 000 000 inhabitants. Samples were collected in sterile containers, transported to the laboratory within 2 h of collection, and processed immediately for the analysis of fecal indicators and bacteriophage isolation, as described below. Six wastewater viromes, three from each wastewater treatment plant were generated with these samples.

Two soil viromes were generated in this study from two pools of soil samples, each consisting of a mixture of five samples (5 g of each sample) from croplands in Barcelona (northeast Spain) and five samples from Asturias (north Spain). For each pool, 25 g was homogenized in 50 ml of PBS using stomacher bags with filters (Afora) in a Masticator (IUL Instruments GmbH) for 2 min. After the homogenization with 50 ml of PBS, phage particles were recovered by the addition of 1 g of beef extract (Sigma Aldrich) to the homogenate and increasing the homogenate pH to 8.5–8.7 with NaOH 0.1 N to detach viruses that can be adhered to soil particles (method adapted from [22]). The homogenate obtained was processed for purification of phage particles as described below.

Food samples

Twenty-nine food virome samples were analyzed, four were generated in this study and the rest were obtained in a previous study from our group [18]. The meat pool included six viromes of chicken meat, two of beef, and two of pork. In the fish category, one pool was obtained from mussels, one of clams (shellfish 2), fresh hake (Atlantic fish 1), sardine (Atlantic fish 2), frozen hake (frozen fish 1), frozen monkfish (frozen fish 2), whiting (Mediterranean fish), salmon (farm fish 1), and trout (farm fish 2). The vegetable pool was comprised of two viromes obtained from lettuce, two from spinach, two from cucumber.

New viromes were generated in the category of vegetables, a nut milk product (horchata). In the category of dairy products, one virome of fresh milk, one of fresh cheese, and one of kefir Twenty-five g of kefir and fresh cheese, and 20 ml of fresh milk and unpasteurized horchata (nut-based milk) were homogenized/mixed for 2 min in 60 ml of phage buffer (22 mM KH_2PO_4, 50 mM Na_2HPO_4, 85 mM NaCl, 1 mM $MgSO_4$, and 0.1 mM $CaCl_2$) using the stomacher bags with filters (Afora) in a Masticator (IUL Instruments GmbH). The homogenates were centrifuged at 4000 $\times g$ and filtered through 0.22 μm polyethersulfone (PES) membranes. All samples were purchased from local retailers in Barcelona (Spain) in 2017–2020, transported in sterile containers, and kept at −80°C until analysis.

Non-human samples

Twenty-four viromes of animal feces were selected from public databases, including four chicken fecal viromes from Spain and China, two additional viromes of chicken feces, eight viromes of pig feces, one virome of cattle feces, two viromes of rabbit feces, three non-human primate viromes, and four viromes of makaka feces.

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

Culturable bacterial and viral fecal indicators

Samples were evaluated for the presence of fecal indicators, either in the present (environmental samples, horchata, and dairy) or in a previous study from our group [18]. Samples from animal feces were not processed in this study as the data for these viromes were obtained from public databases. *Escherichia coli* in samples or homogenates was determined on Chromocult Coliform Agar (Merck). Initial incubation was performed for 2 h at 37°C to acclimatize potentially damaged microorganisms and then overnight at 44°C. Ten per cent of the colonies turned blue, representing potential *E. coli*, and were confirmed with the Indole test and grown on McConkey agar. Somatic coliphages, proposed as viral indicators of fecal pollution [23], were analyzed after homogenate filtration to evaluate the presence of infectious fecal viruses in the samples. The filtrates were decimal diluted in PBS and somatic coliphages in one ml of each dilution were analyzed by the double agar layer method following the ISO standard [24], using *E. coli* strain WG5 (ATCC 700078) as host, and incubation at 37°C for 18 h. Each dilution was analyzed in triplicate. Each homogenate was analyzed in duplicate. Mann–Whitney U-test was performed to assess the statistically significant differences in *E. coli* or in somatic coliphages concentrations between the group of samples with high (human gut, human urine, wastewater 1 and 2) or low fecal load (soil, meat, fish, vegetable, dairy). For statistics, GraphPad Prism (GraphPad Software) was used. A difference was deemed significant when the *P*-value ≤ 0.05.

Purification of phage particles

For the viromes generated in this study, 50 ml of wastewater samples and 50 ml of the homogenate of soil samples, nut milk, and dairy products were filtered through 0.22 μm pore size, low-protein-binding PES membranes (Millipore) to remove bacteria and other particulate material. Filtered homogenates were 20-fold concentrated using polyethylene glycol (PEG) precipitation [25], followed by incubation at 4°C for 12 h, and centrifugation at 16000 $\times g$. Two ml of the PEG-concentrated suspensions were dialyzed and treated twice with chloroform 1:10 (*v:v*), mixed by vigorous vortexing for 5 min, and centrifuged at 16000 $\times g$ for 10 min. Chloroform treatment aims to break bacterial cells and any DNA-containing small vesicles that might have passed through the filters. The supernatant was treated with DNase I (100 units/ml; Sigma-Aldrich) for 1 h at 37°C to eliminate any non-packaged DNA. DNase was inactivated by heating for five min at 75°C. Human samples [19, 20], chicken feces [21], and the rest of food viromes [18]were processed in previous studies using the same protocol described above with minor modifications.

Phage DNA isolation

DNA isolation from the purified particles was performed with QIAamp DNA Blood (Qiagen GmbH), following the manufacturer's instructions, and the obtained DNA was suspended in 200 μl of sterile bi-distilled water. The DNA concentration of each pooled sample was evaluated using a Qubit Fluorometer (Life Technologies), and the DNA quality was further confirmed by the 2100 Bioanalyzer system (Agilent Technologies).

Sequencing

Five μl of DNA at a concentration of 0.2 ng/μl was fragmented and used to prepare libraries for whole genome sequencing using the Nextera XT Kit (Illumina) according to the manufacturer's protocol. The libraries were purified with AmPure beads (Beckman Coulter Inc.), checked for fragment distribution and size, and quantified using a TapeStation 4200 and the Agilent High Sensitivity D1000 ScreenTape system (Agilent Technologies) in a Quantus Fluorometer (Promega). An equimolar pool of the 25 individual phage genomes was separately sequenced on the NextSeq System (Illumina) with a high output run of 300 cycles.

Metagenomic sequence trimming

Metagenomic raw reads generated in this study and those downloaded from public databases were trimmed separately using Trimmomatic (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) [26]. Finally, trimmed metagenomes were pooled by origin (human gut, wastewater, food) as detailed in Supplementary Data 2. The quality of trimmed reads was checked by FastQC [27].

Presence, abundance and diversity of crAss-like phages in different viromes

To assess the presence of crAss-like phages in the different viromes, reads from the viromes in the database, which includes human, animal, and low fecal contamination viromes from this study (Supplementary Data 2), were recruited against the genomes of crAss-like phages. Selected crAss-like phage genomes have been used as subject in the BLASTn database and virome reads were the queries [28]. The presence and relative abundance of the human crAss-like phages were determined using standalone BLASTn with a cutoff of 70% query coverage, an *e*-value ≥0.1, and filtered by the "best hit" option by BlastTab.best_hit_sorted.pl script from the enveomics collection [29]. The default BLASTn parameters of VIRIDIC were used (–word_size 7 -reward 2 -penalty –3 -gapopen 5 -gapextend 2'), specie threshold 95%, and genus 70% of similarity) [28]. After applied the best-hit and query coverage cutoff, the relative abundances of each crAss-like phage in each virome were calculated using BlastTab.seqdepth.pl from enveomics collection [29]. Sequencing depth (SD) values were normalized by dataset size (Gbp) and genome length (Kbp) (SD/Gbp/Kbp). Fragment recruitment data were plotted using the R package from the enveomics collection [29]. The graphs (bars, boxplots, and heatmaps) were drawn using Plotly in R [30]. Finally, the average nucleotide identity of mapped reads (ANIr) of each virome was determined as previously described [31]. Only reads with ≥70% query coverage were considered for the analyses. ANIr and relative abundance heatmaps were drawn by heatmapper [32].

Genome comparison and phylogenetic analysis of human-specific crAss-like phages

Phage genomes were clustered at 100% identity by cd-hit-est [33]. Cd-hit-est does not cluster together genomes that are 100% identical but circularly permuted with different start sites. Therefore, VIRIDIC (BLASTn-based) [28], which checks each comparison for genomic synteny and calculates the intergenomic similarity of two viruses with BLASTn [34], was also used to assess similarities between phage genomes. Whole genome alignments were performed by Mauve of Geneious Prime version 2024.0.5 [35]. ViPTree v 4.0 [34] was used to create a viral "proteomic tree" with the phage genome sequences available in the published databases. The tree is based on a distance matrix of pairwise genome-wide similarities calculated using tBLASTx. The proteomic tree illustrates the global genomic similarity relationships among the selected phages and facilitates alignment visualization for comparative genomics of viruses.

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

Presence and abundance of human crAss-like phages worldwide

To evaluate the geographical distribution of the highest human-specific crAss-like phages, a collection of viral metagenomes was created in this study (updated from a previously reported global human viromes database) [5]. This collection comprised 1260 human gut viromes available in the databases from children and adults across 14 countries (Table 1; see Supplementary Data 3 for more extended information) that were recruited against the 13 human phages. The presence and relative abundance of the human crAss-like phages were determined as described above. Next, crAss-like phage abundances were calculated using sequencing depth (SD) as previously described, values normalized by dataset size (Gbp) and genome length (Kbp) (SD/Gbp/Kbp). Fragment recruitment data were plotted by the enveomics R package in the R statistical tool [29]. The graphs (bars, boxplots, and heatmaps) were drawn with Plotly by R [30].

Results

A crAss-like phage database was constructed in this study using 91 non-redundant complete phage genomes identified as crAssphage or crAss-like phage available in public databases (Supplementary Data 1). Of these, 25 crAssBcn phages were recently isolated in our area (Barcelona) [5]. The only isolated virions represented in this list of genomes were ΦCrAss002, from the *Intestiviridae* family, ΦCrAss001, DAC15, DAC17, and the 25 crAssBcn phages (from the *Steigviridae* family) [5, 7, 9, 10], whereas the rest (Supplementary Data 1) were uncultured viruses.

Presence and relative abundance of crAss-like phage genomes in human viromes and viromes containing human fecal pollution

To ascertain which members of the order Crassvirales were present in humans, the 91 complete genomes were recruited in a set of 60 viromes of human origin or impacted by humans (Supplementary Data 2), comprising 23 human samples, eight environmental samples, and 29 food samples. In the recruitment plots (Fig. 1A), the reads of each virome are mapped against each complete phage genome to verify phage presence/absence in the virome. An intense vertical black signal indicated high recruitment with a read coverage of over 70% (only reads that mapped with 95–100% identity are shown in the recruitment plot) and allowed selection of phage genomes showing high recruitment (Fig. 1). Weak signals indicated that reads only recruited against part of the phage genome and not to the whole phage genome, in which case these phages were considered not to be present in the virome. As the recruitment results showed a similar behavior for all crAssBcn phages (Fig. 1A), only one representative of the group (the genome of phage ϕCrAssBcn3) was used in subsequent analyses, as shown below. Read recruitment against phages was observed in the pool of human feces (comprising seven gut viromes) [19, 20], human urine (comprising 16 urine viromes) [19, 20], and in wastewater viromes, which were distributed in two pools, one containing three viromes from wastewater treatment plant 1 (Besòs) and one containing three viromes from wastewater treatment plant 2 (Prat), both obtained in the present study (Fig. 1A). Heatmap shows the average nucleotide identity of mapped reads (ANIr) from each individual virome for each crAss-like genome separately and revealed high ANIr in soil virome 1 for several crAss-like phages (Fig. 1B, Supplementary Data 2 for numerical data).

The fecal content of all samples used to obtain the viromes was evaluated. Human samples (feces and urine) and urban wastewater showed the highest levels of the bacterial fecal indicator (*E. coli*) and viral fecal indicators (somatic coliphages), reflecting their high fecal load (Supplementary Fig. 1). Differences in fecal indicators between highly polluted samples and low polluted samples (soil and food) were significant by the Mann–Whitney U-test (*E. coli*, $U = 20.0$, $P < 0.0159$ and somatic coliphages $U = 20.0$, $P < 0.0159$). As an effective indicator should also be detectable in samples expected to have a low fecal load, viromes from agricultural soil samples (obtained in this study) and food (vegetables, meat, fish, and dairy products) [50] (Supplementary Data 2), which had significantly less fecal contamination than the human samples (Supplementary Fig. 1), were used to recruit crAss-like phage genomes.

Many but not all of the complete crAss-like phage genomes were detected in samples with high levels of human fecal contamination (Fig. 1). Recruitment against crAss-like phages was higher in human feces and wastewater viromes, and lower in human urine viromes. In the case of crAssBcn phages, they were highly recruited in human feces and wastewater viromes but were less represented in urine viromes compared to other crAss-like phages.

In food samples and agricultural soil, the amount of human fecal pollution was very low (Supplementary Fig. 1). Regardless of the origin of the fecal pollution, crAss-like phages were not detected in these viromes (below recruitment detection limits), or only a few reads were recruited in one of the soil viromes against some of the phages (Fig. 1B).

The relative abundance of each crAss-like phage in each pool (Supplementary Fig. 2) is in accordance with the recruitment plots (Fig. 1A) and the ANIr (Fig. 1B). The highest relative abundance of phages corresponded to the human gut, urine, wastewater and soil viromes, being lower in food viromes. The relative abundance of crAssBcn phages was higher compared to the other crAss-like phages analyzed. In accordance with the recruitment plots (Fig. 1), crAss-like phages were not observed in food viromes that had very low levels of human fecal pollution.

Selecting crAss-like phages with the highest specificity to human viromes or viromes containing human fecal pollution

Using the selection criterion of "presence in human viromes", crAss-like phages with a horizontal genome coverage (defined as the portion of the phage genome covered by metagenomic reads) of over 40% in the human gut and human urine were chosen (Supplementary Data 4). These corresponded to 27 crAss-like phages and all the crAssBcn phages (Supplementary Data 1). As indicated above, ϕCrAssBcn3, one of the crAssBcn phages with a high percentage of recovery (Supplementary Data 1), was used to represent the whole group. In total, the 28 crAss-like phages with the best recruitment results in human viromes were included in the study.

The 28 crAss-like phages selected (Supplementary Data 1 and Fig. 1) were also detected in wastewater viromes. In this case, coverage levels were close to 40% in only 19 phages, 20–30% in four phages, and below 10% in the remaining five phages. Sequence diversity was lower in fecal than in wastewater samples, which are inherently more complex matrices, as discussed below. The higher diversity of wastewater viromes reduces the likelihood of detecting a particular genome in a sample, potentially contributing to a lower recruitment in wastewater viromes, even if phages are abundant in human feces viromes. Therefore, all phages present

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

Table 1. Origin and characteristics of the metaviromes used in this study. Metagenomes in bold correspond to those that recruited against specific crAss-like phages.

Country	# metaG_w notes	# of people	Age	Health state	Clean read Bps by country/sequencing platform	VLP or Bulk	MDA	Sample type	Extraction method	Detected crAss-like phages	Ref.
Ethiopia	**29**	**269**	**Children**	**Diarrheal disease**	**4 225 680 651**	**VLP**	**Yes**	**Feces**	**C, F, N**	**Yes**	[36]
USA	**16**	**10**	**Adults, Teens**	**Fecal transplantation**	**9 373 845 702**	**VLP**	**Yes**	**Feces**	**C, F, CE, N**	**Yes**	[37]
USA	**24**	**24**	**Children**	**IBD, Healthy**	**2 577 552 227**	**VLP**	**Yes**	**Feces**	**C, F, N**	**Yes**	[38]
China	**196**	**196**	**Adults**	**Hypertension, Healthy**	**1.05501 x10 ^{12}**	**Bulk**	**No**	**Feces**	**N/A**	**Yes**	[39]
Finland	**96**	**38**	**Infants**	**Type 1 diabetes, Healthy**	**15 942 577 574**	**VLP**	**Yes**	**Feces**	**C, F, N**	**Yes**	[40]
Ireland	**20**	**20**	**Infants**	**Various birth modes**	**5 488 316 876**	**VLP**	**Yes**	**Feces**	**C, F, P, N**	**Yes**	[41]
Uganda	**65**	**65**	**Adults**	**HIV, Healthy**	**32 402 875 119**	**VLP**	**Yes**	**Feces**	**C, F, N**	**Yes**	[42]
Spain	19	19	Adults	IBD, Healthy	6 246 589	VLP	Yes	Feces	C, F, P, N	Bdl	[43]
Tanzania	**27**	**27**	**Adults**	**Healthy**	**60 267 567 734**	**Bulk**	**No**	**Feces**	**N/A**	**Yes**	[44]
Peru	36	36	Adults	Healthy	1.23152 x10^{11}	Bulk	No	Feces	N/A	Bdl	[44]
Italy	11	11	Adults	Healthy	13 446 170 791	Bulk	No	Feces	N/A	Bdl	[44]
USA	22	22	Adults	Healthy	96 979 018 221	Bulk	No	Feces	N/A	Bdl	[44]
Malawi	**320**	**53**	**Children, Adults**	**Malnutrition, Healthy**	**2 387 517 871**	**VLP**	**Yes**	**Feces**	**C, F, N**	**Yes**	[45]
Ireland	**84**	**13**	**Adults**	**IBD, Healthy**	**1.31496 x10 ^{11}**	**VLP**	**Yes**	**Feces**	**C, F, P, N**	**Yes**	[46]
Cameroon	**63**	**221**	**Infants, Adults**	**Healthy (contact with bats)**	**73 694 302 905**	**VLP**	**Yes**	**Feces**	**C, F, N**	**Yes**	[47]
Estonia	51	4	Infant	Type 1 diabetes, Healthy	19 806 647 068	VLP	Yes	Feces	C, F, N	Bdl	[48]
Finland	169	18	Infant	Type 1 diabetes, Healthy	64 749 010 013	VLP	Yes	Feces	C, F, N	Bdl	[48]
Spain	**7**	**7**	**Adults**	**Antibiotic, Healthy**	**75 043 928**	**VLP**	**No**	**Feces**	**C, F**	**Yes**	[20]
Mexico	**5**	**28**	**Children**	**Obesity**	**885 441 818**	**VLP**	**No**	**Feces**	**C, F**	**Yes**	[49]

MDA: Multiple displacement amplification. VLP: Virus like particles. C: centrifugation. F: filtration. N: nucleases. CE: Centricon concentration. P: PEG precipitation. IBD: Inflammatory bowel disease

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

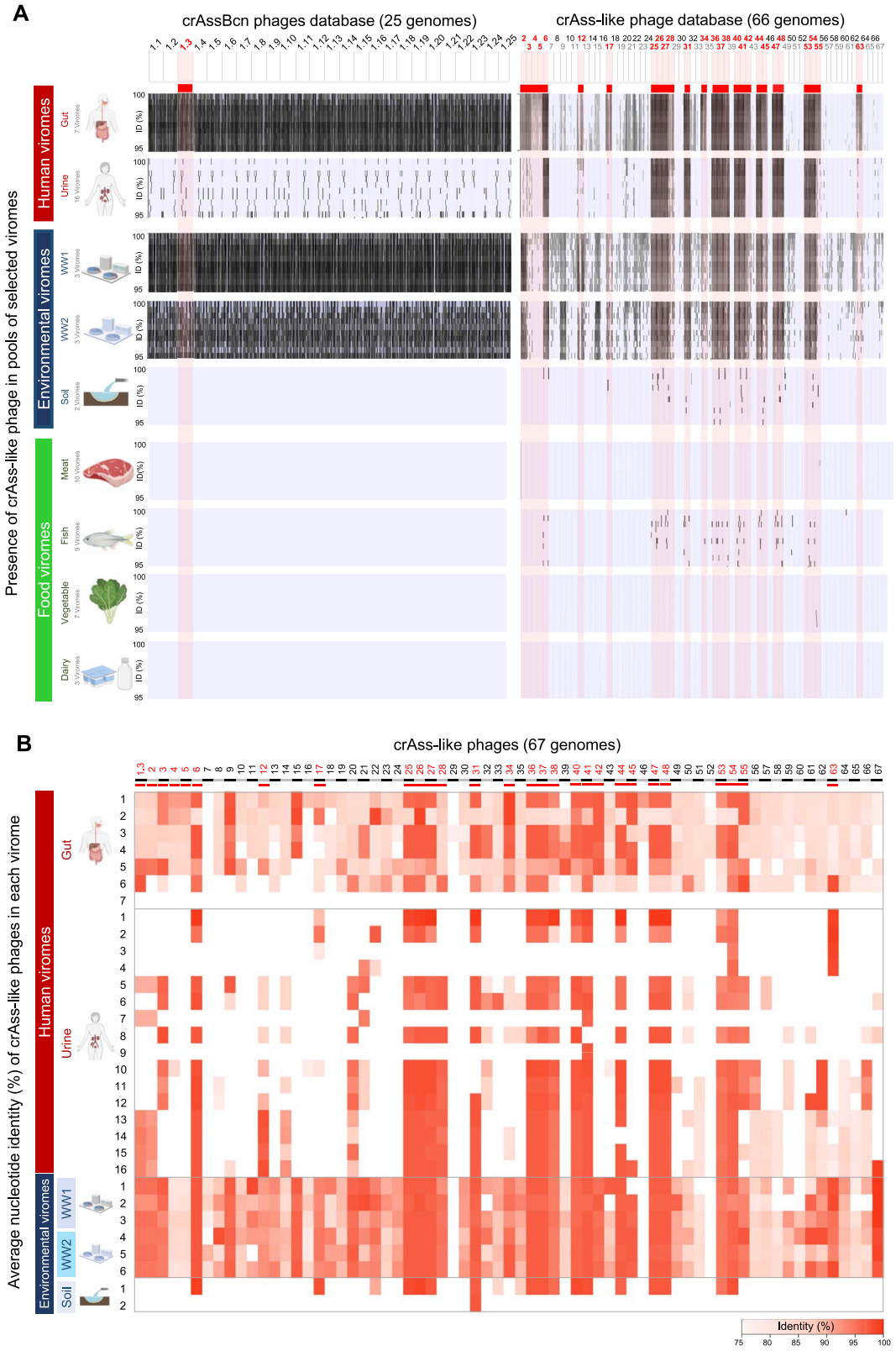


Figure 1. Presence of crAss-like phages in human-impacted viromes. (**A**) Presence of CrAss-like phage in the selected viromes. Reads from various viromes from humans (feces and urine), environmental (urban wastewater from plant 1 and 2(WW) and agricultural soil), meat pool (chicken, pig, and beef meat), fish pool (Atlantic, Mediterranean, frozen, farm fish, and shellfish), vegetables pool (spinach, lettuce, cucumber, and horchata) and dairy products pool (milk, cheese, and kefir), recruited against the genome of 25 crAssBcn phages and 66 crAss-like phages from databases (x-axis). Selected phages are indicated in the top of each chart. Recruitment plots for each virome display mapped reads and the percent identity of the reads for each phage genome are shown on the y-axis (only reads with 95 to 100% identity are shown). (**B**) Heatmap of identity of detected crAss-like phages in viromes. Average nucleotide identity of mapped reads (ANIr) from each virome was calculated for each crAss-like genome separately (only reads recruited over 70% of coverage were considered). Viromes with crAss-like phage relative abundances below 0.00001 SD/Kbp/Gbp were not showed. (see Supplementary Data 2 for numerical data used to build the Heatmap).

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

in human fecal viromes were selected for the following analysis regardless of their relative abundance in wastewater viromes.

Absence of selected crAss-like phages in animal gut viromes

The 28 crAss-like phages selected for their abundance in human feces were analyzed in a collection of 24 animal fecal viromes of chicken, pigs, cows, rabbits, and non-human primates (Supplementary Data 2). As previously, to objectively apply the exclusion criteria ("not detected in animal feces"), we selected the phages with a horizontal genome coverage of below 40% in the animal viromes. No reads of cows, rabbits, and non-human primate viromes were recruited against this set of phages, and only anecdotally in pig viromes. In contrast, consistently high recruitment against diverse crAss-like phages was obtained with two different sets of chicken viromes, one from Spain and the other from China (Fig. 2). Further analysis revealed that 15 crAss-like phages were present in both chicken viromes with more than 40% of genome coverage (Fig. 2B), and these were consequently excluded due to their lack of human-specificity. In contrast, nine crAss-like phages were present in only one of these two chicken viromes (Fig. 2B, marked by orange symbols and labeled "quite" in Table 2). As human fecal contamination of the chicken samples could not be ruled out, and we did not want to exclude phages based on the results of viromes from a single geographical area, we decided to include selected crAss-like phages with less than 40% of genome coverage in at least one of the chicken viromes. Additionally, the four highly human-specific phages not detected (below 40% of genome coverage) in any animal viromes (Fig. 2B, marked by green symbols and labeled "high" in Table 2), were selected.

In total, a list of 13 human crAss-like phages was obtained, four of which stood out as the highest human-specific (Table 2): two (DAC15 and DAC17, both *Wulfhauvirus bangladeshii* family *Steigviridae*) were isolated phages, whereas cr85 (*Kahnovirus oralis*, family *Steigviridae*) and cr17 (*Endlipuvirus intestinihominis*, family *Intestiviridae*) corresponded to uncultured phages. Other isolated phages, namely ΦCrAss001, ΦCrAss002, and ϕCrAssBcn3 (Table 2), were categorized as quite specific for humans, given that ΦCrAss001 and ΦCrAss002 were identified in a Chinese chicken virome and ϕCrAssBcn3 in a Spanish chicken virome (Fig. 2). Of the 13 human phages, four showed high recruitment in wastewater (close to or higher than 40% coverage), phage cr85 being the most recruited in wastewater viromes among the four phages with the highest specificity (Table 2).

Comparing the genomes of human-specific phages

The genomes of the 13 human- crAss-like phages were aligned to evaluate if they shared a core region. This finding would provide insights about common genomic features that could explain their specificity for humans. Moreover, the discovery of common nucleotide sequences in the human-specific phages would facilitate the development of molecular tools for their detection as human fecal markers. However, a genomic comparison by VIRIDIC analysis showed high sequence heterogeneity among the 13 human crAss-like phages. According to the International Committee for the Taxonomy of Viruses [51], the main species demarcation criterion for bacterial and archaeal viruses is currently set as a genome sequence identity of 95% over 85% of the complete genome. Applying this criterion, the 13 human crAss-like phage genomes were distributed in three families, eight genera, and 12 species (Fig. 3A), indicating that human-specificity is not restricted to a given family or species. No core genome region was

Table 2. List of the 13 selected human phages. Selection was based on their presence in samples with human fecal pollution and total or low recruitment in animal samples (coverage: > 40% in black and < 40% in red). In bold, the four phages showing the highest human specificity (High). WW: wastewater.

Phage #	Genome ID	Name		Family	Origin (City, Country)	Phage	Coverage (%)				Human specificity
		Common	Binomial				Human Gut	Urban WW	Poultry (Spain)	Poultry (China)	
1	OQ221538	ϕCrAssBcn3	*Kehishuvirus* sp.	*Steigviridae*	Barcelona, Spain	Isolated	74.0	69.0	26.2	67.8	Quite
2	NC_049977.1	ϕCrAss001	*Kehishuvirus primarius*	*Steigviridae*	Cork, Ireland	Isolated	67.3	67.1	26.4	64.7	Quite
3	MN917146.1	ϕCrAss002	*Jahgtovirus secundus*	*Intestiviridae*	Cork, Ireland	Isolated	90.3	9.8	88.2	7.0	Quite
4	**MT074136.1**	DAC15	*Wulfhauvirus bangladeshii*	*Steigviridae*	Stanford, USA	Isolated	**70.1**	**6.0**	**2.8**	**2.3**	**High**
5	**MT074138.1**	DAC17	*Wulfhauvirus bangladeshii*	*Steigviridae*	Stanford, USA	Isolated	**71.4**	**6.1**	**2.4**	**2.3**	**High**
12	MT774388.1	cr112	*Kehishuvirus splanchnicus*	*Steigviridae*	Cork, Ireland	Uncultured	49.1	51.2	23.4	48.1	Quite
17	MT774396.1	cr53	*Blohavirus americanus*	*Suoliviridae*	Cork, Ireland	Uncultured	66.7	27.8	79.2	12.3	Quite
28	MT774375.1	cr50	*Cohcovirus hiberniae*	*Suoliviridae*	Cork, Ireland	Uncultured	47.6	19.9	82.0	16.6	Quite
34	**MT774390.1**	cr85	*Kahnovirus oralis*	*Steigviridae*	Cork, Ireland	Uncultured	**98.1**	**38.9**	**24.1**	**4.6**	**High**
42	MT774401.1	cr6	*Afonbuvirus faecalis*	*Suoliviridae*	Cork, Ireland	Uncultured	90.9	24.4	96.2	13.7	Quite
45	MZ130492.1	cr12	*Afonbuvirus intestinihominis*	*Suoliviridae*	Cork, Ireland	Uncultured	91.9	25.6	98.0	16.8	Quite
55	MT774377.1	cr107	*Jahgtovirus intestinihominis*	*Intestiviridae*	Cork, Ireland	Uncultured	90.2	9.0	78.1	7.9	Quite
63	**MZ130488.1**	cr17	*Endlipuvirus intestinihominis*	*Intestiviridae*	Cork, Ireland	Uncultured	**93.5**	**7.0**	**11.6**	**0.9**	**High**

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

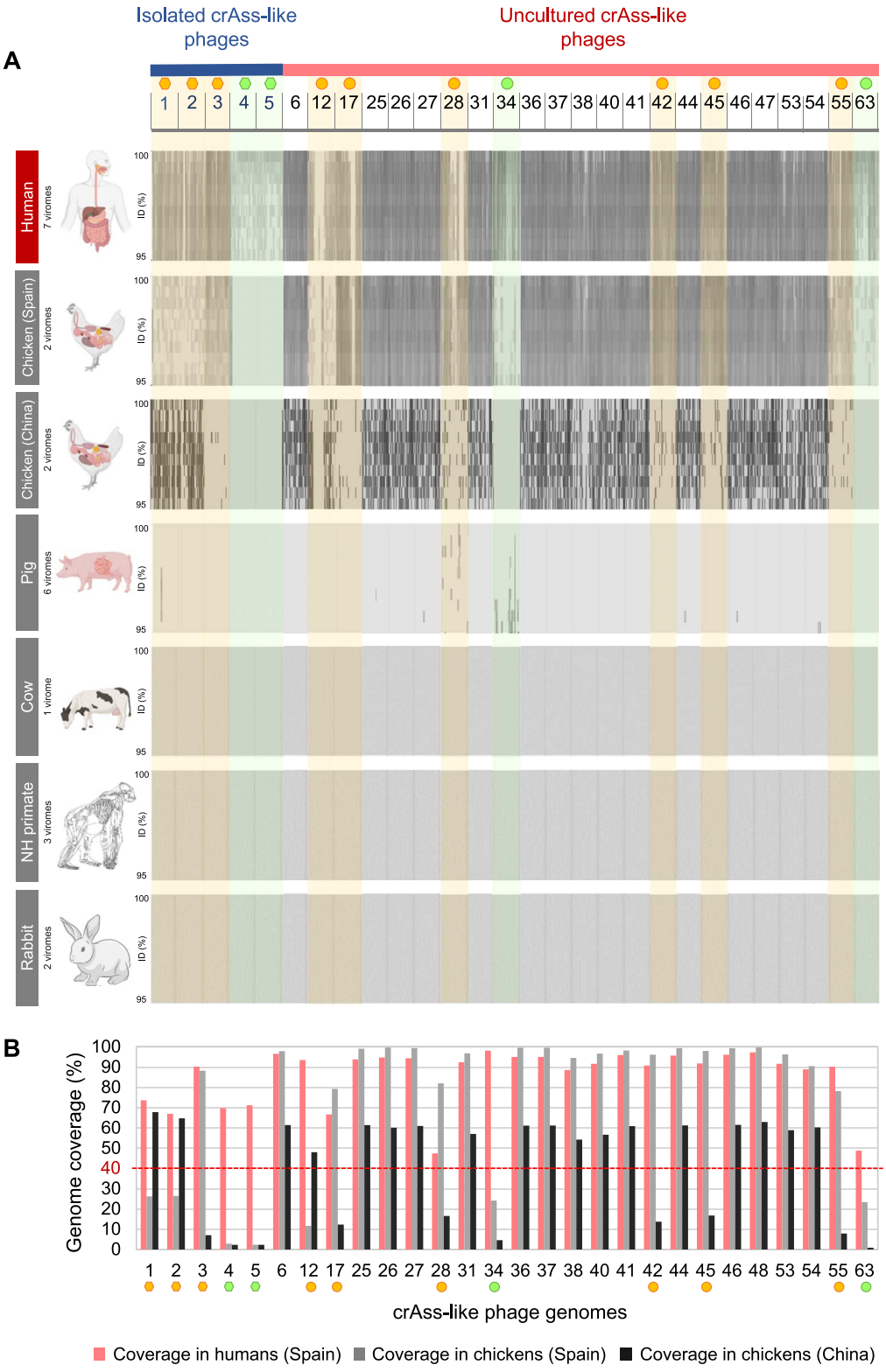


Figure 2. Recruitment plots of crAss-like phages detected in human-impacted viromes with reads in animal viromes. The datasets displayed were obtained from healthy human feces, chicken feces (from Spain and China), pig feces, cow feces, non-human primates (makaka) intestinal content, and rabbit feces (for more information see Supplementary Data 2). (**A**) the recruitment plots display the results of a BLASTn search of metagenomic short reads from animal viromes against the complete genomes of 28 selected crAss-like phages identified in human viromes (X-axis). The upper panel shows the number given to each crAss-like phage genome according to its position in the list of phages used in the study (Supplementary Data 1). Only viromes showing recruitments are included. The recruitment plots for each metagenome display mapped reads and the percent identity of the reads for each phage genome (y-axis) (only reads with 95 to 100% identity). (**B**) the chart indicates the percentage of genome coverage of the 28 selected human phages in viromes of humans, chickens from Spain, and chickens from China. Those with coverage of more than 40% in the human gut and less than 40% in both chicken viromes were selected as highly human phages (green). Those with coverage of more than 40% in the human gut and less than 40% in at least one of the two chicken viromes were considered as quite good human phages (orange). Uncultured phages are marked by a circle and isolated phages are indicated by a hexagonal symbol.

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

observed when aligning the genomes of the 13 human crAss-like phages (Fig. 3B). High identity (over 80% in nucleotides), in accordance with their taxonomical classification, was found between phages CrAssBcn3 and CrAss001 (*Kehishuvirus* sp. and *Kehishuvirus primarius* respectively, family *Steigviridae*), DAC15 and DAC17 (both *Wulfhauvirus bangladeshii*, family *Intestiviridae*), cr6 and cr12 (*Afonbuvirus faecalis* and *Afonbuvirus intestinihominis* respectively, both from family *Suoliviridae*, and CrAss002 and cr107 (*Jahgtovirus secundus* and *Jahgtovirus intestinihominis*, respectively, both from family *Intestiviridae*), but not between the rest. A lack of coincidence was also observed when comparing only the four phages highly specific for humans (Supplementary Fig. 3).

A phylogenomic comparison of all 91 crAss-like phage genomes included in the study revealed that the 13 human-specific crAss-like phages are distributed throughout the order Crassvirales (Fig. 4A). Three of the most human-specific ones are located closest to each other (all three belonging to the family *Steigviridae*), but cr17 (family *Intestiviridae*) is situated in a different branch of the tree. To better understand the phylogenomic relationships among the 13 human-specific phages, they were compared with each other and showed a distribution in three independent clusters that correspond to the three different families, *Steigviridae*, *Intestiviridae*, and *Suoliviridae*) (Fig. 4B).

The distribution detected in this analysis with the subset of 13 crAss-like phages is in accordance with the taxonomic classification. Those that belong to similar genera are closely related, whereas those of different families are only related at the order level, so no nucleotide similarities are expected.

Biogeographical distribution of the human-specific crAss-like phages

Our aim was to evaluate the geographical distribution of the 13 human crAss-like phages, and to assess whether some of them have a wider global distribution than others. This attribute was previously reported for p-crAssphage [15], using DNA genome fragments, and for crAssBcn phages [5]. Read recruitment of a viral metagenomes collection, consisting of 1260 human gut viromes from children and adults from 14 countries was conducted against the whole genomes of the 13 human crAss-like phages (Table 1; Supplementary Data 3). Among these, 77% (971) of the metagenomes, distributed in 11 countries, recruited against at least one of the 13 human phages (Fig. 5), whereas 22.9% (289 gut viromes) were below the detection limits.

The distribution of crAss-like phages varied between countries, with some showing more frequent recruitment and relative abundance than others (Fig. 5A). The country with the lowest diversity was Tanzania, with eight crAss-like phages, followed by China with nine. Mexico and Cameroon both had 12 phages, whereas the other countries were found to have all 13 crAss-like phages. No single crAss-like phage predominated in all the countries. A relatively high abundance was found for cr53 in Finland, Tanzania, USA 2, and Ireland, and for cr12 in Malawi and Mexico. Phage cr112 (*Kehishuvirus splanchnicus*) was predominant in China, and cr17 in USA 1 and Uganda. All these genomes are of uncultured phages. Regarding the isolated phages, ΦCrAss002 was the most predominant overall (the most prevalent in USA 1, Ireland, Spain, Uganda, and Malawi), whereas ΦCrAssBcn3 and ΦCrAss001 predominated in Ethiopia, Cameroon, Finland, and USA 2, and DAC15/DAC17 in China and Tanzania.

The percentage of crAss-like phages identified in viromes around the world (Fig. 5A) allows comparison between different countries. Although the metagenomic sequencing depth (SD) of all viromes was normalized by dataset size (Gbp) and genome length (Kbp), comparisons between countries had limitations. A phage may be very prevalent in one country compared to the other phages in the same country but could show a low global recruitment. To better evaluate phage prevalence globally, comparing the relative abundance of each phage across all countries is more informative (Fig. 5B). Hence, although phage cr112 predominated in China and Ethiopia, its global relative abundance was quite low. Similarly, DAC15 and DAC17 showed the lowest relative abundance overall, yet were identified in several countries. The highest recruitment of DAC15/DAC17 genomes was found in Spain, but they were not the most prevalent phages in this country. By far the highest relative abundance worldwide was shown by phage cr85 (Fig. 5B), followed by cr12, cr6, and ΦCrAss002. The most abundant isolated crAss-like phage was ΦCrAss002 (Fig. 5A).

Determining the presence and abundance of crAss-like phages in each metagenome revealed the most represented crAss-like phage for each country (Fig. 6). However, phage prevalence is not only reflected by relative abundance (Fig. 6, lower chart) but also by the recruitment of the whole genome, which is depicted by black shadowing in the recruitment plots. Plots with only a few black signals indicate that few reads were recruited. Even if the relative abundance bars are high, a phage cannot be assumed to be present in the metagenome. For example, phage cr17 showed a high abundance in USA 1 but was discarded because the reads of this virome were not recruited against the whole phage genome. Given the variations in sequencing depth among the metagenomes (Table 1), the normalized relative abundances are depicted using different Y-axis scales. The metagenome from Spain had the greatest sequencing depth, followed by Finland, Ireland, Mexico, Malawi, and USA 2 (all five being similar), then Uganda and Ethiopia, and USA-1 and Cameroon (both the same). Finally, Tanzania and China metagenomes had the lowest sequencing depths.

Phage cr53 (*Blohavirus americanus,* family *Suoliviridae*) demonstrated high abundance in five countries, including Spain, with high read recruitment, but it was not observed in Malawi, USA 1, Cameroon, Tanzania, and China. Overall, the phage with the highest relative abundance was cr85 (*Kahnovirus oralis,* family *Steigviridae*) (Fig. 5), which is explained by its high prevalence in Spain, but it was below the detection limit in Mexico, Ethiopia, USA-1, Cameroon, Tanzania, and China, and showed low recruitment in Malawi and Ireland. Phages cr6 and cr12 (both of genus *Afonbuvirus*), ranking second in global abundance (Fig. 5), exhibited good levels of recruitment in seven countries, and moderate in three countries. These two phages, which were below the detection limit in only three countries, came the closest to the ideal of universal human-specific crAss-like phage in this study. In summary, whereas certain phages showed high representativity in some countries, they are not recruited in other and there is no human-specific crAss-like phage that has spread to many of the countries or universally.

DNA regions of the prototypical p-crAssphage has been described as having a worldwide distribution [15]. Although p-crAssphage was not included in our study because it was highly recruited in chicken viromes from Spain and China (Supplementary Data 4), we were still interested in exploring the relative abundance of the whole p-crAssphage genome in different metagenomes around the world. In accordance with previous reports [15], the relative abundance of this phage was high in human viromes from Ireland, Spain, and USA 1, but lower in Finland, Malawi, and USA 2 (Supplementary Fig. 4). The entire p-crAssphage genome was below the detection limit in

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

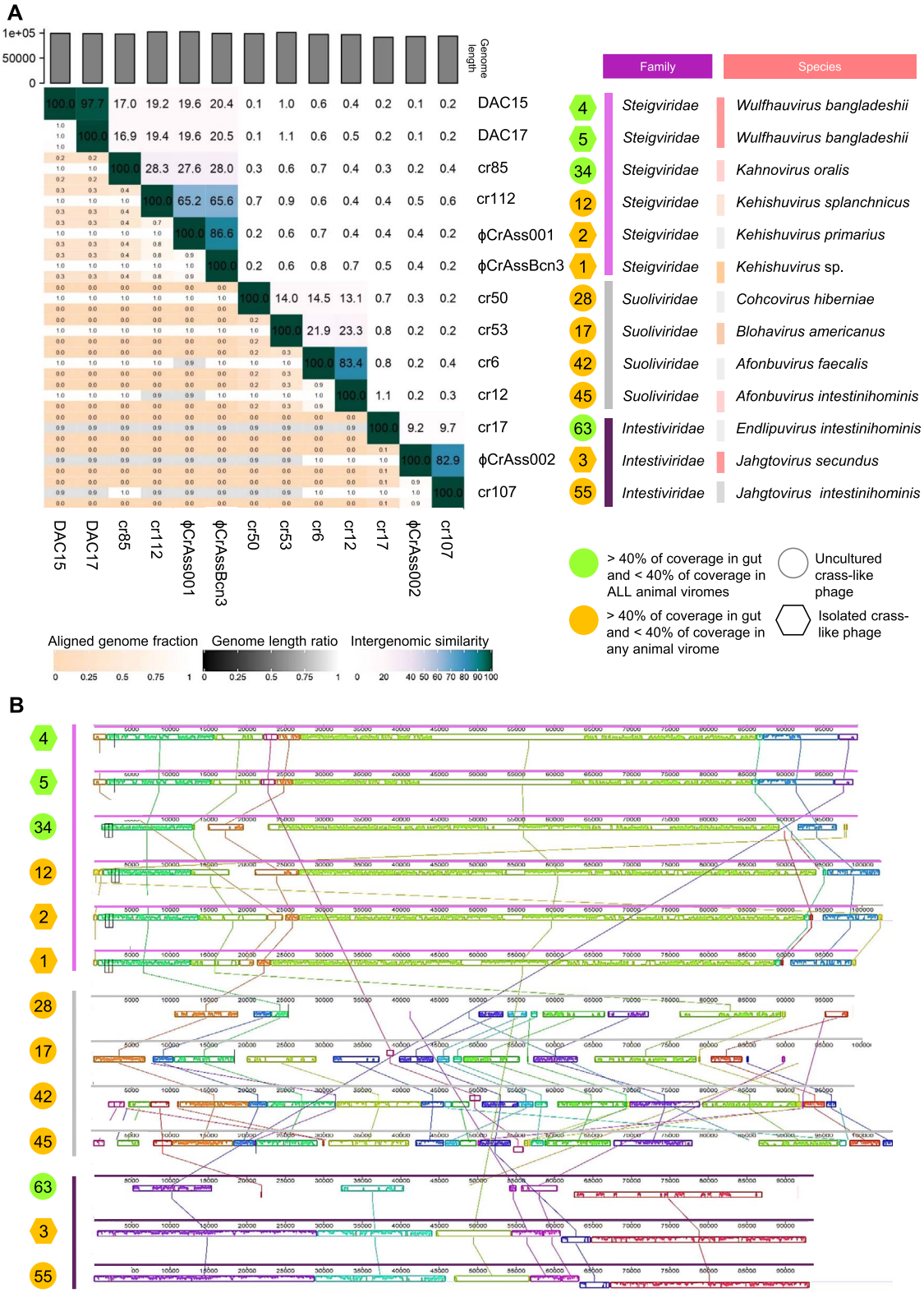


Figure 3. Intergenomic comparison of the 13 human crAss-like phages genomes. (**A**) the heatmap generated by VIRIDIC shows intergenomic similarity values (right half) and alignment indicators (left half). The percent identity between two genomes was determined by BLASTn, integrating intergenomic similarity values with data on genome lengths and aligned genome fractions. In the right half, darker colors signify closer genomic relationships. The numbers inside the map represent the identity values for each genome pair rounded to the first decimal. In the left half, aligned genome lengths are indicated, with darker colors corresponding to smaller aligned genome fractions. The family cluster and species cluster in which each phage is classified are indicated on the right side of the heatmap. (**B**) Multiple genome alignment constructed with the complete genomes of the 13 human-specific crAss-like phages. Boxes with identical colors indicate homologous DNA regions shared by two or more genomes. Uncultured phages are indicated by a circle and isolated phages by a hexagonal symbol. The number of each crAss-like phage indicates its position in the list of crAss-like phages used in the study (Supplementary Data 1).

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

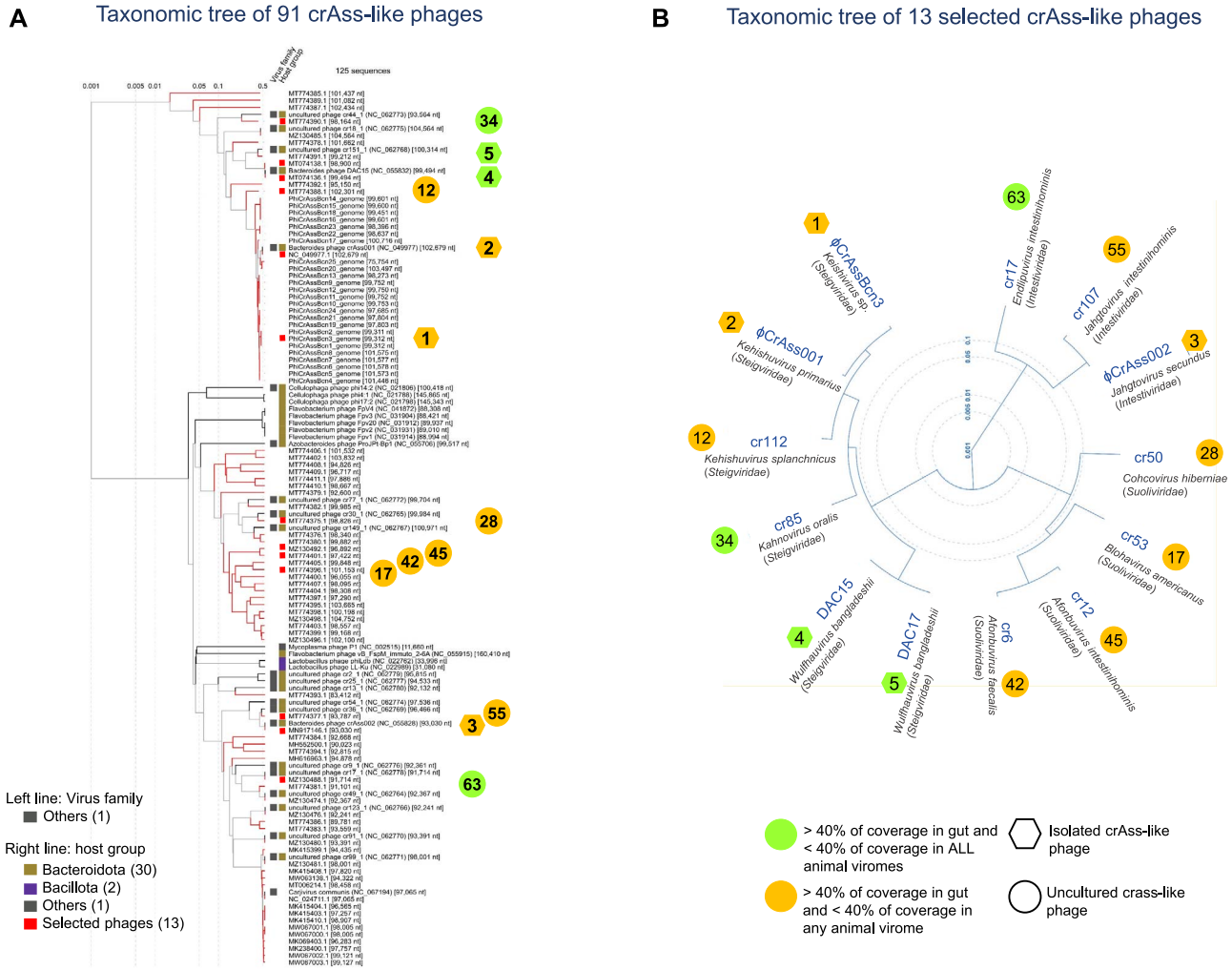


Figure 4. Taxonomical analysis and phylogenomic relationships of the 13 human crAss-like genomes. (**A**) Portion of a viral proteomic tree constructed with ViPTree, where the 13 phage genomes were included as a query and compared with crAss-like phage genomes available in the databases. Branch length is indicated at the top of the tree. (**B**) Phylogenetic tree constructed with the genomes of the 13 human crAss-like phages. Uncultured phages are indicated by a circle and isolated phages by a hexagonal symbol. The number of each crAss-like phage indicates its position in the list of crAss-like phages used in the study (Supplementary Data 1).

metagenomes from five out of 12 countries (Ethiopia, Cameroon, Tanzania, Mexico, and China).

Discussion

In 2014, a new virus named crAssphage (p-crAssphage, assigned to the species *Carjivirus communis*) was identified and defined as highly human-specific [1]. Although it has never been isolated, p-crAssphage was the first member of the order Crassvirales to be identified in human feces and the analysis of fragments of its DNA allowed to determine a wide biogeographical distribution [15], particularly in Western countries [13]. Many studies report that the molecular detection of p-crAssphage genome has proven effective in assessing the abundance of this phage in different ecosystems containing human fecal contamination [3, 52–58]. However, there has been a misconception of what is being detected when referring to crAssphage. Many authors refer to p-crAssphage as a single phage, without considering the diversity of phages within the order Crassvirales. The present study is the first to evaluate which members within the crAss-like group other than p-crAssphage are highly and exclusively present in the human gut and to explore the reasons of this specificity.

p-crAssphage and other members of the Crassvirales are highly abundant and specific to the mammalian gut, particularly in humans [4, 15, 59, 60], but not exclusively. A taxonomical analysis of Metagenome Assembled Genomes (MAGs) of crAss-like phages revealed some environmental (potentially non-human) members within the Epsilon and Zeta groups of the Crassvirales order [6]. Additionally, few previous analyses have already detected p-crAssphage in samples containing exclusively animal feces, albeit at lower concentrations than in human feces or urban sewage [17]. Its presence in animal feces has been attributed to a loss of phage specificity for the human gut and to the use of a qPCR assay targeting a non-human-specific region in the p-crAssphage genome, which might be shared by other crAss-like phages present in animals. This observation prompted us not to use genome fragments in the present study, due to their potential low specificity, and instead to analyze complete genomes of crAss-like phages in human samples, as well as to confirm their presence or absence in animal samples.

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

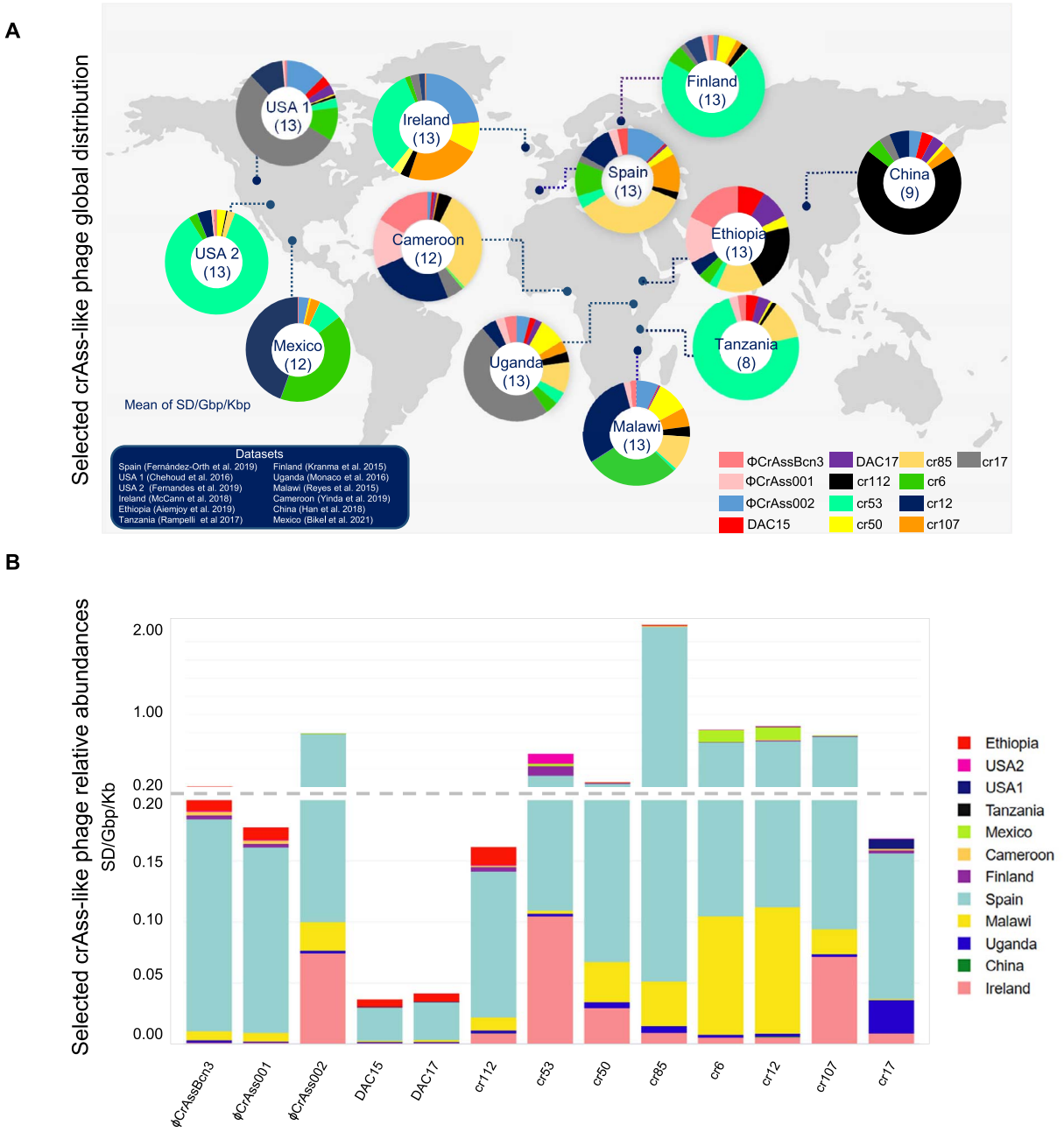


Figure 5. Geographical distribution of the 13 human crAss-like phages. (**A**) the distribution (percentage) of the 13 human crAss-like phages in each country. Only metagenomes showing recruitments are included, which are from Ethiopia, USA (USA 1 and 2), Tanzania, Mexico, Cameroon, Finland, Spain, Malawi, Ireland, Uganda, and China. The number of crAss-like phages found in each country is given in brackets. (**B**) Relative abundance of each of the 13 human phages in the different countries showing recruitments. Only mapped reads with coverage of at least 70% and best-hit reads were considered. Relative abundances of the human crAss-like phages were normalized by metagenome size (Gbp) and phage length (Kbp). World map figure in the background was obtained from Designspot/Freepik.

For recruitment analysis of the highly human-specific phages we used samples with confirmed presence of human pollution from high (human feces, wastewater, sludge) to low fecal content (environmental and food samples). Pools of each type of sample were generated with various datasets to obtain the necessary sequencing effort for nearly complete sequence coverage to detect specific phage genomes. The sequencing effort of all pools (Supplementary Dataset 2) was similar or above the size recommended [61], with the exception of soil viromes. Despite the small size of soil viromes, they were included in the study because some crAss-like phages were detected in soil virome 1 (cutoff ≥70% of query coverage) and with a relative high abundance (Supplementary Fig. 2).

Of the 91 phage genomes analyzed (67 if only one representative of the 25 crAssBcn phages is included), only 41.8% were highly abundant and diverse (ANIr between 85–100%) in the human gut, and some were either not recruited or found in very low concentrations in all viromes. Within this group of human phages, exclusive specificity for the human gut was even less common (only 5.97% showed absolute specificity and 13.4% were partially specific). It is possible that the viromes analyzed may contain new crAss-like phage sequences not yet described and, consequently,

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

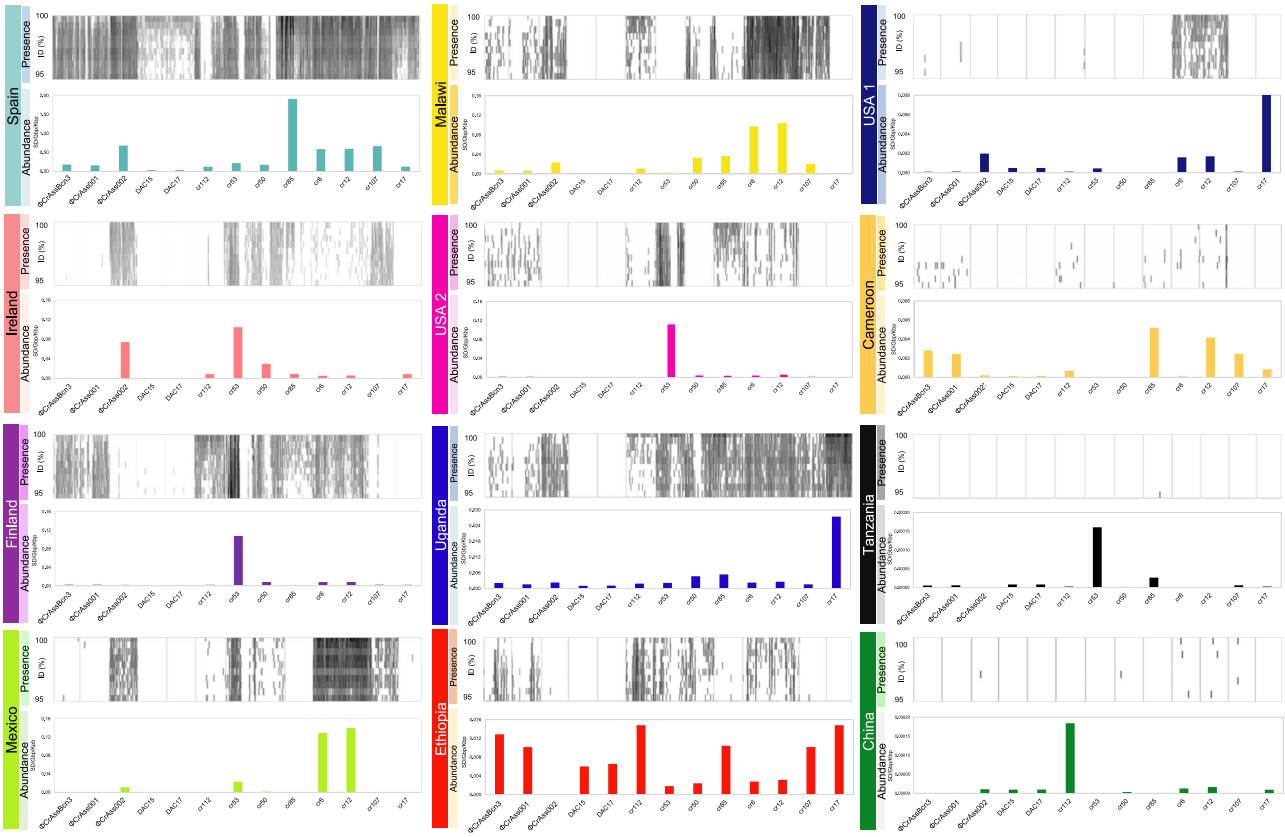


Figure 6. Selected crAss-like phage presence (over 95% ID) and abundance around the globe. Recruitment of the metagenomes from different countries against the 13 human crAss-like phages is displayed. For each country, the upper chart (presence) shows the recruitment plots with mapped reads (X-axis) and percent identity (95 to 100%) (Y-axis). The lower chart (abundance) shows the relative abundance of each of the 13 human crAss-like phage genomes (X-axis) in each country (only best-hit reads recruited ≥70% of query coverage were considered). The relative abundances of the human crAss-like phages were normalized by metagenome size (Gbp) and phage length (Kbp).

not included in our reference database. This study used only previously reported and confirmed complete crAss-like phage genomes.

Several phages found in the human feces with high human-specificity were not among the phages with the highest recruitment levels in wastewater viromes, even if all were present. Of the 13 phage genomes showing higher recruitment in humans, only four were well represented in wastewater viromes (ϕCrAssBcn3, ΦCrAss001, cr112, and cr85, all belonging to the family *Steigviridae*) (Table 2), and only cr112 and cr85, were classified as highly human-specific. In addition, none of the crAss-like phages analyzed were detected by metagenomics in samples with low fecal content, such as soil or food. Ideally, a phage present in fecal samples is expected to be found in human wastewater, containing a high fecal load, and be recruited even in samples with low levels of fecal contamination. Nevertheless, the use of metagenomics has some limitations. As the aim of this study was to evaluate which phages were the most prevalent in humans, metagenomics was a useful tool, allowing us to analyze whole phage genomes. Metagenomic analysis, however, is not suitable for application in real, complex samples (like wastewater) or in those with a low DNA content (soil or food). The high phage diversity of wastewater, reported to be higher than in feces [61–63], reduces the chances of detecting a single phage genome in a wastewater sample. In wastewater samples, as in those with low fecal inputs, the presence of crAss-like phages could be determined using more sensitive detection techniques, such as qPCR amplification. Compared to PCR-based methods, the limit of detection of metagenomics may be a couple of orders of magnitude higher, or even more, depending on DNA yields and sequencing efforts [64]. Considering this, we preferred not to exclude any of the phages showing low relative abundance in wastewater to avoid underestimation of some of the crAss-like phages analyzed and we continued our study with the 28 phages highly recruited in human fecal samples.

It should be considered that the group of 28 phages present in human samples were selected using samples with human contamination from our own geographical area. Because crAss-like phages show geographical variability, it could not be excluded that the use of human samples from other areas may have included some more candidates to the list with higher relative abundances. Even so, the Spanish metavirome has the highest relative abundance among all the countries studied and showed recruitment against all the phages analyzed. In contrast, the use of soil or food viromes from other areas would not have influenced the results because the fecal load of these samples and the limitation of metagenomics is more determinant to identify new candidates than the geographical origin of these samples.

Various animal viromes in public databases were then assessed for the absence of human-related sequences, using only those with high sequence quality (Supplementary Data 2). The animal viromes (from cows, pigs, rabbits, or non-human primates) were largely without crAss-like phages, indicating the absence of human fecal pollution and supporting the preferential residence of Crassvirales in the human gut. Conversely, various crAss-like phages were detected in two chicken viromes, from Spain and China. In these cases, the absence of human fecal contamination

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

cannot be guaranteed, particularly for the samples not analyzed in our laboratory. This finding, together with previous reports of Crassvirales detection in animal samples [16, 17, 54, 65], highlights the need for more research to assess which members of this heterogeneous order are the most exclusive inhabitants of the human gut.

Among different animals, the presence of p-crAssphage in cats could be attributed to contamination from their human companions [16], that transferred the phage or the host strain *B. intestinalis* [16]. Cross-contamination from humans is not that evident in other comparative studies that have also detected high number of crAss-like phages in chicken viromes [17, 65]. In the present work, considering the detection limit of metagenomics, the presence and relative abundance of crAss-like phages in chicken viromes was not negligible (Fig. 2). Overestimation due to cross-reactivity of the qPCR assay can be ruled out, as the whole phage genome was used rather than single DNA fragments. Consequently, in addition to the human gut, the chicken gut could be a reservoir of some crAss-like phages. This possibility is not unexpected, considering the ability of diverse *Bacteroides* species isolated from chickens to adapt to humans and vice-versa [66]. Once the bacterial host is in a new gut environment, the phages infecting this host should have a higher chance of adaptation. However, the phages showed varying prevalence in the two chicken viromes. Thus, if only one of the viromes had been studied, some of the phages would have appeared as highly human-specific. As we did not want to exclude any of the human-specific phages, and we could not verify whether any of the viromes (European or Chinese) were free of human fecal contamination, phages not recruited in one chicken virome were considered to have partial specificity for humans (Fig. 2).

To identify those crAss-like phages most represented in humans, phages that had a minimum of 40% genome coverage by metagenomic reads were considered. Similarly, in animal viromes, phage genomes with less than 40% coverage were regarded as not recruited. This cutoff recruitment value is more stringent than the previously reported 10% coverage by metagenomic reads [67]. Considering that crAss-like phages have a genome size close to 100 kb [5, 59], a single gene could comprise 10% of the genome of a crAss-like phage. This single gene may be shared among several crAss-like phages and be considered within a cutoff of 10% coverage, without being a real phage. We did not want to base our conclusions on a small fraction of the genome that could potentially confuse the results. We decided to increase the requirements for horizontal genomic coverage to consider when a phage genome is present, thereby reducing the risk of overestimating detection and obtaining false positives. In humans, the average genomic coverage was 47%, and in wastewater ~37%. For these reasons, we deemed it safer to increase to a cutoff of 40% the coverage requirement to determine the specificity of crAss-like phages. Based on these thresholds, only 13 human crAss-like phages from the initial group of 91 phage genomes analyzed were selected, whereas p-crAssphage was excluded because it failed to meet the criteria for human specificity and was highly recruited in both chicken viromes (Supplementary Data 4).

Despite the similarities between DAC15 and DAC17, and between ΦCrAssBcn3 and ΦCrAss001, as previously reported [5, 9], a comparison of the 13 human-specific crAss-like phages failed to uncover a shared "core" genome at the nucleotide level that could be correlated to their specificity for the human host. The percentage of identity among the 13 crAss-like phages was minimal, their genomes showing low synteny and very low homology. Moreover, the heterogeneity persisted when comparing only the four crAss-like phages highly specific for humans. This diversity is supported by the phylogenomic analysis of the 13 human phages, that belong to different genera or families of the Crassvirales order. These results suggest that within the Crassvirales group, human specificity is likely not a feature of a common human-resident ancestor before spreading across the globe, but rather was introduced on separate/independent occasions in evolutionary history. The phages belong to different genera and species, and are distributed in three different families (*Steigviridae*, *Intestiviridae*, and *Suoliviridae*), that have evolved independently within the order, possibly achieving human specificity through symbiotic events with their bacterial host strain within the human reservoir. This symbiotic relationship is not limited to a single bacterial species, because in the group of 13 human crAss-like phages at least three different *Bacteroides* species can be identified as hosts (*B. intestinalis, B. tethaiotaomicron*, and *B. xylanisolvens*) [7, 9, 10].

Some crAss-like phage genomes, mostly from the Zeta group, use in some parts of their genome alternative genetic codes, in which stop codons such as TGA or TAG are reassigned and encode an amino acid (tryptophan or glutamine respectively) [6, 68]. In addition to alternative genetic codes, these genomes include massively introns and inteins that seriously complicates annotation [6]. These changes in the genetic code may affect the proteins encoded and should be considered when cataloging phage gene inventories or when analyzing the biological function of the proteins. In our study, we did not annotate phages genomes, but we used the annotation of all crAss-like phage genomes available in the databases, which may have not considered putative read-through stop codons and other elements. For this reason, phage genome comparison was performed at nucleotide level (Fig. 3). The phylogenomic analysis presented in the viral proteomic tree (Fig. 4) were coherent with the comparison performed at nucleotide level.

