

Tesis Doctoral

Estudio metagenómico de comunidades microbianas de barros activados: interacciones tróficas y estrategias de crecimiento bacterianas

Pérez, María Victoria

2017

Este documento forma parte de las colecciones digitales de la Biblioteca Central Dr. Luis Federico Leloir, disponible en bibliotecadigital.exactas.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the digital collection of the Central Library Dr. Luis Federico Leloir, available in bibliotecadigital.exactas.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Pérez, María Victoria. (2017). Estudio metagenómico de comunidades microbianas de barros activados: interacciones tróficas y estrategias de crecimiento bacterianas. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

https://hdl.handle.net/20.500.12110/tesis_n6452_Perez

Cita tipo Chicago:

Pérez, María Victoria. "Estudio metagenómico de comunidades microbianas de barros activados: interacciones tróficas y estrategias de crecimiento bacterianas". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2017.

https://hdl.handle.net/20.500.12110/tesis_n6452_Perez

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

**Estudio metagenómico de comunidades microbianas de barros
activados: interacciones tróficas y estrategias de crecimiento
bacterianas.**

Tesis presentada para optar al título de Doctor de la Universidad de Buenos
Aires en el área de Ciencias Biológicas

María Victoria Pérez

Director de Tesis: Dr. Leonardo Erijman

Directora Asistente: Dra. Eva M. Figuerola

Consejero de Estudios: Dr. Pablo Cerdán

Lugar de trabajo: Instituto de Investigaciones en Ingeniería Genética y Biología
Molecular "Dr. Héctor N. Torres" (INGEBI – CONICET)

Buenos Aires, noviembre de 2017

Fecha de defensa: 4 de diciembre de 2017

ÍNDICE

RESUMEN	4
ABSTRACT	6
AGRADECIMIENTOS	8
INTRODUCCIÓN	10
Tratamiento de efluentes domiciliarios	10
Sistema de barros activados	12
Ecología microbiana de barros activados.....	15
Comunidades microbianas en sistemas de tratamiento de efluentes	16
Interacciones multitróficas en barros activados.....	20
Perspectivas del trabajo de tesis	23
HIPÓTESIS Y OBJETIVOS.....	25
MATERIALES Y MÉTODOS	26
Sistema experimental	26
Descripción de la planta de tratamiento de efluentes.....	26
Muestreo y conservación de las muestras	27
Características operacionales del período de muestreo	29
Microscopía y cuantificación de organismos eucariotas.....	34
Extracción de ADN.....	34
Extracción de ADN total.....	34
Extracción de ADN viral.....	35
Ensayos preliminares.....	36
Electroforesis en gel con gradiente desnaturizante (DGGE)	36
Búsqueda de hospederos de fagos en cultivo	39
Pirosecuenciación de regiones hipervariables del gen ARNr 16S.....	39
Hibridación fluorescente in situ (FISH).....	40

Secuenciación metagenómica de la serie temporal	42
Diseño experimental	42
Secuenciación de amplicones	43
Secuenciación al azar.....	45
RESULTADOS	53
Análisis preliminares de diversidad y dinámica microbianas	53
Diversidad y dinámica microbianas	62
Estructura de la comunidad eucariota	62
Estructura de la comunidad bacteriana	66
Relación especie-tiempo.....	67
Influencia de factores abióticos en la estructura de las comunidades	68
Análisis metagenómico	73
Funciones de los metagenomas	73
Reconstrucción de genomas a partir de metagenomas	75
Bases genómicas de estrategias de crecimiento bacterianas	78
Número de copias del operón ribosomal rrn	78
Índice de replicación (iRep)	79
Análisis del metaviroma	83
Estructura de la comunidad viral.....	83
Interacción fago-hospedero	84
Comparación de perfiles de abundancia.....	84
Búsqueda de CRISPRs	85
Búsqueda de similitud de secuencias	89
DISCUSIÓN	96
CONCLUSIONES	107
GLOSARIO	109
ANEXO A.	111

ANEXO B.....	118
ANEXO C.....	120
ANEXO D.....	128
BIBLIOGRAFÍA	134

RESUMEN

Estudio metagenómico de comunidades microbianas de barros activados: Interacciones tróficas y estrategias de crecimiento bacterianas

El sistema de barros activados es uno de los modelos más estudiados en ecología microbiana. En las últimas décadas se han realizado grandes avances en el conocimiento de la composición de las complejas poblaciones microbianas participantes en dicho proceso. Si bien se ha visto que las variables ambientales y operativas son factores críticos en la estructuración de las comunidades bacterianas, aún son escasos los estudios sobre los diferentes niveles tróficos presentes en el sistema. El objetivo de este trabajo es comprender la influencia de factores bióticos y abióticos en la estructura y dinámica de los diferentes niveles tróficos de comunidades microbianas de un sistema de tratamiento de efluentes cloacales. Para ello, se estudiaron los tres niveles tróficos principales de los barros activados: bacterias, eucariotas y bacteriófagos, en una planta de tratamiento de efluentes domiciliarios de Buenos Aires. El período evaluado comprendió tres años en los que la planta operó bajo un gradiente creciente de tiempo de residencia de sólidos (TRS) y atravesó una etapa de fluctuaciones operacionales. Los microorganismos eucariotas se analizaron por microscopía mientras que la composición bacteriana se estudió por secuenciación de amplicones de la región hipervariable V3-V4 del gen ARNr 16S y secuenciación al azar sobre el ADN metagenómico. La presencia y abundancia de bacteriófagos se siguió por secuenciación al azar del ADN presente en el sobrenadante (metaviroma).

Luego del co-ensamblado de los 60 metagenomas, de la serie temporal, se utilizó una metodología de *binning* basada principalmente en cobertura diferencial con la que se reconstruyeron 257 genomas. Los perfiles de co-ocurrencia de especies bacterianas resultaron altamente dependientes del régimen operativo, donde el período de fluctuaciones y de bajo TRS, fue dominado por filos mayormente compuestos por estrategias-r (copiótrofos); contrastando con el período de alto TRS, en el que predominaron los estrategias-K (oligótrofos). La partición de poblaciones bacterianas en base a

diferencias en las estrategias de crecimiento fue confirmada por la determinación del número de operones ARN ribosomales (*rnn*) estimados a partir de los genomas ensamblados.

Por su parte, el índice biótico de barros (SBI), basado en la abundancia y diversidad de protozoos, correlacionó positivamente con la abundancia total de bacterias oligotróficas, apoyando el uso del SBI como indicador para el control de calidad del tratamiento y reflejando la modulación por depredadores eucariotas.

Finalmente, los cambios en la composición de fagos reflejaron un patrón similar al revelado por las poblaciones bacterianas. La reducción en la abundancia relativa de virus de la familia *Podoviridae* y el correspondiente aumento en *Siphoviridae* que acompañan el aumento en TRS es consistente con el hecho de que los *Podoviridae* infectan principalmente *Proteobacteria*, mientras que sifovirus pueden ser lisogénicos por varias generaciones, un comportamiento que ha sido asimilado a estrategias K.

Las interacciones entre los fagos y las poblaciones bacterianas, detectadas a través de la búsqueda de elementos CRISPR y por similitud de secuencias, permitieron la identificación de mecanismos de interacción tipo Lotka-Volterra depredador-presa, así como de modelos alternativos de interacciones múltiples, en los que un fago infecta a más de un hospedero y/o un hospedero es infectado por más de un tipo de fago.

Se concluye que los parámetros de proceso determinan la composición de poblaciones bacterianas en base a sus estrategias de crecimiento, mientras que sobre esta estructuración, determinada por factores abióticos, se superpone una modulación fina mediada por depredadores.

Palabras clave: barros activados; ecología microbiana; metagenómica; estrategias de crecimiento bacterianas; interacciones tróficas.

ABSTRACT

Metagenomic study of activated sludge microbial communities: Trophic interactions and bacterial growth strategies

Activated sludge system is one of the most studied models in microbial ecology. In the last decades, large progress has been made in the knowledge of the composition of these complex microbial populations. Although it has been seen that environmental and operative variables are critical factors on the bacterial community structuring, there are still few studies regarding to different trophic levels in this system. The aim of this work is to understand the influence of biotic and abiotic factors on the structure and dynamics of different trophic levels of microbial communities of a sewage treatment system. For this, the three main trophic levels in activated sludge: bacteria, eukaryotes and phages, were studied in a wastewater treatment plant of Buenos Aires. The evaluated period comprise three years in which the plant operates under an increasing gradient of sludge retention time (SRT) and went through a stage of operational fluctuations. Eukaryotic organisms were analyzed by microscopy while bacterial composition was studied by amplicon sequencing of the 16S rRNA gene V3-V4 hypervariable region and metagenomic DNA shotgun sequencing. Phages presence and abundance were followed by shotgun sequencing of DNA present in the supernatant (metavirome).

After co-assembling the 60 metagenomes, from the time series, we applied a binning methodology, based mainly on the differential coverage, from which 257 genomes were reconstructed. The co-occurrence profiles of bacterial species resulted highly dependent on the operational regime, where the disturbed period with low SRT was dominated for phyla mostly composed of r-strategist (copiotrophs); contrasting with the high SRT period, where K-strategist (oligotrophs) predominated. The separation of bacterial populations based on differences in their growth strategies was confirmed by determining the ribosomal RNA (*rrn*) operon number, estimated from the assembled genomes.

Meanwhile, the sludge biotic index (SBI), based on protozoa abundance and diversity, correlated positively with total abundance of oligotroph bacteria, supporting the use of the SBI as an indicator of treatment quality control and displaying a modulation by Eukaryotic predators.

Finally, changes in phage composition showed a similar pattern than observed in bacterial populations. Reduction of the relative abundance of *Podoviridae* family viruses and the corresponding increase on *Siphoviridae* along with the increase on SRT is consistent with the fact that *Podoviridae* mainly infect *Proteobacteria*, while siphoviruses could be lysogenic for several generations, a feature assimilated to K-strategists.

Phage-bacteria interactions, detected by searching for CRISPR elements and sequence homologies, allowed to identify prey-predator interactions mechanisms type Lotka-Volterra, as well as multiple interactions alternative models, in which a phage infects more than one host and/or a host is infected by more than one type of phage.

We conclude that bacterial population composition based on their growth strategies is determined by process parameters, while this structuring, determined by abiotic factors, is superposed by a fine modulation mediated by predators.

Keywords: activated sludge; bacterial growth strategies; metagenomics; microbial ecology; trophic interactions.

AGRADECIMIENTOS

A Leonardo Erijman, mi director de tesis, por haberme recibido en su laboratorio, por guiarme excelentemente y siempre con predisposición, y sobre todo por el gran aporte que hizo a mi formación.

A Eva Figuerola, mi directora asistente, por su ayuda permanente y por haberme enseñado con ejemplos y consejos.

A Enrique Calderón, por haber confiado en mí para llevar a cabo este proyecto y sobre todo por transmitirme sus conocimientos y su gran vocación.

A Agua y Saneamientos Argentinos S.A. (AySA) y al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) quienes gracias a la ejecución de un convenio de partes hicieron posible este trabajo. A la Dirección General y a la Dirección de Recursos Humanos de AySA, en especial a su Director Luis Aversano, a Omar Fioravanti y a todo el resto del equipo de AySA que ha colaborado para que este convenio se haya concretado y haya tenido continuidad.

A todos los que formaron parte del grupo del laboratorio de Ecología Microbiana del INGEBI conmigo, por haber sido unos compañeros de trabajo maravillosos y haber estado siempre dispuestos a ayudarme: Daniela Vega, Carol Davies, Florencia Tesoriero, Silvina Rosa, Joaquín Ayarza, Tatiana Spatola Rossi. Particularmente, a Leandro Guerrero, quien colaboró intensamente con el procesamiento de los datos metagenómicos y con los análisis bioinformáticos más sofisticados, a Esteban Orellana, que también colaboró con sus conocimientos en bioinformática directamente y también con muchos consejos diarios, y a Federico Ibarbalz, con quien trabajé a la par durante la mayor cantidad del tiempo en el laboratorio y a quien le debo una gran parte de lo aprendido en estos años.

A Jorge Muschietti, a Agustina Mazzella y a los integrantes de sus equipos, con quienes compartí, junto a mi grupo, muy felizmente mis días dentro y fuera del INGEBI: Mari, Diego, Denise, Sofí, Ana, Martín, Luciana, Julián, Juli, Lau y Regi.

A la Dirección de Saneamiento de AySA y a los jefes de Planta Norte que ejercieron durante los años en los que se llevó a cabo este trabajo, por permitirme realizar el muestreo. En especial a Nahuel Maldonado y a Ángela Cicero, por su enorme colaboración y por seguir confiando en el aporte de este trabajo. A todo el resto de mis compañeros del laboratorio de la planta, que durante todos estos años, una vez por semana me incorporaron como uno más: Leo C., Marce P., Martín, Marcelo D., Iara, Deba, Mauri, Leo V. y Juan, que además se ocuparon de acompañarme a cada muestreo y realizar los análisis diariamente. A todos, por su enorme colaboración, su energía, compromiso y entusiasmo. Y a todo el resto del personal de la planta por recibirme e incluirme como compañera de trabajo.

A la Dirección de Apoyo Logístico de AySA que facilitó mis traslados a la planta, y al Laboratorio Central de AySA por realizar los análisis mensuales de la planta.

A los directivos y a todo el personal del Instituto de Investigaciones en Ingeniería Genética y Biología Molecular (INGEBI) por haber permitido el desarrollo de mi tesis allí y por hacer que el INGEBI sea un cálido lugar de trabajo.

A Cecilia Quiroga, a Mariana Piuri, y su grupo, por sus colaboraciones en temas de bacteriófagos.

A los jurados propuestos, por haber aceptado evaluar mi tesis, y a los que fueron asignados por tomarse el tiempo de hacerlo.

A Cynthia Corroto y a Corina Nigro, por escucharme, aconsejarme, y acompañarme recorriendo este camino desde el principio.

A Jean-Mi, por su paciencia, su interés por mi trabajo y sobre todo por llenar mi vida de amor.

A mis amigas y a mi familia. Sobre todo, a los fuertes pilares de mi vida, a mi papá y a mi hermana, por acompañarme y apoyarme en cada paso, por escucharme, por sostenerme y hacer mi vida feliz.

Y al mejor ejemplo de mi vida, que se que me cuida y me acompaña siempre, mi mamá.

INTRODUCCIÓN

Tratamiento de efluentes domiciliarios

El desarrollo de toda sociedad se encuentra estrechamente ligado a la utilización de agua como recurso. Este uso, ya sea industrial o domiciliario, trae aparejado la generación de aguas residuales que deben ser alejadas de la vida humana para evitar enfermedades, y tratadas previamente a ser vertidas a los cuerpos receptores a fin de evitar la contaminación y/o eutrofización de los mismos.

Los efluentes cloacales pueden contener microorganismos patógenos fecales provenientes de heces contaminadas. Si estos efluentes no se alejan ni se tratan, las consecuencias para la salud pública en cuanto a transmisión de enfermedades hídricas serían aberrantes, lo cual sigue sucediendo aun hoy, según la Organización Mundial de la Salud, en muchos sitios en el mundo que no cuentan con sistemas de saneamiento adecuados (sumado a la inaccesibilidad al agua potable y a la falta de higiene). Las aguas residuales domiciliarias también contienen remanentes de alimentos y de otros residuos que son arrojados a las descargas cloacales. Por un lado, estos compuestos orgánicos al ser vertidos a los cuerpos de agua promueven el crecimiento de bacterias heterótrofas aeróbicas que los degradan, pudiendo agotar el oxígeno disponible y amenazando así al resto de la vida acuática. Por otro lado, un enriquecimiento en nutrientes tales como nitrógeno y, particularmente, fósforo puede generar un crecimiento desmedido de algas y cianobacterias, quienes aprovechan de manera eficaz esos nutrientes, fenómeno que se conoce como eutrofización. El tratamiento de efluentes se desarrolló mundialmente a fines del siglo XIX (Henze *et al.*, 2008), como un campo en la ingeniería sanitaria y ambiental, con el objetivo primordial de evitar enfermedades de origen hídrico y también para preservar la calidad de los ambientes acuáticos que reciben los efluentes, afrontando así estos serios problemas ambientales y de salud pública.

En una planta de tratamiento de efluentes (PTE) domiciliarios, el proceso consta de varias etapas bien establecidas en la actualidad, aunque no siempre se encuentren todas presentes:

- Pretratamiento
- Tratamiento primario
- Tratamiento secundario (tratamiento biológico)
- Tratamiento terciario

En la etapa de pretratamiento se eliminan los sólidos voluminosos que llegan con el efluente, así como también las grasas y arenas. Durante el tratamiento primario, que consta generalmente de una etapa de sedimentación, se eliminan los sólidos en suspensión capaces de decantar por gravedad. El tratamiento secundario, consiste en un tratamiento biológico en el cual los microorganismos son los encargados de disminuir la concentración de materia orgánica suspendida o disuelta en el efluente. El tratamiento terciario puede incorporarse como etapa final a fin de incrementar la calidad del efluente, desinfectando o removiendo nutrientes tales como nitrógeno y fósforo (aunque esto último también puede ser incorporado en el tratamiento secundario).

Los sólidos suspendidos, de naturaleza orgánica e inorgánica, son las impurezas más perceptibles en los efluentes y es por eso que los primeros sistemas de tratamiento, introducidos en el siglo XIX, fueron diseñados como unidades de separación por sedimentación por gravedad (tratamiento primario). Es evidente que esto puede tratar sólo parcialmente las aguas residuales debido a que una gran proporción de la materia orgánica de las mismas se encuentra en estado coloidal o disuelto y por lo tanto no es sedimentable. El tratamiento secundario fue introducido hacia las primeras décadas del siglo XX con el objetivo de mejorar la eficiencia de remoción de la materia orgánica en las plantas de tratamiento de efluentes (Haandel and Lubbe, 2012). El tratamiento secundario se caracteriza por el uso de procesos biológicos. El más utilizado es el sistema de barros activados, aunque también existen otros tipos de tratamiento biológico de efluentes como las lagunas, los lechos percoladores, y los digestores anaeróbicos.

La eficiencia del tratamiento de efluentes se puede estimar como la reducción de la materia orgánica medida en el laboratorio como demanda biológica de oxígeno (DBO). La DBO es una medida del consumo de oxígeno requerido por oxidación microbiana. La medida estándar es la DBO₅ (DBO a 5 días), que expresa el oxígeno consumido por la muestra durante 5 días de incubación a 20°C. Los efluentes domiciliarios suelen contener una DBO₅ de aproximadamente 250 a 400 mg/l. Típicamente la combinación de los tratamientos primario y secundario reduce la DBO original del efluente en un 90-95%. Las normas para desagües cloacales en Argentina, según el Marco Regulatorio incluido en la Ley 26221, indican que el líquido tratado de las planta depuradoras que cuenten con tratamiento primario y secundario, que descarguen a un cuerpo receptor, debe tener una DBO₅ por debajo de los 30 mg/l, entre otros parámetros.

Sistema de barros activados

El proceso de barros activados fue inventado en 1913 en el Reino Unido por dos ingenieros de la Universidad de Manchester, Edward Arden y William Lockett, mientras realizaban investigaciones para Manchester Corporation (Henze *et al.*, 2008). Sus experimentos sobre el tratamiento de aguas residuales en un reactor de drenaje y llenado produjeron un efluente tratado con baja demanda de oxígeno (Arden and Lockett, 1914). El proceso consistía en una secuencia de ciclos de llenado con efluente crudo y drenaje de líquido tratado, manteniendo los sólidos formados en el reactor. Los investigadores mostraron que a medida que aumentaba la cantidad de sólido formado, el tiempo requerido para lograr un efluente tratado de alta calidad disminuía. En términos de proceso lograban concentrar los microorganismos biológicos, disociando el tiempo de retención de líquidos del tiempo de retención de sólidos (TRS). Como referencia, y a falta de mejor nombre, los autores denominaron barro activado al sólido depositado que lograba la oxidación completa del agua residual (Arden and Lockett, 1914) . En 1914 se probó este proceso bajo flujo continuo en una planta piloto y al obtenerse resultados tan satisfactorios se creó la primera planta de barros activados a gran escala en la ciudad de Worcester (Seviour and Nielsen, 2010). Actualmente, con más de

100 años desde su implementación, el proceso, que mantiene el nombre de barros activados, es aún el sistema de tratamiento biológico de efluentes cloacales e industriales más utilizado en el mundo.

El sistema de configuración básico de barros activados (Figura 1) consta de dos etapas que se llevan a cabo en compartimentos diferentes, un reactor biológico aeróbico y un sedimentador. Cabe destacar que existen otras configuraciones con diseños más sofisticados basadas en los objetivos de remoción de nutrientes, como nitrógeno y fósforo.

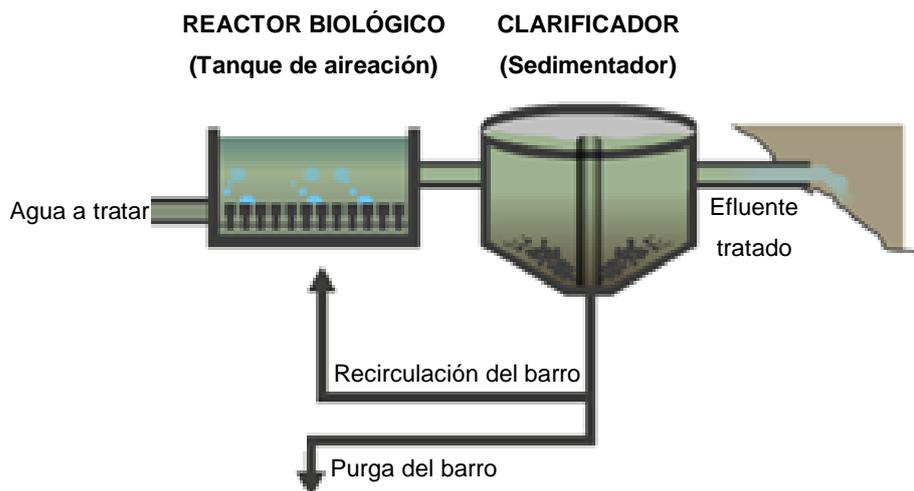


Figura 1. Esquema de configuración básica del sistema de barros activados

El reactor biológico (cámara o tanque de aireación) recibe el líquido a tratar que proviene ya sea de un tratamiento primario, un pretratamiento o directamente líquido crudo. Allí crecen en condiciones aeróbicas, debido a la inyección de aire, microorganismos heterótrofos que oxidan la materia orgánica biodegradable transformándola en más biomasa, CO_2 y H_2O , que a su vez son capaces de formar un agregado microbiano llamado floc. En un segundo compartimento, sedimentador secundario o clarificador, sedimenta el floc del líquido tratado por gravedad. El efluente ya tratado se descarga a un cuerpo receptor de agua.

La mezcla de aguas residuales crudas con la masa biológica que se encuentra dentro del tanque de aireación se conoce como licor mezcla. Típicamente, la concentración de sólidos suspendidos del licor mezcla (SSLM) en un reactor en operación oscila entre 2 y 6 g/l. Esta concentración de sólidos se mantiene en

el sistema mediante la recirculación de parte del barro sedimentado en el clarificador, mientras que el excedente es eliminado del sistema (purga). Debido a que el barro sedimenta y es recirculado, el tiempo de retención hidráulico (tiempo en el que el líquido permanece en el sistema) se diferencia del tiempo de residencia de sólidos (TRS), también llamado edad de barros o tiempo de retención celular, que es el tiempo promedio de permanencia de la biomasa en el sistema y es calculado a partir de la siguiente ecuación:

$$TRS = \frac{V \cdot SSLM}{Q_e \cdot SSe + Q_p \cdot SSp}$$

donde V es el volumen del reactor; SSLM, los sólidos suspendidos totales del licor mezcla; Q_e, el caudal del efluente; SSe, los sólidos suspendidos totales en el efluente; Q_p, el caudal de purga de barros y SSp, los sólidos suspendidos totales en los barros purgados.

El éxito del funcionamiento de los sistemas de barros activados está dado por la capacidad de la comunidad microbiana tanto para metabolizar la materia orgánica del líquido como para formar un agregado sedimentable.

El control operativo se realiza manipulando la cantidad de barro que se elimina del sistema (que define indirectamente la cantidad de barro a recircular y por consiguiente la cantidad de sólidos en el reactor), y el aire inyectado al sistema para mantener un cierto nivel de oxígeno disuelto en los reactores (aproximadamente de 2 a 4 mg/l). La relación entre la carga orgánica del líquido a tratar, que se puede medir como la cantidad de DBO o DQO (demanda química de oxígeno), y los sólidos presentes en el tanque de aireación determina lo que se conoce como relación F/M (por *Food* y *Microbes* en inglés).

$$\frac{F}{M} = \frac{DBO \times Caudal}{SSLM \times V}$$

Es decir que las variables operativas se regulan en definitiva en base a relación F/M, considerando también un suministro apropiado de oxígeno para mantener la aerobiosis. El TRS es un parámetro de diseño que determina las dimensiones de la planta y su capacidad. Su variación es posible dentro de un

rango determinado por la dimensión de la planta y la cantidad del efluente tratado.

Ecología microbiana de barros activados

En un mundo con una población en constante crecimiento, el tratamiento de aguas residuales es un proceso de gran importancia. Una de las principales funciones de la biotecnología ambiental es comprender a las comunidades microbianas que proporcionan servicios esenciales a la sociedad (Erijman *et al.*, 2011). La microbiología del tratamiento de efluentes es uno de los modelos más empleados en ecología microbiana (junto con el microbioma del intestino, de suelos, de diferentes ambientes acuáticos, entre otros ecosistemas) por lo que su estudio también es de gran utilidad para avanzar hacia una profunda comprensión de comunidades microbianas complejas. De hecho, la microbiología de aguas residuales es hoy en día una subdisciplina importante en ecología microbiana (Daims *et al.*, 2006b).