Few studies have evaluated the spread of Crassvirales around the globe [4, 15]. Our analysis of 1260 human gut metagenomes from 14 countries detected crAss-like phages in metagenomes of 12 countries, with the phages showing varying levels of global distribution. The identification of human-specific crAss-like phages in metagenomes of different countries indicates that the members of the large and diverse Crassvirales group have spread differently across the globe [4, 6, 59], but none is globally distributed. In our study, some countries appear to have higher relative abundances of crAss-like phages than others. The depth of the available viromes influenced the relative abundance of the phages. The metagenome of Spain had the highest relative abundance, followed by Finland, Ireland, Mexico, Malawi, and the USA-2. The relative abundances were normalized by the sequencing depth of the metagenome, the dataset size, and the genome length, but even after normalization, the greater depth influenced the observed relative abundances.

When present, p-crAssphage showed very high relative abundances, higher than any of the other crAss-like phages studied. However, the whole p-crAssphage genome failed to be recruited or fell below the detection threshold in numerous countries analyzed. p-crAssphage has been proposed and extensively used as a universal marker of human fecal pollution. Recently, crAssBcn phages were also found to be highly abundant in Spain and other countries [5]. However, some of the crAss-like phages identified in the present study have a wider geographical distribution and are more represented in human gut viromes than p-crAssphage or than crAssBcn phages. The variability in the geographical

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

distribution of p-crAssphage has been attributed to factors such as human migration patterns, higher fitness, or environmental stability [15].

The Crassvirales group show high levels of bacterial host specialization [7]. Therefore, it can be assumed that the specificity of some crAss-like phages for the human gut is the result of their ability to infect human-specific *Bacteroides* strains, rather than an attribute of the phages themselves. The specificity between the *Bacteroides* group and their hosts [69] enables the isolation of phages specifically present in humans or animals by using a human-specific or an animal-specific *Bacteroides* host strains [70, 71]. The presence of *B. intestinalis* in cats [16] or *Bacteroides* spp. in chicken [66] could explain the presence of crAss-like phages in these animals.

Human-specific *Bacteroides* hosts vary in their sensitivity to different phages according to the geographic area where they have been isolated, being more susceptible to those phages from the same geographic area and of the same fecal origin [70, 72]. Because *Bacteroides* shows a strong adaptation to the gut environment [73], these geographical differences in *Bacteroides* can be attributed to various epigenetic factors affecting their human hosts, such as age, health state, and cultural traditions affecting diet. Diet modulates the gut microbiota profiles of different human populations and promote diverse symbiotic relationships between human hosts and their microorganisms [74, 75].

The emergence and dissemination of human symbiotic *Bacteroides* spp. strains may have occurred within human populations inhabiting distinct geographical areas. As a result, human-specific crAss-like phages infecting these strains would have soon emerged. These events might have occurred stochastically and simultaneously in various human populations occupying diverse ecological niches worldwide, thus contributing to the varied geographical distribution of the different human-specific crAss-like phages and the absence of a phylogenetic relationship between them, as observed in this study.

To properly study crAss-like phages and their biological cycle, it is necessary to use isolated virions. Using an isolated phage is desirable, firstly because its genome is not a product of incorrect assembly, and secondly, its ecological distribution and persistence, host-phage relationship, and biological cycle can be assessed. Thus, the isolated phage can be used as a control in experiments to evaluate analytical performance, and their persistence under different treatments [76, 77]. In the group of human crAss-like phages, ΦCrAssBcn3 (*Kehishuvirus* sp.), ΦCrAss001 (*Kehishuvirus primarius*), ΦCrAss002 (*Jahgtovirus secundus*), DAC15, and DAC17 (both *Wulfhauvirus bangladeshii*) are isolated phages with different *Bacteroides* strains as hosts [5, 10]. Among them, DAC15 and DAC17 were the most human-specific, but the least abundant globally. ΦCrAss002 was the most recruited overall but it was not detected in all the countries, and although it was quite specific for humans, it was also found in chicken.

Soon after its discovery, p-crAssphage was proposed as a marker of human fecal contamination. Tracing the origin of fecal pollution (named "Microbial Source Tracking" or MST [14, 72]) is particularly useful for the rapid prediction of health risks associated with exposure to fecal pollution [78], and consequently the biotechnological application of this group of phages is of pivotal relevance. p-crAssphage and other crAss-like phages have been recognized as useful markers to distinguish between human and animal fecal samples [17, 54, 55, 79]. However, our results raise questions about the human specificity and global distribution of p-crAssphage and consequently its use as a universal human marker, particularly if applied without considering its limitations, and advocate for the investigation of alternative candidates. Regrettably, finding a universal, highly human-specific crAss-like phage seems incompatible with the characteristics of this group of phages. In this scenario, care should be taken to avoid amassing data on crAssphage without assessing human specificity in different regions. Otherwise, there is a risk of selecting an unsuitable marker lacking correlation with human fecal microorganisms, ultimately undermining the potential biological applications of this novel human-specific indicator.

After analyzing 91 crAss-like phage genomes, our aim was to identify those found in the human gut, ideally absent in animal feces, and to elucidate whether common traits of the phages are determinants for their human specificity. At this juncture, this work provides a selection of human crAss-like phages and their geographical distribution. The vast majority of the crAss-like phages in the original list were discarded, leaving a small group of unrelated crAss-like phages, confirming that, unlike previously thought, most phages within Crassvirales are not exclusively found in human viromes.

Our conclusions apply to the group of phages analyzed in this study, and if new crAss-like phage genomes become available in the future, they could lead to different outcomes. Searching for new crAss-like phages in the analyzed viromes could help retrieving sequences of new divergent members of the Crassvirales order [80]. However, identification of new crAss-like phages was not the goal of this study, but to determine the human specificity of the ones already described. In addition, it should be considered the difficulty in reconstructing complete genomes of viruses, such as crAss-like phages, from viromes, as the genomic diversity present in this group of viruses complicates their recovery through metagenomic assembly [81]. It has been reported that analysis of bulk metagenomes may provide an advantage for viral detection in comparison to Virus-like particle (VLP)-enriched metagenomes [80]. However, the abundance of cellular sequences in a bulk metagenome may hinder the detection of viral sequences. In our analysis, most bulk metagenomes did not show recruitment against the selected phages (Table 1), whereas hits were obtained using VLP-enriched metagenomes.

With the group of phages analyzed here, our results suggest that the few human crAss-like phages do not derive from a common human-specific ancestor, but they appear in different lineages within the Crassvirales group. In addition, they do not share common genetic traits that could be attributed to their specific residence in the human gut. The lack of common genetic sequences between the 13 human phages difficults the design of molecular tools for their detection, which would be very useful for their biotechnological application as human fecal marker. There is not a uniform global distribution of the human crAss-like phages either, although some of them are more globally widespread than others and they seem to have emerged simultaneously in different geographical areas. Our results indicate that the presence of human-specific crAss-like phages is possibly depending on the residence of their bacterial host strain, *Bacteroides*, that shows a close symbiotic relationship with the different human populations.

Acknowledgements

This research was funded by the Spanish Ministerio de Ciencia, Innovación y Universidades (MICIU). (PID2020-113355GB-I00); the Agencia Estatal de Investigación (AEI) and the European regional fund (ERF), and by the grant CNS2022-135352 funded

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

by MICIU/AEI/10.13039/501100011033 and by the "European Union NextGenerationEU/PRTR". C. G-G has a fellowship from the University of Barcelona. MD.R.B has a Margarita Salas fellowship from the Universitat d' Alacant. L. S-C has a Beatriu de Pinos fellowship, Generalitat de Catalunya. G. V has a FPI grant from the Spanish Ministerio de Universidades. L. R-R. is a Serra-Hunter professor, Generalitat de Catalunya. M. M. is a researcher of the ICREA Academia 2022 program.

Author contributions

M.D. R-B conducted all bioinformatics analyses, prepared figures, and contributed to writing and revising the paper. C. G-G isolated the phage fraction and generated some of the viromes. L. S-C calculated statistical analysis, prepared figures, and contributed to writing the paper. G. V performed the sampling, prepared the DNA from the viromes, and analyzed microbial indicators. L. R-R wrote and revised the paper, and discussed results. M. M. is the corresponding author, supervised and funded the study, designed experiments, analyzed data, and wrote the paper.

Supplementary material

Supplementary material is available at *The ISME Journal* online.

Conflicts of interest

The authors report there are no competing interests to declare.

Funding

None declared.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

References

1. Dutilh BE, Cassman N, McNair K *et al.* A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* 2014;**5**:4498. https://doi.org/10.1038/ncomms5498
2. Stachler E, Akyon B, De Carvalho NA *et al.* Correlation of crAssphage qPCR markers with culturable and molecular indicators of human fecal pollution in an impacted urban watershed. *Environ Sci Technol* 2018;**52**:7505–12. https://doi.org/10.1021/acs.est.8b00638
3. Sabar MA, Honda R, Haramoto E. CrAssphage as an indicator of human-fecal contamination in water environment and virus reduction in wastewater treatment. *Water Res* 2022;**221**:118827. https://doi.org/10.1016/j.watres.2022.118827
4. Guerin E, Shkoporov A, Stockdale SR *et al.* Biology and taxonomy of crAss-like bacteriophages, the most abundant virus in the human gut. *Cell Host Microbe* 2018;**24**:653–664.e6. https://doi.org/10.1016/j.chom.2018.10.002
5. Ramos-Barbero MD, Gómez-Gómez C, Sala-Comorera L *et al.* Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution. *Nat Commun* 2023;**14**:4295. https://doi.org/10.1038/s41467-023-40098-z
6. Yutin N, Benler S, Shmakov SA *et al.* Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nat Commun* 2021;**12**:1044. https://doi.org/10.1038/s41467-021-21350-w
7. Shkoporov AN, Khokhlova EV, Fitzgerald CB *et al.* ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides intestinalis*. *Nat Commun* 2018;**9**:4781. https://doi.org/10.1038/s41467-018-07225-7
8. Turner D, Shkoporov AN, Lood C *et al.* Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch Virol* 2023;**168**:74. https://doi.org/10.1007/s00705-022-05694-2
9. Hryckowian AJ, Merrill BD, Porter NT *et al.* *Bacteroides thetaiotaomicron*-infecting bacteriophage isolates inform sequence-based host range predictions. *Cell Host Microbe* 2020;**28**:371–379.e5. https://doi.org/10.1016/j.chom.2020.06.011
10. Guerin E, Shkoporov AN, Stockdale SR *et al.* Isolation and characterisation of ΦcrAss002, a crAss-like phage from the human gut that infects *Bacteroides xylanisolvens*. *Microbiome* 2021;**9**:89. https://doi.org/10.1186/s40168-021-01036-7
11. Papudeshi B, Vega AA, Souza C *et al.* Host interactions of novel Crassvirales species belonging to multiple families infecting bacterial host, *Bacteroides cellulosilyticus* WH2. *Microb Genom* 2023;**9**:001100. https://doi.org/10.1099/mgen.0.001100
12. Holmfeldt K, Solonenko N, Shah M *et al.* Twelve previously unknown phage genera are ubiquitous in global oceans. *Proc Natl Acad Sci USA* 2013;**110**:12798–803. https://doi.org/10.1073/pnas.1305956110
13. Smith L, Goldobina E, Govi B *et al.* Bacteriophages of the order Crassvirales: what do we currently know about this keystone component of the human gut virome? *Biomol Ther* 2023;**13**:584. https://doi.org/10.3390/biom13040584
14. Stachler E, Bibby K. Metagenomic evaluation of the highly abundant human gut bacteriophage CrAssphage for source tracking of human fecal pollution. *Environ Sci Technol Lett* 2014;**1**:405–9. https://doi.org/10.1021/ez500266s
15. Edwards RA, Vega AA, Norman HM *et al.* Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat Microbiol* 2019;**4**:1727–36. https://doi.org/10.1038/s41564-019-0494-6
16. Li Y, Gordon E, Shean RC *et al.* CrAssphage and its bacterial host in cat feces. *Sci Rep* 2021;**11**:815. https://doi.org/10.1038/s41598-020-80076-9
17. García-Aljaro C, Ballesté E, Muniesa M *et al.* Determination of crAssphage in water samples and applicability for tracking human faecal pollution. *Microb Biotechnol* 2017;**10**:1775–80. https://doi.org/10.1111/1751-7915.12841
18. Blanco-Picazo P, Gómez-Gómez C, Tormo M *et al.* Prevalence of bacterial genes in the phage fraction of food viromes. *Food Res Int* 2022;**156**:111342. https://doi.org/10.1016/j.foodres.2022.111342
19. Blanco-Picazo P, Fernández-Orth D, Brown-Jaque M *et al.* Unravelling the consequences of the bacteriophages in human samples. *Sci Rep* 2020;**10**:6737. https://doi.org/10.1038/s41598-020-63432-7
20. Fernández-Orth D, Miró E, Brown-Jaque M *et al.* Faecal phageome of healthy individuals: presence of antibiotic resistance genes and variations caused by ciprofloxacin treatment. *J Antimicrob Chemother* 2019;**74**:854–64. https://doi.org/10.1093/jac/dky540
21. Blanco-Picazo P, Gómez-Gómez C, Aguiló-Castillo S *et al.* Chicken liver is a potential reservoir of bacteriophages and phage-derived particles containing antibiotic resistance genes.

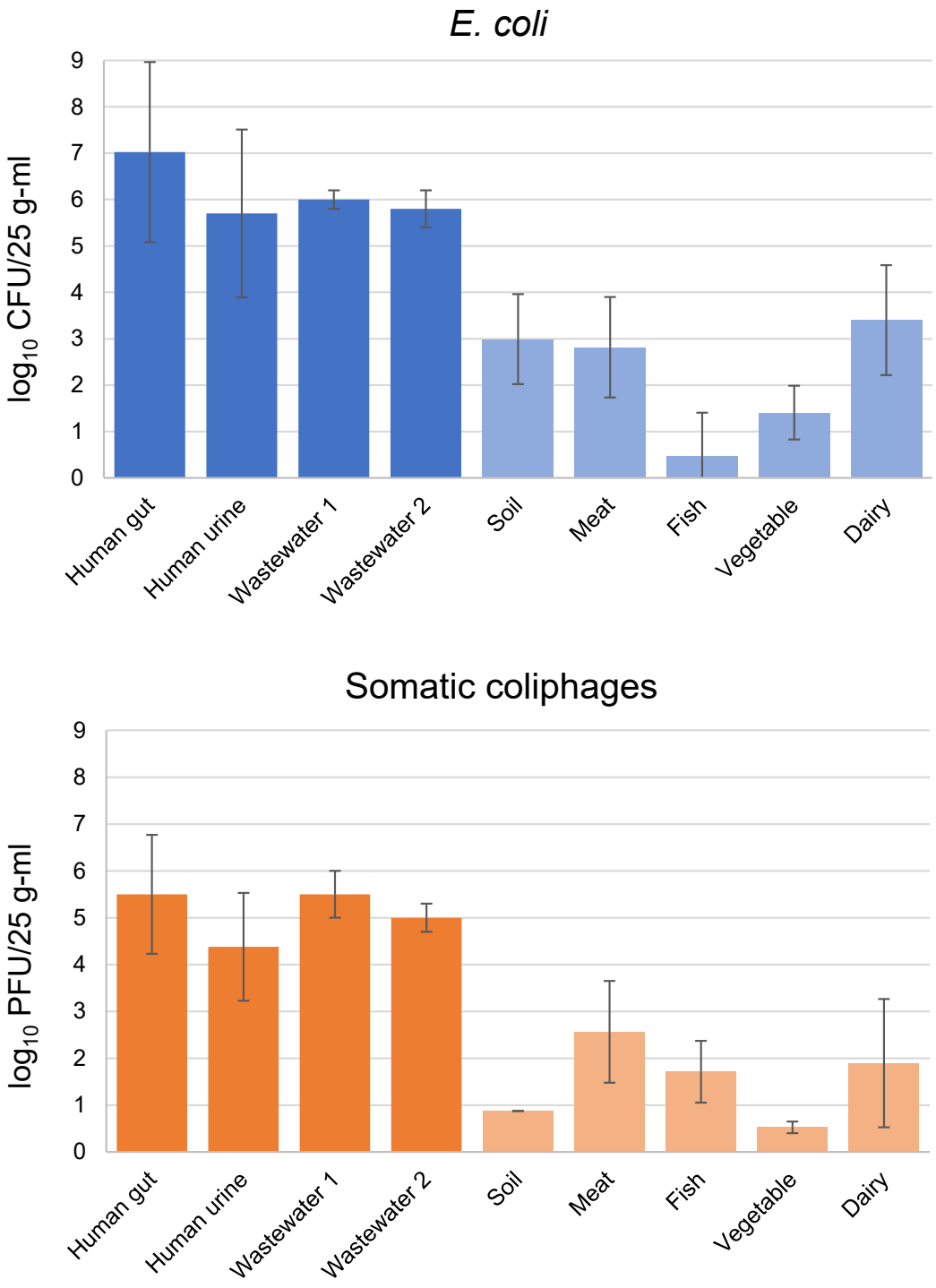
Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

Microb Biotechnol 2022;**15**:2464–75. https://doi.org/10.1111/1751-7915.14056
22. Lasobras J, Dellunde J, Jofre J *et al.* Occurrence and levels of phages proposed as surrogate indicators of enteric viruses in different types of sludges. *J Appl Microbiol* 1999;**86**:723–9. https://doi.org/10.1046/j.1365-2672.1999.00722.x
23. Jofre J. Indicators of waterborne enteric viruses. In: Bosch A (ed). *Human Viruses in Water*. Amsterdam, Netherlands: Elsevier, 2007, pp. 227–49. https://doi.org/10.1016/S0168-7069(07)17011-7.
24. Anonymous. *ISO 10705-2: Water Quality. Detection and Enumeration of Bacteriophages -Part 2: Enumeration of Somatic Coliphages*. Geneva: International Organisation for Standardisation, 2000.
25. Sambrook J, Russell D. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001, 999.
26. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20. https://doi.org/10.1093/bioinformatics/btu170
27. Wingett S, Andrews A. FastQ screen: a tool for multi-genome mapping and quality control. *F1000Res* 2018;**7**:1338. https://doi.org/10.12688/f1000research.15931.1
28. Moraru C, Varsani A, Kropinski AM. VIRIDIC-A novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* 2020;**12**:1268. https://doi.org/10.3390/v12111268
29. Rodriguez-R LM, Konstantinidis KT. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Prepr* 2016;**4**:e1900v1.
30. Plotly – Collaborative Data Science – 2015 Data Storytelling Studio @ MIT. https://datastudio2015.datatherapy.org/index.html%3Fp=745.html. Accessed 22 Mar 2022.
31. Meziti A, Tsementzi D, Rodriguez-R LM *et al.* Quantifying the changes in genetic diversity within sequence-discrete bacterial populations across a spatial and temporal riverine gradient. *ISME J* 2019;**13**:767–79. https://doi.org/10.1038/s41396-018-0307-6
32. Babicki S, Arndt D, Marcu A *et al.* Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* 2016;**44**:W147–53. https://doi.org/10.1093/nar/gkw419
33. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006;**22**:1658–9. https://doi.org/10.1093/bioinformatics/btl158
34. Adriaenssens EM, Rodney BJ. How to name and classify your phage: an informal guide. *Viruses* 2017;**9**:70. https://doi.org/10.3390/v9040070
35. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;**5**:e11147. https://doi.org/10.1371/journal.pone.0011147
36. Aiemjoy K, Altan E, Aragie S *et al.* Viral species richness and composition in young children with loose or watery stool in Ethiopia. *BMC Infect Dis* 2019;**19**:53. https://doi.org/10.1186/s12879-019-3674-3
37. Chehoud C, Dryga A, Hwang Y *et al.* Transfer of viral communities between human individuals during fecal microbiota transplantation. *MBio* 2016;**7**:e00322. https://doi.org/10.1128/mBio.00322-16
38. Fernandes MA, Verstraete SG, Phan T *et al.* Enteric virome and bacterial microbiota in children with ulcerative colitis and Crohn disease. *J Pediatr Gastroenterol Nutr* 2019;**68**:30–6. https://doi.org/10.1097/MPG.0000000000002140
39. Han M, Yang P, Zhong C *et al.* The human gut virome in hypertension. *Front Microbiol* 2018;**9**:3150. https://doi.org/10.3389/fmicb.2018.03150
40. Kramná L, Kolářová K, Oikarinen S *et al.* Gut virome sequencing in children with early islet autoimmunity. *Diabetes Care* 2015;**38**:930–3. https://doi.org/10.2337/dc14-2490
41. McCann A, Ryan FJ, Stockdale SR *et al.* Viromes of one year old infants reveal the impact of birth mode on microbiome diversity. *PeerJ* 2018;**6**:e4694. https://doi.org/10.7717/peerj.4694
42. Monaco CL, Gootenberg DB, Zhao G *et al.* Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. *Cell Host Microbe* 2016;**19**:311–22. https://doi.org/10.1016/j.chom.2016.02.011
43. Pérez-Brocal V, García-López R, Vázquez-Castellanos JF *et al.* Study of the viral and microbial communities associated with Crohn's disease: a metagenomic approach. *Clin Transl Gastroenterol* 2013;**4**:e36. https://doi.org/10.1038/ctg.2013.9
44. Rampelli S, Turroni S, Schnorr SL *et al.* Characterization of the human DNA gut virome across populations with different subsistence strategies and geographical origin. *Environ Microbiol* 2017;**19**:4728–35. https://doi.org/10.1111/1462-2920.13938
45. Reyes A, Blanton LV, Cao S *et al.* Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. *Proc Natl Acad Sci USA* 2015;**112**:11941–6. https://doi.org/10.1073/pnas.1514285112
46. Shkoporov AN, Ryan FJ, Draper LA *et al.* Reproducible protocols for metagenomic analysis of human faecal phageomes. *Microbiome* 2018;**6**:68. https://doi.org/10.1186/s40168-018-0446-z
47. Yinda CK, Vanhulle E, Conceição-Neto N *et al.* Gut Virome analysis of Cameroonians reveals high diversity of enteric viruses, including potential interspecies transmitted viruses. *mSphere* 2019;**4**:e00585. https://doi.org/10.1128/mSphere.00585-18
48. Zhao G, Vatanen T, Droit L *et al.* Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children. *Proc Natl Acad Sci USA* 2017;**114**:E6166–75. https://doi.org/10.1073/pnas.1706359114
49. Bikel S, López-Leal G, Cornejo-Granados F *et al.* Gut dsDNA virome shows diversity and richness alterations associated with childhood obesity and metabolic syndrome. *iScience* 2021;**24**:102900. https://doi.org/10.1016/j.isci.2021.102900
50. Blanco-Picazo P, Morales-Cortes S, Ramos-Barbero MD *et al.* Dominance of phage particles carrying antibiotic resistance genes in the viromes of retail food sources. *ISME J* 2023;**17**:195–203. https://doi.org/10.1038/s41396-022-01338-0
51. Lefkowitz EJ, Dempsey DM, Hendrickson RC *et al.* Virus taxonomy: the database of the international committee on taxonomy of viruses (ICTV). *Nucleic Acids Res* 2018;**46**:D708–17. https://doi.org/10.1093/nar/gkx932
52. Ballesté E, Pascual-Benito M, Martín-Díaz J *et al.* Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. *Water Res* 2019;**155**:233–44. https://doi.org/10.1016/j.watres.2019.02.042
53. Park GW, Ng TFF, Freeland AL *et al.* CrAssphage as a novel tool to detect human fecal contamination on environmental surfaces and hands. *Emerg Infect Dis* 2020;**26**:1731–9. https://doi.org/10.3201/eid2608.200346
54. Ahmed W, Payyappat S, Cassidy M *et al.* Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. *Water Res* 2018;**145**:769–78. https://doi.org/10.1016/j.watres.2018.08.049
55. Sala-Comorera L, Reynolds LJ, Martin NA *et al.* crAssphage as a human molecular marker to evaluate temporal and spatial variability in faecal contamination of urban marine

Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024

bathing waters. *Sci Total Environ* 2021;**789**:147828. https://doi.org/10.1016/j.scitotenv.2021.147828
56. Martin NA, Sala-Comorera L, Gao G *et al.* Inclusion of hydrodynamic properties of bathing waters is critical in selecting faecal indicators to assess public health impacts of faecal contamination. *Water Res* 2023;**242**:120137. https://doi.org/10.1016/j.watres.2023.120137
57. Meuchi Y, Nakada M, Kuroda K *et al.* Applicability of F-specific bacteriophage subgroups, PMMoV and crAssphage as indicators of source specific fecal contamination and viral inactivation in rivers in Japan. *PLoS One* 2023;**18**:e0288454. https://doi.org/10.1371/journal.pone.0288454
58. Chettleburgh C, Ma SX, Swinwood-Sky M *et al.* Evaluation of four human-associated fecal biomarkers in wastewater in southern Ontario. *Sci Total Environ* 2023;**904**:166542. https://doi.org/10.1016/j.scitotenv.2023.166542
59. Koonin EV, Yutin N. The crAss-like phage group: how metagenomics reshaped the human virome. *Trends Microbiol* 2020;**28**:349–59. https://doi.org/10.1016/j.tim.2020.01.010
60. Yutin N, Makarova KS, Gussow AB *et al.* Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat Microbiol* 2018;**3**:38–46. https://doi.org/10.1038/s41564-017-0053-y
61. Rodriguez-R LM, Konstantinidis KT. Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. *Bioinformatics* 2014;**30**:629–35. https://doi.org/10.1093/bioinformatics/btt584
62. Newton RJ, McLellan SL, Dila DK *et al.* Sewage reflects the microbiomes of human populations. *MBio* 2015;**6**:e02574–14. https://doi.org/10.1128/mBio.02574-14
63. Zaheer R, Lakin SM, Polo RO *et al.* Comparative diversity of microbiomes and resistomes in beef feedlots, downstream environments and urban sewage influent. *BMC Microbiol* 2019;**19**:197. https://doi.org/10.1186/s12866-019-1548-x
64. Lindner BG, Gerhardt K, Feistel DJ *et al.* A user's guide to the bioinformatic analysis of shotgun metagenomic sequence data for bacterial pathogen detection. *Int J Food Microbiol* 2024;**410**:110488. https://doi.org/10.1016/j.ijfoodmicro.2023.110488
65. Ahmed W, Lobos A, Senkbeil J *et al.* Evaluation of the novel crAssphage marker for sewage pollution tracking in storm drain outfalls in Tampa. *Florida Water Res* 2018;**131**:142–50. https://doi.org/10.1016/j.watres.2017.12.011
66. Kollarcikova M, Faldynova M, Matiasovicova J *et al.* Different *Bacteroides* species colonise human and chicken intestinal tract. *Microorganisms* 2020;**8**:1483. https://doi.org/10.3390/microorganisms8101483
67. Castro JC, Rodriguez LM, Harvey WT *et al.* imGLAD: accurate detection and quantification of target organisms in metagenomes. *PeerJ* 2018;**6**:e5882. https://doi.org/10.7717/peerj.5882
68. Peters SL, Borges AL, Giannone RJ *et al.* Experimental validation that human microbiome phages use alternative genetic coding. *Nat Commun* 2022;**13**:5710. https://doi.org/10.1038/s41467-022-32979-6
69. Xu J, Bjursell MK, Himrod J *et al.* A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* 1979;**299**:2074–6. https://doi.org/10.1126/science.1080029
70. Payan A, Ebdon J, Taylor H *et al.* Method for isolation of *Bacteroides* bacteriophage host strains suitable for tracking sources of fecal pollution in water. *Appl Environ Microbiol* 2005;**71**:5659–62. https://doi.org/10.1128/AEM.71.9.5659-5662.2005
71. Gómez-Doñate M, Payán A, Cortés I *et al.* Isolation of bacteriophage host strains of *Bacteroides* species suitable for tracking sources of animal faecal pollution in water. *Environ Microbiol* 2011;**13**:1622–31. https://doi.org/10.1111/j.1462-2920.2011.02474.x
72. Jofre J, Blanch AR, Lucena F *et al.* Bacteriophages infecting *Bacteroides* as a marker for microbial source tracking. *Water Res* 2014;**55**:1–11. https://doi.org/10.1016/j.watres.2014.02.006
73. Wexler AG, Goodman AL. An insider's perspective: *Bacteroides* as a window into the microbiome. *Nat Microbiol* 2017;**2**:17026. https://doi.org/10.1038/nmicrobiol.2017.26
74. Wu M, McNulty NP, Rodionov DA *et al.* Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut *Bacteroides*. *Science* 2015;**350**:aac5992. https://doi.org/10.1126/science.aac5992
75. Sonnenburg ED, Zheng H, Joglekar P *et al.* Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations. *Cell* 2010;**141**:1241–52. https://doi.org/10.1016/j.cell.2010.05.005
76. Pascual-Benito M, García-Aljaro C, Casanovas-Massana S *et al.* Effect of hygienization treatment on the recovery and/or regrowth of microbial indicators in sewage sludge. *J Appl Microbiol* 2015;**118**:412–8. https://doi.org/10.1111/jam.12708
77. Sommer R, Lhotsky M, Haider T *et al.* UV inactivation, liquid-holding recovery, and photoreactivation of *Escherichia coli* O157 and other pathogenic *Escherichia coli* strains in water. *J Food Prot* 2000;**63**:1015–20. https://doi.org/10.4315/0362-028X-63.8.1015
78. Harwood VJ, Staley C, Badgley BD *et al.* Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiol Rev* 2014;**38**:1–40. https://doi.org/10.1111/1574-6976.12031
79. Stachler E, Kelty C, Sivaganesan M *et al.* Quantitative CrAssphage PCR assays for human fecal pollution measurement. *Environ Sci Technol* 2017;**51**:9146–54. https://doi.org/10.1021/acs.est.7b02703
80. Gregory AC, Zablocki O, Zayed AA *et al.* The gut virome database reveals age-dependent patterns of virome diversity in the human gut. *Cell Host Microbe* 2020;**28**:724–740.e8. https://doi.org/10.1016/j.chom.2020.08.003
81. Martinez-Hernandez F, Fornas O, Lluesma Gomez M *et al.* Single-virus genomics reveals hidden cosmopolitan and abundant viruses. *Nat Commun* 2017;**8**:15892. https://doi.org/10.1038/ncomms15892

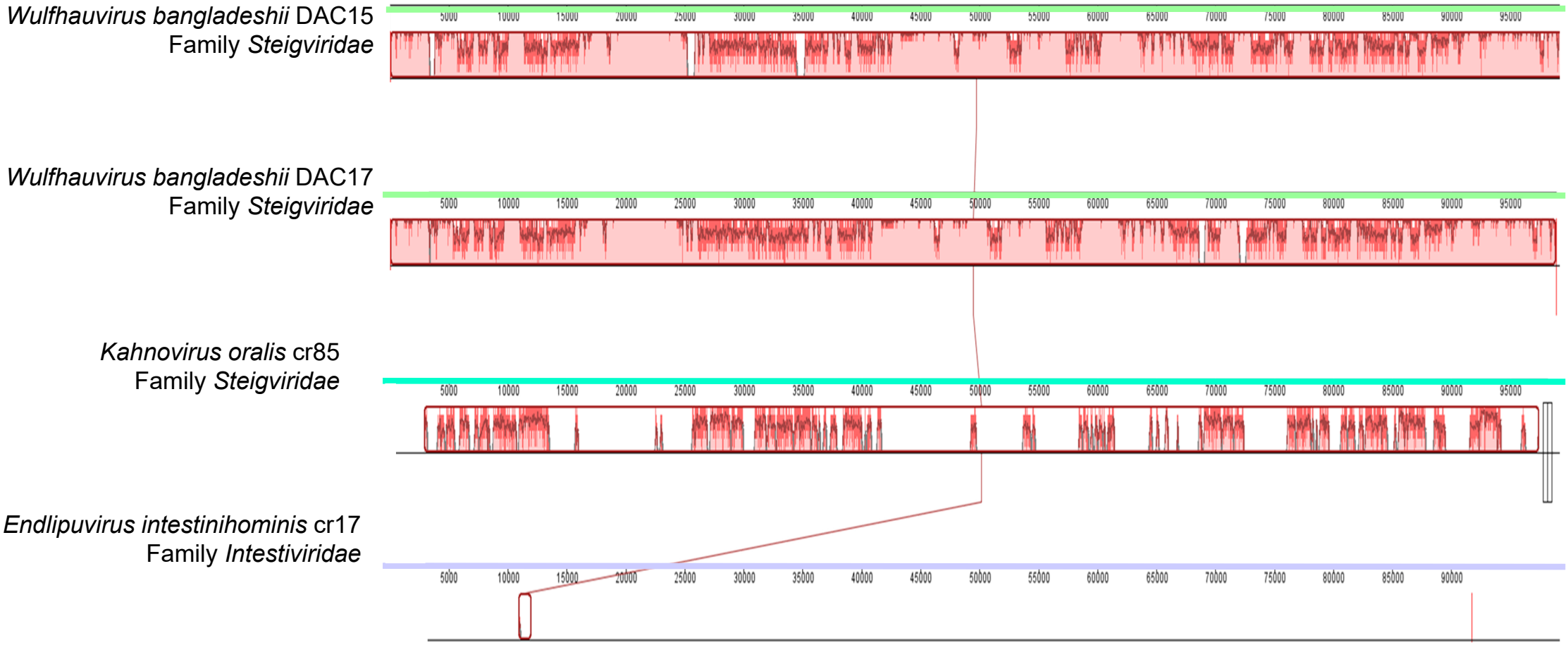
Downloaded from https://academic.oup.com/ismej/article/18/1/wrae192/7810224 by guest on 17 October 2024



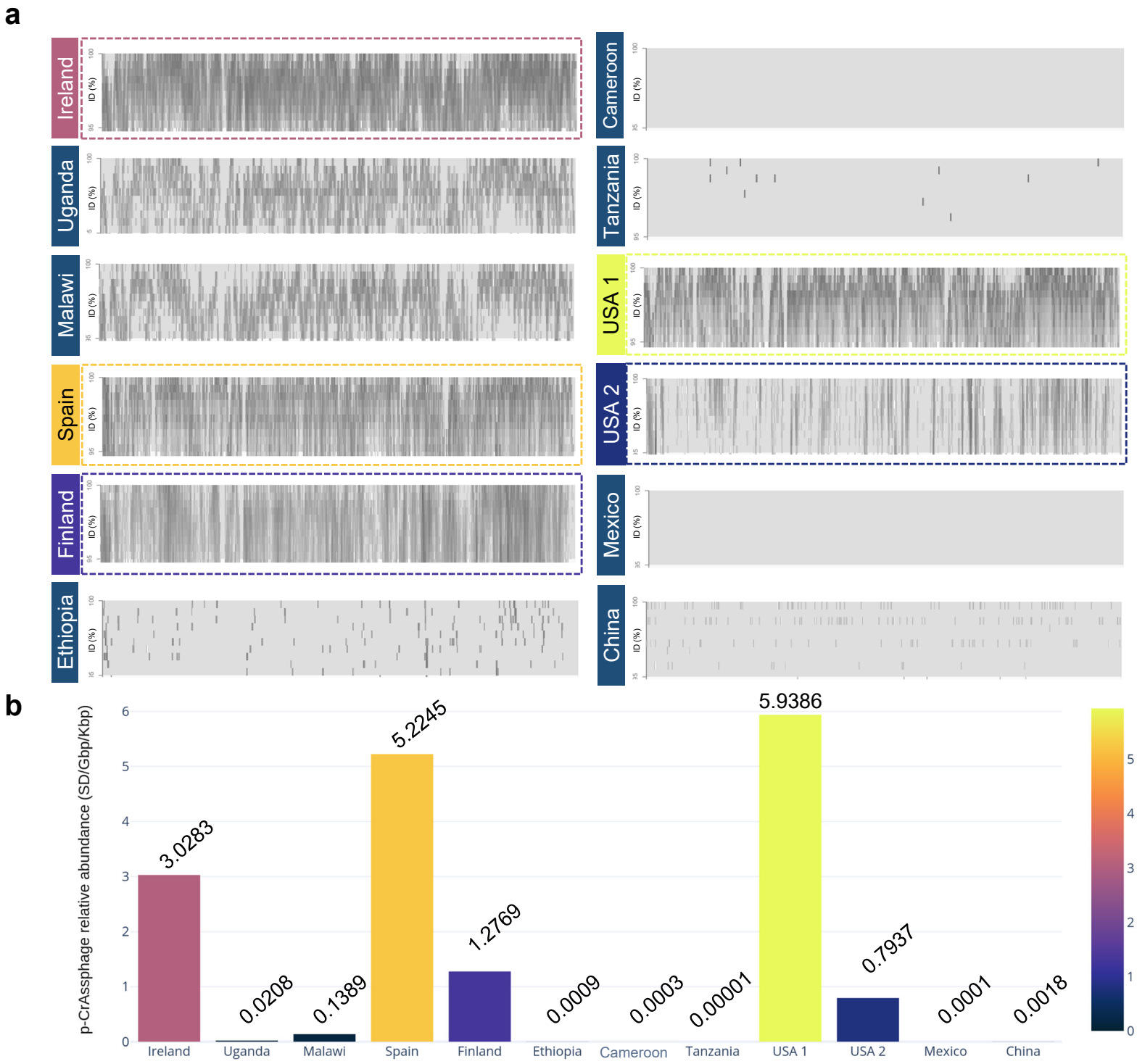
Supplementary Fig. 1. Levels of fecal indicators in the samples used to extract the viromes. *E. coli* (log_{10} CFU/25 g-ml) was used as bacterial fecal indicator while somatic coliphages (log_{10} PFU/25g-ml) were used as viral fecal indicator. Samples were grouped considering their origin.



Supplementary Fig. 2. Relative abundance of all crAss-like phages in the viromes. Each line of the chart shows the relative abundance (normalized by the by metagenome size (Gbp) and phage length (Kbp)) of each of the 91 crAss-like genomes from the databases in the different viromes analyzed. Lower row indicates the averaged relative abundance of each virome (SD/Kb/Gbp). Bdl: Below detection limit. WW. Wastewater treatment plant,



Supplementary Figure 3. Genomic comparison of the genome of the best four selected crAss-like phage. The multialignment of phage genomes were constructed with the complete genomes of the highly human specific crAss-like phages (absent in all animal viromes) by Mauve tool.



Supplementary Figure 4. Geographical distribution of the prototypical p-crAssphage around the globe. (**a)** Fragment recruitment in gut metagenomes from around the globe against the prototypical p-crAssphage **(b)** Relative abundance of the prototypical p-crAssphage in the different countries showing recruitments. Only mapped sequences with coverage of at least 70 % and best-hit reads were considered. Relative abundances of selected crAss-like phages were normalized by metagenome size (Gbp) and phage length (Kbp).

4.4. Capítulo 4:

Persistencia de los fagos crAssBcn en condiciones de inactivación natural y procesos de desinfección, y su potencial como marcadores de trazabilidad de muestras de origen humano

Persistence of crAssBcn phages in conditions of natural inactivation and disinfection process and their potential role as human source tracking markers

Clara Gómez-Gómez, María Dolores Ramos-Barbero, Laura Sala-Comorera, Sara Morales-Cortes, Gloria Vique, Cristina García-Aljaro y Maite Muniesa

Science of the Total Environment (2024); 957, 177450

DOI: 10.1016/j.scitotenv.2024.177450

Introducción

La trazabilidad del origen de la contaminación fecal (MST, por sus siglas en inglés) utiliza indicadores microbianos que permiten identificar fuentes de contaminación fecal en aguas y facilitar la implementación de medidas correctivas (Farnleitner and Blanch, 2017). Los fagos que infectan especies de *Bacteroides* destacan por su especificidad y su incapacidad para replicarse en el ambiente (Jofre et al., 2014), características que los posicionan como marcadores MST efectivos (Tartera et al., 1988; Kator and Rhodes, 1992; Grabow et al., 1995; Muniesa et al., 2012; Jofre et al., 2014). Sin embargo, su aplicación rutinaria enfrenta desafíos como la necesidad de cultivos anaeróbicos de la cepa hospedadora, bajas concentraciones en aguas residuales, la carencia de métodos moleculares eficaces para su detección, así como también la variabilidad geográfica.

Estos obstáculos han llevado al creciente interés en los fagos tipo crAss, que incluyen *p*-crAssphage, descubierto *in silico* en el año 2014 (Dutilh et al., 2014). Debido a su especificidad por el tracto intestinal humano, *p*-crAssphage y otros fagos tipo crAss, miembros del orden Crassvirales (Current ICTV Taxonomy Release | ICTV, 2023), que infectan bacterias del género *Bacteroides,* se han propuesto como indicadores universales de contaminación fecal humana (Ahmed et al., 2018b; García-Aljaro et al., 2017; Stachler et al., 2017). Aunque también han sido detectados en primates (Edwards et al., 2019), y otros animales, en general su presencia en animales es escasa y a niveles menores que en humanos (García-Aljaro et al., 2017). Además, aparentemente estos fagos destacan por su persistencia frente a la

inactivación ambiental (Stachler and Bibby, 2014; Ballesté et al., 2019; Greaves et al., 2020; Ahmed et al., 2021; Zhang et al., 2021) y es posible detectarlos incluso después de procesos de desinfección (Wu et al., 2020; Hamza and Abd-Elmaksoud, 2023).

Aunque se han identificado abundantes secuencias de fagos tipo crAss en diversos metagenomas (Edwards et al., 2019; Yutin et al., 2021) y se han aplicado métodos como qPCR para su detección (García-Aljaro et al., 2017; Morrison et al., 2020; Park et al., 2020; Sabar et al., 2022), en general las técnicas moleculares presentan ciertas limitaciones, ya que no diferencian entre virus infecciosos y no infecciosos, lo que genera preocupación, ya que sólo los virus viables representan un riesgo para la salud pública. Además, la persistencia ambiental de los fagos, esencial para su uso como indicadores, no puede evaluarse adecuadamente con secuencias genéticas, sin analizar el comportamiento de viriones aislados (Guerin et al., 2021). Aun así, estudios previos sugieren que los fagos tipo crAss persisten más tiempo en el ambiente que otros indicadores microbianos (Ballesté et al., 2019).

La reciente disponibilidad de 25 nuevos fagos aislados (crAssBcn) (Ramos-Barbero et al., 2023) ha permitido analizar su persistencia y evaluar su viabilidad como indicadores MST. Esta evaluación se ha realizado mediante enfoques integrados que combinan análisis de infectividad y técnicas moleculares, lo que refuerza su potencial como herramientas clave para el monitoreo de la calidad del agua.

Objetivo

Se evaluó la persistencia de fagos tipo crAss aislados bajo diversas temperaturas y pHs, así como su resistencia frente a métodos tradicionales de inactivación microbiana utilizados en agua y alimentos, incluyendo hipoclorito de sodio, luz UV y temperaturas elevadas. Adicionalmente, se estudió su comportamiento en entornos naturales mediante simulaciones en mesocosmos. El objetivo principal fue comprender los efectos individuales de estos parámetros fisicoquímicos en la viabilidad de los fagos tipo crAss utilizados como potenciales marcadores de MST.

Resultados y discusión

Se evaluó la estabilidad a temperaturas de 4°C, 22 °C y 37°C, representando condiciones ambientales y fisiológicas, de 3 fagos crAssBcn, del fago ΦCrAss001, de un control de fago (GA17-A que infecta *Bacteroides thetaiotaomicron*) y un control bacteriano (*Bacteroides intestinalis*, el hospedador de los fagos crAssBcn). A 4°C y 22°C, todos los fagos infecciosos mostraron una alta estabilidad, con reducciones menores a un $\log_{10}$ tras 30 días de tratamiento. Sin embargo, a 37°C la estabilidad disminuyó, con reducciones de 1,5 a 2 $\log_{10}$ en el mismo período, tanto para los crAssBcn como para el fago GA17-A. En comparación

con otros microorganismos indicadores de contaminación fecal, como los colifagos (Sinton et al., 1999; Allué-Guardia et al., 2012; Mcminn et al., 2014) y bacterias (*Bacteroides intestinalis*), los crAssBcn demostraron mayor persistencia. Al ser evaluados por qPCR, el ADN vírico se pudo detectar en todas las condiciones, sin disminuciones significativas.

El estudio también examinó el efecto de diferentes pHs (3, 7 y 9) a lo largo de 29 días. Durante la exposición a pH 3, los fagos del tipo crAss fueron detectables por qPCR, pero se inactivaron completamente en 15 días (decaimiento menor que los controles de fago y bacteria), atribuyendo este efecto al daño de las proteínas de la cápside debido a las condiciones ácidas del medio (Sharp et al., 1946), descartando la posible formación de agregados ya que no se observó una desagregación tras neutralizar el pH. A pH neutro y alcalino, los fagos conservaron su infectividad y estabilidad del ADN, lo cual confirma ratios de inactivación menores en rangos de pH entre 6 y 9 (Sharp et al., 1946; Kerby et al., 1949; Feng et al., 2003).

En cuanto al tratamiento con hipoclorito de sodio a 10 ppm, utilizado comúnmente en la desinfección de aguas, el resultado fue una reducción significativa de partículas infecciosas, de hasta 4,77 $\log_{10}$ tras 30 minutos, alcanzando un punto de estabilización. Estos hallazgos sugieren que los fagos tipo crAss pueden detectarse en aguas tratadas con cloro que cumplen los criterios de desinfección bacteriológica (Durán et al., 2003). Este comportamiento contrasta con estudios previos en los que otros fagos fueron completamente inactivados en condiciones similares (Allué-Guardia et al., 2012). Del mismo modo, la bacteria hospedadora empleada en este ensayo mostró una inactivación inmediata. Estructuras como las proteínas de cápside, altamente susceptibles al daño por cloro (Zhong et al., 2016), explican la notable reducción en la infectividad de los fagos, mientras que el ADN, al carecer de estos componentes vulnerables, mantuvo una mayor estabilidad molecular (Qiao et al., 2022). La eficacia del cloro puede variar dependiendo de la calidad del agua (Kahler et al., 2010), pero en este estudio se minimizó este efecto utilizando agua destilada para las diluciones de las muestras.

El tratamiento con luz UV mostró que los fagos crAssBcn y ΦcrAss001 fueron más sensibles que GA17-A y *B. intestinalis*, alcanzando la inactivación completa en tan solo 5 minutos, mientras que los controles requirieron 30 minutos de exposición. Estas diferencias en las tasas de inactivación entre los fagos tipo crAss y el fago control, se podrían atribuir a variaciones en la morfología de la cápside (podovirus frente a siphovirus, respectivamente), el contenido de GCs (Sun et al., 2023) y el tamaño del genoma, ya que los crAssphage tienen un genoma más grande, lo que aumenta el número de dianas que pueden verse afectadas por la radiación, facilitando su inactivación por UV. Las diferencias observadas en los resultados de la cepa hospedadora podrían estar relacionadas con la presencia de CPS que protegen a la

bacteria de la radiación UV (Akhlaq et al., 1990; Nunez et al., 2023). El tratamiento con UV fue el único agente que redujo la concentración de genomas víricos, aunque ligeramente, lo que sugiere que el daño principal causado por la radiación UV afecta primero a la cápside (Pecson et al., 2009), y luego al ADN. Los tratamientos secuenciales con UV seguidos de hipoclorito de sodio podrían mejorar la eficiencia de la desinfección (Shang et al., 2007; Zhang et al., 2015).

A temperaturas de 60°C y 70°C, comúnmente utilizadas para eliminar microorganismos en aguas, alimentos y lodos (Mocé-Llivina et al., 2003; Nuanualsuwan and Cliver, 2003; Guglielmotti et al., 2012; Peng et al., 2017), *B. intestinalis* se inactivó completamente a los 30 minutos a ambas temperaturas, mostrando una alta tasa de decaimiento. En contraste, los crAssBcn y el ΦCrAss001 demostraron una mayor resistencia que el fago GA17-A, lo que podría atribuirse a una mayor estabilidad de la partícula, posiblemente debido a la estructura de la cápside (Atamer et al., 2009). Los crAssBcn resultaron prácticamente estables a 60°C, mientras que, a 70°C, la disminución de partículas infecciosas fue de hasta 3 $\log_{10}$ probablemente porque las cápsides se ven afectadas a estas temperaturas elevadas (Nuanualsuwan and Cliver, 2003; Wigginton et al., 2012). Este resultado difiere de la completa inactivación observada en el fago GA17-A que infecta a *B. thetaiotamicron.* Estas diferencias en cuanto a la resistencia térmica entre los fagos confirman que los fagos que infectan especies de *Bacteroides* son un grupo heterogéneo. A altas temperaturas, tal y como se demostró en otros estudios (Nuanualsuwan and Cliver, 2003; Pfaender et al., 2015), la estabilidad del ADN viral se mantuvo intacta.

Se realizaron varios experimentos de mesocosmos exteriores (con tres réplicas en verano y tres en invierno) para evaluar el impacto de las condiciones naturales sobre los diferentes microorganismos. Para ello, los fagos y bacteria se introdujeron en membranas de diálisis que a su vez fueron introducidas a 10 cm de la superficie del agua en una balsa exterior conteniendo agua no clorada y se realizaron experimentos en verano y en invierno. De todos los factores evaluados, los más variables a lo largo de los experimentos fueron la temperatura (tanto externa como del agua de la balsa) y la irradiación solar. Los ensayos realizados durante el verano, con una duración de 24 días, mostraron una mayor reducción de la infectividad, alcanzando hasta 6,5 $\log_{10}$ para los fagos y una inactivación completa para las bacterias. Estos resultados se atribuyen a la contribución de la irradiación solar, un factor clave en la desinfección de superficies acuáticas (Sinton et al., 1999), siendo los rayos UV suficientes para causar daño fotobiológico (Nelson et al., 2018). Durante el invierno, la inactivación fue más lenta, siendo las bacterias (que llegaron a inactivarse totalmente) más afectadas que los fagos (que mostraron una reducción de hasta 4,4 $\log_{10}$). La evaluación del ADN por qPCR mostró una reducción significativa en verano, aunque menor que la de la

infectividad, y menor en los meses de invierno, donde la reducción de copias genómicas fue mucho más sutil. Siendo éste el primer estudio sobre la inactivación natural de fagos tipo crAss aislados, las comparaciones se tuvieron que realizar con los análisis de genomas fágicos, concluyendo que los crAssBcn podrían ser más persistentes que el fago *p*-crAssphage (García-Aljaro et al., 2017; Stachler et al., 2017; Ballesté et al., 2019; Greaves et al., 2020). Por otro lado, la detección de *p*-crAssphage únicamente por métodos moleculares (dado que no existe virión aislado) podría ser más sensible, aunque menos realista. En comparación con otros fagos indicadores, como los colifagos somáticos, la persistencia de los fagos de tipo crAss fue similar durante la temporada más cálida y algo superior en los meses más fríos.

Conclusiones

Los fagos tipo crAss, particularmente los crAssBcn, destacan como prometedores marcadores virales de MST para monitorear la calidad del agua, gracias a su notable estabilidad y persistencia, tanto a procesos de desinfección como a las condiciones ambientales naturales. Su persistencia es comparable o superior a la de virus patógenos, aunque no siempre correlacionan con la presencia de éstos. Su principal ventaja radica en que son más fácilmente detectables en muestras contaminadas fecalmente, ya que se encuentran en mayores concentraciones en el ambiente que los virus patógenos. La investigación continúa enfocada en superar desafíos como el comportamiento de estos fagos en matrices ambientales más complejas, la variabilidad genética dentro del orden Crassvirales y la mejora de los métodos de evaluación, que combinan análisis moleculares y de infectividad (Zhang et al., 2019).

Science of the Total Environment 957 (2024) 177450

Contents lists available at ScienceDirect

Science of the Total Environment

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

Persistence of crAssBcn phages in conditions of natural inactivation and disinfection process and their potential role as human source tracking markers

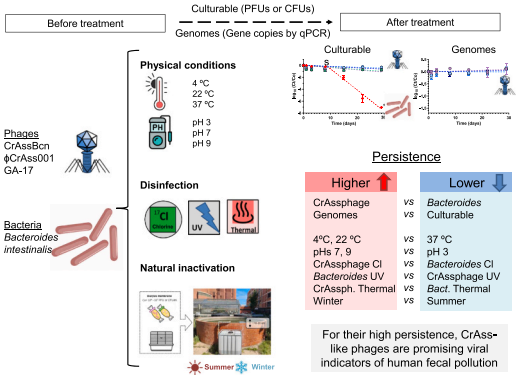
Clara Gómez-Gómez, Maria Dolores Ramos-Barbero, Laura Sala-Comorera, Sara Morales-Cortes, Gloria Vique, Cristina García-Aljaro, Maite Muniesa *

Department of Genetics, Microbiology and Statistics, Section of Microbiology, Virology and Biotechnology, School of Biology, University of Barcelona, Diagonal 643, E-08028 Barcelona, Spain

HIGHLIGHTS

- Infectious crAss-like phages show high persistence to inactivation factors.
- Low pH reduces drastically the infectivity of crAss-like phages.
- CrAssBcn phages and ΦcrAss001 show identical inactivation patterns.
- crAss-like phages persist better than *Bacteroides intestinalis* at all conditions.
- Its environmental persistence makes crAss-like phages good MST indicators.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:
CrAssphage
Bacteriophage
Water disinfection
Microbial source tracking
Natural inactivation
Chlorination
Persistence

ABSTRACT

Due to their abundance in the human gut, human specificity, and global distribution, some crAss-like phages, including the original p-crAssphage, have been proposed as indicators of human fecal pollution suitable for microbial source tracking (MST). The prevalence of crAss-like phages in water, and consequently their usefulness as MST indicators, is determined by their ability to survive various inactivation and disinfection processes. Recently, we isolated new crAss-like phages (named crAssBcn phages) capable of infecting *Bacteroides intestinalis* and exhibiting a wide geographical distribution. Here, we assessed the infectivity and DNA integrity of three crAssBcn phages (ΦCrAssBcn6, 10, and 15) and ΦCrAss001, the first crAss-like phage isolated, at different pHs and temperatures, after UV and chlorine treatments, and under natural conditions. Their bacterial host, *B. intestinalis* and a siphovirus *Bacteroides*-infecting phage GA17-A were used as controls. Infectious crAssBcn phages remained stable for a month at 4, 22, and 37 °C, and at pH 7, but inactivated when exposed to pH 3. Infective crAssBcn phages decreased by 5 $\log_{10}$ after treatment with 10 ppm of chlorine for 1 min and after UV treatment at a fluence of 5.94 mJ/cm^2. However, heat treatment at 60 and 70 °C resulted in only a moderate decrease (<1 $\log_{10}$ and almost 3 $\log_{10}$ units of reduction, respectively). Experiments under natural conditions in

* Corresponding author.
E-mail address: mmuniesa@ub.edu (M. Muniesa).

https://doi.org/10.1016/j.scitotenv.2024.177450
Received 11 August 2024; Received in revised form 16 October 2024; Accepted 6 November 2024
0048-9697/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

outdoor mesocosms revealed that inactivation rates for crAssBcn phages, as for the other microorganisms, were higher in summer (up to 6 $\log_{10}$) than in winter (<4 $\log_{10}$), suggesting a higher incidence of inactivation factors, such as sunlight and temperature, in the warmer months. *B. intestinalis* was significantly more prone to inactivation than phages in most conditions except for the irradiation treatment. In contrast, crAssBcn phage DNA remained stable, with minimal reduction under most of the tested conditions, except in the summer mesocosm and UV assays.

1. Introduction

The first prototypical-crAssphage (p-crAssphage) described *in silico* in 2014 (Dutilh et al., 2014) is highly abundant in the human intestinal tract and has a global distribution (Edwards et al., 2019). Although p-crAssphage has also been detected in the metagenome of primates (Edwards et al., 2019), its presence in the feces of other animal species has been observed only rarely and at lower levels (García-Aljaro et al., 2017). Since the discovery of p-crAssphage, which is classified under the family *Intestiviridae*, species *Carjivirus communis* (Turner et al., 2023), other crAss-like phages, mostly uncultured, have been described, forming the new order of Crassvirales (Current ICTV Taxonomy Release | ICTV, 2023). CrAss-like phages share a similar genomic architecture and a common ancestor with p-crAssphage. Many are also highly abundant in the mammalian gut and other habitats (Guerin et al., 2018; Yutin et al., 2021) and many are exclusively found in the human gut. The Crassvirales group has expanded with the recent isolation by our group of 25 crAss-like phage virions infecting *Bacteroides intestinalis* (named crAssBcn phages) from wastewater in Barcelona, Spain (Ramos-Barbero et al., 2023). Currently, the order Crassvirales comprises four families, 11 subfamilies, 42 genera, and 80 species (https://ictv.global/taxonomy) (Ramos-Barbero et al., 2023). All members of this order appear to infect bacteria in the phylum Bacteroidetes and, according to the few virions isolated to date, they specifically target the *Bacteroides* genus.

Due to their specificity for the human intestinal tract, p-crAssphage and other members of the CrAssvirales have been proposed as universal indicators of human fecal pollution (García-Aljaro et al., 2017; Stachler et al., 2017; Ahmed et al., 2018; Sala-Comorera et al., 2021), despite their genetic variability (Koonin and Yutin, 2020; Yutin et al., 2021; Ramos-Barbero et al., 2023). The use of specific human indicators of fecal pollution is of particular interest for microbial source tracking (MST) (Harwood et al., 2014; Jebri et al., 2017), a methodology aimed at identifying the origin of fecal contamination with microbiological markers. Once identified, mitigation measures can be implemented to correct the problem at the source (Farnleitner and Blanch, 2017). Fecal pollution, especially in water, is a major global public health concern (World Health Organization, 2008). Identifying and tracking the source of fecal contamination can improve pollution management, reduce economic losses, resolve legal issues, and protect human health (Jagals and Grabow, 1996; Blanch et al., 2004).

Bacteriophages infecting *Bacteroides* have previously been proposed for MST (Tartera et al., 1988; Kator and Rhodes, 1992; Grabow et al., 1995; Muniesa et al., 2012; Jofre et al., 2014) due to the high specificity of certain *Bacteroides* species and their phages for human or animal hosts (Payan et al., 2005). However, established methods based on infectious bacteriophages (Araujo et al., 2001; Payan et al., 2005; Ebdon et al., 2007; Muniesa et al., 2012; Jofre et al., 2014) require the culture of anaerobic *Bacteroides* as host strains, which hinders their routine implementation in laboratories. This limitation could be overcome by using *Bacteroides*-infecting phages that can be detected by molecular assays and are strongly associated with human fecal pollution, such as crAss-like phages.

An abundance of crAss-like phage sequences have been identified in various metagenomes (Edwards et al., 2019; Yutin et al., 2021) and an increasing number of studies are using qPCR to detect crAssphages (or crAss-like phages) in different ecosystems (García-Aljaro et al., 2017; Morrison et al., 2020; Park et al., 2020; Sala-Comorera et al., 2021; Sabar et al., 2022). Most of these analyses have relied on molecular methods because p-crAssphage and most crAss-like phages are uncultured viral genomes. To date, only a few, including ΦCrAss001, ΦCrAss002, and the 25 crAsssBcn phages (Ramos-Barbero et al., 2023), are available as isolated viruses.

Current molecular methods for detecting crAss-like phages have limitations because they can be detecting a DNA fragment rather than a complete phage and they do not distinguish between infectious and not infectious phages. Phage persistence in the environment - a requisite for a good indicator - cannot be properly studied without isolated viruses and the use of phage-host pairs (Guerin et al., 2021). The widespread distribution of crAss-like phage genomes across various environments suggests they can persist under diverse conditions. Previous studies have indicated that crAss-like phage genomes may survive in the environment longer than other microbial indicators, such as *Escherichia coli*, and comparably to other phage indicators such as somatic coliphages (Ballesté et al., 2019). The recent isolation of 25 new crAss-like phages (Ramos-Barbero et al., 2023) has enabled us, for the first time, to conduct a more accurate investigation into the persistence and inactivation of this group of phages. Hence, by integrating infectivity and molecular analyses, we were able to better assess the suitability of crAss-like phages as MST indicators.

In the present study, we evaluated the persistence of three crAssBcn phages (6, 10, and 15 representing species I, II and III, respectively), isolated and classified by Ramos-Barbero et al. (2023). We also analyzed ΦCrAss001, the first crAss-like phage isolated (Shkoporov et al., 2018) of the same genus than crAssBcn phages (*Kehishuvirus*). The phage GA17-A, a siphovirus infecting *Bacteroides thetaiotaomicron* (data not shown), was included as a control to evaluate a *Bacteroides* phage of a morphological group different from crAss-like phages. *B. intestinalis*, the host strain of the crAss-like phages in this study, was used as bacterial control. Phage stability at different temperatures and pHs was evaluated as well as their resistance to disinfection when exposed to treatments with chemical (sodium hypochlorite) and physical agents (UV light and heat), which are traditional methods for the rapid and efficient inactivation of viruses in water or food matrices. Additionally, phage behavior was studied under natural conditions through a mesocosm simulation. The primary objective of the persistence and inactivation assays was to understand the individual effects of each physical-chemical parameter on these potential viral markers.

2. Material and methods

2.1. Bacteria, bacteriophages, strains, and media

CrAssBcn bacteriophages ΦCrAssBcn6, ΦCrAssBcn10, and ΦCrAssBcn15 were isolated in previous studies (Ramos-Barbero et al., 2023). ΦCrAss001 bacteriophage (DSM 109066) (Shkoporov et al., 2018) and *B. thetaiotaomicron*-infecting phage GA17-A (this study) were included as controls. *B. intestinalis* strain APC919/174 (DSM 108646) was used as a control for bacterial persistence and as the host strain to evaluate the infectivity of crAssBcn phages and phage ΦCrAss001. *B. thetaiotaomicron* strain GA17 (Payan et al., 2005), was used as the host for phage GA17-A.

Anaerobe basal broth (ABB) (Thermo Scientific, Wilmington, MA, USA) was used for the growth of the *Bacteroides* strains. For

bacteriophage enumeration, ABB medium containing 0.7 % (*w/v*) agar-agar for soft agar or 1.4 % (w/v) agar-agar for agar plates was used. Phage buffer (100 mM NaCl, 10 mM $MgCl_2$, 50 mM Tris-HCl, 0.01 % gelatin, pH 7.5) and phosphate-buffered saline 1× (PBS) were used to make dilutions of phages (1:100) and bacteria (1:10), respectively.

Manipulation was performed in aerobic conditions, whereas incubation of cultures and agar plates was done at 37 °C for 24 h or 48 h in anaerobic jars (GasPak; BD-BBL, Madrid, Spain) with CO_2 atmosphere generators (Anaerocult A; Sigma-Aldrich. Saint Louis, MI, US) and anaerobic indicator strips (Anerotest; Merck Life Science, Madrid, Spain).

2.2. Preparation of bacteriophage suspensions

Phage lysates were prepared as previously described (Ramos-Barbero et al., 2023). Briefly, 0.1 ml of phages from a stock was used to infect 1 ml of *Bacteroides* cultures grown to the mid-exponential growth phase (OD_{600} of 0.3) in 8 ml of ABB medium and incubated anaerobically at 37 °C for 24 h. After incubation, phage suspensions were purified by filtration of the cultures through low protein-binding 0.22-μm pore size membrane filters (Millex-GP; Millipore Bedford, MA, USA), and phages in the filtrate were enumerated by plaque assays. Phage suspensions were diluted with phage buffer to obtain the same phage density for all phages ($10^{6–7}$ plaque forming units (PFU)/ml) at the beginning of the experiments. To ensure no inactivation occurred other than that due to the applied treatment, fresh phage suspensions were prepared at the beginning of each experiment, and then diluted and quantified to obtain the same phage densities.

2.3. Bacteriophage and bacteria enumeration

The infectivity of the crAss-like phages and GA17-A after all treatments was evaluated at different time intervals using the double agar layer plaque assay, as previously described (Adams, 1959). Briefly, 0.1 ml of serial decimal dilutions of the phage suspensions, or the phages in the treated samples, were used to infect 1 ml of the respective *Bacteroides* cultures (*B. intestinalis* for CrAssBcn phages and *B. thetaiotaomicron* for GA17-A) grown at the mid-exponential growth phase (OD_{600} of 0.3) in 2.5 ml of ABB soft agar, gently mixed, and poured on ABB agar plates. Phage enumeration of each suspension was performed in triplicate. Plates were incubated in anaerobic jars at 37 °C for 18-24 h. Medium, bacterial and phage negative controls were used in each experiment.

The enumeration of the bacterial control *B. intestinalis* was performed by colony counts using ABB agar plates. Briefly, 0.1 ml of serial decimal dilutions of the bacterial cultures at the different times of the treatment were plated in triplicate on ABB agar plates and incubated under anaerobic conditions in anaerobic jars (GasPak; BD-BBL) at 37 °C for 24 h. After incubation, white colonies generated by *B. intestinalis* were enumerated. Since the bacteria was analyzed in pure cultures, only a single morphology was observed. Medium and bacterial controls were used in each experiment.

2.4. Phage DNA isolation

Phage DNA was extracted using the QIAamp DNA blood mini kit (Qiagen GmbH, Hilden, Germany), following the protocol indicated by the manufacturer. One negative extraction control was introduced using ABB broth. The DNA was suspended in a final volume of 200 μl of sterile double-distilled water. The DNA concentration of each sample was measured using a Qubit Fluorometer and Qubit 1× dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, US), and the DNA quality was further confirmed by the 2100 Bioanalyzer system (Agilent Technologies, Santa Clara, CA, US).

2.5. qPCR assays

A qPCR assay specific for the three crAssBcn phages analyzed and also detected phage ΦCrAss001 was used to evaluate the decay of phage DNA throughout the treatments. This qPCR assay incorporated an forward primer (5′-ACCTGCTTCTACACTTTCCTT-3′) a reverse primer (5′-AGTGCTCCAGAATAGGATTGT-3′) and a TaqMan hydrolysis probe (6FAM-ATATCTAACCCAGCTC-MGBNFQ) (Ramos-Barbero et al., 2023) was designed from the ΦCrAss001 genome (NC_049977.1) targeting a gene coding for a hypothetical protein (ORF 46) also present in crAssBcn phages, with an amplicon size of 75 pb.

qPCR amplifications were carried out using the StepOne Real Time PCR System (Applied Biosystems) in a 20 μl reaction mixture with TaqMan Environmental Master Mix 2.0 (Thermo Fisher Scientific), 18 μM of a mixture of both primers and 5 μM of the TaqMan probe. The reaction contained 9 μl of the sample DNA, controls or standards. The thermal-cycler conditions were 10 min at 95 °C, 40 denaturation cycles of 15 s at 95 °C, and 1 min of annealing and extension at 60 °C. Standard curve was prepared with a stock with known gene copy (GC) number of gBlocks Gene Fragments (Integrated DNA Technologies, Coralville, USA) of 490 bp including the hypothetical protein located in crAssBcn and ΦCrAss001 phages described above. Serial decimal dilutions of this stock were made in double-distilled water that were aliquoted and stored at −80 °C until use. Values of the standard curve and limits of detection were calculated by performing 10 replicates of each dilution of the standard, in three independent qPCR reactions and by triplicate. Three replicates of each dilution were added to each qPCR reaction. The results were analyzed with the Applied Biosystems StepOne Instrument program. The qPCR assay showed a range of the curve from 10 to 10^6 gene copies (GC) per reaction, a detection limit of 1.61 GC (threshold cycle of 35.07), an efficiency (*E*) of 96.49 %, a slope (−3.409), and a R^2 of 0.992. All samples were run in triplicate, as well as the standards, and positive and negative controls. The number of gene copies (GC) was defined as the mean of the triplicate data obtained. DNA extraction negative control as well as qPCR negative (water) and positive (standard of known concentration) controls were included in each run.