El licor mezcla de los reactores aeróbicos del sistema de barros activados consiste en comunidades complejas de microorganismos, constituidas principalmente por bacterias, principales responsables de la degradación, que comprenden aproximadamente el 95% del total de la biomasa, virus, protozoos y metazoos. Estos microorganismos se encuentran agregados formando el floc biológico (Figura 2).

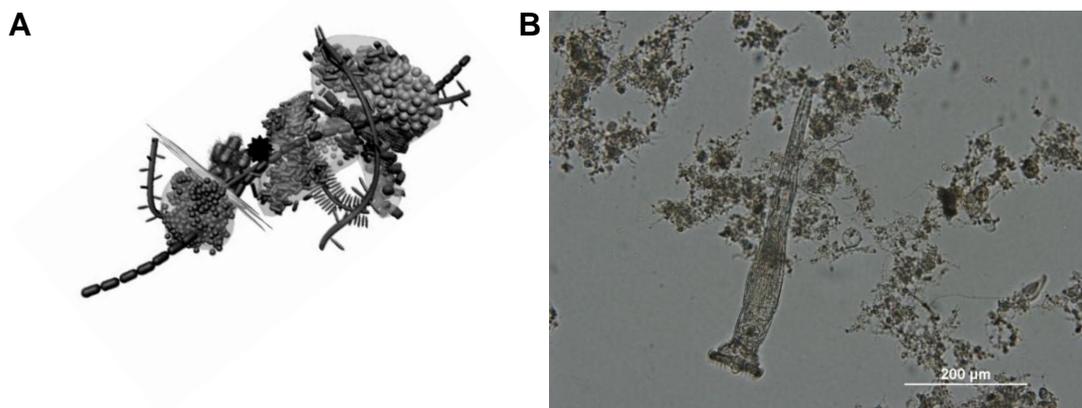


Figura 2. Floc biológico de barros activados. **A)** Diagrama representativo de la estructura de un floc (Microbial Ecology of Activated Sludge, 2010) **B)** Micrografía de un floc (Magnificación 200X)

El floc biológico de los barros activados es una estructura heterogénea compleja, donde los microorganismos están embebidos en sustancias poliméricas extracelulares (SPE) (Zartarian *et al.*, 1994). La composición de la matriz polimérica es altamente variable y aunque no se encuentra caracterizada completamente se ha visto que está formada por proteínas, polisacáridos, y productos de secreción celular que derivan de los propios microorganismos (Sutherland, 2001). Muchas bacterias se pueden encontrar como microcolonias unidas a la estructura del floc, como por ejemplo bacterias nitrificantes y organismos acumuladores de fosfatos (PAO, por sus siglas en inglés). Otras bacterias están presentes como células individuales, y algunas formando filamentos (Seviour and Nielsen, 2010). Los organismos eucariotas se ubican en el floc según sus formas de vida: ciertos protozoos se encuentran anclados al floc biológico (como los ciliados sésiles), otros se arrastran sobre el floc (como los ciliados rastreros) y otros se encuentran libres en el licor mezcla (como los ciliados de nado libre, flagelados, ameboides y algunos metazoos).

Comunidades microbianas en sistemas de tratamiento de efluentes

El microscopio ha sido la primera herramienta que utilizaron los microbiólogos para observar los flóculos y estudiar la biología de los sistemas de barros activados. La observación al microscopio permite obtener detalles de la estructura microscópica del floc, y realizar la cuantificación de bacterias filamentosas. Asimismo, los microorganismos eucariotas, al ser fácilmente observables al microscopio, fueron históricamente los primeros microorganismos en ser estudiados en muestras de barros activados, y la sucesión de protozoos y metazoos sigue siendo utilizada como indicador de funcionamiento del proceso desde sus inicios (Curds and Cockburn, 1970).

No obstante, se ha mencionado que son las bacterias las principales responsables del éxito del tratamiento y las más abundantes en el sistema. En las últimas décadas, las técnicas moleculares han permitido identificar actores bacterianos claves en el tratamiento (por ejemplo en la remoción de nutrientes, y en la formación de espuma y abultamiento de barros) y han proporcionado

considerable información sobre la estructura, diversidad, función bacteriana (Erijman *et al.*, 2011; Daims *et al.*, 2006b).

La eliminación de nutrientes, como el nitrógeno y fósforo, resulta importante para evitar el deterioro de la calidad del cuerpo de agua que recibe los efluentes. Durante la eliminación biológica del fósforo se aprovecha la capacidad de ciertas bacterias para acumular polifosfato intracelularmente. Estos organismos se seleccionan durante fases secuenciales anaeróbico-aeróbicas en lo que se conoce como eliminación biológica de fósforo aumentada o EBPR, por sus siglas en inglés (Blackall *et al.*, 2002). Los actores claves en este proceso son los organismos acumuladores de polifosfato (PAO) y los organismos acumuladores de glucógeno (GAO) quienes compiten con los PAOs por los recursos a expensas de la eficiencia de eliminación de fósforo (Stokholm-Bjerregaard *et al.*, 2017). Se han realizado numerosos estudios a fin de identificar y analizar la ecofisiología de estos tipos de organismos y se han hallado varios candidatos para cada uno de ellos. Algunos ejemplos de géneros propuestos de PAO son la betaproteobacteria 'Candidatus *Accumulibacter phosphatis*' (Hesselmann *et al.*, 1999) y organismos relacionados con *Rhodocyclus* (Kong *et al.*, 2004; Crocetti *et al.*, 2000) y en el caso de los GAO 'Candidatus *Competibacter phosphatis*' (Blackall *et al.*, 2002).

La eliminación de nitrógeno de las aguas residuales comprende tres procesos principales: nitrificación, desnitrificación y oxidación anaeróbica de amonio (anammox). Los organismos modelo de estos procesos han sido estudiados extensivamente (Daims *et al.*, 2001; Maixner *et al.*, 2006), lo que permitió revelar los principales responsables de la nitrificación y desnitrificación. En cuanto a la oxidación anaeróbica de amonio, desde que se hallaron los organismos pertenecientes al filo *Planctomycetes* capaces de llevar a cabo este proceso (Strous *et al.*, 1999) se ha logrado implementar el proceso para su aplicación a escala en el tratamiento de efluentes. Fue en el año 2002 cuando se construyó el primer reactor a gran escala que utiliza el proceso anammox (Abma *et al.*, 2007). Por su parte, la nitrificación, ha sido considerada siempre como un proceso en dos etapas en el que el amoníaco se oxida primero a nitrito por bacterias oxidantes del amoníaco (AOB) y/o arqueas (AOA), y posteriormente a nitrato por bacterias oxidantes de nitritos (NOB). La

oxidación completa de amoníaco (Commamox) a nitratos llevada a cabo por un sólo organismo estaba postulada, pero aun no se conocían organismos que realizaran este proceso. Recientemente se reportó el hallazgo de dos especies bacterianas completamente nitrificantes del género *Nitrospira*, aisladas de bioreactores (van Kessel *et al.*, 2015). En los genomas de estas especies de *Nitrospira* se encontraron los genes que codifican para todas las enzimas necesarias para la oxidación del amoníaco a nitrato, por vía de nitrito.

La formación de espuma y el abultamiento de barros significan serios problemas de sedimentación en la operación de plantas de barros activados, frecuentemente causados por el exceso de bacterias filamentosas. Como se ha mencionado, las buenas propiedades de sedimentación del barro son cruciales para la separación del líquido tratado y el barro. Los flocs con excesiva cantidad de filamentos son menos densos y no sedimentan adecuadamente, reduciendo la calidad del tratamiento. Los flocs que contienen abundantes filamentos con superficie celular hidrofóbica tienden a unirse a burbujas de aire, formando una espuma estable que flota sobre la superficie del barro. A principios de los años setenta, utilizando métodos tradicionales basados en morfología y tinciones, se describieron más de 30 morfotipos de bacterias filamentosas diferentes en sistemas de barros activados que tratan principalmente aguas residuales municipales (Eikelboom, 1975). Muchas otras bacterias filamentosas se han reportado en plantas de tratamiento industriales (Eikelboom and Geurkink, 2002). La combinación del uso de técnicas moleculares, con estudios enfocados a comprender la identidad, ecofisiología y nichos ecológicos de estos organismos (Nielsen *et al.*, 2000; Kragelund *et al.*, 2006; Levantesi *et al.*, 2006; Kragelund *et al.*, 2007b, 2007a) demostraron que existe una gran diversidad filogenética y una versatilidad fisiológica de estas bacterias.

La aplicación de técnicas moleculares, tales como electroforesis en gel con gradiente desnaturante (DGGE), hibridación fluorescente in situ (FISH) y otras, han proporcionado considerable información sobre la estructura, diversidad y función microbianas. Con el avance de las nuevas técnicas de secuenciación, junto con la disminución de sus costos y el empleo de etiquetas (lo que permite el procesamiento de muchas muestras en simultáneo) fue

posible aumentar la profundidad de la resolución, ayudando a contestar aspectos aún no resueltos de la relación entre la estructura de las comunidades microbianas y la función de los procesos. Numerosos estudios basados en la secuenciación de amplicones de regiones del gen ARNr 16S han mostrado que las comunidades bacterianas de barros activados son altamente diversas y dinámicas, con una gran cantidad de microorganismos desconocidos; y que las bacterias, además de ser los microorganismos mayoritarios, son los principales responsables de la degradación (Sanapareddy *et al.*, 2009; Xia *et al.*, 2010; Kwon *et al.*, 2010; Wells *et al.*, 2011; Ye and Zhang, 2013; Wang *et al.*, 2012; Kim *et al.*, 2013). Estos estudios además mostraron la enorme diversidad bacteriana que se encuentra en estos ambientes, y que existe una consistencia en su composición, y los grupos bacterianos que más comúnmente se encuentran son: *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, *Bacteroidetes*, *Firmicutes*, *Actinobacteria* y *Acidobacteria*. También existen evidencias que muestran que el tipo de efluente es un fuerte determinante en la composición de la comunidad bacteriana (Ibarbalz *et al.*, 2013, 2016) y que la estructura y la diversidad de la comunidad microbiana de barros activados influyen sobre el desempeño y la estabilidad del tratamiento de efluentes (Cyzdik-Kwiatkowska and Zielińska, 2016).

Resulta de suma importancia la comprensión de las influencias ambientales y biológicas sobre la dinámica de las comunidades microbianas. Se ha estudiado extensamente como la estructuración de la comunidad bacteriana de barros activados se produce como resultado de la selección ecológica debida a presiones abióticas ejercidas por las condiciones ambientales y parámetros operativos (TRS, temperatura, pH, OD, características del efluente, etc). En este sentido se han realizado muchos estudios de comunidades bacterianas de muestras de barros activados en diversas plantas de tratamiento para estudiar estas influencias (Saikaly *et al.*, 2005; Pholchan *et al.*, 2010; Wells *et al.*, 2011; Wang *et al.*, 2012; Vuono *et al.*, 2015; Ju and Zhang, 2015; Lee *et al.*, 2015). Por un lado, los cambios en la estructura de la comunidad se han asociado con el OD, la temperatura, los sólidos (SSLM), el caudal y diferentes compuestos del efluente (Wells *et al.*, 2011). Por otro lado, se ha visto que además de los parámetros operacionales, las características del efluente y la ubicación

geográfica pueden contribuir a la estructuración de la comunidad bacteriana (Wang *et al.*, 2012). También se observó que cambios en ciertos parámetros operativos (cantidad de carga orgánica y composición de la alimentación) y en la configuración del sistema (régimen de alimentación continuo o en *batch*) afectan la diversidad de las comunidades microbianas y el rendimiento general de un sistema a escala de laboratorio (Pholchan *et al.*, 2010).

Particularmente la influencia del TRS ha sido estudiada en varias ocasiones (Saikaly *et al.*, 2005; Vuono *et al.*, 2015). Vuono y sus colaboradores se basaron en la relación que existe entre el TRS con las tasas de crecimiento para plantear la hipótesis de que su manipulación controlaría la composición y la diversidad microbiana. Considerando que sólo los microorganismos que tienen tiempos de duplicación menores a un correspondiente TRS serían capaces de permanecer en el reactor sin ser lavados del sistema, la manipulación del TRS implicaría que los microorganismos se seleccionen en base a las estrategias de crecimiento, la base de la teoría de selección r-K, que es lo que finalmente observaron. Si bien existen alternativamente otros estudios, principalmente realizados en ambientes marinos, que buscan conocer las bases genómicas respecto a las estrategias de crecimiento de bacterias (Lauro *et al.*, 2009; Haggerty and Dinsdale, 2017; Roller *et al.*, 2016), es necesario aumentar este conocimiento en función de las características ambientales que las seleccionan.

Interacciones multitróficas en barros activados

Muchos procesos, incluido el tratamiento de aguas residuales, están afectados por interacciones microbianas multitróficas mediadas por depredadores que ejercen una selección ecológica. Hasta qué punto estas interacciones dependen del contexto en el funcionamiento de estos procesos ha sido poco estudiado (Saleem and Moe, 2014). Las interacciones depredador-presa son un tema fundamental de la teoría ecológica. Los conocimientos actuales se basan principalmente en el estudio de plantas y animales, mientras que aún se conoce muy poco de las interacciones entre microorganismos en el ambiente. Los tres grupos de micro-depredadores incluyen protistas, bacterias depredadoras y bacteriófagos, cuyo tamaño,

selectividad y estrategias de depredación difieren notablemente (Julia *et al.*, 2014).

Los protistas son generalistas y constituyen uno de los principales factores que explican la mortalidad de los procariontes en el ambiente (Sherr and Sherr, 2002) debido a que son capaces de eliminar un gran número de bacterias en un corto período de tiempo.

Los bacteriófagos, por otra parte, en ecosistemas acuáticos pueden alcanzar densidades de entre 10 y 100 veces más altas que las bacterias (Breitbart, 2012), siendo la lisis viral responsable de hasta el 71% de la mortalidad bacteriana dependiendo de la estación del año (Thomas *et al.*, 2011). Los fagos son, a diferencia de los protozoos, muy específicos de sus hospederos, una característica que puede conducir a oscilaciones del depredador y de la presa como sugiere la hipótesis de "matar al ganador" ("*kill the winner*") (Thingstad, 2000). Al mismo tiempo, las cepas bacterianas desarrollan resistencia a bacteriófagos y estos a su vez evolucionan rápidamente para lograr infectar a su hospedero nuevamente, lo que sucesivamente lleva a un proceso que se ha equiparado a una carrera armamentista, siguiendo un modelo conocido como hipótesis de la Reina Roja (Weitz *et al.*, 2005). La dinámica de no equilibrio de dominancia alternada es considerada un factor importante en el mantenimiento de la biodiversidad (Vandermeer, 2006).

Son muy escasos los estudios experimentales diseñados para desentrañar el papel que desempeñan los virus bacterianos en la dinámica microbiana de los sistemas de barros activados. Los primeros estudios de virus en sistemas de tratamiento de efluentes se basaron en análisis de fago-hospedero mediante técnicas de cultivo (Ewert and Paynter, 1980; Hantula *et al.*, 1991; Khan *et al.*, 2002) y cuantificación de poblaciones de virus utilizando microscopio electrónico (Ewert and Paynter, 1980). Para esto último también se utilizaron técnicas de microscopía de fluorescencia (Ottawa *et al.*, 2007; Wu and Liu, 2009) y más recientemente citometría de flujo (Brown *et al.*, 2015), coincidiendo todas en que la concentración de virus en muestras de licor mezcla de barros activados está en el orden de 10^9 ml^{-1} , es decir un orden superior a la de bacterias. Estos y otros estudios, realizados en base a técnicas clásicas de

cultivo o microscopía, infieren que los fagos son una parte activa y abundante de este ecosistema y que podrían desempeñar un papel muy importante en el funcionamiento del proceso (Rosenberg *et al.*, 2010; Barr *et al.*, 2010).

También se ha discutido sobre el uso de bacteriófagos como herramienta de biocontrol en el proceso (Petrovski *et al.*, 2011; Kotay *et al.*, 2011), aunque no es algo que aún se haya llevado a la práctica. Hace poco menos de una década han comenzado a utilizarse técnicas metagenómicas en el estudio de las comunidades virales (Kunin *et al.*, 2008; Parsley *et al.*, 2010; Tamaki *et al.*, 2012; Alhamlan *et al.*, 2013; Aw *et al.*, 2014). Los ensayos sugieren que las comunidades virales son muy diversas, y que existe una enorme diversidad viral por descubrir.

Uno de los pocos reportes que mostraron interacciones fago-hospedero en una comunidad microbiana compleja durante un período prolongado fue en un sistema de tratamiento de aguas residuales en el que se analizaron cambios en la composición de la comunidad bacteriana de un biorreactor durante 462 días usando una serie de bibliotecas de clones de genes ARNr 16S, y se evaluó la interacción con fagos mediante técnicas de cultivo (Shapiro *et al.*, 2010). Los resultados preliminares fueron consistentes con un modelo de Lotka-Volterra explicado de acuerdo a la hipótesis de "matar al ganador".

Sin embargo, las infecciones virales pueden ser líticas o lisogénicas, por lo que las determinaciones de interacciones realizadas en cultivo se encuentran sesgadas tanto a los organismos cultivables como a las infecciones líticas. De este modo, la lisogenia se encuentra estudiada a nivel celular pero en ecosistemas sigue siendo subexplorada en comparación con las infecciones líticas (Howard-Varona *et al.*, 2017). En las últimas décadas, los nuevos métodos independientes de cultivo han dado la posibilidad de inferir cuestiones sobre ambas interacciones fago-hospedero (Fouts, 2006; Anderson *et al.*, 2011; Mizuno *et al.*, 2013; Allers *et al.*, 2013) abriendo la posibilidad de su aplicación para su estudio en ecosistemas complejos.

Las experiencias actuales más relevantes sobre interacciones multitróficas fueron obtenidas en ambientes acuáticos (Ory *et al.*, 2010; Rodríguez-Brito *et al.*, 2010; Bouvy *et al.*, 2011; Chow *et al.*, 2014; Labonté *et al.*, 2015). Se ha visto principalmente que en ecosistemas estables las funciones metabólicas permanecen en el tiempo (gracias a la redundancia funcional de especies equivalentes) mientras que existen fluctuaciones de comunidades de bacterias y fagos a lo largo del tiempo tal como sugiere la hipótesis de “matar al ganador” (Rodríguez-Brito *et al.*, 2010). También se sabe y se ha demostrado en ambientes marinos mediante redes de co-ocurrencia que las interacciones fago-bacteria son más específicas que las de protista-bacteria (Chow *et al.*, 2014). Sin embargo son aún escasos los estudios que muestran interacciones específicas de a pares depredador-presa en ecosistemas. Labonté y su grupo, utilizando datos de secuenciación de una sola célula fueron capaces de encontrar al menos seis interacciones fago-hospedero específicas en 20 familias de fagos correspondientes a 58 genomas bacterianos aislados de ambientes marinos (Labonté *et al.*, 2015). Desentrañar la dinámica de las interacciones fago-hospedero en ecosistemas es aún un desafío.

Perspectivas del trabajo de tesis

Comprender las bases genómicas de la selección biótica y abiótica de bacterias en barros activados requiere el análisis de un sistema a escala real, y un monitoreo frecuentemente durante un largo período de tiempo. El diseño experimental del trabajo de Tesis incluye el estudio de los tres niveles tróficos principales existentes en barros activados de una planta de tratamiento de efluentes domiciliarios: bacterias, eucariotas y bacteriófagos. El rango de tiempo estudiado de tres años estuvo caracterizado por un gradiente creciente de TRS, que oscila (entre 4 y 10 días) e incluyó un período de aproximadamente seis meses de perturbaciones esporádicas debidas a obras civiles llevadas a cabo en la planta de tratamiento. Para el estudio de los tres niveles tróficos se hizo uso de diversas técnicas, como la microscopía para el análisis de los organismos eucariotas, secuenciación de amplicones para la comunidad bacteriana, y se aprovecharon las grandes ventajas que otorga la secuenciación masiva al azar para estudiar el metaviroma y el metagenoma.

Esto último permitió realizar la reconstrucción de genomas microbianos en los metagenomas

HIPÓTESIS Y OBJETIVOS

OBJETIVO GENERAL

El objetivo general de la tesis es comprender la influencia de factores abióticos y bióticos en la estructura de las comunidades microbianas en un sistema de tratamiento de efluentes cloacales, enfocando el análisis en la elucidación de las bases genómicas de selección de microorganismos en función de las condiciones ambientales, y en el estudio de la dinámica de interacción entre bacterias y sus depredadores.

OBJETIVOS ESPECÍFICOS E HIPÓTESIS DE TRABAJO

La hipótesis del trabajo es que la diversidad y estructura de las comunidades microbianas de barros activados están determinadas por factores bióticos y abióticos, y que esta estructuración es importante para el funcionamiento y la estabilidad de dichos ecosistemas.

El primer objetivo de este estudio fue determinar la composición y diversidad de diferentes taxones microbianos (eucariotas, bacterias y virus) presentes en barros activados de una planta de tratamiento de efluentes domiciliarios a lo largo del tiempo.

El segundo objetivo fue evaluar la influencia de factores abióticos, en particular el tiempo de retención de sólidos y el efecto de perturbaciones, sobre la estructura de la comunidad microbiana.

El tercer objetivo consistió en analizar mediante redes tróficas la interacción entre los tres principales grupos microbianos (virus, bacterias, y eucariotas).

El cuarto objetivo residió en explorar las bases genómicas de selección de bacterias en función de las variables operativas.

Como quinto y último objetivo se buscó comprender como la interacción depredador-presa modula la dinámica bacteriana.

MATERIALES Y MÉTODOS

Sistema experimental

Descripción de la planta de tratamiento de efluentes

La planta de tratamiento de efluentes cloacales se encuentra ubicada en la localidad de San Fernando, Provincia de Buenos Aires. La planta depuradora trata los líquidos residuales domiciliarios provenientes de partidos de Tigre, San Fernando y San Isidro, y vierte el líquido tratado al Río Reconquista. La planta consiste en un diseño tipo modular, previsto para cuatro módulos idénticos, de los cuales el primer módulo (módulo I) se encuentra en servicio desde el año 1998; y el segundo (módulo II), desde el año 2014. Durante el año 2013 se llevaron a cabo las obras civiles correspondientes a la construcción del módulo II. Cada módulo trata un caudal medio de 78.000 m³/día, lo que representa una población equivalente de 300.000 habitantes.

El proceso de tratamiento de aguas residuales consta de tres etapas: pretratamiento, tratamiento primario y tratamiento secundario (Figura 3). También cuenta con una línea de tratamiento de los barros generados durante estas tres etapas, línea de la que no se hará mención en este trabajo.

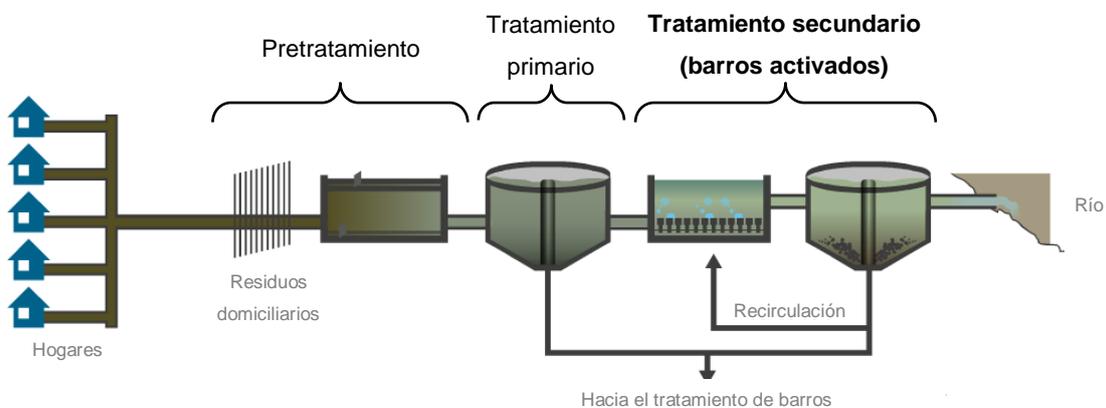


Figura 3. Esquema de las etapas de tratamiento de líquidos de la planta de tratamiento de efluentes estudiada.

En la etapa de **pretratamiento** se retienen ciertos residuos del efluente. En una fosa de gruesos se retienen los sólidos pesados y voluminosos. Luego el líquido pasa por diversas rejillas que retienen los sólidos de menor tamaño; y posteriormente, por un sector de desarenado y desengrasado. En la etapa de **tratamiento primario** se elimina la materia en suspensión del líquido por sedimentación en decantadores primarios. La etapa de **tratamiento secundario** consiste en un reactor biológico de barros activados (cámara de aireación), en donde la materia orgánica disuelta y coloidal se transforma en biomasa sedimentable que se separa del efluente tratado en decantadores secundarios o clarificadores.



Figura 4. Fotografía de las unidades operativas correspondientes a la etapa de tratamiento secundario (barros activados) pertenecientes al módulo I de la planta en estudio. Cada módulo consta de cuatro cámaras de aireación y cuatro sedimentadores secundarios.