2.6. Influence of temperature and pH on crAssBcn phages

Ten milliliters of phage lysates containing $10^{6–7}$ PFU/ml in phage buffer and $10^{6–7}$ CFU/ml of *Bacteroides* in PBS were incubated in 10 ml tubes in darkness for 1, 3, 8, 15, 22, and 29 days at 4, 22, and 37 °C. Ten milliliters of phage lysates containing $10^{6–7}$ PFU/ml in phage buffer and $10^{6–7}$ CFU/ml of *Bacteroides* in PBS were adjusted to different pHs (3, 7, and 9) by adding HCl 37 % for pH 3 or NaOH 1 N for pH 9 and incubated in 10 ml tubes in darkness for 1, 3, 8, 15, 22, and 29 days at 4 °C. The pH in each tube was confirmed before conducting the analysis. At each interval time, crAss-like phages were evaluated by plaque assay (infectious phages) and by real time qPCR (phage DNA (GC)). The control virulent phage of *Bacteroides* GA17-A was enumerated by double agar layer and the number of cells of the bacterial control *B. intestinalis* was enumerated on ABB agar plates as indicated above. Negative controls of medium, bacteria and phages were included. To evaluate if aggregation was responsible of the reduction in the number of phages at pH 3, the experiment was repeated with all the phages and after 8 days of incubation, the pH of the phage suspensions was adjusted back to pH 7 by adding NaOH 1 N. Neutralized phage suspensions were then incubated for 1 h under agitation at 4 °C, and infectious phages were evaluated by double agar layer.

2.7. Inactivation conditions

Aliquots of phage lysates containing approximately $10^{6–7}$ PFU/ml of phages in phage buffer and $10^{6–7}$ CFU/ml of *Bacteroides* in PBS were used for the inactivation experiments. Negative controls of medium, bacteria and phages were included in all the conditions assayed.

2.7.1. Heat inactivation

One milliliter of phage lysates and the bacterial culture in 1.5 ml tubes were incubated (Fisherbrand Isotemp Digital Dry Baths/Block Heaters, Thermo Fisher Scientific) at 60 °C and 70 °C for 30 and 60 min, temperatures close to those used for pasteurization.

2.7.2. Chlorine inactivation

One milliliter of phage lysates and the bacterial culture were diluted 1:100 in double-distilled water, to minimize interferences with chlorine, and treated with chlorine supplied as sodium hypochlorite to a calculated final concentration of 10 ppm, confirmed using the DPD (*N*,*N*-diethyl-*p*-phenylenediamine) colorimetric technique (Palin, 1957) with a Chlorine Test Kit following the manufacturer's instructions (Merck, Darmstadt, Germany). Free chlorine measurements were completed within 1 min of sample collection, to minimize interference from chloramines, as recommended by Jensen and Johnson, 1990. Each sample was incubated for 1, 5, 10, 20, and 30 min at room temperature (21 °C). After 30 min of incubation, the levels of free chlorine in the 30 min tubes were measured with the DPD colorimetric technique showed the presence of 3 to 4 ppm in the different replicates. Following incubation, and before conducting the infectivity and quantification assays, the residual chlorine was neutralized by adding 3 % (*w/v*) sodium thiosulphate (Durán et al., 2003).

2.7.3. UV treatment

Five milliliters of phage lysates and the bacterial culture were placed in open petri dishes, 10 cm below an 8-W germicidal UV lamp (model G30T8 operating at 253.7 nm wavelength; Sankyo Denki, Tokyo, Japan). Plates were exposed to UV for 1 to 30 min (Allué-Guardia et al., 2014). The lamp was warmed up for 15 to 30 min before starting the experiments. The UV fluence rate was calculated using the eq. $D = I \times T$, where D is the dose, T is the exposure time (60, 300, 600, and 1800 s), and I is the intensity received by the samples placed at 10 cm from the lamp (0.099 mW/cm^2).

2.8. Natural inactivation in an outdoor mesocosm

An outdoor pond maintained in the University campus for experimental purposes containing 60 m^3 of non-chlorinated water from a well was used to simulate inactivation under natural conditions. Twenty milliliters of clear suspensions containing approximately 10^{6-7} PFU/ml of phages in phage buffer and 10^{6-7} CFU/ml of *Bacteroides* in PBS were used for the natural inactivation experiments. Thirty replicates of suspensions containing each microorganism were separately introduced into dialysis membranes (cut-off 12–14 kDa) (Medicell Membranes Ltd., London, UK) and sealed with reusable closures. Dialysis membranes were transparent and were placed within a cage (a metal bars cage 50 cm × 30 cm × 30 cm) and immersed in the outdoor pond at a depth of 10–20 cm from the surface to replicate the sunlight conditions that microorganisms shall receive under natural conditions. Three tubes for each microorganism were removed at various time intervals (1, 3, 7, 10, 14, 17, 20, and 24 days in summer and 1, 3, 8–10, 15–17, 24, 31, 38, and 45 days in winter) to evaluate the abundance and decay rate of the phage lysates and the bacterial culture, analyzed by culture and molecular methods. A negative control of the well water was monitored in each interval to confirm the absence of the microorganisms analyzed in the pond. The experiments performed in summer lasted only for 24 days due to membrane deterioration (June-August 2023: minimal and maximal temperature of 20.9 °C–36.1 °C; average solar radiation of 23.6 MJ/m^2), and those in winter lasted for 45 days (January-March 2024: temperature of 6.1–22.6 °C; average solar radiation of around 12 MJ/m^2). Temperature and irradiation data were obtained from the regional weather forecast service database (*Servei Metereologic de Catalunya*, Spain). The physical-chemical characteristics of the water pond (temperature, pH, dissolved oxygen concentration, and conductivity) were recorded during the assays (Table 1), using a multiparameter waterproof meter (Hanna Instruments, Romania). The pH (~8), salinity, and dissolved oxygen of the pond remained stable throughout the experiments.

2.9. Data analysis

Data computation, statistical tests, and the creation of charts were done using GraphPad Prism 10 (GraphPad Software, San Diego, CA, US). Predicted (dashed lines) and observed values (data points), normalized by the initial concentration, were plotted on figures to represent logarithmic inactivation $[log(N_T/N_0)]$ over time. Data from triplicates were combined. Phage and bacterial decay rates were assumed to follow first-order decay kinetics according to Chick (1908). Three types of decay models were employed for each experiment: linear regression, biphasic decay (determined using a segmented pairwise regression) (Cerf, 1977; Brouwer et al., 2017), and monophasic nonlinear exponential decay (Fritz, 2014). Firstly, the linear regression is described by the following formula:

$$N_t = N_0 e^{-kt}$$

where N_t corresponds to a microbial concentration or GC at the specific time (t), N_0 is the initial concentration, and k is the inactivation rate constant (t^{-1}). k was set as equal to the slope of the line with 95 % confidence intervals. Secondly, for the biphasic model, two different decay rates were identified with a break point between segments (Lee et al., 2001),

$$N_t = N_{f0} e^{-kft} + Ns_0 e^{-kst}$$

where k_f corresponds to the fast decay rate and k_s to the slow decay rate. K_f (fast decay rate) and k_s (slow decay rate) can also be defined as k_1 and k_2 respectively. N_{f0} and N_{s0} are the initial concentrations in the fast and slow decay periods, respectively. Finally, when the initial concentration N_0 decays in one phase to a plateau (P), which is the asymptotic value that the concentration variable approaches as time tends to infinity, a one-phase nonlinear decay model would be the best fit.

$$N_t = (N_0 - P)e^{-kt} + P$$

The adequacy of the one-phase decay model was assessed using r^2 values, indicating the percentage of variability explained by the independent variable within the model, whereas the root mean square error (RMSE) quantifies the discrepancy between observed values and predicted values.

The time required for a 1-log inactivation of the initial viral concentration, known as T_{90}, can be estimated by solving (t) from the above equations using the most probable parameter values.

Table 1
Values for the physical and chemical parameters measured during natural inactivation assays (values shown are means ± standard deviations).

Season	Water temperature (°C)[a]	External temperature mean (°C)[b]	Daily irradiance (MJ/m^2)[b]	DO (ppm)[a]	pH[a]	Conductivity (mS/cm)[a]
Summer	29.11 ± 1.41	26.77 ± 1.33	23.64 ± 4.73	7.10 ± 0.65	8.04 ± 0.20	6.58 ± 0.16
Winter	12.97 ± 0.66	12.93 ± 1.11	11.95 ± 4.05	10.4 ± 1.27	8.17 ± 0.05	5.08 ± 0.07

[a] Physical-chemical factors measured with the multiparameter waterproof meter.
[b] Information provided by Catalunya Forecast Service© Meteocat 2023/2024.

2.10. Statistical analyses

GraphPad Prism was also used to perform two-way ANOVA with Tukey's multiple comparison tests, to perform an evaluation of the samples across the different disinfection treatments, as well as to assess the impact of the mesocosm period (summer and winter) by comparing the decay rates of each sample (with a significance of $p < 0.05$).

3. Results and discussion

3.1. Stability of crAss-like phages at different temperatures and pHs

3.1.1. Temperature

We evaluated the stability of infectious phages and the integrity of their DNA at different temperatures and pHs. Using initial concentrations of all microorganisms ranging from 6.8 to 7.7 $\log_{10}$ PFU/ml or CFU/ml, the results are expressed as the logarithm of the ratio of the final microorganism concentration to the initial concentration of the experiment. This normalization allows direct comparison between different datasets. When evaluating the microorganisms, most of them showed a linear regression, whereas at certain temperatures some exhibited a biphasic decay profile with two inactivation rates (k_1 and k_2) (Table 2 and Supplementary Table 1).

The three temperatures tested represent seasonal fluctuations in our region: 4 °C to simulate winter conditions, 22 °C to assess inactivation at room temperature, and 37 °C to represent both the maximum summer temperature and physiological temperature inside the human body. At 4 °C and 22 °C, all phages, showed remarkably similar behavior, remaining stable (<1-log reduction) throughout the 30 days of incubation (Fig. 1), with no significant differences between them. At 37 °C, all tested phages showed lower stability, with a decrease of 1.5 to 2 $\log_{10}$ after one month of storage. At this temperature, the decay rates of all phages (Supplementary Table 1) indicated slightly higher deterioration in the initial phase ($k = 0.34–0.82$ days^{-1}, $T_{90} = 1.22–2.91$ days), followed by a very slow phase ($k = 0.01–0.04$ days^{-1}, T_{90} 28.45–32.84 days). The biphasic kinetics of thermal inactivation in phages have been attributed to the presence of mutant viral species in heterogeneous populations (Nowak, 1992). Thus, while a portion of the population was sensitive to the temperature exposure, another portion was relatively resistant (Yesil et al., 2024). The tested phages showed greater persistence under all conditions compared to other phages previously described as indicators of human fecal contamination, although it should be emphasized that our samples were clear phage suspensions, free from other microorganisms or external contaminants. For instance, a phage infecting *Bacteroides fragilis* strain GB-124 was reported to reduce its concentration in sewage samples by 3 to 4 logarithms in the first week of treatment at three temperatures (Mcminn et al., 2014). In the study by Sinton et al. (1999), the *B. fragilis* HSP40-infecting phage B40-8 was more persistent in contaminated seawater at temperatures between 14 and 20 °C than somatic coliphages and F-specific coliphages. However, based on T_{90} values, it was inactivated more rapidly than the phages in the present study. Very similar T_{90} values were reported by Allué-Guardia et al. (2012) for a *E. coli*-infecting phage at the same temperatures.

When evaluating phage genomes, considering that GC values can be extrapolated to the quantity of crAss-like phage DNA in each sample, the decay profile was linear for all conditions, with no significant differences between them ($p > 0.05$). Long-term treatment at all three temperatures did not affect the integrity of crAss-like phage DNA, as GC values remained stable after one month of treatment.

The inactivation of the bacterial control, *B. intestinalis*, was generally greater than that of phages, with a reduction of up to 4 $\log_{10}$ units at 4 °C after one month of storage ($k = 0.14$ days^{-1}, $T_{90} = 7.05$ days), and complete inactivation at 22 °C within the same period. Similar temperature effects have been observed for other *Bacteroides* strains in the first 8 days of incubation (Ballesté and Blanch, 2010). At 37 °C, complete inactivation occurred at the third week of testing. An increase of bacterial CFUs was observed at the beginning of the experiment until day 13. This was attributed to bacterial regrowth, as the tubes were closed, ABB contains reducing agents that facilitate the generation of moderate anaerobic conditions, and the study began with an exponential culture. Following this initial increase, the culture started to decline drastically, with the greatest decay rate ($k = 0.89$ days^{-1} and a $T_{90} = 1.12$ days) attributed to increased metabolism in an environment with limited nutrients (Sinton et al., 1999).

3.1.2. pH

Another key factor affecting phage stability is the pH level of the medium. The progressive decline in microbial concentrations over 29 days of exposure to three different pHs (3, 7, and 9) (Fig. 2) was assessed at 4 °C, as at this temperature the reduction in viable microorganisms and impact on stability would be minimal. Linear regression parameters, indicative of reasonable fits with r^2 values ranging from 0.9 to 0.99, along with estimated first-order decay rate constants, are summarized in Supplementary Table 2, with the T_{90} values represented in Table 2.

Exposure to pH 3 caused the largest reductions in all phages, consistent with previous results reported for various phages (Sharp et al., 1946; Kerby et al., 1949; Allué-Guardia et al., 2012; Allué-Guardia

Table 2
T_{90} (days or min) with 95 % confidence intervals (CI) for infectious crAss-like phages, GA17-A and *B. intestinalis*, obtained from the regression lines under all tested disinfection conditions.

Treatment	Value	ΦCrAssBcn6	ΦCrAssBcn10	ΦCrAssBcn15	ΦCrAss001	GA17-A	*B. intestinalis*
Temp (days)	4 °C	s	s	s	s	s	7.05 (6.57 to 7.61)
	22 °C	s	s	s	s	s	2.81 (2.34 to 3.39)[b]
	37 °C	1.95 (NA to 4.62)[a] 29.10 (22.56 to 41.05)[b]	2.91 (NA to 7.36)[a] 28.45 (23.75 to 35.56)[b]	2.64 (1.53 to 7.30)[a] 32.84 (25.13 to 47.37)[b]	1.68 (0.89 to 5.71)[a] 29.02 (19.14 to 60.02)[b]	1.22 (NA to 2.30)[a] 73.86 (38.12 to NA)[b]	Increase 1.12 (1.08 to 1.17)[b]
pH (days)	3	1.82 (1.70 to 1.97)	1.77 (1.60 to 1.99)	1.84 (1.67 to 2.05)	1.8 (1.62 to 2.03)	0.15 (0.14 to 0.15)	0.14
	7	s	s	s	s	s	7.09 (6.59 to 7.68)
	9	s	s	s	s	s	1.09 (1.02 to 0.16)
Chlorination (min)	10 ppm	0.13 (NA to 0.37)	0.06 (NA to 0.19)	0.08 (NA to 0.29)	0.11 (NA to 0.31)	0.09 (NA to 0.18)	0.17 (0.16 to 0.17)
UV (min)		0.18 (9.16 to 0.19)	0.19 (0.17 to 0.21)	0.13	0.44 (0.41 to 0.48)	1.49 (1.36 to 1.64)	1.49 (1.27 to 1.78)
High Temp (min)	60 °C	s	s	s	s	9.54 (8.11 to 11.59)	3.57
	70 °C	21.92 (19.31 to 25.34)	20.61 (18.47 to 23.30)	18.42 (16.46 to 20.90)	21.55 (19.18 to 24.59)	3.77 (3.66 to 3.89)	3.57
Natural Inactivation (days)	Summer	3.44 (3.14 to 3.79)	4.22 (3.58 to 5.14)	4.15 (3.66 to 4.78)	3.65 (3.31 to 4.07)	4.39 (3.79 to 5.20)	0.98 (0.88 to 1.11)
	Winter	19.54 (16.52 to 23.92)	16.89 (15.1 to 19.16)	11.56 (10.16 to 13.42)	14.29 (12.25 to 17.15)	17.34 (15.33 to 19.95)	1.87 (1.67 to 2.12)

NA. Not achieved. s. stable. [a, b] Biphasic results, being [a] Decay rate 1 (k_1) and [b] decay rate 2 (k_2).

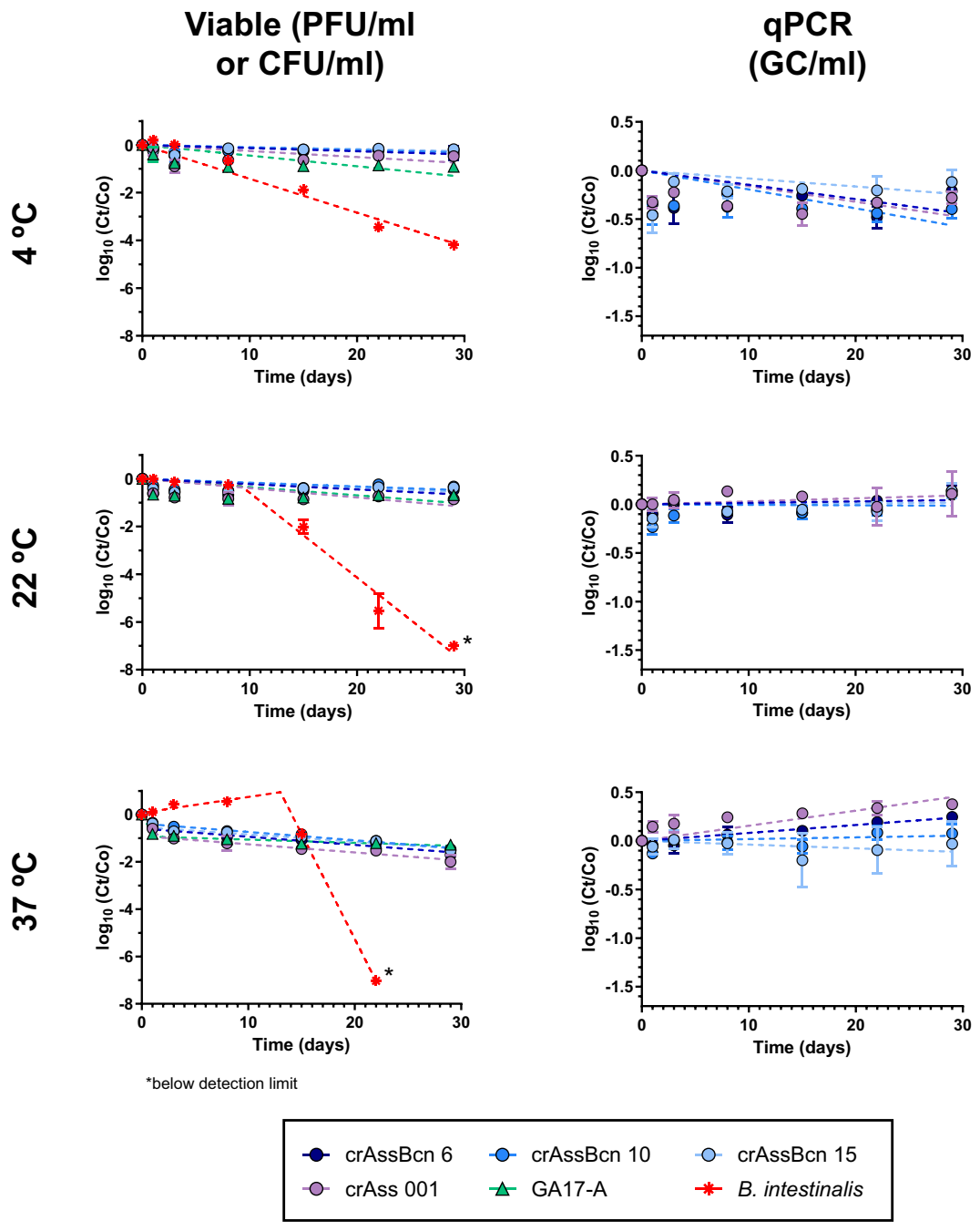


Fig. 1. Stability of infectious crAssBcn phages, ΦCrAss001, GA17-A, and *B. intestinalis* (PFU/ml and CFU/ml) and their genome copies (GC/ml) at different temperatures (4, 22, and 37 °C) after 29 days of storage. Results are presented as the logarithm of the ratio between the concentration at a given time (Ct) and the initial concentration at time 0 (Co), normalized to compare different data sets. Data points indicate the mean values obtained from triplicate experiments. The dashed lines represent the models fitted to the observed data, and the error bars denote the standard deviation. Asterisks indicate values below the detection limit.

et al., 2014). Both crAssBcn phages and ΦcrAss001 were totally inactivated after 15 days of treatment, with decay rates (k) averaging at around 0.5 days^{-1}. Viruses are quantified by counting infectious units, such as PFU, where each unit may consist of either a single viral particle or an aggregate containing multiple viral particles. A decrease in the number of infectious units can result from either inactivation (loss of infectivity) or aggregation. A higher concentration of hydrogen ions (low pH) may promote phage aggregation, especially when the pH is equal to or lower than their isoelectric point (Langlet et al., 2007). Additional experiments aiming at disaggregating viruses showed that when phage suspensions at pH 3 were neutralized, the number of phages did not differ significantly ($p > 0.05$) from the non-neutralized suspensions. This suggested that aggregates were likely not formed by the low pH, and points to viral inactivation as the most plausible cause of the PFU decrease. Viral inactivation can likely be attributed to the damage caused by the highly acidic environment to the capsid proteins, leading to reduced infectivity (Sharp et al., 1946).

At pH 3 the decay of crAss-like phages was lower than that of GA17-A and *B. intestinalis* (Supplementary Table 2) that showed a complete loss of viability after one day of treatment (Fig. 2). Some phages inactivate much faster than others at acidic pH levels; for example, at pH 3 T7 phage became completely inactive after 1 h (Kerby et al., 1949), and the somatic coliphage SOM23 lost its activity on the first day (Allué-Guardia et al., 2012). Despite the substantial reduction in viable particles at pH

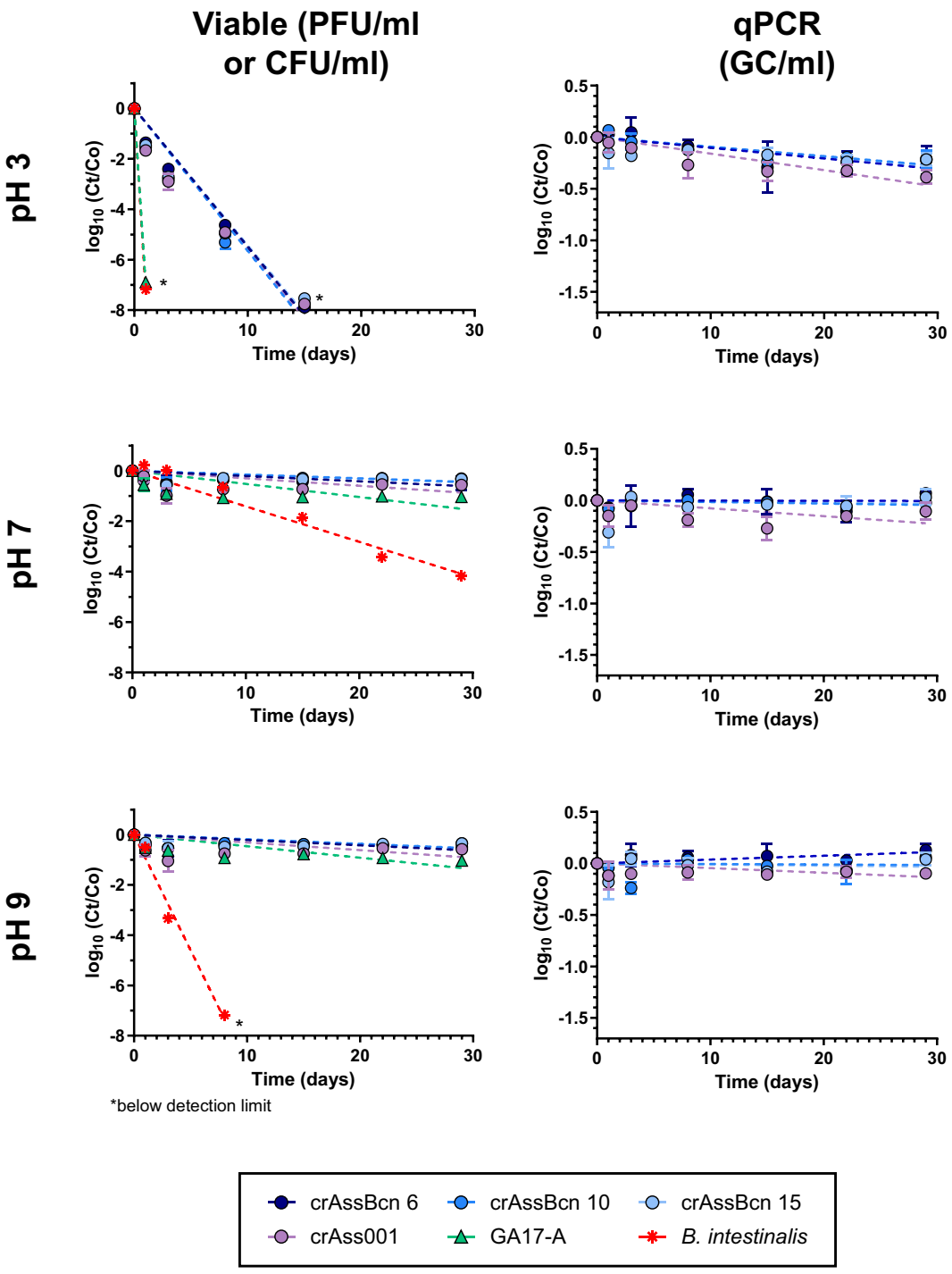


Fig. 2. Influence of different pH treatments (3, 7, and 9) on the persistence of crAssBcn phages, ΦCrAss001, and phage and bacterial controls (GA17-A and *B. intestinalis*, respectively) during 29 days of storage at 4 °C. The charts on the left show logarithmic values of the ratio Ct/Co in PFU/ml or CFU/ml, indicative of viability through time, while on the right, DNA inactivation (decay) measured by qPCR is expressed in GC/ml. Data points indicate the mean values obtained from triplicate experiments. The dashed lines represent the models fitted to the observed data, and the error bars denote the standard deviation. Asterisks indicate values below the detection limit.

3, we observed a non-significant ($p > 0.05$) decrease of phage DNA by qPCR (<0.5 $\log_{10}$) (Fig. 2), in agreement with previous research (Calero-Cáceres and Muniesa, 2016). This could be attributed to the shielding of dsDNA provided by the capsid, which minimizes DNA degradation (Pérez-Illana et al., 2021). In contrast, proteins are highly sensitive to denaturation by acidic pH, which affects phage infectivity.

At pH 7 and pH 9, the infectivity of all phages and the stability of phage DNA remained constant after one month (Fig. 2), with no significant differences among them ($p > 0.05$). Phage inactivation rates are reported to be lower at a pH range of 6 to 9 over long storage periods (Sharp et al., 1946; Kerby et al., 1949; Feng et al., 2003). The decay of *B. intestinalis* at the neutral or alkaline pH was less drastic compared to acidic conditions, but still more pronounced than for phages, showing a reduction of 4.2 $\log_{10}$ after one month at pH 7, and total inactivation within 8 days at pH 9 (Fig. 2). Comparing the bacterial inactivation rate across the three conditions, the slowest decay rate was at pH 7 ($k = 0.14$

days^{-1}), requiring 7.09 days to achieve a 90 % reduction in concentration (Supplementary Table 2). At pH 3 and 9, only 0.14 and 1.09 days (k = 0.92 days^{-1}), respectively, were needed for the same level of reduction. The mechanisms underlying the lower persistence of bacteria compared to phages at low pH are multifaceted. Acidic pH causes damage to nucleic acids, alters membrane permeability at the expense of ATP consumption, induces cellular changes, and leads to protein denaturation or activation of acid-induced proteins, causing structural and metabolic changes in the cell. These effects result in altered molecular composition and impact the overall growth of the bacteria (Lund et al., 2020).

3.2. Inactivation treatments

The three crAssBcn phages, ΦcrAss001, and the two controls, GA17-A and *B. intestinalis*, were subjected to three independent inactivation experiments, using chemical agents (chlorination) and physical agents (UV light treatment and high temperatures).

3.2.1. Chlorination

Chlorination is a widely used method for water disinfection, being the sole method employed for drinking water in some parts of the world. In this study, the microorganisms were treated with a final concentration of 10 ppm of chlorine, recommended for disinfection and reuse of wastewater (United States Environmental Protection Agency, 2012), leading to a significant decrease in the densities of infectious particles. This dosage is to achieve the final concentration of 10 ppm, the amount of sodium hypochlorite applied was adjusted taking into account the reducing effect on free chlorine of a component in the ABB medium. Except for *B. intestinalis*, whose inactivation followed a linear regression model, the decay rates of the phages were calculated using nonlinear exponential kinetics to determine the best-fit line (Supplementary Table 3). This model shows how rapidly the concentration decreased to the asymptote, where values remain constant as time approaches infinity. *B. intestinalis* was completely inactivated within the first minute of treatment (k = 6.06 min^{-1}) (Fig. 3 and Supplementary Table 3). In contrast, after 30 min, crAss-like phages were reduced by 4.61 to 4.77 $\log_{10}$ and GA17-A phage was inactivated by 3.85 $\log_{10}$ units, reaching a stabilization point. This persistence was not observed in a previous study also using 10 ppm of chlorine (Allué-Guardia et al., 2012), where somatic coliphages were completely inactivated after 30 min of chlorine treatment. In agreement with the literature, our study confirms that crAss-like phages can be detected in chlorine-treated waters that meet bacteriological disinfection criteria (Durán et al., 2003).

The time required to achieve 1-log inactivation in all assays ranged between 0.06 and 0.17 min (Table 2). This decay was not observed in phage DNA, which showed a negligible reduction of 0.3 $\log_{10}$ units or less. Studies on the use of 10 ppm of chlorine for inactivation bacteriophages infecting *E. coli* have shown genomic reduction below 0.5 $\log_{10}$ units after 20 min (Allué-Guardia et al., 2014) or <1 $\log_{10}$ unit after 30 min (Allué-Guardia et al., 2012). Viruses are more susceptible to chlorine because they depend on easily damaged components like proteins (capsid) for their function (Zhong et al., 2016). Once these are compromised, the viral genome is exposed and quickly degraded. In contrast, naked DNA is more resistant because it lacks these vulnerable structures, and though chlorine can damage nucleic acids, the overall stability of the DNA molecule itself gives it more resilience (Qiao et al., 2022). Furthermore, the disinfection efficiency of chlorine is directly influenced by the quality and source of the water (Kahler et al., 2010), which may contain varying amounts of substances that deactivate chlorine ions. Although this effect could not be completely ruled out, in the present study we tried to minimize this effect by diluting phages and bacteria in double-distilled water for chlorination experiments.

3.2.2. UV treatment

The UV fluences at 10 cm from the UV lamp were estimated to be 5.94, 29.7, 59.4, and 178.2 mJ/cm^2 for each exposure time of 1, 5, 10, and 30 min, respectively (Fig. 3). Infectious microorganisms exhibited a nonlinear exponential decay profile in 5 out of the 6 samples, with ΦcrAssBcn15 following a linear regression model (Fig. 3 and Supplementary Table 3). Although this type of irradiation cannot be directly compared to solar UV light, especially as most germicide lamps produce only UV-C, it has been demonstrated that UV irradiation affects the capsid by breaking down its proteins, thereby hindering viral genome injection and reducing infectivity (Wigginton et al. (2012). The crAssBcn and ΦcrAss001 phages were significantly ($p < 0.05$) more sensitive to UV treatment than GA17-A and *B. intestinalis*, achieving complete inactivation after just 5 min of irradiation (Fig. 3), a faster rate compared to somatic coliphages, which need up to 59.4 mJ/cm^2 (after 10 min of treatment) to be completely inactivated (Calero-Cáceres and Muniesa, 2016). The decay rate constants varied among the samples: k values were 0.73–0.79 min^{-1} for ΦCrAssBcn6 and 10; 7.56 min^{-1} for ΦCrAssBcn15, and 0.23 min^{-1} for ΦCrAss001 (Supplementary Table 3), with T_{90} values ranging from 0.13 to 0.44 min (Table 2). GA17-A and *B. intestinalis* required 30 min to be completely inactivated (k = 0.1 min^{-1}, T_{90} = 1.49 min and k = 0.08 min^{-1}, T_{90} = 1.49 min, respectively).

The different inactivation rates of crAss-like phages and GA17-A could be attributed to differences in morphology, as crAss-like phages are podoviruses (Ramos-Barbero et al., 2023), whereas GA17-A is a siphovirus (data not shown), like the majority of virulent phages infecting *Bacteroides* (Queralt et al., 2003). Differences can also be attributed to genome size; crAss-like viruses have a relatively large double-stranded circular DNA genome (~100 kb), which presents a larger target for UV light inactivation. In contrast, GA17-A, though not yet sequenced, is a *B. fragilis*-infecting siphovirus, which typically have genomes of 45 kb (Tariq and Carding, 2021; Bakuradze et al., 2023). Furthermore, as reported by Sun et al. (2023), the GC content can also affect the bactericidal effect of UV radiation.

B. intestinalis is more resistant to UV irradiation than other bacteria, such as *E. coli* (Calero-Cáceres and Muniesa, 2016), which could be due to the presence of capsular polysaccharides. Akhlaq et al. (1990) observed that UV light at 250 J/m^2 had barely any degradative effect on capsular polysaccharides, supporting the findings of Nunez et al. (2023), who demonstrated that bacterial encapsulation conferred greater resistance to UV irradiation.

Unlike the culturable microorganisms, all four phages exhibited linear DNA decay under UV treatment (Fig. 3 and Table 3). UV light was the only disinfectant agent observed to have a significant impact on phage nucleic acids ($p < 0.05$), although the reduction in DNA concentration was slight, with a decline of 0.95 to 1.65 $\log_{10}$ units after 30 min of irradiation. Consequently, these samples exhibited low DNA inactivation rates (k = 0.035–0.059 min^{-1}, T_{90} = 16.95–28.38 min) (Table 3 and Supplementary Table 4), supporting the hypothesis that in phages UV light primarily damages the capsid, and high irradiation levels are required to achieve significant DNA inactivation, that should be protected by the capsid against UV exposure (Pecson et al., 2009; McKinney and Pruden, 2012). Nevertheless, the DNA decay was greater compared to other disinfectants (Wigginton et al., 2012). Nucleic acid removal and disinfection efficiency could be improved by applying a sequential treatment of UV irradiation followed by sodium hypochlorite (Shang et al., 2007; Zhang et al., 2015).

3.2.3. Thermal treatment

Thermal treatments are among the most commonly used methods for eliminating microorganisms in water, food, and sludge (Mocé-Llivina et al., 2003; Nuanualsuwan and Cliver, 2003; Guglielmotti et al., 2012; Peng et al., 2017). These treatments induce morphological changes in phages, such as the release of genetic material, aggregation of phage tails, or even structural breakdown, where the head separates from the tail (Atamer et al., 2009). High-temperature inactivation assays were conducted at 60 and 70 °C, representing the lowest and optimal

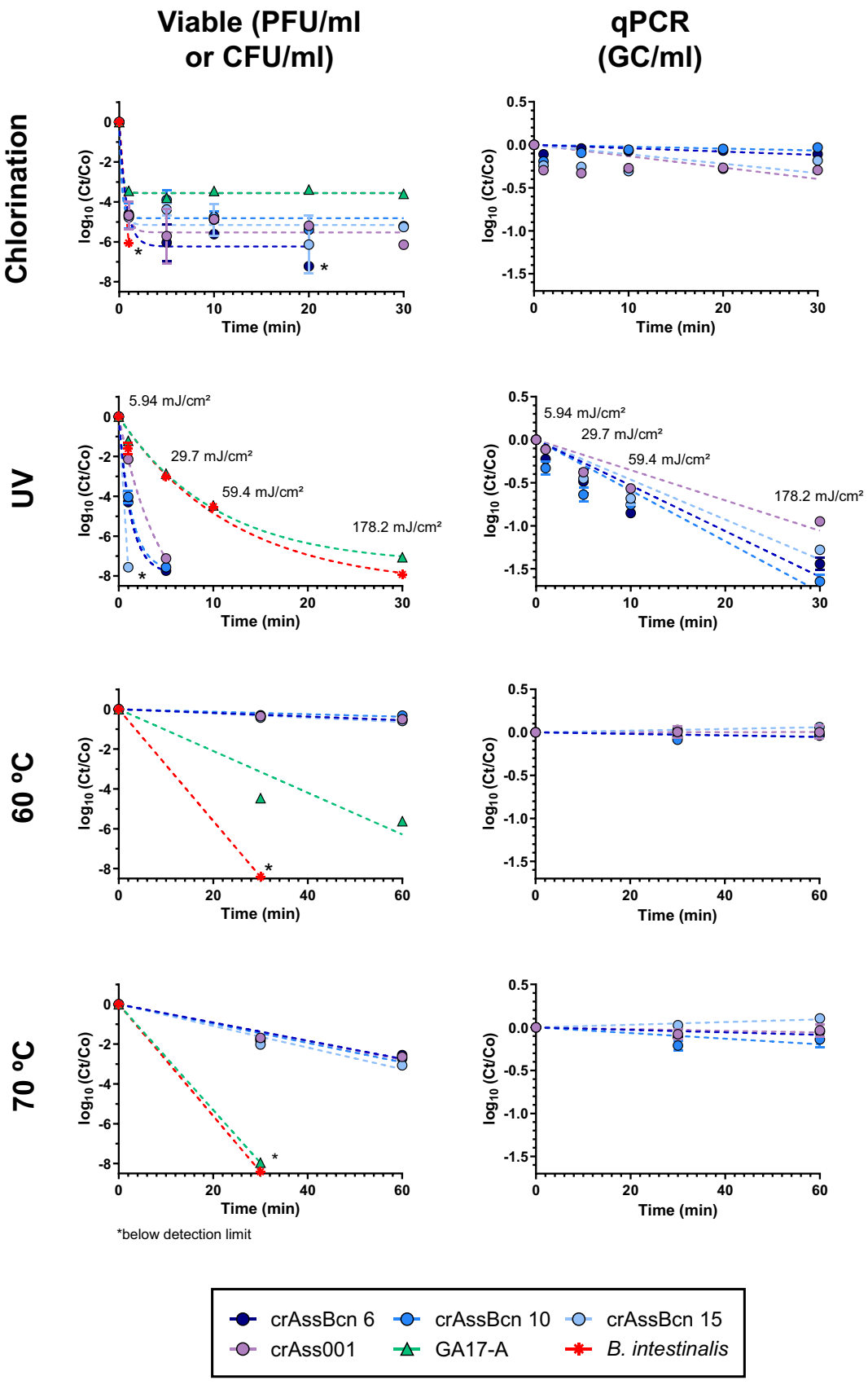


Fig. 3. Comparative inactivation (decay) rates of crAssBcn phages, ΦCrAss001, GA17-A phage, and *B. intestinalis* (infectivity and DNA), after 60 min of various treatments: chlorination, UV irradiation, and high temperatures (60 and 70 °C). Data points indicate the mean values obtained from triplicate experiments. The dashed lines represent the models fitted to the observed data, and the error bars denote the standard deviation. Asterisks indicate values below the detection limit.

Table 3
T_{90} (days or min) with 95 % confidence intervals (CI) for crAss-like phages DNA, obtained from the regression lines under natural inactivation in summer and UV light treatment.

Microorganism	UV inactivation (min)	Summer inactivation (days)
ΦCrAssBcn6	18.85 (16.10 to 22.74)	9.36 (7.81 to 11.70)
ΦCrAssBcn10	16.95 (14.63 to 20.15)	7.88 (6.41 to 10.24)
ΦCrAssBcn15	21.56 (18.91 to 25.08)	6.89 (6.09 to 7.94)
ΦCrAss001	28.38 (24.28 to 34.15)	7.77 (6.40 to 9.88)

temperatures used for pasteurization, respectively. Over a 60-minute period, all samples exhibited a linear decay profile (Fig. 3; Table 2; Supplementary Table 5). At both temperatures, *B. intestinalis* was completely inactivated within 30 min, demonstrating the highest decay rate ($k = 0.28\ min^{-1}$, $T_{90} = 3.57$ min) (Supplementary Table 5; Table 2). The phages displayed a lower decay rate under these conditions, indicating greater persistence compared to their bacterial host. Exposure of crAssBcn and ΦCrAss001 phages to 60 °C resulted in stable infectious particle titers, with <0.5 log_{10} units of reduction. However, at 70 °C, there was a reduction of 2.5 to 3 log_{10} units in viable particles and a higher decay rate ($k = 0.05\ min^{-1}$, $T_{90} = 18.42–21.92$ min) (Supplementary Table 5; Table 2), attributed to the high temperature causing severe damage to the viral capsid (Nuanualsuwan and Cliver, 2003; Wigginton et al., 2012). These reductions are comparable to those previously reported for somatic coliphages, which showed minimal decay at 60 °C, but a decrease of 2 to 2.5 log_{10} at 70–80 °C (Mocé-Llivina et al., 2003; Cunault et al., 2011; Calero-Cáceres and Muniesa, 2016).

In contrast, the *Bacteroides*-infecting phage GA17-A, a siphovirus, was completely inactivated at 70 °C within just 30 min, reaching values below the detection limit (3.77 min for a 90 % reduction). At 60 °C, the inactivation of GA17-A was lower (5.6 log_{10}), but still higher compared to the crAss-like viruses, with a T_{90} value of 9.54 min (Table 2). The crAss-like phages were much less sensitive to high-temperature treatments than GA17-A, possibly due to greater stability of packaged phage DNA or overall particle stability (Atamer et al., 2009), potentially related to differences in capsid structure. The differences in thermal treatment resistance between crAss-like phages, *Bacteroides*-infecting phage GA17-A, and other previously tested phages such as B40-8 and RYC2056 (which infect *B. fragilis* strains and also are siphoviruses) (Mocé-Llivina et al., 2003), suggest that phages infecting *Bacteroides* are a heterogeneous group in terms of thermal resistance.

The crAss-like phage DNA remained stable throughout the assay, with reductions of <0.2 log_{10} units under both conditions and at all-time intervals (Fig. 3). This stability indicates that the viral DNA remained intact within the phage capsid during the treatment, as demonstrated in previous studies (Nuanualsuwan and Cliver, 2003; Pfaender et al., 2015).

3.3. Natural inactivation

To evaluate the impact of natural conditions on the three crAssBcn phages, ΦCrAss001, GA17-A, and the host *B. intestinalis*, an outdoor mesocosm experiment was conducted during summer and winter seasons. Dialysis membranes were used for their beneficial characteristics in mimicking natural conditions. Their pore size (>12–14 kDa) prevents the diffusion of nutrients and larger microorganisms, while they allow water and molecules smaller than 14 kDa to pass (Maraccini et al., 2016). Additionally, being transparent allows the transmission of light.

Biotic parameters, such as grazing by protozooplankton, were not measured in our experiment because the use of dialysis membranes prevents contact with particles and microorganisms larger than 12–14 kDa. Grazing significantly affects the survival of *Bacteroides* species, particularly those of fecal origin (Gonzalez et al., 1992; Menon et al., 2003; Bell et al., 2009), although its effect on viruses is not well understood. Finally, the regrowth of the bacterial strain would be negligible as the anaerobic metabolism of *Bacteroides* prevents its environmental growth.

The most variable factors in the mesocosm were temperature (both environmental and water) and solar irradiation (Fig. 4). Other factors, such as salinity, oxygen concentration, and pH, remained stable throughout the experiment. All results were represented using a linear regression model. The winter assays lasted 45 days compared to a maximum of 24 days for the assays conducted in summer, when the dialysis membranes containing the samples deteriorated faster.

Overall, the reduction in viable microorganisms for all samples was significantly ($p < 0.05$) greater in summer than in winter (Supplementary Table 6), likely because higher temperatures and solar irradiation lead to increased inactivation (Sinton et al., 1999; Maraccini et al., 2016). Since solar irradiation is a key factor in natural disinfection on the surface of aquatic environments (Sinton et al., 1999), we propose this factor contributes more to the decay of the microorganisms than water temperature that did not raise >20 °C along the summer experiment, and synergistic effects of photoinactivation and temperatures occurs above 30 °C (Nelson et al., 2018). As the samples were positioned approximately 10 cm below the surface, where turbidity is much lower than at greater depths, UV rays were strong enough to cause notable photo-biological damage (Nelson et al., 2018).

During the summer period, *B. intestinalis* was the most affected microorganism, declining below the detection limit after eight days ($k = 1.02\ days^{-1}$, $T_{90} = 0.98$ days) (Supplementary Table 7; Table 2). However, the phages, including crAss-like phages and the control GA17-A, experienced a notable reduction in viable particles, though not as drastic as the bacterial host. All phages followed a similar trend, with reductions between 5.6 and 6.5 log_{10} units after 24 days (Fig. 4). First-order decay rates ranged from 0.22 to 0.29 $days^{-1}$, and T_{90} values from 3.44 to 4.39 days (Supplementary Table 7; Table 2).

Genomic copy number evaluation by qPCR showed a smaller yet significant ($p < 0.05$) reduction of up to 3 log_{10} GC/ml, with estimated decay rates between 0.11 and 0.15 $days^{-1}$ ($T_{90} = 6.89$-9.36 days) (Fig. 4; Table 3; Supplementary Table 4). This suggests that factors such as solar irradiation may damage the phage capsids, leading to the degradation of free DNA molecules.

During the winter season, the inactivation of infectious microorganisms was less pronounced. Nonetheless, as observed in summer, the bacteria were significantly ($p < 0.05$) more affected than the other samples (Supplementary Table 6), with levels reduced to below the detection limit after 15 days ($k = 0.54\ days^{-1}$, $T_{90} = 1.87$ days) (Table 2; Supplementary Table 7). This reinforces previous observations that *Bacteroides* species are more rapidly inactivated in the environment compared to fecal coliforms and enterococci, suggesting they could be categorized as indicators of recent contamination, as noted by Ballesté and Blanch (2010). The crAss-like phages and GA17-A exhibited a similar pattern, with a significantly ($p < 0.05$) lower reduction in winter, ranging from 3.2 to 4.4 log_{10} and 2.4 log_{10} PFU/ml, respectively (Fig. 4). Their inactivation slopes ranged from 0.05 to 0.09 $days^{-1}$, and T_{90} values ranged from 11.56 to 19.54 days (Supplementary Table 7; Table 2).

Again, the phages evaluated by qPCR showed a very subtle reduction in GC, equal to or <1 log_{10} units, with a stable decay rate (Fig. 4). As this is the first study on the natural inactivation of isolated crAss-like phage virions, the results were compared with previous research based on the analysis of phage genomes. Ahmed et al. (2021) observed greater inactivation of crAss-like phages at 25 °C (representative of summer temperatures), the values decreasing to below the quantification limit (approximately 5 log_{10}) by 21 days, compared to 15 °C (typical of winter), where reductions were <4 log_{10} by the end of the trial. The persistence values obtained in the present study with isolated virions are higher than those previously found for p-crAssphage using real-time PCR assays (García-Aljaro et al., 2017; Ballesté et al., 2019) with a T_{90} of 2.41 days. Greaves et al. (2020) reported mean k values of 1.01 $days^{-1}$, indicating a steeper decay slope compared to crAssBcn phages.

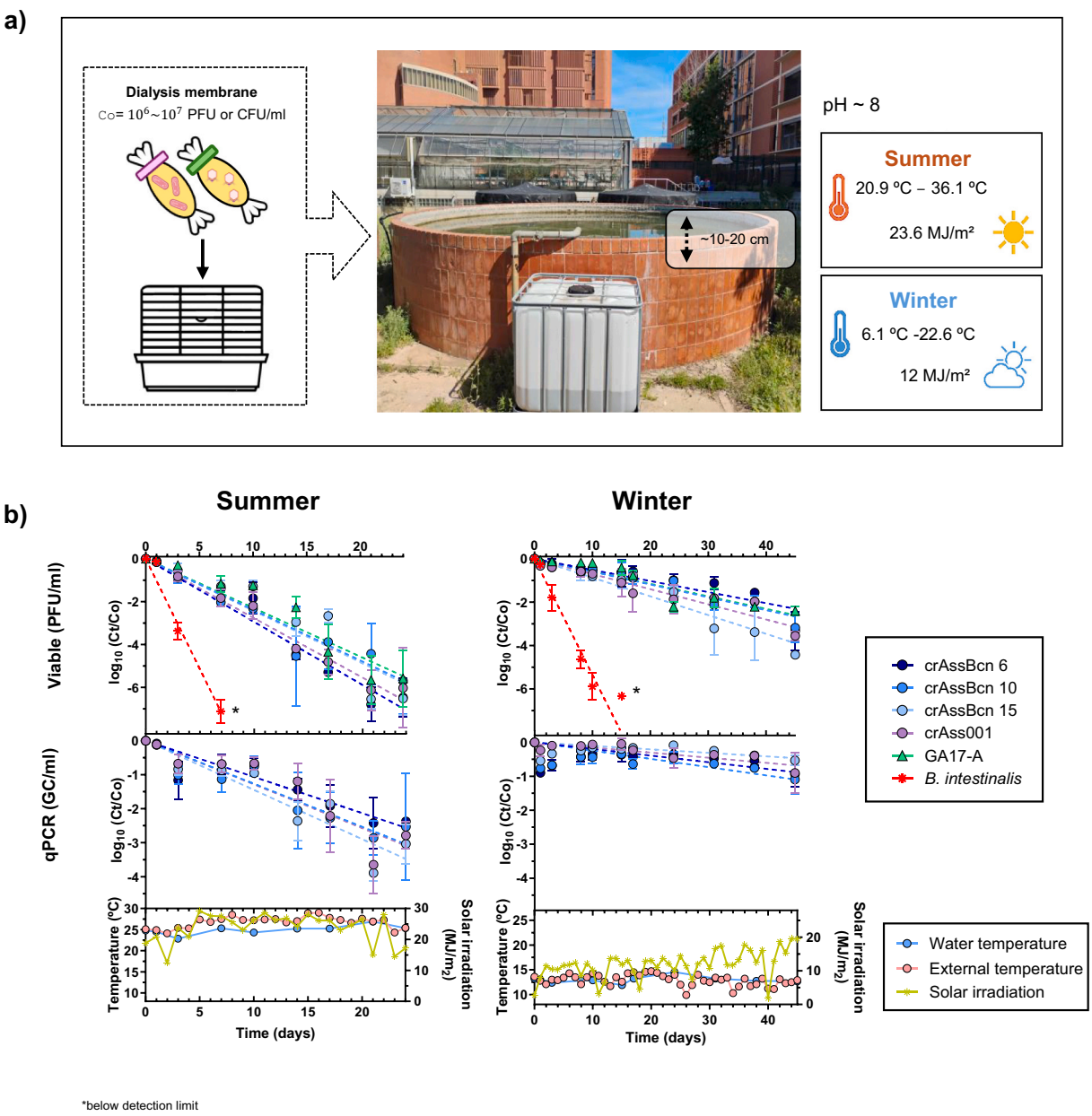


Fig. 4. Natural inactivation of crAssBcn phages, ΦCrAss001, GA17-A, and *B. intestinalis* (PFU/ml and CFU/ml) and their genome copies (GC/ml) assessed by an *in situ* mesocosm experiment in summer and winter. a) Illustration of the experimental setup in the water pond, showing the average pH, temperature, and irradiation. b) Charts representing the inactivation (decay) of crAssBcn phages, ΦCrAss001, GA17-A, and *B. intestinalis* (infectivity and DNA): 24 days in summer (on the left) and 45 days in winter (on the right). Variations in water, external temperatures, and solar irradiation during the trials are also represented at the bottom part of the charts. Logarithmic reduction is measured as the difference between the logarithm of PFU, CFU, or GC/ml at each time point (Ct) and the logarithm at the beginning of the experiment (Co). Data points indicate the mean values obtained from triplicate experiments. The dashed lines represent the models fitted to the observed data, and the error bars denote the standard deviation. Asterisks indicate values below the detection limit.

These comparisons suggest that crAssBcn phages show higher persistence compared to p-crAssphage detected by qPCR (Stachler et al., 2017). The differences could be because crAssBcn phages are more persistent in the environment than p-crAssphage (or similar phages detected by the qPCR assay (Stachler et al., 2017)) or, more likely, they may reflect variability in the efficiency and sensitivity of the molecular methods used.

Comparing the persistence values obtained for somatic coliphages (Ballesté et al., 2019) with those of cultivable crAssBcn phages isolated in Barcelona, both in summer (T_{90} of 4.84 *versus*. an average of 3.93, respectively) and winter (T_{90} of 9.9 *versus* 15.99, respectively), suggests they share a similar environmental persistence, at least during the warmer season.

MST uses microbial indicators specific to the gastrointestinal tract of certain animal species to identify sources of fecal contamination in water bodies. Viral markers are particularly effective as they provide a more accurate reflection of pathogenic viruses in contaminated waters. Bacteriophages, specifically those infecting species of the genus *Bacteroides*, are considered promising indicators due to the high specificity of their bacterial host to the human and animal gut and their inability to replicate in the environment (Jofre et al., 2014). Phages infecting *B. fragilis* strain HSP40 (Jofre et al., 1986), *B. fragilis* strain GB-124 (Ebdon et al., 2007; Mcminn et al., 2014), or *B. thetaiotaomicron* strain GA17 (Payan et al., 2005) have been evaluated as potential human markers. However, limitations such as low densities in wastewater or the lack of available qPCR detection methods have prompted the search for

alternative viral markers.

p-crAssphage, discovered through metagenomic tools (Dutilh et al., 2014), has emerged as a promising MST marker because of its high prevalence in human feces and sewage (Crank et al., 2020). Studies have shown that p-crAssphage can indicate human fecal contamination, is found in high concentrations, and is more resilient to environmental degradation/inactivation in water compared to other bacterial and viral markers (Stachler and Bibby, 2014; Ballesté et al., 2019; Greaves et al., 2020; Ahmed et al., 2021; Zhang et al., 2021). Additionally, p-crAssphage is more resistant to inactivation processes and environmental stressors than human viruses, making it an effective indicator of viral persistence in contaminated waters (Wu et al., 2020; Hamza and Abd-Elmaksoud, 2023). However, even though five different virions have been isolated so far (Shkoporov et al., 2018; Hryckowian et al., 2020; Guerin et al., 2021; Papudeshi et al., 2023), most crAss-like phage persistence assays have been conducted using molecular methods, generating data that do not correlate with the viability of viral particles (Aw and Rose, 2012; Greaves et al., 2020; Wu et al., 2020). Since only viable viruses, not nucleic acids residues, can cause disease, this limitation poses a significant concern for public health risk assessment. The major limitation in the environmental water application of p-crAssphage is that the qPCR signal remains stable after standard disinfection methods (Botes et al., 2013).

High specificity and abundance in human feces mean that p-crAssphage is more reliable for detecting human fecal pollution in environmental waters than traditional markers. Numerous studies have documented the correlation between p-crAssphage and various waterborne diseases caused by viruses such as norovirus, polyomavirus, or adenovirus (Crank et al., 2019; Farkas et al., 2019; Wu et al., 2020; Hamza and Abd-Elmaksoud, 2023). However, assessing the persistence of crAss-like phages, a crucial trait for an effective microbial fecal indicator, requires isolating both the viruses and their hosts.

Based on the extensive information provided here about the persistence and inactivation of crAss-like phages in comparison with other bacterial and viral indicators such as *Bacteroides*-infecting phages, it can be concluded that crAss-like phages, particularly crAssBcn, show promising characteristics as MST viral markers for monitoring water quality and safety. They exhibited a high stability and resistance to different disinfection processes and under conditions mimicking those of the natural environment. The persistence of crAss-like phages in the environment aligns with that of other phages used as indicators and is, in general, equal to or greater than that of pathogenic viruses, particularly enveloped ones (Grabow et al., 1983; Mocé-Llivina et al., 2005; Bertrand et al., 2012; Emmoth et al., 2017). Viral indicators serve as effective tools for detecting human fecal pollution; however, they do not always correlate with the presence of fecal pathogens. The correlation between phages and enteric viruses is comparable to the correlation among various enteric viruses (Jofre et al., 2016). In addition, since phages are usually found in larger numbers in the environment, they are easier to detect in fecal pollution than viral pathogens, which may be absent. Continued research on crAss-like phage behavior in complex environmental matrices will further validate their utility and refine their application in water quality assessments. Challenges include the variability in decay rates due to genetic diversity among different crAss-like phage strains, comparison with other enteric viruses, and the need for consistent methodologies to accurately assess their persistence, integrating both molecular and infectivity assays (Zhang et al., 2019).

CRediT authorship contribution statement

Clara Gómez-Gómez: Writing – original draft, Investigation, Formal analysis. **Maria Dolores Ramos-Barbero:** Writing – review & editing, Software, Data curation. **Laura Sala-Comorera:** Writing – review & editing, Visualization, Investigation. **Sara Morales-Cortes:** Writing – review & editing, Investigation, Formal analysis. **Gloria Vique:** Methodology, Investigation. **Cristina García-Aljaro:** Writing – review & editing, Resources, Conceptualization. **Maite Muniesa:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgements

This research was funded by grant number PID2020-113355GB-I00 funded by MICIU/AEI/10.13039/501100011033 and by "ERDF A way of making Europe". C. G-G has a fellowship from the University of Barcelona. L. S-C has a Beatriu de Pinos grant from the Government of Catalonia. S. M-C has a grant from Colciencias (Republic of Colombia). G. V. has a grant PRE2021-100745 funded by MICIU/AEI/10.13039/501100011033 and by "ESF+". M. M. is a researcher of the ICREA Academia 2022 program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.177450.

Data availability

Data will be made available on request.

References

Adams, M.H., 1959. Bacteriophages. New York, Interscience Publishers, New York.
Ahmed, W., Payyappat, S., Cassidy, M., Besley, C., Power, K., 2018. Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. Water Res. 145, 769–778. https://doi.org/10.1016/j.watres.2018.08.049.
Ahmed, W., Toze, S., Veal, C., Fisher, P., Zhang, Q., Zhu, Z., Staley, C., Sadowsky, M.J., 2021. Comparative decay of culturable faecal indicator bacteria, microbial source tracking marker genes, and enteric pathogens in laboratory microcosms that mimic a sub-tropical environment. Sci. Total Environ. 751, 141475. https://doi.org/10.1016/j.scitotenv.2020.141475.
Akhlaq, M.S., Schuchmann, H.P., Von Sonntag, C., 1990. Degradation of the polysaccharide alginic acid: a comparison of the effects of UV light and ozone. Environ. Sci. Technol. 24, 379–383.
Allué-Guardia, A., Jofre, J., Muniesa, M., 2012. Stability and infectivity of cytolethal distending toxin type V gene-carrying bacteriophages in a water mesocosm and under different inactivation conditions. Appl. Environ. Microbiol. 78, 5818–5823. https://doi.org/10.1128/AEM.00997-12.
Allué-Guardia, A., Martínez-Castillo, A., Muniesa, M., 2014. Persistence of infectious shiga toxin-encoding bacteriophages after disinfection treatments. Appl. Environ. Microbiol. 80, 2142–2149. https://doi.org/10.1128/AEM.04006-13.
Araujo, R., Muniesa, M., Méndez, J., Puig, A., Queralt, N., Lucena, F., Jofre, J., 2001. Optimisation and standardisation of a method for detecting and enumerating bacteriophages infecting *Bacteroides fragilis*. J. Virol. Methods 93, 127–136.
Atamer, Z., Dietrich, J., Müller-Merbach, M., Neve, H., Heller, K.J., Hinrichs, J., 2009. Screening for and characterization of *Lactococcus lactis* bacteriophages with high thermal resistance. Int. Dairy J. 19, 228–235. https://doi.org/10.1016/j.idairyj.2008.10.012.
Aw, T.G., Rose, J.B., 2012. Detection of pathogens in water: from phylochips to qPCR to pyrosequencing. Curr. Opin. Biotechnol. 23, 422–430. https://doi.org/10.1016/j.copbio.2011.11.016.
Bakuradze, N., Merabishvili, M., Kusradze, I., Ceyssens, P.J., Onsea, J., Metsemakers, W. J., Grdzelishvili, N., Natroshvili, G., Tatrishvili, T., Lazvliashvili, D., Mitskevich, N., Pirnay, J.P., Chanishvili, N., 2023. Characterization of a bacteriophage GEC_vB_Bfr_UZM3 active against *Bacteroides fragilis*. Viruses 15, 1042. https://doi.org/10.3390/v15051042.
Ballesté, E., Blanch, A.R., 2010. Persistence of *Bacteroides* species populations in a river as measured by molecular and culture techniques. Appl. Environ. Microbiol. 76, 7608–7616. https://doi.org/10.1128/AEM.00883-10.
Ballesté, E., Pascual-Benito, M., Martín-Díaz, J., Blanch, A.R., Lucena, F., Muniesa, M., Jofre, J., García-Aljaro, C., 2019. Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. Water Res. 155, 233–244. https://doi.org/10.1016/j.watres.2019.02.042.
Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., Sayler, G.S., 2009. Factors influencing the persistence of fecal *Bacteroides* in stream water. J. Environ. Qual. 38, 1224–1232. https://doi.org/10.2134/jeq2008.0258.

Bertrand, I., Schijven, J.F., Sánchez, G., Wyn-Jones, P., Ottoson, J., Morin, T., Muscillo, M., Verani, M., Nasser, A., de Roda Husman, A.M., Myrmel, M., Sellwood, J., Cook, N., Gantzer, C., 2012. The impact of temperature on the inactivation of enteric viruses in food and water: a review. J. Appl. Microbiol. 112 (6), 1059–1074.

Blanch, A.R., Belanche-Muñ Oz, L., Bonjoch, X., Ebdon, J., Gantzer, C., Lucena, F., Ottoson, J., Kourtis, C., Iversen, A., Kü, I., Moce, L., Muniesa, M., Schwartzbrod, J., Skraber, S., Papageorgiou, G., Taylor, H.D., Wallis, J., Jofre, J., 2004. Tracking the origin of faecal pollution in surface water: an ongoing project within the European Union research programme. J. Water Health 2, 249–260.

Botes, M., De Kwaadsteniet, M., Cloete, T.E., 2013. Application of quantitative PCR for the detection of microorganisms in water. Anal. Bioanal. Chem. 405, 91–108. https://doi.org/10.1007/s00216-012-6399-3.

Brouwer, A.F., Eisenberg, M.C., Remais, J.V., Collender, P.A., Meza, R., Eisenberg, J.N.S., 2017. Modeling biphasic environmental decay of pathogens and implications for risk analysis. Environ. Sci. Technol. 51, 2186–2196. https://doi.org/10.1021/acs.est.6b04030.

Calero-Cáceres, W., Muniesa, M., 2016. Persistence of naturally occurring antibiotic resistance genes in the bacteria and bacteriophage fractions of wastewater. Water Res. 95, 11–18. https://doi.org/10.1016/j.watres.2016.03.006.

Cerf, 1977. Tailing of survival curves of bacterial spores. J. Appl. Bacteriol. 42, 1–19.

Chick, H., 1908. An investigation of the laws of disinfection. J. Hyg. 8, 92–158. https://doi.org/10.1017/S0022172400006987.

Crank, K., Petersen, S., Bibby, K., 2019. Quantitative microbial risk assessment of swimming in sewage impacted waters using crAssphage and pepper mild mottle virus in a customizable model. Environ. Sci. Technol. Lett. 6, 571–577. https://doi.org/10.1021/acs.estlett.9b00468.

Crank, K., Li, X., North, D., Ferraro, G.B., Iaconelli, M., Mancini, P., La Rosa, G., Bibby, K., 2020. CrAssphage abundance and correlation with molecular viral markers in Italian wastewater. Water Res. 184, 116161. https://doi.org/10.1016/j.watres.2020.116161.

Cunault, C., Pourcher, A.M., Burton, C.H., 2011. Using temperature and time criteria to control the effectiveness of continuous thermal sanitation of piggery effluent in terms of set microbial indicators. J. Appl. Microbiol. 111, 1492–1504. https://doi.org/10.1111/j.1365-2672.2011.05144.x.

Current ICTV Taxonomy Release | ICTV, 2023. Current ICTV Taxonomy Release | ICTV [WWW Document]. https://ictv.global/taxonomy. URL. https://ictv.global/taxonomy (accessed 7.2.24).

Durán, A.E., Muniesa, M., Mocé-Llivina, L., Campos, C., Jofre, J., Lucena, F., 2003. Usefulness of different groups of bacteriophages as model micro-organisms for evaluating chlorination. J. Appl. Microbiol. 95, 29–37.

Dutilh, B.E., Cassman, N., McNair, K., Sanchez, S.E., Silva, G.G.Z., Boling, L., Barr, J.J., Speth, D.R., Seguritan, V., Aziz, R.K., Felts, B., Dinsdale, E.A., Mokili, J.L., Edwards, R.A., 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat. Commun. 5, 4498. https://doi.org/10.1038/ncomms5498.

Ebdon, J., Muniesa, M., Taylor, H., 2007. The application of a recently isolated strain of *Bacteroides* (GB-124) to identify human sources of faecal pollution in a temperate river catchment. Water Res. 41, 3683–3690. https://doi.org/10.1016/j.watres.2006.12.020.

Edwards, R.A., Vega, A.A., Norman, H.M., Ohaeri, M., Levi, K., Dinsdale, E.A., Cinek, O., Aziz, R.K., McNair, K., Barr, J.J., Bibby, K., Brouns, S.J.J., Cazares, A., de Jonge, P. A., Desnues, C., Díaz Muñoz, S.L., Fineran, P.C., Kurilshikov, A., Lavigne, R., Mazankova, K., McCarthy, D.T., Nobrega, F.L., Reyes Muñoz, A., Tapia, G., Trefault, N., Tyakht, A.V., Vinuesa, P., Wagemans, J., Zhernakova, A., Aarestrup, F. M., Ahmadov, G., Alassaf, A., Anton, J., Asangba, A., Billings, E.K., Cantu, V.A., Carlton, J.M., Cazares, D., Cho, G.S., Condeff, T., Cortés, P., Cranfield, M., Cuevas, D. A., De la Iglesia, R., Decewicz, P., Doane, M.P., Dominy, N.J., Dziewit, L., Elwasila, B.M., Eren, A.M., Franz, C., Fu, J., Garcia-Aljaro, C., Ghedin, E., Gulino, K. M., Haggerty, J.M., Head, S.R., Hendriksen, R.S., Hill, C., Hyöty, H., Ilina, E.N., Irwin, M.T., Jeffries, T.C., Jofre, J., Junge, R.E., Kelley, S.T., Khan Mirzaei, M., Kowalewski, M., Kumaresan, D., Leigh, S.R., Lipson, D., Lisitsyna, E.S., Llagostera, M., Maritz, J.M., Marr, L.C., McCann, A., Molshanski-Mor, S., Monteiro, S., Moreira-Grez, B., Morris, M., Mugisha, L., Muniesa, M., Neve, H., Nguyen, N. phuong, Nigro, O.D., Nilsson, A.S., O'Connell, T., Odeh, R., Oliver, A., Piuri, M., Prussin, A.J., Qimron, U., Quan, Z.X., Rainetova, P., Ramírez-Rojas, A., Raya, R., Reasor, K., Rice, G.A.O., Rossi, A., Santos, R., Shimashita, J., Stachler, E.N., Stene, L.C., Strain, R., Stumpf, R., Torres, P.J., Twaddle, A., Ugochi Ibekwe, M.A., Villagra, N., Wandro, S., White, B., Whiteley, A., Whiteson, K.L., Wijmenga, C., Zambrano, M.M., Zschach, H., Dutilh, B.E., 2019. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. Nat. Microbiol. 4, 1727–1736. https://doi.org/10.1038/s41564-019-0494-6.

Emmoth, E., Rovira, J., Rajkovic, A., Corcuera, E., Wilches Pérez, D., Dergel, I., Ottoson, J.R., Widén, F., 2017. Inactivation of viruses and bacteriophages as models for swine Hepatitis E virus in food matrices. Food Environ Virol 9 (1), 20–34. https://doi.org/10.1007/s12560-016-9268-y.

Farkas, K., Adriaenssens, E.M., Walker, D.I., McDonald, J.E., Malham, S.K., Jones, D.L., 2019. Critical evaluation of crAssphage as a molecular marker for human-derived wastewater contamination in the aquatic environment. Food Environ Virol 11, 113–119. https://doi.org/10.1007/s12560-019-09369-1.

Farnleitner, A., Blanch, A., 2017. Microbial indicators - "workhorses" in the field of health-related water quality testing. In: Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project). https://doi.org/10.14321/waterpathogens.5.

Feng, Y.Y., Ong, S.L., Hu, J.Y., Tan, X.L., Ng, W.J., 2003. Effects of pH and temperature on the survival of coliphages MS2 and Qβ. J. Ind. Microbiol. Biotechnol. 30, 549–552. https://doi.org/10.1007/s10295-003-0080-y.

Fritz, M.S., 2014. An exponential decay model for mediation. Prev. Sci. 15, 611–622. https://doi.org/10.1007/s11121-013-0390-x.

García-Aljaro, C., Ballesté, E., Muniesa, M., Jofre, J., 2017. Determination of crAssphage in water samples and applicability for tracking human faecal pollution. J. Microbial. Biotechnol. 10, 1775–1780. https://doi.org/10.1111/1751-7915.12841.

Gonzalez, J.M., Iriberri, J., Egea, L., Barcina, I., 1992. Characterization of culturability, protistan grazing, and death of enteric bacteria in aquatic ecosystems. Appl. Environ. Microbiol. 58, 998–1004.

Grabow, W.O.K., Gauss-Muller, V., Prozesky, O.W., Deinhardt, F., 1983. Inactivation of Hepatitis A virus and indicator organisms in water by free chlorine residuals. Appl. Environ. Microbiol. 46 (3), 619–624.

Grabow, W.O.K., Neubrech, T.E., Holtzhausen, C.S., Jofre, J., 1995. *Bacteroides fragilis* and *Escherichia coli* bacteriophages: excretion by humans and animals. Wat Sci Technol 31 (5–6), 223–230.

Greaves, J., Stone, D., Wu, Z., Bibby, K., 2020. Persistence of emerging viral fecal indicators in large-scale freshwater mesocosms. Water Res X 9, 100067. https://doi.org/10.1016/j.wroa.2020.100067.

Guerin, E., Shkoporov, A., Stockdale, S.R., Clooney, A.G., Ryan, F.J., Sutton, T.D.S., Draper, L.A., Gonzalez-Tortuero, E., Ross, R.P., Hill, C., 2018. Biology and taxonomy of crAss-like bacteriophages, the most abundant virus in the human gut. Cell Host Microbe 24, 653–664. https://doi.org/10.1016/j.chom.2018.10.002.

Guerin, E., Shkoporov, A.N., Stockdale, S.R., Comas, J.C., Khokhlova, E.V., Clooney, A. G., Daly, K.M., Draper, L.A., Stephens, N., Scholz, D., Ross, R.P., Hill, C., 2021. Isolation and characterisation of ΦcrAss002, a crAss-like phage from the human gut that infects *Bacteroides xylanisolvens*. Microbiome 9, 1–21. https://doi.org/10.1186/s40168-021-01036-7.

Guglielmotti, D.M., Mercanti, D.J., Reinheimer, J.A., Quiberoni, A. del L., 2012. Review: efficiency of physical and chemical treatments on the inactivation of dairy bacteriophages. Front. Microbiol. 2, 282. https://doi.org/10.3389/fmicb.2011.00282.

Hamza, I.A., Abd-Elmaksoud, S., 2023. Applicability of crAssphage as a performance indicator for viral reduction during activated sludge wastewater treatment. Environ. Sci. Pollut. Res. Int. 30, 50723–50731. https://doi.org/10.1007/s11356-023-25824-w.

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K., Korajkic, A., 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. FEMS Microbiol. Rev. 38, 1–40. https://doi.org/10.1111/1574-6976.12031.

Hryckowian, A.J., Merrill, B.D., Porter, N.T., Van Treuren, W., Nelson, E.J., Garlena, R. A., Russell, D.A., Martens, E.C., Sonnenburg, J.L., 2020. *Bacteroides thetaiotaomicron*-infecting bacteriophage isolates inform sequence-based host range predictions. Cell Host Microbe 28, 371–379. https://doi.org/10.1016/j.chom.2020.06.011.

Jagals, P., Grabow, W., 1996. An evaluation of sorbitol-fermenting bifidobacteria as specific indicators of human faecal pollution of environmental water. Water SA 22, 235–238.

Jebri, S., Muniesa, M., Jofre, J., 2017. General and host-associated bacteriophage indicators of faecal pollution. In: Global Water Pathogen Project: Part 2. Indicators and Microbial Source Tracking Markers. Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management. Michigan State University. https://doi.org/10.14321/waterpathogens.7.

Jensen, J.N., Johnson, J.D., 1990. Interferences by monochloramine and organic chloramines in free available chlorine methods. 2. N,N-Diethyl-p-phenylenediamine. Environ. Sci. Technol. 24, 985–990.

Jofre, J., Bosch, A., Lucena, F., Girones, R., Tartera, C., 1986. Evaluation of *Bacteroides fragilis* bacteriophages as indicators of the virological quality of water. Wat Sci Tech 18, 167–173.

Jofre, J., Blanch, A.R., Lucena, F., Muniesa, M., 2014. Bacteriophages infecting *Bacteroides* as a marker for microbial source tracking. Water Res. 55, 1–11. https://doi.org/10.1016/j.watres.2014.02.006.

Jofre, J., Lucena, F., Blanch, A.R., Muniesa, M., 2016. Coliphages as model organisms in the characterization and management of water resources. Water 8 (5), 199. https://doi.org/10.3390/w8050199.

Kahler, A.M., Cromeans, T.L., Roberts, J.M., Hill, V.R., 2010. Effects of source water quality on chlorine inactivation of adenovirus, coxsackievirus, echovirus, and murine norovirus. Appl. Environ. Microbiol. 76, 5159–5164. https://doi.org/10.1128/AEM.00869-10.

Kator, H., Rhodes, M., 1992. Evaluation of *Bacteroides fragilis* bacteriophage, a candidate human-specific indicator of fecal contamination for shellfish-growing waters. In: Final Report. Virginia Institute of Marine Sciences, College of Williams and Mary, Gloucester Point, VA.

Kerby, G.P., Gowdy, R.A., Dillon, E.S., Dillon, M.L., Csâky, T.Z., Sharp, D.G., Beard, J.W., 1949. Purification, pH stability and sedimentation properties of the T7 bacteriophage of *Escherichia coli*. J. Immunol. 63, 93–107. https://doi.org/10.4049/jimmunol.63.1.93.

Koonin, E.V., Yutin, N., 2020. The crAss-like phage group: how metagenomics reshaped the human virome. Trends Microbiol. 28, 349–359. https://doi.org/10.1016/j.tim.2020.01.010.

Langlet, J., Gaboriaud, F., Gantzer, C., 2007. Effects of pH on plaque forming unit counts and aggregation of MS2 bacteriophage. J. Appl. Microbiol. 103, 1632–1638. https://doi.org/10.1111/j.1365-2672.2007.03396.x.

Lee, D.U., Heinz, V., Knorr, D., 2001. Biphasic inactivation kinetics of *Escherichia coli* in liquid whole egg by high hydrostatic pressure treatments. Biotechnol. Prog. 17, 1020–1025. https://doi.org/10.1021/bp010095o.

Lund, P.A., De Biase, D., Liran, O., Scheler, O., Mira, N.P., Cetecioglu, Z., Fernández, E. N., Bover-Cid, S., Hall, R., Sauer, M., O'Byrne, C., 2020. Understanding how microorganisms respond to acid pH is central to their control and successful exploitation. Front. Microbiol. 11, 556140. https://doi.org/10.3389/fmicb.2020.556140.

Maraccini, P.A., Mattioli, M.C.M., Sassoubre, L.M., Cao, Y., Griffith, J.F., Ervin, J.S., Van De Werfhorst, L.C., Boehm, A.B., 2016. Solar inactivation of *Enterococci* and *Escherichia coli* in natural waters: effects of water absorbance and depth. Environ. Sci. Technol. 50, 5068–5076. https://doi.org/10.1021/acs.est.6b00505.

McKinney, C.W., Pruden, A., 2012. Ultraviolet disinfection of antibiotic resistant bacteria and their antibiotic resistance genes in water and wastewater. Environ. Sci. Technol. 46, 13393–13400. https://doi.org/10.1021/es303652q.

Mcminn, B.R., Korajkic, A., Ashbolt, N.J., 2014. Evaluation of *Bacteroides fragilis* GB-124 bacteriophages as novel human-associated faecal indicators in the United States. Lett. Appl. Microbiol. 59, 115–121. https://doi.org/10.1111/lam.12252.

Menon, P., Billen, G., Servais, P., 2003. Mortality rates of autochthonous and fecal bacteria in natural aquatic ecosystems. Water Res. 37, 4151–4158. https://doi.org/10.1016/S0043-1354(03)00349-X.

Mocé-Llivina, L., Muniesa, M., Pimenta-Vale, H., Lucena, F., Jofre, J., 2003. Survival of bacterial indicator species and bacteriophages after thermal treatment of sludge and sewage. Appl. Environ. Microbiol. 69, 1452–1456. https://doi.org/10.1128/AEM.69.3.1452-1456.2003.

Mocé-Llivina, L., Lucena, F., Jofre, J., 2005. Enteroviruses and bacteriophages in bathing waters. Appl. Environ. Microbiol. 71 (11), 6838–6844.

Morrison, C.M., Betancourt, W.Q., Quintanar, D.R., Lopez, G.U., Pepper, I.L., Gerba, C.P., 2020. Potential indicators of virus transport and removal during soil aquifer treatment of treated wastewater effluent. Water Res. 177, 115812. https://doi.org/10.1016/j.watres.2020.115812.

Muniesa, M., Lucena, F., Blanch, A.R., Payán, A., Jofre, J., 2012. Use of abundance ratios of somatic coliphages and bacteriophages of *Bacteroides thetaiotaomicron* GA17 for microbial source identification. Water Res. 46, 6410–6418. https://doi.org/10.1016/j.watres.2012.09.015.

Nelson, K.L., Boehm, A.B., Davies-Colley, R.J., Dodd, M.C., Kohn, T., Linden, K.G., Liu, Y., Maraccini, P.A., McNeill, K., Mitch, W.A., Nguyen, T.H., Parker, K.M., Rodriguez, R.A., Sassoubre, L.M., Silverman, A.I., Wigginton, K.R., Zepp, R.G., 2018. Sunlight-mediated inactivation of health-relevant microorganisms in water: a review of mechanisms and modeling approaches. Environ Sci Process Impacts 20, 1089–1122. https://doi.org/10.1039/c8em00047f.

Nowak, M.A., 1992. What is a quasispecies? Trends Ecol. Evol. 7, 118–121.

Nuanualsuwan, S., Cliver, D.O., 2003. Capsid functions of inactivated human picornaviruses and feline calicivirus. Appl. Environ. Microbiol. 69, 350–357. https://doi.org/10.1128/AEM.69.1.350-357.2003.

Nunez, C., Kostoulias, X., Peleg, A., Short, F., Qu, Y., 2023. A comprehensive comparison of biofilm formation and capsule production for bacterial survival on hospital surfaces. Biofilm 5, 100105. https://doi.org/10.1016/j.bioflm.2023.100105.

Palin, A.T., 1957. The determination of free and combined chlorine in water by the use of diethyl-p-phenylene diamine. J. Am. Water Works Assoc. 49, 873–880.

Papudeshi, B., Vega, A.A., Souza, C., Giles, S.K., Roach, M.J., An, M., Jacobson, N., McNair, K., Fernanda Mora, M., Pastrana, K., Boling, L., Leigh, C., Plewa, W.S., Grigson, S.R., Bouras, G., Luque, A., Droit, L., Handley, S.A., Wang, D., Dinsdale, E. A., Edwards, R.A., 2023. Host interactions of novel Crassvirales species belonging to multiple families infecting bacterial host, *Bacteroides cellulosilyticus* WH2. Microb Genom 9, 001100. https://doi.org/10.1101/2023.03.05.531146.

Park, G.W., Ng, T.F.F., Freeland, A.L., Marconi, V.C., Boom, J.A., Staat, M.A., Montmayeur, A.M., Browne, H., Narayanan, J., Payne, D.C., Cardemil, C.V., Treffiletti, A., Vinjé, J., 2020. CrAssphage as a novel tool to detect human fecal contamination on environmental surfaces and hands. Emerg. Infect. Dis. 26, 1731–1739. https://doi.org/10.3201/eid2608.200346.

Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G.T., Blanch, A. R., Lucena, F., Jofre, J., Muniesa, M., 2005. Method for isolation of *Bacteroides* bacteriophage host strains suitable for tracking sources of fecal pollution in water. Appl. Environ. Microbiol. 71, 5659–5662. https://doi.org/10.1128/AEM.71.9.5659-5662.2005.

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.

Peng, J., Tang, J., Barrett, D.M., Sablani, S.S., Anderson, N., Powers, J.R., 2017. Thermal pasteurization of ready-to-eat foods and vegetables: critical factors for process design and effects on quality. Crit. Rev. Food Sci. Nutr. 57, 2970–2995. https://doi.org/10.1080/10408398.2015.1082126.

Pérez-Illana, M., Martín-González, N., Hernando-Pérez, M., Condezo, G.N., Gallardo, J., Menéndez, M., San Martín, C., de Pablo, P.J., 2021. Acidification induces condensation of the adenovirus core. Acta Biomater. 135, 534–542. https://doi.org/10.1016/j.actbio.2021.08.019.

Pfaender, S., Brinkmann, J., Todt, D., Riebesehl, N., Steinmann, Joerg, Steinmann, Jochen, Pietschmann, T., Steinmann, E., 2015. Mechanisms of methods for hepatitis C virus inactivation. Appl. Environ. Microbiol. 81, 1616–1621. https://doi.org/10.1128/AEM.03580-14.

Qiao, Z., Ye, Y., Szczuka, A., Harrison, K.R., Dodd, M.C., Wigginton, K.R., 2022. Reactivity of viral nucleic acids with chlorine and the impact of virus encapsidation. Environ. Sci. Technol. 56 (1), 218–227. https://doi.org/10.1021/acs.est.1c04239.

Queralt, N., Jofre, J., Araujo, R., Muniesa, M., 2003. Homogeneity of the morphological groups of bacteriophages infecting *Bacteroides fragilis* strain HSP40 and strain RYC2056. Curr. Microbiol. 46, 163–168. https://doi.org/10.1007/s00284-002-3813-7.

Ramos-Barbero, M.D., Gómez-Gómez, C., Sala-Comorera, L., Rodríguez-Rubio, L., Morales-Cortes, S., Mendoza-Barberá, E., Vique, G., Toribio-Avedillo, D., Blanch, A. R., Ballesté, E., Garcia-Aljaro, C., Muniesa, M., 2023. Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution. Nat. Commun. 14, 4295. https://doi.org/10.1038/s41467-023-40098-z.

Sabar, M.A., Honda, R., Haramoto, E., 2022. CrAssphage as an indicator of human-fecal contamination in water environment and virus reduction in wastewater treatment. Water Res. 221, 118827. https://doi.org/10.1016/j.watres.2022.118827.

Sala-Comorera, L., Reynolds, L.J., Martin, N.A., O'Sullivan, J.J., Meijer, W.G., Fletcher, N.F., 2021. Decay of infectious SARS-CoV-2 and surrogates in aquatic environments. Water Res. 201, 117090. https://doi.org/10.1016/j.watres.2021.117090.

Shang, C., Cheung, L.M., Liu, W., 2007. MS2 coliphage inactivation with UV irradiation and free chlorine/monochloramine. Environ. Eng. Sci. 24, 1321–1332. https://doi.org/10.1089/ees.2006.0261.

Sharp, D.G., Hook, A.E., Taylor, A.R., Beard, D., Beard, J.W., 1946. Sedimentation characters and pH stability of the T2 bacteriophage of *Escherichia coli*. J. Biol. Chem. 165, 259–270.

Shkoporov, A.N., Khokhlova, E.V., Fitzgerald, C.B., Stockdale, S.R., Draper, L.A., Ross, R. P., Hill, C., 2018. ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides intestinalis*. Nat. Commun. 9, 4781. https://doi.org/10.1038/s41467-018-07225-7.

Sinton, L.W., Finlay, R.K., Lynch, P.A., 1999. Sunlight inactivation of fecal bacteriophages and bacteria in sewage-polluted seawater. Appl. Environ. Microbiol. 65, 3605–3613.

Stachler, E., Bibby, K., 2014. Metagenomic evaluation of the highly abundant human gut bacteriophage crAssphage for source tracking of human fecal pollution. Environ. Sci. Technol. Lett. 1, 405–409. https://doi.org/10.1021/ez500266s.

Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., Shanks, O.C., 2017. Quantitative crAssphage PCR assays for human fecal pollution measurement. Environ. Sci. Technol. 51, 9146–9154. https://doi.org/10.1021/acs.est.7b02703.

Sun, W., Jing, Z., Zhao, Z., Yin, R., Santoro, D., Mao, T., Lu, Z., 2023. Dose-response behavior of pathogens and surrogate microorganisms across the ultraviolet-C spectrum: inactivation efficiencies, action spectra, and mechanisms. Environ. Sci. Technol. 57, 10891–10900. https://doi.org/10.1021/acs.est.3c00518.

Tariq, M.A., Carding, S.R., 2021. Complete genome sequence of a *Bacteroides fragilis* bacteriophage, vB_BfrS_NCTC. Microbiol Resour Announc 10, 10. https://doi.org/10.1128/mra.00548-21.

Tartera, C., Bosch, A., Jofre, J., 1988. The inactivation of bacteriophages infecting *Bacteroides fragilis* by chlorine treatment and UV-irradiation. FEMS Microbiol. Lett. 56, 313–316. https://doi.org/10.1111/j.1574-6968.1988.tb03198.x.

Turner, D., Shkoporov, A.N., Lood, C., Millard, A.D., Dutilh, B.E., Alfenas-Zerbini, P., van Zyl, L.J., Aziz, R.K., Oksanen, H.M., Poranen, M.M., Kropinski, A.M., Barylski, J., Brister, J.R., Chanisvili, N., Edwards, R.A., Enault, F., Gillis, A., Knezevic, P., Krupovic, M., Kurtböke, I., Kushkina, A., Lavigne, R., Lehman, S., Lobocka, M., Moraru, C., Moreno Switt, A., Morozova, V., Nakavuma, J., Reyes Muñoz, A., Rümnieks, J., Sarkar, B., Sullivan, M.B., Uchiyama, J., Wittmann, J., Yigang, T., Adriaenssens, E.M., 2023. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. Arch. Virol. 168, 74. https://doi.org/10.1007/s00705-022-05694-2.

United States Environmental Protection Agency, 2012. Chapter 5. Disinfection processes in wastewater treatment. In: In Guidelines for Water Reuse.

Wigginton, K.R., Pecson, B.M., Sigstam, T., Bosshard, F., Kohn, T., 2012. Virus inactivation mechanisms: impact of disinfectants on virus function and structural integrity. Environ. Sci. Technol. 46, 12069–12078. https://doi.org/10.1021/es3029473.

World Health Organization, 2008. The global burden of disease 2004 update. https://www.who.int/publications/i/item/9789241563710.

Wu, Z., Greaves, J., Arp, L., Stone, D., Bibby, K., 2020. Comparative fate of crAssphage with culturable and molecular fecal pollution indicators during activated sludge wastewater treatment. Environ. Int. 136, 105452. https://doi.org/10.1016/j.envint.2019.105452.

Yesil, M., Kasler, D.R., Huang, E., Yousef, A.E., 2024. Thermal inactivation of *Escherichia* phage OSYSP and host strain *Escherichia coli* O157:H7 EDL933: a comparative kinetic analysis. J. Food Prot. 87, 100215. https://doi.org/10.1016/j.jfp.2023.100215.

Yutin, N., Benler, S., Shmakov, S.A., Wolf, Y.I., Tolstoy, I., Rayko, M., Antipov, D., Pevzner, P.A., Koonin, E.V., 2021. Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. Nat. Commun. 12, 1044. https://doi.org/10.1038/s41467-021-21350-w.

Zhang, Yingying, Zhuang, Y., Geng, J., Ren, H., Zhang, Yan, Ding, L., Xu, K., 2015. Inactivation of antibiotic resistance genes in municipal wastewater effluent by chlorination and sequential UV/chlorination disinfection. Sci. Total Environ. 512–513, 125–132. https://doi.org/10.1016/j.scitotenv.2015.01.028.

Zhang, Y., Qu, S., Xu, L., 2019. Progress in the study of virus detection methods: the possibility of alternative methods to validate virus inactivation. Biotechnol. Bioeng. 116, 2095–2102. https://doi.org/10.1002/bit.27003.

Zhang, Y., Wu, R., Li, W., Chen, Z., Li, K., 2021. Occurrence and distributions of human-associated markers in an impacted urban watershed. Environ. Pollut. 275, 116654. https://doi.org/10.1016/j.envpol.2021.116654.

Zhong, Q., Carratala, A., Nazarov, S., Guerrero-Ferreira, R.C., Piccinini, L., Bachmann, V., Leiman, P.G., Kohn, T., 2016. Genetic, structural, and phenotypic properties of MS2 coliphage with resistance to ClO2 disinfection. Environ. Sci. Technol. 50 (24), 13520–13528. https://doi.org/10.1021/acs.est.6b04170.

Supplementary Table 1.- Inactivation parameters of infectious crAss-like phages. GA17-A and *B. intestinalis* at different temperatures (4, 22 and 37 ºC). First order decay rates were estimated by linear regression (k_1) and by segmented linear regression (k_1 and k_2), both expressed in days. TP, turning point: the day at which k_1 changes to k_2; 95%CI: 95 % confidence intervals; s, stable for the duration of the experiment (29 days); s*, stable for the first decay rate; NA, not achieved; root mean square error, RMSE.

	4 ºC				22 ºC					37 ºC				
Microorganism	r^2	RMSE	k ($days^{-1}$) (95%CI)	TP	r^2	RMSE	k_1 ($days^{-1}$) (95%CI)	k_2 ($days^{-1}$) (95%CI)	TP	r^2	RMSE	k_1 ($days^{-1}$) (95%CI)	k_2 ($days^{-1}$) (95%CI)
ϕCrAssBcn6			s				s		1.24	0.91	0.16	0.51 (0.21 to NA)	0.03 (0.02 to 0.04)
ϕCrAssBcn10			s				s		1.24	0.95	0.11	0.34 (0.14 to NA)	0.04 (0.03 to 0.04)
ϕCrAssBcn15			s				s		1.54	0.91	0.15	0.38 (0.14 to 0.66)	0.03 (0.02 to 0.04)
ϕCrAss001			s				s		1.63	0.83	0.28	0.59 (0.18 to 1.12)	0.03 (0.02 to 0.05)
GA17-A			s				s		1.15	0.80	0.21	0.82 (0.44 to NA)	0.01 (0.001 to 0.03)
B. intestinalis	0.96	0.35	0.14 (0.13 to 0.15)	9.25	0.96	0.54	0.04 (s*) (-0.15 to 0.30)	0.36 (0.3 to 0.43)	13.03	0.99	0.13	-0.07 (-0.1 to 0.04)	0.89 (0.85 to 0.93)

Supplementary Table 2.- Inactivation parameters of infectious crAss-like phages, GA17-A and *B. intestinalis* at different pHs (3, 7 and 9). First order decay rates were estimated by linear regression for all samples, expressed in days. s, stable for the duration of the experiment (29 days); 95%CI: 95 % confidence intervals; root mean square error, RMSE.

	pH 3			pH 7			pH 9		
Microorganism	r^2	RMSE	k ($days^{-1}$) (95 % CI)	r^2	RMSE	k ($days^{-1}$) (95 % CI)	r^2	RMSE	k ($days^{-1}$) (95 % CI)
ϕCrAssBcn6	0.961	0.568	0.548 (0.507 to 0.589)			s	s		
ϕCrAssBcn10	0.908	0.867	0.565 (0.503 to 0.627)			s	s		
ϕCrAssBcn15	0.921	0.774	0.543 (0.487 to 0.598)			s	s		
ϕCrAss001	0.901	0.878	0.555 (0.492 to 0.618)			s	s		
GA17-A	0.999	0.102	6.894 (6.743 to 7.046)			s	s		
B. intestinalis	0.999	0.0443	7.18 (7.114 to 7.246)	0.955	0.360	0.141 (0.130 to 0.152)	0.982	0.399	0.918 (0.859 to 0.977)

Supplementary Table 3.- Inactivation parameters of infectious crAss-like phages, GA17-A and *B. intestinalis* after chlorination and UV irradiation. First order decay rates were estimated by linear regression and by nonlinear exponential decay, both expressed in min. P, plateau: referring to the asymptotic value that the concentration variable approaches as time tends to infinity; NA, not achieved; 95%CI: 95 % confidence intervals; root mean square error, RMSE.

	Chlorination					UV				
Microorganism	Best fitting model	r^2	RMSE	k (min^{-1}) (95%CI)	Plateau (95%CI)	Best fitting model	r^2	RMSE	k (min^{-1}) (95%CI)	Plateau (95%CI)
ϕCrAssBcn6	One phase decay	0.94	0.67	1.34 (0.72 to 4.27)	-6.23 (-4.27 to -0.72)	One phase decay	0.99	0.2	0.79 (0.68 to 0.9)	-7.76 (-8.25 to -7.55)
ϕCrAssBcn10	One phase decay	0.90	0.62	3.73 (1.38 to NA)	-4.81 (-5.32 to -4.34)	One phase decay	0.99	0.27	0.73 (0.6 to 0.88)	-7.76 (-8.28 to -7.29)
ϕCrAssBcn15	One phase decay	0.83	0.94	2.61 (0.9 to NA)	-5.15 (-6.004 to -4.34)	Linear regression	1		7.56	
ϕCrAss001	One phase decay	0.89	0.85	1.88 (0.81 to NA)	-5.52 (-6.49 to -4.57)	Linear regression	0.99	0.09	0.23 (0.19 to 0.27)	-10.45 (-11.86 to -9.5)
GA17-A	One phase decay	0.98	0.21	3.49 (1.99 to NA)	-3.56 (-3.73 to -3.39)	One phase decay	0.99	0.27	0.1 (0.08 to 0.12)	-7.42 (-7.98 to -6.92)
B. intestinalis	Linear regression	0.99	0.12	6.06 (5.78 to 6.34)		One phase decay	0.97	0.51	0.08 (0.06 to 0.11)	-8.54 (-9.99 to -7.51)

Supplementary Table 4.- Inactivation parameters of crAss-like phages DNA (evaluated by qPCR) under UV light assay and under natural inactivation experiment in summer. First order decay rates were estimated by linear regression, expressed in days and min, respectively. 95%CI: 95% confidence intervals; root mean square error, RMSE.

	UV			Summer mesocosms		
Microorganism	r^2	RMSE	k (min^{-1}) (95%CI)	r^2	RMSE	k (min^{-1}) (95%CI)
ϕCrAssBcn6	0.808	0.235	0.053 (0.04 to 0.06)	0.573	0.685	0.11 (0.09 to 0.13)
ϕCrAssBcn10	0.826	0.242	0.059 (0.05 to 0.07)	0.499	0.942	0.13 (0.09 to 0.16)
ϕCrAssBcn15	0.876	0.168	0.046 (0.04 to 0.05)	0.788	0.615	0.15 (0.13 to 0.16)
ϕCrAss001	0.809	0.154	0.035 (0.03 to 0.04)	0.610	0.886	0.13 (0.1 to 0.16)

Supplementary Table 5.- Inactivation parameters of infectious crAss-like phages, GA17-A and *B. intestinalis* in high temperature inactivation treatment (60 and 70 ºC). First order decay rates were estimated by linear regression, expressed in min. s, stable for the duration of the experiment (60 min); 95%CI: 95% confidence intervals; root mean square error, RMSE.

			60 ºC			70 ºC
Microorganism	r^2	RMSE	k (min^{-1}) (95%CI)	r^2	RMSE	k (min^{-1}) (95%CI)
ϕCrAssBcn6			s	0.93	0.31	0.05 (0.04 to 0.05)
ϕCrAssBcn10			s	0.95	0.28	0.05 (0.04 to 0.05)
ϕCrAssBcn15			s	0.94	0.33	0.05 (0.05 to 0.06)
ϕCrAss001			s	0.94	0.29	0.05 (0.04 to 0.05)
GA17-A	0.87	0.93	0.11 (0.09 to 0.12)	0.99	0.16	0.27 (0.26 to 0.27)
B. intestinalis	1		0.28	1		0.28

Supplementary Table 6.- Summary of two-way analysis of variance (ANOVA) with Tukey's multiple comparison test comparing the inactivation of each microorganism in summer and winter. The results represent the adjusted p-value outcomes. Values in bold indicate statistically significant difference ($p<0.05$).

Tukey's comparison test		ϕCrAssBcn6		ϕCrAssBcn10		ϕCrAssBcn15		ϕCrAss001		GA17-A		*B. intestinalis*	
		S	W	S	W	S	W	S	W	S	W	S	W
ϕCrAssBcn6	S	1	**<0.0001**	>0.9999	**0.0046**	>0.9999	**0.0105**	>0.9999	**0.0263**	>0.9999	**0.0052**	**<0.0001**	**<0.0001**
	W		1	0.0653	>0.9999	**0.0311**	>0.9999	**0.0125**	>0.9999	0.0584	>0.9999	**<0.0001**	**<0.0001**
ϕCrAssBcn10	S			1	**<0.0001**	>0.9999	**0.0395**	0.9995	0.0907	>0.9999	**0.0206**	**<0.0001**	**<0.0001**
	W				1	**0.0081**	>0.9999	**0.0031**	0.9995	**0.0161**	>0.9999	**<0.0001**	**<0.0001**
ϕCrAssBcn15	S					1	**<0.0001**	>0.9999	**0.0442**	>0.9999	**0.0092**	**<0.0001**	**<0.0001**
	W						1	**0.0072**	>0.9999	**0.0352**	>0.9999	**<0.0001**	**<0.0001**
ϕCrAss001	S							1	**<0.0001**	0.9998	**0.0035**	**<0.0001**	**<0.0001**
	W								1	0.0815	0.9998	**<0.0001**	**<0.0001**
GA17-A	S									1	**<0.0001**	**<0.0001**	**<0.0001**
	W										1	**<0.0001**	**<0.0001**
B. intestinalis	S											1	**<0.0001**
	W												1

S, summer; W, winter.

Supplementary Table 7.- Inactivation parameters of infectious crAss-like phages, GA17-A and *B. intestinalis* in the natural inactivation experiments (summer and winter). First order decay rates were estimated by linear regression, expressed in days. 95%CI: 95% confidence intervals; root mean square error, RMSE.

	Mesocosms summer			Mesocosms winter		
Microorganism	r^2	RMSE	k ($days^{-1}$) (95%CI)	r^2	RMSE	k ($days^{-1}$) (95%CI)
ϕCrAssBcn6	0.9	0.85	0.29 (0.26 to 0.32)	0.68	0.51	0.05 (0.04 to 0.06)
ϕCrAssBcn10	0.67	1.33	0.24 (0.20 to 0.28)	0.83	0.38	0.06 (0.05 to 0.07)
ϕCrAssBcn15	0.83	1.004	0.24 (0.21 to 0.27)	0.8	0.65	0.09 (0.08 to 0.1)
ϕCrAss001	0.86	0.91	0.27 (0.25 to 0.3)	0.67	0.63	0.07 (0.06 to 0.08)
GA17-A	0.76	1.15	0.22 (0.19 to 0.26)	0.82	0.41	0.06 (0.05 to 0.07)
B. intestinalis	0.95	0.7	1.02 (0.9 to 1.13)	0.9	0.82	0.54 (0.47 to 0.6)

5. DISCUSIÓN

El estudio de la microbiota intestinal y sus interacciones con el hospedador resulta esencial para poder comprender su impacto en la salud humana y desarrollar estrategias terapéuticas efectivas (Nicholson et al., 2012). Esta microbiota constituye un ecosistema sumamente complejo y diverso (Kamada et al., 2013; Kostic et al., 2013; Ley et al., 2006) que presenta variaciones tanto entre individuos como dentro de cada persona a lo largo de su vida (Sommer and Bäckhed, 2013; Trosvik et al., 2015), y en el que bacterias, bacteriófagos y otros microorganismos coexisten e interactúan de manera dinámica (Lozupone et al., 2012). En particular, los bacteriófagos, incluyendo aquellos que interactúan con bacterias del género *Bacteroides*, desempeñan un papel crucial en la regulación del microbioma intestinal, ya que influyen en la estabilidad, composición y funcionalidad de dicho ecosistema (Koskella and Brockhurst, 2014; Weitz and Wilhelm, 2012). La modulación de las poblaciones bacterianas a través de la depredación y la transferencia genética (Shuwen and Kefeng, 2022) no solo promueve escenarios de coexistencia beneficiosa, sino que también evidencia una interdependencia evolutiva compleja (Rodríguez-Rubio et al., 2020), que permite la propagación del fago mientras las bacterias mejoran sus capacidades competitivas y adaptativas al entorno. A nivel poblacional, los fagos ejercen una presión constante que promueve la diversidad bacteriana, generando nuevas variantes fenotípicas y genotípicas, ya sea a través de resistencias permanentes o mediante la selección fluctuante de estas (Mirzaei and Maurice, 2017; Ogilvie and Jones, 2015; Shkoporov et al., 2022). Sin embargo, las relaciones precisas entre fagos y bacterias en el intestino aún no se comprenden por completo debido a limitaciones metodológicas (Browne et al., 2016) y al desconocimiento de la diversidad de la microbiota intestinal en su totalidad. No obstante, se plantea que los fagos podrían desempeñar un papel crucial no solo en el mantenimiento de una microbiota intestinal saludable, sino también en la regulación del sistema inmune del hospedador animal (Sausset et al., 2020).

El descubrimiento de crAssphage, posiblemente los fagos más abundantes del intestino y con una amplia distribución biogeográfica, ha revolucionado el estudio de las interacciones dentro de la microbiota intestinal. Estos virus, cuyo hospedador bacteriano pertenece al género *Bacteroides* (Dutilh et al., 2014; Edwards et al., 2019), ofrecen nuevas perspectivas para comprender la dinámica y la influencia del viroma en el entorno intestinal. Además de proporcionar una comprensión más profunda de estas sinergias, se ha identificado que crAssphage podría tener aplicaciones prácticas relevantes. Entre ellas, destaca su uso como marcador en el diagnóstico de diversas enfermedades (Cervantes-Echeverría et al., 2023; Chu et al., 2023; Clooney et al., 2019; Draper et al., 2018; Imai et al., 2022; Kwan et al., 2024; Norman et al., 2015; Tomofuji et al., 2022) y como herramienta para la detección de contaminación fecal de origen humano en aguas (García-Aljaro et al., 2017; Sala-Comorera

et al., 2021; Stachler et al., 2017). Su prevalencia en heces humanas y su aparente notable persistencia a los procesos de desinfección (Ahmed et al., 2021; Ballesté et al., 2019; Crank et al., 2020; Greaves et al., 2020; Hamza and Abd-Elmaksoud, 2023; Stachler and Bibby, 2014; Wu et al., 2020; Zhang et al., 2021), han llevado a considerarlo un posible indicador universal de contaminación fecal humana.

El uso de fagos como indicadores de contaminación fecal responde a la necesidad de controlar la transmisión de patógenos virales propagados a través de la ruta fecal-oral (Ishii and Sadowsky, 2008; Jofre, 2007; Jofre et al., 2016), un aspecto clave para la salud pública y el cumplimiento de los estándares de calidad del agua (Dias et al., 2018). En este contexto, los marcadores virales han demostrado ser más eficaces que las bacterias indicadoras para representar la posible presencia de virus patógenos en ambientes acuáticos (Jung et al., 2011; Love et al., 2014; Lucena et al., 2006), dado que las bacterias presentan características que difieren significativamente de las de los virus (Byappanahalli et al., 2012; Ferguson and Signoretto, 2011; Harwood et al., 2014; Winfield and Groisman, 2003). En particular, diversos estudios han señalado que *p*-crAssphage muestra una mayor persistencia a los factores ambientales en comparación con virus humanos y que, en algunos casos, su presencia se ha correlacionado con brotes de enfermedades causadas por virus de transmisión hídrica (Crank et al., 2019; Farkas et al., 2019; Hamza and Abd-Elmaksoud, 2023; Wu et al., 2020). Esto, junto a su abundancia en muestras con contaminación fecal humana, lo ha posicionado como un indicador robusto de virus de origen fecal humano en aguas contaminadas. En este sentido, los bacteriófagos que infectan especies de *Bacteroides* ya han sido propuestos como marcadores para los estudios de MST, debido a su elevada especificidad por su hospedador bacteriano (Jofre et al., 2014; Payan et al., 2005; Tartera et al., 1989). Esto se debe a que algunas especies de *Bacteroides* residen exclusivamente en el tracto intestinal humano o en el de animales (Carson et al., 2005; Gómez-Doñate et al., 2011), y junto a ellos, los fagos que los infectan son también específicos de huésped, lo que permite discriminar entre distintas fuentes de contaminación fecal.

Los fagos tipo crAss representan una alternativa más precisa para rastrear contaminación fecal humana frente a otros fagos de *Bacteroides* ya que presentan ventajas, como su alta especificidad por hospedadores humanos, su abundancia en aguas residuales, la existencia de métodos moleculares para su detección y su mayor persistencia a la inactivación ambiental. Sin embargo, su aplicación aún enfrenta diversos desafíos que requieren soluciones. Por ejemplo, aunque son detectables mediante qPCR y metagenómica, son difíciles de aislar utilizando técnicas de cultivo convencionales. Además, uno de los principales inconvenientes del uso de qPCR es que sólo detecta un pequeño fragmento del ADN fágico, por lo que se puede producir una identificación errónea. Por otro lado, la detección de un fragmento de

genoma no se corresponde con la infectividad de las partículas virales (Aw and Rose, 2012; Greaves et al., 2020; McKee et al., 2015; Wu et al., 2020). Esto se debe a que la señal de qPCR puede mantenerse estable incluso después de aplicar métodos estándares de desinfección (Botes et al., 2013). Por tanto, la persistencia de un virus se debe realizar usando partículas infecciosas y monitorizando cómo los diferentes tratamientos afectan a dicha infectividad. Sin embargo, la dificultad de aislar los viriones de fagos tipo crAss limita significativamente la evaluación de la persistencia e inactivación de estos fagos en el ambiente, así como el estudio de la interacción fago-hospedador, aspectos clave para determinar la idoneidad de un indicador de contaminación fecal. Además, el uso exclusivo de genomas obtenidos de bases de datos conlleva el riesgo de trabajar con productos de ensamblaje erróneos, lo que compromete la precisión de los análisis y la validez de las conclusiones.

Desde el descubrimiento del *p*-crAssphage, se han aislado algunos viriones, como el ΦCrAss001, DAC15/DAC17, ΦCrAss002, el fago φ14:2 y otras 3 especies nuevas en 2023 (Guerin et al., 2021; Holmfeldt et al., 2013; Hryckowian et al., 2020; Papudeshi et al., 2023; Shkoporov et al., 2018). Dada la amplia experiencia de nuestro grupo de investigación en el estudio y uso de fagos como indicadores de contaminación fecal en aguas (Araujo et al., 2001; Durán et al., 2002; Gómez-Doñate et al., 2011; Jofre et al., 2016, 2014; Mocé-Llivina et al., 2003; Muniesa et al., 2003; Queralt et al., 2003), creímos ser capaces de aislar nuestros propios fagos tipo crAss. Ante este reto, el primer capítulo de esta tesis se centró en el aislamiento de fagos tipo crAss en la región de Cataluña. El objetivo principal tras el aislamiento fue profundizar en el estudio de su ciclo de vida y abordar algunas de las limitaciones inherentes a los estudios metagenómicos, como la alta diversidad genética de estos virus.

El asilamiento exitoso de los fagos tipo crAss requiere el cumplimiento de tres condiciones fundamentales: la selección de la muestra adecuada, la identificación de un hospedador idóneo y la implementación de ensayos sensibles para su detección. En primer lugar, la elección de la muestra de partida es un factor determinante. Estudios previos habían empleado tanto muestras fecales (Guerin et al., 2021; Shkoporov et al., 2018) como efluentes primarios de EDARs (Hryckowian et al., 2020) para el aislamiento de viriones. En este estudio, nos decantamos por utilizar aguas residuales de diferentes orígenes debido a su mayor diversidad microbiana, en comparación con las muestras de heces, al ser representativas de un mayor tamaño poblacional (Newton et al., 2015), lo que incrementa la probabilidad de detectar los fagos. El segundo aspecto clave es la selección de un hospedador bacteriano adecuado. Para ello, se emplearon diversas cepas del género *Bacteroides*, tanto de origen ambiental como clínico, procedentes de nuestro banco de cepas. Esta selección se

fundamentó en la evidencia previa que señala a *Bacteroides* como el principal hospedador de los fagos tipo crAss (Dutilh et al., 2014). Por último, la detección eficiente de estos fagos requiere la implementación de ensayos moleculares específicos. Con este objetivo, se evaluaron tres ensayos distintos de qPCR: uno diseñado específicamente para este estudio, basado en el genoma de ΦCrAss001, el primer virión aislado, y dos previamente descritos en la literatura (García-Aljaro et al., 2017; Stachler et al., 2017). Aunque podrían haberse empleado otras estrategias, como la secuenciación metagenómica de virus presentes en las muestras de aguas, estas no garantizan el aislamiento de los fagos. Por este motivo, se optó por el método de la qPCR como el enfoque más adecuado y sensible para su detección y estudio.

Inicialmente, se intentaron aislar estos fagos directamente de aguas residuales mediante la técnica de la doble capa de agar, sin una fase previa de propagación. Sin embargo, estos intentos resultaron infructuosos. Se determinó que la baja proporción de fagos tipo crAss en las muestras analizadas en comparación con otros fagos de carácter más virulento, situaba la concentración de crAssphage por debajo del límite de detección de la técnica empleada, obstaculizando su aislamiento. Además, se observó que las calvas de lisis generadas eran extremadamente tenues y difíciles de detectar. Ante estas dificultades, se optó por una estrategia alternativa basada en la realización de tres ciclos consecutivos de propagación con el objetivo de seleccionar e incrementar la proporción de fagos tipo crAss en la muestra y facilitar su aislamiento. Además, se introdujo la técnica de hibridación de calvas de lisis usando una sonda de ADN contra el genoma del ΦCrAss001 para facilitar la detección de las calvas de lisis y el aislamiento del fago tipo crAss a partir de éstas. Tras evaluar un total de 47 cepas hospedadoras, se identificó que *B. intestinalis* mostró la mejor respuesta en el proceso de propagación que se monitorizó con un ensayo de qPCR que empleaba una sonda diseñada utilizando como referencia el genoma de ΦCrAss001 (Shkoporov et al., 2018). Este ensayo se destacó por su mayor eficacia en la detección de estos fagos, por lo que esto favoreció el aislamiento de fagos relacionados con ΦCrAss001 a partir de calvas detectadas mediante la hibridación de calvas de lisis empleando la misma sonda que la del ensayo de qPCR.

Mediante la metodología empleada en este estudio, se logró aislar un total de 25 fagos tipo crAss, que fueron denominados crAssBcn. Tras su análisis genómico se determinó que todos ellos pertenecen al género *Kehishuvirus,* presentan una morfología característica de podovirus, son altamente abundantes en España y se clasificaron en seis especies distintas. Esta clasificación fue posible gracias a su similitud con el fago ΦCrAss001, dado que no se encontró coincidencia con otros fagos previamente descritos dentro del mismo género. Este hallazgo resalta la necesidad de establecer criterios más precisos para la clasificación de los fagos tipo crAss, empleando una longitud representativa del genoma o, preferiblemente,

genomas completos de fagos aislados (Yutin et al., 2021). La dificultad en el aislamiento de estos virus ha llevado a que muchas de las secuencias depositadas en bases de datos provengan de ensamblajes metagenómicos construidos a partir de múltiples individuos, lo que introduce variaciones significativas en las secuencias obtenidas (Brown et al., 2021). En otros casos, la identificación se ha basado en fragmentos genómicos de longitud reducida, lo que compromete su representatividad y dificulta una clasificación robusta. Para abordar esta problemática, en el Capítulo 3 de esta tesis se emplearon estrategias optimizadas para la detección de los fagos tipo crAss basado en el reclutamiento de genomas completos, para garantizar una mejor interpretación de su diversidad genética y su detección en diversos viromas.

A pesar de compartir un mismo hospedador, los fagos crAssBcn aislados exhiben diferencias genómicas tanto entre sí como en comparación con otros fagos tipo crAss del mismo género. Las principales variaciones se observaron en proteínas clave, como las endonucleasas, las lectinas de tipo C o las proteínas y espículas de la cola. Estas diferencias parecen estar impulsadas por una intensa presión selectiva que fomenta la adaptación de los fagos a las estrategias de resistencia adquiridas por las bacterias (Brown et al., 2021; Lee et al., 2018; McMahon et al., 2005; Siranosian et al., 2020; Wilson and Edgell, 2009). En consecuencia, estos fagos deben mantener una alta variabilidad genética para adaptarse constantemente a dichos cambios, escapando de las defensas de las bacterias y preservando su capacidad infectiva en un entorno cambiante. Estas variantes pueden originarse a partir de un proceso evolutivo divergente desde un origen común, como es el caso de las endonucleasas, o alternativamente, también podrían derivar de distintos eventos de especiación (Gabaldón and Koonin, 2013), que han dado lugar a una distribución filogenética diferenciada de ciertas proteínas, como las espículas de la cola, ubicándolas en distintas ramas del árbol evolutivo.

Un hallazgo particularmente relevante fue la identificación de diferencias en la actividad de reparación exonucleasa 3’-5’ de la ADN polimerasa I de la familia A en los fagos crAssBcn, específicamente entre las especies V y VI en comparación con el resto. Este descubrimiento revela la coexistencia de dos variantes de esta enzima dentro de un mismo grupo de fagos, lo que cuestiona la idoneidad de las polimerasas como genes marcadores para la identificación de secuencias virales en muestras metagenómicas y para la clasificación de fagos (Liang et al., 2016). La variabilidad observada pone de manifiesto la importancia de considerar esta diversidad en el diseño de nuevos ensayos de qPCR y subraya la necesidad de utilizar genes más conservados. Esto garantizaría una identificación más precisa y fiable, reduciendo la generación de datos erróneos que podrían no corresponder con crAssphage y dificultar la identificación de estos virus. Sorprendentemente, estas variaciones observadas en la ADN polimerasa I no parecen comprometer su funcionalidad, lo que sugiere que ambas

variantes coexisten y desempeñan la misma función en el ciclo de vida viral. Este hecho refuerza la hipótesis de que su origen podría estar asociado a eventos de recombinación con otros fagos filogenéticamente relacionados con crAssphage (Filée et al., 2002). Asimismo, estos resultados ponen de manifiesto la notable plasticidad genética y dinámica evolutiva de estos virus, que podrían beneficiarse de la selección de módulos funcionales intercambiables, un mecanismo evolutivo previamente propuesto en otros grupos de fagos (Botstein, 1980).

Los experimentos de propagación revelaron un comportamiento atípico en la interacción entre los fagos tipo crAss y su bacteria hospedadora. De acuerdo con estudios previos (Shkoporov et al., 2018, 2021), durante el proceso de infección, no se observó una lisis completa del cultivo bacteriano, sino que, a medida que los fagos se propagaban, la población bacteriana seguía manteniéndose estable, sin que su concentración disminuyera considerablemente. Este fenómeno sugiere la presencia de una población de células heterogénea, en la que una fracción de la población bacteriana es susceptible a la infección, mientras que otra muestra resistencia a los fagos, lo que impide la reducción excesiva del número de células del cultivo, asegurando la supervivencia global de la comunidad (Cortés-Martín et al., 2024).

Dado que los análisis genéticos no detectaron genes relacionados con la lisogenia, se concluyó que los fagos tipo crAss son estrictamente virulentos. No obstante, la coexistencia dinámica entre el fago y su hospedador sugiere la existencia de mecanismos alternativos que favorecen la persistencia de la bacteria. Diversas hipótesis previas han propuesto explicaciones para esta dinámica, tales como la infección en estado portador, la pseudolisogenia (Mäntynen et al., 2021; Shkoporov et al., 2021), así como estrategias de resistencia bacteriana (Georjon and Bernheim, 2023) tanto a nivel extracelular (Hesse et al., 2020), mediante la alteración de los receptores de superficie (Porter et al., 2020), como a nivel intracelular, interfiriendo en alguna etapa de la replicación viral (Garb et al., 2022; Garneau et al., 2010; Patel et al., 2024). La necesidad de comprender en profundidad estas interacciones llevó al planteamiento del segundo capítulo de esta tesis doctoral, donde se exploraron los factores que determinan la coexistencia entre los fagos tipo crAss y su hospedador bacteriano. El propósito fue descifrar los mecanismos que regulan esta relación y comprender sus implicaciones dentro del ecosistema microbiano.

En los experimentos de infección con fagos crAssBcn sobre cultivos bacterianos líquidos, se observó que algunas células adquirían resistencia, aunque de manera inestable, ya que revertían a su estado sensible en ausencia de fagos. Mediante cultivos consecutivos, se lograron seleccionar dos mutantes que adquirieron una resistencia estable, la cual se mantuvo durante sucesivos cultivos sin revertir incluso en ausencia de fagos. Además, se aisló otro mutante que presentó una mayor sensibilidad a la infección de los fagos crAssBcn en

comparación con la cepa salvaje, lo que permitió realizar comparaciones genéticas entre los distintos mutantes obtenidos y así tratar de identificar los mecanismos subyacentes a esta diferente dinámica fago-hospedador. El análisis de las secuencias genómicas descartó la participación de mecanismos de defensa intracelulares, como los sistemas CRISPR-Cas (Garneau et al., 2010) o BREX (Yuan et al., 2024). El estudio se centró entonces en la resistencia basada en sistemas de variación de fase o PVRs (Krinos et al., 2001), un mecanismo común en el género *Bacteroides.* Este sistema permite la reversión espontánea entre fenotipos sensibles y resistentes mediante la modificación dinámica de la superficie celular, principalmente a través de la expresión diferencial de CPSs (Fletcher et al., 2007). Este mecanismo, de carácter inespecífico, actúa sobreexpresando ciertos CPSs que enmascaran los receptores utilizados por distintos fagos, dificultando así su adhesión e infección (Carasso et al., 2024; Porter et al., 2020; Shkoporov et al., 2021). Estos CPS no solo influyen en la interacción con los fagos, sino que también tienen un impacto múltiple en la fisiología bacteriana, como la formación de biofilms y la modulación o evasión del sistema inmune del hospedador, entre otros (Béchon et al., 2020; Coyne and Comstock, 2008; Neff et al., 2016; Patrick et al., 2010; Porter et al., 2017; Round et al., 2011).

La investigación de los genomas de los distintos mutantes nos llevó a identificar una inversión diferencial en la región PVR7, responsable de la expresión del CPS7. Esta inversión variaba entre los mutantes sensibles (en posición OFF, inhibiendo el operón CPS7) y los resistentes (en posición ON promoviendo la transcripción del operón y la expresión del CPS7). El uso de técnicas de proteómica y RT-qPCR permitió detectar un leve aumento en la expresión de ciertos ORFs del operón CPS7 en las cepas resistentes. Este aumento de expresión fue algo menor en comparación con estudios previos (Cortés-Martín et al., 2024) pero quizás suficiente para inhibir la expresión de otros operones de CPS, favoreciendo la transcripción predominante de CPS7 en la mayoría de las células (Saba et al., 2024). En nuestro estudio, a diferencia de otros estudios previos (Cortés-Martín et al., 2024; Shkoporov et al., 2021) no se detectaron inversiones en otras zonas del genoma.

Se quiso comprobar si la inversión de esta región genética era un factor determinante para conferir resistencia a la infección por fagos. Para evaluar su impacto, se clonó dicha región reversible en orientación ON y se transfirió al mutante sensible. Resultados preliminares indicaron que la introducción del plásmido con el inserto en orientación ON provocó la expresión del CPS7 promovida por el *locus* PVR7. Esta expresión confirió resistencia al mutante sensible, no solo frente a fagos crAssBcn, sino también a otros fagos no relacionados con éstos, evidenciando el papel clave de esta inversión en la regulación de la resistencia bacteriana. Este hallazgo, confirma por primera vez la implicación de CPS7 en la resistencia a fagos crAssBcn y ΦCrAss001, e indica que el mecanismo de resistencia no es específico

para un solo tipo de fago, lo que implica que se trataría de un mecanismo independiente del receptor, y sugiere que la expresión capsular produce la cobertura a los distintos receptores fágicos, confiriendo la resistencia a la infección.

Tal como se mencionó en la introducción, la alta abundancia de los fagos tipo crAss en el intestino humano (Dutilh et al., 2014; Guerin et al., 2018; Koonin and Yutin, 2020; Yutin et al., 2018) y, en consecuencia, su excreción en aguas residuales, junto con su marcada ubicuidad en países occidentales (Edwards et al., 2019; Smith et al., 2023), llevó a la propuesta de *p*-crAssphage como un posible indicador universal de contaminación fecal de origen humano. La detección por qPCR de *p*-crAssphage ha sido ampliamente reportada en diversos estudios (Ahmed et al., 2018b; Ballesté et al., 2019; Chettleburgh et al., 2023; Martin et al., 2023; Meuchi et al., 2023; Park et al., 2020; Sabar et al., 2022; Sala-Comorera et al., 2021). Sin embargo, considerando los criterios esenciales que debe cumplir un buen indicador (World Health Organization, 2017), nuestro grupo de investigación ha planteado una serie de cuestiones fundamentales. En primer lugar, no se ha logrado su aislamiento en condiciones de laboratorio, y los pocos fagos aislados hasta la fecha (Guerin et al., 2021; Holmfeldt et al., 2013; Hryckowian et al., 2020; Papudeshi et al., 2023; Ramos-Barbero et al., 2023; Shkoporov et al., 2018) presentan una homología mínima con *p*-crAssphage (Guerin et al., 2021). Sin disponer de un aislamiento de este fago, resulta imposible llevar a cabo ensayos fiables sobre su persistencia e inactivación tanto en condiciones naturales como frente a distintos tratamientos de desinfección. En segundo lugar, surge la incertidumbre respecto a la verdadera especificidad de los fagos tipo crAss y en particular de *p*-crAssphage hacia hospedadores humanos y los factores que la determinan. En particular, *p*-crAssphage se ha detectado en muestras exclusivas de animales, aunque en concentraciones menores que en heces humanas (Ahmed et al., 2018; Ahmed et al., 2018; García-Aljaro et al., 2017). Esto podría deberse a una pérdida parcial de especificidad hacia el intestino humano, mutaciones que permiten adaptarse a entornos intestinales animales o, alternativamente, a que los ensayos de qPCR hayan sido dirigidos a regiones genómicas del fago que comparte con otros fagos de tipo crAss presentes en animales. Además, se han identificado algunos miembros del orden Crassvirales en ambientes naturales (Yutin et al., 2021), lo que sugiere que no todos los fagos tipo crAss tienen un origen estrictamente humano. Finalmente, se ha tratado a *p*-crAssphage como una única entidad viral, cuando en realidad se trata de un grupo abundante y extremadamente heterogéneo de fagos dentro del orden Crassvirales (Yutin et al., 2018). Dado que existen múltiples fagos similares a crAss, tanto identificados como aún por descubrir, su enorme diversidad y su distribución dependiente del área geográfica podrían representar un desafío significativo para la selección de un indicador humano de contaminación fecal fiable y universal. Por ello, resulta fundamental abordar una cuestión

clave: entre la gran diversidad de estos virus, ¿cuál es el candidato más adecuado, representativo y más ampliamente distribuido a nivel global para ser utilizado como indicador de contaminación fecal de origen humano y para el estudio de la epidemiología basada en aguas residuales? ¿Es realmente *p*-crAssphage el mejor candidato, o simplemente se usa por ser el primero que se descubrió?

De este modo, estructuramos el tercer capítulo en torno a varios criterios clave para poder seleccionar el mejor candidato entre los fagos tipo crAss. Estos criterios incluyen: alta abundancia y especificidad por viromas humanos o su asociación con contaminación fecal de origen humano, la ausencia en viromas intestinales de animales y una distribución biogeográfica global uniforme. Además, se pretendía seleccionar preferiblemente un fago que haya sido aislado.

Como se mencionó previamente, se implementó una estrategia innovadora que, por primera vez, usa para el análisis genomas completos de fagos tipo crAss y no fragmentos genómicos. Esta aproximación reduce el riesgo de clasificaciones erróneas al eliminar la posibilidad de sobreestimaciones por reactividades cruzadas y supera las limitaciones de estudios previos que utilizaban únicamente fragmentos genómicos, ya que la reconstrucción de los genomas de fagos crAss representa un desafío significativo debido a su elevada diversidad genética, dificultando su recuperación mediante ensamblaje metagenómico (Martinez-Hernandez et al., 2017).

Se seleccionaron 91 genomas completos de fagos tipo crAss, previamente reportados en bases de datos y confirmados para el estudio, y se reclutaron con viromas provenientes de diversas fuentes con contaminación fecal humana confirmada de nuestra región geográfica. El reclutamiento implica el reconocimiento e identificación (más del 90 % de similitud de secuencia) de un número significativo (> 70%) de los *reads* del genoma del fago, en los reads que componen el viroma evaluado. Además, los *reads* del fago que se detectan en el viroma deben cubrir, al menos, el 40 % de la longitud del genoma del fago, para evitar que sólo se detecte intensamente una zona pequeña del genoma del fago. En cuanto a los viromas analizados, las muestras de las que se obtuvieron incluían matrices con alta carga fecal, como heces humanas, aguas residuales y lodos, así como muestras con menor contenido fecal, como muestras ambientales y alimentos. De los 91 genomas fágicos analizados, los viromas humanos reclutaron todos los crAssBcn (representados en un único genoma para análisis posteriores) y 27 fagos tipo crAss, con una cobertura mínima del 40% del genoma en estos viromas humanos, un criterio más exigente que el 10% previamente reportado (Castro et al., 2018). Esto sugiere que estos 28 fagos estuvieron altamente representados en las muestras fecales humanas. Dado que las muestras provenían de nuestra región geográfica, cabe

remarcar que la variabilidad geográfica de crAssphage podría haber llevado a la inclusión de otros candidatos si se hubieran analizado muestras de diferentes regiones. Sin embargo, optamos por usar los metaviromas de mejor calidad y con mayores abundancias relativas en comparación con otros países, que correspondían a nuestras muestras. No obstante, enfrentamos la primera limitación en cuanto a la detección de fagos en muestras con alta diversidad (Newton et al., 2015; Rodriguez-R and Konstantinidis, 2014; Zaheer et al., 2019) o en matrices con bajo contenido fecal. La metagenómica, aunque esencial para analizar genomas completos de fagos y detectar gran cantidad de genes, tiene un límite de detección más alto que otros métodos más sensibles, como la qPCR (Crossette et al., 2021; Lindner et al., 2024), especialmente en muestras complejas o con baja carga de ADN. Estas limitaciones redujeron la probabilidad de identificar ciertos genomas de fagos crAss en muestras de aguas residuales, suelos o alimentos. Esta limitación entra en conflicto con el criterio de un buen indicador, que debería estar presente incluso en muestras con bajos niveles de contaminación fecal (Ballesté et al., 2021; Harwood et al., 2014). Por ello no descartamos los fagos con baja abundancia relativa en estas matrices, considerando los términos de las limitaciones de la técnica.

Para confirmar la preferencia por el intestino humano de los 28 fagos de tipo crAss seleccionados, evaluamos su presencia en viromas animales mediante reclutamiento. En el caso de los viromas de vacas, cerdos, conejos y primates no humanos, todos los reclutamientos fueron negativos, es decir, baja detección con una cobertura horizontal del genoma por debajo del 40%. Sin embargo, los resultados fueron distintos al analizar los dos viromas de pollo, procedentes de España y China. De los 28 fagos previamente seleccionados, cuatro estuvieron ausentes, nueve se detectaron en solo uno de los viromas y el resto aparecieron en ambos. Los fagos que aparecieron en ambos viromas de pollo fueron descartados ya que se confirmó su presencia en animales, esto incluyó el primer fago descrito, *p*-crAssphage, ya que su presencia y abundancia en los viromas de pollo no fueron despreciables. La detección de estos fagos en viromas de pollo también podría indicar una contaminación cruzada de origen humano, atribuible a diversos factores. Entre ellos, la similitud en las comunidades microbianas de humanos y animales, influenciada por aspectos dietéticos, o la posible ingestión de residuos humanos por parte de los animales (Ahmed et al., 2016; Feng and McLellan, 2019). Estos hallazgos también sugieren que el intestino de los pollos podría actuar como un reservorio de fagos tipo crAss, lo que cuestiona su exclusividad en el viroma humano y apunta a una distribución más amplia de lo que se pensaba. Esta posibilidad se ve respaldada por la notable capacidad de ciertas cepas de *Bacteroides* para adaptarse tanto a humanos como a pollos (Kollarcikova et al., 2020), lo que podría facilitar la persistencia y transferencia de estos fagos entre ambos hospedadores. Si sólo se hubiera

considerado uno de los dos viromas de pollo, algunos fagos habrían sido clasificados erróneamente como altamente específicos de humanos. Asimismo, al no poder descartar por completo la contaminación cruzada, decidimos llegar a una solución de consenso, descartamos los fagos presentes en viromas de pollo de ambos orígenes, pero clasificamos aquellos fagos detectados en un solo viroma de pollo (China o España), como específicos parciales para humanos. Por lo tanto, se seleccionaron 13 fagos específicos de humanos, los cuatro genomas completamente ausentes en viromas de pollo y los nueve con especificidad parcial por encontrarse sólo en un viroma de pollo.

El análisis comparativo a nivel de nucleótidos de los 13 fagos seleccionados como específicos de humanos, clasificados en 3 familias distintas (*Steigviridae*, *Intestiviridae* y *Suoliviridae*), reveló una homología mínima entre ellos. Esto sugiere la ausencia de una relación filogenética directa, lo que indica que estos fagos pertenecen a linajes evolutivamente independientes dentro del mismo grupo. Lamentablemente, no se identificaron regiones conservadas entre ellos, lo que habría sido fundamental para el diseño de una herramienta molecular común (qPCR por ejemplo) destinada a su detección. Esta falta de homología genética refuerza la hipótesis de que su especificidad por el hospedador humano no es el resultado de la herencia de un ancestro común, sino que probablemente haya surgido de manera independiente y simultánea en distintos eventos evolutivos. Dicho fenómeno parece estar más relacionado con la estrecha interacción de estos fagos con su bacteria hospedadora y de la bacteria con el intestino humano (Shkoporov et al., 2018), en lugar de ser una característica inherente a los propios fagos. En este sentido, la especificidad de las especies de *Bacteroides* por su hospedador humano o animal, ampliamente reportada (Gómez-Doñate et al., 2011; Payan et al., 2005; Wexler and Goodman, 2017; Xu et al., 2003) podría ser el principal factor que determina la presencia de estos fagos en sus respectivos reservorios

Finalmente, se evaluó la distribución global de esos 13 fagos seleccionados, utilizando un total de 1260 metagenomas de intestinos humanos procedentes de 14 países diferentes. Los resultados obtenidos a partir del análisis de las abundancias relativas, normalizadas según la profundidad de secuenciación de cada metagenoma, indicaron que, a pesar de la gran diversidad del grupo Crassvirales y su dispersión por el mundo, probablemente influenciada por eventos de migración humana (Edwards et al., 2019), ninguno mostró una distribución global uniforme. En nuestro estudio previo (capítulo 1), se observó que los fagos crAssBcn están altamente representados en metagenomas intestinales de varios países, incluyendo a España, incluso en mayor medida que ΦCrAss001. Sin embargo, a pesar de su abundancia en ciertos países, solo se detectaron en el 51% de los metagenomas analizados, lo que impide atribuirles una distribución global uniforme. Incluso los viriones ya aislados, como ΦcrAss002 (el más reclutado) o DAC15 y DAC17 (los más específicos para humanos, aunque menos

abundantes), no presentaron una distribución homogénea a nivel global. Estas diferencias geográficas también se podrían justificar por la capacidad de adaptación de *Bacteroides* al entorno intestinal (Xu et al., 2003), cuyas características pueden variar según la región geográfica, influenciadas principalmente por factores epigenéticos, como la dieta, que tiene un impacto significativo en la modulación del microbioma intestinal (Bubier et al., 2021; Jończyk et al., 2011; Ley et al., 2008; Mirzaei and Maurice, 2017; Shkoporov and Hill, 2019; Sonnenburg et al., 2010; Wu et al., 2015). En consecuencia, dependiendo de las cepas bacterianas presentes en una determinada zona, éstas serán más susceptibles a fagos presentes en la misma zona (Jofre et al., 2014).

Y ¿qué sucede con *p*-crAssphage, que ha sido ampliamente utilizado como marcador de MST y considerado como un marcador universal? En primer lugar, este fago fue descartado en los análisis relacionados con su presencia en viromas de pollo, lo que plantea su presencia más allá del ser humano. Otros estudios también lo han detectado en muestras animales, aunque mediante diferentes ensayos de qPCR (Sabar et al., 2022). Además, aunque su genoma completo es altamente abundante en los países en los que está presente, su detección no fue uniforme en otros países incluidos en el estudio, estando ausente en algunos de ellos, por ejemplo, en Asia. De hecho, otros fagos tipo crAss mostraron una distribución geográfica mucho más amplia que *p*-crAssphage. Estos hallazgos evidencian que *p*-crAssphage podría no ser un marcador universal y resaltan la necesidad de evaluar otros candidatos con una mayor representatividad y estabilidad en diferentes entornos, que posiblemente todavía no se han identificado debido al límite de detección de técnicas como la metagenómica o por la falta de marcadores de qPCR.

Sin embargo, hasta el momento no hemos logrado identificar un candidato universal entre los genomas analizados que supere al *p*-crAssphage, aunque hay que tener en cuenta que la aparición de nuevas secuencias podría generar resultados completamente distintos a los obtenidos. Dado que la distribución geográfica de estos fagos sugiere que su presencia está influenciada por factores locales, nuestros hallazgos respaldan la necesidad de buscar candidatos alternativos en aquellas regiones donde *p*-crAssphage esté poco representado y considerar el uso de fagos tipo crAss específicos para cada región geográfica como indicadores de contaminación fecal humana. Para ello, resulta fundamental aislar viriones autóctonos y llevar a cabo estudios detallados con su genoma completo además de los estudios sobre su persistencia, inactivación, distribución y ciclo de vida, con el objetivo de evaluar su viabilidad como marcadores.

En nuestro caso, al haber logrado aislar un mayor número de viriones de la región de Barcelona (crAssBcn), hemos sido capaces de realizar ensayos más precisos y confiables

sobre su persistencia, combinando mediciones de infectividad con análisis moleculares. Los resultados de estos estudios se presentan en el cuarto capítulo de esta tesis y proporcionan información clave para la evaluación de su potencial como marcadores de MST.

Para determinar la estabilidad de estos fagos en distintos entornos, se evaluó su persistencia bajo diferentes condiciones de temperatura, pH y su inactivación con tratamientos de desinfección, incluyendo exposición a cloro, luz ultravioleta (UV) y temperaturas extremas. Además, se llevaron a cabo ensayos en condiciones naturales, constituyendo el primer estudio sobre la inactivación natural de fagos tipo crAss aislados, un avance significativo en la comprensión de las características que los hacen buenos indicadores. Para ello, se diseñaron experimentos en mesocosmos dentro de una balsa ubicada dentro del recinto universitario, destinada a fines científicos. Con el objetivo de simular el entorno acuático, se emplearon membranas de diálisis siguiendo el mismo protocolo utilizado en estudios previos del grupo (Allué-Guardia et al., 2014, 2012; Ballesté et al., 2019; Calero-Cáceres and Muniesa, 2016), lo que permitió recrear de manera controlada las condiciones del ecosistema. Los experimentos se realizaron con tres fagos representativos de los crAssBcn, el fago ΦCrAss001 y dos controles: un fago que infecta *Bacteroides thetaiotaomicron* GA17 (fago GA17-A) y el hospedador bacteriano de los fagos crAssBcn, *B. intestinalis.*

Los resultados indicaron que los principales factores que comprometen la infectividad de los fagos tipo crAss son el pH ácido (pH 3.0), la cloración y la radiación UV. En particular, la exposición a pH bajo provoca daños en las proteínas de la cápside (Sharp et al., 1946). La cloración, ampliamente utilizada en la desinfección del agua, disminuyó significativamente la infectividad de lo fagos, aunque se confirmó que aún podían detectarse en aguas tratadas con cloro que cumplen los estándares de desinfección bacteriológica (Durán et al., 2003). En cuanto a la radiación UV se observó una inactivación más drástica en los fagos tipo crAss en comparación con los controles evaluados y otros indicadores virales, como los colifagos. Estas diferencias podrían atribuirse a variaciones en la morfología de los viriones, el tamaño del genoma o el contenido de GC (Sun et al., 2023). Es importante resaltar que, si bien la radiación UV tuvo un impacto más significativo en la cápside viral que en el ADN, como han señalado estudios previos (McKinney and Pruden, 2012; Pecson et al., 2009), otros tratamientos como la cloración, no parecieron comprometer la integridad del genoma. Esto se debe a la alta estabilidad estructural de los ácidos nucleicos frente a ciertos desinfectantes químicos (Qiao et al., 2022; Wigginton et al., 2012). Esto pone de manifiesto que el análisis de persistencia de indicadores frente a diferentes tratamientos utilizando genomas puede arrojar resultados que no se correspondan con la persistencia real del indicador, pareciendo altamente persistente ya que mantiene la integridad del ADN frente a los tratamientos, cuando en realidad el indicador se ha inactivado totalmente.