Muestreo y conservación de las muestras

Durante un período de tres años, comenzando en el mes de mayo del año 2012 y finalizando en el mes de septiembre del año 2015 se tomaron muestras semanales del licor mezcla de una de las cámaras de aireación del módulo I de la planta de tratamiento de efluentes. Las muestras fueron tomadas siempre en el mismo sector de la cámara, cercano a la salida, con un recipiente muestreador de 2 litros, y fueron transportadas hasta el laboratorio

dentro de las dos horas posteriores a la toma de muestra, a temperatura ambiente, en un recipiente plástico dejando una cámara de aire durante el traslado.

Una vez en el laboratorio se conservaron diferentes alícuotas de las muestras:

- 2 alícuotas de 1,5 ml de muestra a -20°C
- 1 alícuota de 1,5 ml de muestra a -80°C
- 2 alícuotas de 1,5 ml de muestra en glicerol 50% a -80°C
- 1 alícuota de 300 µl de muestra fijada en paraformaldehído (PFA) 4% (ver protocolo de fijación en PFA a continuación) a -20°C
- 2 alícuotas de 15 ml de sobrenadante filtrado en glicerol 20% (ver protocolo de conservación de muestras para análisis de virus a continuación), una a -20°C y otra a -80°C

Protocolo de fijación en PFA

300 µl de muestra fresca fueron incubados con 900 µl de paraformaldehído 4% a 4°C durante 1 hora. Posteriormente la muestra se lavó 2 veces en PBS 1X y se resuspendió en 300 µl de PBS 1X y 300 µl de etanol. La muestra fijada se conservó a -20°C para su posterior análisis.

Protocolo de conservación de muestras para análisis de virus

En un tubo se colocaron 50 ml de muestra fresca, se agitó en vortex por 3 minutos y se centrifugó 2 veces a 4000 rpm por 15 minutos. El sobrenadante se pasó por un filtro de 0,2 µm y en esterilidad se conservaron 2 alícuotas de 15 ml del sobrenadante filtrado en glicerol 20%.

Medición de parámetros operacionales

Durante el muestreo se midió la temperatura y el oxígeno disuelto (OD) *in situ*. En el laboratorio de la planta se realizaron las siguientes mediciones para realizar las curvas de sedimentabilidad: se homogeneizó la muestra, se colocó 1 litro en una probeta (de 1 y 2 litros), y se registró el volumen que ocupa el barro sedimentado a los 5, 10, 15, 20, 25, 30, 40, 50 y 60 minutos. El volumen ocupado por el sólido sedimentado a los 30 minutos se conoce como sólidos sedimentables a los 30 minutos (SS30').

Por otro lado los operadores de planta realizaron periódicamente mediciones de pH, potencial de óxido-reducción (ORP) y sólidos suspendidos totales del licor mezcla (MLSS) sobre las muestra de barros activados. Estos y otros parámetros (como DBO, DQO y turbidez) medidos por los analistas diariamente sobre distintos puntos de muestreo de la planta fueron registrados para su posterior análisis. A su vez, sobre muestras de ingreso y salida, diversos parámetros de calidad como metales, DBO, DQO, fósforo, nitrógeno total Kjeldahl (NTK), nitratos, nitritos, amonio, entre otros, fueron medidos mensualmente por un laboratorio externo.

Características operacionales del período de muestreo

La siguiente tabla resume las medidas estadísticas de los principales parámetros operacionales de la planta depuradora durante el período muestreado.

Parámetro	Unidad	Media	Desvío Estándar	Mínimo	Máximo
TRS	día	7,1	2,0	3,6	11,5
TRH	h	3,4	0,9	2,2	6,0
Caudal de ingreso (promedio mensual)	m ³ /día	112.838	23.065	80.352	172.800
T*	°C	22,0	2,9	13,2	28,7
pH*		7,3	0,1	7,0	7,8
SSLM*	mg/l	1739	526	360	4420
IVB*	ml/g	235	129	69	923
DBO ingreso	mg/l	176	59	51	822
DQO ingreso	mg/l	384	121	129	1276
NH₄⁺ ingreso	mg/l	16,5	6,4	3,3	30,0
NTK ingreso	mg/l	25,0	6,9	6,1	45,0
NO₃⁻ ingreso	mg/l	1,74	1,77	0,50	9,60
NO₂⁻ ingreso	mg/l	0,043	0,061	0,004	0,310
P ingreso	mg/l	3,24	1,06	0,56	5,00
DBO salida	mg/l	17	4	<10	36
DQO salida	mg/l	50	13	8	113
NH₄⁺ salida	mg/l	7,9	5,8	3,0	23,0
TKN salida	mg/l	11,2	8,1	3,0	29,0
NO₃⁻ salida	mg/l	7,4	4,6	0,9	20,1
NO₂⁻ salida	mg/l	1,34	2,07	0,02	8,70

P salida	mg/l	2,01	1,60	0,63	10,00
-----------------	------	------	------	------	-------

* Medida sobre la cámara de aireación

Tabla 1. Resumen de los parámetros operacionales de la planta estudiada en el período de muestreo.
Todos los valores registrados durante el período fueron tenidos en cuenta.

La variación temporal de los parámetros operativos más relevantes de la planta de tratamiento estudiada se puede observar en la Figura 5.

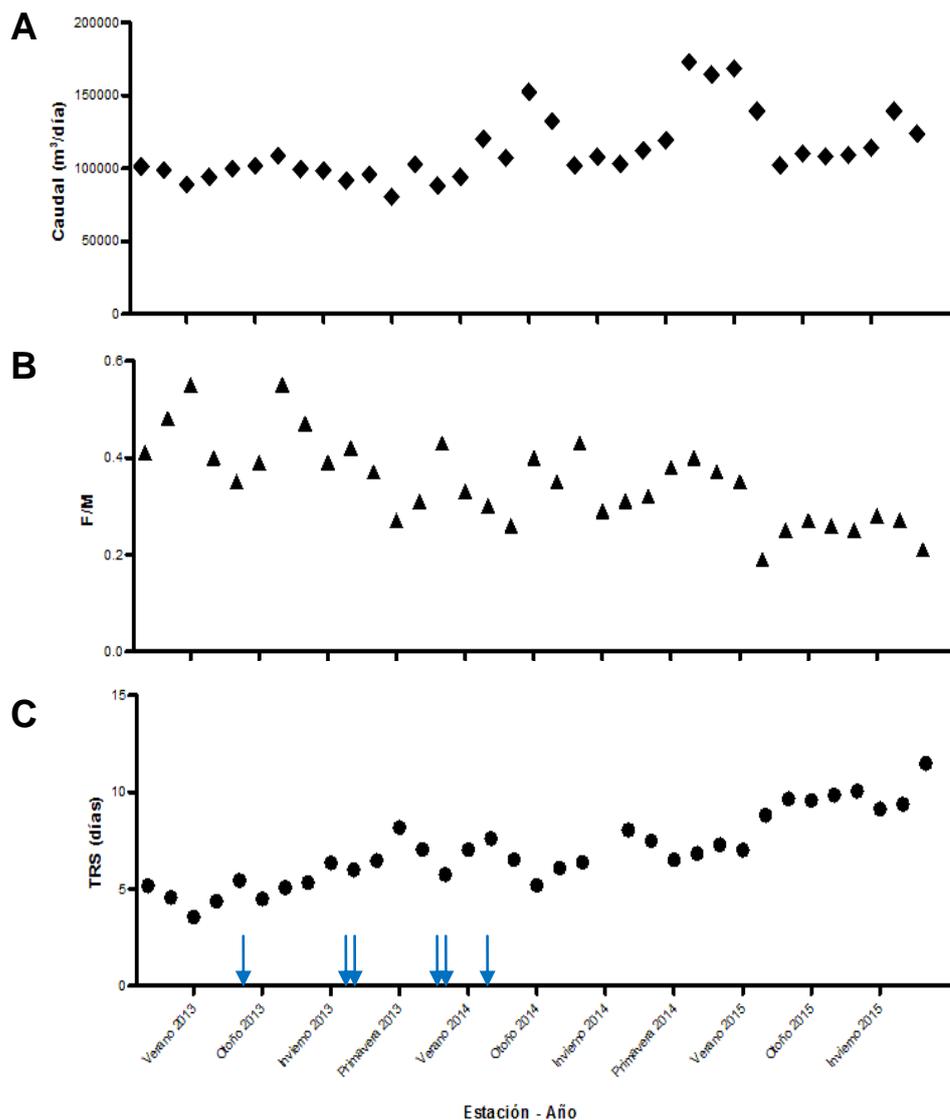


Figura 5. Parámetros operativos mensuales de la planta de tratamiento de efluentes estudiada durante el período de muestreo. **A)** Caudal de ingreso (promedio mensual). **B)** Relación F/M. **C)** Tiempo de retención de sólidos (las flechas azules indican operativos de planta realizados durante la construcción del módulo II).

Durante el período investigado, el caudal de ingreso de líquido a tratar aumentó a lo largo del tiempo debido a la extensión de la red cloacal (Figura 5 A). Superpuesto a este incremento, ocurrieron variaciones estacionales explicadas por el mayor uso del agua durante los meses de verano. El aumento de la capacidad de la planta fue posible gracias a la construcción y el comienzo de operación del segundo módulo. Durante la puesta en marcha del mismo, la incorporación en funcionamiento de las unidades operativas se fue realizando sucesivamente de acuerdo al incremento de los caudales de ingreso. Esto explica que, durante el período investigado la planta de tratamiento de efluentes operara bajo un gradiente creciente de tiempo de retención de sólidos (TRS), de entre 4 y 10 días.

En la Figura 5 C se pueden observar los valores promedio del TRS estimados a partir de los promedios mensuales del caudal de entrada, de los SSLM y del volumen correspondiente a las unidades operativas en funcionamiento en el correspondiente mes (número de cámaras de aireación en operación).

La relación F/M, que expresa la carga en el reactor (kgDBO/kgMLSS.día), disminuyó en el tiempo de muestreo como se puede observar en la Figura 5 B, de manera opuesta al TRS.

La etapa de construcción y puesta en marcha del nuevo módulo durante el año 2013 se caracterizó por la existencia de fluctuaciones operativas debidas a obras de interconexión, que incluyó paradas ocasionales durante varias horas de ingreso de efluente a tratar, indicadas con flechas azules en la Figura 5 C.

En la Figura 6 se graficaron las principales variables medidas sobre la cámara de aireación en función del tiempo. Durante el período de perturbaciones mencionado las celdas de aireación atravesaron ocasiones de anoxia (Figura 6 A) y también una amplia variación de la concentración de sólidos, que se estabilizó en el período final de muestreo (Figura 6 B).

Los cambios de temperatura claramente responden a la variación estacional, alcanzando temperaturas *in situ* de 17°C durante los meses de invierno y hasta 28°C durante los meses de verano (Figura 6 C).

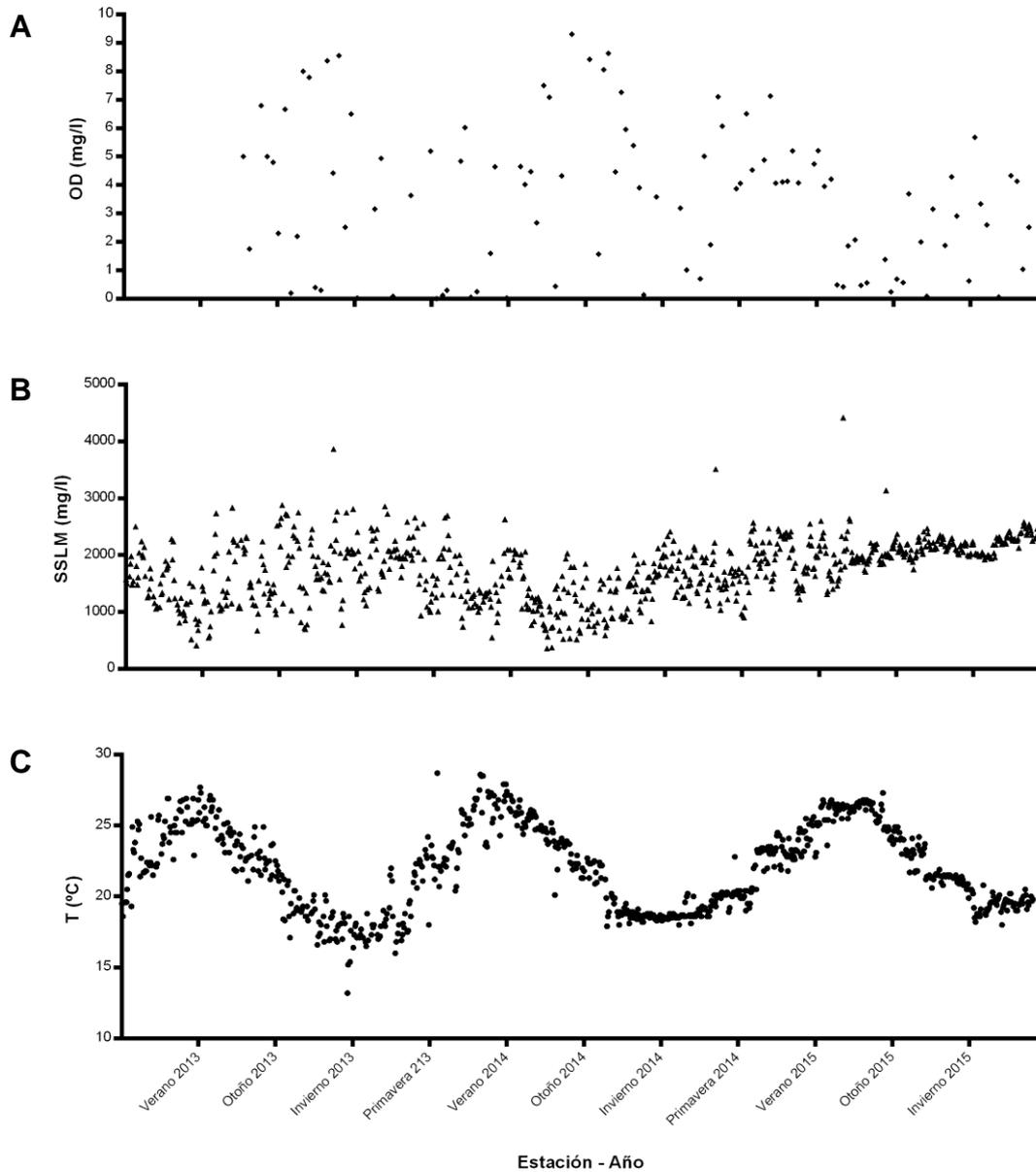


Figura 6. Variables medidas sobre la cámara de aireación de la planta de tratamiento de efluentes estudiada durante el período de muestreo. **A)** Oxígeno disuelto (medición puntual in situ) **B)** Sólidos suspendidos totales del licor mezcla **C)** Temperatura in situ

La figura 7 muestra los principales parámetros de calidad de la planta de tratamiento durante el período de muestreo: la DBO de ingreso y de salida (Figura 7 A); y las concentraciones de algunos compuestos nitrogenados

(Figura 7 B) que muestran que hacia fines del año 2014 la planta comenzó a nitrificar.

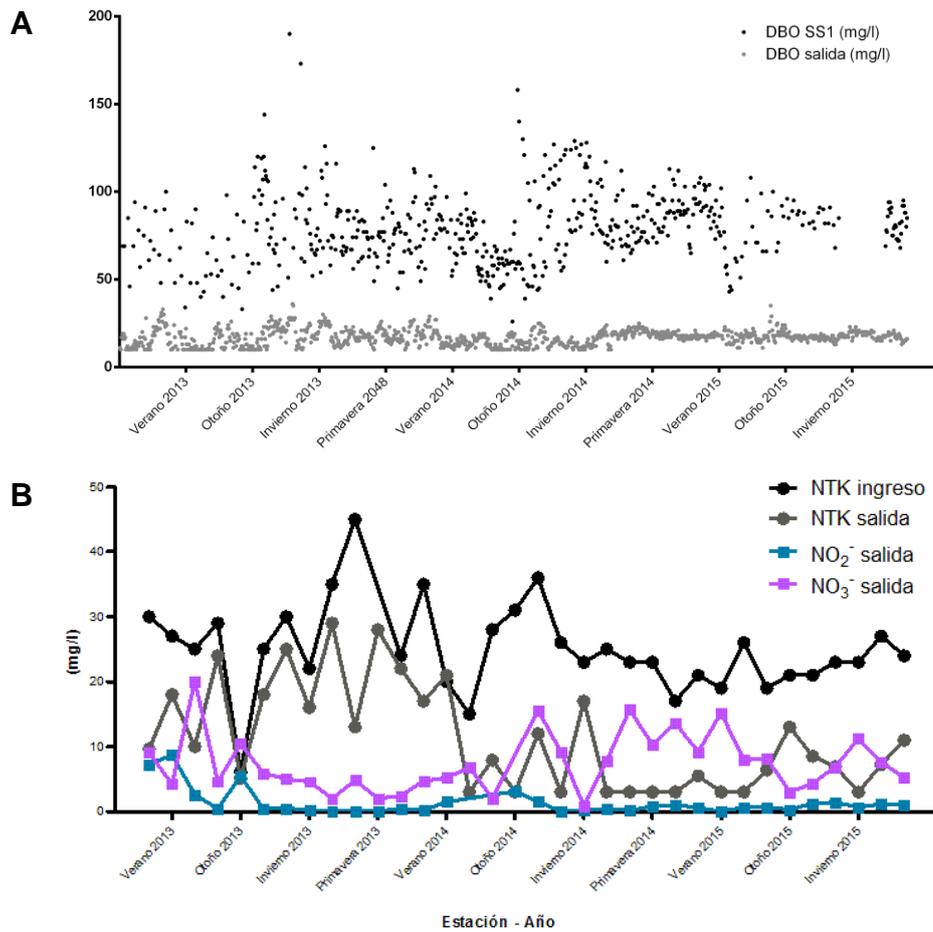


Figura 7. Parámetros de calidad de la planta durante el período de muestreo. **A)** Valores de DBO diarios tomados en la salida del sedimentador primario (SS1), líquido que ingresa a la cámara de aireación, y en la salida de planta. **B)** Valores mensuales de compuestos nitrogenados al ingreso y a la salida de planta.

En base a las características funcionales descriptas durante el período investigado, gradiente creciente de tiempo de retención de sólidos y período perturbado, se han definido 3 períodos operativos, que a partir de ahora se denominan: **PP**: período perturbado (de bajo TRS y con perturbaciones operacionales), **PE**: período estable (de alta estabilidad y alto TRS) y **PI**: período intermedio (de estabilidad y bajo TRS).

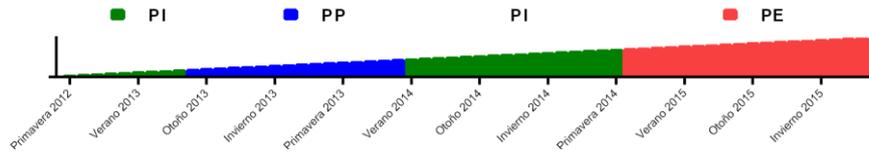


Figura 8. Períodos definidos según las características observadas. El gradiente representa de manera esquemática el incremento en el TRS

Microscopía y cuantificación de organismos eucariotas

Todas las muestras fueron observadas en microscopio óptico (OLYMPUS B41X) en campo claro el mismo día de muestreo. Sobre cada muestra fresca se registraron las características del floc: morfología, abundancia de filamentos y efectos de filamentos en la estructura del floc (Jenkins *et al.*, 2004). Sobre un volumen de 20 μ l de muestra se realizó una cuantificación de organismos eucariotas (protozoos y metazoos) realizando un recorrido a lo largo del portaobjetos con una magnitud de 200x. Los metazoos se identificaron a nivel de filo (rotíferos, nematodos, gastrótricos) y los protozoos a nivel de género dentro de los tres grupos morfológicos que se encuentran comúnmente en estos ambientes: ciliados (caminadores, nadadores libres y sésiles), flagelados y amebas (Serrano *et al.*, 2008; Berk and Gunderson, 1993; Gerardi and Horsfall, 1995; Richard, 1989). A partir de estos datos se estimó el índice biótico de barro, SBI por sus siglas en inglés (Madoni, 1994).

Extracción de ADN

Extracción de ADN total

El ADN total fue extraído mediante el método CTAB modificado (Wagner-Döbler *et al.*, 1999), un procedimiento de lisis directa que involucra disrupción mecánica de las células y purificación. Se centrifugó 1,5 ml de muestra y se resuspendió en 500 μ l de buffer TENP pH 7,6 (50 mM Tris-HCl, 10 mM EDTA, 10 mM NaCl, 1% polivinilpirrolidona (PVP)) y se transfirió a tubos de 2 ml con tapa a rosca con 200 μ l de esferas de zirconia/silica de 0,5 mm (BioSpec Products, Inc.). Se agregaron 50 μ l de dodecilsulfato sódico 10% (SDS) y 1 μ l

de RNAsa A (100 mg/ml) y se incubaron por 1 hora a 37°C. Luego se agregaron 3 µl de proteinasa K (20 mg/ml) y se incubaron a 37°C por 1 hora. Luego de la incubación, se rompieron las células en un homogeneizador celular (Precellys 24, Bioamerica) a 5000 rpm en 3 ciclos de 30 segundos. A las suspensiones obtenidas se les agregaron 120 µl de NaCl 5 M y 80 µl de bromuro de cetiltrimetilamonio (CTAB) 10% en NaCl 0,7 M y se incubó por 10 minutos a 65°C. El ADN fue extraído con un volumen equivalente de fenol equilibrado con buffer TE (pH 7.5 a 8.0) y centrifugado por 10 minutos a 13.200 rpm a temperatura ambiente. La fase acuosa se transfirió a un nuevo tubo y se extrajo dos veces con iguales volúmenes de cloroformo isoamilalcohol (24:1). El ADN se precipitó incubando toda la noche a -20°C con 0,7 volúmenes de isopropanol y 0,1 volúmenes de acetato de amonio 3 M. Se centrifugó 40 minutos a 13.000 rpm a 4°C. El ADN precipitado se lavó con etanol 70 % y se dejó secar. Finalmente, el ADN se resuspendió en buffer TE (pH 8) o en agua grado biología molecular, para el caso de las muestras que fueron enviadas a secuenciar, y se conservó a -20°C.

Extracción de ADN viral

Previamente a la extracción del ADN viral, las muestras conservadas para análisis de virus fueron dializadas y concentradas. Para ello las alícuotas de 15 ml de muestra filtrada conservada en 20% de glicerol se incubaron a temperatura ambiente en membranas de diálisis (Dialysis tubing cellulose membrane SIGMA-ALDRICH) en buffer TENP diluido (1:4) o agua, toda la noche en agitación, con el fin de eliminar el glicerol y las sales disueltas. Esto se repitió dos veces en un recipiente con un volumen total de 2 litros de buffer o agua (para la primera y segunda diálisis respectivamente) conteniendo 7 muestras. El dializado se concentró en *Speed Vac* (FTS Systems) hasta un volumen de 200 µl. Finalmente, la extracción de ADN viral se realizó sobre estas muestras con el kit de Qiagen MinElute Virus Spin Kit, según protocolo del fabricante.

Ensayos preliminares

Electroforesis en gel con gradiente desnaturalizante (DGGE)

DGGE para bacterias

PCR

La reacción de amplificación se llevó a cabo en un volumen total de 50 µl con una concentración final de buffer Taq polimerasa de 1X, MgCl₂ 3 mM, dimetilsulfóxido 5%, BSA 0,25 µg/µl, dNTPs 0,2 mM, 0,5 µM de cada cebador, 2,5 unidades de Taq ADN polimerasa y aproximadamente 1 ng de ADN. El programa de amplificación fue el siguiente: 5 minutos a 94°C, 35 ciclos de 45 segundos a 94°C, 45 segundos a 55°C y 45 segundos a 72°C, y finalmente 10 minutos a 72°C. Se llevó a cabo la amplificación de un fragmento de 193 pb correspondiente al dominio hipervariable V3 del gen de ARN ribosomal 16S utilizando los cebadores F341-GC (5'-CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCGCCCG-CCTACGGGAGGCAGCAG-3') y R534 (5'-ATTACCGCGGCTGCTGG-3') (Muyzer *et al.*, 1993).

DGGE

El procedimiento de electroforesis en gel con gradiente desnaturalizante (DGGE) se realizó de acuerdo a los procedimientos establecidos por Muyzer y colaboradores (Muyzer *et al.*, 1993). Utilizando un armador de gradiente (Hofer Scientific Instruments) se preparó un gel de poliacrilamida 8% con un gradiente desnaturalizante de 40 a 65% (donde el 100% de desnaturalizante corresponde a una solución de urea 7 M y formamida 40 %). 45 µl de los productos de amplificación fueron concentrados en *Speed Vac* (FTS Systems) hasta 10 µl para ser sembrados en el gel junto con 4 µl de *loading buffer*. La corrida electroforética se llevó a cabo en un equipo de DGGE (CBS Scientific) a 65 V durante 16 horas a 60°C en buffer TAE 1X. Los geles fueron teñidos con SYBR Gold (Invitrogen) 1X en TAE 1X durante 30 minutos, y lavados en TAE 1X durante 10 minutos. Se visualizaron en transiluminador UV con cámara acoplada (Syngene). Se utilizó en programa GelCompareII (Applied Maths, Belgium) para procesar las imágenes de los geles obtenidos.