Los ensayos de estabilidad térmica mostraron que los fagos tipo crAss mantienen su infectividad a temperaturas bajas y moderadas, mientras que a 37ºC (temperatura máxima de los veranos en la región estudiada), se observó una leve disminución en su concentración. Esta tendencia concuerda con los resultados obtenidos en los ensayos de mesocosmos, donde la persistencia de los fagos fue mayor en invierno que en verano. En el caso de *B. intestinalis*, la inactivación fue aún más rápida, lo que refuerza su clasificación como indicador de contaminación fecal reciente (Ballesté and Blanch, 2010). Probablemente, el principal factor que contribuye al declive de la concentración viral en los meses cálidos es la irradiación solar (Maraccini et al., 2016; Sinton et al., 1999). Se ha demostrado que la radiación ultravioleta juega un papel clave en la desinfección natural de cuerpos de agua superficiales (Sinton et al., 1999), induciendo daños foto-bióticos (Nelson et al., 2018) tanto en las cápsides proteicas como en el genoma viral. Este efecto es consistente con los resultados de inactivación por UV en laboratorio, que mostraron que la degradación inicial de la cápside expone el ADN viral, facilitando su posterior degradación.

Los datos obtenidos indican que los fagos crAssBcn presentan una mayor persistencia frente a diversos tratamientos en comparación con otros fagos previamente descritos como indicadores de contaminación fecal humana. En particular, se observó que son más persistentes que el fago GA17-A y, en ciertos tratamientos, incluso superan la persistencia de los colifagos somáticos y F-específicos (Sinton et al., 1999). Por ejemplo, en los ensayos de cloración y temperaturas elevadas de pasteurización, los fagos crAss mostraron una notable estabilidad, en contraste con la mayor susceptibilidad observada en los colifagos (Allué-Guardia et al., 2012; Calero-Cáceres and Muniesa, 2016; Cunault et al., 2011; Mocé-Llivina et al., 2003). No obstante, en los experimentos de inactivación en condiciones naturales, la persistencia de los crAssBcn fue comparable a la de los colifagos somáticos, especialmente en temporadas cálidas (Ballesté et al., 2019). Además, también mostraron niveles de persistencia superiores a los observados para virus entéricos y a los reportados en estudios previos sobre *p*-crAssphage mediante métodos moleculares (Ahmed et al., 2021; Ballesté et al., 2019; García-Aljaro et al., 2017; Greaves et al., 2020; Stachler et al., 2017). Esto sugiere que los fagos crAssBcn podrían ser igual o más persistentes o que las diferencias en la eficiencia y sensibilidad de los métodos de detección podrían influir en la percepción de su persistencia. En conjunto, los resultados obtenidos respaldan el potencial de los fagos crAssBcn, en concreto, como indicadores MST para el monitoreo de la calidad y seguridad del agua, al menos en la región estudiada.

A lo largo de esta tesis, hemos explorado los desafíos asociados con la búsqueda de un indicador universal de contaminación fecal humana. Los intentos previos de establecer un marcador único y universal han demostrado ser infructuosos, en gran parte debido a la enorme

diversidad de las poblaciones humanas y a las variaciones geográficas y ambientales que influyen en la microbiota intestinal y sus componentes virales. Sin embargo, en este contexto, los fagos tipo crAss han surgido como los candidatos más prometedores, a pesar de las limitaciones que se han discutido a lo largo de este trabajo y que aún deben ser abordadas. El uso de los fagos tipo crAss como indicadores de contaminación fecal ofrece ventajas significativas en comparación con otros enfoques tradicionales. Su especificidad por la microbiota humana, su alta concentración en aguas residuales y su persistencia ambiental los convierte en herramientas valiosas para el monitoreo de la calidad del agua y la implementación de estrategias MST. No obstante, para consolidar su uso como biomarcadores fiables, es fundamental evaluar su desempeño en matrices ambientales más complejas que representen de manera más realista las condiciones naturales. Estos ensayos permitirán refinar su aplicación y determinar con mayor precisión su capacidad para reflejar la presencia de contaminación fecal en distintos escenarios ecológicos.

Además de su potencial como indicadores ambientales, los fagos tipo crAss han despertado un creciente interés en la comunidad científica debido a su posible correlación con ciertas enfermedades. Su abundante presencia en la microbiota intestinal humana y su relación con la estabilidad del ecosistema microbiano sugiere que podrían desempeñar un papel clave en la regulación de la salud intestinal. Investigaciones recientes han planteado la posibilidad de que estos virus no sean meros pasajeros del tracto digestivo, sino que podrían influir activamente en la composición y funcionalidad de la microbiota (Gao et al., 2021; Yang et al., 2020).

Este descubrimiento abre nuevas líneas de investigación con implicaciones tanto ecológicas como biomédicas. A medida que se profundiza en el estudio de crAssphage y su interacción con su hospedador bacteriano, se plantea la posibilidad de explorar su uso en aplicaciones terapéuticas. El uso de fagos específicos para modular la microbiota podría representar una estrategia innovadora en el tratamiento de enfermedades gastrointestinales y la restauración de comunidades microbianas alteradas, especialmente tras el uso de antibióticos (Baaziz et al., 2022). Este enfoque, conocido actualmente como “rehabilitación mediante fagos” (Baaziz et al., 2022), abre nuevas posibilidades en el manejo de desequilibrios microbianos.

En conclusión, crAssphage no solo representa un prometedor indicador de contaminación fecal, sino que también forma parte integral del viroma intestinal humano, con implicaciones que van mucho más allá del monitoreo ambiental. Su estudio se encuentra en una etapa de auge, con un creciente número de investigaciones que buscan comprender su papel ecológico, su persistencia en diferentes entornos y su potencial impacto en la salud humana. En el futuro, una comprensión más profunda de estos virus podría redefinir no solo la vigilancia

de la calidad del agua o los alimentos, sino también la manera en que abordamos el equilibrio y la funcionalidad de nuestra microbiota intestinal.

6. CONCLUSIONES

Las conclusiones extraídas de los resultados obtenidos en esta tesis doctoral son:

- El aislamiento exitoso de fagos tipo crAss requirió una estrategia metodológica que combina la selección de una muestra adecuada, un hospedador eficaz y métodos de detección sensibles, además de una propagación previa para favorecer al fago tipo crAss más virulento.
- El aislamiento de 25 nuevos fagos tipo crAss en Cataluña (denominados crAssBcn), todos del género *Kehishuvirus* y agrupados en seis especies diferentes, ha ampliado el repertorio de Crassvirales y muestra la gran variabilidad genética y funcional de este grupo de fagos, pese a compartir hospedador, lo que refleja su adaptación evolutiva.
- Se identificaron dos variantes funcionales de la ADN polimerasa I en los fagos crAssBcn, lo que evidencia su plasticidad genética y cuestiona el uso de la polimerasa como marcador genético para clasificar fagos, destacando la necesidad de usar genes más conservados.
- Los fagos crAssBcn, estrictamente virulentos, no producen la lisis completa del cultivo, lo que sugiere un modelo de coexistencia con subpoblaciones bacterianas resistentes y susceptibles, clave para la estabilidad del sistema fago-hospedador.
- Los mutantes conseguidos de *B. intestinalis,* con resistencia o sensibilidad estables, permitieron analizar la dinámica fago-hospedador, descartando defensas intracelulres (CRISPR-Cas, BREX) y resaltando el mecanismo de variación de fase como resposble de la resistencia a los fagos.
- La resistencia bacteriana a los fagos se asoció a una variación de fase estable en PVR7 que activa el operón CPS7 que, al expresarse, genera una barrera física indepenciente del receptor, que bloquea la adhesión de fagos crAssBcn y otros no relacionados.
- No hay un sólo fago tipo crAss que pueda considerarse un indicador universal de contaminación fecal humana. De 91 genomas completos analizados, sólo 13 fagos presentan especificidad humana total o parcial. Dado que presentan baja homología y carecen de regiones conservadas, no se puede diseñar un único ensayo molecular para su detección.
- Ninguno de los 13 fagos seleccionados mostró una distribución global uniforme y su variabilidad geográfica parece depender de factores ambientales, dietéticos o regionales.

- Aunque *p*-crAssphage es el candidato más usado, presenta limitaciones importantes (no se ha aislado, se detecta en pollos y no presenta una distribución mundial uniforme), por lo que convendría aislar viriones autóctonos y evaluar su persistencia para definir marcadores regionales más fiables.
- El aislamiento de múltiples fagos crAssBcn permitió evaluar su persistencia e inactivación frente a diferentes condiciones y tratamientos mediante métodos moleculares y de infectividad, aportando evidencias sólidas sobre su buena persistencia ambiental.
- Los crAssBcn mostraron mayor resistencia que otros fagos indicadores frente a cloración, altas temperaturas y factores ambientales, aunque la radiación UV resultó ser el tratamiento más eficaz de inactivación. En condiciones naturales persistieron más en invierno que en verano, lo que confirma el papel de la radiación solar como un factor determinante en la inactivación ambiental.
- Los resultados obtenidos en esta tesis respaldan el uso de fagos crAssBcn como indicadores MST en la región de estudio por su persistencia ante factores ambientales y tratamientos de inactivación.

7. BIBLIOGRAFÍA

Abedon, S.T., Murray, K.L., 2013. Archaeal viruses, not archaeal phages: An archaeological dig. Archaea 2013, 251245. https://doi.org/10.1155/2013/251245

Ackermann, H.W., 2009. Basic phage electron microscopy, in: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization and Interactions. pp. 113–126.

Adriaenssens, E.M., Cowan, D.A., 2014. Using signature genes as tools to assess environmental viral ecology and diversity. Appl Environ Microbiol 80, 4470–4480. https://doi.org/10.1128/AEM.00878-14

Ahmed, W., Bivins, A., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2022. Distribution of human fecal marker genes and their association with pathogenic viruses in untreated wastewater determined using quantitative PCR. Water Res 226, 119093. https://doi.org/10.1016/j.watres.2022.119093

Ahmed, W., Hughes, B., Harwood, V.J., 2016. Current status of marker genes of bacteroides and related taxa for identifying sewage pollution in environmental waters. Water (Switzerland) 8, 231. https://doi.org/10.3390/w8060231

Ahmed, W., Lobos, A., Senkbeil, J., Peraud, J., Gallard, J., Harwood, V.J., 2018a. Evaluation of the novel crAssphage marker for sewage pollution tracking in storm drain outfalls in Tampa, Florida. Water Res 131, 142–150. https://doi.org/10.1016/j.watres.2017.12.011

Ahmed, W., Payyappat, S., Cassidy, M., Besley, C., Power, K., 2018b. Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. Water Res 145, 769–778. https://doi.org/10.1016/j.watres.2018.08.049

Ahmed, W., Toze, S., Veal, C., Fisher, P., Zhang, Q., Zhu, Z., Staley, C., Sadowsky, M.J., 2021. Comparative decay of culturable faecal indicator bacteria, microbial source tracking marker genes, and enteric pathogens in laboratory microcosms that mimic a sub-tropical environment. Science of the Total Environment 751, 141475. https://doi.org/10.1016/j.scitotenv.2020.141475

Ahmed, W., Zhang, Q., Kozak, S., Beale, D., Gyawali, P., Sadowsky, M.J., Simpson, S., 2019. Comparative decay of sewage-associated marker genes in beach water and sediment in a subtropical region. Water Res 149, 511–521. https://doi.org/10.1016/j.watres.2018.10.088

Akhlaq, M.S., Schuchmann, H.P., Von Sonntag, C., 1990. Degradation of the polysaccharide alginic acid: a comparison of the effects of UV light and ozone. Environ Sci Technol 24, 379–383. https://doi.org/10.1021/es00073a014

Allué-Guardia, A., Jofre, J., Muniesa, M., 2012. Stability and infectivity of cytolethal distending toxin type V gene-carrying bacteriophages in a water mesocosm and under different inactivation conditions. Appl Environ Microbiol 78, 5818–5823. https://doi.org/10.1128/AEM.00997-12

Allué-Guardia, A., Martínez-Castillo, A., Muniesa, M., 2014. Persistence of infectious shiga toxin-encoding bacteriophages after disinfection treatments. Appl Environ Microbiol 80, 2142–2149. https://doi.org/10.1128/AEM.04006-13

Amarasiri, M., Kitajima, M., Nguyen, T.H., Okabe, S., Sano, D., 2017. Bacteriophage removal efficiency as a validation and operational monitoring tool for virus reduction in wastewater reclamation: Review. Water Res 121, 258–269. https://doi.org/10.1016/j.watres.2017.05.035

Anderson, K.L., Whitlock, J.E., Harwood, V.J., 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl Environ Microbiol 71, 3041–3048. https://doi.org/10.1128/AEM.71.6.3041-3048.2005

Araujo, R., Muniesa, M., Méndez, J., Puig, A., Queralt, N., Lucena, F., Jofre, J., 2001. Optimisation and standardisation of a method for detecting and enumerating bacteriophages infecting *Bacteroides fragilis*. J Virol Methods 93, 127–136. https://doi.org/10.1016/S0166-0934(01)00261-0

Armon, R., Kott, Y., 1995. Distribution comparison between coliphages and phages of anaerobic bacteria (*Bacteroides fragilis*) in water sources, and their reliability as fecal pollution indicators in drinking water. Water Science and Technology 31, 215–222. https://doi.org/10.1016/0273-1223(95)00269-S

Atamer, Z., Dietrich, J., Müller-Merbach, M., Neve, H., Heller, K.J., Hinrichs, J., 2009. Screening for and characterization of *Lactococcus lactis* bacteriophages with high thermal resistance. Int Dairy J 19, 228–235. https://doi.org/10.1016/j.idairyj.2008.10.012

Atwill, E.R., 2005. Microbial pathogens excreted by livestock and potentially transmitted to humans through water, Veterinary Medicine Teaching and Research Center, School of Veterinary Medicine, University of California, Davis, published online at http://nature. berkeley. Veterinary Medicine Teaching and Research Center, School of Veterinary Medicine.

Aw, T.G., Rose, J.B., 2012. Detection of pathogens in water: From phylochips to qPCR to pyrosequencing. Curr Opin Biotechnol 23, 422–430. https://doi.org/10.1016/j.copbio.2011.11.016

Azzazy, H.M., Highsmith Jr, W.E., 2002. Phage display technology: clinical applications and recent innovations. Clin Biochem 35, 425–445. https://doi.org/10.1016/S0009-9120(02)00343-0

Baaziz, H., Baker, Z.R., Franklin, H.C., Hsu, B.B., 2022. Rehabilitation of a misbehaving microbiome: phages for the remodeling of bacterial composition and function. iScience 25, 104146. https://doi.org/10.1016/j.isci

Ballesté, E., Blanch, A.R., 2010. Persistence of *Bacteroides* species populations in a river as measured by molecular and culture techniques. Appl Environ Microbiol 76, 7608–7616. https://doi.org/10.1128/AEM.00883-10

Ballesté, E., Blanch, A.R., Mendez, J., Sala-Comorera, L., Maunula, L., Monteiro, S., Farnleitner, A.H., Tiehm, A., Jofre, J., García-Aljaro, C., 2021. Bacteriophages are good estimators of human viruses present in water. Front Microbiol 12, 619495. https://doi.org/10.3389/fmicb.2021.619495

Ballesté, E., Pascual-Benito, M., Martín-Díaz, J., Blanch, A.R., Lucena, F., Muniesa, M., Jofre, J., García-Aljaro, C., 2019. Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. Water Res 155, 233–244. https://doi.org/10.1016/j.watres.2019.02.042

Barr, J.J., Auro, R., Furlan, M., Whiteson, K.L., Erb, M.L., Pogliano, J., Stotland, A., Wolkowicz, R., Cutting, A.S., Doran, K.S., Salamon, P., Youle, M., Rohwer, F., 2013. Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc Natl Acad Sci U S A 110, 10771–10776. https://doi.org/10.1073/pnas.1305923110

Barrow, P.A., 2001. The use of bacteriophages for treatment and prevention of bacterial disease in animals and animal models of human infection. Journal of Chemical Technology and Biotechnology 76, 677–682. https://doi.org/10.1002/jctb.436

Baughn A.D., Malamy M.H., 2004. The strict anaerobe *Bacteroides fragilis* grows in and benefits from nanomolar concentrations of oxygen. Nature 427, 441–444. https://doi.org/10.1038/nature02245

Bauza, V., Madadi, V., Ocharo, R.M., Nguyen, T.H., Guest, J.S., 2019. Microbial source tracking using 16S rRNA amplicon sequencing identifies evidence of widespread contamination from young children's feces in an Urban Slum of Nairobi, Kenya. Environ Sci Technol 53, 8271–8281. https://doi.org/10.1021/acs.est.8b06583

Bayfield, O.W., Shkoporov, A.N., Yutin, N., Khokhlova, E. V., Smith, J.L.R., Hawkins, D.E.D.P., Koonin, E. V., Hill, C., Antson, A.A., 2023. Structural atlas of a human gut crassvirus. Nature 617, 409–416. https://doi.org/10.1038/s41586-023-06019-2

Bazan, J., Całkosiński, I., Gamian, A., 2012. Phage display - A powerful technique for immunotherapy. 1. Introduction and potential of therapeutic applications. Hum Vaccin Immunother 8, 1817–1828. https://doi.org/10.4161/hv.21703

Béchon, N., Mihajlovic, J., Vendrell-Fernández, S., Chain, F., Langella, P., Beloin, C., Ghigo, J.-M., 2020. Capsular polysaccharide cross-regulation modulates *Bacteroides thetaiotaomicron* biofilm formation. mBio 11, 10–1128. https://doi.org/10.1128/mBio

Ben-Assa, N., Coyne, M.J., Fomenkov, A., Livny, J., Robins, W.P., Muniesa, M., Carey, V., Carasso, S., Gefen, T., Jofre, J., Roberts, R.J., Comstock, L.E., Geva-Zatorsky, N., 2020. Analysis of a phase-variable restriction modification system of the human gut symbiont *Bacteroides fragilis*. Nucleic Acids Res 48, 11040–11053. https://doi.org/10.1093/nar/gkaa824

Benler, S., Cobián-Güemes, A.G., McNair, K., Hung, S.H., Levi, K., Edwards, R., Rohwer, F., 2018. A diversity-generating retroelement encoded by a globally ubiquitous *Bacteroides* phage. Microbiome 6, 1–10. https://doi.org/10.1186/s40168-018-0573-6

Benler, S., Yutin, N., Antipov, D., Rayko, M., Shmakov, S., Gussow, A.B., Pevzner, P., Koonin, E. V., 2021. Thousands of previously unknown phages discovered in whole-community human gut metagenomes. Microbiome 9, 1–17. https://doi.org/10.1186/s40168-021-01017-w

Bjursell, M.K., Martens, E.C., Gordon, J.I., 2006. Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. Journal of Biological Chemistry 281, 36269–36279. https://doi.org/10.1074/jbc.M606509200

Blanch, A.R., Belanche-Muñ Oz, L., Bonjoch, X., Ebdon, J., Gantzer, C., Lucena, F., Ottoson, J., Kourtis, C., Iversen, A., Kü, I., Moce, L., Muniesa, M., Schwartzbrod, J., Skraber, S., Papageorgiou, G., Taylor, H.D., Wallis, J., Jofre, J., 2004. Tracking the origin of faecal pollution in surface water: an ongoing project within the European Union research programme. J Water Health 2, 249–260. https://doi.org/10.2166/wh.2004.0022

Blanch, A.R., Belanche-Muñoz, L., Bonjoch, X., Ebdon, J., Gantzer, C., Lucena, F., Ottoson, J., Kourtis, C., Iversen, A., Kühn, I., Mocé, L., Muniesa, M., Schwartzbrod, J., Skraber, S., Papageorgiou, G.T., Taylor, H., Wallis, J., Jofre, J., 2006. Integrated analysis of established and novel microbial and chemical methods for microbial source tracking. Appl Environ Microbiol 72, 5915–5926. https://doi.org/10.1128/AEM.02453-05

Bondy-Denomy, J., Qian, J., Westra, E.R., Buckling, A., Guttman, D.S., Davidson, A.R., Maxwell, K.L., 2016. Prophages mediate defense against phage infection through diverse mechanisms. ISME Journal 10, 2854–2866. https://doi.org/10.1038/ismej.2016.79

Borodovich, T., Shkoporov, A.N., Ross, R.P., Hill, C., 2022. Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. Gastroenterol Rep (Oxf) 10, 1–12. https://doi.org/10.1093/gastro/goac012

Botes, M., De Kwaadsteniet, M., Cloete, T.E., 2013. Application of quantitative PCR for the detection of microorganisms in water. Anal Bioanal Chem 405, 91–108. https://doi.org/10.1007/s00216-012-6399-3

Botstein, D., 1980. A theory of modular evolution for bacteriophages. Annals of the New York Academy of Sciences 354, 484–490. https://doi.org/10.1111/j.1749-6632.1980.tb27987.x

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J., Salamon, P., Rohwer, F., 2003. Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol 185, 6220–6223. https://doi.org/10.1128/JB.185.20.6220-6223.2003

Broecker, F., Russo, G., Klumpp, J., Moelling, K., 2017. Stable core virome despite variable microbiome after fecal transfer. Gut Microbes 8, 214–220. https://doi.org/10.1080/19490976.2016.1265196

Brown, B.P., Chopera, D., Havyarimana, E., Wendoh, J., Jaumdally, S., Nyangahu, D.D., Gray, C.M., Martin, D.P., Varsani, A., Jaspan, H.B., 2021. crAssphage genomes identified in fecal samples of

an adult and infants with evidence of positive genomic selective pressure within tail protein genes. Virus Res 292, 198219. https://doi.org/10.1016/j.virusres.2020.198219

Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., Goulding, D., Lawley, T.D., 2016. Culturing of "unculturable" human microbiota reveals novel taxa and extensive sporulation. Nature 533, 543–546. https://doi.org/10.1038/nature17645

Brussaard, C.P., 2009. Enumeration of bacteriophages using flow cytometry, in: Bacteriophages: Methods and Protocols, Volume 1: Isolations, Characterization, and Interactions. pp. 97–111. https://doi.org/10.1007/978-1-60327-164-6_11

Bryan, W.F., 1977. Diseases transmitted by foods contaminated. J Food Prot 40, 45–56. https://doi.org/10.4315/0362-028X-40.1.45

Bubier, J.A., Chesler, E.J., Weinstock, G.M., 2021. Host genetic control of gut microbiome composition. Mammalian Genome 32, 263–281. https://doi.org/10.1007/s00335-021-09884-2

Byappanahalli, M.N., Nevers, M.B., Korajkic, A., Staley, Z.R., Harwood, V.J., 2012. *Enterococci* in the Environment. Microbiology and Molecular Biology Reviews 76, 685–706. https://doi.org/10.1128/mmbr.00023-12

Calero-Cáceres, W., Muniesa, M., 2016. Persistence of naturally occurring antibiotic resistance genes in the bacteria and bacteriophage fractions of wastewater. Water Res 95, 11–18. https://doi.org/10.1016/j.watres.2016.03.006

Camarillo-Guerrero, L.F., Almeida, A., Rangel-Pineros, G., Finn, R.D., Lawley, T.D., 2021. Massive expansion of human gut bacteriophage diversity. Cell 184, 1098-1109.e9. https://doi.org/10.1016/j.cell.2021.01.029

Campos, C., 2013. Personnel communication. Claudia Campos (Universidad Javeriana).

Carasso, S., Zaatry, R., Hajjo, H., Kadosh-Kariti, D., Ben-Assa, N., Naddaf, R., Mandelbaum, N., Pressman, S., Chowers, Y., Gefen, T., Jeffrey, K.L., Jofre, J., Coyne, M.J., Comstock, L.E., Sharon, I., Geva-Zatorsky, N., 2024. Inflammation and bacteriophages affect DNA inversion states and functionality of the gut microbiota. Cell Host Microbe 32, 322-334.e9. https://doi.org/10.1016/J.CHOM.2024.02.003

Carding, S.R., Davis, N., Hoyles, L., 2017. Review article: the human intestinal virome in health and disease. Aliment Pharmacol Ther 46, 800–815. https://doi.org/10.1111/apt.14280

Carson, C.A., Christiansen, J.M., Yampara-Iquise, H., Benson, V.W., Baffaut, C., Davis, J. V., Broz, R.R., Kurtz, W.B., Rogers, W.M., Fales, W.H., 2005. Specificity of a *Bacteroides thetaiotaomicron* marker for human feces. Appl Environ Microbiol 71, 4945–4949. https://doi.org/10.1128/AEM.71.8.4945-4949.2005

Caruso, G., 2013. Microbes and their use as indicators of pollution. Journal of Pollution Effects and Control 1, e102. https://doi.org/10.4172/2375-4397.1000e102

Casjens, S.R., Hendrix, R.W., 2015. Bacteriophage lambda: early pioneer and still relevant. Virology 479, 310–330. https://doi.org/10.1016/j.virol.2015.02.010

Castro, J.C., Rodriguez, L.M., Harvey, W.T., Weigand, M.R., Hatt, J.K., Carter, M.Q., Konstantinidis, K.T., 2018. ImGLAD: Accurate detection and quantification of target organisms in metagenomes. PeerJ 6, e5882. https://doi.org/10.7717/peerj.5882

Cervantes-Echeverría, M., Gallardo-Becerra, L., Cornejo-Granados, F., Ochoa-Leyva, A., 2023. The two-faced role of crAssphage subfamilies in obesity and metabolic syndrome: between good and evil. Genes (Basel) 14, 139. https://doi.org/10.3390/genes14010139

Chatzidaki-Livanis, M., Coyne, M.J., Roche-Hakansson, H., Comstock, L.E., 2008. Expression of a uniquely regulated extracellular polysaccharide confers a large-capsule phenotype to *Bacteroides fragilis*. J Bacteriol 190, 1020–1026. https://doi.org/10.1128/JB.01519-07

Chatzidaki-Livanis, M., Weinacht, K.G., Comstock, L.E., 2010. *Trans* locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont *Bacteroides fragilis*. Proc Natl Acad Sci U S A 107, 11976–11980. https://doi.org/10.1073/pnas.1005039107

Chettleburgh, C., Ma, S.X., Swinwood-Sky, M., McDougall, H., Kireina, D., Taggar, G., McBean, E., Parreira, V., Goodridge, L., Habash, M., 2023. Evaluation of four human-associated fecal biomarkers in wastewater in Southern Ontario. Science of the Total Environment 904, 166542. https://doi.org/10.1016/j.scitotenv.2023.166542

Chow, C.E.T., Suttle, C.A., 2015. Biogeography of viruses in the sea. Annu Rev Virol 2, 41–66. https://doi.org/10.1146/annurev-virology-031413-085540

Chu, Y., Meng, Q., Yu, J., Zhang, J., Chen, J., Kang, Y., 2023. Strain-level dynamics reveal regulatory roles in atopic eczema by gut bacterial phages. Microbiol Spectr 11, e04551-22. https://doi.org/10.1128/spectrum.04551-22

Chung, H., Sobsey, M.D., 1993. Comparative survival of indicator viruses and enteric viruses in seawater and sediment. Water Science and Technology 27, 425–428. https://doi.org/10.2166/wst.1993.0385

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

Clooney, A.G., Sutton, T.D.S., Shkoporov, A.N., Holohan, R.K., Daly, K.M., O'Regan, O., Ryan, F.J., Draper, L.A., Plevy, S.E., Ross, R.P., Hill, C., 2019. Whole-virome analysis sheds light on viral dark matter in inflammatory bowel disease. Cell Host Microbe 26, 764–778. https://doi.org/10.1016/j.chom.2019.10.009

Colomer-Lluch, M., Jofre, J., Muniesa, M., 2011. Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples. PLoS One 6, e17549. https://doi.org/10.1371/journal.pone.0017549

Cortés-Martín, A., Buttimer, C., Maier, J.L., Tobin, C.A., Draper, L.A., Ross, R.P., Kleiner, M., Hill, C., Shkoporov, A.N., 2024. Adaptations in gut *Bacteroidales* facilitate stable co-existence with their lytic bacteriophages. Gut Microbes 17, 2507775. https://doi.org/10.1101/2024.11.17.624012

Costán-Longares, A., Montemayor, M., Payán, A., Méndez, J., Jofre, J., Mujeriego, R., Lucena, F., 2008. Microbial indicators and pathogens: Removal, relationships and predictive capabilities in water reclamation facilities. Water Res 42, 4439–4448. https://doi.org/10.1016/j.watres.2008.07.037

Coyne, M.J., Comstock, L.E., 2016. A new pillar in pilus assembly. Cell 165, 520–521. https://doi.org/10.1016/j.cell.2016.04.024

Coyne, M.J., Comstock, L.E., 2008. Niche-specific features of the intestinal *Bacteroidales*. J Bacteriol 190, 736–742. https://doi.org/10.1128/JB.01559-07

Coyne, M.J., Weinacht, K.G., Krinos, C.M., Comstock, L.E., 2003. Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. Proceedings of the National Academy of Sciences 100, 10446–10451. https://doi.org/doi.org/10.1073/pnas.1832655100

Crank, K., Li, X., North, D., Ferraro, G.B., Iaconelli, M., Mancini, P., La Rosa, G., Bibby, K., 2020. CrAssphage abundance and correlation with molecular viral markers in Italian wastewater. Water Res 184, 116161. https://doi.org/10.1016/j.watres.2020.116161

Crank, K., Petersen, S., Bibby, K., 2019. Quantitative microbial risk assessment of swimming in sewage impacted waters using crAssphage and pepper mild mottle virus in a customizable model. Environ Sci Technol Lett 6, 571–577. https://doi.org/10.1021/acs.estlett.9b00468

Crossette, E., Gumm, J., Langenfeld, K., Raskin, L., Duhaime, M., Wigginton, K., 2021. Metagenomic quantification of genes with internal standards. mBio 12, 1–12. https://doi.org/10.1128/mBio.03173-20

Cuevas-Ferrando, E., Pérez-Cataluña, A., Falcó, I., Randazzo, W., Sánchez, G., 2022. Monitoring human viral pathogens reveals potential hazard for treated wastewater discharge or reuse. Front Microbiol 13, 836193. https://doi.org/10.3389/fmicb.2022.836193

Cunault, C., Pourcher, A.M., Burton, C.H., 2011. Using temperature and time criteria to control the effectiveness of continuous thermal sanitation of piggery effluent in terms of set microbial indicators. J Appl Microbiol 111, 1492–1504. https://doi.org/10.1111/j.1365-2672.2011.05144.x

Current ICTV Taxonomy Release | ICTV, 2023. Current ICTV Taxonomy Release | ICTV [WWW Document]. https://ictv.global/taxonomy. URL https://ictv.global/taxonomy (accessed 7.2.24).

David, L.A., Materna, A.C., Friedman, J., Campos-Baptista, M.I., Blackburn, M.C., Perrotta, A., Erdman, S.E., Alm, E.J., 2014. Host lifestyle affects human microbiota on daily timescales. Genome Biol 15, 1–15. https://doi.org/10.1186/gb-2014-15-7-r89

De Sordi, L., Lourenço, M., Debarbieux, L., 2019. “I will survive”: A tale of bacteriophage-bacteria coevolution in the gut. Gut Microbes 10, 92–99. https://doi.org/10.1080/19490976.2018.1474322

Devoto, A.E., Santini, J.M., Olm, M.R., Anantharaman, K., Munk, P., Tung, J., Archie, E.A., Turnbaugh, P.J., Seed, K.D., Blekhman, R., Aarestrup, F.M., Thomas, B.C., Banfield, J.F., 2019. Megaphages infect *Prevotella* and variants are widespread in gut microbiomes. Nat Microbiol 4, 693–700. https://doi.org/10.1038/s41564-018-0338-9

Dias, E., Ebdon, J., Taylor, H., 2018. The application of bacteriophages as novel indicators of viral pathogens in wastewater treatment systems. Water Res 129, 172–179. https://doi.org/10.1016/j.watres.2017.11.022

Dion, M.B., Oechslin, F., Moineau, S., 2020. Phage diversity, genomics and phylogeny. Nat Rev Microbiol. https://doi.org/10.1038/s41579-019-0311-5

Donaldson, G.P., Lee, S.M., Mazmanian, S.K., 2015. Gut biogeography of the bacterial microbiota. Nat Rev Microbiol 14, 20–32. https://doi.org/10.1038/nrmicro3552

Draper, L.A., Ryan, F.J., Smith, M.K., Jalanka, J., Mattila, E., Arkkila, P.A., Ross, R.P., Satokari, R., Hill, C., 2018. Long-term colonisation with donor bacteriophages following successful faecal microbial transplantation. Microbiome 6, 1–9. https://doi.org/10.1186/s40168-018-0598-x

Duerkop, B.A., Clements, C. V., Rollins, D., Rodrigues, J.L.M., Hooper, L. V., 2012. A composite bacteriophage alters colonization by an intestinal commensal bacterium. Proc Natl Acad Sci U S A 109, 17621–17626. https://doi.org/10.1073/pnas.1206136109

Durán, A.E., Muniesa, M., Méndez, X., Valero, F., Lucena, F., Jofre, J., 2002. Removal and inactivation of indicator bacteriophages in fresh waters. J Appl Microbiol 92, 338–347. https://doi.org/10.1046/j.1365-2672.2002.01536.x

Durán, A.E., Muniesa, M., Mocé -Llivina, L., Campos, C., Jofre, J., Lucena, F., 2003. Usefulness of different groups of bacteriophages as model micro-organisms for evaluating chlorination. J Appl Microbiol 95, 29–37. https://doi.org/10.1046/j.1365-2672.2003.t01-1-01948.x

Dutilh, B.E., Cassman, N., McNair, K., Sanchez, S.E., Silva, G.G.Z., Boling, L., Barr, J.J., Speth, D.R., Seguritan, V., Aziz, R.K., Felts, B., Dinsdale, E.A., Mokili, J.L., Edwards, R.A., 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat Commun 5, 4498. https://doi.org/10.1038/ncomms5498

Dy, R.L., Rigano, L.A., Fineran, P.C., 2018. Phage-based biocontrol strategies and their application in agriculture and aquaculture. Biochem Soc Trans 46, 1605–1613. https://doi.org/10.1042/BST20180178

Ebdon, J., Muniesa, M., Taylor, H., 2007. The application of a recently isolated strain of *Bacteroides* (GB-124) to identify human sources of faecal pollution in a temperate river catchment. Water Res 41, 3683–3690. https://doi.org/10.1016/j.watres.2006.12.020

Ebdon, J., Sellwood, J., Shore, J., Taylor, H., 2012. Phages of *Bacteroides* (GB-124): A novel tool for viral waterborne disease control? Environ Sci Technol 46, 1163–1169. https://doi.org/10.1021/es202874p

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., Relman, D.A., 2005. Diversity of the human intestinal microbial flora. Science (1979) 308, 1635–1638. https://doi.org/10.1126/science.1110591

Edwards, R.A., Vega, A.A., Norman, H.M., Ohaeri, M., Levi, K., Dinsdale, E.A., Cinek, O., Aziz, R.K., McNair, K., Barr, J.J., Bibby, K., Brouns, S.J.J., Cazares, A., de Jonge, P.A., Desnues, C., Díaz Muñoz, S.L., Fineran, P.C., Kurilshikov, A., Lavigne, R., Mazankova, K., McCarthy, D.T., Nobrega, F.L., Reyes Muñoz, A., Tapia, G., Trefault, N., Tyakht, A. V., Vinuesa, P., Wagemans, J., Zhernakova, A., Aarestrup, F.M., Ahmadov, G., Alassaf, A., Anton, J., Asangba, A., Billings, E.K., Cantu, V.A., Carlton, J.M., Cazares, D., Cho, G.S., Condeff, T., Cortés, P., Cranfield, M., Cuevas, D.A., De la Iglesia, R., Decewicz, P., Doane, M.P., Dominy, N.J., Dziewit, L., Elwasila, B.M., Eren, A.M., Franz, C., Fu, J., Garcia-Aljaro, C., Ghedin, E., Gulino, K.M., Haggerty, J.M., Head, S.R., Hendriksen, R.S., Hill, C., Hyöty, H., Ilina, E.N., Irwin, M.T., Jeffries, T.C., Jofre, J., Junge, R.E., Kelley, S.T., Khan Mirzaei, M., Kowalewski, M., Kumaresan, D., Leigh, S.R., Lipson, D., Lisitsyna, E.S., Llagostera, M., Maritz, J.M., Marr, L.C., McCann, A., Molshanski-Mor, S., Monteiro, S., Moreira-Grez, B., Morris, M., Mugisha, L., Muniesa, M., Neve, H., Nguyen, N. phuong, Nigro, O.D., Nilsson, A.S., O'Connell, T., Odeh, R., Oliver, A., Piuri, M., Prussin, A.J., Qimron, U., Quan, Z.X., Rainetova, P., Ramírez-Rojas, A., Raya, R., Reasor, K., Rice, G.A.O., Rossi, A., Santos, R., Shimashita, J., Stachler, E.N., Stene, L.C., Strain, R., Stumpf, R., Torres, P.J., Twaddle, A., Ugochi Ibekwe, M.A., Villagra, N., Wandro, S., White, B., Whiteley, A., Whiteson, K.L., Wijmenga, C., Zambrano, M.M., Zschach, H., Dutilh, B.E., 2019. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. Nat Microbiol 4, 1727–1736. https://doi.org/10.1038/s41564-019-0494-6

Farkas, K., Adriaenssens, E.M., Walker, D.I., McDonald, J.E., Malham, S.K., Jones, D.L., 2019. Critical evaluation of crAssphage as a molecular marker for human-derived wastewater contamination in the aquatic environment. Food Environ Virol 11, 113–119. https://doi.org/10.1007/s12560-019-09369-1

Farnleitner, A., Blanch, A., 2017. Microbial indicators - "Workhorses" in the field of health-related water quality testing, in: Global Water Pathogen Project. (Part 2: Indicators and Microbial Source Tracking Markers). Michigan State University, E. Lansing, MI, UNESCO, https://www.waterpathogens.org/book/editorial.

Feng, S., McLellan, S.L., 2019. Highly specific sewage-derived Bacteroides quantitative PCR assays target sewage-polluted waters. Appl Environ Microbiol 85, e02696-18. https://doi.org/10.1128/AEM.02696-18

Feng, Y.Y., Ong, S.L., Hu, J.Y., Tan, X.L., Ng, W.J., 2003. Effects of pH and temperature on the survival of coliphages MS2 and Qβ. J Ind Microbiol Biotechnol 30, 549–552. https://doi.org/10.1007/s10295-003-0080-y

Ferguson, D., Signoretto, C., 2011. Environmental persistence and naturalization of fecal indicator organisms. Microbial Source Tracking: Methods, Applications, and Case Studies 379–397. https://doi.org/10.1007/978-1-4419-9386-1_17

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res 41, 3517–3538. https://doi.org/10.1016/j.watres.2007.06.056

Filée, J., Forterre, P., Sen-Lin, T., Laurent, J., 2002. Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. J Mol Evol 54, 763–773. https://doi.org/10.1007/s00239-001-0078-x

Fitzgerald, C.B., Shkoporov, A.N., Upadrasta, A., Khokhlova, E. V., Ross, R.P., Hill, C., 2021. Probing the “dark matter” of the human gut phageome: culture assisted metagenomics enables rapid discovery and host-linking for novel bacteriophages. Front Cell Infect Microbiol 11, 616918. https://doi.org/10.3389/fcimb.2021.616918

Fletcher, C.M., Coyne, M.J., Bentley, D.L., Villa, O.F., Comstock, L.E., 2007. Phase-variable expression of a family of glycoproteins imparts a dynamic surface to a symbiont in its human intestinal ecosystem. PNAS 104, 2413–2418. https://doi.org/10.1073/PNAS.0608797104

Fortier, L.C., Sekulovic, O., 2013. Importance of prophages to evolution and virulence of bacterial pathogens. Virulence 4, 354–365. https://doi.org/10.4161/viru.24498

Gabaldón, T., Koonin, E. V., 2013. Functional and evolutionary implications of gene orthology. Nat Rev Genet 14, 360–366. https://doi.org/10.1038/nrg3456

Gantzer, C., Maul, A., Audic, J.M., Schwartzbrod, A.L., 1998. Detection of infectious enteroviruses, enterovirus genomes, somatic coliphages, and *Bacteroides fragilis* phages in treated wastewater. Appl Environ Microbiol 64, 4307–4312. https://doi.org/10.1128/AEM.64.11.4307-4312.1998

Gao, R., Zhu, Yefei, Kong, C., Xia, K., Li, H., Zhu, Yin, Zhang, X., Liu, Y., Zhong, H., Yang, R., Chen, C., Qin, N., Qin, H., 2021. Alterations, interactions, and diagnostic potential of gut bacteria and viruses in colorectal cancer. Front Cell Infect Microbiol 11, 657867. https://doi.org/10.3389/fcimb.2021.657867

Garb, J., Lopatina, A., Bernheim, A., Zaremba, M., Siksnys, V., Melamed, S., Leavitt, A., Millman, A., Amitai, G., Sorek, R., 2022. Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. Nat Microbiol 7, 1849–1856. https://doi.org/10.1038/S41564-022-01207-8

García-Aljaro, C., Ballesté, E., Muniesa, M., Jofre, J., 2017. Determination of crAssphage in water samples and applicability for tracking human faecal pollution. Microb Biotechnol 10, 1775–1780. https://doi.org/10.1111/1751-7915.12841

Garneau, J.E., Dupuis, M.È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67–71. https://doi.org/10.1038/nature09523

Georjon, H., Bernheim, A., 2023. The highly diverse antiphage defence systems of bacteria. Nature Reviews Microbiology 21, 686–700. https://doi.org/10.1038/s41579-023-00934-x

Girones, R., Ferrús, M.A., Alonso, J.L., Rodriguez-Manzano, J., Calgua, B., de Abreu Corrêa, A., Hundesa, A., Carratala, A., Bofill-Mas, S., 2010. Molecular detection of pathogens in water - The pros and cons of molecular techniques. Water Res 44, 4325–4339. https://doi.org/10.1016/j.watres.2010.06.030

Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., Afik, S., Ofir, G., Sorek, R., 2015. BREX is a novel phage resistance system widespread in microbial genomes. EMBO J 34, 169–183. https://doi.org/10.15252/EMBJ.201489455

Gómez-Doñate, M., Payán, A., Cortés, I., Blanch, A.R., Lucena, F., Jofre, J., Muniesa, M., 2011. Isolation of bacteriophage host strains of *Bacteroides* species suitable for tracking sources of animal faecal pollution in water. Environ Microbiol 13, 1622–1631. https://doi.org/10.1111/j.1462-2920.2011.02474.x

Grabow, W., 2001. Bacteriophages: Update on application as models for viruses in water. Water SA 27, 251–68. https://doi.org/10.4314/wsa.v27i2.4999

Grabow, W., Holtzhausen, C., De Villiers, J.C., 1993. Research on bacteriophages as indicators of water quality. Pretoria, South Africa. Project Report, 321. Water Research Commission.

Grabow, W.O.K., Neubrech, T.E., Holtzhausen, C.S., Jofre, J., 1995. Bacteroides fragilis and Escherichia coli bacteriophages: excretion by humans and animals. Water Science and Technology 31, 223–230. https://doi.org/10.1016/0273-1223(95)00270-W

Greaves, J., Stone, D., Wu, Z., Bibby, K., 2020. Persistence of emerging viral fecal indicators in large-scale freshwater mesocosms. Water Res X 9, 100067. https://doi.org/10.1016/j.wroa.2020.100067

Guelin, A., 1948. “Etude quantitative de bacteriophages de la mer.” Ann Inst Pateur 74, 104–112.

Guerin, E., Shkoporov, A., Stockdale, S.R., Clooney, A.G., Ryan, F.J., Sutton, T.D.S., Draper, L.A., Gonzalez-Tortuero, E., Ross, R.P., Hill, C., 2018. Biology and taxonomy of crAss-like bacteriophages, the most abundant virus in the human gut. Cell Host Microbe 24, 653–664. https://doi.org/10.1016/j.chom.2018.10.002

Guerin, E., Shkoporov, A.N., Stockdale, S.R., Comas, J.C., Khokhlova, E. V., Clooney, A.G., Daly, K.M., Draper, L.A., Stephens, N., Scholz, D., Ross, R.P., Hill, C., 2021. Isolation and characterisation of ΦcrAss002, a crAss-like phage from the human gut that infects *Bacteroides xylanisolvens*. Microbiome 9, 1–21. https://doi.org/10.1186/s40168-021-01036-7

Guglielmotti, D.M., Mercanti, D.J., Reinheimer, J.A., Quiberoni, A. del L., 2012. Review: Efficiency of physical and chemical treatments on the inactivation of dairy bacteriophages. Front Microbiol 2, 282. https://doi.org/10.3389/fmicb.2011.00282

Gulyaeva, A., Garmaeva, S., Ruigrok, R.A.A.A., Wang, D., Riksen, N.P., Netea, M.G., Wijmenga, C., Weersma, R.K., Fu, J., Vila, A.V., Kurilshikov, A., Zhernakova, A., 2022. Discovery, diversity, and functional associations of crAss-like phages in human gut metagenomes from four Dutch cohorts. Cell Rep 38, 110204. https://doi.org/10.1016/j.celrep.2021.110204

Guo, H., Arambula, D., Ghosh, P., Miller, J.F., 2014. Diversity-generating retroelements in phage and bacterial genomes. Mobile DNA III 2, 1237–1252. https://doi.org/10.1128/microbiolspec.mdna3-0029-2014

Gutiérrez, D., Rodríguez-Rubio, L., Fernández, L., Martínez, B., Rodríguez, A., García, P., 2017. Applicability of commercial phage-based products against *Listeria monocytogenes* for improvement of food safety in Spanish dry-cured ham and food contact surfaces. Food Control 73, 1474–1482. https://doi.org/10.1016/j.foodcont.2016.11.007

Guzmán, C., Jofre, J., Montemayor, M., Lucena, F., 2007. Occurrence and levels of indicators and selected pathogens in different sludges and biosolids. J Appl Microbiol 103, 2420–2429. https://doi.org/10.1111/j.1365-2672.2007.03487.x

Hamza, I.A., Abd-Elmaksoud, S., 2023. Applicability of crAssphage as a performance indicator for viral reduction during activated sludge wastewater treatment. Environmental Science and Pollution Research 30, 50723–50731. https://doi.org/10.1007/s11356-023-25824-w

Harshey, R.M., 2015. Transposable phage Mu. Mobile DNA III 669–691. https://doi.org/10.1128/microbiolspec.mdna3-0007-2014

Harwood, V.J., Boehm, A.B., Sassoubre, L.M., Vijayavel, K., Stewart, J.R., Fong, T.T., Caprais, M.P., Converse, R.R., Diston, D., Ebdon, J., Fuhrman, J.A., Gourmelon, M., Gentry-Shields, J., Griffith, J.F., Kashian, D.R., Noble, R.T., Taylor, H., Wicki, M., 2013. Performance of viruses and bacteriophages for fecal source determination in a multi-laboratory, comparative study. Water Res 47, 6929–6943. https://doi.org/10.1016/j.watres.2013.04.064

Harwood, V.J., Levine, A.D., Scott, T.M., Chivukula, V., Lukasik, J., Farrah, S.R., Rose, J.B., 2005. Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. Appl Environ Microbiol 71, 3163–3170. https://doi.org/10.1128/AEM.71.6.3163-3170.2005

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K., Korajkic, A., 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. FEMS Microbiol Rev 38, 1–40. https://doi.org/10.1111/1574-6976.12031

Hedžet, S., Rupnik, M., Accetto, T., 2022. Broad host range may be a key to long-term persistence of bacteriophages infecting intestinal *Bacteroidaceae* species. Sci Rep 12, 21098. https://doi.org/10.1038/s41598-022-25636-x

Hehemann, J.H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., Michel, G., 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. Nature 464, 908–912. https://doi.org/10.1038/nature08937

Heilmann, S., Sneppen, K., Krishna, S., 2012. Coexistence of phage and bacteria on the boundary of self-organized refuges. Proc Natl Acad Sci U S A 109, 12828–12833. https://doi.org/10.1073/pnas.1200771109

Hernando-Amado, S., Coque, T.M., Baquero, F., Martínez, J.L., 2019. Defining and combating antibiotic resistance from One Health and Global Health perspectives. Nat Microbiol 4, 1432–1442. https://doi.org/10.1038/s41564-019-0503-9

Hesse, S., Rajaure, M., Wall, E., Johnson, J., Bliskovsky, V., Gottesman, S., Adhya, S., 2020. Phage resistance in multidrug-resistant *Klebsiella pneumoniae* ST258 evolves via diverse mutations that culminate in impaired adsorption. mBio 11, 10–1128. https://doi.org/10.1128/MBIO.02530-19

Hewitson, J.P., Maizels, R.M., 2014. Vaccination against helminth parasite infections. Expert Rev Vaccines. https://doi.org/10.1586/14760584.2014.893195

Hoffmann, C., Dollive, S., Grunberg, S., Chen, J., Li, H., Wu, G.D., Lewis, J.D., Bushman, F.D., 2013. Archaea and Fungi of the Human Gut Microbiome: Correlations with Diet and Bacterial Residents. PLoS One 8, e66019. https://doi.org/10.1371/journal.pone.0066019

Holcomb, D.A., Stewart, J.R., 2020. Microbial indicators of fecal pollution: recent progress and challenges in assessing water quality. Curr Environ Health Rep 7, 311–324. https://doi.org/10.1007/s40572-020-00278-1

Holm, R.H., Nagarkar, M., Yeager, R.A., Talley, D., Chaney, A.C., Rai, J.P., Mukherjee, A., Rai, S.N., Bhatnagar, A., Smith, T., 2022. Surveillance of RNase P, PMMoV, and CrAssphage in wastewater as indicators of human fecal concentration across urban sewer neighborhoods, Kentucky. FEMS Microbes 3, 1–12. https://doi.org/10.1093/femsmc/xtac003

Holmfeldt, K., Solonenko, N., Shah, M., Corrier, K., Riemann, L., VerBerkmoes, N.C., Sullivan, M.B., 2013. Twelve previously unknown phage genera are ubiquitous in global oceans. Proc Natl Acad Sci U S A 110, 12798–12803. https://doi.org/10.1073/pnas.1305956110

Honap, T.P., Sankaranarayanan, K., Schnorr, S.L., Ozga, A.T., Warinner, C., Lewis, C.M., 2020. Biogeographic study of human gut-associated crAssphage suggests impacts from industrialization and recent expansion. PLoS One 15, e0226930. https://doi.org/10.1371/journal.pone.0226930

Hooper, L. V., MacPherson, A.J., 2010. Immune adaptations that maintain homeostasis with the intestinal microbiota. Nat Rev Immunol 10, 159–169. https://doi.org/10.1038/nri2710

Horvath, T.D., Ihekweazu, F.D., Haidacher, S.J., Ruan, W., Engevik, K.A., Fultz, R., Hoch, K.M., Luna, R.A., Oezguen, N., Spinler, J.K., Haag, A.M., Versalovic, J., Engevik, M.A., 2022. *Bacteroides ovatus* colonization influences the abundance of intestinal short chain fatty acids and neurotransmitters. iScience 25, 104158. https://doi.org/10.1016/j.isci.2022.104158

Howard-Varona, C., Hargreaves, K.R., Abedon, S.T., Sullivan, M.B., 2017. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. ISME Journal 11, 1511–1520. https://doi.org/10.1038/ismej.2017.16

Hryckowian, A.J., Merrill, B.D., Porter, N.T., Van Treuren, W., Nelson, E.J., Garlena, R.A., Russell, D.A., Martens, E.C., Sonnenburg, J.L., 2020. *Bacteroides thetaiotaomicron*-infecting bacteriophage isolates inform sequence-based host range predictions. Cell Host Microbe 28, 371–379. https://doi.org/10.1016/j.chom.2020.06.011

Hsu, B.B., Gibson, T.E., Yeliseyev, V., Liu, Q., Lyon, L., Bry, L., Silver, P.A., Gerber, G.K., 2019. Dynamic modulation of the gut microbiota and metabolome by bacteriophages in a mouse model. Cell Host Microbe 25, 803-814.e5. https://doi.org/10.1016/j.chom.2019.05.001

Hunter, M., Fusco, D., 2022. Superinfection exclusion: A viral strategy with short-term benefits and long-term drawbacks. PLoS Comput Biol 18, e1010125. https://doi.org/10.1371/journal.pcbi.1010125

Hurwitz, B.L., U'Ren, J.M., 2016. Viral metabolic reprogramming in marine ecosystems. Curr Opin Microbiol 31, 161–168. https://doi.org/10.1016/j.mib.2016.04.002

Huttenhower, C., Gevers, D., Knight, R., Koren, O., Koren, S., Kota, K.C., Kovar, C.L., Kyrpides, N.C., La Rosa, P.S., Lee, S.L., Lemon, K.P., Lennon, N., Lewis, C.M., Lewis, L., Ley, R.E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C.C., Lozupone, C.A., Dwayne Lunsford, R., Madden, T., Mahurkar, A.A., Mannon, P.J., Mardis, E.R., Markowitz, V.M., Mavromatis, K., McCorrison, J.M., McDonald, D., McEwen, J., McGuire, A.L., McInnes, P., Mehta, T., Mihindukulasuriya, K.A., Miller, J.R., Minx, P.J., Newsham, I., Nusbaum, C., Oglaughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S.M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K.S., Pop, M., Priest, M.E., Proctor, L.M., Qin, X., Raes, J., Ravel, J., Reid, J.G., Rho, M., Rhodes, R., Riehle, K.P., Rivera, M.C., Rodriguez-Mueller, B., Rogers, Y.H., Ross, M.C., Russ, C., Sanka, R.K., Sankar, P., Fah Sathirapongsasuti, J., Schloss, J.A., Schloss, P.D., Schmidt, T.M., Scholz, M., Schriml, L., Schubert, A.M., Segata, N., Segre, J.A., Shannon, W.D., Sharp, R.R., Sharpton, T.J., Shenoy, N., Sheth, N.U., Simone, G.A., Singh, I., Smillie, C.S., Sobel, J.D., Sommer, D.D., Spicer, P., Sutton, G.G., Sykes, S.M., Tabbaa, D.G., Thiagarajan, M., Tomlinson, C.M., Torralba, M., Treangen, T.J., Truty, R.M., Vishnivetskaya, T.A., Walker, J., Wang, L., Wang, Z., Ward, D. V., Warren, W., Watson, M.A., Wellington, C., Wetterstrand, K.A., White, J.R., Wilczek-Boney, K., Wu, Y., Wylie, K.M., Wylie, T., Yandava, C., Ye, L., Yooseph, S., Youmans, B.P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J.D., Birren, B.W., Gibbs, R.A., Highlander, S.K., Methé, B.A., Nelson, K.E., Petrosino, J.F., Weinstock, G.M., Wilson, R.K., The Human Microbiome Project Consortium, 2012. Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214. https://doi.org/10.1038/nature11234

Hyman, P., Abedon, S.T., 2009. Practical methods for determining phage growth parameters, in: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization and Interactions. Humana Press, pp. 175–202.

IAWPRC Study group on health related water microbiology, 1991. Bacteriophages as model viruses in water quality control. Wtaer Research 25, 529–45. https://doi.org/10.1016/0043-1354(91)90126-B

Imai, T., Inoue, R., Nishida, A., Yokota, Y., Morishima, S., Kawahara, M., Kusada, H., Tamaki, H., Andoh, A., 2022. Features of the gut prokaryotic virome of Japanese patients with Crohn's disease. J Gastroenterol 57, 559–570. https://doi.org/10.1007/s00535-022-01882-8

Imamovic, L., Ballesté, E., Martínez-Castillo, A., García-Aljaro, C., Muniesa, M., 2016. Heterogeneity in phage induction enables the survival of the lysogenic population. Environ Microbiol 18, 957–969. https://doi.org/10.1111/1462-2920.13151

Ishii, S., Sadowsky, M.J., 2008. *Escherichia coli* in the environment: Implications for water quality and human health. Microbes Environ 23, 101–108. https://doi.org/10.1264/jsme2.23.101

Jagals, P., Grabow, W., 1996. An evaluation of sorbitol-fermenting bifidobacteria as specific indicators of human faecal pollution of environmental water. Water SA 22, 235–238. https://doi.org/10.1111/j.1365-2672.1983.tb01331.x

Jamal, M., Bukhari, S.M.A.U.S., Andleeb, S., Ali, M., Raza, S., Nawaz, M.A., Hussain, T., Rahman, S. u., Shah, S.S.A., 2019. Bacteriophages: an overview of the control strategies against multiple bacterial infections in different fields. J Basic Microbiol 59, 123–133. https://doi.org/10.1002/jobm.201800412

Jebri, S., Muniesa, M., Jofre, J., 2017. General and host-associated bacteriophage indicators of faecal pollution, in: Global Water Pathogen Project. Part 2. Indicators and Microbial Source Tracking Markers. Michigan State University, E. Lansing, MI, UNESCO. https://doi.org/10.14321/waterpathogens.7

Jebri, S., Yahya, M., 2013. Unité d'application agricole et médicales des technologies nucléaires, Technopole de Sidi Thabet, Tunisia. Personnel communication.

Jennings, W.C., Gálvez-Arango, E., Prieto, A.L., Boehm, A.B., 2020. CrAssphage for fecal source tracking in Chile: Covariation with norovirus, HF183, and bacterial indicators. Water Res X 9, 100071. https://doi.org/10.1016/j.wroa.2020.100071

Jiang, S.C., Chu, W., Olson, B.H., He, J.W., Choi, S., Zhang, J., Le, J.Y., Gedalanga, P.B., 2007. Microbial source tracking in a small southern California urban watershed indicates wild animals and growth as the source of fecal bacteria. Appl Microbiol Biotechnol 76, 927–934. https://doi.org/10.1007/s00253-007-1047-0

Jofre, J., 2009. Is the replication of somatic coliphages in water environments significant? J Appl Microbiol 106, 1059–1069. https://doi.org/10.1111/j.1365-2672.2008.03957.x

Jofre, J., 2007. Indicators of waterborne enteric viruses. Perspect Med Virol 17, 227–249. https://doi.org/10.1016/S0168-7069(07)17011-7

Jofre, J., Blanch, A.R., Lucena, F., Muniesa, M., 2014. Bacteriophages infecting Bacteroides as a marker for microbial source tracking. Water Res 55, 1–11. https://doi.org/10.1016/j.watres.2014.02.006

Jofre, J., Bosch, A., Lucena, F., Girones, R., Tartera, C., 1986. Evaluation of *Bacteroides fragilis* bacteriophages as indicators of the virological quality of water. Wat. Sci. Tech 18, 167–173.

Jofre, J., Lucena, F., Blanch, A.R., Muniesa, M., 2016. Coliphages as model organisms in the characterization and management of water resources. Water 8(5), 199. https://doi.org/10.3390/w8050199

Jończyk, E., Kłak, M., Międzybrodzki, R., Górski, A., 2011. The influence of external factors on bacteriophages-review. Folia Microbiol (Praha) 56, 191–200. https://doi.org/10.1007/s12223-011-0039-8

Jung, J.H., Yoo, C.H., Koo, E.S., Kim, H.M., Na, Y., Jheong, W.H., Jeong, Y.S., 2011. Occurrence of norovirus and other enteric viruses in untreated groundwaters of Korea. J Water Health 9, 544–555. https://doi.org/10.2166/wh.2011.142

Jurzik, L., Hamza, I.A., Puchert, W., Überla, K., Wilhelm, M., 2010. Chemical and microbiological parameters as possible indicators for human enteric viruses in surface water. Int J Hyg Environ Health 213, 210–216. https://doi.org/10.1016/j.ijheh.2010.05.005

Kahler, A.M., Cromeans, T.L., Roberts, J.M., Hill, V.R., 2010. Effects of source water quality on chlorine inactivation of adenovirus, coxsackievirus, echovirus, and murine norovirus. Appl Environ Microbiol 76, 5159–5164. https://doi.org/10.1128/AEM.00869-10

Kamada, N., Seo, S.U., Chen, G.Y., Núñez, G., 2013. Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol 13, 321–335. https://doi.org/10.1038/nri3430

Karlsson, F.H., Ussery, D.W., Nielsen, J., Nookaew, I., 2011. A closer look at *Bacteroides*: phylogenetic relationship and genomic implications of a life in the human gut. Microb Ecol 61, 473–485. https://doi.org/10.1007/s00248-010-9796-1

Kator, H., Rhodes, M., 1992. Evaluation of *Bacteroides fragilis* bacteriophage, a candidate human-specific indicator of fecal contamination for shellfish-growing waters. Final Report. Gloucester Point, VA: Virginia Institute of Marine Sciences, College of Williams and Mary.

Katzenelson, E., Buium, I., Shuval, H.I., 1976. Risk of communicable disease infection associated with wastewater irrigation in agricultural settlements. New Series 194, 944–946. https://doi.org/10.1126/science.982051

Kerby, G.P., Gowdy, R.A., Dillon, E.S., Dillon, M.L., Csâky, T.Z., Sharp, D.G., Beard, J.W., 1949. Purification, pH stability and sedimentation properties of the T7 bacteriophage of *Escherichia coli*. The Journal of Immunology 63, 93–107. https://doi.org/10.4049/jimmunol.63.1.93

Kilaru, P., Hill, D., Anderson, K., Collins, M.B., Green, H., Kmush, B.L., Larsen, D.A., 2023. Wastewater surveillance for infectious disease: a systematic review. Am J Epidemiol 192, 305–322. https://doi.org/10.1093/aje/kwac175

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., Wuertz, S., 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales:* A Bayesian approach. Water Res 41, 3701–3715. https://doi.org/10.1016/j.watres.2007.06.037

Kim, M.S., Bae, J.W., 2018. Lysogeny is prevalent and widely distributed in the murine gut microbiota. ISME Journal 12, 1127–1141. https://doi.org/10.1038/s41396-018-0061-9

Knecht, L.E., Veljkovic, M., Fieseler, L., 2020. Diversity and function of phage encoded depolymerases. Front Microbiol 10, 2949. https://doi.org/10.3389/fmicb.2019.02949

Knowles, B., Silveira, C.B., Bailey, B.A., Barott, K., Cantu, V.A., Cobian-Guëmes, A.G., Coutinho, F.H., Dinsdale, E.A., Felts, B., Furby, K.A., George, E.E., Green, K.T., Gregoracci, G.B., Haas, A.F., Haggerty, J.M., Hester, E.R., Hisakawa, N., Kelly, L.W., Lim, Y.W., Little, M., Luque, A., McDole-Somera, T., McNair, K., De Oliveira, L.S., Quistad, S.D., Robinett, N.L., Sala, E., Salamon, P., Sanchez, S.E., Sandin, S., Silva, G.G.Z., Smith, J., Sullivan, C., Thompson, C., Vermeij, M.J.A., Youle, M., Young, C., Zgliczynski, B., Brainard, R., Edwards, R.A., Nulton, J., Thompson, F., Rohwer, F., 2016. Lytic to temperate switching of viral communities. Nature 531, 466–470. https://doi.org/10.1038/nature17193

Kollarcikova, M., Faldynova, M., Matiasovicova, J., Jahodarova, E., Kubasova, T., Seidlerova, Z., Babak, V., Videnska, P., Cizek, A., Rychlik, I., 2020. Different *Bacteroides* species colonise human and chicken intestinal tract. Microorganisms 8, 1–14. https://doi.org/10.3390/microorganisms8101483

Kongprajug, A., Mongkolsuk, S., Sirikanchana, K., 2019. CrAssphage as a potential human sewage marker for microbial source tracking in Southeast Asia. Environ Sci Technol Lett 6, 156–164. https://doi.org/10.1021/acs.estlett.9b00041

Koonin, E. V., Yutin, N., 2020. The crAss-like phage group: How metagenomics reshaped the human virome. Trends Microbiol 28, 349–359. https://doi.org/10.1016/j.tim.2020.01.010

Korajkic, A., Wanjugi, P., Brooks, L., Cao, Y., Harwood, V.J., 2019. Persistence and decay of fecal microbiota in aquatic habitats. Microbiology and Molecular Biology Reviews 83, e00005-19.

Koskella, B., Brockhurst, M.A., 2014. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev 38, 916–931. https://doi.org/10.1111/1574-6976.12072

Kostic, A.D., Howitt, M.R., Garrett, W.S., 2013. Exploring host-microbiota interactions in animal models and humans. Genes Dev 27, 701–718. https://doi.org/10.1101/gad.212522.112

Kotay, S., Chai, W., Guilford, W., Barry, K., Mathers, A.J., 2017. Spread from the sink to the patient: In situ study using green fluorescent protein (GFP)-expressing *Escherichia coli* to model bacterial dispersion from hand-washing sink-trap reservoirs. Appl Environ Microbiol 83, e03327-16. https://doi.org/10.1128/AEM.03327-16

Krinos, C.M., Coyne, M.J., Weinacht, K.G., Tzianabos, A.O., Kasper, D.L., Comstock, L.E., 2001. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature 414, 555–558.