DGGE para fagos de la familia Myoviridae

PCR

La reacción de amplificación se realizó en un volumen total de 50 µl con una concentración final de buffer Taq polimerasa de 1X, MgCl₂ 3 mM, dimetilsulfóxido 5%, BSA 0,25 µg/µl, dNTPs 0,2 mM, 0,5 µM de cada cebador, 2,5 unidades de Taq ADN polimerasa y aproximadamente 1 ng de ADN. El programa de amplificación fue el siguiente: 3 minutos a 94°C, 35 ciclos de 45 segundos a 94°C, 1 minuto a 50°C y 45 segundos a 72°C, y finalmente 5 minutos a 72°C. Se llevó a cabo la amplificación de un fragmento de entre 500 y 600 pb correspondiente al gen que codifica para g23, proteína principal de la cápside de los fagos tipo-T4, que comprenden la gran mayoría de los bacteriófagos pertenecientes a la familia *Myoviridae*. Para ello se utilizó un conjunto de cebadores de PCR degenerados: MZIA1bis (5'-GATATTTGIGGIGTTCAGCCATGA-3') y MZIA6 (5'-CGCGGTTGATTTCCAGCATGATTC-3') (Fujii *et al.*, 2008; Filee *et al.*, 2005)

DGGE

El procedimiento de DGGE se realizó de acuerdo a lo explicado en el DGGE para bacterias utilizando un gradiente desnaturante de 20 a 65% puesto a punto teniendo como referencia un experimento similar realizado en suelos (Fujii *et al.*, 2008).

Secuenciación de los productos de PCR

Con el fin de verificar que los productos de PCR correspondan a las secuencias de bacteriófagos se realizó un clonado del producto de PCR de cinco de las muestras y 17 de ellos se enviaron a secuenciar.

Se realizó la reacción de amplificación tal como se explicó anteriormente sobre el ADN total de cinco muestras (cada una correspondiente a los meses de febrero, abril, mayo, agosto y septiembre del año 2014). Los productos de amplificación se purificaron con el kit de Qiagen (QIAquick PCR Purification Kit) según protocolo. Luego se llevó a cabo la reacción de A-tailing incubando 4 µl del producto de amplificación purificado con 1 µl de buffer Taq polimerasa 10X, 0,25 µl de MgCl₂ 50 mM, 1 µl de dATP, 1 µl de enzima Taq polimerasa

(Invitrogen) 5 U/μl y 3 μl de agua libre de nucleasas durante 30 minutos a 70°C. Posteriormente se ligó en vector pGEM T-Easy (Promega). Cada reacción de ligación se llevó a cabo en un volumen total de 10 μl con una relación molar entre inserto y vector de 3 a 1 (5 μl de Buffer 2x, 1 μl de vector, 3 μl de ADN, 1 μl de ligasa T4 (3 Unidades Weiss), y 1 μl agua libre de nucleasas) incubando toda la noche a 4°C. Se frenó la ligación incubando 15 minutos a 65°C. Finalmente se emplearon células *E. coli* termocompetentes (DH5α) para llevar a cabo la transformación con el vector ya ligado. Se realizó una dilución 1/100 de bacterias inoculadas toda la noche en 10 ml de medio LB y se inoculó 2 horas a 37°C. Una alícuota de 1,5 ml por cada transformación se centrifugó a 5000 rpm por 2 minutos y el *pellet* se resuspendió en 1 ml de CaCl₂ 0,1M estéril y frío. Se centrifugó nuevamente, el *pellet* se resuspendió en 0,1 ml de CaCl₂ 0,1M etéril y frío y se agregaron 2 μl de plásmido ligado y se incubó una hora en hielo, luego 90 segundos a 42°C y finalmente 2 minutos en hielo. Se agregaron 0,9 ml de medio LB y se incubó 1 hora a 37°C. Luego se centrifugó 2 minutos a 5000 rpm, se descartó el sobrenadante dejando entre 100 y 200 μl y se resuspendió nuevamente. Seguidamente se sembró en placas de petri con agar LB con ampicilina 100 μg/ml, IPTG (100 μl de una solución 100mM por placa) y X-Gal (20 μl de una solución 50 mg/ml por placa). Se realizaron controles de bacteria sin plásmido. Se incubaron toda la noche a 37°C. (Rendimiento 5×10^6 - 2×10^7 colonias transformadas por μg de plásmido superenrollado). El *screening* de los clones positivos (colonias blancas que contienen el vector con inserto interrumpiendo el gen de la β-galactosidasa) se realizó mediante la PCR utilizada para el DGGE y se realizó un DGGE para visualizar la posición de la banda correspondiente luego de la corrida.

Se extrajo el ADN plasmídico de las colonias transformantes. Los amplicones correspondientes a los productos de PCR introducidos en los plásmidos purificados se secuenciaron bajo el servicio de MacroGen Inc.

Búsqueda de hospederos de fagos en cultivo

Con el objetivo de hallar, mediante técnicas clásicas de cultivo, bacterias que sean hospederos de bacteriófagos se procedió a realizar un aislamiento de bacterias en medio de cultivo R2A. Con cada cultivo bacteriano puro se realizó un ensayo de placas de lisis utilizando el sobrenadante filtrado como fuente de fagos.

Aislamiento de bacterias

Se disgregaron en *vortex* 500 µl de muestra de barros activados con esferas de zirconia/silica de 0,5 mm (BioSpec Products, Inc) por 60 segundos. Se sembraron 150 µl de cada una de las diluciones (10^{-4} , 5×10^{-5} y 10^{-5}) en solución fisiológica estéril sobre placas de Petri con agar R2A y se incubaron a 30°C. Se aislaron por re-estriado en placa colonias de bacterias que crecieron entre las 24 hs y los 5 días.

Ensayo de placa

Se realizó un ensayo de placa sobre todos los aislamientos bacterianos obtenidos crecidos en cultivo líquido R2A. En un tubo de cultivo de vidrio con 1,5 ml de top agar R2A (1,2% de agar) a 55°C, se agregaron 500 µl de cultivo bacteriano y 1 ml de fuente de fagos (sobrenadante de la muestra filtrado en filtro de 0,2 µl). Rápidamente se volcó en una placa de Petri con una base de agar R2A y se incubó a 30°C. Durante los días posteriores se observaron las playas de lisis formadas.

El ADN genómico de los aislamientos bacterianos a los que se les encontró alguna playa de lisis en el ensayo de placa se enviaron a MacroGen Inc. para secuenciación del gen ARNr 16S completo.

Pirosecuenciación de regiones hipervariables del gen ARNr 16S

PCR y secuenciación

Se seleccionaron 12 muestras tomadas con intervalos mensuales desde mayo de 2012 a abril de 2013.

Las regiones hipervariables V1–V3 y V4 del gen ARNr 16S fueron amplificadas por duplicado del ADN total de todas las muestras en el tiempo con los cebadores universales F8 (5'-AGAGTTTGATCCTGGCTCAG-3') y R534 (5'-WTTACCGCGGCTGCTGG-3'), y F563 (5'-AYTGGGYDTAAAGNG-3') y R907 (5'-CCGTCAATTCMTTTRAGT-3') respectivamente para cada región. Los 48 amplicones resultantes (12 muestras en el tiempo, 2 regiones hipervariables y 2 réplicas técnicas) fueron secuenciados en un equipo Roche 454 GS FLX en Microbiome Core Facility (North Carolina, USA).

Procesamiento de las secuencias y análisis de datos

Se generaron un total de 281.102 lecturas crudas con un largo promedio de 451 pb para la región V1–V3, y 255.145 lecturas con un largo promedio de 364 pb para la región V4. Luego del procedimiento de filtrado, este número se redujo a 161.822 y 140.753 lecturas, respectivamente.

Se utilizó el programa Mothur v.1.31.2 (Schloss *et al.*, 2009) para eliminar el ruido, recortar, filtrar y alinear secuencias, encontrar quimeras, asignar secuencias a unidades taxonómicas operacionales (OTUs) con un 97% de similitud y describir la α -diversidad, siguiendo el procedimiento operativo estándar sugerido por el autor. A fin de evitar el sesgo causado por las diferencias en la profundidad de secuenciación en la estimación de la alfa y la beta-diversidad, se realizó el submuestreo al azar de 2125 secuencias por muestra. Las secuencias fueron clasificadas utilizando la base de datos de ARNr 16S trainset9 de RDP (Cole *et al.*, 2014).

Hibridación fluorescente in situ (FISH)

Protocolo de hibridación

Sobre las mismas 12 muestras a las que se les realizó el ensayo de pirosecuenciación de regiones hipervariables del gen ARNr 16S se realizó un protocolo de hibridación fluorescente in situ para una posterior cuantificación (Daims, 2009) de bacterias del género *Thiothrix*. Los portaobjetos fueron

pretratados sumergiéndolos en una solución de gelatina 0,1 % y KCrSO_4 0,01 %, a 70°C. En cada pocillo del portaobjetos pretratado se colocaron 5 μl de muestra fijada en PFA 4% y se dejó secar. Este procedimiento se repitió tres veces. Una vez seca la muestra, se sumergió el portaobjetos en una solución de agarosa 0,75% y se dejó solidificar en frío. Luego se deshidrató sumergiendo por tres minutos el portaobjetos en soluciones de etanol 50%, 80% y 100% sucesivamente. Se dejó secar y se agregaron 9 μl de mezcla de hibridación (8 μl de buffer de hibridación y 1 μl de sonda de 25 ng/ μl) sobre cada preparado y se incubó durante 2 horas a 46°C en cámara húmeda, en un horno de hibridación. Se utilizó una sonda específica para el género *Thiothrix* marcada con Cy3 (5'-CCTTCCGATCTCTATGCA-3') en presencia de cantidades equimolares de su competidor sin marcar (5'-CCTTCCGATCTCTACGCA-3'), en 40 % formamida (Kanagawa *et al.*, 2000). Se realizaron controles negativos utilizando la sonda NONEUB (59-ACTCCTACGGGAGGCAGC-39) y otros controles sin agregado de sonda.

Luego de la hibridación, los portaobjetos fueron sumergidos en el buffer de lavado correspondiente durante 20 minutos a 46°C. Se enjuagó con agua y se dejó secar en oscuridad. Luego se colocaron sobre cada muestra 5 μl de solución de DAPI (4',6-diamidino-2-fenilindol diclorhidrato) 1 ng/ μl y se incubó por 5 minutos en oscuridad. Se lavó con etanol 80 %, se enjuagó con agua y se dejó secar. Finalmente, los preparados fueron montados con una solución de n-propilgalato 2% en glicerol (Giloh and Sedat, 1982) a fin de prevenir el fotoblanqueo.

Adquisición de imágenes y análisis

Las muestras se observaron en un Microscopio Confocal (Olympus Fluoview FV1000) utilizando un objetivo de magnificación de 600X (UPLSAPO 60X WNA:1.20) acoplado a una cámara CCD. Se adquirieron las imágenes digitales de 30 campos de visión en posiciones elegidas al azar. Para cada campo, se capturó una imagen de la señal de la sonda específica para *Thiothrix* utilizada (Cy3) y una imagen con la señal de DAPI. En cada campo se determinó la proporción de *Thiothrix* en el floc, como la relación entre las señales entre Cy3

y DAPI, estimada a partir del porcentaje de volumen que el taxon detectado ocupa en la muestra de barro (biovolumen). Para cada muestra se consideró el promedio del valor estimado en los 30 campos. El análisis de las imágenes y el cálculo del biovolumen se llevó a cabo con el programa Daime 2.0 (Daims *et al.*, 2006a), descargado desde <http://www.microbial-ecology.net/daime/>.

Secuenciación metagenómica de la serie temporal

Diseño experimental

Se seleccionaron 60 muestras espaciadas quincenalmente durante el período de tres años de muestreo, comenzando en noviembre de 2012 y finalizando en septiembre de 2015. Se extrajo el ADN total del barro activado de todas las muestras. Los sobrenadantes conservados para análisis de virus correspondientes a estas muestras fueron unificados en grupos de 5 muestras consecutivas para obtener 12 muestras finales a las que se les extrajo el ADN viral. La siguiente tabla muestra las fechas de muestreo correspondientes a las muestras utilizadas.

Número de muestra	Fecha de muestreo	Número de muestra	Fecha de muestreo	Número de muestra	Fecha de muestreo
01	28/11/2012	21	08/11/2013	41	17/10/2014
02	12/12/2012	22	29/11/2013	42	05/11/2014
03	02/01/2013	23	11/12/2013	43	26/11/2014
04	16/01/2013	24	27/12/2013	44	10/12/2014
05	30/01/2013	25	22/01/2014	45	29/12/2014
06	19/02/2013	26	07/02/2014	46	16/01/2015
07	13/03/2013	27	19/02/2014	47	28/01/2015
08	27/03/2013	28	06/03/2014	48	19/02/2015
09	17/04/2013	29	27/03/2014	49	05/03/2015
10	30/04/2013	30	08/04/2014	50	19/03/2015
11	22/05/2013	31	29/04/2014	51	09/04/2015
12	05/06/2013	32	15/05/2014	52	30/04/2015
13	19/06/2013	33	05/06/2014	53	07/05/2015
14	10/07/2013	34	19/06/2014	54	28/05/2015
15	24/07/2013	35	01/07/2014	55	18/06/2015

16	14/08/2013	36	23/07/2014	56	02/07/2015
17	28/08/2013	37	13/08/2014	57	23/07/2015
18	11/09/2013	38	29/08/2014	58	06/08/2015
19	02/10/2013	39	17/09/2014	59	03/09/2015
20	18/10/2013	40	01/10/2014	60	24/09/2015

Tabla 2. Fechas de muestreo que corresponden a las 60 muestras de la serie temporal del período de tres años utilizadas en el análisis.

Secuenciación de amplicones

PCR y Secuenciación

El ADN total extraído del barro activado de las 60 muestras fue enviado a MACROGEN Inc. para la secuenciación de amplicones de la región V3-V4 del gen ARNr 16S en un sistema MiSeq de Illumina utilizando los cebadores b341F y Bakt805R, obteniendo lecturas apareadas de 300 pb.

Cebador	Secuencia (5'-3')	Posición	Longitud (pb)
b341F	CCTACGGGAGGCAGCAG	341-357	17
Bakt805R	GACTACHVGGGTATCTAATCC	785-805	21

Tabla 3. Cebadores utilizados en la secuenciación de amplicones de la región V3-V4 de las 60 muestras de la serie temporal.

Análisis de datos

Se obtuvieron 7.293.800 pares de lecturas crudas con un mínimo de 61.320 y un máximo de 172.007 por muestra. Luego de realizar controles de calidad de las secuencias (Fastqc), las lecturas fueron filtradas con Trimmomatic (Bolger *et al.*, 2014) quedando 4.581.297 pares de lecturas totales con un mínimo de 36.932 y un máximo de 105.766 por muestra.

Se utilizó USEARCH (Edgar, 2013) para juntar las secuencias apareadas, considerando un solapamiento mínimo de 32 pb entre los pares de secuencias, un mínimo de longitud de 400 pb y un máximo de 450 pb, quedando un total de 2.498.186 secuencias apareadas. Se utilizó también USEARCH para filtrar, y

definir las unidades taxonómicas operacionales al 97% de similitud. Se realizó un submuestreo considerando la muestra con el menor número de secuencias, quedando así un total de 15481 secuencias totales por cada muestra. Las secuencias fueron clasificadas con Mothur (Schloss *et al.*, 2009) usando la base de datos trainset14 de RDP (Cole *et al.*, 2014).

Red de co-ocurrencia

Para determinar las asociaciones entre especies y dependencia de factores ambientales se utilizó el análisis de similitud local (LSA, por sus siglas en inglés) (Xia *et al.*, 2013, 2011). El LSA es un enfoque computacional que permite establecer relaciones entre variables en series temporales. El método de LSA es mucho más ventajoso que cualquier otro a la hora de estudiar series temporales en metagenómica ya que detecta patrones de asociación y de co-ocurrencia locales y desplazados en el tiempo que no podrían identificarse mediante los análisis de correlación habituales. Se consideraron significativas las correlaciones positivas con un *score* mayor a 0,7 y las negativas con un *score* menor a -0,7, con un p-valor < 0.01. Se utilizó el programa Cytoscape (Shannon *et al.*, 2003) para esquematizar las redes de interacción.

Se utilizó el paquete *vegan* 2.0-7 en R 3.0.2 para realizar un análisis de coordenadas principales restringido utilizando la matriz de abundancia de taxones (OTUs) abundantes para todas las muestras y las variables ambientales T, TRS y SSLM estandarizadas. Se usó una prueba de permutación similar a ANOVA para evaluar la significancia de las restricciones.

Relación especies-tiempo

La relación especies-tiempo ($S = cT^w$) se estimó como la riqueza acumulada de especies en función del tiempo. La tasa de reemplazo bacteriana y eucariota (*w*) se obtuvo a partir de las pendientes de las regresiones lineales de las funciones logarítmicas. En el caso de bacterias, se eliminaron las OTUs representadas por una sola secuencia en el conjunto completo de datos.

Secuenciación al azar

Secuenciación

El ADN total extraído de las 60 muestras de barro activado y el ADN viral extraído de los 12 sobrenadantes se enviaron al Instituto de Agrotecnología Rosario (INDEAR) para su secuenciación al azar (secuenciación tipo “*shotgun*”). Se realizó una corrida de secuenciación rápida (2 *lanes*) en un sistema HiSeq 1500 de Illumina generando lecturas apareadas (PE) de un largo de 250 pb. Las estadísticas generales de la corrida se detallan en la siguiente tabla:

Procedencia del ADN	Número de Muestra	Tamaño promedio de la biblioteca (pb)	Total de lecturas apareadas (M) obtenidas	Total de bases secuenciadas (Gb)
Barro	01	936	10,148	2,547
Barro	02	614	11,392	2,859
Barro	03	1066	11,552	2,900
Barro	04	697	11,135	2,795
Barro	05	693	11,859	2,977
Barro	06	569	11,206	2,813
Barro	07	623	11,665	2,928
Barro	08	1048	9,864	2,476
Barro	09	701	10,818	2,715
Barro	10	598	11,483	2,882
Barro	11	517	10,558	2,650
Barro	12	517	10,946	2,747
Barro	13	513	11,533	2,895
Barro	14	529	11,826	2,968
Barro	15	523	11,178	2,806
Barro	16	555	10,548	2,647
Barro	17	828	8,575	2,152
Barro	18	802	8,767	2,200
Barro	19	828	11,816	2,966
Barro	20	784	10,620	2,666
Barro	21	805	12,653	3,176
Barro	22	653	11,011	2,764
Barro	23	622	10,430	2,618
Barro	24	699	9,173	2,302
Barro	25	804	9,504	2,385
Barro	26	906	9,831	2,468
Barro	27	1019	9,851	2,473
Barro	28	3360	1,469	0,369
Barro	29	1121	10,099	2,535

Barro	30	851	10,816	2,715
Barro	31	665	17,977	4,512
Barro	32	636	10,515	2,639
Barro	33	1103	12,664	3,179
Barro	34	2003	9,940	2,495
Barro	35	1040	13,299	3,388
Barro	36	691	11,617	2,916
Barro	37	795	10,116	2,539
Barro	38	998	10,102	2,536
Barro	39	778	9,632	2,418
Barro	40	828	11,166	2,803
Barro	41	919	11,387	2,858
Barro	42	894	11,521	2,892
Barro	43	1020	10,679	2,680
Barro	44	1223	8,629	2,166
Barro	45	937	13,247	3,325
Barro	46	1459	10,055	2,524
Barro	47	789	10,049	2,522
Barro	48	813	16,800	4,217
Barro	49	833	9,222	2,315
Barro	50	1144	10,091	2,533
Barro	51	1511	9,127	2,291
Barro	52	1351	9,773	2,453
Barro	53	890	12,594	3,161
Barro	54	835	13,841	3,474
Barro	55	1285	11,451	2,874
Barro	56	1083	13,771	3,457
Barro	57	867	16,614	4,170
Barro	58	710	9,220	2,314
Barro	59	2350	7,288	1,829
Barro	60	856	11,430	2,869
Sobrenadante	01-05	703	11,050	2,773
Sobrenadante	06-10	618	10,687	2,683
Sobrenadante	11-15	653	12,842	3,223
Sobrenadante	16-20	460	10,845	2,722
Sobrenadante	21-25	666	10,400	2,610
Sobrenadante	26-30	872	10,742	2,696
Sobrenadante	31-35	713	11,361	2,852
Sobrenadante	36-40	754	10,296	2,584
Sobrenadante	41-45	720	12,075	3,031
Sobrenadante	46-50	781	10,484	2,631
Sobrenadante	51-55	604	12,369	3,105
Sobrenadante	56-60	554	10,381	2,606

Tabla 4. Estadísticas generales de la corrida de secuenciación al azar de las 60 muestras del metagenoma y de las 12 muestras del metaviroma correspondientes a la serie temporal.

Procesamiento de datos del metagenoma

Anotación automática (MGRAST)

Los archivos en formato fastq de las lecturas *forward* (R1) crudas de los 60 metagenomas de barros activados fueron subidos a Metagenome Rapid Annotation using Subsystem Technology (MG-RAST) y analizadas utilizando los parámetros por defecto (Meyer *et al.*, 2008). Para la notación se utilizaron las bases de datos RefSeq, COG, y Subsistemas de la estructura jerárquica SEED (nivel 1: funciones metabólicas generales).

Reconstrucción de genomas a partir del metagenoma

Las lecturas crudas de las 60 muestras metagenómicas fueron filtradas y recortadas para remover bases ambiguas y asegurar un valor de calidad mínimo de Q=30. Las lecturas filtradas de las 60 muestras fueron combinadas en un solo archivo para su posterior ensamblado. El ensamblado de las secuencias se realizó con MEGAHIT (Li *et al.*, 2015) con múltiple longitud de k-mer (95 a 99) utilizando los parámetros por defecto. Para la identificación de genomas individuales a partir del pool de secuencias ensambladas se eliminaron los *contigs* menores a 1500pb y se agruparon (proceso denominado *binning*) utilizando el programa MetaBAT (Kang *et al.*, 2015). El agrupamiento se realizó utilizando información de la frecuencia de tetranucleótidos y el patrón de abundancia de las secuencias en cada una de las 60 muestras. Este último criterio es posible utilizarlo de manera ventajosa debido a la cantidad de muestras utilizadas (n=60) y a la variación de la abundancia de los distintos genomas a lo largo del tiempo. La limpieza final de los *bines* obtenidos se realizó manualmente considerando que los *contigs* dentro de un *bin* (genoma) deben ser similares en porcentaje de GC, en frecuencia de tetranucleótidos y en cobertura (abundancia) a lo largo de las muestras. Para realizar la limpieza manual de los *bines* se utilizó el programa Cytoscape (Shannon *et al.*, 2003) a fin de visualizar los criterios tenidos en cuenta:

- Porcentajes de GC de los *contigs*.

- Correlación de Pearson entre la distribución de frecuencia de tetranucleótidos para cada par de *contigs* del bin.
- Correlación de Pearson entre los perfiles de abundancia a lo largo de las 60 muestras para cada par de *contigs*.

Para considerar un *bin* como un solo genoma, las secuencias que lo componen debían contener valores similares de GC% (± 5 % del porcentaje de GC promedio), y estar todas interconectadas según correlación positiva de Pearson significativa (tetranucleótidos y abundancia). Posteriormente las secuencias que conformaban cada *bin* fueron extraídas y reensambladas utilizando el programa MEGAHIT según lo descrito anteriormente.

La calidad de los genomas bacterianos ensamblados (completitud y contaminación) se estimó con el programa CheckM (Parks *et al.*, 2015).

Categorización por cobertura

Para estimar la abundancia por muestra, las lecturas filtradas fueron mapeadas a los *contigs* utilizando Bowtie2 (Langmead and Salzberg, 2012). Luego se calculó la cobertura con el programa MetaBAT, relativizando la cobertura a la longitud del *contig*. La abundancia de cada genoma en las 60 muestras se estimó como el promedio de las relaciones entre la cobertura y la longitud de cada uno de los *contigs* que lo forman, como resume la siguiente ecuación, donde i es cada *contig* que compone el genoma y n es el total de los mismos:

$$\text{Abundancia del genoma} = \frac{\sum_i^n \frac{\text{cobertura}_i}{\text{largo}_i}}{n}$$

Asignación taxonómica de genomas de bacteria

La asignación taxonómica de los genomas se realizó utilizando AMPHORA2 (Wu and Scott, 2012). Amphora utiliza 31 proteínas codificantes como genes marcadores de bacterias, y 104 de arqueas para determinar el filotipo genómico y metagenómico. La mayoría de estos son genes de copia única, por lo tanto

este método es adecuado para la estimación de la composición taxonómica de las comunidades bacterianas y de arqueas a partir de los datos de secuenciación.

Estimación del índice de replicación (iRep)

Se aplicó un algoritmo recientemente desarrollado (Brown *et al.*, 2016) que utiliza secuencias de alta calidad de un genoma reconstruido para inferir las tasas de replicación bacteriana en un solo punto de tiempo. El algoritmo calcula un índice de replicación (iRep) a partir de diferencias de cobertura a lo largo del genoma que resultan de la replicación bidireccional del mismo desde un único origen de replicación. Se utilizó este método para los genomas bacterianos reconstruidos, más de 75% completos, calculándose el índice de replicación (iRep) utilizando los parámetros por defecto.

Estimación del número de operones del gen ARNr 16S

A partir de las secuencias crudas de las 60 muestras metagenómicas se recuperaron aquellas lecturas que corresponden a la subunidad pequeña del gen ribosomal (ARNr 16S) utilizando Metaxa2 (Bengtsson-Palme *et al.*, 2015). Posteriormente, empleando las lecturas recuperadas y el programa EMIRGE (Miller *et al.*, 2011), se reconstruyeron las secuencias completas de los genes ARNr 16S, que fueron luego clasificadas utilizando la base de datos Silva (Quast *et al.*, 2013). Una vez reconstruidos los genes ribosomales se identificó el correspondiente a cada genoma mediante los perfiles de abundancia a lo largo de las 60 muestras verificando luego la coincidencia en la clasificación taxonómica del genoma y del gen 16S ARNr. Se utilizó el programa Cytoscape para la visualización de los criterios utilizados en la asignación. El número de copias del operón ribosomal 16S se estimó como la relación de coberturas del 16S y la del resto del genoma. A modo de ejemplo se muestra la siguiente figura en donde se observa el perfil de abundancias a lo largo del tiempo de uno de los genomas y de su correspondiente gen ribosomal:

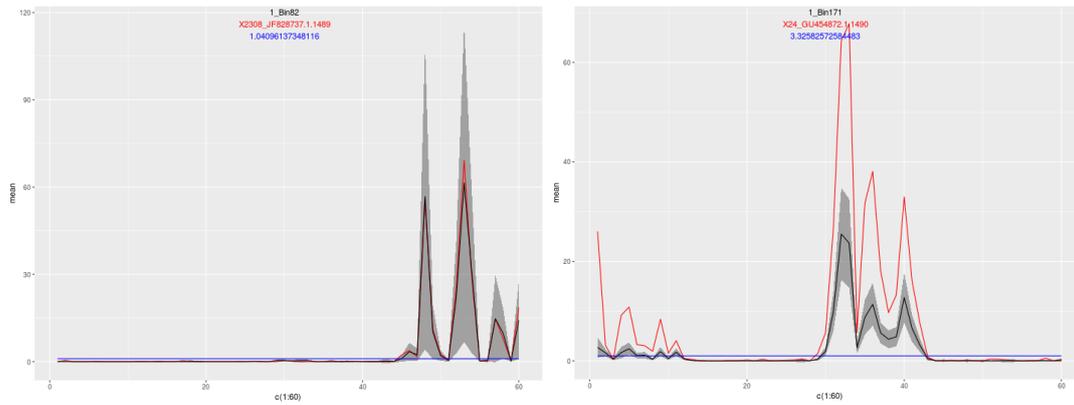


Figura 9. Ejemplos de estimación del número de operones rrn. En los gráficos se observa la cobertura a lo largo del tiempo de todos los *contigs* de un *Bin* (gris, en negro se observa su promedio) y la de su correspondiente secuencia rrn (rojo). El cociente entre estas coberturas corresponde al número de operones rrn, 1 y 3 en los ejemplos mostrados.

Procesamiento de datos del metaviroma

Ensamblado de viromas a partir del metaviroma

Las lecturas crudas de las 12 muestras metaviroma fueron filtradas y recortadas para remover bases ambiguas y asegurar un valor de calidad mínimo de Q=30. Las lecturas filtradas de las 12 muestras fueron juntadas para el posterior ensamblado. El ensamblado de las secuencias fue realizado con MEGAHIT, con múltiple longitud de k-mer (95 a 99) utilizando los parámetros por defecto.

Para estimar la abundancia por muestra, las lecturas filtradas fueron mapeadas a los *contigs* utilizando Bowtie2. La abundancia de fagos fue calculada como el número de lecturas coincidentes y se las dividió por la longitud del *contig*. Este mapeo se realizó tanto con las lecturas de las 12 muestras del metaviroma como con las de las 60 muestras del metagenoma, de manera de estimar la abundancia de los *contigs* virales en las 12 muestras correspondientes al sobrenadante (a partir de las cuales se ensamblaron estos *contigs*) y en las 60 muestras de la serie temporal de donde provienen.

Estimación de la contaminación de ADN bacteriano en metaviroma

Se estimó la contaminación del ADN bacteriano en las 12 muestras correspondientes al metaviroma mediante tres estrategias:

- Se mapearon las secuencias de los genomas de bacteria reconstruidos a partir del metagenoma total sobre las lecturas obtenidas de la secuenciación de las 12 muestras del metaviroma utilizando Bowtie2.
- Se buscaron genes marcadores de bacteria en *contigs* mayores a 2000 pb ensamblados a partir del metaviroma utilizando Amphora2.

Asignación taxonómica de contigs de fagos

La clasificación taxonómica de los *contigs* de fagos se realizó utilizando la base de datos de Grupos Ortólogos de Virus Procariotas (pVOGs) (Kristensen *et al.*, 2013) basada en la presencia de genes pVOGs específicos de familia en los *contigs*, identificados utilizando BLASTp de las secuencias traducidas con GeneMarkS (Besemer, 2001). Esto se realizó para los *contigs* del metaviroma mayores a 5000 pb.

Búsqueda de elementos CRISPRs

Se realizó una búsqueda de elementos CRISPRs en los genomas ensamblados utilizando MetaCRT descargado de <http://omics.informatics.indiana.edu/CRISPR/>. Luego se realizó un alineamiento (BLASTn) entre las secuencias de los espaciadores de los elementos CRISPRs encontrados en los genomas y los *contigs* de fagos mayores a 2000 pb. Sólo los alineamientos con 0 y 1 *mismatch* se consideraron para analizar interacciones de a pares fago-bacteria.

Las secuencias de los genomas a los que se les encontró al menos un espaciador similar en algún *contig* de fago se subieron a <http://omics.informatics.indiana.edu/CRISPRone> (Zhang and Ye, 2017) a fin de buscar y anotar loci pertenecientes a sistemas CRISPRs. Para estos mismos

genomas también se estimó la población de bacterias con o sin CRISPR (población resistente o sensible respectivamente) en las muestras en donde la relación de cobertura y longitud del *contig* del genoma fue mayor a 5. Para ello se mapearon las lecturas de todas las muestras al *contig* conteniendo el CRISPR utilizando Bowtie2. Las abundancias de las poblaciones bacterianas en cada muestra se estimaron de la siguiente manera: a) la población total bacteriana, como la cobertura total del genoma, b) la población bacteriana resistente, como la cobertura de la región que contiene el CRISPR, y c) la población sensible, como la diferencia entre la población total y la resistente.

Alineamiento de secuencias nucleotídicas de genomas y contigs de fagos

Se realizó un alineamiento de las secuencias nucleotídicas (BLASTn) entre los genomas bacterianos ensamblados y los *contigs* mayores a 5000 pb ensamblados a partir del metaviroma. Se consideraron sólo los alineamientos mayores a 1000 pb con un porcentaje de similitud de entre 99 a 100%, y se tuvieron en cuenta sólo los *contigs* de fagos que lograron clasificarse como fagos. Se consideraron profagos los alineamientos coincidentes con la longitud total del *contig* correspondiente al fago. Los *contigs* de los fagos y genomas de bacterias fueron subidos a la plataforma Rapid Annotation using Subsystem Technology (RAST) para su anotación (Brettin *et al.*, 2015; Overbeek *et al.*, 2014; Aziz *et al.*, 2008).

Detección de señales virales en los genomas bacterianos ensamblados

Se utilizó VirSorter 1.0.3 (Roux *et al.*, 2015) para detectar señales virales en todos los genomas ensamblados utilizando la base de datos de VirSorter, Virome db, que incluye una base de datos de RefSeq (con 114.297 proteínas de virus que infectan bacterias o arqueas) más secuencias de virus de muestras de agua dulce, agua de mar e intestino humana, pulmón y saliva.

RESULTADOS

Análisis preliminares de diversidad y dinámica microbianas

Uno de los objetivos específicos de este trabajo fue determinar la composición y diversidad de los diferentes taxones microbianos presentes en las muestras de barros activados a lo largo del tiempo. Previamente a los experimentos de secuenciación masiva se realizaron ensayos con el fin de estudiar de manera preliminar la diversidad e interacciones de comunidades de bacterias y bacteriófagos. Estos experimentos consistieron en electroforesis en gel con gradiente desnaturalizante (DGGE) para bacterias y para bacteriófagos, y búsquedas de pares fago-bacteria mediante técnicas de cultivo.

Diversidad bacteriana

El ensayo de DGGE de bacteria total se realizó sobre 12 muestras tomadas con un intervalo semanal. Este ensayo tuvo dos objetivos principales: por un lado, verificar que el perfil de diversidad bacteriana corresponde adecuadamente a muestras de barros activados, es decir como un proxy de la calidad de las muestras de ADN extraído. Por otro lado, realizar un análisis preliminar de la dinámica de las poblaciones bacterianas, de modo de seleccionar el intervalo de tiempo entre muestras para la serie de tres años. En la Figura 10 se muestra el gel obtenido para las doce muestras analizadas, en donde cada banda representa un taxón; y su intensidad, la abundancia relativa del taxón en la comunidad. A partir de allí se estimaron los índices de diversidad de Shannon de cada muestra y con las distancias euclidianas se realizó un agrupamiento (Figura 11). Los resultados muestran índices de diversidad de las comunidades bacterianas dentro del rango de los comúnmente encontrados en comunidades de barros activados (Jo *et al.*, 2016; Ibarbalz *et al.*, 2014; Ye and Zhang, 2013; Yi *et al.*, 2012). El análisis de agrupamiento indica, en general, que los principales cambios ocurren entre muestras separadas por más de un intervalo semanal. En base a este

resultado, se seleccionó un intervalo quincenal para la toma de muestra subsiguiente.

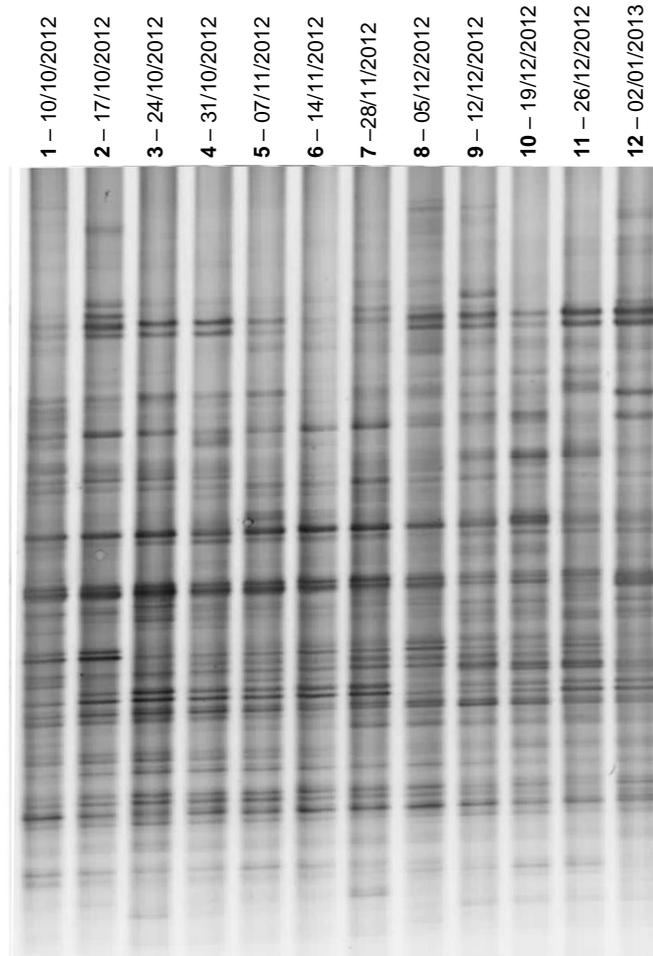


Figura 10. DGGE de bacteria total realizado sobre 12 muestras de barros activados tomadas semanalmente

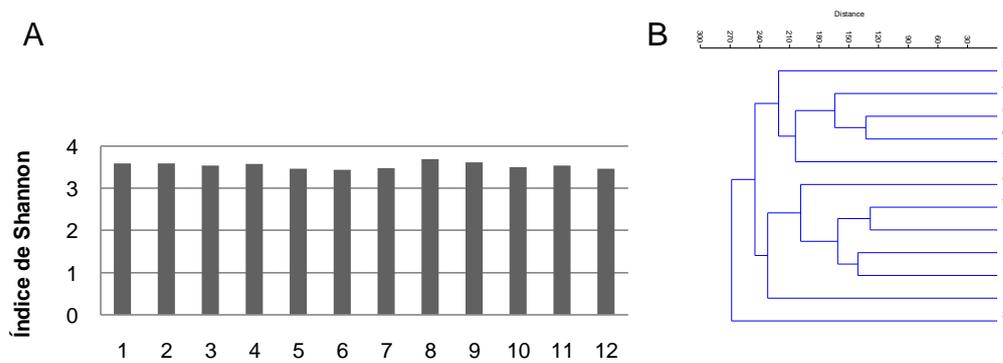


Figura 11. Índices de Shannon **(A)** y agrupamiento realizado a partir de distancias euclídeas **(B)** estimados a partir del DGGE de bacterias en el programa Past (Ryan *et al.*, 2001).

Diversidad de bacteriófagos de la familia *Myoviridae*

Es bien conocida la dificultad del estudio molecular de la diversidad de virus y fagos debido a la ausencia de un marcador filogenético universal, a diferencia del caso de los dominios *Bacteria* y *Eukarya*. Diversos estudios han desarrollado el uso de ciertos genes firma (en general los que codifican proteínas estructurales, genes del metabolismo auxiliar o genes de polimerasa) para investigar la diversidad de grupos de virus específicos (Adriaenssens and Cowan, 2014). Luego de explorar las posibilidades de algunos de estos genes para estudiar la diversidad viral, se puso a punto un protocolo de DGGE para bacteriófagos tipo T4 de la familia *Myoviridae*. Esto permitió obtener una estimación preliminar de la diversidad de una de las familias más abundantes de fagos en muestras de barros activados (Tamaki *et al.*, 2012).

Con el objetivo de confirmar que los amplicones observados en las bandas pertenecen a secuencias correspondientes a proteínas g23 de la cápside de bacteriófagos tipo T4 de la familia *Myoviridae* se construyó una biblioteca de clones con los productos de PCR de 5 de las muestras analizadas en DGGE. Se secuenciaron 17 clones de la biblioteca construida a partir de los amplicones, y se determinó que todos ellos contenían insertos correspondientes a g23 (Tabla 5).

Número de clon	% GC	Longitud (pb)	Resultado del blast (Score máximo)
1, 2, 3, 4 y 5	48,66	409	Uncultured Myoviridae gp23 gene for major capsid protein, partial cds, clone: FW-Ca-14
6, 7 y 8	42,54	409	Uncultured Myoviridae gp23 gene for major capsid protein, partial cds, clone: FW-Ca-15
9 y 10	47,61	397	Uncultured Myoviridae clone W415 major capsid protein (g23) gene, partial cds
11 y 12	52,36	403	Uncultured Myoviridae clone W415 major capsid protein (g23) gene, partial cds
13	45,31	448	Uncultured Myoviridae clone s2-8 major capsid protein-like (g23) gene, partial sequence
14	49	451	Uncultured Myoviridae clone 11-6 major capsid protein-like (g23) gene, partial sequence
15	49,46	372	Uncultured Myoviridae clone W411 major capsid protein (g23) gene, partial cds
16	48,31	445	Uncultured Myoviridae clone 11-3 major capsid protein-like (g23) gene, partial sequence
17	46,48	454	Uncultured Myoviridae clone 9-15 major capsid protein-like (g23) gene, partial sequence

Tabla 5. Características de las secuencias de los productos de PCR clonados utilizados en el protocolo de DGGE para bacteriófagos de la familia *Myoviridae*

Fujii colaboradores (Fujii *et al.*, 2008) habían determinado que el DGGE era una herramienta poderosa para obtener diferentes fragmentos de genes g23 a partir

suelos. Sin embargo, postularon que los análisis de diversidad a partir de los patrones de DGGE estaban sesgados como consecuencia de las diferentes longitudes de los fragmentos. Teniendo en cuenta esto, el DGGE de este grupo de virus permite una obtener una estimación, aunque sesgada, de su diversidad. Se realizó un ensayo de DGGE sobre las 12 muestras de ADN viral extraído de los sobrenadantes correspondientes a las muestras de la serie temporal que se enviaron posteriormente para su secuenciación al azar.

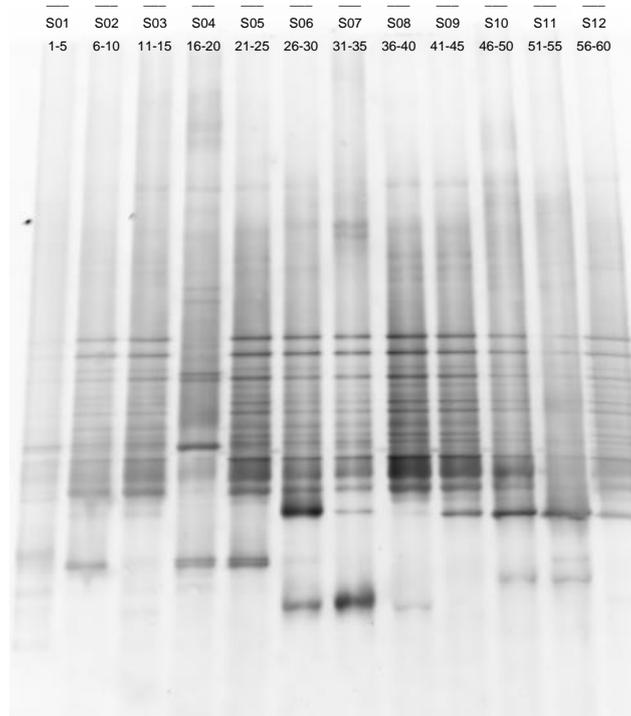


Figura 12. DGGE para fagos de la familia *Myoviridae*

Con este ensayo se pudo verificar la calidad del ADN extraído y se determinó en forma preliminar que el grupo de fagos tipo T4 de la familia *Myoviridae* muestra alta diversidad y que existen variaciones de distinta extensión entre muestras tomadas a distintos tiempos. Estos resultados resaltan el uso de la técnica de DGGE como un ensayo rápido y sencillo, de valor para el análisis de este grupo de bacteriófagos abundantes en muestras ambientales.

Búsqueda de hospederos de fagos con técnicas de cultivo

Se aislaron 30 cultivos bacterianos puros en medio R2A a partir de las muestras de barros activados. En aproximadamente 30% de esos aislamientos bacterianos (11 de 30) se encontró al menos una playa de lisis en el ensayo de placa. Es decir, de las bacterias que fueron capaces de crecer en medio R2A en las condiciones de cultivo utilizadas, el 30% de ellas posee al menos un fago lítico en 1 ml de sobrenadante. El resultado del alineamiento de las secuencias completas del gen ARNr 16S de cada una de estas bacterias hospederas (denominadas H1-H11) de fagos líticos halladas en cultivo se muestra en la siguiente tabla.

Hospedero del fago	Número de playas de lisis	Características de playas de lisis	Identidad
H1	2	Con halos	<i>Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxallaceae, Enhydrobacter</i>
H2	2	Con halos	<i>Proteobacteria, Gammaproteobacteria, Pseudomonadales, Moraxallaceae, Enhydrobacter</i>
H3	2	Pequeñas	<i>Proteobacteria, Alphaproteobacteria, Rhizobiales</i>
H4	6	Diversos tamaños. Algunas con halos	<i>Proteobacteria, Alphaproteobacteria, Caudobacterales, Caudobacteraceae, Phenylobacterium</i>
H5	24	Tamaños parejos	<i>Proteobacteria, Alphaproteobacteria, Rhizobiales, Xanthobacteraceae, Xanthobacter</i>
H6	38	Con halos y sin halos	NA
H7	400	-	NA
H8	13	Con halos	<i>Proteobacteria, Alphaproteobacteria, Rhizobiales, Bradyrhizobiaceae, Bosea</i>
H9	1	-	<i>Proteobacteria, Gammaproteobacteria</i>
H10	Incontables	-	<i>Proteobacteria, Gammaproteobacteria</i>
H11	31	-	<i>Proteobacteria, Betaproteobacteria, Burkholderiales, Comamonadaceae, Acidovorax</i>

Tabla 6. Características de las playas de lisis observadas e identidad de cada una de las bacterias hospederas de fagos aisladas en medio de cultivo R2A a partir de las muestras

Las bacterias hospederas de fagos halladas en cultivo pertenecen a las clases *Alphaproteobacteria*, *Betaproteobacteria* y *Gammaproteobacteria*. Shapiro y colaboradores encontraron 18 bacterias hospederas de fagos de un total de 68 aislamientos en cultivo analizados, las cuales pertenecían a tres familias de esta misma clase de proteobacterias (Shapiro *et al.*, 2010). Esto no significa necesariamente que sólo las proteobacterias son hospederos de fagos en este

tipo de ecosistemas, sino que probablemente se deba al hecho de que el filo *Proteobacteria* es seleccionado preferencialmente por su capacidad de crecer fácilmente con técnicas clásicas de cultivo.

Estos resultados representan sólo algunos ejemplos de fagos líticos cuyos hospederos son cultivables en el medio de cultivo utilizado. Sin embargo, esta metodología se encuentra limitada debido a la dificultad de cultivar una proporción grande de bacterias a partir de barros activados (Amann *et al.*, 1995). Más allá de estas limitaciones, es interesante tener el conocimiento de hospederos de fagos líticos presentes en las muestras y tener la posibilidad de contar con los aislamientos de los fagos y sus hospederos. En algunos casos se insinúa la posibilidad de que exista más de un fago para un hospedero, debido a las características diferentes de las playas de lisis observadas (como en H4 y en H6). Debido a que se percibió que el título de los fagos se perdía durante el almacenamiento de las muestras del sobrenadante, dificultando la cuantificación a lo largo del tiempo, se decidió no continuar con esta estrategia. En su reemplazo se optó por realizar una investigación más profunda y abarcadora de interacciones fago-hospedero a través del estudio del metagenoma y del metaviroma (PARTE IV).

Evaluación de sesgos de las regiones secuenciadas en la secuenciación de amplicones.

El uso de técnicas independientes de cultivo basadas en filogenias moleculares del gen ARNr 16S ha posibilitado ampliar enormemente el conocimiento de la diversidad bacteriana en diferentes ecosistemas. Los avances más importantes se obtuvieron a partir de estudios basados en secuenciación de amplicones, en los que un par de cebadores universales de PCR se dirigen a regiones conservadas que flanquean alguna de las nueve regiones hipervariables (V1-V9) del gen ARNr 16S (Figura 13).

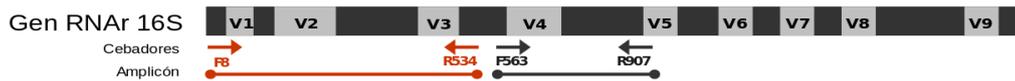


Figura 13. Representación de las 9 regiones variables del gen rRNA 16S y cebadores utilizados

Una limitación de esta técnica es que los resultados se encuentran afectados por el sesgo asociado con la PCR (Reysenbach *et al.*, 1992; Polz and Cavanaugh, 1998; Smith *et al.*, 2002; Sipos *et al.*, 2007; Hong *et al.*, 2009). Existen diversos estudios en los cuales se han evaluado los sesgos que otorgan los distintos conjuntos de cebadores (Baker *et al.*, 2003; Klindworth *et al.*, 2013; Fredriksson *et al.*, 2013) y también se han puesto a prueba con el fin de intentar seleccionar los pares que ofrecen una estimación más precisa de la diversidad (Kumar *et al.*, 2011; Cai *et al.*, 2013). Como en este trabajo en particular se intenta comprender la dinámica temporal bacteriana interesa evaluar este sesgo en la caracterización de la dinámica temporal de las comunidades bacterianas bajo la hipótesis de que a pesar de que la amplificación de diferentes regiones del gen rRNA 16S produce una estimación sesgada de la composición de la comunidad, permite cuantificar de manera precisa de la dinámica temporal de la población. Para probar esta hipótesis, se realizó una secuenciación de amplicones correspondientes a dos regiones variables diferentes, V1-V3 y V4, del gen rRNA 16S en sendos duplicados de muestras de barros activados de una serie de tiempo, y se comparó el rendimiento de ambos pares de cebadores en su capacidad para describir cambios en la diversidad bacteriana.

Si bien la abundancia relativa de cada taxón bacteriano resultó diferente para los dos conjuntos de datos de secuenciación, ambos compartieron un grupo de taxones centrales que se observaron a través de 12 muestras tomadas con un intervalo mensual (Ibarbalz, 2016). Un caso particular fue el observado en las bacterias filamentosas del género *Thiothrix*, donde los cambios en su abundancia relativa causaron una gran distorsión en los índices de diversidad de dos de las doce muestras en el conjunto de datos de secuenciación de la región V1-V3. No obstante, la abundancia relativa de *Thiothrix* en el tiempo obtenida con los datos de secuenciación de ambas regiones correlacionó

significativamente con la estimación de su abundancia determinada por FISH (Figura 14).

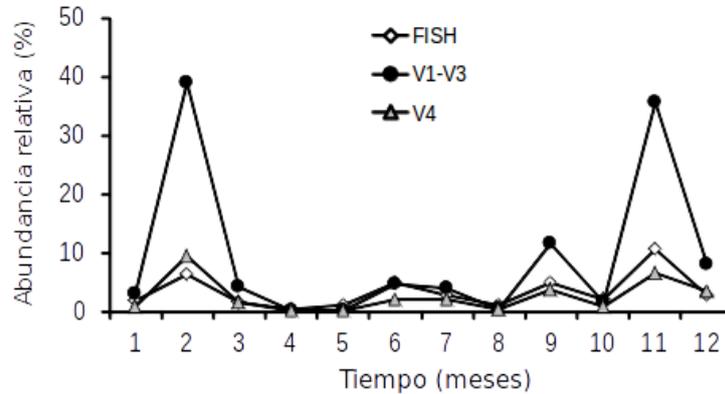


Figura 14. Abundancia relativa de *Thiothrix* estimada por FISH y por secuenciación de amplicones de la región V1-V3 y V4 del gen ARNr 16S. Los coeficientes de correlación de Pearson entre la abundancia calculada con FISH y por secuenciación de amplicones son 0,87 (región V1-V3) y 0,81 (región V4), en ambos casos con p-valores menores a 0,01.