Krishnamurthy, S.R., Wang, D., 2017. Origins and challenges of viral dark matter. Virus Res 239, 136–142. https://doi.org/10.1016/j.virusres.2017.02.002

Kwan, S.-Y., Sabotta, C.M., Cruz, L.R., Wong, M.C., Ajami, N.J., McCormick, J.B., Fisher-Hoch, S.P., Beretta, L., 2024. Gut phageome in Mexican Americans: a population at high risk for metabolic dysfunction-associated steatotic liver disease and diabetes. mSystems 9, e00434-24. https://doi.org/10.1128/msystems.00434-24

Lee, S.M., Donaldson, G.P., Mikulski, Z., Boyajian, S., Ley, K., Mazmanian, S.K., 2013. Bacterial colonization factors control specificity and stability of the gut microbiota. Nature 501, 426–429. https://doi.org/10.1038/nature12447

Lee, S.Y., Yang, J., Lee, J.H., 2023. Improvement of crAssphage detection/quantification method and its extensive application for food safety. Front Microbiol 14, 1185788. https://doi.org/10.3389/fmicb.2023.1185788

Lee, Y.J., Dai, N., Walsh, S.E., Müller, S., Fraser, M.E., Kauffman, K.M., Guan, C., Corrêa, I.R., Weigele, P.R., 2018. Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses. Proc Natl Acad Sci U S A 115, E3116–E3125. https://doi.org/10.1073/pnas.1714812115

Ley, R.E., Hamady, M., Lozuopone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., Gordon, J.I., 2008. Evolution of mammals and their gut microbiomes. Science (1979) 320, 1643–1647. https://doi.org/10.1126/science.1155390

Ley, R.E., Peterson, D.A., Gordon, J.I., 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124, 837–848. https://doi.org/10.1016/j.cell.2006.02.017

Li, E., Saleem, F., Edge, T.A., Schellhorn, H.E., 2024. Assessment of crAssphage as a human fecal source tracking marker in the lower Great Lakes. Science of the Total Environment 912, 168840. https://doi.org/10.1016/j.scitotenv.2023.168840

Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J.R., Prifti, E., Nielsen, T., Juncker, A.S., Manichanh, C., Chen, B., Zhang, W., Levenez, F., Wang, Juan, Xu, X., Xiao, L., Liang, S., Zhang, D., Zhang, Z., Chen, W., Zhao, H., Al-Aama, J.Y., Edris, S., Yang, H., Wang, Jian, Hansen, T., Nielsen, H.B., Brunak, S., Kristiansen, K., Guarner, F., Pedersen, O., Doré, J., Ehrlich, S.D., Bork, P., Wang, Jun, Pons, N., Le Chatelier, E., Batto, J.M., Kennedy, S., Haimet, F.,

Winogradski, Y., Pelletier, E., Lepaslier, D., Artiguenave, F., Bruls, T., Weissenbach, J., Turner, K., Parkhill, J., Antolin, M., Casellas, F., Borruel, N., Varela, E., Torrejon, A., Denariaz, G., Derrien, M., van Hylckama Vlieg, J.E.T., Viega, P., Oozeer, R., Knoll, J., Rescigno, M., Brechot, C., M'rini, C., Mérieux, A., Yamada, T., Tims, S., Zoetendal, E.G., Kleerebezem, M., de Vos, W.M., Cultrone, A., Leclerc, M., Juste, C., Guedon, E., Delorme, C., Layec, S., Khaci, G., van de Guchte, M., Vandemeulebrouck, G., Jamet, A., Dervyn, R., Sanchez, N., Blottière, H., Maguin, E., Renault, P., Tap, J., Mende, D.R., 2014. An integrated catalog of reference genes in the human gut microbiome. Nat Biotechnol 32, 834–841. https://doi.org/10.1038/nbt.2942

Li, Y., Gordon, E., Shean, R.C., Idle, A., Deng, X., Greninger, A.L., Delwart, E., 2021. CrAssphage and its bacterial host in cat feces. Sci Rep 11, 815. https://doi.org/10.1038/s41598-020-80076-9

Liang, Y., Jin, X., Huang, Y., Chen, S., 2018. Development and application of a real-time polymerase chain reaction assay for detection of a novel gut bacteriophage (crAssphage). J Med Virol 90, 464–468. https://doi.org/10.1002/jmv.24974

Liang, Y.Y., Zhang, W., Tong, Y.G., Chen, S.P., 2016. CrAssphage is not associated with diarrhoea and has high genetic diversity. Epidemiol Infect 144, 3549–3553. https://doi.org/10.1017/S095026881600176X

Lin, J., Ganesh, A., 2013. Water quality indicators: Bacteria, coliphages, enteric viruses. Int J Environ Health Res 23, 484–506. https://doi.org/10.1080/09603123.2013.769201

Lindner, B.G., Gerhardt, K., Feistel, D.J., Rodriguez-R, L.M., Hatt, J.K., Konstantinidis, K.T., 2024. A user's guide to the bioinformatic analysis of shotgun metagenomic sequence data for bacterial pathogen detection. Int J Food Microbiol 410, 110488. https://doi.org/10.1016/j.ijfoodmicro.2023.110488

Liu, J., Dehbi, M., Moeck, G., Arhin, F., Banda, P., Bergeron, D., Callejo, M., Ferretti, V., Ha, N., Kwan, T., McCarty, J., Srikumar, R., Williams, D., Wu, J.J., Gros, P., Pelletier, J., DuBow, M., 2004. Antimicrobial drug discovery through bacteriophage genomics. Nat Biotechnol 22, 185–191. https://doi.org/10.1038/nbt932

Liu, J., Kabir, F., Manneh, J., Lertsethtakarn, P., Begum, S., Gratz, J., Becker, S.M., Operario, D.J., Taniuchi, M., Janaki, L., Platts-Mills, J.A., Haverstick, D.M., Kabir, M., Sobuz, S.U., Nakjarung, K., Sakpaisal, P., Silapong, S., Bodhidatta, L., Qureshi, S., Kalam, A., Saidi, Q., Swai, N., Mujaga, B., Maro, A., Kwambana, B., Dione, M., Antonio, M., Kibiki, G., Mason, C.J., Haque, R., Iqbal, N., Zaidi, A.K.M., Houpt, E.R., 2014. Development and assessment of molecular diagnostic tests for 15 enteropathogens causing childhood diarrhoea: A multicentre study. Lancet Infect Dis 14, 716–724. https://doi.org/10.1016/S1473-3099(14)70808-4

Liu, J., Platts-Mills, J.A., Juma, J., Kabir, F., Nkeze, J., Okoi, C., Operario, D.J., Uddin, J., Ahmed, S., Alonso, P.L., Antonio, M., Becker, S.M., Blackwelder, W.C., Breiman, R.F., Faruque, A.S.G., Fields, B., Gratz, J., Haque, R., Hossain, A., Hossain, M.J., Jarju, S., Qamar, F., Iqbal, N.T., Kwambana, B., Mandomando, I., McMurry, T.L., Ochieng, C., Ochieng, J.B., Ochieng, M., Onyango, C.,

Panchalingam, S., Kalam, A., Aziz, F., Qureshi, S., Ramamurthy, T., Roberts, J.H., Saha, D., Sow, S.O., Stroup, S.E., Sur, D., Tamboura, B., Taniuchi, M., Tennant, S.M., Toema, D., Wu, Y., Zaidi, A., Nataro, J.P., Kotloff, K.L., Levine, M.M., Houpt, E.R., 2016. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. The Lancet 388, 1291–1301. https://doi.org/10.1016/S0140-6736(16)31529-X

Loc-Carrillo, C., Abedon, S.T., 2011. Pros and cons of phage therapy. Bacteriophage 1, 111–114. https://doi.org/10.4161/bact.1.2.14590

Lourenço, M., Chaffringeon, L., Lamy-Besnier, Q., Pédron, T., Campagne, P., Eberl, C., Bérard, M., Stecher, B., Debarbieux, L., De Sordi, L., 2020. The spatial heterogeneity of the gut limits predation and fosters coexistence of bacteria and bacteriophages. Cell Host Microbe 28, 390-401.e5. https://doi.org/10.1016/j.chom.2020.06.002

Lourenço, M., De Sordi, L., Debarbieux, L., 2018. The diversity of bacterial lifestyles hampers bacteriophage tenacity. Viruses 10, 327. https://doi.org/10.3390/v10060327

Love, D.C., Rodriguez, R.A., Gibbons, C.D., Griffith, J.F., Yu, Q., Stewart, J.R., Sobsey, M.D., 2014. Human viruses and viral indicators in marine water at two recreational beaches in Southern California, USA. J Water Health 12, 136–150. https://doi.org/10.2166/wh.2013.078

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., Knight, R., 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489, 220–230. https://doi.org/10.1038/nature11550

Lucena, F., Duran, A.E., Morón, A., Calderón, E., Campos, C., Gantzer, C., Skraber, S., Jofre, J., 2004. Reduction of bacterial indicators and bacteriophages infecting faecal bacteria in primary and secondary wastewater treatments. J Appl Microbiol 97, 1069–1076. https://doi.org/10.1111/j.1365-2672.2004.02397.x

Lucena, F., Ribas, F., Duran, A.E., Skraber, S., Gantzer, C., Campos, C., Morón, A., Calderón, E., Jofre, J., 2006. Occurrence of bacterial indicators and bacteriophages infecting enteric bacteria in groundwater in different geographical areas. J Appl Microbiol 101, 96–102. https://doi.org/10.1111/j.1365-2672.2006.02907.x

Malla, B., Ghaju Shrestha, R., Tandukar, S., Sherchand, J.B., Haramoto, E., 2019. Performance evaluation of human-specific viral markers and application of pepper mild mottle virus and crAssphage to environmental water samples as fecal pollution markers in the Kathmandu Valley, Nepal. Food Environ Virol 11, 274–287. https://doi.org/10.1007/s12560-019-09389-x

Manrique, P., Bolduc, B., Walk, S.T., van der Oost, J., de Vos, W.M., Young, M.J., 2016. Healthy human gut phageome. Proceedings of the National Academy of Sciences 113, 10400–10405. https://doi.org/10.1073/pnas.1601060113

Manrique, P., Dills, M., Young, M.J., 2017. The human gut phage community and its implications for health and disease. Viruses 9, 141. https://doi.org/10.3390/v9060141

Mäntynen, S., Laanto, E., Oksanen, H.M., Poranen, M.M., Díaz-Muñoz, S.L., 2021. Black box of phage-bacterium interactions: exploring alternative phage infection strategies. Open Biol 11, 210188. https://doi.org/10.1098/rsob.210188

Maraccini, P.A., Mattioli, M.C.M., Sassoubre, L.M., Cao, Y., Griffith, J.F., Ervin, J.S., Van De Werfhorst, L.C., Boehm, A.B., 2016. Solar inactivation of Enterococci and *Escherichia coli* in natural waters: Effects of water absorbance and depth. Environ Sci Technol 50, 5068–5076. https://doi.org/10.1021/acs.est.6b00505

Mari, L., Bertuzzo, E., Righetto, L., Casagrandi, R., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., 2012. Modelling cholera epidemics: The role of waterways, human mobility and sanitation. J R Soc Interface 9, 376–388. https://doi.org/10.1098/rsif.2011.0304

Marsalek, J., Rochfort, Q., 2004. Urban wet-weather flows: Sources of fecal contamination impacting on recreational waters and threatening drinking-water sources. J Toxicol Environ Health A 67, 1765–1777. https://doi.org/10.1080/15287390490492430

Martens, E.C., Roth, R., Heuser, J.E., Gordon, J.I., 2009. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. Journal of Biological Chemistry 284, 18445–18457. https://doi.org/10.1074/jbc.M109.008094

Martin, N.A., Sala-Comorera, L., Gao, G., Corkery, A., Reynolds, L.J., Nolan, T.M., Whitty, M., O'Sullivan, J.J., Meijer, W.G., 2023. Inclusion of hydrodynamic properties of bathing waters is critical in selecting faecal indicators to assess public health impacts of faecal contamination. Water Res 242, 120137. https://doi.org/10.1016/j.watres.2023.120137

Martinez-Hernandez, F., Fornas, O., Lluesma Gomez, M., Bolduc, B., De La Cruz Peña, M.J., Martínez, J.M., Anton, J., Gasol, J.M., Rosselli, R., Rodriguez-Valera, F., Sullivan, M.B., Acinas, S.G., Martinez-Garcia, M., 2017. Single-virus genomics reveals hidden cosmopolitan and abundant viruses. Nat Commun 8, 15892. https://doi.org/10.1038/ncomms15892

McCall, C., Wu, H., Miyani, B., Xagoraraki, I., 2020. Identification of multiple potential viral diseases in a large urban center using wastewater surveillance. Water Res 184, 116160. https://doi.org/10.1016/j.watres.2020.116160

McKee, A.M., Spear, S.F., Pierson, T.W., 2015. The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. Biol Conserv 183, 70–76. https://doi.org/10.1016/j.biocon.2014.11.031

McKinney, C.W., Pruden, A., 2012. Ultraviolet disinfection of antibiotic resistant bacteria and their antibiotic resistance genes in water and wastewater. Environ Sci Technol 46, 13393–13400. https://doi.org/10.1021/es303652q

Mclaughlin, M.R., Rose, J.B., 2006. Application of *Bacteroides fragilis* phage as an alternative indicator of sewage pollution in Tampa Bay, Florida. Estuaries and Coasts 29, 246–256. https://doi.org/10.1007/BF02781993

McMahon, S.A., Miller, J.L., Lawton, J.A., Kerkow, D.E., Hodes, A., Marti-Renom, M.A., Doulatov, S., Narayanan, E., Sali, A., Miller, J.F., Ghosh, P., 2005. The C-type lectin fold as an evolutionary solution for massive sequence variation. Nat Struct Mol Biol 12, 886–892. https://doi.org/10.1038/nsmb992

Mcminn, B.R., Korajkic, A., Ashbolt, N.J., 2014. Evaluation of *Bacteroides fragilis* GB-124 bacteriophages as novel human-associated faecal indicators in the United States. Lett Appl Microbiol 59, 115–121. https://doi.org/10.1111/lam.12252

Meays, C.L., Broersma, K., Nordin, R., Mazumder, A., 2004. Source tracking fecal bacteria in water: A critical review of current methods. J Environ Manage 73, 71–79. https://doi.org/10.1016/j.jenvman.2004.06.001

Méndez, J., Audicana, A., Isern, A., Llaneza, J., Moreno, B., Tarancón, M.L., Jofre, J., Lucena, F., 2004. Standardised evaluation of the performance of a simple membrane filtration-elution method to concentrate bacteriophages from drinking water. J Virol Methods 117, 19–25. https://doi.org/10.1016/j.jviromet.2003.11.013

Meuchi, Y., Nakada, M., Kuroda, K., Hanamoto, S., Hata, A., 2023. Applicability of F-specific bacteriophage subgroups, PMMoV and crAssphage as indicators of source specific fecal contamination and viral inactivation in rivers in Japan. PLoS One 18, e0288454. https://doi.org/10.1371/journal.pone.0288454

Minot, S., Bryson, A., Chehoud, C., Wu, G.D., Lewis, J.D., Bushman, F.D., 2013. Rapid evolution of the human gut virome. Proc Natl Acad Sci U S A 110, 12450–12455. https://doi.org/10.1073/pnas.1300833110

Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S.A., Wu, G.D., Lewis, J.D., Bushman, F.D., 2011. The human gut virome: Inter-individual variation and dynamic response to diet. Genome Res 21, 1616–1625. https://doi.org/10.1101/gr.122705.111

Mirzaei, M.K., Maurice, C.F., 2017. Ménage à trois in the human gut: Interactions between host, bacteria and phages. Nat Rev Microbiol 15, 397–408. https://doi.org/10.1038/nrmicro.2017.30

Mocé-Llivina, L., Lucena, F., Jofre, J., 2005. Enteroviruses and bacteriophages in bathing waters. Appl Environ Microbiol 71(11), 6838–6844. https://doi.org/10.1128/AEM.71.11.6838-6844.2005

Mocé-Llivina, L., Muniesa, M., Pimenta-Vale, H., Lucena, F., Jofre, J., 2003. Survival of bacterial indicator species and bacteriophages after thermal treatment of sludge and sewage. Appl Environ Microbiol 69, 1452–1456. https://doi.org/10.1128/AEM.69.3.1452-1456.2003

Modi, S.R., Lee, H.H., Spina, C.S., Collins, J.J., 2013. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. Nature 499, 219–222. https://doi.org/10.1038/nature12212

Momba, M., Ebdon, J., Kamika, I., Verbyla, M., 2019. Using indicators to assess microbial treatment and disinfection efficacy, in: Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project). Michigan State University. https://doi.org/10.14321/waterpathogens.9

Morimoto, D., Tominaga, K., Takabe, H., Sulcius, S., Yoshida, T., 2022. Viral nature of the aquatic ecosystems. The biological role of a virus 3–25. https://doi.org/10.1007/978-3-030-85395-2_1

Morozova, V., Fofanov, M., Tikunova, N., Babkin, I., Morozov, V. V., Tikunov, A., 2020. First crAss-like phage genome encoding the diversity-generating retroelement (DGR). Viruses 12, 573. https://doi.org/10.3390/v12050573

Morrison, C.M., Betancourt, W.Q., Quintanar, D.R., Lopez, G.U., Pepper, I.L., Gerba, C.P., 2020. Potential indicators of virus transport and removal during soil aquifer treatment of treated wastewater effluent. Water Res 177, 115812. https://doi.org/10.1016/j.watres.2020.115812

Muniesa, M., Lucena, F., Blanch, A.R., Payán, A., Jofre, J., 2012. Use of abundance ratios of somatic coliphages and bacteriophages of *Bacteroides thetaiotaomicron* GA17 for microbial source identification. Water Res 46, 6410–6418. https://doi.org/10.1016/j.watres.2012.09.015

Muniesa, M., Mocé-Llivina, L., Katayama, H., Jofre, J., 2003. Bacterial host strains that support replication of somatic coliphages. Antonie Van Leeuwenhoek 83, 305–315. https://doi.org/10.1023/a:1023384714481

Muniesa, M., Serra-Moreno, R., Jofre, J., 2004. Free Shiga toxin bacteriophages isolated from sewage showed diversity although the stx genes appeared conserved. Environ Microbiol 6, 716–725. https://doi.org/10.1111/j.1462-2920.2004.00604.x

Murray, P.R., Baron, E.J., 2007. Manual of clinical microbiology, Clinical Infectious Diseases. https://doi.org/10.1086/524076

Nakayama-Imaohji, H., Hirota, K., Yamasaki, H., Yoneda, S., Nariya, H., Suzuki, M., Secher, T., Miyake, Y., Oswald, E., Hayashi, T., Kuwahara, T., 2016. DNA inversion regulates outer membrane vesicle production in *Bacteroides fragilis*. PLoS One 11, e0148887. https://doi.org/10.1371/journal.pone.0148887

Nam, S.J., Kim, D.W., Lee, S.H., Koo, O.K., 2023. Assessment of microbial source tracking marker and fecal indicator bacteria on food-contact surfaces in school cafeterias. J Food Prot 86, 100035. https://doi.org/10.1016/j.jfp.2022.100035

Naughton, C., Mihelcic, J., 2017. Introduction to the importance of sanitation. Part one: The health hazard of excreta: theory and control. United States.

Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S., Kyrpides, N.C., 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol 39, 578–585. https://doi.org/10.1038/s41587-020-00774-7

Neff, C.P., Rhodes, M.E., Arnolds, K.L., Collins, C.B., Donnelly, J., Nusbacher, N., Jedlicka, P., Schneider, J.M., McCarter, M.D., Shaffer, M., Mazmanian, S.K., Palmer, B.E., Lozupone, C.A., 2016. Diverse intestinal bacteria contain putative zwitterionic capsular polysaccharides with anti-inflammatory properties. Cell Host Microbe 20, 535–547. https://doi.org/10.1016/j.chom.2016.09.002

Nelson, K.L., Boehm, A.B., Davies-Colley, R.J., Dodd, M.C., Kohn, T., Linden, K.G., Liu, Y., Maraccini, P.A., McNeill, K., Mitch, W.A., Nguyen, T.H., Parker, K.M., Rodriguez, R.A., Sassoubre, L.M., Silverman, A.I., Wigginton, K.R., Zepp, R.G., 2018. Sunlight-mediated inactivation of health-relevant microorganisms in water: a review of mechanisms and modeling approaches. Environ Sci Process Impacts 20, 1089–1122. https://doi.org/10.1039/c8em00047f

Newton, R.J., McLellan, S.L., Dila, D.K., Vineis, J.H., Morrison, H.G., Murat Eren, A., Sogin, M.L., 2015. Sewage reflects the microbiomes of human populations. mBio 6, 10–1128. https://doi.org/10.1128/mBio.02574-14

Nguyen, K.H., Senay, C., Young, S., Nayak, B., Lobos, A., Conrad, J., Harwood, V.J., 2018. Determination of wild animal sources of fecal indicator bacteria by microbial source tracking (MST) influences regulatory decisions. Water Res 144, 424–434. https://doi.org/10.1016/j.watres.2018.07.034

Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., Pettersson, S., 2012. Host-gut microbiota metabolic interactions. Science (1979) 336, 1262–1267. https://doi.org/10.1126/science.1224203

Norman, J.M., Handley, S.A., Baldridge, M.T., Droit, L., Liu, C.Y., Keller, B.C., Kambal, A., Monaco, C.L., Zhao, G., Fleshner, P., Stappenbeck, T.S., McGovern, D.P.B., Keshavarzian, A., Mutlu, E.A., Sauk, J., Gevers, D., Xavier, R.J., Wang, D., Parkes, M., Virgin, H.W., 2015. Disease-specific alterations in the enteric virome in inflammatory bowel disease. Cell 160, 447–460. https://doi.org/10.1016/j.cell.2015.01.002

North Carolina Administration, 2011. North Carolina Administrative Code 15A NCAC 2U. Raleigh, USA.

Nuanualsuwan, S., Cliver, D.O., 2003. Capsid functions of inactivated human picornaviruses and feline calicivirus. Appl Environ Microbiol 69, 350–357. https://doi.org/10.1128/AEM.69.1.350-357.2003

Nunez, C., Kostoulias, X., Peleg, A., Short, F., Qu, Y., 2023. A comprehensive comparison of biofilm formation and capsule production for bacterial survival on hospital surfaces. Biofilm 5, 100105. https://doi.org/10.1016/j.bioflm.2023.100105

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

Oliveira, J., Cunha, A., Castilho, F., Romalde, J.L., Pereira, M.J., 2011. Microbial contamination and purification of bivalve shellfish: Crucial aspects in monitoring and future perspectives - A mini-review. Food Control 22, 805–816. https://doi.org/10.1016/j.foodcont.2010.11.032

Ortmann, A.C., Suttle, C.A., 2009. Determination of virus abundance by epifluorescence microscopy, in: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interacions. pp. 87–95. https://doi.org/10.1007/978-1-60327-164-6_10

O'Shea, Y.A., Boyd, E.F., 2002. Mobilization of the *Vibrio* pathogenicity island between *Vibrio cholerae* isolates mediated by CP-T1 generalized transduction . FEMS Microbiol Lett 214, 153–157. https://doi.org/10.1111/j.1574-6968.2002.tb11339.x

Owen, S. V., Canals, R., Wenner, N., Hammarlöf, D.L., Kröger, C., Hinton, J.C.D., 2020. A window into lysogeny: Revealing temperate phage biology with transcriptomics. Microb Genom 6, e000330. https://doi.org/10.1099/mgen.0.000330

Pal, C., Maciá, M.D., Oliver, A., Schachar, I., Buckling, A., 2007. Coevolution with viruses drives the evolution of bacterial mutation rates. Nature 450, 1079–1081. https://doi.org/10.1038/nature06350

Papp, K., Moser, D., Gerrity, D., 2020. Viral surrogates in potable reuse applications: evaluation of a membrane bioreactor and full advanced treatment. Journal of Environmental Engineering 146, 04019103. https://doi.org/10.1061/(asce)ee.1943-7870.0001617

Papudeshi, B., Vega, A.A., Souza, C., Giles, S.K., Mallawaarachchi, V., Roach, M.J., An, M., Jacobson, N., McNair, K., Mora, M.F., Pastrana, K., Boling, L., Leigh, C., Harker, C., Plewa, W.S., Grigson, S.R., Bouras, G., Decewicz, P., Luque, A., Droit, L., Handley, S.A., Wang, D., Segall, A.M., Dinsdale, E.A., Edwards, R.A., 2023. Host interactions of novel Crassvirales species belonging to multiple families infecting bacterial host, Bacteroides cellulosilyticus WH2. Microb Genom 9, 001100. https://doi.org/10.1099/mgen.0.001100

Parfrey, L.W., Walters, W.A., Knight, R., 2011. Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. Front Microbiol 2, 153. https://doi.org/10.3389/fmicb.2011.00153

Pargin, E., Roach, M.J., Skye, A., Papudeshi, B., Inglis, L.K., Mallawaarachchi, V., Grigson, S.R., Harker, C., Edwards, R.A., Giles, S.K., 2023. The human gut virome: composition, colonization, interactions, and impacts on human health. Front Microbiol 14, 963173. https://doi.org/10.3389/fmicb.2023.963173

Park, G.W., Ng, T.F.F., Freeland, A.L., Marconi, V.C., Boom, J.A., Staat, M.A., Montmayeur, A.M., Browne, H., Narayanan, J., Payne, D.C., Cardemil, C. V., Treffiletti, A., Vinjé, J., 2020. 1CrAssphage as a novel tool to detect human fecal contamination on environmental surfaces and hands. Emerg Infect Dis 26, 1731–1739. https://doi.org/10.3201/eid2608.200346

Patel, P.H., Taylor, V.L., Zhang, C., Getz, L.J., Fitzpatrick, A.D., Davidson, A.R., Maxwell, K.L., 2024. Anti-phage defence through inhibition of virion assembly. Nat Commun 15, 1644. https://doi.org/10.1038/S41467-024-45892-X

Patrick, S., Blakely, G.W., Houston, S., Moore, J., Abratt, V.R., Bertalan, M., Cerdeño-Tárraga, A.M., Quail, M.A., Corton, N., Corton, C., Bignell, A., Barron, A., Clark, L., Bentley, S.D., Parkhill, J., 2010. Twenty-eight divergent polysaccharide loci specifying within- and amongst-strain capsule diversity

in three strains of *Bacteroides fragilis*. Microbiology (N Y) 156, 3255–3269. https://doi.org/10.1099/mic.0.042978-0

Paul, J.H., 2008. Prophages in marine bacteria: Dangerous molecular time bombs or the key to survival in the seas? ISME Journal 2, 579–589. https://doi.org/10.1038/ismej.2008.35

Paul, J.H., Sullivan, M.B., Segall, A.M., Rohwer, F., 2002. Marine phage genomics. Comp Biochem Physiol B Biochem Mol Biol 133, 463–476. https://doi.org/10.1016/S1096-4959(02)00168-9

Payan, A., 2006. Bacteriófagos como modelo en el origen de la contaminación fecal y en aguas regeneradas. University of Barcelona, Barcelona, Spain.

Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G.T., Blanch, A.R., Lucena, F., Jofre, J., Muniesa, M., 2005. Method for isolation of *Bacteroides* bacteriophage host strains suitable for tracking sources of fecal pollution in water. Appl Environ Microbiol 71, 5659–5662. https://doi.org/10.1128/AEM.71.9.5659-5662.2005

Payment, P., Locas, A., 2011. Pathogens in water: value and limits of correlation with microbial indicators. Ground Water 49, 4–11. https://doi.org/10.1111/j.1745-6584.2010.00710.x

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

Peng, J., Tang, J., Barrett, D.M., Sablani, S.S., Anderson, N., Powers, J.R., 2017. Thermal pasteurization of ready-to-eat foods and vegetables: Critical factors for process design and effects on quality. Crit Rev Food Sci Nutr 57, 2970–2995. https://doi.org/10.1080/10408398.2015.1082126

Peterson, D.A., McNulty, N.P., Guruge, J.L., Gordon, J.I., 2007. IgA response to symbiotic bacteria as a mediator of gut homeostasis. Cell Host Microbe 2, 328–339. https://doi.org/10.1016/j.chom.2007.09.013

Pfaender, S., Brinkmann, J., Todt, D., Riebesehl, N., Steinmann, Joerg, Steinmann, Jochen, Pietschmann, T., Steinmann, E., 2015. Mechanisms of methods for hepatitis C virus inactivation. Appl Environ Microbiol 81, 1616–1621. https://doi.org/10.1128/AEM.03580-14

Pietilä, M.K., Laurinavičius, S., Sund, J., Roine, E., Bamford, D.H., 2010. The single-stranded DNA genome of novel archaeal virus Halorubrum pleomorphic virus 1 is enclosed in the envelope decorated with glycoprotein spikes. J Virol 84, 788–798. https://doi.org/10.1128/jvi.01347-09

Pinar-Méndez, A., Galofré, B., Blanch, A.R., García-Aljaro, C., 2022. Culture and molecular methods as complementary tools for water quality management. Science of the Total Environment 848, 157789. https://doi.org/10.1016/j.scitotenv.2022.157789

Pinart, M., Dötsch, A., Schlicht, K., Laudes, M., Bouwman, J., Forslund, S.K., Pischon, T., Nimptsch, K., 2022. Gut microbiome composition in obese and non-obese persons: A systematic review and meta-analysis. Nutrients 14, 12. https://doi.org/10.3390/nu14010012

Platts-Mills, J.A., Babji, S., Bodhidatta, L., Gratz, J., Haque, R., Havt, A., McCormick, B.J.J., McGrath, M., Olortegui, M.P., Samie, A., Shakoor, S., Mondal, D., Lima, I.F.N., Hariraju, D., Rayamajhi, B.B., Qureshi, S., Kabir, F., Yori, P.P., Mufamadi, B., Amour, C., Carreon, J.D., Richard, S.A., Lang, D., Bessong, P., Mduma, E., Ahmed, T., Lima, A.A.A.M., Mason, C.J., Zaidi, A.K.M., Bhutta, Z.A., Kosek, M., Guerrant, R.L., Gottlieb, M., Miller, M., Kang, G., Houpt, E.R., Chavez, C.B., Trigoso, D.R., Flores, J.T., Vasquez, A.O., Pinedo, S.R., Acosta, A.M., Ahmed, I., Alam, D., Ali, A., Rasheed, M., Soofi, S., Turab, A., Yousafzai, A.K., Bose, A., Jennifer, M.S., John, S., Kaki, S., Koshy, B., Muliyil, J., Raghava, M.V., Ramachandran, A., Rose, A., Sharma, S.L., Thomas, R.J., Pan, W., Ambikapathi, R., Charu, V., Dabo, L., Doan, V., Graham, J., Hoest, C., Knobler, S., Mohale, A., Nayyar, G., Psaki, S., Rasmussen, Z., Seidman, J.C., Wang, V., Blank, R., Tountas, K.H., Swema, B.M., Yarrot, L., Nshama, R., Ahmed, A.M.S., Tofail, F., Hossain, I., Islam, M., Mahfuz, M., Chandyo, R.K., Shrestha, P.S., Shrestha, R., Ulak, M., Black, R., Caulfield, L., Checkley, W., Chen, P., Lee, G., Murray-Kolb, L.E., Schaefer, B., Pendergast, L., Abreu, C., Costa, H., Moura, A. Di, Filho, J.Q., Leite, Á., Lima, N., Maciel, B., Moraes, M., Mota, F., Oriá, R., Quetz, J., Soares, A., Patil, C.L., Mahopo, C., Mapula, A., Nesamvuni, C., Nyathi, E., Barrett, L., Petri, W.A., Scharf, R., Shrestha, B., Shrestha, S.K., Strand, T., Svensen, E., 2015. Pathogen-specific burdens of community diarrhoea in developing countries: A multisite birth cohort study (MAL-ED). Lancet Glob Health 3, e564–e575. https://doi.org/10.1016/S2214-109X(15)00151-5

Porter, N.T., Canales, P., Peterson, D.A., Martens, E.C., 2017. A subset of polysaccharide capsules in the human symbiont *Bacteroides thetaiotaomicron* promote increased competitive fitness in the mouse gut. Cell Host Microbe 22, 494. https://doi.org/10.1016/J.CHOM.2017.08.020

Porter, N.T., Hryckowian, A.J., Merrill, B.D., Fuentes, J.J., Gardner, J.O., Glowacki, R.W.P., Singh, S., Crawford, R.D., Snitkin, E.S., Sonnenburg, J.L., Martens, E.C., 2020. Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in *Bacteroides thetaiotaomicron*. Nat Microbiol 5, 1170–1181. https://doi.org/10.1038/s41564-020-0746-5

Porter, N.T., Luis, A.S., Martens, E.C., 2018. *Bacteroides thetaiotaomicron*. Trends Microbiol 26, 966–967. https://doi.org/10.1016/j.tim.2018.08.005

Puig, A., Araujo, R., Jofre, J., Frias-Lopez, J., 2001. Identification of cell wall proteins of *Bacteroides fragilis* to which bacteriophage B40-8 binds specifically. Microbiology (N Y) 147, 281–288.

Puig, A., Queralt, N., Jofre, J., Araujo, R., 1999. Diversity of *Bacteroides fragilis* strains in their capacity to recover phages from human and animal wastes and from fecally polluted wastewater. Appl Environ Microbiol 65, 1772–1776.

Pumbwe, L., Skilbeck, C.A., Wexler, H.M., 2006. The *Bacteroides fragilis* cell envelope: quarterback, linebacker, coach-or all three? Anaerobe 12, 211–220. https://doi.org/10.1016/j.anaerobe.2006.09.004

Qiao, Z., Ye, Y., Szczuka, A., Harrison, K.R., Dodd, M.C., Wigginton, K.R., 2022. Reactivity of viral nucleic acids with chlorine and the impact of virus encapsidation. Environ Sci Technol 56(1), 218–227. https://doi.org/10.1021/acs.est.1c04239

Quan, Y., Zhang, K.X., Zhang, H.Y., 2023. The gut microbiota links disease to human genome evolution. Trends in Genetics 39, 451–461. https://doi.org/10.1016/j.tig.2023.02.006

Queralt, N., Jofre, J., Araujo, R., Muniesa, M., 2003. Homogeneity of the morphological groups of bacteriophages infecting *Bacteroides fragilis* strain HSP40 and strain RYC2056. Curr Microbiol 46, 163–168. https://doi.org/10.1007/s00284-002-3813-7

Raghavan, V., Lowe, E.C., Townsend, G.E., Bolam, D.N., Groisman, E.A., 2014. Tuning transcription of nutrient utilization genes to catabolic rate promotes growth in a gut bacterium. Mol Microbiol 93, 1010–1025. https://doi.org/10.1111/mmi.12714

Ramírez-Castillo, F.Y., Loera-Muro, A., Jacques, M., Garneau, P., Avelar-González, F.J., Harel, J., Guerrero-Barrera, A.L., 2015. Waterborne pathogens: Detection methods and challenges. Pathogens 4, 307–334. https://doi.org/10.3390/pathogens4020307

Ramos-Barbero, M.D., Gómez-Gómez, C., Sala-Comorera, L., Rodríguez-Rubio, L., Morales-Cortes, S., Mendoza-Barberá, E., Vique, G., Toribio-Avedillo, D., Blanch, A.R., Ballesté, E., Garcia-Aljaro, C., Muniesa, M., 2023. Characterization of crAss-like phage isolates highlights Crassvirales genetic heterogeneity and worldwide distribution. Nat Commun 14, 4295. https://doi.org/10.1038/s41467-023-40098-z

Remesh, A.T., Viswanathan, R., 2024. CrAss-like phages: from discovery in human fecal metagenome to application as a microbial source tracking marker. Food Environ Virol 16, 121–135. https://doi.org/10.1007/s12560-024-09584-5

Renz, H., Brandtzaeg, P., Hornef, M., 2012. The impact of perinatal immune development on mucosal homeostasis and chronic inflammation. Nat Rev Immunol 12, 9–23. https://doi.org/10.1038/nri3112

República de Colombia, 2014. Decreto nº1287. Criterios para el uso de biosólidos generados en plantas de tratamientos de aguas residuales municipales. Bogotá.

Reyes, A., Semenkovich, N.P., Whiteson, K., Rohwer, F., Gordon, J.I., 2012. Going viral: Next-generation sequencing applied to phage populations in the human gut. Nat Rev Microbiol 10, 607–617. https://doi.org/10.1038/nrmicro2853

Reyes, A., Wu, M., McNulty, N.P., Rohwer, F.L., Gordon, J.I., 2013. Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. Proc Natl Acad Sci U S A 110, 20236–20241. https://doi.org/10.1073/pnas.1319470110

Risse, J., Thomson, M., Patrick, S., Blakely, G., Koutsovoulos, G., Blaxter, M., Watson, M., 2015. A single chromosome assembly of *Bacteroides fragilis* strain BE1 from Illumina and MinION nanopore sequencing data. Gigascience 4, s13742-015. https://doi.org/10.1186/s13742-015-0101-6

Rochelle-Newall, E., Nguyen, T.M.H., Le, T.P.Q., Sengtaheuanghoung, O., Ribolzi, O., 2015. A short review of fecal indicator bacteria in tropical aquatic ecosystems: Knowledge gaps and future directions. Front Microbiol 6, 308. https://doi.org/10.3389/fmicb.2015.00308

Rodó, X., Pascual, M., Doblas-Reyes, F.J., Gershunov, A., Stone, D.A., Giorgi, F., Hudson, P.J., Kinter, J., Rodríguez-Arias, M.À., Stenseth, N.C., Alonso, D., García-Serrano, J., Dobson, A.P., 2013. Climate change and infectious diseases: Can we meet the needs for better prediction? Clim Change 118, 625–640. https://doi.org/10.1007/s10584-013-0744-1

Rodriguez-R, L.M., Konstantinidis, K.T., 2014. Nonpareil: A redundancy-based approach to assess the level of coverage in metagenomic datasets. Bioinformatics 30, 629–635. https://doi.org/10.1093/bioinformatics/btt584

Rodríguez-Rubio, L., Blanco-Picazo, P., Muniesa, M., 2020. Are phages parasites or symbionts of bacteria?, in: Biocommunication of Phages. Springer International Publishing, pp. 143–162. https://doi.org/10.1007/978-3-030-45885-0

Rodriguez-Valera, F., Martín-Cuadrado, A.-B., Rodriguez-Brito, B., Pasic, L., Thingstad, T.F., Rohwer, F., Mira, A., 2009. Explaining microbial population genomics through phage predation. Nature Precedings 1–1. https://doi.org/10.1038/npre.2009.3489.1

Rossi, A., Treu, L., Toppo, S., Zschach, H., Campanaro, S., Dutilh, B.E., 2020. Evolutionary study of the crAssphage virus at gene level. Viruses 12, 1035. https://doi.org/10.3390/v12091035

Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A., Mazmanian, S.K., 2011. The toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. Science (1979) 332, 974–977. https://doi.org/10.1126/science.1206095

Roux, S., Hallam, S.J., Woyke, T., Sullivan, M.B., 2015. Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. Elife 4, e08490. https://doi.org/10.7554/eLife.08490.001

Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., Tran, B.Q., Barry, N.A., Zheng, H., Peterson, S.B., Chou, S., Gonen, T., Goodlett, D.R., Goodman, A.L., Mougous, J.D., 2014. A type VI secretion-related pathway in *Bacteroidetes* mediates interbacterial antagonism. Cell Host Microbe 16, 227–236. https://doi.org/10.1016/j.chom.2014.07.007

Saba, J., Flores, K., Marshall, B., Engstrom, M.D., Peng, Y., Garje, A.S., Comstock, L.E., Landick, R., 2024. *Bacteroides* expand the functional versatility of a conserved transcription factor and transcribed DNA to program capsule diversity. Nature Communications 15, 10862. https://doi.org/10.1038/s41467-024-55215-9

Sabar, M.A., Honda, R., Haramoto, E., 2022. CrAssphage as an indicator of human-fecal contamination in water environment and virus reduction in wastewater treatment. Water Res 221, 118827. https://doi.org/10.1016/j.watres.2022.118827

Sala-Comorera, L., Reynolds, L.J., Martin, N.A., Pascual-Benito, M., Stephens, J.H., Nolan, T.M., Gitto, A., O'Hare, G.M.P., O'Sullivan, J.J., García-Aljaro, C., Meijer, W.G., 2021. crAssphage as a human molecular marker to evaluate temporal and spatial variability in faecal contamination of urban marine bathing waters. Science of the Total Environment 789, 147828. https://doi.org/10.1016/j.scitotenv.2021.147828

Salyers, A., 1984. Bacteroides of the human lower intestinal tract. Annu Rev Microbiol 38, 293–313.

Samson, J.E., Magadán, A.H., Sabri, M., Moineau, S., 2013. Revenge of the phages: Defeating bacterial defences. Nat Rev Microbiol 11, 675–687. https://doi.org/10.1038/nrmicro3096

Sangkaew, W., Kongprajug, A., Chyerochana, N., Ahmed, W., Rattanakul, S., Denpetkul, T., Mongkolsuk, S., Sirikanchana, K., 2021. Performance of viral and bacterial genetic markers for sewage pollution tracking in tropical Thailand. Water Res 190, 116706. https://doi.org/10.1016/j.watres.2020.116706

Sausset, R., Petit, M.A., Gaboriau-Routhiau, V., De Paepe, M., 2020. New insights into intestinal phages. Mucosal Immunol 13, 205–215. https://doi.org/10.1038/s41385-019-0250-5

Säwström, C., Lisle, J., Anesio, A.M., Priscu, J.C., Laybourn-Parry, J., 2008. Bacteriophage in polar inland waters. Extremophiles 12, 167–175. https://doi.org/10.1007/s00792-007-0134-6

Scanlan, P.D., 2017. Bacteria–bacteriophage coevolution in the human gut: implications for microbial diversity and functionality. Trends Microbiol 25, 614–623. https://doi.org/10.1016/j.tim.2017.02.012

Scarpellini, E., Ianiro, G., Attili, F., Bassanelli, C., De Santis, A., Gasbarrini, A., 2015. The human gut microbiota and virome: Potential therapeutic implications. Digestive and Liver Disease 47, 1007–1012. https://doi.org/10.1016/j.dld.2015.07.008

Schaper, M., Jofre, J., Uys, M., Grabow, W., 2002. Distribution of genotypes of F-specific RNA bacteriophages in human and non-human sources of faecal pollution in South Africa and Spain. J Appl Microbiol 92, 657–667.

Setubal, J.C., 2021. Metagenome-assembled genomes: concepts, analogies, and challenges. Biophys Rev 13, 905–909. https://doi.org/10.1007/s12551-021-00865-y

Shang, C., Cheung, L.M., Liu, W., 2007. MS2 coliphage inactivation with UV irradiation and free chlorine/monochloramine. Environ Eng Sci 24, 1321–1332. https://doi.org/10.1089/ees.2006.0261

Shang, J., Jiang, J., Sun, Y., 2021. Bacteriophage classification for assembled contigs using graph convolutional network. Bioinformatics 37, i25–i33. https://doi.org/10.1093/bioinformatics/btab293

Shareefdeen, H., Hynes, A.P., 2021. Does over a century of aerobic phage work provide a solid framework for the study of phages in the gut? Anaerobe 68, 102319. https://doi.org/10.1016/j.anaerobe.2021.102319

Sharma, S., Chatterjee, S., Datta, S., Prasad, R., Dubey, D., Prasad, R.K., Vairale, M.G., 2017. Bacteriophages and its applications: an overview. Folia Microbiol 62, 17–55. https://doi.org/10.1007/s12223-016-0471-x

Sharma, S., Sachdeva, P., Virdi, J.S., 2003. Emerging water-borne pathogens. Appl Microbiol Biotechnol 61, 424–428. https://doi.org/10.1007/s00253-003-1302-y

Sharp, D.G., Hook, A.E., Taylor, A.R., Beard, D., Beard, J.W., 1946. Sedimentation characters and pH stability of the T2 bacteriophage of *Escherichia coli*. Journal of Biological Chemistry 165, 259–270. https://doi.org/10.1016/S0021-9258(17)41228-2

Shin, J.H., Tillotson, G., MacKenzie, T.N., Warren, C.A., Wexler, H.M., Goldstein, E.J.C., 2024. *Bacteroides* and related species: The keystone taxa of the human gut microbiota. Anaerobe 85, 102819. https://doi.org/10.1016/j.anaerobe.2024.102819

Shkoporov, A.N., Clooney, A.G., Sutton, T.D.S., Ryan, F.J., Daly, K.M., Nolan, J.A., McDonnell, S.A., Khokhlova, E. V., Draper, L.A., Forde, A., Guerin, E., Velayudhan, V., Ross, R.P., Hill, C., 2019. The Human gut virome is highly diverse, stable, and individual specific. Cell Host Microbe 26, 527-541.e5. https://doi.org/10.1016/j.chom.2019.09.009

Shkoporov, A.N., Hill, Colin., 2019. Bacteriophages of the human gut: the “known unknown” of the microbiome. Cell Host Microbe 25, 195–209. https://doi.org/10.1016/j.chom.2019.01.017

Shkoporov, A.N., Khokhlova, E. V., Fitzgerald, C.B., Stockdale, S.R., Draper, L.A., Ross, R.P., Hill, C., 2018. ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides intestinalis*. Nat Commun 9, 4781. https://doi.org/10.1038/s41467-018-07225-7

Shkoporov, A.N., Khokhlova, E. V., Stephens, N., Hueston, C., Seymour, S., Hryckowian, A.J., Scholz, D., Ross, R.P., Hill, C., 2021. Long-term persistence of crAss-like phage crAss001 is associated with phase variation in *Bacteroides intestinalis*. BMC Biol 19, 1–16. https://doi.org/10.1186/s12915-021-01084-3

Shkoporov, A.N., Turkington, C.J., Hill, C., 2022. Mutualistic interplay between bacteriophages and bacteria in the human gut. Nat Rev Microbiol 20, 737–749. https://doi.org/10.1038/s41579-022-00755-4

Shoemaker, N.B., Vlamakis, H., Hayes, K., Salyers, A.A., 2001. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. Appl Environ Microbiol 67, 561–568. https://doi.org/10.1128/AEM.67.2.561-568.2001

Shuval, H.I., 1993. Investigation of typhoid fever and cholera transmission by raw wastewater irrigation in Santiago, Chile. Waf. Sci. Tech 27, 167–174.

Shuwen, H., Kefeng, D., 2022. Intestinal phages interact with bacteria and are involved in human diseases. Gut Microbes 14, 2113717. https://doi.org/10.1080/19490976.2022.2113717

Silveira, C.B., Rohwer, F.L., 2016. Piggyback-the-winner in host-associated microbial communities. NPJ Biofilms Microbiomes 2, 1–5. https://doi.org/10.1038/npjbiofilms.2016.10

Simmonds, P., Adams, M.J., Benk, M., Breitbart, M., Brister, J.R., Carstens, E.B., Davison, A.J., Delwart, E., Gorbalenya, A.E., Harrach, B., Hull, R., King, A.M.Q., Koonin, E. V., Krupovic, M., Kuhn, J.H., Lefkowitz, E.J., Nibert, M.L., Orton, R., Roossinck, M.J., Sabanadzovic, S., Sullivan, M.B., Suttle, C.A., Tesh, R.B., Van Der Vlugt, R.A., Varsani, A., Zerbini, F.M., 2017. Consensus statement: Virus taxonomy in the age of metagenomics. Nat Rev Microbiol 15, 161–168. https://doi.org/10.1038/nrmicro.2016.177

Sims, N., Kasprzyk-Hordern, B., 2020. Future perspectives of wastewater-based epidemiology: monitoring infectious disease spread and resistance to the community level. Environ Int 139, 105689. https://doi.org/10.1016/j.envint.2020.105689

Sinigalliano, C.D., Fleisher, J.M., Gidley, M.L., Solo-Gabriele, H.M., Shibata, T., Plano, L.R.W., Elmir, S.M., Wanless, D., Bartkowiak, J., Boiteau, R., Withum, K., Abdelzaher, A.M., He, G., Ortega, C., Zhu, X., Wright, M.E., Kish, J., Hollenbeck, J., Scott, T., Backer, L.C., Fleming, L.E., 2010. Traditional and molecular analyses for fecal indicator bacteria in non-point source subtropical recreational marine waters. Water Res 44, 3763–3772. https://doi.org/10.1016/j.watres.2010.04.026

Sinton, L.W., Finlay, R.K., Lynch, P.A., 1999. Sunlight inactivation of fecal bacteriophages and bacteria in sewage-polluted seawater. Appl Environ Microbiol 65, 3605–3613. https://doi.org/10.1128/AEM.65.8.3605-3613.1999

Siranosian, B.A., Tamburini, F.B., Sherlock, G., Bhatt, A.S., 2020. Acquisition, transmission and strain diversity of human gut-colonizing crAss-like phages. Nat Commun 11, 280. https://doi.org/10.1038/s41467-019-14103-3

Sirikanchana, K., 2013. Center of advanced materials for the purification of water with systems. Department of Civil and Environmental Engineering. Personnel Communication .

Smith, L., Goldobina, E., Govi, B., Shkoporov, A.N., 2023. Bacteriophages of the order Crassvirales: what do we currently know about this keystone component of the human gut virome? Biomolecules 13, 584. https://doi.org/10.3390/biom13040584

Smug, B.J., Szczepaniak, K., Rocha, E.P.C., Dunin-Horkawicz, S., Mostowy, R.J., 2023. Ongoing shuffling of protein fragments diversifies core viral functions linked to interactions with bacterial hosts. Nat Commun 14, 7460. https://doi.org/10.1038/s41467-023-43236-9

Soller, J.A., Schoen, M.E., Bartrand, T., Ravenscroft, J.E., Ashbolt, N.J., 2010. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res 44, 4674–4691. https://doi.org/10.1016/j.watres.2010.06.049

Sommer, F., Bäckhed, F., 2013. The gut microbiota-masters of host development and physiology. Nat Rev Microbiol 11, 227–238. https://doi.org/10.1038/nrmicro2974

Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., Sonnenburg, J.L., 2010. Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations. Cell 141, 1241–1252. https://doi.org/10.1016/j.cell.2010.05.005

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., Gordon, J.I., 2005. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science (1979) 307, 1955–1959. https://doi.org/10.1126/science.1108397

Stachler, E., Akyon, B., De Carvalho, N.A., Ference, C., Bibby, K., 2018. Correlation of crAssphage qPCR markers with culturable and molecular indicators of human fecal pollution in an impacted urban watershed. Environ Sci Technol 52, 7505–7512. https://doi.org/10.1021/acs.est.8b00638

Stachler, E., Bibby, K., 2014. Metagenomic evaluation of the highly abundant human gut bacteriophage crAssphage for source tracking of human fecal pollution. Environ Sci Technol Lett 1, 405–409. https://doi.org/10.1021/ez500266s

Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., Shanks, O.C., 2017. Quantitative crAssphage PCR assays for human fecal pollution measurement. Environ Sci Technol 51, 9146–9154. https://doi.org/10.1021/acs.est.7b02703

Stecher, B., Hardt, W.D., 2011. Mechanisms controlling pathogen colonization of the gut. Curr Opin Microbiol 14, 82–91. https://doi.org/10.1016/j.mib.2010.10.003

Stewart, J.R., Boehm, A.B., Dubinsky, E.A., Fong, T.T., Goodwin, K.D., Griffith, J.F., Noble, R.T., Shanks, O.C., Vijayavel, K., Weisberg, S.B., 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. Water Res 47, 6829–6838. https://doi.org/10.1016/j.watres.2013.04.063

Sun, W., Jing, Z., Zhao, Z., Yin, R., Santoro, D., Mao, T., Lu, Z., 2023. Dose-response behavior of pathogens and surrogate microorganisms across the ultraviolet-C spectrum: inactivation efficiencies, action spectra, and mechanisms. Environ Sci Technol 57, 10891–10900. https://doi.org/10.1021/acs.est.3c00518

Suttle, C.A., 2005. Viruses in the sea. Nature 437, 356–361. https://doi.org/10.1038/nature04160

Sutton, T.D.S., Hill, C., 2019. Gut bacteriophage: current understanding and challenges. Front Endocrinol 10, 784. https://doi.org/10.3389/fendo.2019.00784

Tandukar, S., Sherchan, S.P., Haramoto, E., 2020. Applicability of crAssphage, pepper mild mottle virus, and tobacco mosaic virus as indicators of reduction of enteric viruses during wastewater treatment. Sci Rep 10, 3616. https://doi.org/10.1038/s41598-020-60547-9

Tartera, C., Bosch, A., Jofre, J., 1988. The inactivation of bacteriophages infecting *Bacteroides fragilis* by chlorine treatment and UV-irradiation. FEMS Microbiol Lett 56, 313–316. https://doi.org/10.1111/j.1574-6968.1988.tb03198.x

Tartera, C., Jofre, J., 1987. Bacteriophages active against *Bacteroides fragilis* in sewage-polluted waters. Appl Environ Microbiol 53, 1632–1637. https://doi.org/10.1128/aem.53.7.1632-1637.1987

Tartera, C., Lucena, F., Jofre, J., 1989. Human origin of *Bacteroides fragilis* bacteriophages present in the environment. Appl Environ Microbiol 55, 2696–2701. https://doi.org/10.1128/aem.55.10.2696-2701.1989

Teaf, C.M., Flores, D., Garber, M., Harwood, V.J., 2018. Toward forensic uses of microbial source tracking. Microbiol Spectr 6, 115–141. https://doi.org/10.1128/microbiolspec.emf-0014-2017

The European Parliament and the Council of the European Union Regulation (EU), 2020. Minimum requirements for water reuse, 2020/741. Off. J. Eur. Union L177/32-L177/55.

Tomofuji, Y., Kishikawa, T., Maeda, Y., Ogawa, K., Nii, T., Okuno, T., Oguro-Igashira, E., Kinoshita, M., Yamamoto, K., Sonehara, K., Yagita, M., Hosokawa, A., Motooka, D., Matsumoto, Y., Matsuoka, H., Yoshimura, M., Ohshima, S., Nakamura, S., Inohara, H., Mochizuki, H., Takeda, K., Kumanogoh, A., Okada, Y., 2022. Whole gut virome analysis of 476 Japanese revealed a link between phage and autoimmune disease. Ann Rheum Dis 81, 278–288. https://doi.org/10.1136/annrheumdis-2021-221267

Toribio-Avedillo, D., Blanch, A.R., Muniesa, M., Rodríguez-Rubio, L., 2021. Bacteriophages as fecal pollution indicators. Viruses 13, 1089. https://doi.org/10.3390/v13061089

Troeger, C., Blacker, B.F., Khalil, I.A., Rao, P.C., Cao, S., Zimsen, S.R., Albertson, S.B., Stanaway, J.D., Deshpande, A., Abebe, Z., Alvis-Guzman, N., Amare, A.T., Asgedom, S.W., Anteneh, Z.A., Antonio, C.A.T., Aremu, O., Asfaw, E.T., Atey, T.M., Atique, S., Avokpaho, E.F.G.A., Awasthi, A., Ayele, H.T., Barac, A., Barreto, M.L., Bassat, Q., Belay, S.A., Bensenor, I.M., Bhutta, Z.A., Bijani, A., Bizuneh, H., Castañeda-Orjuela, C.A., Dadi, A.F., Dandona, L., Dandona, R., Do, H.P., Dubey, M., Dubljanin, E., Edessa, D., Endries, A.Y., Eshrati, B., Farag, T., Feyissa, G.T., Foreman, K.J., Forouzanfar, M.H., Fullman, N., Gething, P.W., Gishu, M.D., Godwin, W.W., Gugnani, H.C., Gupta, R., Hailu, G.B., Hassen, H.Y., Hibstu, D.T., Ilesanmi, O.S., Jonas, J.B., Kahsay, A., Kang, G., Kasaeian, A., Khader, Y.S., Khan, E.A., Khan, M.A., Khang, Y.H., Kissoon, N., Kochhar, S., Kotloff, K.L., Koyanagi, A., Kumar, G.A., Magdy Abd El Razek, H., Malekzadeh, R., Malta, D.C., Mehata, S., Mendoza, W., Mengistu, D.T., Menota, B.G., Mezgebe, H.B., Mlashu, F.W., Murthy, S., Naik, G.A., Nguyen, C.T., Nguyen, T.H., Ningrum, D.N.A., Ogbo, F.A., Olagunju, A.T., Paudel, D., Platts-Mills, J.A., Qorbani, M., Rafay, A., Rai, R.K., Rana, S.M., Ranabhat, C.L., Rasella, D., Ray, S.E., Reis, C., Renzaho, A.M., Rezai, M.S., Ruhago, G.M., Safiri, S., Salomon, J.A., Sanabria, J.R., Sartorius, B., Sawhney, M., Sepanlou, S.G., Shigematsu, M., Sisay, M., Somayaji, R., Sreeramareddy, C.T., Sykes, B.L., Taffere, G.R., Topor-Madry, R., Tran, B.X., Tuem, K.B., Ukwaja, K.N., Vollset, S.E., Walson, J.L., Weaver, M.R., Weldegwergs, K.G., Werdecker, A., Workicho, A.,

Yenesew, M., Yirsaw, B.D., Yonemoto, N., El Sayed Zaki, M., Vos, T., Lim, S.S., Naghavi, M., Murray, C.J., Mokdad, A.H., Hay, S.I., Reiner, R.C., 2018. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis 18, 1211–1228. https://doi.org/10.1016/S1473-3099(18)30362-1

Trosvik, P., De Muinck, E.J., Stenseth, N.C., 2015. Biotic interactions and temporal dynamics of the human gastrointestinal microbiota. ISME Journal 9, 533–541. https://doi.org/10.1038/ismej.2014.147

Turner, D., Shkoporov, A.N., Lood, C., Millard, A.D., Dutilh, B.E., Alfenas-Zerbini, P., van Zyl, L.J., Aziz, R.K., Oksanen, H.M., Poranen, M.M., Kropinski, A.M., Barylski, J., Brister, J.R., Chanisvili, N., Edwards, R.A., Enault, F., Gillis, A., Knezevic, P., Krupovic, M., Kurtböke, I., Kushkina, A., Lavigne, R., Lehman, S., Lobocka, M., Moraru, C., Moreno Switt, A., Morozova, V., Nakavuma, J., Reyes Muñoz, A., Rūmnieks, J., Sarkar, B., Sullivan, M.B., Uchiyama, J., Wittmann, J., Yigang, T., Adriaenssens, E.M., 2023. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. Arch Virol 168, 74. https://doi.org/10.1007/s00705-022-05694-2

UNESCO, 2017. Wastewater: the untapped resource. The United Nations World Water Development Report 2017.

U.S. EPA, 2016. In: 2016 Coliphage experts workshop: discussion topics and findings. Office of Water. EPA 823-F-16F001.

van Houte, S., Buckling, A., Westra, E.R., 2016. Evolutionary ecology of prokaryotic immune mechanisms. Microbiology and Molecular Biology Reviews 80, 745–763. https://doi.org/10.1128/mmbr.00011-16

Weidhaas, J.L., Macbeth, T.W., Olsen, R.L., Sadowsky, M.J., Norat, D., Harwood, V.J., 2010. Identification of a *Brevibacterium* marker gene specific to poultry litter and development of a quantitative PCR assay. J Appl Microbiol 109, 334–347. https://doi.org/10.1111/j.1365-2672.2010.04666.x

Weitz, J.S., Wilhelm, S.W., 2012. Ocean viruses and their effects on microbial communities and biogeochemical cycles. F1000 Biol Rep 4, 17. https://doi.org/10.3410/B4-17

Wen, X., Chen, F., Lin, Y., Zhu, H., Yuan, F., Kuang, D., Jia, Z., Yuan, Z., 2020. Microbial indicators and their use for monitoring drinkingwater quality-A review. Sustainability (Switzerland) 12, 2249. https://doi.org/10.3390/su12062249

Western Australian Government, 2012. Western Australian guidelines for biosolids management. Perth: Western Australian Government. Department of Environment and Conservation.

Wexler, Goodman, A.L., 2017. An insider's perspective: *Bacteroides* as a window into the microbiome. Nat Microbiol 2, 1–11. https://doi.org/10.1038/nmicrobiol.2017.26

Wexler, H.M., 2007. *Bacteroides:* The good, the bad, and the nitty-gritty. Clin Microbiol Rev 20, 593–621. https://doi.org/10.1128/CMR.00008-07

Wicki, M., Auckenthaler, A., Felleisen, R., Tanner, M., Baumgartner, A., 2011. Novel *Bacteroides* host strains for detection of human- and animal-specific bacteriophages in water. J Water Health 9, 159–168. https://doi.org/10.2166/wh.2010.165

Wiggins, B.A., 1996. Discriminant analysis of antibiotic resistance patterns in fecal Streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters. Appl Environ Microbiol 62, 3997–4002. https://doi.org/10.1128/aem.62.11.3997-4002.1996

Wigginton, K.R., Pecson, B.M., Sigstam, T., Bosshard, F., Kohn, T., 2012. Virus inactivation mechanisms: Impact of disinfectants on virus function and structural integrity. Environ Sci Technol 46, 12069–12078. https://doi.org/10.1021/es3029473

Williamson, K.E., Fuhrmann, J.J., Wommack, K.E., Radosevich, M., 2017. Viruses in soil ecosystems: an unknown quantity within an unexplored territory. Annu Rev Virol 4, 201–219. https://doi.org/10.1146/annurev-virology

Williamson, K.E., Radosevich, M., Wommack, K.E., 2005. Abundance and diversity of viruses in six Delaware soils. Appl Environ Microbiol 71, 3119–3125. https://doi.org/10.1128/AEM.71.6.3119-3125.2005

Wilson, G.W., Edgell, D.R., 2009. Phage T4 mobE promotes *trans* homing of the defunct homing endonuclease I-TevIII. Nucleic Acids Res 37, 7110–7123. https://doi.org/10.1093/nar/gkp769

Winfield, M.D., Groisman, E.A., 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli.* Appl Environ Microbiol 69, 3687–3694. https://doi.org/10.1128/AEM.69.7.3687-3694.2003

Winter, C., Bouvier, T., Weinbauer, M.G., Thingstad, T.F., 2010. Trade-offs between competition and defense specialists among unicellular planktonic organisms: the “Killing the Winner” hypothesis revisited. Microbiology and Molecular Biology Reviews 74, 42–57. https://doi.org/10.1128/mmbr.00034-09

World Health Organization, 2019. Water, sanitation, hygiene and health: a primer for health professionals. Geneve.

World Health Organization, 2017. Guidelines for drinking-water quality. WHO Chron 38, 104–8.

World Health Organization, 2003. Guidelines for safe recreational water environments: Coastal and fresh waters (Vol. 1). World Health Organization.

Wu, H., Juel, M.A.I., Eytcheson, S., Aw, T.G., Munir, M., Molina, M., 2023. Temporal and spatial relationships of CrAssphage and enteric viral and bacterial pathogens in wastewater in North Carolina. Water Res 239, 120008. https://doi.org/10.1016/j.watres.2023.120008

Wu, M., McNulty, N.P., Rodionov, D.A., Khoroshkin, M.S., Griffin, N.W., Cheng, J., Latreille, P., Kerstetter, R.A., Terrapon, N., Henrissat, B., Osterman, A.L., Gordon, J.I., 2015. Genetic determinants of *in vivo* fitness and diet responsiveness in multiple human gut *Bacteroides*. Science (1979) 350. https://doi.org/10.1126/science.aac5992

Wu, Z., Greaves, J., Arp, L., Stone, D., Bibby, K., 2020. Comparative fate of crAssphage with culturable and molecular fecal pollution indicators during activated sludge wastewater treatment. Environ Int 136, 105452. https://doi.org/10.1016/j.envint.2019.105452

Xihui, Z., Luqi, T., Zhiwei, W., Feng, X., Wei, Z., 2024. A comparative analysis of phage classification methods in light of 2 the recent ICTV taxonomic revisions. Available at SSRN 4673864. https://doi.org/10.1016/j.virol.2024.110016

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L. V, Gordon, J.I., 2003. A genomic view of the human: *Bacteroides thetaiotaomicron* symbiosis. New Series 299, 2074–2076.

Yadav, P., Singh, R.P., Gupta, R.K., Pradhan, T., Raj, A., Singh, S.K., Kaushalendra, Pandey, K.D., Kumar, A., 2023. Contamination of soil and food chain through wastewater application, in: Advances in Chemical Pollution, Environmental Management and Protection. Elsevier , pp. 109–132. https://doi.org/10.1016/bs.apmp.2022.11.001

Yan, W., Hall, A.B., Jiang, X., 2022. *Bacteroidales* species in the human gut are a reservoir of antibiotic resistance genes regulated by invertible promoters. NPJ Biofilms Microbiomes 8, 1. https://doi.org/10.1038/s41522-021-00260-1

Yang, H., Gan, D., Li, Y., Wang, X., Jin, L., Qin, K., Cui, C., Wu, J., Wang, Z., 2020. Quyushengxin formula causes differences in bacterial and phage composition in ulcerative colitis patients. Evidence-based Complementary and Alternative Medicine 2020, 5859023. https://doi.org/10.1155/2020/5859023

Yuan, Y., DeMott, M.S., Byrne, S.R., Flores, K., Poyet, M., Groussin, M., Conservancy, G.M., Berdy, B., Comstock, L., Alm, E.J., Dedon, P.C., 2024. Phosphorothioate DNA modification by BREX Type 4 systems in the human gut microbiome. bioRxiv. https://doi.org/10.1101/2024.06.03.597175

Yutin, N., Benler, S., Shmakov, S.A., Wolf, Y.I., Tolstoy, I., Rayko, M., Antipov, D., Pevzner, P.A., Koonin, E. V., 2021. Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. Nat Commun 12, 1044. https://doi.org/10.1038/s41467-021-21350-w

Yutin, N., Makarova, K.S., Gussow, A.B., Krupovic, M., Segall, A., Edwards, R.A., Koonin, E. V., 2018. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. Nat Microbiol 3, 38–46. https://doi.org/10.1038/s41564-017-0053-y

Zaheer, R., Lakin, S.M., Polo, R.O., Cook, S.R., Larney, F.J., Morley, P.S., Booker, C.W., Hannon, S.J., Van Domselaar, G., Read, R.R., McAllister, T.A., 2019. Comparative diversity of microbiomes and

resistomes in beef feedlots, downstream environments and urban sewage influent. BMC Microbiol 19, 1–17. https://doi.org/10.1186/s12866-019-1548-x

Zhang, Y., Qu, S., Xu, L., 2019. Progress in the study of virus detection methods: The possibility of alternative methods to validate virus inactivation. Biotechnol Bioeng 116, 2095–2102. https://doi.org/10.1002/bit.27003

Zhang, Y., Wu, R., Li, W., Chen, Z., Li, K., 2021. Occurrence and distributions of human-associated markers in an impacted urban watershed. Environmental Pollution 275, 116654. https://doi.org/10.1016/j.envpol.2021.116654

Zhang, Yingying, Zhuang, Y., Geng, J., Ren, H., Zhang, Yan, Ding, L., Xu, K., 2015. Inactivation of antibiotic resistance genes in municipal wastewater effluent by chlorination and sequential UV/chlorination disinfection. Science of the Total Environment 512–513, 125–132. https://doi.org/10.1016/j.scitotenv.2015.01.028

Zhong, Q., Carratala, A., Nazarov, S., Guerrero-Ferreira, R.C., Piccinini, L., Bachmann, V., Leiman, P.G., Kohn, T., 2016. Genetic, structural, and phenotypic properties of MS2 coliphage with resistance to ClO2 disinfection. Environ Sci Technol 50 (24), 13520–13528. https://doi.org/10.1021/acs.est.6b04170

Zhu, Y., Shang, J., Peng, C., Sun, Y., 2022. Phage family classification under Caudoviricetes: A review of current tools using the latest ICTV classification framework. Front Microbiol 13, 1032186.

Zinke, M., Schröder, G.F., Lange, A., 2022. Major tail proteins of bacteriophages of the order Caudovirales. Journal of Biological Chemistry 298 (1). https://doi.org/10.1016/j.jbc.2021.101472

8. ANEXOS

8.1. Anexo 1: Otras publicaciones de la doctoranda no incluidas en la tesis

A continuación, se enumeran cronológicamente otras publicaciones en las que la doctoranda ha participado, pero que no forman parte del contenido de esta tesis:

- Larrañaga, O., Brown-Jaque, M., Quiros, P., **Gómez-Gómez, C.,** Blanch, A. R., Rodríguez-Rubio, L., & Muniesa, M. (2018). Phage particles harboring antibiotic resistance genes in fresh-cut vegetables and agricultural soil. *Environment international*, *115*, 133-141. https://doi.org/10.1016/j.envint.2018.03.019
  Factor de impacto: 7,943; Q1 (*Environmental Sciences*).

- **Gómez-Gómez, C.,** Blanco-Picazo, P., Brown-Jaque, M., Quirós, P., Rodríguez-Rubio, L., Cerdà-Cuellar, M., & Muniesa, M. (2019). Infectious phage particles packaging antibiotic resistance genes found in meat products and chicken feces. *Scientific reports*, *9*(1), 13281. https://doi.org/10.1038/s41598-019-49898-0
  Factor de impacto: 3,998; Q1 (*Multidisciplinary Sciences*).

- **Gómez-Gómez, C**., Méndez, J., & Muniesa, M. (2020). Efficacy of daily hygiene routine using cleansing wipes in the reduction of eyelid bacterial load. *Research Square.* https://doi.org/10.21203/rs.3.rs-33635/v1

- Blanco-Picazo, P., Roscales, G., Toribio-Avedillo, D., **Gómez-Gómez, C**., Avila, C., Ballesté, E., Muniesa, M., & Rodríguez-Rubio, L. (2020). Antibiotic resistance genes in phage particles from Antarctic and Mediterranean seawater ecosystems. *Microorganisms*, *8*(9), 1293. https://doi.org/10.3390/microorganisms8091293
  Factor de impacto: 4,128 (2020); Q2 (*Microbiology*).

- Blanco-Picazo, P., **Gómez-Gómez, C**., Morales-Cortes, S., Muniesa, M., & Rodríguez-Rubio, L. (2022). Antibiotic resistance in the viral fraction of dairy products and a nut-based milk. *International Journal of Food Microbiology*, *367*, 109590. https://doi.org/10.1016/j.ijfoodmicro.2022.109590
  Factor de impacto: 5,911 (2022); Q1 (*Food Science & Technology*).

- Blanco-Picazo, P., **Gómez-Gómez, C.,** Aguiló-Castillo, S., Fernández-Orth, D., Cerdà-Cuéllar, M., Muniesa, M., & Rodríguez-Rubio, L. (2022). Chicken liver is a potential reservoir of bacteriophages and phage-derived particles containing antibiotic resistance genes. *Microbial biotechnology*, *15*(9), 2464-2475. https://doi.org/10.1111/1751-7915.14056
  Factor de impacto: 6,575 (2022); Q1 (*Biotechnology & Applied Microbiology*).

- Blanco-Picazo, P., **Gómez-Gómez, C.**, Tormo, M., Ramos-Barbero, M. D., Rodríguez-Rubio, L., & Muniesa, M. (2022). Prevalence of bacterial genes in the phage fraction of food viromes. *Food Research International*, *156*, 111342. https://doi.org/10.1016/j.foodres.2022.111342
  Factor de impacto: 8,1 (2022); Q1 (*Food Science & Technology*).

- Quirós, P., Sala-Comorera, L., **Gómez-Gómez, C.**, Ramos-Barbero, M. D., Rodríguez-Rubio, L., Vique, G., Yance-Chávez, T., Atarés, S., García-Gutierrez, S., García-Marco, S., Vallejo, A., Salaet, I., & Muniesa, M. (2023). Identification of a virulent phage infecting species of *Nitrosomonas*. *The ISME Journal*, *17*(5), 645-648. https://doi.org/10.1038/s41396-023-01380-6
  Factor de impacto: 10,8 (2023); Q1 (*Microbiology, Ecology, Evolution, Behavior and Systematics*).

- Toribio-Avedillo, D., **Gómez-Gómez, C.**, Sala-Comorera, L., Rodríguez-Rubio, L., Carcereny, A., García-Pedemonte, D., Pintó, R. M., Guix, S., Galofré, B., Bosch, A., Merino, S., & Muniesa, M. (2023). Monitoring influenza and respiratory syncytial virus in wastewater. Beyond COVID-19. *Science of the Total Environment*, *892*, 164495. https://doi.org/10.1016/j.scitotenv.2023.164495
  Factor de impacto: 8,2 (2023); Q1 (*Environmental Sciences*).

- Toribio-Avedillo, D., **Gómez-Gómez, C.**, Sala-Comorera, L., Galofré, B., & Muniesa, M. (2024). Adapted methods for monitoring influenza virus and respiratory syncytial virus in sludge and wastewater. *Science of The Total Environment*, *918*, 170636. https://doi.org/10.1016/j.scitotenv.2024.170636
  Factor de impacto: 8 (2024); Q1 (*Environmental Sciences*).

- Morales-Cortés, S., Sala-Comorera, L., **Gómez-Gómez, C.**, Muniesa, M., & García-Aljaro, C. (2024). CrAss-like phages are suitable indicators of antibiotic resistance genes found in abundance in fecally polluted samples. *Environmental Pollution*, *359*, 124713. https://doi.org/10.1016/j.envpol.2024.124713
  Factor de impacto: 7,6 (2024); Q1 (*Environmental Sciences*).

Environment International 115 (2018) 133–141

Contents lists available at ScienceDirect

Environment International

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

Phage particles harboring antibiotic resistance genes in fresh-cut vegetables and agricultural soil

Olatz Larrañaga, Maryury Brown-Jaque, Pablo Quirós, Clara Gómez-Gómez, Anicet R. Blanch, Lorena Rodríguez-Rubio, Maite Muniesa*

Department of Genetics, Microbiology and Statistics, University of Barcelona, Diagonal 643, Annex, Floor 0, E-08028 Barcelona, Spain

ARTICLE INFO

Keywords:
Bacteriophages
Antibiotic resistance genes
Cropland
Lettuce
Cucumber
Spinach

ABSTRACT

Bacteriophages are ubiquitously distributed prokaryotic viruses that are more abundant than bacteria. As a consequence of their life cycle, phages can kidnap part of their host's genetic material, including antibiotic resistance genes (ARGs), which released phage particles transfer in a process called transduction. The spread of ARGs among pathogenic bacteria currently constitutes a serious global health problem. In this study, fresh vegetables (lettuce, spinach and cucumber), and cropland soil were screened by qPCR for ten ARGs (bla_{TEM}, $bla_{CTX\text{-}M\text{-}1}$ group, $bla_{CTX\text{-}M\text{-}9}$ group, $bla_{OXA\text{-}48}$, bla_{VIM}, *mecA*, *sul1*, *qnrA*, *qnrS* and *armA*) in their viral DNA fraction. The presence of ARGs in the phage DNA was analyzed before and after propagation experiments in an *Escherichia coli* host strain to evaluate the ability of the phage particles to infect a host. ARGs were found in the phage DNA fraction of all matrices, although with heterogeneous values. ARG prevalence was significantly higher in lettuce and soil, and the most common overall were β-lactamases. After propagation experiments, an increase in ARG densities in phage particles was observed in samples of all four matrices, confirming that part of the isolated phage particles were infectious. This study reveals the abundance of free, replicative ARG-containing phage particles in vegetable matrices and cropland soil. The particles are proposed as vehicles for resistance transfer in these environments, where they can persist for a long time, with the possibility of generating new resistant bacterial strains. Ingestion of these mobile genetic elements may also favor the emergence of new resistances, a risk not previously considered.

1. Introduction

The consumption of vegetables is accepted as part of a healthy lifestyle, especially when freshly cut. Vegetables cultivated in open fields are susceptible to contamination by microorganisms (Berger et al., 2010), mainly via soil, organic fertilizers, and irrigation water (Semenov and Franz, 2008; Tyler and Triplett, 2008). Moreover, as these foods are often consumed raw, microorganisms can persist on their surfaces despite good washing practices. Although not considered as contaminants, bacteriophages, or more precisely phage particles, have recently been explored as elements involved in horizontal gene transfer among bacterial species. The persistence of phage particles in different environmental conditions, greater than that of their bacterial hosts (Calero-Cáceres and Muniesa, 2016; Lucena et al., 2004; Skraber et al., 2009), allows them to act as environmental reservoirs of the genes they mobilize.

Among genes of interest for human health, those encoding antibiotic resistances are of increasing concern because of their ability to generate new multiresistant strains, rendering antibiotic treatments ineffective (World Health Organization, 2011). Resistance to antibiotics is as old as the antimicrobials themselves, which have existed in bacterial ecosystems for billions of years (von Wintersdorff et al., 2016), forming part of the competitive and antagonistic relations among microorganisms. However, over the last decades, resistances have become an increasingly serious global problem, arising from their overuse, globalization and industrialization (Ventola, 2015). Excessive application of antibiotics in areas such as medicine, livestock and even aquaculture has led to the excretion of antibiotics virtually unchanged into the environment, where they exert selective pressures on bacterial communities. Several species of environmental bacteria are known to be responsible for the transfer of resistance to pathogenic bacteria, as is the case of the genes $bla_{CTX\text{-}M}$ with species of *Kluyvera* (Humeniuk et al., 2002), *qnrA* with *Shewanella algae* (Poirel et al., 2005), *qnrS* with *Vibrio splendicus* (Cattoir et al., 2007), and $bla_{OXA\text{-}48}$ in other species of *Shewanella* (Cummings et al., 2011), etc. Wastewater from populated areas and treatment plants and the sludges they produce are considered “hot

* Corresponding author.
E-mail address: mmuniesa@ub.edu (M. Muniesa).

https://doi.org/10.1016/j.envint.2018.03.019
Received 9 January 2018; Received in revised form 2 March 2018; Accepted 13 March 2018
Available online 20 March 2018
0160-4120/ © 2018 Elsevier Ltd. All rights reserved.

SCIENTIFIC REPORTS
nature research

OPEN

Infectious phage particles packaging antibiotic resistance genes found in meat products and chicken feces

Received: 13 May 2019
Accepted: 29 August 2019
Published online: 16 September 2019

Clara Gómez-Gómez[1], Pedro Blanco-Picazo[1], Maryury Brown-Jaque[1], Pablo Quirós[1], Lorena Rodríguez-Rubio[1], Marta Cerdà-Cuellar[2] & Maite Muniesa[1]

Bacteriophages can package part of their host's genetic material, including antibiotic resistance genes (ARGs), contributing to a rapid dissemination of resistances among bacteria. Phage particles containing ARGs were evaluated in meat, pork, beef and chicken minced meat, and ham and mortadella, purchased in local retailer. Ten ARGs (*bla*$_{TEM}$, *bla*$_{CTX-M-1}$, *bla*$_{CTX-M-9}$, *bla*$_{OXA-48}$, *bla*$_{VIM}$, *qnrA*, *qnrS*, *mecA*, *armA* and *sul1*) were analyzed by qPCR in the phage DNA fraction. The genes were quantified, before and after propagation experiments in *Escherichia coli*, to evaluate the ability of ARG-carrying phage particles to infect and propagate in a bacterial host. According to microbiological parameters, all samples were acceptable for consumption. ARGs were detected in most of the samples after particle propagation indicating that at least part of the isolated phage particles were infectious, being *sul1*the most abundant ARG in all the matrices followed by β-lactamase genes. ARGs were also found in the phage DNA fraction of thirty-seven archive chicken cecal samples, confirming chicken fecal microbiota as an important ARG reservoir and the plausible origin of the particles found in meat. Phages are vehicles for gene transmission in meat that should not be underestimated as a risk factor in the global crisis of antibiotic resistance.

Antibiotic resistance is a major threat to public health and food security. A decreasing production of new antimicrobial compounds and a reduced effectiveness of the existing ones caused by the emergence of resistant bacteria is resulting in higher mortality rates[1]. The uncontrolled use of antibiotics and the ability of bacteria to mutate or acquire external genes is leading to the emergence of resistant strains. Epidemiological studies confirm that antibiotic consumption is directly correlated with the emergence and dissemination of resistance[2].

Natural antibiotics have existed for billions of years in bacterial ecosystems as a strategy for adaptation and defense against other bacteria[3], resulting in a simultaneous evolution of antibiotic resistance genes (ARGs)[4]. The resistance originated in environmental reservoirs and transmitted to human pathogens[5] has been exacerbated by exposure to antibiotics present in clinical and agricultural settings[6,7]. Indeed, most of the antibiotics used in medicine, livestock and aquaculture are excreted practically unaltered to the environment, generating a pressure that selects for naturally occurring resistant strains and their accompanying ARG[8]. Bacteria can be intrinsically resistant to antibiotics or can become resistant by mutation or by acquisition of ARGs by horizontal gene transfer[9,10]. While conjugation and transformation are well-studied mechanisms, transduction, which is understood as gene transfer mediated by a bacterial virus or bacteriophage/phage, is attracting increasing attention and has perhaps been underestimated[11,12].

Three mechanisms of transduction have been described so far; specialized transduction[13], mediated by temperate phages that lysogenize the bacterial cell and mobilize the genes adjacent to their insertion site. Generalized transduction[14] mediated by virulent phages that package bacterial DNA instead of phage DNA. And the most recent mechanism described, lateral transduction mediated by temperate phages that do not excise from the bacterial chromosome after induction and generate capsids able to package only bacterial DNA located downstream

[1]Department of Genetics, Microbiology and Statistics, University of Barcelona, Diagonal 643, Annex, Floor 0, E-08028, Barcelona, Spain. [2]IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autonoma de Barcelona, 08193, Bellaterra, Barcelona, Spain. Correspondence and requests for materials should be addressed to M.M. (email: mmuniesa@ub.edu)

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Efficacy of daily hygiene routine using cleansing wipes in the reduction of eyelid bacterial load.

Clara Gómez-Gómez
Universitat de Barcelona
Javier Méndez
Universitat de Barcelona
Maite Muniesa (mmuniesa@ub.edu)
Universitat de Barcelona

Research article

Keywords: Eyelid hygiene, Cleansing wipes, Eyelid microbiota, Aloe barbadensis, Endophthalmitis, Blepharitis

Posted Date: June 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-33635/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

microorganisms

MDPI

Article

Antibiotic Resistance Genes in Phage Particles from Antarctic and Mediterranean Seawater Ecosystems

Pedro Blanco-Picazo [1], Gabriel Roscales [1], Daniel Toribio-Avedillo [1], Clara Gómez-Gómez [1], Conxita Avila [2], Elisenda Ballesté [1], Maite Muniesa [1,*] and Lorena Rodríguez-Rubio [1,*]

[1] Department of Genetics, Microbiology and Statistics, University of Barcelona, Diagonal 643, Annex, Floor 0, 08028 Barcelona, Spain; p.blanco@ub.edu (P.B.-P.); gascheb@gmail.com (G.R.); d.toribio.avedillo@ub.edu (D.T.-A.); claragg1995@gmail.com (C.G.-G.); eballeste@ub.edu (E.B.)

[2] Department of Evolutionary Biology, Ecology and Environmental Sciences (BEECA), Faculty of Biology, and Biodiversity Research Institute (IrBIO), University of Barcelona, Diagonal 643, 08028 Barcelona, Spain; conxita.avila@ub.edu

* Correspondence: mmuniesa@ub.edu (M.M.); lorenarodriguez@ub.edu (L.R.-R.)

Received: 29 July 2020; Accepted: 19 August 2020; Published: 24 August 2020



Abstract: Anthropogenic activities are a key factor in the development of antibiotic resistance in bacteria, a growing problem worldwide. Nevertheless, antibiotics and resistances were being generated by bacterial communities long before their discovery by humankind, and might occur in areas without human influence. Bacteriophages are known to play a relevant role in the dissemination of antibiotic resistance genes (ARGs) in aquatic environments. In this study, five ARGs (bla_{TEM}, $bla_{CTX\text{-}M\text{-}1}$, $bla_{CTX\text{-}M\text{-}9}$, *sul1* and *tetW*) were monitored in phage particles isolated from seawater of two different locations: (i) the Mediterranean coast, subjected to high anthropogenic pressure, and (ii) the Antarctic coast, where the anthropogenic impact is low. Although found in lower quantities, ARG-containing phage particles were more prevalent among the Antarctic than the Mediterranean seawater samples and Antarctic bacterial communities were confirmed as their source. In the Mediterranean area, ARG-containing phages from anthropogenic fecal pollution might allow ARG transmission through the food chain. ARGs were detected in phage particles isolated from fish (Mediterranean, Atlantic, farmed, and frozen), the most abundant being β-lactamases. Some of these particles were infectious in cultures of the fecal bacteria *Escherichia coli*. By serving as ARG reservoirs in marine environments, including those with low human activity, such as the Antarctic, phages could contribute to ARG transmission between bacterial communities.

Keywords: bacteriophages; shellfish; fish; transduction; ARG; horizontal gene transfer

1. Introduction

Antibiotic resistance is one of the biggest threats to public health this century. The growing environmental presence of antibiotics and their derivative metabolites is mainly due to anthropogenic activities, such as medicine, agriculture, and industry. Antibiotics have been overused both in clinical settings and the livestock industry, including for prophylaxis, although their use as growth supplements has been prohibited in Europe since 2006 (EC Regulation No. 1831/2003) [1–3]. Epidemiological studies have verified that antibiotic consumption is directly connected with the emergence and dissemination of resistances [2].

Depending on their transport and persistence in soil or water, antibiotics can spread to surface water and groundwater, exerting a selective pressure on aquatic bacteria and the microbiota of aquatic animals, and leading to the emergence of resistant bacteria and the dissemination of antibiotic resistance genes (ARGs) [4,5].

An update to this article is included at the end

International Journal of Food Microbiology 367 (2022) 109590

Contents lists available at ScienceDirect

International Journal of Food Microbiology

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

Antibiotic resistance in the viral fraction of dairy products and a nut-based milk

Pedro Blanco-Picazo, Clara Gómez-Gómez, Sara Morales-Cortes, Maite Muniesa, Lorena Rodríguez-Rubio *

Department of Genetics, Microbiology and Statistics, University of Barcelona, Diagonal 643, Annex, Floor 0, 08028 Barcelona, Spain

ARTICLE INFO

Keywords:
Bacteriophage
Transduction
ARG
Horchata

ABSTRACT

Phages, the most abundant biological entities in the biosphere, can carry different bacterial genes, including those conferring antibiotic resistance. In this study, dairy products were analyzed by qPCR for the presence of phages and phage particles containing antibiotic resistance genes (ARGs). Eleven ARGs were identified in 50 samples of kefir, yogurt, milk, fresh cheese and nut-based milk (horchata), purchased from local retailers in Barcelona. Propagation experiments showed that at least some of the phages isolated from these samples infected *Escherichia coli* WG5, which was selected as the host strain because it does not contain prophages or ARGs in its genome. Electron microscopy revealed that the phage particles showed morphologies compatible with the *Myoviridae* and *Siphoviridae* families. Our results show that dairy products contain ARGs within infectious phage particles and may therefore serve as a reservoir of ARGs that can be mobilized to susceptible hosts, both in the food matrix and in the intestinal tract after ingestion.

1. Introduction

Antibiotic resistance is recognized as a worldwide threat to human health. The overuse of antibiotics in human and veterinary medicine has led to the emergence of antibiotic resistances and epidemiological studies verify that antibiotic consumption is directly connected with this emergence and the dissemination of resistances (More, 2020; Ventola, 2019).

In livestock, although their application as growth supplements has been banned in Europe since 2006 (The European Parliament and the Council of the European Union, 2003), antibiotics are still widely used for prophylaxis and/or treatment of infectious diseases (Cabello, 2006; Kemper, 2008). The continuous exposure of animals to antibiotics promotes the development of antimicrobial resistance. Under persistent selective pressure, commensal bacteria in livestock become reservoirs of antibiotic resistance genes (ARGs), which can be transferred to pathogenic bacteria. Resistant bacteria can be transmitted through the consumption of meat and other livestock products (e.g., milk and eggs) and be disseminated in the environment via animal wastes (Gomes and Henriques, 2016; Manyi-Loh et al., 2018). Additionally, antibiotic residues have been detected in animal tissues and milk (Landers et al., 2012; Sachi et al., 2019).

Agents of antimicrobial resistance, including ARGs, have been reported in different genera of lactic acid bacteria (LAB) isolated from dairy products, where they are the most abundant microorganisms (Erginkaya et al., 2018; Wang et al., 2019). Although most LAB strains are not pathogenic, the ARGs they harbor may be disseminated to other microorganisms. In addition, antibiotic resistance has been detected in viromes isolated from the air environment of dairy processing plants as well as in agricultural slurry (Colombo et al., 2018; Cook et al., 2021), which therefore represent additional reservoirs of antimicrobial resistance.

Antibiotics are naturally produced by microorganisms as a protective mechanism against competitors (Checcucci et al., 2020). While resistance to antibiotics is intrinsic to bacteria, it may also be acquired by mutation or horizontal gene transfer (Von Wintersdorff et al., 2016). The most studied mechanisms of DNA mobilization are conjugation and transformation, but transduction events or gene transfer mediated by phages have been recently revealed as another important route (Balcázar, 2020; Lekunberri et al., 2017).

Bacteriophages are the most abundant and ubiquitous biological entities in the world, outnumbering bacteria by an estimated ten-fold (Suttle, 2007). They are present in every ecosystem where bacteria exist (Pujato et al., 2019) and have a long persistence in the

* Corresponding author.
E-mail address: lorenarodriguez@ub.edu (L. Rodríguez-Rubio).

https://doi.org/10.1016/j.ijfoodmicro.2022.109590
Received 17 December 2021; Received in revised form 15 February 2022; Accepted 16 February 2022
Available online 20 February 2022
0168-1605/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Check for updates

microbial biotechnology

Open Access

Chicken liver is a potential reservoir of bacteriophages and phage-derived particles containing antibiotic resistance genes

Pedro Blanco-Picazo,[1] Clara Gómez-Gómez,[1] Sergi Aguiló-Castillo,[2] Dietmar Fernández-Orth,[3] Marta Cerdà-Cuéllar,[4] Maite Muniesa[1] and Lorena Rodríguez-Rubio[1]*

[1]*Department de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Diagonal 643, Planta 0, Barcelona, 08028, Spain.*

[2]*Spanish National Bioinformatics Institute (INB)/ELIXIR-ES, Barcelona Supercomputing Center, Barcelona, Spain.*

[3]*Department of Bioinformatics and Molecular Biology, Cerba Internacional Pl. Ramon Llull, 7-10, Sabadell, Barcelona, 08203, Spain.*

[4]*IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, Bellaterra, Barcelona, 08193, Spain.*

Summary

Poultry meat production is one of the most important agri-food industries in the world. The selective pressure exerted by widespread prophylactic or therapeutic use of antibiotics in intensive chicken farming favours the development of drug resistance in bacterial populations. Chicken liver, closely connected with the intestinal tract, has been directly involved in food-borne infections and found to be contaminated with pathogenic bacteria, including *Campylobacter* and *Salmonella*. In this study, 74 chicken livers, divided into sterile and non-sterile groups, were analysed, not only for microbial indicators but also for the presence of phages and phage particles containing antibiotic resistance genes (ARGs). Both bacteria and phages were detected in liver tissues, including those dissected under sterile conditions. The phages were able to infect *Escherichia coli* and showed a Siphovirus morphology. The chicken livers contained from 10^3 to 10^6 phage particles per g, which carried a range of ARGs (*bla*$_{TEM}$, *bla*$_{CTx\text{-}M\text{-}1}$, *sul1*, *qnrA*, *armA* and *tetW*) detected by qPCR. The presence of phages in chicken liver, mostly infecting *E. coli*, was confirmed by metagenomic analysis, although this technique was not sufficiently sensitive to identify ARGs. In addition, ARG-carrying phages were detected in chicken faeces by qPCR in a previous study of the group. Comparison of the viromes of faeces and liver showed a strong coincidence of species, which suggests that the phages found in the liver originate in faeces. These findings suggests that phages, like bacteria, can translocate from the gut to the liver, which may therefore constitute a potential reservoir of antibiotic resistance genes.

Introduction

The European Union (EU) is one of the world's largest producers of poultry meat and a net exporter of poultry products, with an annual production of around 13.4 million tonnes (Poultry | European Commission). Poultry is the only meat category whose production has expanded during the Covid-19 pandemic and the only one expected to grow between 2020 and 2030 in the EU (European Commission). The consumption of poultry meat is also projected to increase on a global scale (OECD-FAO AGRICULTURAL OUTLOOK 2018–2027), as it is regarded as a cheap, healthy and sustainable product by consumers.

Intensive poultry production, one of the fastest growing industries in the world, strongly depends on antimicrobials to prevent or treat diseases. In several countries, including Spain, the poultry production sectors are committed to a responsible use of antibiotics and are gradually reducing their application (Programa REDUCE Pollos Broiler | PRAN). Nevertheless, in other parts of the world, the usage of antibiotics as prophylactics, or even as growth promoters, a forbidden practice in the

Received 9 December, 2021; accepted 15 March, 2022.
For correspondence. *E-mail lorenarodriguez@ub.edu; Tel. +34934034625; Fax +34934039047.
Microbial Biotechnology (2022) **15**(9), 2464–2475
doi:10.1111/1751-7915.14056
Funding information
This work was supported by the Spanish Ministerio de Innovación y Ciencia (AGL2016-75536-P and PID2020-113355GB-I00), the Agencia Estatal de Investigación (AEI) and the European regional fund (ERF). The study was partially supported by the Union Iberoamericana de Universidades (UIU-OHOW) and the Generalitat de Catalunya (2017SGR170). CERCA Programme from the Generalitat de Catalunya is also acknowledged. P. B-P has a grant from the Spanish Ministerio de Economía, Industria y Competitividad (BES-2017-081296). L. R.-R is a lecturer of the Serra-Húnter program, Generalitat de Catalunya.

© 2022 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Food Research International 156 (2022) 111342

Contents lists available at ScienceDirect

Food Research International

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

Prevalence of bacterial genes in the phage fraction of food viromes

Pedro Blanco-Picazo [a], Clara Gómez-Gómez [a], Marc Tormo [b,c], Maria Dolores Ramos-Barbero [a], Lorena Rodríguez-Rubio [a,*], Maite Muniesa [a,*]

[a] Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Diagonal 643, Annex, Floor 0, E-08028 Barcelona, Spain
[b] Genomics Core Facility, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, 08003 Barcelona, Spain
[c] Scientific IT Core Facility, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, 08003 Barcelona, Spain

ARTICLE INFO

Keywords:
Antibiotic resistance gene
Transduction
Bacteriophages
Viromes
Food safety

ABSTRACT

Antibiotic resistance genes (ARGs) have been identified in viral DNA isolated from different kinds of food, but little is known about their origin. In this study, twenty-one viromes were analyzed from samples of food previously reported to carry ARGs, including meat (poultry, veal, and pork), fish (Mediterranean, Atlantic, frozen, farmed and shellfish) and vegetables (lettuce, cucumber, and spinach). Classification of the contigs by Kraken revealed a large percentage of unclassified contigs (43.7–98.2%) in all the viromes. Only 0.05–7.1% of the contigs were identified as viral and of these, more than 91% belonged to different bacteriophage families, Podophages and Siphophages being the most prevalent. According to VirSorter, the largest number of viral contigs were derived from viromes of shellfish, followed by spinach. Spinach viromes also included the largest number of phage sequences identified by PHASTER. The abundant presence of bacterial genes in the viromes, including *16S rRNA* genes, was attributed to the phage packaging of the bacterial genome fragments, as no bacterial DNA was found outside the viral capsids. The detection of *16S rRNA* genes in the different viromes allowed diverse phage bacterial hosts to be identified. The three major functional groups of genes determined were related to metabolism, detoxification/resistance, and above all, biosynthesis. Various ARGs were quantified in the viromes by qPCR, the most prevalent being β-lactamases, particularly bla_{TEM}. Analysis of ARG diversity in the viromes by Prokka and CARD revealed various resistance-related genes, whereas a more restrictive search by ResFinder identified bla_{TEM} in all the food viromes, bla_{OXA} in Atlantic fish-1 and spinach-2, *oqxB* in lettuce-1, and *dfr* in spinach-2. The presence of ARGs in the food viromes points to bacterial DNA mobilization by transduction mechanisms. Transduction of resistances by phage particles may therefore contribute to the emergence of resistant strains along the food chain and should be monitored.

1. Introduction

The study of viromes is an expanding field, and the development of more reliable shotgun sequencing techniques has provided new insights into the role of viruses in microbial communities (Norman et al., 2015; Reyes et al., 2010; Roux et al., 2016). Microbiome sequencing mainly generates data about the bacterial fraction, as most of the total DNA extracted from a microbiome is bacterial. Although progress in virome research has fed the genetic databases with new sequences, viral genomes remain under-represented (Chatterjee & Duerkop, 2018; Gregory et al., 2020; Pérez-Cataluña et al., 2021; Wang, 2020).

When analyzing a virome, purification of viral particles and removal of external non-viral DNA is essential to enrich viral sequences. Although bioinformatic analysis can help to remove non-viral sequences, the large number of unidentified viral contigs (also known as viral dark matter) implies that only a small percentage of viral sequences can be identified, particularly in complex samples, and most of the viral sequences correspond to bacteriophages (Cantalupo et al., 2011; Pérez-Cataluña et al., 2021).

Phages can mobilize a single bacterial gene inserted in the phage genome (specialized transduction) or a large fragment of bacterial DNA encapsidated during phage assembly (generalized or lateral transduction) (Chiang et al., 2019). When analyzing the extent to which transduction events in a virome are driven by bacteriophages,

* Corresponding authors.
E-mail addresses: lorenarodriguez@ub.edu (L. Rodríguez-Rubio), mmuniesa@ub.edu (M. Muniesa).

https://doi.org/10.1016/j.foodres.2022.111342
Received 24 February 2022; Received in revised form 15 April 2022; Accepted 3 May 2022
Available online 7 May 2022
0963-9969/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

ISME

www.nature.com/ismej

ARTICLE

Identification of a virulent phage infecting species of *Nitrosomonas*

Pablo Quirós[1,2], Laura Sala-Comorera[1], Clara Gómez-Gómez[1], María Dolores Ramos-Barbero[1,3], Lorena Rodríguez-Rubio[1], Gloria Vique[1], Tula Yance-Chávez[2], Sergio Atarés[2], Sandra García-Gutierrez[4,5], Sonia García-Marco[4,5], Antonio Vallejo[4,5], Ignasi Salaet[2] and Maite Muniesa[1]✉

© The Author(s), under exclusive licence to International Society for Microbial Ecology 2023

In the first and limiting step of nitrification, ammonia (NH_3) is oxidised to nitrite (NO_2^-) by the action of some prokaryotes, including bacteria of the *Nitrosomonas* genus. A potential approach to nitrification inhibition would be through the application of phages, but until now this method has been unexplored and no virulent phages that infect nitrifying bacteria have been described. In this study, we report the isolation of the first phage infecting some *Nitrosomonas* species. This polyvalent virulent phage (named ΦNF-1) infected *Nitrosomonas europaea, Nitrosomonas communis*, and *Nitrosomonas nitrosa*. Phage ΦNF-1 has the morphology of the *Podoviridae* family, a dsDNA genome of 41,596 bp and a 45.1 % GC content, with 50 predicted open reading frames. Phage ΦNF-1 was found to inhibit bacterial growth and reduce NH_4^+ consumption in the phage-treated cultures. The application of phages as biocontrol agents could be a useful strategy for nitrification inhibition without the restrictions associated with chemical inhibitors.

The ISME Journal (2023) 17:645–648; https://doi.org/10.1038/s41396-023-01380-6

INTRODUCTION

In the nitrogen (N) cycle, nitrification can be mediated by the activity of canonical ammonia-oxidizing bacteria (AOB) and archaea (AOA), which oxidize ammonia (NH_3) to nitrite (NO_2^-), and then nitrate (NO_3^-) is produced by nitrite-oxidizing bacteria (NOB). Additionally, some members of the genus *Nitrospira* perform complete NH_3 oxidation to NO_3^- (comammox) in a single step [1]. Among the AOB, *Nitrosomonas*, *Nitrosospira*, and *Nitrosococcus* are the most relevant genera [2, 3].

Nitrification occurs in soils, sediments, and aquatic environments. It plays an important role in wastewater treatment systems by contributing to excess N removal, and in agriculture, where it determines the availability of fertilizer N, required for a high plant productivity [4]. In some cases, the transformation of NH_3 by ammonia-oxidizers before uptake by plants renders N fertilization inefficient and encourages the addition of excess fertilizer. This has deleterious ecological effects through the volatilization of NH_3, production of N_2O [5], or N leaching to water bodies [4, 6].

To improve the efficiency of N fertilization, a novel approach would be the application of phages that act against nitrifying prokaryotes. Bacteriophages (phages) have been proposed as biocontrol tools because they infect and lyse bacterial cells and can be applied in different fields [7–11] with a minimum impact on the microbial ecology of each biome. This study presents the first description of a phage that infects representatives of the genus *Nitrosomonas* and has potential application to suppress bacterial nitrifying activity.

RESULTS AND DISCUSSION

Cultures of *Nitrosomonas europaea*, *Nitrosomonas communis*, and *Nitrosomonas nitrosa*, selected for their high abundance in wastewater [2], were grown at 28 °C. As the slow-growing *Nitrosomonas* do not generate sufficient bacterial mass to monitor their growth spectrophotometrically, pH variations of bacterial cultures and qPCR were used instead. After 25 days in the dark (the time required to reach exponential growth), the cultures had a pH of 8.0 and were inoculated (day 0 of infection) with the phages purified from six wastewater samples collected in 2019 from four urban wastewater treatment plants in Catalonia (NE Spain) (supplementary online material). At day 4 of infection, the pH was readjusted to 8 (Fig. S1-A) to avoid bacterial growth inhibition by pH reduction and the cultures were incubated until day 7.

Control cultures exhibited a pH decrease, possibly due to bacterial growth, while the pH of phage-infected cultures remained steady and close to 8.0, particularly in those cultures infected with phage suspensions #1, #2, #4, and #5, which was attributed to phage-induced lysis of bacteria (Fig. S1-A).

As AOB do not form confluent growth on agar plates and plaques of lysis could not be generated, phages from the four suspensions were used to re-infect new cultures in five successive rounds, selecting the phage most successful in infecting each host strain. Suspension #1 (from which phage ΦNF-1 was isolated) exhibited the highest infectivity in all host strains. While the pH of the uninfected control cultures decreased 1.8–2.0 units, the pH of

[1]Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona (UB), Diagonal 643, Annex, Floor 0, E-08028 Barcelona, Spain. [2]Departamento de I+D+i de Fertinagro Biotech S.L, Polígono Industrial La Paz, Teruel, Spain. [3]Departmento de Fisiología, Genética y Microbiología, Universidad de Alicante (UA), 03080 Alicante, Spain. [4]Departamento de Química y Tecnología de Alimentos, ETSI Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid (UPM), Madrid, Spain. [5]Centro de Estudios e Investigación para la Gestión de Riesgos Agrarios y Medioambientales (CEIGRAM), Universidad Politécnica de Madrid (UPM), Madrid, Spain. ✉email: mmuniesa@ub.edu

Received: 1 July 2022 Revised: 1 February 2023 Accepted: 3 February 2023
Published online: 9 February 2023

Downloaded from https://academic.oup.com/ismej/article/17/5/645/7468535 by UGI Biblioteca Lletres user on 16 April 2025

Science of the Total Environment 892 (2023) 164495

Contents lists available at ScienceDirect

Science of the Total Environment

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

Monitoring influenza and respiratory syncytial virus in wastewater. Beyond COVID-19

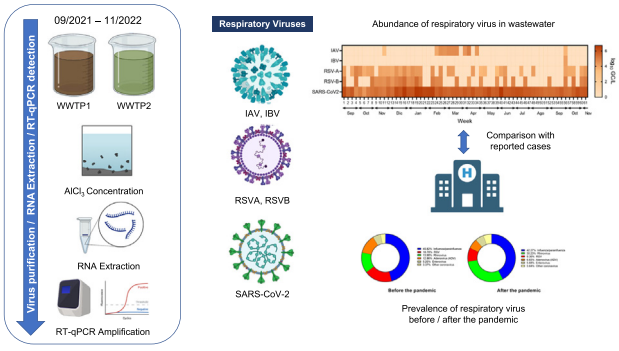
Daniel Toribio-Avedillo [a], Clara Gómez-Gómez [a], Laura Sala-Comorera [a], Lorena Rodríguez-Rubio [a], Albert Carcereny [b,c], David García-Pedemonte [b,c], Rosa Maria Pintó [b,c], Susana Guix [b,c], Belén Galofré [d], Albert Bosch [b,c], Susana Merino [a,c], Maite Muniesa [a,*]

[a] MARS Group (Health Related Water Microbiology Group), Department of Genetics, Microbiology and Statistics, Section of Microbiology, Virology and Biotechnology, School of Biology, University of Barcelona, Diagonal 643, E-08028 Barcelona, Spain
[b] Enteric Virus Laboratory, Department of Genetics, Microbiology and Statistics, Section of Microbiology, Virology and Biotechnology, School of Biology, University of Barcelona, Diagonal 643, E-08028 Barcelona, Spain
[c] Research Institute of Nutrition and Food Safety (INSA-UB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain
[d] Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, General Batet 1-7, Barcelona 08028, Spain

HIGHLIGHTS

- RSV and influenza virus in wastewater from Barcelona city show a seasonal distribution.
- SARS-CoV-2 was more prevalent in wastewater than RSV and RSV more than influenza.
- WBS is epidemiologically useful for monitoring virus circulation and outbreaks.
- WBS efficacy anticipating the peaks of infection of respiratory viruses may vary.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:
Influenza
Respiratory syncytial virus (RSV)
SARS-CoV-2
Wastewater-based epidemiology (WBE)
RT-qPCR
Surveillance

ABSTRACT

Wastewater-based surveillance can be a valuable tool to monitor viral circulation and serve as an early warning system. For respiratory viruses that share similar clinical symptoms, namely SARS-CoV-2, influenza, and respiratory syncytial virus (RSV), identification in wastewater may allow differentiation between seasonal outbreaks and COVID-19 peaks. In this study, to monitor these viruses as well as standard indicators of fecal contamination, a weekly sampling campaign was carried out for 15 months (from September 2021 to November 2022) in two wastewater treatment plants that serve the entire population of Barcelona (Spain). Samples were concentrated by the aluminum hydroxide adsorption-precipitation method and then analyzed by RNA extraction and RT-qPCR. All samples were positive for SARS-CoV-2, while the positivity rates for influenza virus and RSV were significantly lower (10.65 % for influenza A (IAV), 0.82 % for influenza B (IBV), 37.70 % for RSV-A and 34.43 % for RSV-B). Gene copy concentrations of SARS-CoV-2 were often approximately 1 to 2 logarithmic units higher compared to the other respiratory viruses. Clear peaks of IAV H3:N2 in February and March 2022 and RSV in winter 2021 were observed, which matched the chronological incidence of infections recorded in the Catalan Government clinical database. In conclusion, the data obtained from wastewater surveillance provided new information on the abundance of respiratory viruses in the Barcelona area and correlated favorably with clinical data.

* Corresponding author.
E-mail address: mmuniesa@ub.edu (M. Muniesa).

http://dx.doi.org/10.1016/j.scitotenv.2023.164495
Received 12 April 2023; Received in revised form 24 May 2023; Accepted 24 May 2023
Available online 26 May 2023
0048-9697/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Science of the Total Environment 918 (2024) 170636

Contents lists available at ScienceDirect

Science of the Total Environment

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

Adapted methods for monitoring influenza virus and respiratory syncytial virus in sludge and wastewater

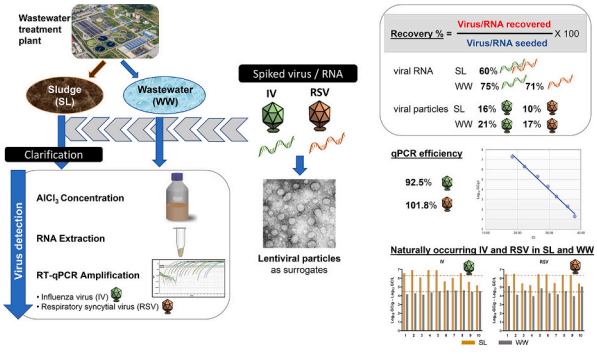
Daniel Toribio-Avedillo [a], Clara Gómez-Gómez [a], Laura Sala-Comorera [a], Belén Galofré [b], Maite Muniesa [a,*]

[a] Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Diagonal 643 Annex, Floor 0, E-08028 Barcelona, Spain
[b] Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, General Batet 1-7, Barcelona 08028, Spain

HIGHLIGHTS

- An efficient method targeting enveloped viruses in sludge is presented.
- Lentiviral particles are useful surrogates for IV and RSV monitoring.
- The protocol allows quantification of naturally occurring IV and RSV in sludge.
- Sludge is a significant reservoir of IV and RSV.
- Winter wastewater and sludge samples were positive for IV and RSV.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:
Influenza virus (IV)
Respiratory syncytial virus (RSV)
Sludge
Wastewater-based epidemiology (WBE)
RT-qPCR
Surveillance

ABSTRACT

Wastewater-based surveillance constitutes a valuable methodology for the continuous monitoring of viral circulation, with the capacity to function as an early warning system. It holds particular significance in scenarios where respiratory viruses exhibit overlapping clinical presentations, as occurs with SARS-CoV-2, influenza virus (IV), and respiratory syncytial virus (RSV), and allows seasonal virus outbreaks to be distinguished from COVID-19 peaks. Furthermore, sewage sludge, given it harbors concentrated human waste from a large population, serves as a substantial reservoir for pathogen detection.

To effectively integrate wastewater-based epidemiology into infectious disease surveillance, the detection methods employed in wastewater samples must be adapted to the distinct characteristics of sludge matrices. In this study, we adapted and applied protocols for the detection of IV and RSV in sewage sludge, comparing their performance with the results obtained in wastewater. To assess the efficiency of these protocols, sludge and wastewater samples were spiked with IV and RSV RNA, either free or incorporated in lentiviral particles. Samples were concentrated using the aluminum hydroxide adsorption-precipitation method before viral RNA extraction. Absolute virus quantification was carried out by RT-qPCR, including an internal control to monitor potential inhibitory factors. Recovery efficiencies for both free IV and RSV RNA were 60 % in sludge, and 75 % and 71 % respectively in wastewater, whereas the values for IV and RSV RNA in lentiviral particles were 16 % and 10 % in sludge and 21 % and 17 % in wastewater respectively. Additionally, the protocol enabled the quantification of

* Corresponding author.
E-mail address: mmuniesa@ub.edu (M. Muniesa).

https://doi.org/10.1016/j.scitotenv.2024.170636
Received 30 November 2023; Received in revised form 10 January 2024; Accepted 31 January 2024
Available online 6 February 2024
0048-9697/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

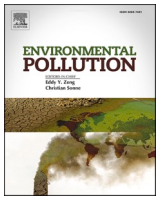
Environmental Pollution 359 (2024) 124713



Contents lists available at ScienceDirect

Environmental Pollution

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





CrAss-like phages are suitable indicators of antibiotic resistance genes found in abundance in fecally polluted samples☆

Sara Morales-Cortés, Laura Sala-Comorera, Clara Gómez-Gómez, Maite Muniesa, Cristina García-Aljaro *

Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Diagonal 643, Prevosti Building Floor 0, E-08028, Barcelona, Spain

ARTICLE INFO

Keywords:
Antibiotic resistance genes
crAssphage
crAssBcn phage
crAss-like phages
Bacteriophages
Indicators
Microbial source tracking
Emerging contaminants

ABSTRACT

Antibiotic resistance genes (ARGs) have been extensively observed in bacterial DNA, and more recently, in phage particles from various water sources and food items. The pivotal role played by ARG transmission in the proliferation of antibiotic resistance and emergence of new resistant strains calls for a thorough understanding of the underlying mechanisms. The aim of this study was to assess the suitability of the prototypical p-crAssphage, a proposed indicator of human fecal contamination, and the recently isolated crAssBcn phages, both belonging to the Crassvirales group, as potential indicators of ARGs. These crAss-like phages were evaluated alongside specific ARGs (bla_{TEM}, $bla_{CTX\text{-}M\text{-}1}$, $bla_{CTX\text{-}M\text{-}9}$, bla_{VIM}, $bla_{OXA\text{-}48}$, *qnrA*, *qnrS*, *tetW* and *sul1*) within the total DNA and phage DNA fractions in water and food samples containing different levels of fecal pollution. In samples with high fecal load (>10^3 CFU/g or ml of *E. coli* or somatic coliphages), such as wastewater and sludge, positive correlations were found between both types of crAss-like phages and ARGs in both DNA fractions. The strongest correlation was observed between *sul1* and crAssBcn phages (rho = 0.90) in sludge samples, followed by $bla_{CTX\text{-}M\text{-}9}$ and p-crAssphage (rho = 0.86) in sewage samples, both in the phage DNA fraction. The use of crAssphage and crAssBcn as indicators of ARGs, considered to be emerging environmental contaminants of anthropogenic origin, is supported by their close association with the human gut. Monitoring ARGs can help to mitigate their dissemination and prevent the emergence of new resistant bacterial strains, thus safeguarding public health.

1. Introduction

The growing resistance to antibiotics is a serious public health concern as it is undermining the effectiveness of treatments for bacterial diseases. Recognized as a global problem (Murray et al., 2022), antibiotic resistance is increasing the morbidity and mortality associated with bacterial infections as well as healthcare costs. The level of antimicrobial resistance, particularly multiple resistance in widely disseminated bacterial strains, has reached unprecedented levels, driven by the pressure of increased antibiotic usage (in both human and animal medicine), greater population movement, and industrialization (Prestinaci et al., 2015; Holmes et al., 2016).

Antibiotic resistance genes (ARGs) are abundant in environmental bacteria, and it has been proposed that many of them have an environmental origin. It is feasible that ARGs can transfer, via various mechanisms, from environmental strains to pathogenic bacteria found in clinical settings. Horizontal gene transfer constitutes the single most important mechanism that accelerates ARG dispersal, and the most studied mobile genetic elements involved are plasmids, transposons, and phages (Muniesa et al., 2013). Although ARGs have usually been detected in the bacterial DNA fraction, recent studies have focused on the potential role of bacteriophages as a gene transfer mechanism (Colavecchio et al., 2017; Gabashvili et al., 2020) and there is growing evidence for the occurrence of ARGs in the phage DNA fraction of environmental, food, and human samples (Anand et al., 2016; Subirats et al., 2016; Barrios et al., 2021; Calero-Cáceres and Balcázar, 2019; Colomer-Lluch et al., 2011; Sala-Comorera et al., 2021; Strange et al., 2021). Considering that bacterial cells and phages exhibit different levels of persistence in the environment (Allué-Guardia et al., 2014; Calero-Cáceres and Muniesa, 2016; Calero-Cáceres et al., 2017), it is

☆ This paper has been recommended for acceptance by Wen-Xiong Wang.

* Corresponding author.
E-mail addresses: smoralco10@alumnes.ub.edu (S. Morales-Cortés), laurasala@ub.edu (L. Sala-Comorera), claragomez@ub.edu (C. Gómez-Gómez), mmuniesa@ub.edu (M. Muniesa), crgarcia@ub.edu (C. García-Aljaro).

https://doi.org/10.1016/j.envpol.2024.124713
Received 27 May 2024; Received in revised form 23 July 2024; Accepted 9 August 2024
Available online 10 August 2024
0269-7491/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

8.2. Anexo 2: Medios de cultivo y tampones

Medios de cultivo

Medios utilizados para el cultivo de *Bacteroides* y la enumeración de bacteriófagos que infectan especies de *Bacteroides:*

BBE agar (*Bacteroides Bile Esculin Agar*):

- 20 g Bilis bovina
- 40 g TSA (*Tryptic Soy Agar*)
- 1 g Esculina
- 0,5 g Citrato amónico férrico
- 1 L H_2O destilada
- 2 ml Solución stock hemina
- 2,5 ml Gentamicina (40mg/ml)

Mezclar todos los componentes, a excepción de la hemina y la gentamicina, y esterilizar en el autoclave a 121°C durante 15 min. Una vez el medio se haya enfriado, añadir la solución de hemina y la gentamicina. Ajustar el pH a 7 en condiciones de esterilidad. Almacenar a 4°C.

BPRM (*Bacteroides Phage Recovery Medium*)

	BPRM agar	**BPRM caldo**	**BPRM ss (semisólido)**
H_2O destilada	1 L	100 ml	100 ml
BPRM	29,42 g	2,94 g	2,94 g
$CaCl_2$ (5%)	1 ml	100 µl	100 µl
Agar	15 g	-	0,75 g
	Disolver por agitación y esterilizar en el autoclave a 121°C durante 15 min. Una vez enfriado, añadir el resto de los componentes y ajustar el pH en condiciones de esterilidad.		
pH	6,3 – 7,3	6,3 – 7,3	6,3 – 7,3
Hemina	10 ml	1 ml	1 ml
Na_2CO_3 (1M)	25 ml	2,5 ml	2,5 ml
Kanamicina (stock 50 mg/ml)	2 ml	200 µl	200 µl
Ácido nalidíxico (stock 25 mg/ml)	4 ml	400 µl	400 µl
HCl (37%)	3,35 ml	335 µl	335 µl

ABB (*Anaerobe Basal Broth*)

	ABB agar	**ABB caldo**	**ABB ss (semisólido)**
H_2O destilada	1 L	100 ml	100 ml
ABB comercial	35,4 g	3,54 g	3,54 g
Agar	15 g	-	0,75 g

Disolver por agitación y esterilizar en el autoclave a 121°C durante 15 min.

Soluciones complementarias para el cultivo de *Bacteroides*

Solución stock de hemina

Para el medio BBE

- 0,05 g Hemina
- 10 ml NaOH 1M
- 90 ml H_2O bd

Para el medio BPRM

- 0,1 g Hemina
- 0,5 ml NaOH 1M
- 99,5 ml H_2O bd

Disolver durante 30-60 min todos los componentes utilizando un agitador magnético. Esterilizar por filtración mediante un filtro de tamaño de poro 0,22 µm de ésteres de celulosa de alta adsorción proteica. Guardar a temperatura ambiente no más de 6 meses.

Solución stock de $CaCl_2$ (5%)

- 5 g $CaCl_2 \cdot 2H_2O$
- 100 ml H_2O bd

Solución stock de Na_2Co_3 (1M)

- 10,6 g Na_2CO_3
- 100 ml H_2O bd

Crioprotector BSA + sacarosa:

- 10 g BSA
- 20 g Sacarosa
- 100 ml H_2O bd

Disolver por agitación con una mosca durante 1 h. Filtrar con filtros de 0,22 µm de ésteres de celulosa de alta adsorción proteica y utilizar inmediatamente. No almacenar.

Medios utilizados para el cultivo de *E. coli* WG5 y DH5α

	LB agar	**LB caldo**	**LB ss**
LB comercial	25 g	2,5 g	2,5 g
Agar	15 g	-	0,75 g
H_2O destilada	1 L	100 ml	100 ml

Esterilizar mediante autoclave a 121°C durante 15 min y guardar las botellas o las placas de agar a temperatura ambiente o 4°C.

Medios para la detección y mantenimiento de mutantes

LB agar + X-Gal/IPTG + ampicilina (100 µg/ml):

- 2,5 g LB comercial
- 80 µl X-Gal (stock de 50 mg/ml)
- 175 µl IPTG (stock 0,1M)
- 1,5 g Agar
- 100 µl Ampicilina (stock de 100 mg/ml)
- 100 ml H_2O destilada

LB agar + espectinomicina (50 µg/ml) (para el mantenimiento de la cepa *helper E. coli* HB101/pRK2073)

- 2,5 g LB comercial
- 1,5 g Agar
- 250 µl Espectinomicina (stock 20 mg/ml)
- 100 ml H_2O destilada

ABB agar + X-Gal/IPTG + eritromicina (300 µg/ml)

- 3,54 g ABB
- 80 µl X-Gal (stock de 50 mg/ml)
- 175 µl IPTG (stock 0,1M)
- 1,5 g Agar
- 600 µl eritromicina (stock de 50 mg/ml)
- 100 ml H_2O destilada

Antibióticos empleados en los estudios:

Antibiótico	Diluido en	Concentración stock (mg/ml)	Concentraciones finales usadas (µg/ml)
Ácido nalidíxico	H_2O bd	25	100
Kanamicina	H_2O bd	50	100
Ampicilina[a]	H_2O bd	100	100
Gentamicina[a]	H_2O bd	40	100
Eritromicina	Etanol	50	50 – 300[b]
Espectinomicina	H_2O bd	20	50

[a] Almacenar en la nevera durante máximo un mes. Mejor si se prepara nuevo cada vez que se deba emplear.

[b] La concentración de 50 µg/ml se emplea para cultivos en líquido y la de 300 µg/ml para cultivos en placas de agar.

Tampones

PBS (*Phosphate buffered saline*)

- 8 g NaCl
- 0,2 g KCl
- 0,2 g KH_2PO_4
- 1,15 g $Na_2HPO_4 \cdot 7\ H_2O$
- 1 L H_2O bd

Ajustar el pH a 7,1 – 7,2. Esterilizar mediante autoclave a 121°C durante 15 min y guardar a temperatura ambiente.

Tampón para fagos

- 3,5 g Na_2HPO_4
- 1,5 g KH_2PO_4
- 2,5 g NaCl
- 0,5 L de H_2O bd

Esterilizar mediante autoclave a 121°C durante 15 min. Una vez frío suplementar, de manera aséptica, cada 0,5 L de tampón con:

- 5 ml $MgSO_4$ (0,1M)
- 5 ml $CaCl_2$ (0,01M)

Guardar a temperatura ambiente.

$CaCl_2$ 0,01M

- 0,11 g $CaCl_2$
- 100 ml H_2O bd

Disolver por agitación. Esterilizar por filtración con filtros de ésteres de celulosa de alta adsorción proteica con un diámetro de poro de 0,22 µm. Almacenar a 4°C.

$MgSO_4$ 0,1M

- 2,47 g $MgSO_4 \cdot 7H_2O$
- 100 ml H_2O bd

Disolver por agitación. Esterilizar por filtración con filtros de ésteres de celulosa de alta adsorción proteica con un diámetro de poro de 0,22 µm. Almacenar a 4°C.

TBE (*Tris Boric EDTA Buffer*) 10X

- 109 g Tris base
- 55,6 g Ácido bórico
- 9,2 g EDTA
- 1 L H_2O bd

Esterilizar mediante autoclave a 121°C durante 15 min y guardar a temperatura ambiente.

TBE 1X

- 100 ml TBE 10X
- 900 ml H_2O bd

Guardar a temperatura ambiente

Gel de agarosa 1,5%

- 150 ml TBE 1X
- 2,25 g Agarosa

El porcentaje de agarosa variará dependiendo del tamaño del fragmento de ADN. Calentar la mezcla en el microondas y agitar varias veces durante el calentamiento evitando la ebullición para la correcta disolución de la agarosa.

ADNasa (10mg/ml)

- 10 mg de ADNasa (Roche®)
- 1 ml H_2O bd estéril

Disolver por agitación. Almacenar a -20°C.

Tampón proteinase K

- 2 ml Tris HCl 1M pH 8
- 2 ml EDTA 0,5 M pH 8
- 10 ml SDS 10%
- 100 ml de H_2O bd

Esterilizar mediante autoclave a 121°C durante 15 minutos y almacenar a temperatura ambiente. El tampón proteinasa K se usa junto con la Proteinasa K de Roche® (concentración final 20 mg/ml), almacenada a -20°C.

Etanol 70%

- 70 ml Etanol (100%)
- 30 ml H_2O bd estéril

Guardar a temperatura ambiente sin esterilizar.

Tampón de lavado (membranas de diálisis)

- 20 g $NaHCO_3$
- 0,37 g EDTA
- 1 L H_2O bd

Disolver por agitación y ajustar el pH a 8. Esterilizar mediante autoclave a 121°C durante 15 minutos. Se prepara antes de usar.

Tampón EDTA (membranas de diálisis)

- 0,37 g EDTA
- 1 L H_2O bd

Disolver por agitación y ajustar el pH a 8. Esterilizar mediante autoclave a 121°C durante 15 minutos. Se prepara antes de usar.

Tampón de diálisis

- 12,11 g Tris Base

- 0,074 g EDTA (ácido etilendiaminotetraacético) (Titriplex®)
- 1 L H_2O bd

Disolver por agitación y ajustar el pH a 8. Esterilizar mediante autoclave a 121°C durante 15 minutos. Se prepara antes de usar.

Molibdato de amonio 2%

- 0,2 mg Molibdato de amonio
- 10 ml H_2O bd estéril

Ajustar el pH a 7,0. Esterilizar por filtración mediante filtros de ésteres de celulosa de alta adsorción proteica con un diámetro de poro de 0,22 µm. Agitar y guardar a 4°C en oscuridad.

Glicerol (870 g/L)

- 69 ml Glicerol (1261 g/L)
- 31 ml H_2O bd

Esterilizar mediante autoclave a 121°C durante 15 min y guardar a temperatura ambiente en oscuridad.

Tampones utilizados en el protocolo de hibridación

Tampón de desnaturalización

- 20 g NaOH
- 87,66 g NaCl
- 1 L H_2O bd estéril

No esterilizar, guardar a temperatura ambiente.

Tampón de neutralización

- 87,66 g NaCl
- 500 ml Tris HCl (1M, pH 7,4)
- 1 L de H_2O bd

Esterilizar mediante autoclave a 121°C 15 min y guardar a temperatura ambiente.

Tampón 1 de hibridación:

- 100 ml Tris HCl (1M, pH 7,5)
- 100 ml NaCl (1,5M)
- 800 ml H_2O bd

No esterilizar, guardar a temperatura ambiente.

Tampón 2 de hibridación

- 5 g Agente bloqueante (Roche®)
- 1 L Tampón 1 de hibridación

Disolver por agitación magnética con calor, y guardar a -20°C.

Tampón 3 de hibridación

- 100 ml Tris HCl (1M, pH 9,5)
- 100 ml NaCl (1M)
- 50 ml $MgCl_2$ (1M)
- 750 ml H_2O bd

Preparar inmediatamente antes de su uso con soluciones previamente esterilizadas.

Tampón de anticuerpo anti-digoxigenina (DIG)

- 10 µl Anticuerpo anti-DIG (Roche®)
- 50 ml Tampón 2

Tampón de prehibridación

- 50 ml SSC ("*saline sodium citrate*") 20X
- 2 ml N- lauryl sarcosina al 10%
- 0,4 ml SDS al 10%
- 2 g Agente bloqueante (Roche®)
- 145 ml H_2O bd

Disolver por agitación magnética con calor, y guardar a -20°C.

Tampón de hibridación

- 50 ml Tampón de prehibridación
- 10 µl Sonda marcada con dUTP (Roche®)

Guardar a 4°C y se puede reutilizar varias veces.

Tampón SDS 10%

- 10 g SDS
- 100 ml H_2O bd

Esterilizar mediante autoclave a 121°C 15 min y guardar a temperatura ambiente.

N-lauryl sarcosina 10%

- 10 g N-lauryl sarcosina
- 100 ml H_2O bd

Esterilizar mediante autoclave a 121°C 15 min y guardar a temperatura ambiente.

Tampón SSC 20X

- 175,32 g NaCl
- 88,23 g Citrato sódico
- 1 L de H_2O bd

Esterilizar mediante autoclave a 121°C 15 min y guardar a temperatura ambiente.

2X SSC + 0,1% SDS

- 50 ml SSC 20X
- 5 ml SDS 10%
- 445 ml H_2O bd

No esterilizar, preparar a partir de soluciones estériles y guardar a temperatura ambiente.

0,04X SSC + 0,1%SDS

- 1 ml SSC 20X
- 5 ml SDS 10%
- 494 ml H_2O bd

No esterilizar, preparar a partir de soluciones estériles y guardar a temperatura ambiente.

2X SSC + 1% SDS

- 50 ml SSC 20X
- 50 ml SDS 10%
- 400 ml H_2O bd

No esterilizar, preparar a partir de soluciones estériles y guardar a temperatura ambiente.

Tritón NaOH

- 0,2 ml Tritón X-100
- 100 ml NaOH 0,5N

Esterilizar mediante el autoclave a 121°C 15 min y guardar a temperatura ambiente.

Tampón de lavado

- 3 g Tween 20
- 1 L Tampón 1 de hibridación

Agitar y guardar a temperatura ambiente.

Tampones para geles de acrilamida

Gel acrilamida *resolving* al 12% (para 1 gel)

Añadir los componentes en este orden:

- 8,4 ml H_2O bd
- 6 ml *Resolving* buffer (1,5M Tris HCl, 0,4% SDS, pH 8,8)
- 9,6 ml Acrilamida al 30% (29,2% acrilamida + 0,8% bis-acrilamida; para 500 ml: 146 g acrilamida + 4 g bis-acrilamida)
- 150 µl APS 10%
- 35 µl TEMED

Gel acrilamida *stacking* al 5% (para 1 gel)

Añadir los componentes en este orden:

- 2,9 ml H_2O bd
- 1,25 ml *Stacking* buffer (0,5M Tris HCl, 0,4% SDS, pH 6,8)
- 850 µl Acrilamida al 30%
- 35 µl APS
- 10 µl TEMED

Running buffer 10x

- 15,14 g Tris (0,25M)
- 72,08 g Glicina (1,92M)
- 5 g SDS (1%)

Añadir H_2O bd hasta un volumen de 500 ml sin ajustar el pH. Guardar a 4°C.

Sample buffer (para la extracción de LPS)

- 5 g Glicerol (10% p/v)
- 2,5 ml *β*-mercaptoetanol (5% v/v)
- 1,15 g SDS (2,3%)
- 0,38 g Tris (0,0625 M)

Añadir H_2O bd hasta un volumen final de 50 ml y un par de gotas de azul de bromofenol al 0,1%. Ajustar el pH con HCl hasta 6,8 y conservar a temperatura ambiente.

Solución de lavado *overnight* para geles de acrilamida

- 55 ml H_2O bd
- 40 ml Etanol al 80%
- 5 ml Ácido acético

Solución de lavado *post-overnight* para geles de acrilamida

- 55 ml H_2O bd
- 40 ml Etanol al 80%
- 5 ml Ácido acético
- 0,7 g Ácido periódico

Tinción con nitrato de plata (para dos geles)

- 28 ml NaOH 0,1N (0,112 g en 28 ml de H_2O – hacer el cálculo pertinente)
- 115 ml H_2O bd
- 1,9 ml NH_4 al 30%
- 1 g de NO_3Ag en 5 ml de H_2O bd

Solución de revelado

- 100 ml H_2O bd
- 50 µl Ácido cítrico al 10%
- 50 µl Formaldehído

8.3. Anexo 3: Protocolos realizados

Aislamiento de cepas de *Bacteroides* (Payan et al., 2005):

- Partiendo de muestras de agua residual, realizar diluciones decimales seriadas y sembrar en placas de BBE agar. Incubar a 37°C durante 24-48 h, utilizando jarras de anaerobiosis (GasPak; BBL) con generadores de atmósfera de CO_2 (Anaerocult A; Merck) y una tira indicadora de anaerobiosis (Anaerotest®)
- Seleccionar aquellas colonias que sean de color negro o que presenten un halo negro a su alrededor.
- Resembrarlas por duplicado en placas de BBE agar, haciendo pequeñas estrías.
- Ambos duplicados se incuban a 37°C durante 24-48 h, pero uno en condiciones aeróbicas y el otro en condiciones anaeróbicas.
- Seleccionar aquellas colonias que sean estrictamente anaerobias.
- Realizar la tinción de Gram, seleccionando aquellas que sean Gram negativas con morfología de bacilo.
- En tubos de vidrio completamente llenos con medio BPRM suplementado, inocular las colonias seleccionadas y cerrar bien el tubo con la rosca para generar el ambiente anaeróbico. Incubar a 37°C durante 18-24 h en estático.
- Las cepas con buen crecimiento se seleccionan para posteriores estudios.

Preparación de cultivos stock de cepas de *Bacteroides* (ISO 10705-4 (2001):

- Con un isopo estéril sembrar una buena cantidad de cultivo de *Bacteroides* (ya sea un cultivo comercial o un cultivo antiguo) en una placa de medio BPRM agar.
- Incubar en las jarras de anaerobiosis a 37°C durante 48 h.
- Si el crecimiento de la placa es abundante, sembrar 1/8 en tubos de vidrio con 10 ml de caldo BPRM suplementado; si el crecimiento es escaso, sembrar la mitad del contenido de la placa. Incubar a 37°C durante 18-24 h.
- Para el almacenamiento, mezclar el cultivo líquido crecido con el crioprotector (BSA + sacarosa), en una proporción de 1:1. Mezclar bien evitando burbujas y distribuir en viales de congelación, y guardar a -80°C.

Titulación de fagos de *Bacteroides* mediante la doble capa de agar (ISO 10705-4 (2001)) y test de la gota:

- Previamente, partiendo de 1 ml del cultivo de 18 h inoculado en 9 ml de medio BPRM, se prepara un cultivo exponencial de la cepa de *Bacteroides* correspondiente, incubando a 37°C en anaerobiosis durante 3-5 h en estático.
- Si se parte de aguas residuales, filtrar la muestra mediante filtros de 0,22 µm de baja adsorción proteica.
- Fundir, suplementar y dispensar 2,5 ml del BPRM ss en tubos de 10 ml (tantos como sea necesario para el experimento), dejándolos en el baño previamente calentado alrededor de 55°C.
- Secar bien y rotular las placas de BPRM agar.
- Una vez el cultivo esté a una DO_{600} nm entre 0,3 y 0,5, proceder con la titulación.
- En cada tubo de 10 ml que contiene los 2,5 ml de BPRM ss, añadir 1 ml de la muestra filtrada y 1 ml de la cepa bacteriana exponencial. Agitar suavemente y verter sobre las placas de agar correspondientes. Repartir todo el volumen por la placa de manera homogénea y dejar enfriar en una superficie llana.
- Conjuntamente, realizar controles de la cepa (2,5 ml BPRM ss + 1 ml del cultivo), un control de BPRM ss (2,5 ml de BPRM ss) y, en caso de haber hecho diluciones, un control del PBS o $Ringer_{1/4}$ (2,5 ml BPRM ss + 1ml del tampón).
- Al secarse las placas, colocarlas en los GasPaks™ con el Anaerocult® para generar la anaerobiosis y una tira de Anaerotest® para controlar que se realiza correctamente.
- Incubar a 37°C durante 18-24 h.
- Comprobar que se ha producido la anaerobiosis correctamente y contar el número de calvas (ufp/ml).

Para determinar si los fagos presentes en una muestra son infectivos bastaría con realizar un test de la gota, una evaluación rápida pero que no permite la cuantificación. El protocolo es muy similar:

- Poner a crecer la cepa hospedadora en caldo ABB o BPRM a 37°C en estático y en condiciones anaeróbicas hasta alcanzar una DO de 0.3-0.5, con lo cual el cultivo estará en fase de crecimiento exponencial.

- Añadir en un tubo de 10 ml, 1 ml de bacteria exponencial + 2.5 ml de ABB ss o BPRM ss, previamente fundido y a 55°C. Agitar suavemente evitando la formación de burbujas y verter en una placa de ABB o BPRM agar.
- Esperar hasta que el agar semisólido esté solidificado y entonces se pone una gota de 10 µl de la suspensión fágica sobre éste. Esperar 10 min para que la gota se absorba e incubar 18-24 h a 37°C en anaerobiosis.

Lo óptimo es realizar diluciones seriadas de la suspensión fágica para garantizar que la zona de lisis observada sea el resultado de la infección por fagos virulentos y no de una lisis inespecífica causada por una sobrecarga de partículas fágicas. En este último caso, la adhesión simultánea de un gran número de fagos a la superficie bacteriana causa la lisis celular por degradación de la pared celular sin que se complete un ciclo normal de infección.

Propagación de bacteriófagos de *Bacteroides* en cultivo líquido

En tubos de vidrio de 10 ml mezclar los siguientes componentes:

- 1 ml de la suspensión fágica si partimos de aguas residuales, o bien 100 µl si partimos de un stock muy concentrado de fagos.
- 1 ml de la cepa de *Bacteroides* a una DO_{600} nm entre 0,3 y 0,5.
- 8 ml de caldo BPRM o ABB.

Llenar hasta arriba los tubos para generar las condiciones anaeróbicas e incubar a 37°C en estático durante 18-24 h. Tras la incubación, centrifugar a 3000 *xg* durante 15 min y filtrar el sobrenadante por filtros PES de baja adsorción proteica de 0,22 µm (Millex-GP, Merck-Millipore).

Marcaje de sondas de hibridación con DIG:

Antes de empezar con la hibridación, es necesario preparar la sonda marcada con DIG del gen diana. Para ello, se usa una reacción de PCR con el PCR DIG Synthesis Kit (Roche®), siguiendo las instrucciones del fabricante que se detallan a continuación (volumen final de 50 µl):

Reactivos	Volumen (µl)
Enzima polimerasa	0,75
DIG Mix	5
Tampón + $MgCl_2$	5
Cebador Up (10x)	5
Cebador Lp (10x)	5
H_2O bd estéril	25,25
Muestra ADN	4

Se usa el mismo programa de termociclador que el que se usa para una PCR convencional. En cuanto a la teoría de este método, los dUTPs, que son análogos de los dTTPs, van unidos a una molécula de DIG. Estos dUTPs son incorporados por la polimerasa en la amplificación por lo que el amplímero queda marcado en diversas posiciones con DIG. Este amplímero con DIG es más pesado que el amplímero con nucleótidos normales, por lo que al realizar una electroforesis con un gel de agarosa se pueden apreciar diferencias en su peso, viendo que la banda marcada se encuentra más arriba (pesa más) que la banda no marcada.