Los análisis de la dinámica de las comunidades microbianas para las dos regiones estudiadas produjeron patrones comparables. Por un lado, los cambios observados dentro de las comunidades en intervalos de tiempo fijos fueron similares y, por otro, también lo fue la tasa de reemplazo de especies estimada a partir de la relación especie-tiempo. Esto demostró que a pesar del sesgo introducido por la PCR, es posible utilizar las regiones analizadas para evaluar cuantitativamente la dinámica y para describir cualitativamente los taxones de la comunidad bacteriana.

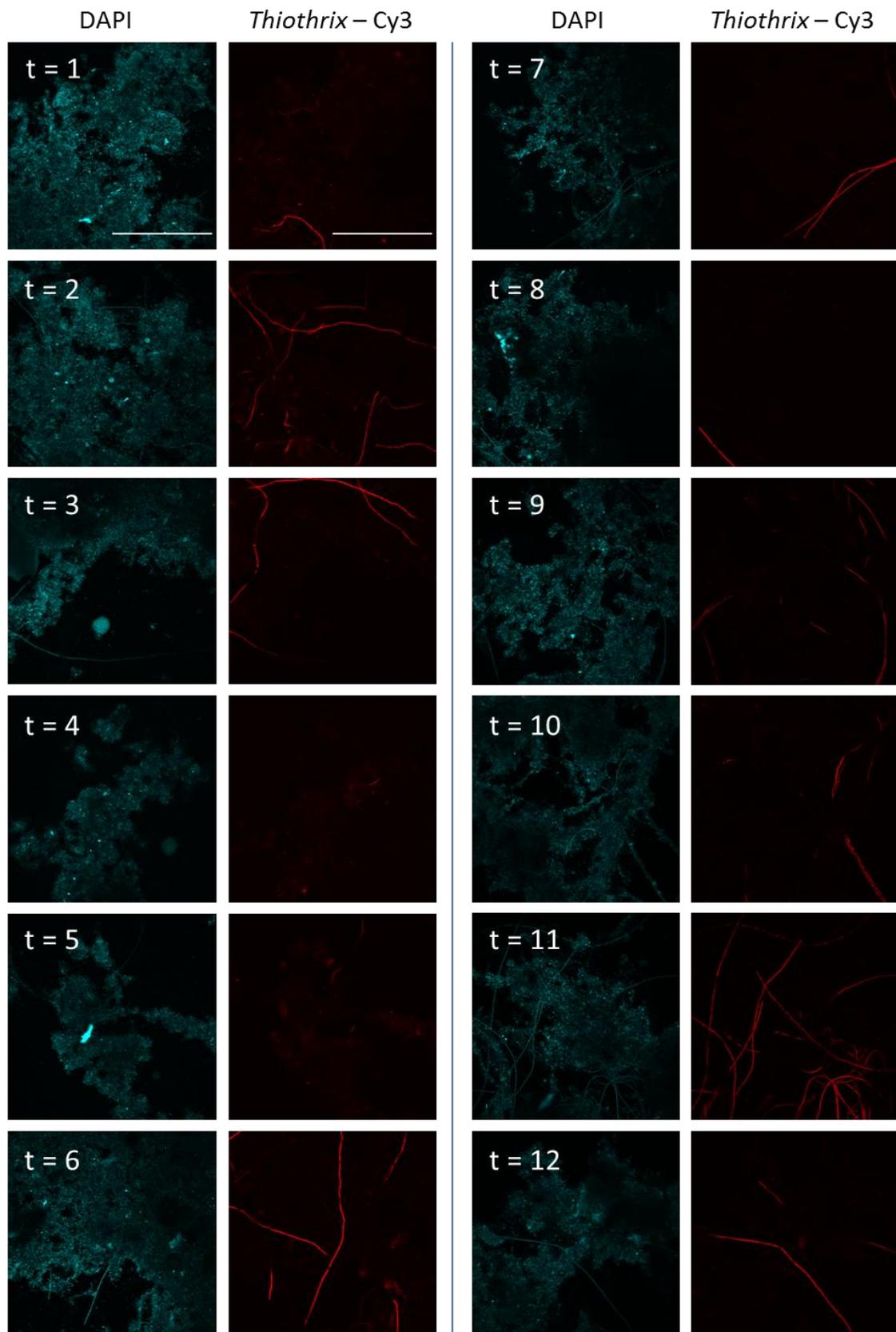


Figura 15. Imágenes de FISH para *Thiothrix* representativas de cada una de las 12 muestras

Diversidad y dinámica microbianas

La diversidad y dinámica temporal, a lo largo del período de tres años, de bacterias se evaluó por secuenciación de las regiones hipervariables V3-V4 del gen ARNr 16S del ADN extraído del barro de las muestras. Los organismos eucariotas se cuantificaron por microscopía.

Estructura de la comunidad eucariota

Una colección representativa de los organismos eucariotas presentes en las muestras se puede observar en las imágenes de la siguiente figura:

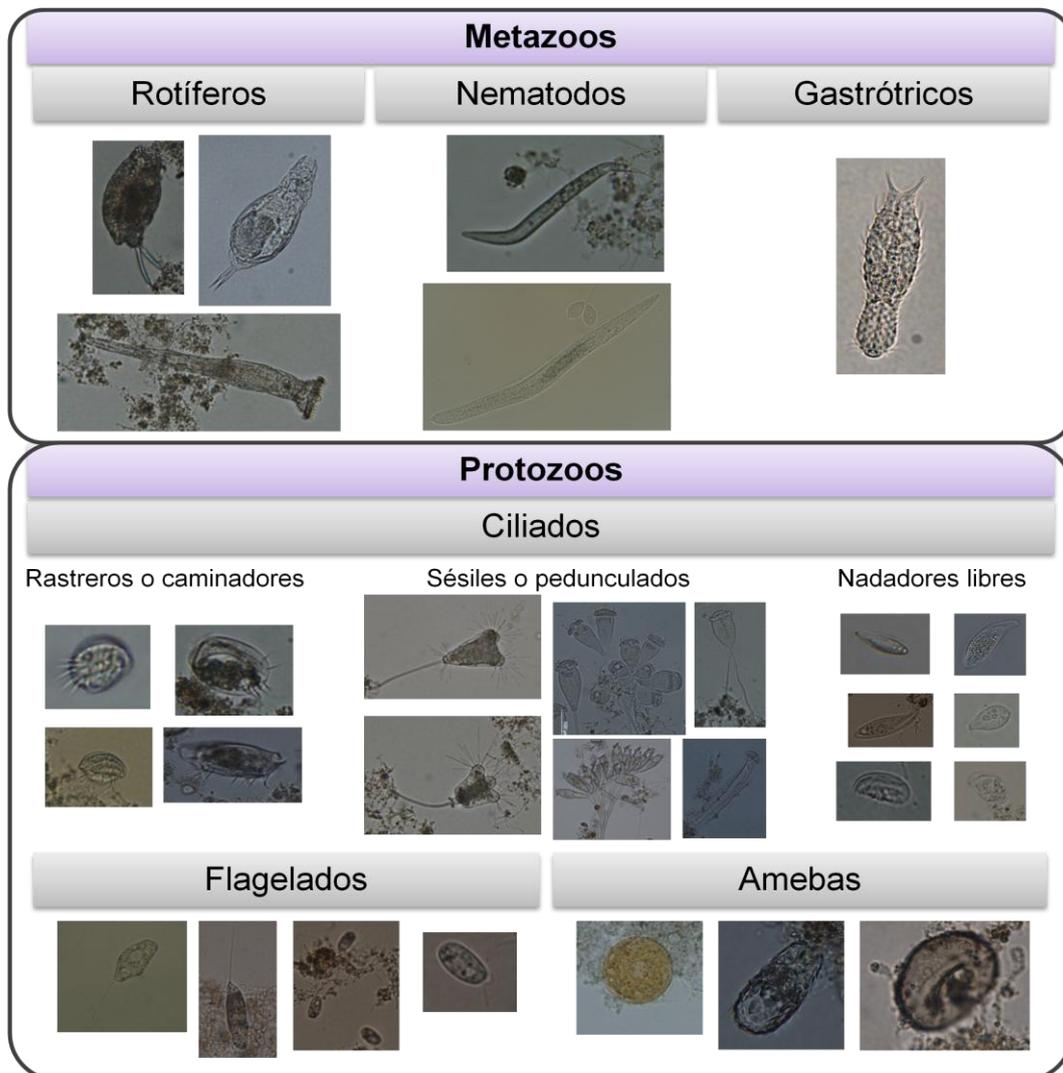


Figura 16. Micrografías de organismos eucariotas tomadas sobre muestras frescas de barros activados a lo largo del período de muestreo en microscopio óptico (200X y 400X). No se muestran las escalas.

La Figura 17 B muestra la distribución de abundancias de protozoos y metazoos en las muestras examinadas a lo largo de tres años. Los protozoos más abundantes y más frecuentemente observados en el total de las muestras fueron los ciliados del género *Epistylis*, con una abundancia relativa promedio de 40,5% y presentes en el 96,7% de las muestras (n=60) y *Aspidisca*, con un 14,8% de abundancia promedio y presente en 55 de las 60 muestras. Estas observaciones son consistentes con la bibliografía, que indica que la comunidad de protozoos de una planta de barros activados en buen funcionamiento está dominada por organismos de las subclases *Peritrichia* (como *Vorticella* y *Epistylis*) e *Hypotrichia* (como *Aspidisca*) (Madoni, 2011).

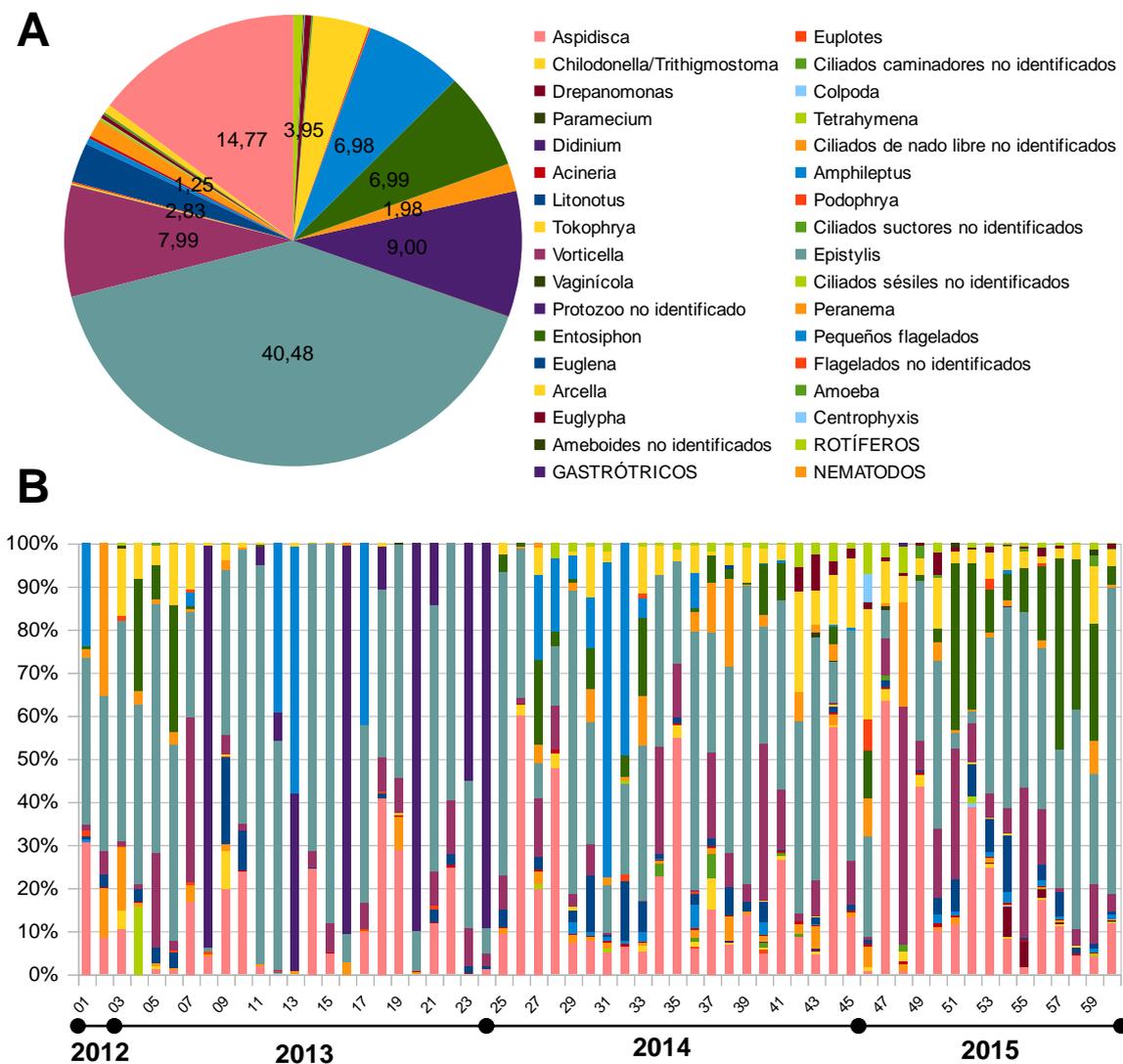


Figura 17. Abundancias relativas de organismos eucariotas calculadas por microscopía. **A)** Abundancia relativa porcentual en el total de las muestras observadas. **B)** Distribución de abundancias de organismos eucariotas en las 60 muestras de la serie temporal

Las amebas de los géneros *Arcella* y *Euglypha* fueron principalmente observadas durante el período de tiempo de retención de sólidos alto. Esto también concuerda con lo reportado en la literatura, que relaciona la presencia de estos microorganismos con un buen desempeño del sistema de tratamiento asociado con bajo contenido de materia orgánica, suficiente oxigenación y nitrificación (Foissner, 2016).

Durante el período de fluctuaciones operacionales, entre verano de 2012/2013 y la primavera de 2013, en el que los valores de DQO de salida fueron mayores al promedio de todo el período investigado (Figura 18), se observó un protozoo pequeño de aproximadamente unos 10 μm , con características particulares de movilidad, rápida y con movimientos espiralados. Su presencia se asoció a valores bajos de OD y posibles condiciones sépticas, pero lamentablemente no pudo ser identificado.

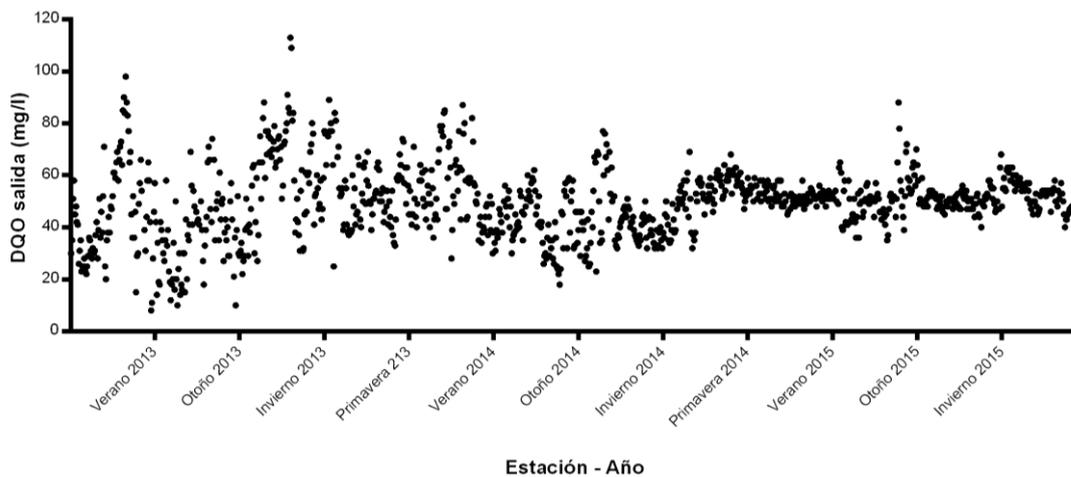


Figura 18. Valores de DQO diarios de salida de planta.

A partir de la abundancia relativa de los organismos se estimó el índice de diversidad de Shannon (Figura 19). De estos resultados es posible concluir que la diversidad de microorganismos eucariotas fue acentuadamente menor durante el período perturbado.

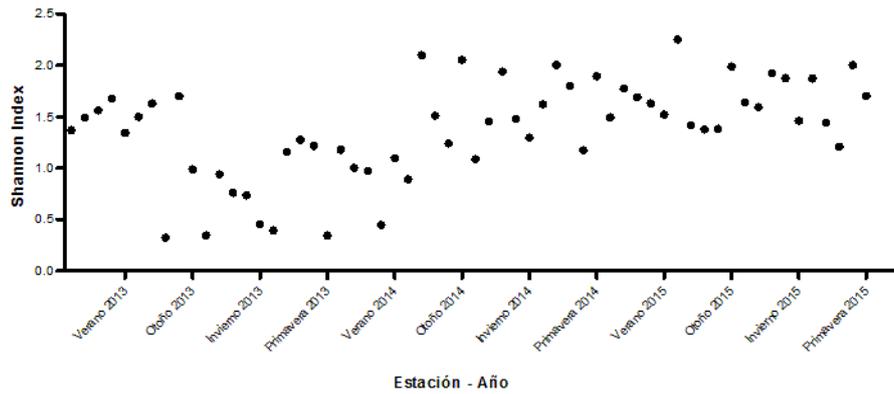


Figura 19. Índices de diversidad de Shannon estimados para organismos eucariotas en todas las muestras (semanales) tomadas durante el período investigado

Las observaciones de protozoos al microscopio fueron utilizadas también para estimar el índice biótico de barros (SBI) (Figura 20). Este índice se estima teniendo en cuenta el número total de organismos por unidad de volumen, la riqueza de especies, el grupo o los grupos predominantes y la cantidad de pequeños flagelados presentes. Los valores de SBI agrupan en cuatro clases que representan la calidad biológica del barro activado. Se ha postulado que la clase I (SBI=8-10) es indicadora de un muy buen rendimiento, de un barro bien colonizado, estable y de excelente actividad biológica, y las clases II (SBI = 6-7), III (SBI=4-5) y IV (SBI=1-3) serían indicadoras de calidades sucesivamente más deficientes y con menores rendimientos de depuración (Madoni, 1994).

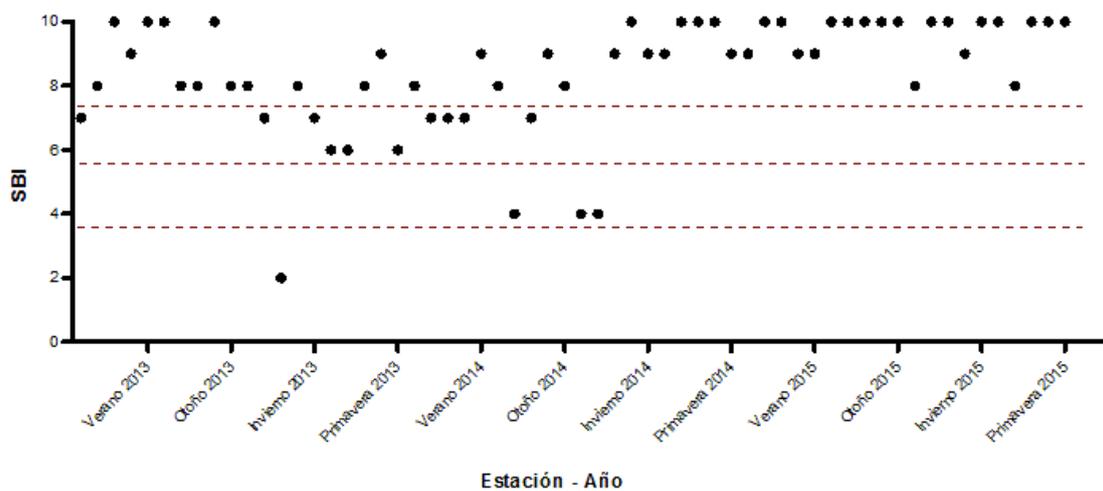


Figura 20. Índice biótico de barros calculado sobre las muestras frescas de la serie temporal

Estructura de la comunidad bacteriana

En la Figura 21 se observa la distribución de abundancias a nivel de filos bacterianos, con el filo *Proteobacteria* separado en clases.

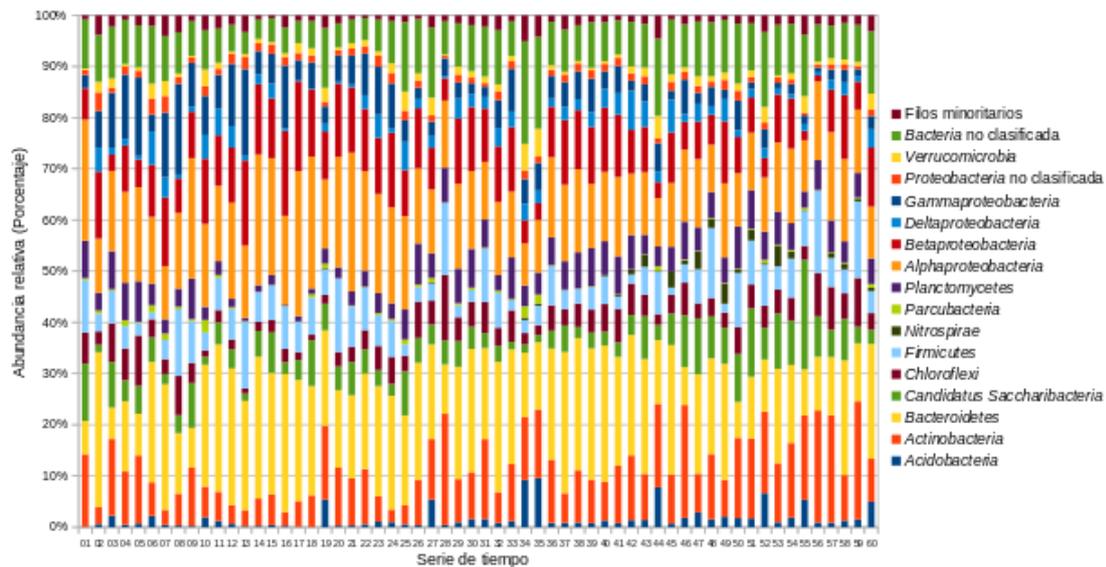


Figura 21. Abundancia relativa de filos (y clases de proteobacterias) en las muestras de la serie temporal de barros activados

El filo bacteriano más abundante en las muestras fue *Proteobacteria* (44%), seguido por *Bacteroidetes* (19%) y *Actinobacteria* (10%). Las clases de proteobacterias más abundantes fueron *Alphaproteobacteria* (16%) y *Betaproteobacteria* (10%). Si bien existe consistencia en la estructura de la comunidad a lo largo del tiempo, es posible percibir cambios en las abundancias relativas de ciertos filos según los períodos operacionales discriminados. El filo *Nitrospirae* aumenta en el período de alto TRS, junto con un incremento en la abundancia relativa de otros filos como *Acidobacteria*, *Actinobacteria* y *Candidatus Saccharibacteria*. La abundancia relativa de *Gammaproteobacteria* y *Bacteroidetes* es mayor en los períodos de bajo TRS. Particularmente durante el período de fluctuaciones operacionales la abundancia relativa de los filos *Chloroflexi* y *Planctomycetes* se ve reducida, mientras que lo contrario sucede con la de las clases *Alphaproteobacteria* y *Betaproteobacteria*.

Relación especie-tiempo

La relación especie-tiempo describe cómo la riqueza de especies de una comunidad aumenta con el lapso de tiempo durante el cual se observa. Esta medida se aplicó para cuantificar la dinámica temporal de las comunidades de bacterias (Figura 22) y de eucariotas. La riqueza de las especies aumentó con el tiempo de acuerdo con el modelo de ley de potencia $S=cT^w$, con un exponente $w=0.20$ para bacterias y $w=0.17$ para eucariotas.

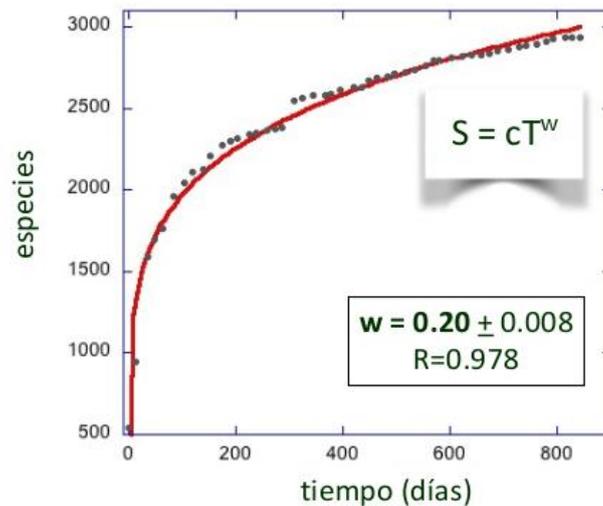


Figura 22. Relación especie-tiempo. Riqueza de especies (OTUs) acumulada en función del tiempo.

Las tasas de reemplazo comparables y los resultados de las redes de interacción sugieren que los organismos eucariotas y las bacterias se encuentran afectados de manera similar por el ambiente.

Influencia de factores abióticos en la estructura de las comunidades

Se realizó un análisis de similitud local entre los taxones bacterianos más abundantes, dando lugar a una red de interacción. Las correlaciones locales significativas se pueden observar en la Figura 23.