Hibridación colonial (*colony blot*, (Nizetic et al., 1991) **y protocolo indicado por Roche® para la detección de la sonda:**

Transferencia de colonias y lisis bacteriana

- Transferir las colonias bacterianas a una membrana de nailon con gran afinidad por los ácidos nucleicos (Hybond N+, Amersham®). Mantener la placa de agar con colonias/estrías en contacto con la membrana durante 2 min, para favorecer una completa transferencia.
- Colocar la membrana (siempre boca arriba) en papel Whatman® 3mm empapado con tampón de desnaturalización durante 20 min y luego secar la membrana sobre papel de nitrocelulosa 5 min.
- Transferir la membrana sobre nuevo papel Whatman® con Tritón-NaOH durante 10 min, y dejar secar de nuevo 5 min sobre papel.
- Empapar el papel Whatman® con tampón de neutralización y colocar encima la membrana 10 min. Secar la membrana durante 5 min.
- Lavar la membrana 5 min con 2XSSC en agitación y luego dejar secar 5 min más.

- En este punto, añadir un control positivo en un extremo de la membrana (1-2 µl de producto de PCR positivo para la región donde la sonda debe unirse). Volver a dejar secar la membrana.
- Fijar el ADN con luz UV, 2 min cada cara (empezando por la de abajo). Si no se continúa el análisis, guardar las membranas a 4°C envueltas en papel de celofán.

A continuación, eliminar los restos celulares con una serie de lavados:

- Lavar la membrana con 2XSSC+1% SDS a 50°C en agitación 20 min.
- Segundo lavado con 2XSSC+0,1% SDS, 10 min a temperatura ambiente en agitación.
- Finalmente, último lavado con 2XSSC, 10 min a temperatura ambiente en agitación.

A partir de aquí, empieza el proceso de hibridación, aplicando el correspondiente tampón que contiene la sonda específica marcada con DIG.

Pre-hibridación

Realizar la pre-hibridación a 68°C al menos 1 h en tampón de pre-hibridación (10 ml de tampón por cada membrana, en tubos de hibridación) y en agitación. Esto permite no sólo equilibrar la membrana para la hibridación, sino que la presencia de un agente bloqueante en el propio tampón impide la unión inespecífica de la sonda en zonas de la membrana libres de ADN, impidiendo la aparición de falsos positivos.

Hibridación

Eliminar el tampón de pre-hibridación y añadir 7 ml del tampón de hibridación (10 µl sonda específica / 50 ml tampón de pre-hibridación) por cada membrana. La hibridación se realiza a la temperatura adecuada de la sonda. En este caso se realizó a 53°C, en agitación y durante toda la noche (18 h).

Lavados y reacción colorimétrica

Realizar el revelado con un anticuerpo específico para la DIG (anti-DIG), unido covalentemente a una enzima, la fosfatasa alcalina, que en presencia de una solución cromática (NBT/BCIP, Roche®) permite el viraje de la misma a azul-lila. Por tanto, se observa coloración en aquellas zonas que hayan hibridado con la sonda (reacción positiva).

- Realizar dos lavados de 5 min a temperatura ambiente y agitación con 2XSSC+0,1% SDS.
- Después, lavar la membrana 2 veces a 64°C en 0,04XSSC+0,1% SDS, 15 min.

- A continuación, tratar 5 min con tampón 1, en agitación.
- Entonces, poner en tampón 2 e incubar 30 min, también en agitación.
- La detección de la sonda consiste en poner las membranas en tampón 2 + anticuerpo anti-DIG (10 µl anti-dioxigenin Roche® por cada 50 ml de tampón 2) e incubar 30 min en agitación.
- Para eliminar el exceso de anticuerpo, lavar 2 veces durante 15 min con tampón de lavado.
- Poner la membrana en tampón 3 (preparado al momento) durante 5 min para equilibrarse.
- Para la reacción colorimétrica, mezclar 2,5 ml de tampón 3 y 50 µl de NBT/BCIP (Roche®) y añadir a cada membrana (en una bolsa de plástico o una placa de Petri estéril). Incubar al menos 2 h en oscuridad para poder visualizar la reacción colorimétrica. Las zonas de la membrana que se corresponden con colonias/estrías positivas viran a un color azul-violeta.

Hibridación de calvas de lisis (*plaque blot*), (Sambrook and Russell, 2001)**:**

El proceso es prácticamente igual al "*colony blot*", transfiriendo la muestra a una membrana e hibridar con la sonda específica para luego detectarla con una solución cromática, pero con la excepción de que no se realizan algunos de los lavados intermedios, necesarios en el caso de tener colonias aisladas. Así pues:

- Colocar la membrana de nailon (Hybond N+, Amersham®) sobre la placa durante 5 min para la transferencia de las calvas de lisis.
- Después, colocar la membrana sobre papel Whatman® de 3 mm empapado en tampón de desnaturalización 5 min y luego secar sobre papel de nitrocelulosa 5 min.
- Transferir la membrana a un nuevo papel Whatman® con tampón de neutralización durante 5 min. Dejar secar de nuevo 5 min.
- Lavar la membrana con 2XSSC 5 min y luego dejar secar.
- En este punto, añadir el control positivo (1-2 µl de producto de PCR) en un extremo de la membrana y dejar secar.

- Fijar el ADN con UV 2 min por cada cara (primero la cara inferior). Si el proceso se para aquí, guardar la membrana a 4°C.

A partir de este punto, el protocolo sigue los mismos pasos que el descrito para la hibridación colonial, a partir de la pre-hibridación.

Purificación de bacteriófagos por gradiente de densidades de CsCl (Sambrook and Russell, 2001).

Lo primero de todo es la preparación de los stocks de los diferentes gradientes de densidades de CsCl, que posteriormente se esterilizan con filtros de ésteres de celulosa de 0,22 µm de alta adsorción proteica.

Densidad (g/ml)	CsCl por muestra (g)	H_2O por muestra (ml)
1,3	1,124	0,875
1,5	0,908	1,09
1,7	0,624	1,375

- Añadir 0,5 g de CsCl por ml de la suspensión de bacteriófagos y disolver por agitación.
- En un tubo Ultra-Clear BECKMAN®, añadir 2 ml de sacarosa al 20%.
- A continuación, realizar el gradiente de CsCl con las tres soluciones de diferente densidad añadiendo 1 ml de la de menor densidad en el fondo de la pipeta Pasteur y, a continuación, ir rellenando poco a poco con las otras dos sucesivamente. Marcar con una señal la interfase entre los gradientes. (En caso de aparecer dos o más bandas demasiado juntas unas de otras, puede hacerse añadiendo 2 ml de cada solución. De esta forma será más fácil separarlas).
- Retirar suavemente la pipeta Pasteur y añadir en la parte superior la suspensión de bacteriófagos (con el CsCl disuelto).
- Centrifugar los tubos a 54000 *xg* a 4°C durante 2 h (ultracentrífuga Beckman, rotor SW41 de ángulo variable), realizando un correcto equilibrado de la centrífuga. Habrá que ajustar el volumen de alguno de los tubos, usar H_2O bd estéril y añadirla muy lentamente sobre la parte superior para que se mantengan los gradientes.
- Después de centrifugar, recuperar las bandas grises (correspondientes cada una a una fracción fágica de la muestra) que deben aparecer a densidades aproximadas de

1,45 g/ml, mediante punción del tubo con una aguja hipodérmica estéril, dejándola gotear sobre un tubo estéril de 1,5 ml.

Diálisis de bandas obtenidas de la purificación de bacteriófagos por gradiente de CsCl

Una vez se han recuperado las bandas obtenidas por gradiente de CsCl, es necesario eliminar los restos de CsCl antes de destinarlas a cualquier protocolo posterior: observación por microscopía electrónica, test de la gota, titulación y/o extracción de ADN fágico para su secuenciación. Para esto es necesario dializar la banda como se describe a continuación:

- Añadir el volumen de la banda a una membrana de diálisis que debe sellarse completamente con pinzas de diálisis o hilo, para evitar la fuga de la muestra durante el proceso.
- Sumergir la membrana de diálisis en 2 L de tampón de diálisis en agitación durante 1-2 h.
- Al cabo de este tiempo, renovar el tampón de diálisis y dejar en agitación 18 h.
- Recuperar el volumen de banda dializado en un tubo de 1,5 ml.

Tratamiento de membranas de diálisis (para la diálisis de fagos y ensayos en mesocosmos)

- Cortar unas 25 tiras de membrana de diálisis (MWCO - 12-14000 Da) (Medicell Membranes Ltd).
- Hervir 10 min en tampón de lavado.
- Lavar con H_2O destilada.
- Hervir 10 min en tampón EDTA.
- Conservar a 4°C en tampón EDTA.

Preparación de muestras de fagos para microscopía electrónica

- Usar rejillas de cobre/níquel de 200 mesh recubiertas con una película de formvar al 5% con recubrimiento de carbón (se preparan en los Servicios científico-técnicos de la Universitat de Barcelona).

- Irradiar las rejillas durante 5 min con luz ultravioleta para facilitar una mejor extensión de la muestra sobre la rejilla.
- Depositar sobre la cara limpia de un trozo de Parafilm®, 10 µl de la muestra, 3 gotas (10 µl cada una) de H_2O bd estéril y 1 gota (10 µl) de molibdato de amonio al 2% (p/v) a pH 7,0.
- Contactar la rejilla con la muestra por la parte donde se encuentra la membrana de formvar durante 1 min.
- Lavar la rejilla pasándola sobre las 3 gotas de H_2O bd estéril durante un tiempo total de 1 min (20 s por gota).
- Poner la rejilla en contacto con el agente de tinción (molibdato de amonio al 2% pH: 7) durante 2 min. Eliminar el exceso de líquido con ayuda de un trozo de papel de filtro.
- Dejar secar la rejilla sobre papel filtro dentro de una placa de Petri a temperatura ambiente (secar durante 18 h, el tiempo necesario para el secado completo de la rejilla antes de observar la rejilla en el microscopio electrónico).
- Observar la preparación en el microscopio electrónico de transmisión del tipo JEOL JEM 1010 (JEOL) operando a 100 kv entre 60000-150000 X.

Extracción y purificación de ADN total para secuenciación (QIAmp® DNA Blood Mini Kit, QIAGEN) - instrucciones del fabricante:

- Se parte de un volumen de 200 µl de muestra.
- Añadir 20 µl de proteinasa K (QIAGEN Protease) y 200 µl de tampón AL. Agitar con el vórtex hasta que la solución sea homogénea.
- Incubar a 56°C durante 10 min y hacer una centrifugación rápida a 16000 *xg* 10 s para recuperar todo el volumen.
- Añadir 200 µl de etanol 100% y mezclar bien.
- Colocar la muestra en una columna (QIAamp Mini spin column) y centrifugar a 6000 *xg* durante 1 min. Colocar la columna en un nuevo tubo de 2 ml.
- Añadir 500 µl de tampón AW1 y centrifugar de nuevo 1 min a 6000 *xg*. Volver a cambiar el tubo de 2 ml.
- Añadir 500 µl de tampón AW2 y centrifugar 3 min a 6000 *xg*.

- Descartar el filtrado y centrifugar 1 min para eliminar los restos de tampón.
- Colocar la columna en un nuevo tubo de 1,5 ml y añadir 200 µl de tampón AE o H_2O bd estéril.
- Incubar 5 min a temperatura ambiente y centrifugar 1 min a 6000 *xg* para obtener el ADN bacteriano.
- Guardar a 4°C para un uso inmediato o a -20°C o -80°C para almacenamientos más prolongados.

Extracción y purificación de ARN (Invitrogen™ PureLink™ RNA Mini Kit) - instrucciones del fabricante:

Antes de empezar, es necesario preparar tampón de lisis fresco que contenga un 1% de 2-mercaptoetanol. Por lo tanto, añadir 10 µl de 2-mercaptoetanol por cada 1 ml de tampón de lisis.

- Centrifugar las células a 2000 *xg* por 5 min a 4°C. Descartar el sobrenadante.
- Añadir 0,6 ml de tampón de lisis con 2-mercaptoetanol y mezclar bien con el vórtex hasta que el pellet se resuspenda.
- Homogenizar a temperatura ambiente pasando el lisado de 5 a 10 veces por una jeringa con una aguja de calibre 18-21.
- Para la purificación del ARN, añadir al homogenado 0,6 ml de etanol al 70% y mezclar bien con el vórtex dispersando cualquier precipitado que se haya formado con la adición del etanol.
- Transferir la muestra a una columna con su tubo colector (máximo 700 µl) y centrifugar a 12000 *xg* durante 15 s a temperatura ambiente. Descartar la fracción que ha pasado a través de la columna y reinsertar la columna en el mismo tubo colector.
- Repetir los dos pasos anteriores hasta que se haya procesado toda la muestra.
- Añadir 700 µl de *Wash Buffer* I a la columna y centrifugar a 12000 *xg* durante 5 s a temperatura ambiente. Colocar la columna en un nuevo tubo colector.
- Añadir 500 µl de *Wash Buffer* II (previamente suplementado con etanol) y centrifugar a 12000 *xg* durante 15 s. Descartar la fracción que ha atravesado la columna.
- Repetir el paso anterior una vez más.

- Centrifugar la columna a 12000 *xg* por 1-2 min para secar la membrana con el ARN enlazado. Descartar el tubo colector y colocar la columna en un tubo de recuperación
- Añadir en el centro de la membrana 30-100 µl de H_2O libre de ARNasa.
- Incubar a temperatura ambiente por 1 min.
- Centrifugar por 2 min a más de 12000 *xg* a temperatura ambiente para eluir el ARN de la membrana.
- Almacenar el ARN extraído a -20°C.

Cuantificación de ADN y ARN

La concentración final de ADN o ARN extraídos se determina utilizando el equipo Qubit™ 4 Fluorometer (Invitrogen, Thermofisher). Para el ADN se usa el kit Qubit™ 1X dsDNA HS y para el ARN, el Invitrogen™ Qubit™ RNA High Sensitivity (HS) kit, siguiendo las instrucciones del fabricante, detalladas a continuación:

- Añadir 190 µl de la solución de trabajo Qubit™ y 10 µl de cada estándar (ST1 y ST2) a dos tubos para ensayos Qubit™.
- Añadir 2 µl de ADN purificado y 198 µl de la solución de trabajo Qubit™ a un tubo de ensayo para Qubit™. Utilizar un tubo por muestra.
- Agitar todos los tubos con vórtex durante 2-3 s.
- Incubar los tubos a temperatura ambiente durante 2 min.
- Realizar una calibración con el ST1 y el ST2. Verificar que el valor del ST2 es al menos 10 veces superior al del ST1 para una correcta calibración.
- Medir la concentración de ADN o ARN de las muestras.

Cuantificación de proteínas

Para la cuantificación de la concentración de proteínas extraídas se usa el Invitrogen™ Qubit™ Protein kit, el cual requiere de 3 estándares.

- Preparar la Solución de trabajo diluyendo 1 volumen del *Protein Reagent* en 200 ml de tampón de proteína por cada muestra a analizar. Hacer los cálculos pertinentes para un número de muestras mayor.

- Añadir 190 µl de la solución de trabajo a cada tubo usado para los estándares.
- Añadir 10 µl de cada estándar al tubo apropiado, mezclar con ayuda del vórtex durante 2-3 s. Intentar no generar burbujas.
- Añadir 198 µl de la Solución de trabajo a cada tubo de muestra y añadir 2 µl de muestra (el volumen final debe ser de 200 µl), mezclar con ayuda del vórtex durante 2-3 s.
- Incubar todos los tubos durante 15 min a temperatura ambiente y proceder con la lectura de los resultados.

Proceso de clonación empleado en esta tesis

1) PCR con dianas de restricción

Amplificar el fragmento de ADN que se quiere insertar en el plásmido añadiendo las dianas de restricción pertinentes en los cebadores. En este caso, se añadió la diana para la enzima *BamH*1 a cada lado del inserto.

2) Electroforesis comprobación

Mediante un gel de agarosa al 1,5%, correr 5 µl del producto de PCR para comprobar que el tamaño del inserto es el correcto y confirmar que solo hay una banda. Utilizar RedSafe™ para la tinción del ácido nucleico (7,5 µl por cada 150 ml de gel).

3) Purificación de amplímeros de PCR a partir del tubo de reacción (PureLink® Quick Gel Extraction and PCR Purification Combo Kit, Invitrogen)

Recomendado para purificar amplímeros de PCR cuando en el gel de agarosa se ha observado una única banda de amplificación. Las instrucciones descritas en el kit se indican a continuación:

Antes de empezar añadir etanol del 96-100% al *Wash Buffer* (W1) e isopropanol al *Binding Buffer* (B2) de acuerdo con las etiquetas de las botellas.

- Añadir 4 volúmenes de *Binding Buffer* (B2) con el isopropanol incorporado a 1 volumen de muestra de PCR (50-100 µl). Mezclar bien.
- Añadir la muestra a la columna PureLink® Clean-up Spin en un tubo recolector. Centrifugar la columna a > 10000 *xg* por 1 min. Descartar el sobrenadante y volver a colocar la columna en el mismo tubo.
- Añadir 650 µl de Wash Buffer (W1) con etanol. Centrifugar la columna a > 10000 *xg* por 1 min.

- Descartar el sobrenadante y volver a colocar la columna en el mismo tubo y centrifugar la columna a máxima velocidad por 2-3 min.
- Colocar la columna en un nuevo tubo de elución de 1,7 ml. Añadir 50 µl de *Elution Buffer* (E1) a la columna.
- Incubar a temperatura ambiente por 1 min, y posteriormente, centrifugar a máxima velocidad durante 1 min.
- El tubo de elución contiene el producto de PCR purificado. Almacenar el ADN purificado a 4°C para un uso inmediato o a -20°C para un almacenamiento a largo plazo.

4) Cuantificación ADN

Mediante el Qubit™, cuantificar la concentración de ADN amplificado, tal y como se especifica en el apartado de cuantificación de ADN.

5) Ligación con el vector pGEM-T Easy (Promega)

Este paso de clonación en el plásmido pGEM-T Easy se realizó para generar múltiples copias del inserto (ya que el pGEM es un plásmido multicopia) y mejorar la eficiencia tras el corte con enzimas de restricción, en lugar de utilizar únicamente el fragmento amplificado. Los pasos para la ligación se siguieron según el fabricante:

- Centrifugar brevemente para recoger todo el volumen del plásmido en el fondo del tubo. La cantidad de inserto de ADN necesario para la ligación debe cumplir una ratio de 3:1 (inserto:vector) cuando se trata de un fragmento con extremos protuberantes. En caso de extremos romos la ratio sería de 5:1. En este caso, se trata de un fragmento con extremos protuberantes.

$$\frac{50ngvector \cdot kbdelinserto}{3kbvector} \cdot \frac{3}{1}(ratio\frac{inserto}{vector})$$

  Si la concentración no es suficiente, se puede concentrar utilizando el secador de vacío (SpeedVac) hasta secar completamente la muestra y luego resuspender en un menor volumen. O bien volver al primer paso y obtener un volumen mayor de producto de PCR.

- Poner a punto la reacción de ligación, descrita a continuación:

2x Tampón de ligación	5 µl
Vector pGEM-T Easy	1 µl
ADN purificado	x*
Ligasa T4	1 µl
H_2O bd estéril	Añadir el volumen necesario para llegar a 10 µl de volumen final de la reacción

*El volumen según los resultados del cálculo empleando la fórmula anterior.

- Mezclar bien la reacción con la pipeta e incubar 18 h a 4°C para un mayor número de transformantes.

6) Preparación de células electrocompetentes y clonación en *E. coli* DH5α

- Inocular 0,5 ml de un cultivo de 18 h de *E. coli* DH5α en 5 ml de caldo LB. Incubar a 37°C durante 2 h con agitación.
- Centrifugar a 1500 *xg* durante 15 min a 4°C y eliminar el sobrenadante. Poner siempre las muestras en hielo para mantenerlas frías.
- Añadir al pellet de bacterias 4,5 ml de H_2O bd fría, siempre estéril, resuspender bien y centrifugar de nuevo a 1500 *xg* durante 10 min a 4°C y eliminar el sobrenadante. Repetir este proceso mínimo 3 veces.
- Añadir 1 ml de H_2O bd fría y resuspender el pellet en un tupo de 1,5 ml.
- Tras el último lavado, centrifugar a 16000 *xg* durante 3 min, eliminar el sobrenadante y resuspender en 100 µl de H_2O bd fría.
- Una vez preparadas las células electrocompetentes, añadir al cultivo 10 µl de la reacción de ligación preparada el día anterior. Pasar la muestra a una cubeta de electroporación de 0,2 cm (previamente mantenida en hielo) y electroporar las células (2kV, 25 F , 200 Ω).
- Recuperar el volumen de la cubeta con 1 ml de caldo LB e incubarlo a 37°C durante al menos 1 h en agitación.
- Sembrar 100 µl del cultivo en placas de LB con X-Gal, IPTG y el antibiótico correspondiente (en este caso ampicilina a 100 µg/ml) e incubar a 37°C 18 h. El volumen restante del cultivo electroporado se centrifuga a 2000 *xg* durante 5 min y el pellet se resuspende en 100 µl de caldo LB y se plaquea en el mismo tipo de placas.

De esta manera se concentra la muestra para tener más probabilidad de encontrar células con el vector recombinante.

Las colonias de color blanco corresponden a aquellas que, potencialmente, han incorporado el plásmido recombinante, permitiéndoles crecer en presencia del antibiótico. La interrupción del gen lacZ por la inserción del fragmento de interés impide la síntesis de B-galactosidasa, lo que imposibilita la hidrólisis del X-Gal, un sustrato cromogénico que, en condiciones normales, genera un producto de color azul al ser degradado. En consecuencia, la ausencia de actividad enzimática mantiene la coloración blanca de estas colonias. En contraste, las colonias no recombinantes, es decir, aquellas que han incorporado el plásmido, pero carecen del inserto, conservan la integridad del gen lacZ, permitiendo la expresión de la B-galactosidasa y la consecuente hidrólisis del X-Gal, lo que da lugar a la coloración azul característica de las colonias.

- A continuación, se puede hacer un paso de comprobación opcional en el que se resiembran en forma de estría aquellas colonia que presentan un color blanco en las mismas placas de LB con X-Gal, IPTG y el antibiótico requerido. Incubar a 37°C 18 h.
- Resembrar las colonias blancas en placas de LB con el antibiótico correspondiente e incubar a 37°C 18 h.

7) Confirmación por PCR de las colonias blancas

Para verificar la presencia del vector recombinante en las colonias de color blanco y la orientación del inserto, realizar una PCR utilizando los cebadores del fragmento de interés con su respectiva electroforesis. Los cebadores utilizados se indican en el anexo 4.

8) Extracción del plásmido (Zyppy™ Plasmid Miniprep Kit)

Una vez confirmada la presencia del vector con el inserto, se preparó un cultivo de 18 h de la cepa clonada. Además, también se puso un cultivo de la cepa que contenía el vector pG106 (Addgene) para *Bacteroides*. Para la extracción y purificación de ambos plásmidos, se siguieron las instrucciones del fabricante del kit que se indican a continuación:

- Añadir 600 µl de cultivo bacteriano, previamente crecido en medio LB, a un tubo de microcentrífuga.
- Añadir 100 µl del tampón de lisis 7X (azul) y mezclar mediante 4-6 inversiones del tubo. Proceder con el siguiente paso después de 2 min.
- Añadir 350 µl del tampón de neutralización frío y mezclar bien. La muestra se volverá de color amarillo cuando la neutralización se haya completado, formándose, además,

un precipitado del mismo color. Invertir unas 2 o 3 veces adicionales para asegurar la completa neutralización.

- Centrifugar a 11000 – 16000 *xg* por 2 min.
- Transferir el sobrenadante a una columna Zymo-Spin™ II. Evitar recoger el pellet de restos celulares.
- Colocar la columna en un tubo colector y centrifugar 15 s.
- Descartar la fracción que ha pasado la membrana y volver a colocar la columna en el mismo tubo colector.
- Añadir 200 µl de tampón *Endo-Wash* a la columna. Centrifugar por 15 s. En este paso no hace falta vaciar el tubo colector.
- Añadir 400 µl del tampón *Zyppy™ Wash* a la columna. Centrifugar por 30 s.
- Transferir la columna a un tubo limpio de microcentrífuga y añadir 30 µl de H_2O bd estéril o de tampón *Zyppy™ Elution* y dejar de 1 a 4 min a temperatura ambiente.
- Finalmente, centrifugar por 15 s para eluir el ADN plasmídico extraído y purificado.
- Por último, cuantificar la cantidad de plásmido extraída en cada caso, ya que para el siguiente paso de digestión es necesario saber la concentración.

9) Digestión de los plásmidos (el pGEM con el inserto y pG106, que viene circularizado) con la enzima de restricción

Por cada 300 – 400 ng de ADN plasmídico, añadir:

- 1 µl Enzima de restricción
- 2 µl Tampón correspondiente a la enzima de restricción utilizada
- 2 µl BSA 10X
- X µl H_2O pura de PCR, hasta llegar a los 20 µl de volumen final de reacción

Incubar durante 2 h a 37°C. A continuación, la muestra de plásmido pG106 digerido, purificarla con el QIAquick® PCR Purification Kit (QIAGEN), mientras que la muestra de pGEM digerido, correrla en un gel de agarosa del 1,5% para extraer el inserto, mediante la purificación a partir de una banda en gel de agarosa.

10) Purificación ADN a partir de una banda visualizada en el gel de agarosa PureLink® Quick Gel Extraction and PCR Purification Combo Kit, Invitrogen)

Recomendado cuando en el gel de agarosa se observa, además de la banda que se quiere amplificar, otras bandas, en este caso, la del plásmido. Las instrucciones descritas en el kit se indican a continuación:

Añadir etanol al *Wash Buffer* (W1) de acuerdo a la etiqueta de la botella.

- Utilizar un bisturí para recortar un área mínima del gel que contenga el fragmento de ADN de interés.
- Pesar el gel que contiene el ADN.
- Añadir el *Gel S*olubilization Buffer (L3) al gel cortado colocado en un tubo. Si el porcentaje de agarosa es menor o igual a 2 la proporción es de 3:1 (L3:gel). Si el porcentaje de agarosa es mayor de 2, la proporción es de 6:1.
- Incubar el tubo a 50°C al menos 10 min e ir invirtiendo el tubo cada 3 min. Después de que el gel se haya disuelto completamente, incubar 5 min adicionales. De manera opcional, añadir 1 volumen de gel de isopropanol al gel disuelto y mezclar bien.
- Pasar el gel disuelto a una columna dentro de un tubo colector. Centrifugar la columna a >12000 *xg* por 1 min. Descartar el sobrenadante y volver a colocar en el mismo tubo colector.
- Añadir a la columna 500 µl de *Wash Buffer* (W1) que contiene etanol. Centrifugar a >12000 *xg* durante 1 min. Descartar el sobrenadante y centrifugar otro 2-3 min a máxima velocidad.
- Colocar la columna en un tubo de recuperación. Añadir 50 µl de *Elution Buffer* (E1) a la columna. Incubar el tubo durante 1 min a temperatura ambiente. Centrifugar a >12000 *xg* durante 1 min.
- El volumen eluido contiene el ADN purificado. Almacenar el ADN purificado a 4°C para un uso inmediato o a -20°C para un almacenamiento a largo plazo.

Finalmente cuantificar los productos digeridos y purificados, para el siguiente paso de ligación.

11) Ligación con pG106

Seguidamente, se aplica la misma fórmula que la mencionada en el quinto paso, cambiando el tamaño del vector, que ahora es de 12,948 kb, para poder calcular el volumen de inserto purificado que se tiene que añadir a la reacción, que consta de:

- X µl del inserto
- 50 ng de plásmido (hacer el cálculo pertinente dependiendo de la concentración tras la extracción).
- 1 µl Ligasa
- 3 µL Tampón de ligación 5x
- H_2O hasta llegar a un volumen de reacción de 15 µl.

Dejar durante 18 h a unos 20°C.

12) Clonación en *E. coli* DH5α

Seguir los mismos pasos que en los puntos 6 y 7. La única diferencia es que el antibiótico empleado en este caso es la eritromicina a una concentración de 300 µg/ml.

13) Conjugación triparental

A partir de los crecimientos en placas de las diferentes cepas, se preparan las siguientes muestras:

- Cepa receptora (*Bacteroides*): resuspender en 150 µl de medio (en este caso ABB).
- Cepa helper (*E. coli* HB101/pRK2073): resuspender 50 µl de medio.
- Cepa donadora (*E.coli* DH5α con el plásmido pG106 clonado): resuspender en 50 µl de medio.

En un tubo de 1,5 ml mezclar: 50 µl de la cepa receptora, 10 µl de la helper y 10 µl de la donadora. Todo el volumen de la mezcla se pone en forma de una gota en una placa de medio (en este caso ABB) sin antibiótico. En cada placa se pueden poner un par de gotas. Incubar 18 h a 37°C boca arriba.

Tras el tiempo requerido para la conjugación, se añade 1 ml de medio líquido a la placa y se recoge todo el cultivo con una asa estéril. Hacer un banco de diluciones, incluyendo la muestra concentrada y extender con una asa 100 µl de cada dilución en una placa de BBE con eritromicina (50 µg/ml). Incubar durante 48 h a 37°C. Una vez pasado el tiempo, seleccionar las colonias con halo negro y resembrarlas de nuevo en una placa de BBE con eritromicina (50 o 300 µg/ml) para su confirmación.

14) PCR confirmación

Para verificar la presencia del vector recombinante en las colonias con halo negro y la orientación del inserto, realizar una PCR utilizando los cebadores del fragmento de interés y los del plásmido, con su respectiva electroforesis.

15) Secuenciación Sanger de muestras "*premixed*"

- Mezclar el ADN que se quiere analizar (lo óptimo son entre 25 – 100 ng por cada kb de ADN molde) con el cebador (3,2 – 5 pmoles, siendo µM equivalente a pmol/µl) en un tubo de PCR. En este caso por cada muestra se realizaron dos reacciones, una para el cebador Up y otra para el cebador Lp.
- Dejar el tubo abierto y secar en un termociclador a 80°C hasta la evaporación total de la muestra.
- Tapar las muestras desecadas, rotular bien y guardar a 4°C o a -20°C hasta su procesamiento en el servicio científico técnico de la UB (CCIT UB).

Extracción de proteínas de membrana (Thein et al., 2010):

- Partir de 1 L de cultivo bacteriano. Centrifugar a 10000 *xg* durante 15 min a 4°C, para bajar el pellet bacteriano y resuspender en 10 ml de 0,2M Tris-HCl pH 8, 1M sacarosa, 1mM EDTA.
- Añadir lisozima a una concentración final de 1 mg/ml. Mezclar bien la suspensión y dejar por 5 min a temperatura ambiente.
- Añadir 40 ml de H_2O y colocar las muestras en hielo (se produce *esferoblasting* que se puede observar al microscopio).
- Centrifugar las células a 200000 *xg* (rotor 70Ti Ultracentrifuga = 52214 rpm) durante 45 min a 4°C. El sobrenadante contiene la fracción periplasmática.
- Resuspender el pellet en 7,5 ml en 10 mM Tris-HCl pH 7,5, 5 mM EDTA, 0,2 mM DTT suplementado con 50 µl de DNAsa (1 mg/mL). En este caso, no utilizamos el DTT ya que no fue necesario.
- A continuación, romper las células mediante French Press, haciendo dos pases a 10^8 Pa.

- Bajar las células que no se han roto mediante centrifugación a 3000 *xg* por 10 min a 4°C.
- Seguidamente, centrifugar el sobrenadante a 280000-300000 *xg* por 2-4 h a 4°C. El sobrenadante contiene la fracción citoplasmática y el pellet contiene las membranas crudas.
- Para limpiar las proteínas de membrana, resuspender el pellet en 2 ml de 50 mM Tris-HCl pH 8, 2% (w/v), Tritón X-100, 10 mM $MgCl_2$ e incubar de 30 min a 1 h a 4°C en agitación. Centrifugar a 105000 *xg* durante 1 h a 4°C. El sobrenadante contiene las proteínas de membrana.

Extracción LPS

Los geles de LPS pueden realizarse a partir de células enteras o bien a partir de LPS purificado o de fracciones.

- Preparar cultivos de 18 h en placa de las bacterias a analizar o en medio líquido si la bacteria crece bien. En el caso de *Bacteroides*, se optó por el cultivo en placa.
- Coger con un palillo una pequeña cantidad del cultivo sólido e inocular 150 µl de PBS 1x y centrifugar a 3000 *xg* durante 5-10 min. En el caso de trabajar con muestras líquidas, centrifugar directamente 150 µl del cultivo.
- Resuspender el pellet seco en 25 µl de *Sample Buffer*. Mezclar bien con el vórtex y dejar 15 min a 100°C.
- Dejar que enfríe un poco y añadir 25 µl de proteinasa K (1 mg/ml en Tris-Mg a pH 7,3) e incubar entre 1,5 y 2 h a 60°C.
- Una vez terminados todos los pasos, y después de enfriarse la muestra, colocar las muestras en los pocillos del gel de acrilamida. Si no, se puede almacenar a -20°C hasta su procesamiento. Si se opta por lo último, cuando se quiera utilizar la muestra, se deberá descongelar y tratar 10 min a 100°C antes de correr el gel.
- Una vez colocadas las muestras conectar la fuente a 9,3-10 miliamperios hasta que las muestras pasen la zona del *stacking gel* y seguidamente, subirlo hasta 20 miliamperios hasta que llegue al final la electroforesis.
- En caso de correr dos geles a la vez en el mismo aparato, duplicar la corriente.
- Finalizada la electroforesis iniciar la tinción del gel.

Montaje y preparación de los geles de acrilamida

Resolving gel:

- Lavar las placas de vidrio y de teflón con cuidado sin rayarlas, sumergir en alcohol un cierto tiempo.
- Dejarlas secar y limpiarlas bien con papel suave.
- Untar los separadores con silicona y colocarlos en los laterales, entre el vidrio y el teflón. Colocar un par de pinzas en cada lado para que el separador quede sujeto.
- Una vez montado, pasar a preparar el *resolving gel buffer*. Seguidamente, con ayuda de una pipeta de 5 ml colocar el *resolving gel buffer* entre la placa de vidrio y la teflón, procurando que no se formen burbujas. Llenar hasta que falte un poco para llegar arriba pues más adelante se colocará el *stacking gel buffer* en el que se incorporará el peine para cargar las muestras.
- Una vez colocado el *resolving gel buffer*, poner sobre él una pequeña cantidad de agua saturada con butanol, para que al polimerizar quede recto y dejar que polimerice entre 20 - 30 min.
- Seguidamente, vaciar el contenido de butanol del gel y lavar con un poco de H_2O destilada.
- Estos geles con el *resolving gel* preparado, pueden guardarse a 4°C en una cámara húmeda durante un cierto tiempo.
- El *stacking gel buffer* debe preparase en el momento de usarlo.

Stacking gel:

- Sujetar el gel por los separadores laterales con una pinzas y rellenar con aproximadamente 1 ml de *stacking gel*, la zona superior.
- Una vez lleno, colocar un peine, para formar los pocillos en los que se cargarán las muestras.
- Dejar el peine colocado durante 10-15 min y sacarlo con cuidado.
- Lavar los pocillos con H_2O desionizada y marcar cada uno de ellos para localizarlos una vez montado el gel en el aparato para correrlo.

Tinción de geles de acrilamida

Durante toda esta tinción debe vigilarse estrictamente de no tocar el gel con las manos, cualquier manipulación debe realizarse con guantes. Además, toda la tinción debe realizarse con una suave agitación.

- Sumergir el gel durante 2 h como mínimo (normalmente se deja 18 h o durante toda la noche) en la solución de lavado *overnight*.
- Al día siguiente sumergir el gel durante 5-7 min en la solución de lavado *post-overnight*.
- Lavar 3 veces con H_2O destilada, con una duración de 10 min cada vez.
- Incubar durante 10 min exactos en la mezcla de nitrato de plata. Esta mezcla, sobre todo si el amoniaco es del 30% es mejor realizarla por duplicado, es decir, multiplicando todas las cantidades por dos. En el Anexo 2 ya se detallan las cantidades duplicadas.
- Lavar tres veces con H_2O destilada, con una duración de 5 min cada vez.
- Para el revelado, sumergir el gel en la solución de revelado. Esta reacción debe controlarse.
- Una vez desarrollada la reacción se parará sumergiendo el gel durante 5-10 min en metilamina diluida en H_2O destilada.
- Seguidamente, sumergir el gel durante 10 min en H_2O destilada y luego dejar secar.
- En caso de que no se pretenda secar el gel el mismo día, se puede sumergir en una solución (36,5 ml H_2O bd + 12,5 ml etanol + 1 ml glicerol) para que no se hidrate excesivamente.

Inactivación de fagos

Para determinar si la expresión diferencial del receptor del fago es responsable del fenotipo resistente se realizaron ensayos de inactivación de fagos utilizando la fracción de membrana de las células bacterianas previamente extraídas. Se deben poner a punto:

- 1 control negativo: 100 µl de fago + 100 µl medio ABB
- 1 control positivo: 100 µl del fago + 100 µl de bacteria inactivada por sonicación.
- Muestras para analizar: 100 µl del fago + 100 µl de las fracciones de membrana correspondientes.

Todas las muestras y los controles se incuban a 20°C durante 20 min en estático. Posteriormente se tratan con 10 µl de cloroformo, que luego se elimina por centrifugación. Finalmente, se usan 100 µl de cada muestra para llevar a cabo el análisis de fagos por doble capa de agar y así cuantificar el número de fagos infectivos que no se han adherido a la membrana celular.

Se deben utilizar suspensiones fágicas con una concentración aproximada de 100 fagos/ml. Por ejemplo, si se parte de una suspensión con 5 x 10^8 ufp/ml, se realiza una dilución de 1/10^5, obteniendo una concentración de 5 x 10^3 ufp/ml. Para el ensayo, se emplean 100 µl de esta suspensión (equivalente a 500 ufp/ml) y se mezclan con 100 µl de medio o membranas. Al sembrar 100 µl de la mezcla en la placa, se espera obtener alrededor de 250 ufp por placa, siendo posible su recuento.

Experimentos de inactivación y persistencia

Todas las muestras obtenidas tras los diversos experimentos de inactivación y persistencia fueron posteriormente utilizadas para analizar los fagos infectivos mediante la técnica de la doble capa y para cuantificar el ADN fágico mediante ensayos de qPCR.

Influencia de temperaturas y pH

- Para la temperatura: Llenar tubos de vidrio de 10 ml estériles con las muestras correspondientes (suspensiones fágicas o bacterianas) y colocarlos a 4°C, 22°C y 37°C, apartados de la luz. Asegurar que los tubos queden bien cerrados para garantizar la anaerobiosis en el caso de las muestras que contengan *Bacteroides.*
- Para el pH: el pH de las muestras fue ajustado a 3, 7 y 9 usando ácido clorhídrico al 37% o hidróxido de sodio 1N. una vez ajustado, llenar los tubos de vidrio estériles y colocar a 4°C en oscuridad.
- A cada día de muestreo (1, 3, 8, 15, 22 y 29), remover bien la muestra, coger una alícuota y verificar el pH en el caso correspondiente.

Cloración

- Diluir 100 veces la suspensión fágica o del cultivo bacteriano utilizando H_2O bd estéril, con el objetivo de minimizar las interferencias del medio con el cloro.
- Añadir hipoclorito de sodio hasta una concentración final de 10 ppm de cloro libre. Confirmar esta concentración utilizando la técnica colorimétrica DPD (N,N-diethyl-p-

phenylenediamina). Para ello, se puede usar del Chlorine Test Kit (Merck) siguiendo las instrucciones del fabricante.

- Tomar alícuotas después de 1, 5, 10, 20 y 30 min de tratamiento a temperatura ambiente. A los 30 min volver a confirmar la concentración de cloro libre en las muestras.
- A cada tiempo y antes de analizar las muestras, neutralizar el cloro residual añadiendo tiosulfato de sodio al 3% (p/v).

Tratamiento térmico

- Añadir 1 ml de cada muestra por cuadruplicado en tubos de 1,5 ml.
- Incubar dos tubos de cada muestra a 60°C y los otros dos a 70°C, utilizando el Thermo Block (Fisherbrand Isotemp Digital Dry Baths/Block Heaters).
- Sacar un tubo de cada temperatura a los 30 min de ensayo y los otros al cabo de 1 h.

Radiación UV

Utilizar una fuente de radiación UV. En este caso se usó una lámpara germicida de 8W, modelo G30T8, 0,099 mW/cm2, con radiación a 253,7 nm (Sankyo Denki, Tokyo, Japón).

- Calentar la lámpara durante 15-30 min antes de iniciar el experimento.
- Pasar 10 ml de muestra a una placa de Petri de plástico de 90 mm de diámetro.
- Colocar la placa a 10 cm de la fuente de radiación, y removerlas después de 1, 5, 10 y 30 min de ensayo.

Inactivación natural mediante mesocosmos

Los experimentos de mesocosmos se realizaron en verano y en invierno. Se cuantificaron diversos parámetros ambientales (Meteocat): temperatura media exterior y del agua y la irradiación solar, además de otros factores fisicoquímicos de la balsa como la conductividad, el pH y el oxígeno disuelto (sonda multiparámetro).

- Usar membranas de diálisis (14kDa), tratadas de la misma manera que las que se usan tras la purificación de fagos por gradiente de cloruro de cesio.
- Rellenar las membranas con alícuotas de las muestras pertinentes, sellarlas con pinzas y protegerlas con una jaula a una profundidad de 10 cm.
- Mezclar bien las alícuotas después de 1, 3, 7, 10, 14, 17, 20 y 24 días en verano y después de 1, 3, 8-10, 15-17, 24, 31, 38 y 45 días en invierno.

Tinciones para el estudio de la cápsula bacteriana

Evaluación de la formación de biofilm con cristal violeta (Béchon et al., 2020).

- Diluir un cultivo de 18 h a una DO de 600 nm de 0,05 en 100 µl de caldo ABB.
- Inocular por duplicado en placas de 96 pocillos de fondo redondo. Llenar los pocillos en el borde de las placas con 200 µl de H_2O para evitar la evaporación. Incubar a 37 °C en condiciones anaerobias durante 48 h.
- Añadir directamente 25 µl de solución de Bouin (ácido pícrico 0,9%, formaldehído 9% y ácido acético 5%) a cada pocillo antes de retirar el sobrenadante, con el fin de evitar la eliminación de grandes agregados durante el lavado.
- Fijar los cultivos durante 10 min y, posteriormente, lavar los pocillos una vez con H_2O por inmersión y agitación.
- Teñir el biofilm con 125 µl de cristal violeta al 1% durante 10 min.
- Eliminar la solución de cristal violeta por agitación, y lavar los biofilms dos veces con H_2O.
- Finalmente, resuspender los biofilms teñidos en una mezcla de 1:4 acetona:etanol, y medir la absorbancia a 575 nm utilizando un lector de placas.

Tinción de cápsula con tinta china (Sagar Aryal, 2022).

- Coloca una pequeña gota de un colorante de tinción negativa (tinta China, Rojo Congo, Nigrosina o Eosina) en el portaobjetos. El Rojo Congo es más fácil de visualizar, pero no funciona bien con algunas cepas. La tinta China suele ser efectiva, pero contiene partículas diminutas que presentan movimiento browniano y deben diferenciarse de las bacterias. La Nigrosina puede necesitar dilución o aplicarse en una capa muy fina.
- Añadir el cultivo bacteriano al portaobjetos con una asa se siembra estéril y mezclar con el colorante. Utilizar otro portaobjetos para extender la mezcla de tinta y bacterias en una película delgada a lo largo del primer portaobjetos. Deja secar al aire (no fijar con calor).
- Cubrir el frotis con cristal violeta (esto teñirá las células, pero no las cápsulas) durante aproximadamente 1 min. Drenar el exceso de colorante inclinando el portaobjetos 45° y dejar que la tinción escurra hasta que se seque al aire.

- Examinar el frotis al microscopio con un aumento de 100X en busca de células encapsuladas, las cuales se observarán con zonas claras a su alrededor.

8.4. Anexo 4: Reacciones de PCR, qPCR, RT- qPCR y cebadores utilizados

Para el diseño de todos los cebadores se utilizó la herramienta PrimerQuest de IDT (Integrated DNA Technologies, s.f.).

PCR convencional

Reactivo	Volumen (µl)
DreamTaq Green PCR Master Mix (2x) (Thermo Fisher Scientific)	12,5
Cebadores UP y LP (10 µM)	1
ADN	2
H_2O pura de PCR	8,5
Volumen final	25

La PCR se llevó a cabo utilizando el equipo Thermal Cycler 2720 (Applied Biosystems). El ciclo del termociclador fue el siguiente:

- Un ciclo inicial de desnaturalización a 95°C / 5 min
- 30 ciclos de: desnaturalización a 95°C / 45 s ; anillamiento a 52°C / 30 s ; elongación a 72°C / 1 min
- Un último ciclo de elongación a 72°C / 10 min.

La temperatura de anillamiento depende de la temperatura de fusión (Tm) de los cebadores, por lo tanto, se debe ajustar según los cebadores empleados.

qPCR utilizando sonda de hidrólisis (TaqMan®)

Para el diseño de los cebadores y sondas se deben seguir una serie de criterios:

- La secuencia diana deber tener un tamaño entre 50 y 150 pb.
- El tamaño de la sonda debe ser mayor al de los cebadores, siendo la Tm de los cebadores entre 58-60°C y la de la sonda entre 68-70°C.
- Evitar que hayan más de tres repeticiones de una base en los cebadores, sobre todo que no hayan más de 3 guaninas (G) o citosinas (C) en los extremos 3’.
- Impedir que haya 3 guaninas juntas o la combinación GA (guanina-adenina) en el extremo 5’.

Reactivo	Volumen (µl)
TaqMan® Environmental Master Mix 2.0 (Applied Biosystems) 2x	10
Cebadores UP y LP, y sonda TaqMan® (Single Tube customizado) 20x	1
ADN	9
Volumen final	20

La qPCR se llevó a cabo utilizando el equipo StepOne Real Time PCR System (Applied Biosystems). El ciclo del termociclador fue el siguiente:

- Paso inicial a 50°C / 2 min
- Un ciclo de activación a 95°C / 10 min
- 40 ciclos de: desnaturalización a 95°C / 15 s; elongación a 60°C / 1 min

qPCR utilizando SYBR™ Green

SYBR™ Green es un colorante fluorescente de unión de ADN bicatenario que permite la detección del producto de PCR que se genera y acumula durante la reacción. La unión del SYBR™ Green al ADN produce un cambio estructural en la molécula del colorante, emitiendo la fluorescencia, que aumenta con la acumulación del producto de PCR. La detección de la fluorescencia permite la cuantificación del gen diana.

Reactivo	Volumen (µl)
Power SYBR™ Green PCR Master Mix (Applied Biosystems) 2x	10
Cebadores UP y LP (3 µM)	2
ADN	6
Volumen final	20

Dependiendo de la Tm de los cebadores empleados, se usará un ciclo u otro del termociclador:

Tm ≥ 60°C

- Paso inicial a 50°C / 2 min
- Un ciclo de activación a 95°C / 2 min
- 40 ciclos de: desnaturalización a 95°C / 15; elongación a 60°C / 1 min

Tm < 60°C

- Paso inicial a 50°C / 2 min
- Un ciclo de activación a 95°C / 2 min
- 40 ciclos de: desnaturalización a 95°C / 15 s ; anillamiento a 55-60°C / 15 s ; elongación a 72°C / 1 min

También es necesaria la generación de curvas de disociación (Melting curves) para detectar una posible amplificación inespecífica que interfiriera con el correcto resultado de la reacción (95°C /15 s; 60°C / 1 min; 95°C / 15 s).

RT-qPCR utilizando SYBR Green

La amplificación de ARN mediante qPCR re realizó utilizando el sistema StepOne™ Real-Time PCR (Applied Biosystems™). La reacción se llevó a cabo en un volumen total de 20 µl, empleando el kit Power SYBR™ Green RNA-to-C_t 1-Step (Applied Biosystems™). Esta metodología de RT-qPCR en un solo paso permite la conversión de ARN a ADNc y la amplificación por qPCR en la misma reacción, utilizando 2 µl de ARN como molde.

Reactivo	Volumen (µl)
Power SYBR® Green RT-PCR Mix (2x) (Applied Biosystems)	10
RT Enzyme Mix (125x)	0,16
Cebadores Up y Lp (3 µM)	1
ADN	6
H_2O pura de PCR	1,84
Volumen final	20

Condiciones del termociclador:

- Un ciclo de transcripción inversa a 48°C / 30 min.
- Un ciclo de activación de la polimerasa a 95°C / 10 min.
- 40 ciclos de: desnaturalización a 95°C / 15 s; anillamiento y elongación a 60°C / 1 min.
- Curva de disociación opcional: desnaturalización a 95°C / 15 s; anillamiento a 60°C / 15 s; desnaturalización a 95°C / 15 s.

	Cebador	Secuencia (5'-3')	Tamaño amplímero
Cebadores de *p*-crAssphage (García-Aljaro et al., 2017)	Crass-2-F	AGGAGAAAGTGAACGTGGAAACA	78 pb
	Crass-2-R	AACGAGCACCAACTTTAAGCTTTA	
	Crass-2-S	AGGATTTGGAGAAGGAA	
Cebadores de crAss001	Crass001-F	ACCTGCTTCTACACTTTCCTT	75pb
	Crass001-R	AGTGCTCCAGAATAGGATTGT	
	Crass001-S	ATATCTAACCCAGCTC	
Confirmación del fenotipo Sensible	Int-S	TTCAACCTGTATCTGTCCTAAG	409 pb
	Ext1	CCACCTACCTATTCATATAAAGTGTGTGTA	
Confirmación del fenotipo Resistente	Int-R	GTGCACAAAGATAATGTCTTGA	418 pb
	Ext1	CCACCTACCTATTCATATAAAGTGTGTGTA	
Confirmación del inserto PVR7	Ext2	CGATAAGTGGCACGCATAGC	696 pb
	Ext1	CCACCTACCTATTCATATAAAGTGTGTGTA	
Confirmación inserto PVR7-UpxY	7upx	GCACGCTGAAGAGTGATTGC	1252 pb
	Ext1	CCACCTACCTATTCATATAAAGTGTGTGTA	
Clonación del inserto PVR7 con dianas BamH1	Ext2-bam	CGCGGATCCCGATAAGTGGCACGCATAGC	714 pb
	Ext1-bam	CGCGGATCCCCACCTACCTATTCATATAAAGTGTGTGTA	
Clonación del inserto PVR7-UpxY con dianas BamH1	7upx-bam	CGCGGATCCGCACGCTGAAGAGTGATTGC	1270 pb
	Ext1-bam	CGCGGATCCCCACCTACCTATTCATATAAAGTGTGTGTA	
pG106	M13-F	CAGGAAACAGCTATGAC	800 pb – PVR7
	M13-R	TGTAAAACGACGGCCAGT	1362 pb – PVR7UpxY

Confirmación orientación resistente	M13-F	CAGGAAACAGCTATGAC	425 pb – PVR7
(ON) del inserto en pG106	Int-R	GTAATATGTGACACTGA	986 pb – PVR7UpxY
Confirmación orientación sensible	M13-F	CAGGAAACAGCTATGAC	522 pb – PVR7
(OFF) del inserto en pG106	Int-S	GTGCACAAAGATAATGT	1083 pb – PVR7UpxY
ORF II	ORFII-F	CTAATGTCTGTCTTGCCCTAC	119 pb
	ORFII-R	AGGGAGAGCATCAAGAAGA	
ORF III	ORFIII-F	GGGATATGCCAAGGATTATGT	108 pb
	ORFIII-R	GCAAACTCACGGACTGTAT	
ORF V	ORFV-F	GCTACCTCTATCGATGTTTACAA	112 pb
	ORFV-R	CCACCTCTGTCCAAGAATTAC	
ORF VI	ORFVI-F	GGAGACTCAAACCTGAAGATG	112 pb
	ORFVI-R	GTAGCTTCTCCTGTAAGAACAC	
ORF VIII	ORFVIII-F	TATGGATGTACGCCTGAAATG	110 pb
	ORFVIII-R	TCACGGATAGAAAGACCTGTA	
ORF X	ORFX-F	AATGTGTGGATTGCTGATAAAG	120 pb
	ORFX-R	CCCTGCGGCTAAACAATA	
ORF XVIII	ORFXVIII-F	TCGCAAGCTTGTCACTTT	110 pb
	ORFXVIII-R	CCTCGACCAATACATTACCTG	
Bacteroides 16S	16S-F	GTAGCGGTGAAATGCTTAGA	102 pb
	16S-R	CCTGTTTGATACCCACACTT	

Secuencia de la sonda de hibridación empleada en esta tesis para la detección de fagos crAssBcn, basada en la secuencia del fago ΦcrAss001

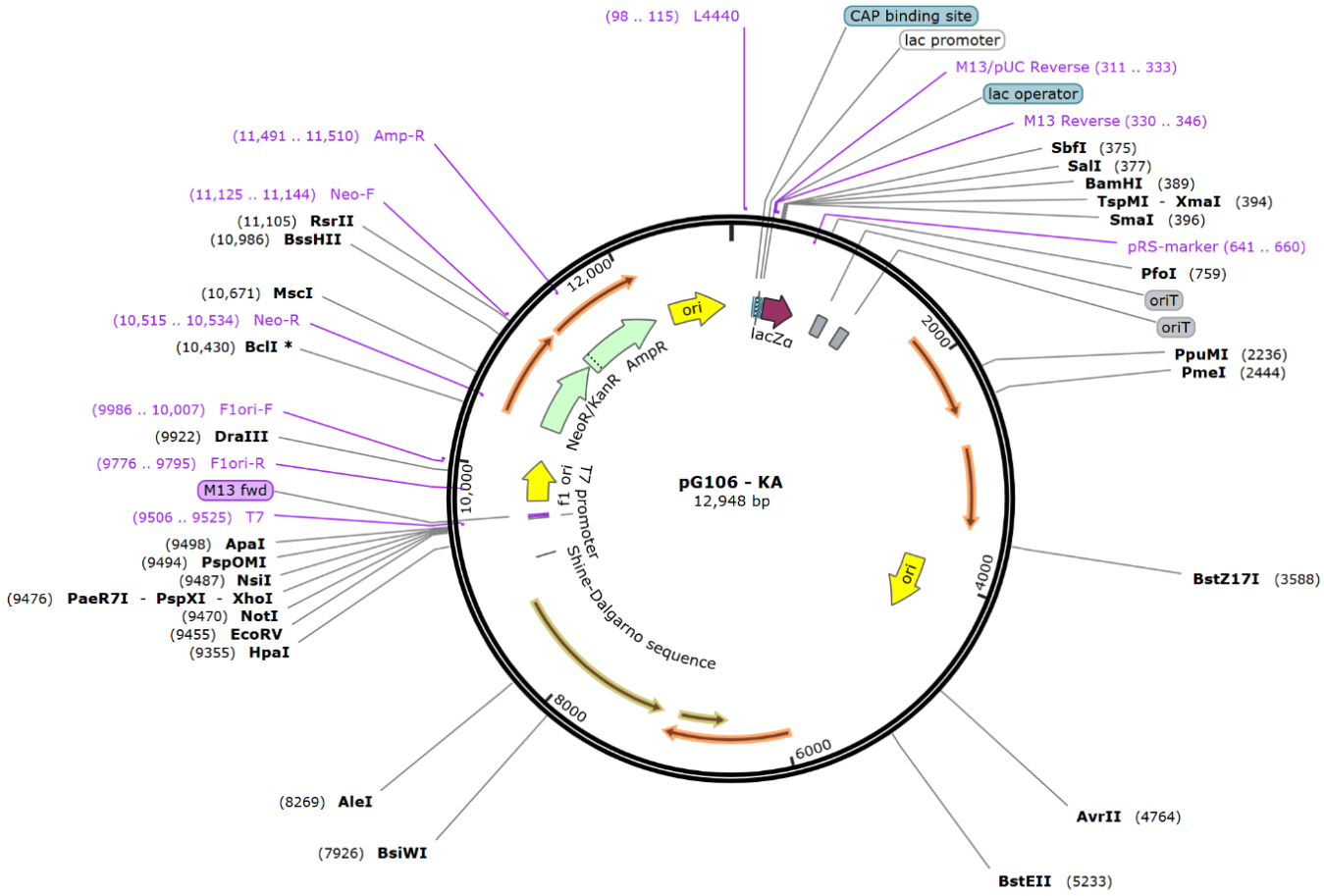
5’ -ACCTGCTTCTACACTTTCCTTAGATGAACTAATATCTAACCCAGCTCTAT- 3’

Secuencia del GBlock empleado para la preparación de los estándares de qPCR para los fagos crAssBcn

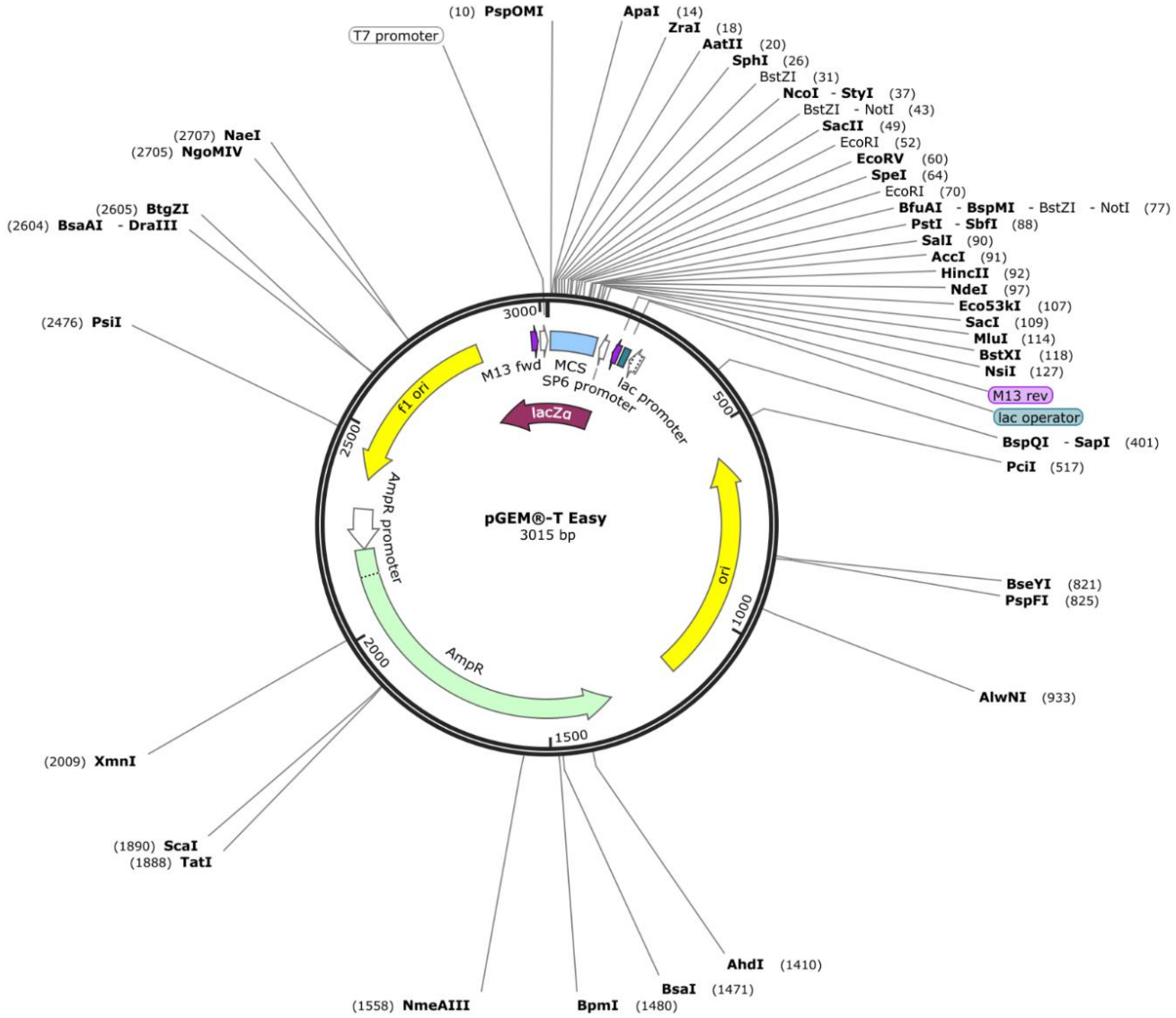
5’-
TGTACAAACAATATTCTAATGACTTGGCTGCACAAGCTGAGTCATTAGCTAAACAAGGACTTACTCCTGCAAGTAGACAGGGATTGATTGAT
ATGAAAAGAAGATACTCTTCTGAGATTGTTCCTATAGAACAGGCTTATAAGAGAAGACAAGAGTTAGTAGATGAACAAAGGAAATTACAGGC
TCAGGATAGTACACTGTTATTTGATAGACCTGCTTCTACACTTTCCTTAGATGAACTAATATCTAACCCAGCTCTATCACCACAATCCTATTCT
GGAGCACTATTATCCAAACAAGTGGGTACTGCTGCACAGAATTTAGCTAAGGAAGTAAGAGAGAACCCAAGAAAGTGGAGAACAATCTTAG
GTAATCAATACTATGAAACCATCATGCAGAAGGGATTCAGACCTGAGGAAATTATGCAGGCTGTACAGAATAATCCTGAGGCTTCTCCTATA
CTTCAAGGTATTGTGGAAGATGCAGTAGG -3’

8.5. Anexo 5: Plásmidos utilizados

pG106-KA

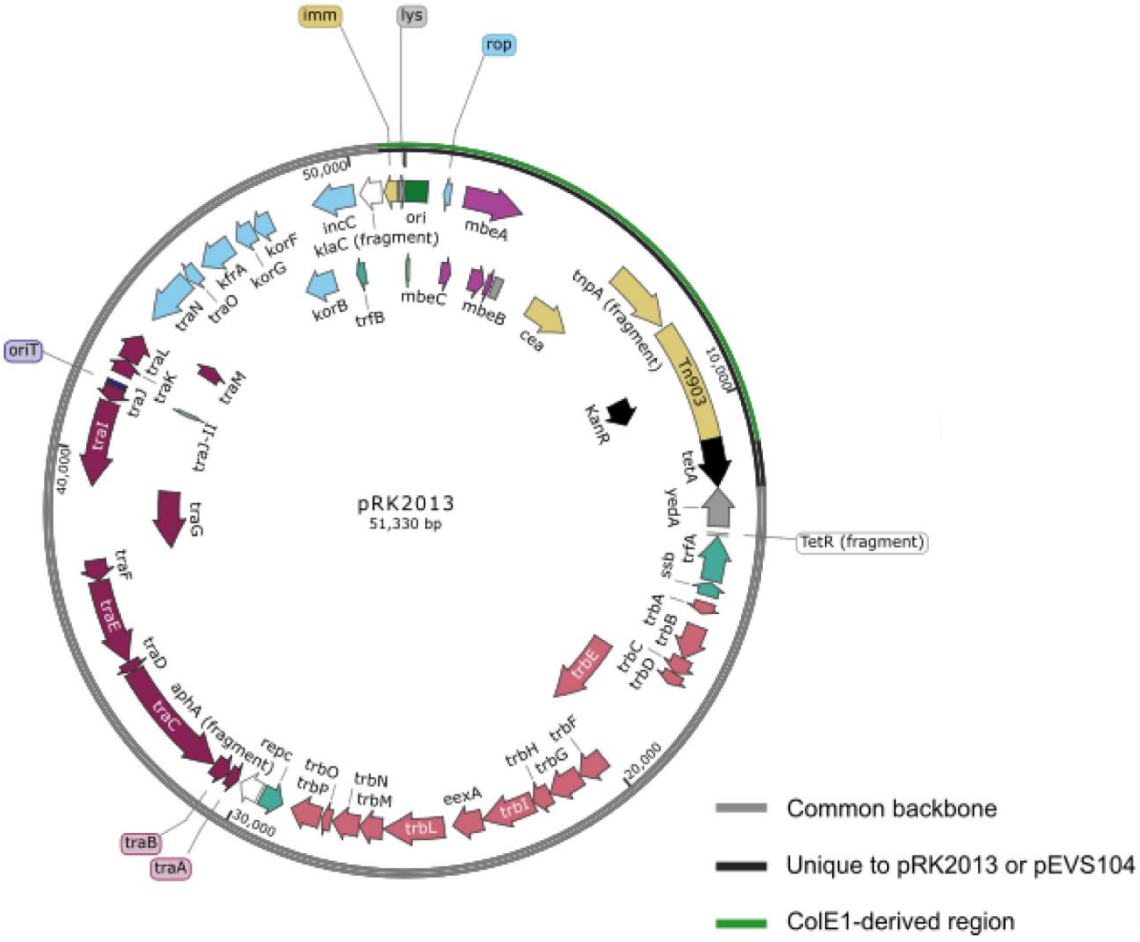


p-GEM® t-EASY


pGEM®-T Easy
3015 bp
T7 promoter
M13 fwd
MCS
SP6 promoter
lac promoter
lacZα
f1 ori
AmpR promoter
AmpR
ori
M13 rev
lac operator
3000
500
1000
1500
2000
2500
(10) PspOMI
ApaI (14)
ZraI (18)
AatII (20)
SphI (26)
BstZI (31)
NcoI - StyI (37)
BstZI - NotI (43)
SacII (49)
EcoRI (52)
EcoRV (60)
SpeI (64)
EcoRI (70)
BfuAI - BspMI - BstZI - NotI (77)
PstI - SbfI (88)
SalI (90)
AccI (91)
HincII (92)
NdeI (97)
Eco53kI (107)
SacI (109)
MluI (114)
BstXI (118)
NsiI (127)
BspQI - SapI (401)
PciI (517)
BseYI (821)
PspFI (825)
AlwNI (933)
AhdI (1410)
BsaI (1471)
BpmI (1480)
(1558) NmeAIII
(2009) XmnI
(1890) ScaI
(1888) TatI
(2476) PsiI
(2605) BtgZI
(2604) BsaAI - DraIII
(2707) NaeI
(2705) NgoMIV

Plásmido pRK2073 de la cepa helper (*E. coli* HB101)

Plásmido derivado de pRK2013, que facilita la conjugación (Ditta et al., 1985).



Referencias de los Anexos

- Béchon, N., Mihajlovic, J., Vendrell-Fernández, S., Chain, F., Langella, P., Beloin, C., & Ghigo, J.-M. (2020). Capsular polysaccharide cross-regulation modulates Bacteroides thetaiotaomicron biofilm formation. MBio, 11(3), 10–1128. https://doi.org/10.1128/mBio

- Ditta, G., Schmidhauser, T., Yakobson, E., Lu, P., Liang, X.-W., Finlay, D. R., Guiney, D., & Helinsiu, D. R. (1985). Plasmids related to the broad host range vector, pRK290, useful for gene cloning and for monitoring gene expression. PLASMID, 13, 149–153.

- García-Aljaro, C., Ballesté, E., Muniesa, M., & Jofre, J. (2017). Determination of crAssphage in water samples and applicability for tracking human faecal pollution. Microbial Biotechnology, 10(6), 1775–1780. https://doi.org/10.1111/1751-7915.12841

- Nizetic, D., Drmanac, R., & Lehrach, H. (1991). An improved bacterial colony lysis procedure enables direct DNA hybridisation using short (10, 11 bases) oligonucleotides to cosmids. Nucleic Acids Research, 19, 182.

- Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G. T., Blanch, A. R., Lucena, F., Jofre, J., & Muniesa, M. (2005). Method for isolation of Bacteroides bacteriophage host strains suitable for tracking sources of fecal pollution in water. Applied and Environmental Microbiology, 71(9), 5659–5662. https://doi.org/10.1128/AEM.71.9.5659-5662.2005

- Sagar Aryal. (2022). Capsule staining. https://microbiologyinfo.com/capsule-staining-principle-reagents-procedure-and-result/. https://microbiologyinfo.com/capsule-staining-principle-reagents-procedure-and-result/

- Sambrook, J., & Russell, D. W. (2001). Molecular cloning: a laboratory manual. Cold Spring Harb. Lab. Press. Cold Spring Harb. NY 999. www.cshlpress.org

- Thein, M., Sauer, G., Paramasivam, N., Grin, I., & Linke, D. (2010). Efficient subfractionation of gram-negative bacteria for proteomics studies. Journal of Proteome Research, 9(12), 6135–6147. https://doi.org/10.1021/PR1002438