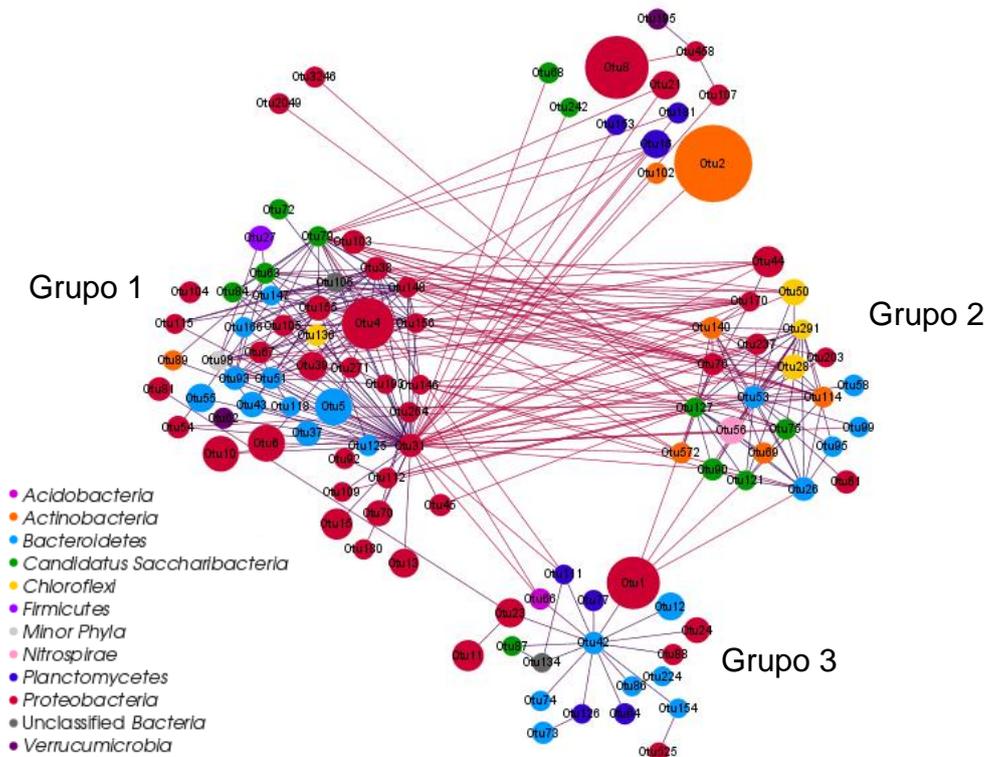


Figura 23. Red de interacción entre taxones bacterianos (OTUs). Las líneas azules indican correlaciones positivas y las rojas, negativas. El tamaño de los nodos esquematiza la abundancia relativa de cada OTU en el total de muestras.

El LSA permitió establecer patrones de asociación en el tiempo, con las especies bacterianas (OTUs) agrupadas según su co-ocurrencia, en tres grupos principales. Para comprender cuáles son los taxones bacterianos abundantes en cada uno de los períodos se graficó la suma de la abundancia relativa de los taxones que forman cada uno de los tres grupos de la red de la Figura 23.

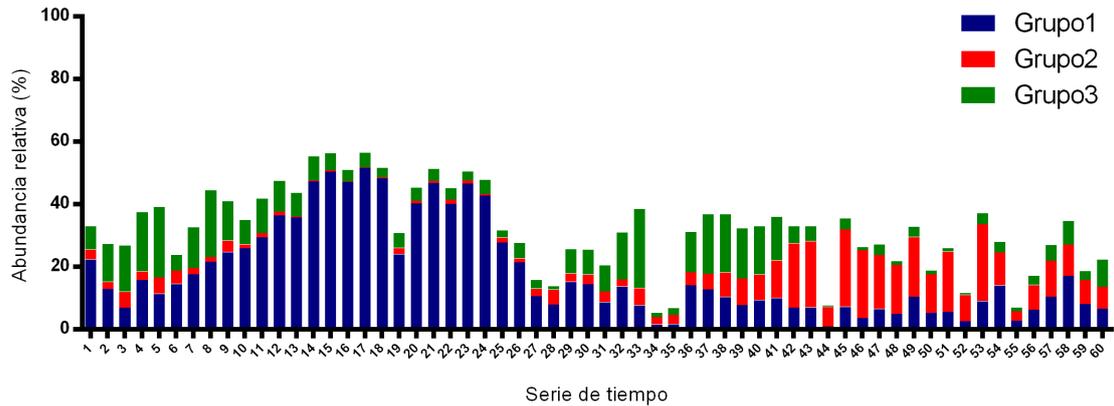


Figura 24. Abundancia relativa de OTUs abundantes en los grupos 1,2 y 3 definidos a partir de la red de interacción de la figura 23.

Como se puede observar, las especies correspondientes al grupo 1, principalmente pertenecientes a los filos *Bacteroidetes* y *Proteobacteria*, son abundantes en el período con fluctuaciones operacionales. Las especies bacterianas del grupo 2 son abundantes en el período estable de alto TRS; y las del grupo 3, predominan en el período intermedio, caracterizado como período estable de relativamente bajo TRS.

De acuerdo a los períodos operacionales anteriormente definidos a partir de las características operacionales (Figura 8), podemos asociar estos tres grupos, y las muestras, a cada uno de ellos:

- PP: período perturbado. Período de bajo TRS, caracterizado por perturbaciones operacionales. (Muestras 10 a 25 que forman el grupo 1)
- PE: período estable (Muestras 42 a 60). Período de alta estabilidad y alto TRS. (Muestras 42 a 60 que forman el grupo 2)
- PI: período intermedio. Período de estabilidad y bajo TRS. (Muestras intermedias, que forman el grupo 3)

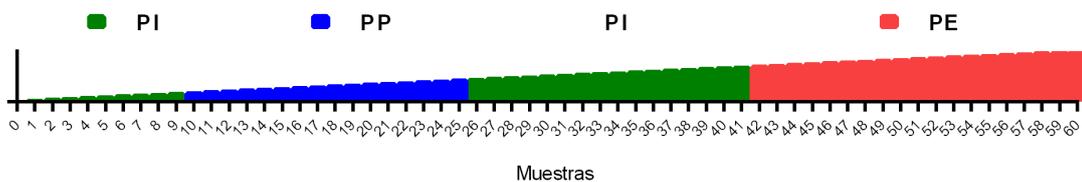


Figura 25. Períodos definidos según las características observadas en función de las muestras de la serie temporal. El gradiente representa de manera esquemática el incremento en el TRS.

La abundancia total de todos los filos bacterianos y clases del filo *Proteobacteria* de las muestras seleccionadas del PP (grupo 1) y del PE (grupo 2) se compararon en el programa STAMP (Parks *et al.*, 2014), realizando un test estadístico (test t de Welch) a fin de evaluar las diferencias significativas entre estos grupos taxonómicos (Figura 26). En el PP los filos *Proteobacteria* y *Bacteroidetes* resultaron ser significativamente más abundantes que en el PE. Estos filos, usualmente clasificados como estrategias-r por su alta tasa de reproducción y oportunismo ecológico, suelen competir favorablemente en ambientes perturbados o inestables (Fierer *et al.*, 2007).

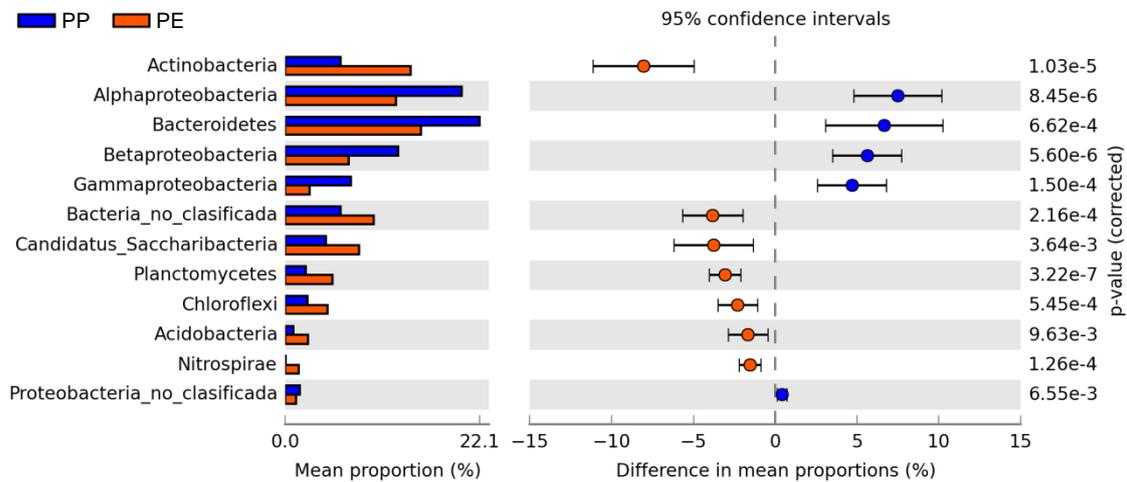


Figura 26. Filos bacterianos y clases de *Proteobacteria* con diferencias significativas de abundancia en los períodos comparados.

Por el contrario en el PE se observó una proporción significativamente mayor de especies pertenecientes a los filos *Actinobacteria*, *Planctomycetes*, *Candidatus Saccharibacteria* (TM7), *Chloroflexi*, *Actinobacteria* y *Nitrospirae*. Algunos de estos taxones han sido relacionados con baja velocidad de crecimiento y alta afinidad por el sustrato, es decir estrategias-K, que suelen predominar en ambientes estables (Fierer *et al.*, 2007).

Los resultados del LSA entre los taxones bacterianos y los organismos eucariotas se muestran en la siguiente figura.

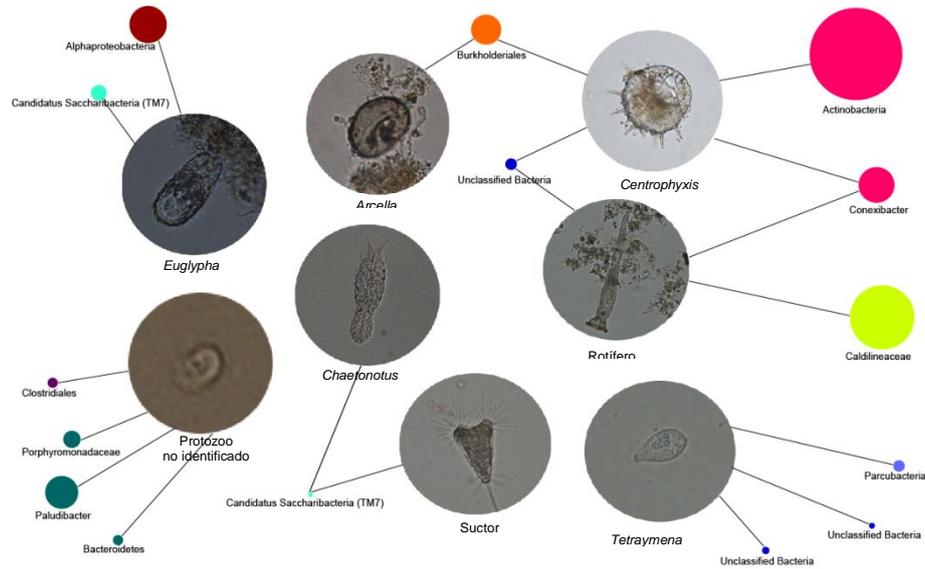


Figura 27. Rede de interacção entre organismos eucariotas e espécies bacterianas (OTUs). Os tamanhos de círculos de espécies bacterianas representam sua abundância relativa em las muestras y colores similares indican pertenencia a un mismo filo.

Si estos resultados se combinan con los del LSA realizado para taxones bacterianos incluyendo a los eucariotas en la red de la Figura 23, los rotíferos y amebas (*Arcella*, *Euglypha* y *Centrophyxis*) formarían parte del grupo 2, que representa al PE; y el protozoo no identificado, al grupo 1, PP. Hay otros eucariotas que se relacionan con algunas otras especies bacterianas, pero no caracterizan a ningún período en particular. Ya se había mencionado que las amebas y rotíferos habían sido mayormente observados en el período de alto TRS, y que el protozoo no identificado se lo asociaba a períodos de anoxia (existentes durante el PP). En todos los casos es importante aclarar que la co-ocurrencia no permite afirmar que exista una relación de depredación protista-bacteria.

Con el fin de confirmar que la estructura de la comunidad bacteriana discriminada en tres grupos de taxones es el resultado de las presiones de selección ambientales, se realizó un análisis de similitud local entre las variables operativas y las abundancias sumadas de las especies capaces de diferenciar cada uno de los grupos. La abundancia relativa de las especies

presentes en el “grupo 2” correlacionan negativamente con el TRS (p -valor < 0.01). Por su parte, el índice volumétrico de lodos (IVB) mostró una correlación positiva significativa con las especies de este grupo, y negativa con las del “grupo1”. El IVB indica el volumen que ocupa un gramo de barro, y es utilizado rutinariamente en las PTE como indicador de las propiedades de sedimentación del barro. Valores altos, generalmente mayores a 150ml/g pueden causar problemas de sedimentación (Jenkins *et al.*, 2004).

El análisis de coordenadas principales restringido realizado se muestra en la Figura 28. El test de permutación mostró que la temperatura y el TRS poseen una correlación significativa ($p < 0.01$) con la comunidad bacteriana.

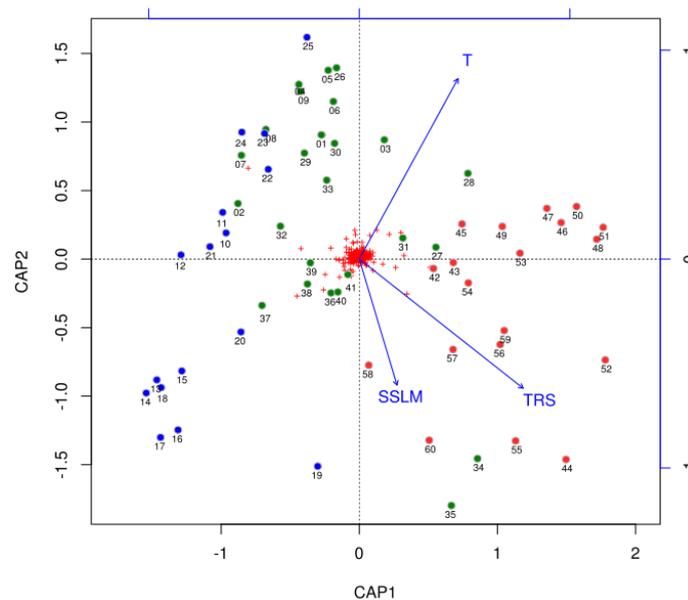


Figura 28. Análisis de coordenadas principales restringido. Las muestras se colorearon según los períodos definidos en la Figura 24.

Estos resultados confirman el TRS es un determinante de la estructura de la comunidad bacteriana.

Análisis metagenómico

Con la finalidad de comprender la adaptación de las comunidades bacterianas a las variaciones operacionales y ambientales durante el proceso de tratamiento, nos propusimos investigar las bases genómicas de las diferentes estrategias de crecimiento bacterianas. Para ello se realizó la secuenciación masiva al azar, seguida por el análisis de las funciones de los metagenomas y finalmente la reconstrucción y el análisis de genomas ensamblados a partir de los metagenomas.

Funciones de los metagenomas

En primer lugar se realizó un análisis de coordenadas principales a partir de las funciones anotadas de los 60 metagenomas utilizando la base de datos COG (nivel 2) y Subsistemas (nivel 1).

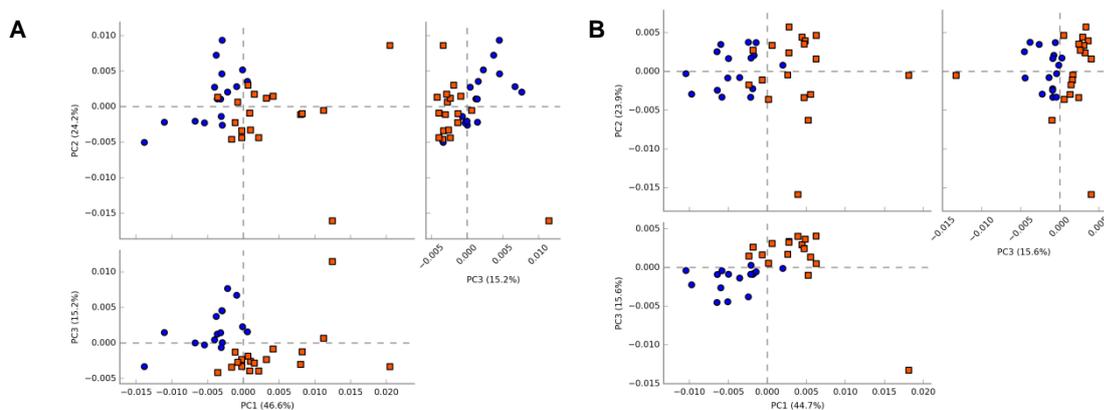


Figura 29. Análisis de coordenadas principales (PCA) realizado en el programa STAMP sobre las funciones de los períodos PP (azul) y PE (rojo) anotadas en MGRAST a partir de las secuencias crudas de los metagenomas utilizando distintas bases de datos: **A)** COG (nivel 2) y **B)** Subsistemas (nivel 1).

Este resultado mostró que las funciones presentes en los metagonomas son capaces de discriminar los períodos PP y PE (Figura 29). Las funciones que muestran diferencias significativas en ambos períodos se observan en la Figura 30.

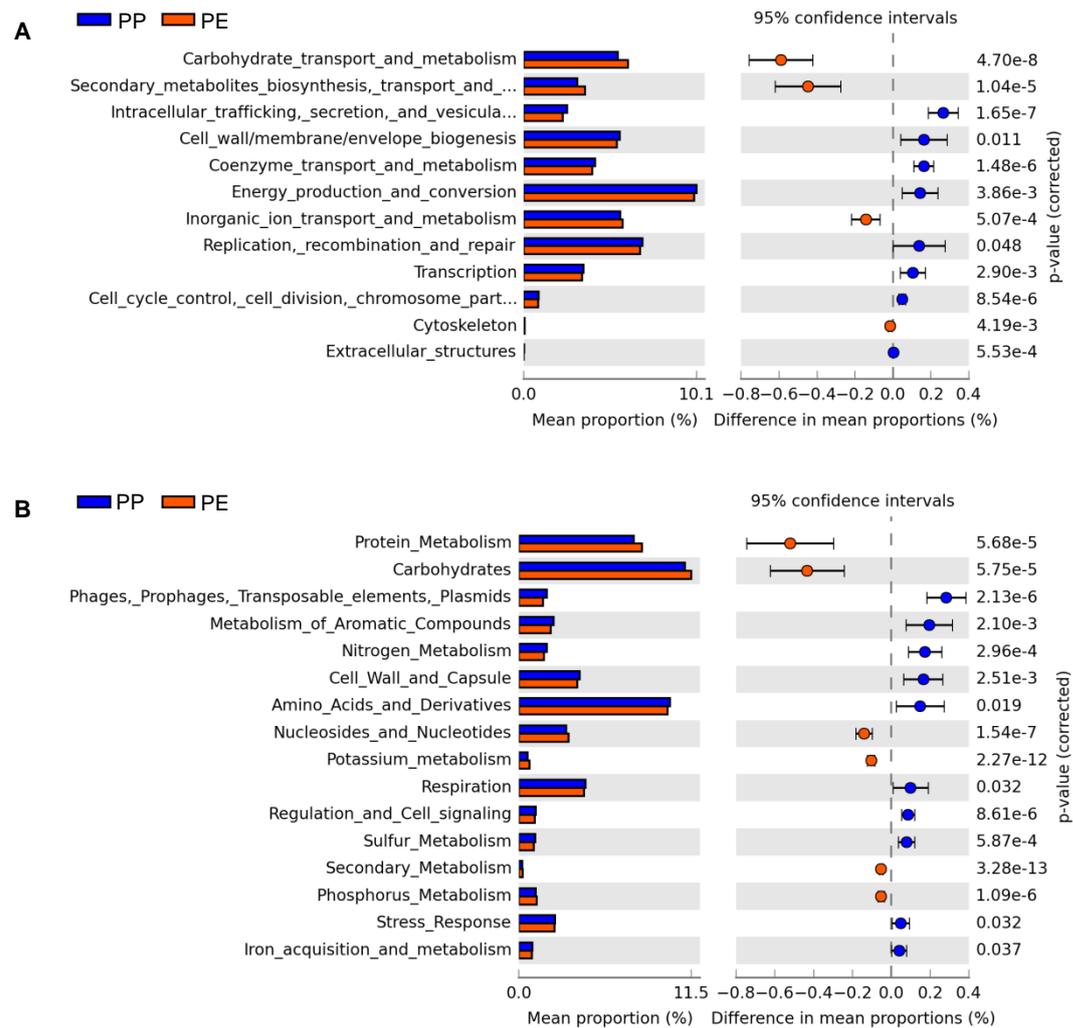


Figura 30. Funciones con diferencias significativas en los períodos PP (barras azules) y PE (barras rojas) anotadas en MGRAST a partir de las secuencias crudas del metagenoma utilizando distintas bases de datos: **A)** COG (a nivel 1) y **B)** Subsistemas (a nivel 2).

Los metagenomas de barros activados pertenecientes al PE mostraron estar enriquecidos principalmente en funciones relacionadas con el metabolismo de proteínas, carbohidratos y con el metabolismo secundario.

Las funciones sobrerrepresentadas en los metagenomas del PP fueron sobre todo las involucradas en la replicación, transcripción, producción y conversión de energía, entre algunas otras. También fue significativamente más abundante la respuesta a estrés y el metabolismo de sulfuros, que podrían estar reflejando la adaptación a las perturbaciones y a etapas de anoxia, respectivamente.

Las funciones involucradas en el metabolismo secundario han sido observadas con mayor abundancia en genomas de bacterias adaptadas a condiciones oligotróficas (estrategas-K); y una alta proporción de COG relacionados con la transcripción, en genomas de bacterias copiotróficas (estrategas-r) (Lauro *et al.*, 2009).

Reconstrucción de genomas a partir de metagenomas

Se pudieron reconstruir 257 genomas bacterianos utilizando la información combinada de los 60 metagenomas de barros activados. Un total de 67 genomas reconstruidos se encuentran más de 90% completos, entre ellos 63 que poseen menos de 5% de contaminación, por lo cual pueden considerarse genomas de alta calidad (Bowers *et al.*, 2017). En la Figura 31 se observa el grado de completitud y de contaminación de todos los genomas (bins) ensamblados, estimado con CheckM y en el ANEXO A se listan las características de cada uno de ellos.

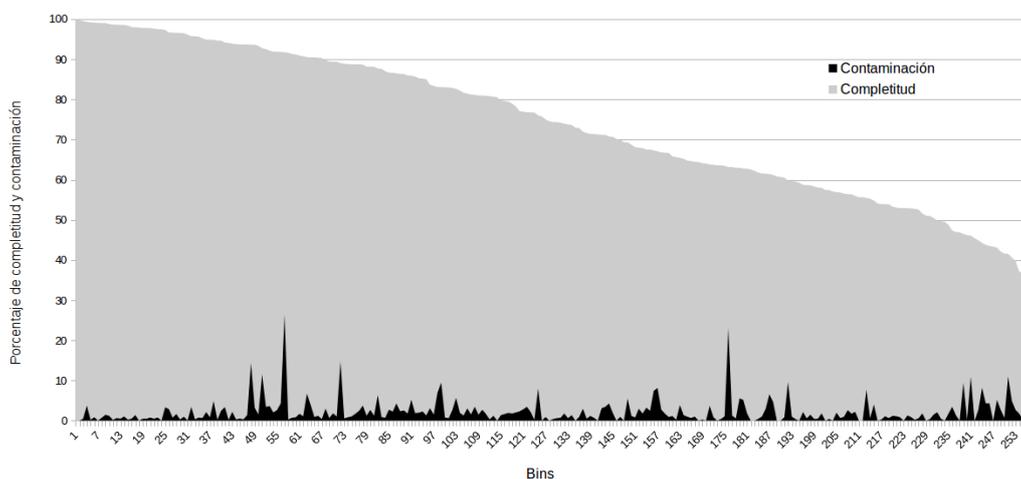


Figura 31. Porcentaje de completitud y de contaminación de los 257 genomas ensamblados ordenados por completitud descendente.

La abundancia en las muestras de cada uno de los genomas de graficó en la Figura 32.

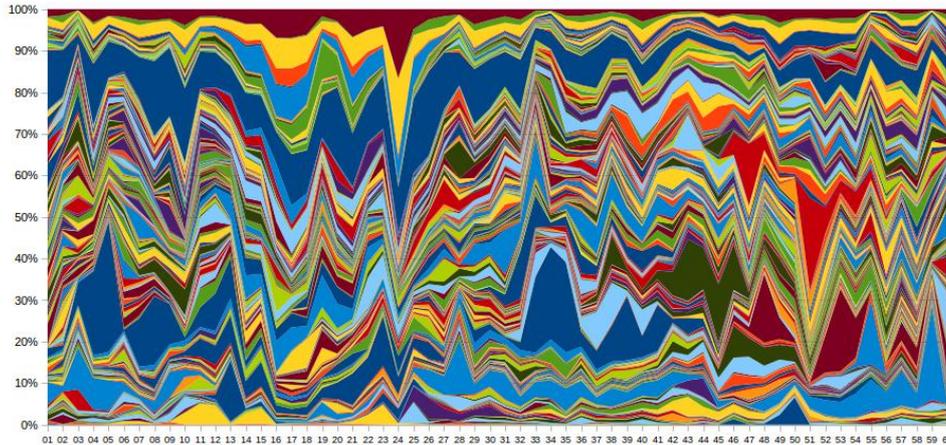


Figura 32. Abundancia de los 257 genomas bacterianos reconstruidos a partir de los metagenomas de barros activados

Durante la reconstrucción de genomas bacterianos a partir de metagenomas existe una enorme dificultad para lograr ensamblar secuencias del gen ARNr 16S (Miller *et al.*, 2011). No obstante, existen herramientas computacionales que permiten la reconstrucción *de novo* de este gen a partir de las lecturas crudas de metagenomas (Miller *et al.*, 2011; Yuan *et al.*, 2015; Pericard *et al.*, 2017; Zeng *et al.*, 2017). Utilizando una de ellas (EMIRGE), en este trabajo se lograron ensamblar 453 secuencias diferentes del gen ARNr 16S, de las cuales las 218 más abundantes representan el 80% del total (Figura 33). Esto último indica que los genomas ensamblados son una buena representación de la comunidad.

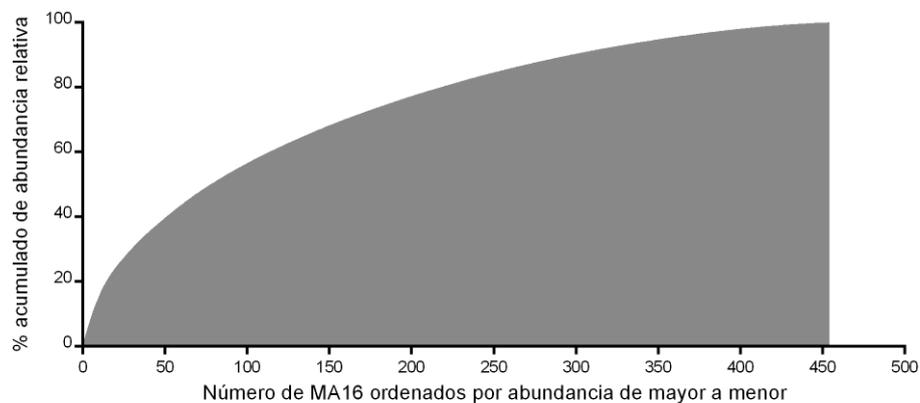


Figura 33. Porcentaje acumulado de la abundancia relativa en el total de las muestras metagenómicas de los 453 genes ARNr 16S ensamblados a partir de los metagenomas (MA16S).

Para evaluar la representatividad de las subcomunidades obtenidas a partir de la reconstrucción de genomas en los metagenomas, se comparó la metacomunidad bacteriana determinada por los diferentes métodos de secuenciación y procedimientos de análisis, que a partir de aquí se definen como:

- MAG: genomas ensamblados a partir del metagenoma (MAG, por su siglas en inglés).
- MA16S: genes ARNr 16S ensamblados a partir del metagenoma (por su analogía con MAG).
- Amp16S: secuenciación de amplicones de la región V3-V4

En la Figura 34 se compara la abundancia de filos y clases de proteobacterias en el total de las muestras. En el ANEXO B se detallan los resultados y se muestra también la abundancia en cada una de las 60 muestras. Si bien se observan algunas diferencias, los tres métodos mostraron consistencia en la composición y coinciden en la determinación de los filos mayoritarios presentes.

La comunidad bacteriana determinada por Amp16S mostró una mayor proporción de filos minoritarios, debido a la mayor profundidad de este tipo de secuenciación. En MAG y en mayor proporción en Amp16S, se obtuvieron taxones sin clasificar, lo que no sucedió con MA16S, ya que en general, es posible una asignación taxonómica confiable del gen 16S completo, al menos a nivel de filo y clase. Otros factores que explican las diferencias en los resultados son: el sesgo cuantitativo de la amplificación en Amp16S, la falta de especificidad de los cebadores “universales” utilizados en la secuenciación de amplicones, el sesgo en la estimación de la abundancia dado por el diferente número de copias del gen ARN 16S en los dos métodos basados en la secuencia del gen ARNr 16S, y la ausencia en MAG, de taxones que no son necesariamente minoritarios, pero cuyos genomas no pudieron ser reconstruidos a partir del metagenoma.

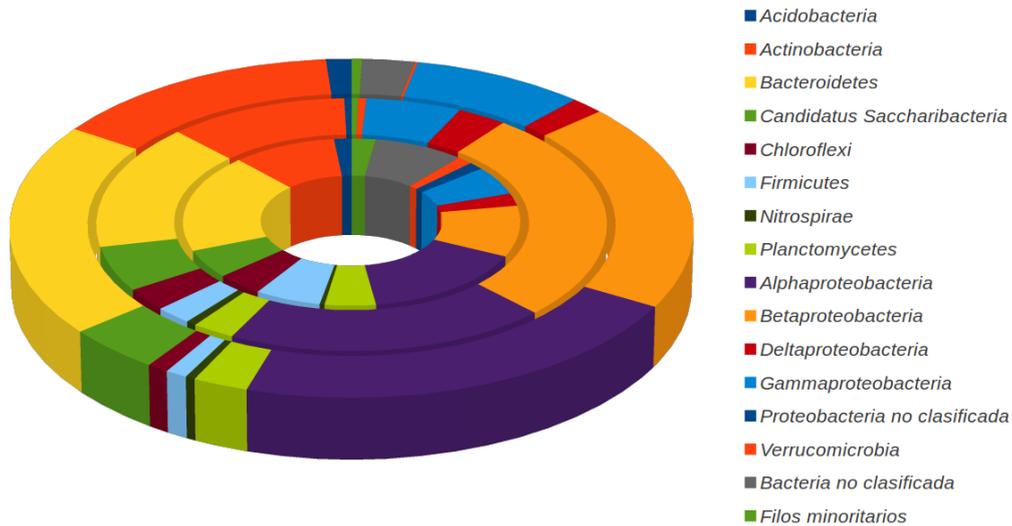


Figura 34. Comparación de la abundancia de filos y clases de proteobacterias obtenidas a partir de Amp16S (anillo interior), MA16S (anillo medio) y MAG (anillo exterior).

Estos resultados confirman que el conjunto de genomas ensamblados a partir de los metagenomas representan adecuadamente las comunidades bacterianas presentes en el sistema.

Bases genómicas de estrategias de crecimiento bacterianas

Número de copias del operón ribosomal rrn

Utilizando la metodología descrita en la sección M&M se estimó el número de copias del operón ribosomal en cada genoma reconstruido, en base a la relación entre la coberturas del gen ribosomal ARNr 16S recuperado a partir de las lecturas crudas y la cobertura del resto del genoma. Este procedimiento fue posible aplicarlo a 85 de los genomas reconstruidos. La Figura 35 ilustra la abundancia relativa de bacterias en la serie temporal con diferente número de copias del ARN ribosomal.

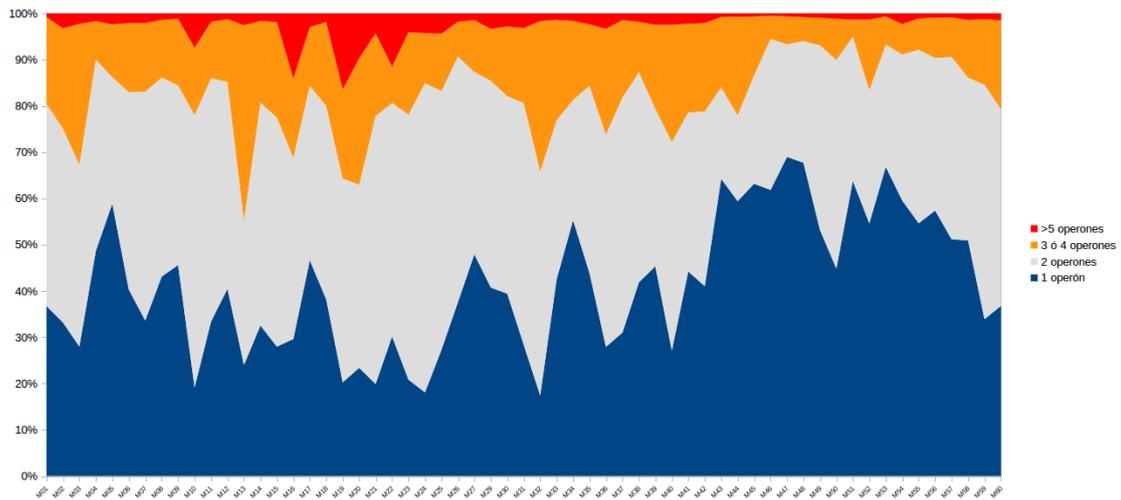


Figura 35. Abundancia relativa de bacterias con 1 (azul), 2 (gris), 3 ó 4 (naranja) y más de 5 (rojo) copias del operón *rrn* en su genoma.

En el PE predominaron los genomas con una sola copia del operón *rrn*, mientras que en el PP se vio aumentada la abundancia de los genomas con más de 3 o 4 número de copias del operón *rrn*. En reportes de bibliografía se ha mostrado que el número copias del operón *rrn* correlaciona con la velocidad a la que las bacterias responden a la disponibilidad de recursos (Klappenbach *et al.*, 2000) y además, se demostró que permite predecir características de reproducción bacterianas (Roller *et al.*, 2016). En base a estos antecedentes concluimos que los cambios en condiciones ambientales, TRS y perturbaciones, seleccionaron diferentes poblaciones bacterianas en base a sus estrategias de crecimiento.

Por su parte, el índice biótico de barros (SBI), que se había estimado a partir de la cuantificación de protozoos, correlacionó positivamente ($R=0.52$, $p<0.0001$) con la abundancia total de bacterias con 1 y 2 operones ribosomales.

Índice de replicación (iRep)

El iRep es un método desarrollado recientemente que permite determinar las tasas de replicación bacteriana a partir de los genomas reconstruidos a partir de un metagenoma y las secuencias crudas del metagenoma (Brown *et al.*,

2016). Este índice tiene en cuenta las diferencias de cobertura entre las regiones del genoma cercanas al índice de replicación y las del resto del genoma. El iRep es mayor en genomas de bacterias de crecimiento rápido debido a que la cobertura cerca del origen de replicación se encuentra incrementada.

A fin de comparar las tasas de replicación de los genomas bacterianos reconstruidos en los diferentes períodos se calculó el iRep para todos los genomas más de 75% completos en aquellas muestras que disponían suficiente cobertura del genoma (Figura 36).

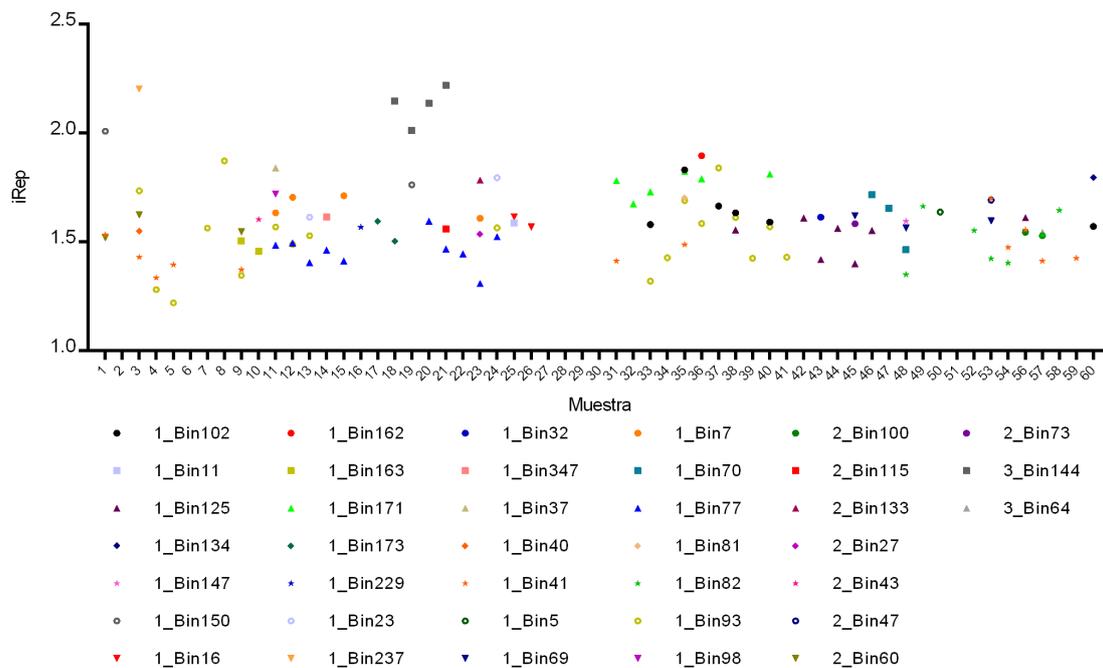


Figura 36. Índices de replicación de genomas calculados para cada muestra

No se observó una relación entre el iRep y los resultados presentados previamente, que mostraron selección de bacterias según sus estrategias de crecimiento en los diferentes períodos. A fin de establecer si existe alguna relación entre el iRep y el número de copias del operón *rnm*, se relacionaron los valores de iRep calculados para los genomas con el número copias del operón *rnm* correspondiente al genoma (Figura 37).

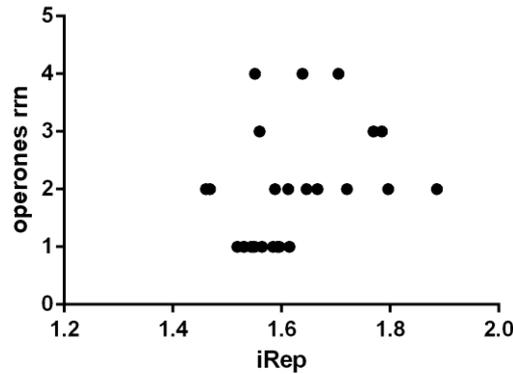


Figura 37. Número de copias del operon de ARN ribosomal en función del iRep calculados para genomas bacterianos reconstruidos a partir del metagenoma de barros activados.

Si bien se percibe que para los genomas que contienen una sola copia del operón *rrn* los valores calculados de iRep son menores en comparación con los que poseen más de tres copias, la correlación de Pearson, fue claramente positiva, pero no resultó significativa (p -valor = 0,08). De todos modos, este análisis estuvo limitado por el bajo número de genomas con alto número de copias a los cuales, a su vez, se les ha podido calcular el iRep.

Como ejemplo de funciones codificadas en los genomas reconstruidos que fueran distintivas de las diferentes condiciones operativas, se analizaron los genes relacionados a la biosíntesis de tiamina. Se ha reportado que la biosíntesis de tiamina *de novo* tiene lugar preferentemente en bacterias con alto número de operones *rrn*, mientras que bacterias con pocas copias del operón *rrn* participan en reciclado de la tiamina, que les permitiría ahorrar energía en ambientes de escasos recursos empleando las reservas exógenas de este cofactor (Roller *et al.*, 2016). A fin de explorar esta observación se realizó un LSA entre la abundancia en el tiempo de genomas bacterianos y de las funciones de los metagenomas correspondientes a esta vía metabólica.

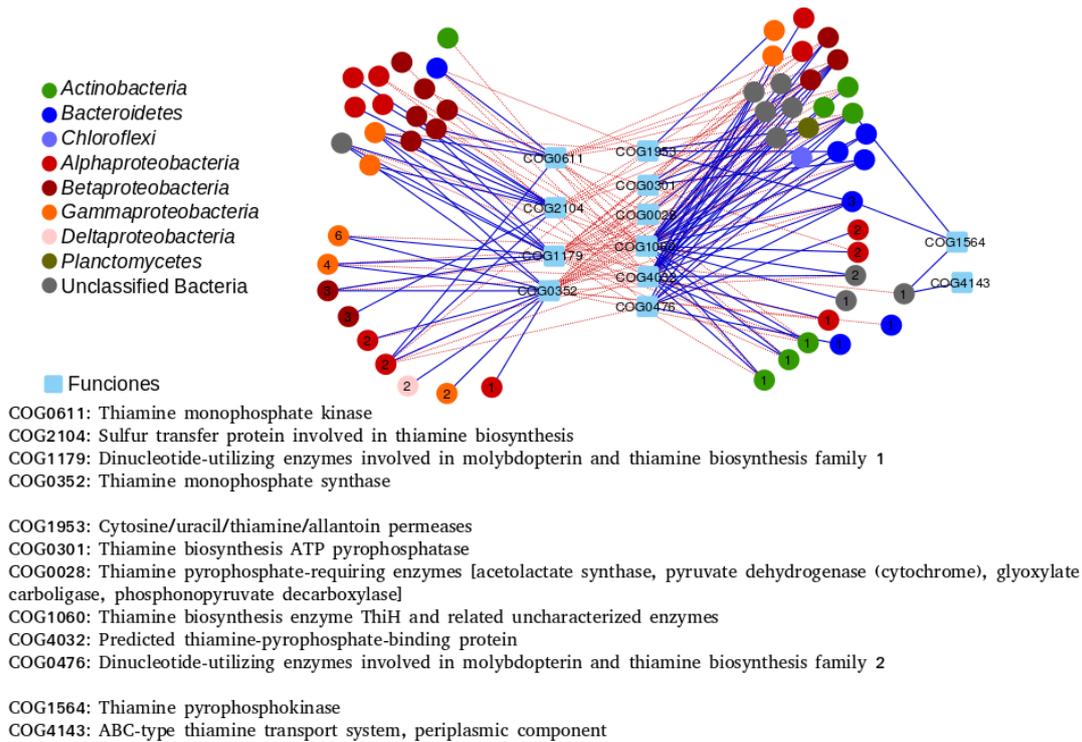


Figura 38. Red de co-ocurrencia entre genomas y funciones relacionadas a la tiamina. Las líneas azules muestran las correlaciones positivas y las rojas las negativas. Los números indican el número de copias del gen ARNr 16S del genoma, en aquellos donde pudo ser calculado.

La Figura 38 muestra que las bacterias que predominan en el PP (izquierda), con mayor número de operones *rrn* y las mayoritarias en el PE (derecha), con menor número de operones *rrn*, estarían seleccionando diferentes caminos para la biosíntesis de tiamina.

Análisis del metaviroma

Estructura de la comunidad viral

A partir de la secuenciación de las muestras de los filtrados de los barros activados, se obtuvieron 32572 *contigs* de un largo de más de 2000 pb, de los cuales 6263 *contigs* tienen un largo de más de 5000 pb. Estos últimos se clasificaron con la base de datos pVOG, y 4957 fueron asignados a virus procariotas. Se pudo estimar la contaminación de DNA bacteriano en el metaviroma entre 1-15%. Para los análisis subsiguientes se utilizaron sólo *contigs* clasificados como bacteriófagos.

Las abundancias de aquellos que lograron clasificarse hasta nivel de familia se muestran en la Figura 39.

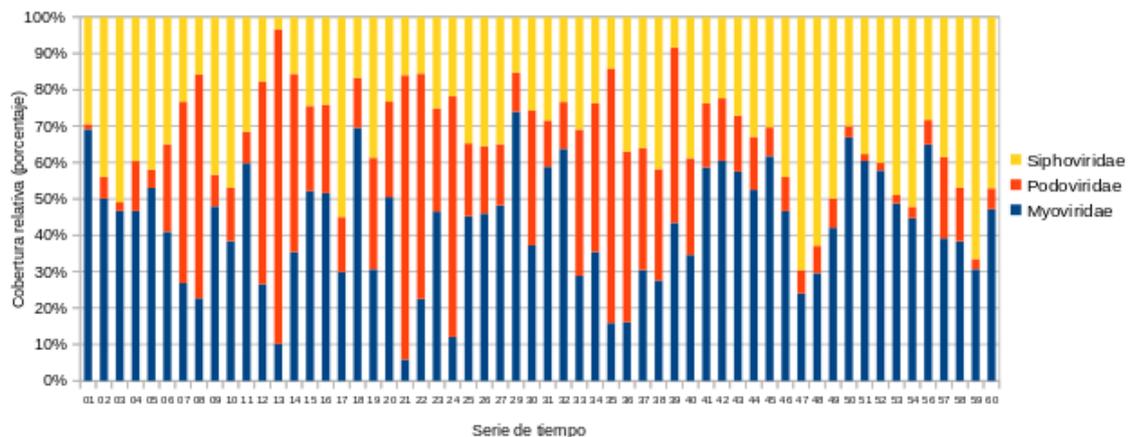


Figura 39. Abundancia de familias de *Caudovirales* en las muestras

Estas familias pertenecen a *Caudovirales*, el orden de bacteriófagos más comúnmente encontrado en general en diversos ecosistemas, como agua de río, intestino humano, piel y aguas residuales, y también en los barros activados (Tamaki *et al.*, 2012; Aw *et al.*, 2014; Cai *et al.*, 2016; Hannigan *et al.*, 2015). Los cambios observados en la composición de fagos reflejan un patrón similar al revelado por las poblaciones bacterianas. La reducción en la abundancia relativa de virus de la familia *Podoviridae* y el correspondiente aumento en *Siphoviridae* que acompañan el aumento en TRS es consistente con el hecho de que los *Podoviridae* infectan principalmente *Proteobacteria*,

mientras que sifovirus pueden ser lisogénicos por varias generaciones, un comportamiento que ha sido asimilado a estrategias K (Sime-Ngando, 2014).

Interacción fago-hospedero

Para estudiar interacciones fago-hospedero en metagenomas es posible recurrir al uso de herramientas computacionales que permitan inferir el hospedero del virus. Entre las posibles estrategias para lograrlo, se encuentran: la búsqueda de similitud de secuencias, la presencia de espaciadores CRISPRs, la comparación de los perfiles de abundancia en el tiempo, la similitud de perfiles de k-meros y % de G+C, entre el huésped y el hospedero (Edwards *et al.*, 2016). En este trabajo se ha utilizado la búsqueda de elementos CRISPR y la similitud de secuencias, complementando dicha información con la comparación de los perfiles de abundancia temporal.

Comparación de perfiles de abundancia

Se realizó inicialmente un LSA considerando las abundancias de las especies bacterianas y de los *contigs* de los fagos.

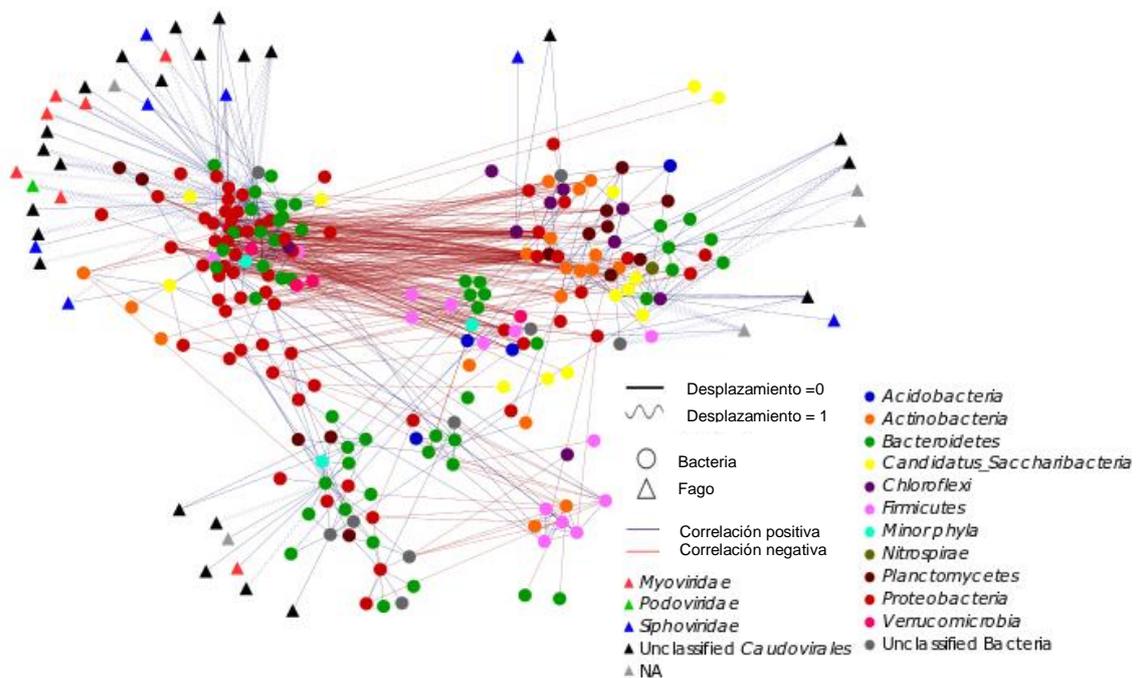


Figura 40. Red de co-ocurrencia fago-bacteria realizada a partir de LSA entre OTUs de bacterias y *contigs* ensamblados a partir del metaviroma.

Se observó que en el PP, mayormente compuesto por bacterias del filo *Proteobacteria* y *Bacteroidetes*, predominan fagos pertenecientes a familias *Myoviridae*. Los myovirus suelen ser mayormente líticos y estar asociados a estrategias-r y a ambientes fluctuantes (Sime-Ngando, 2014). Si bien esta red de co-ocurrencia identificó pocos fagos en el PE, los escasos que pudieron ser clasificados a nivel de familia, corresponden a la familia *Siphoviridae*.

Debido a que la co-ocurrencia no es prueba directa de interacción, esta estrategia por sí sola no permite la identificación de pares fago-hospedero.

Búsqueda de CRISPRs

Se encontraron elementos CRISPRs en 157 genomas, es decir el 61% de los genomas bacterianos ensamblados. Estos contenían en total 4976 espaciadores que fueron alineados con los *contigs* ensamblados a partir del metaviroma. Se hallaron 50 espaciadores coincidentes entre genomas y fagos, incluidos en 23 genomas bacterianos. Estos resultados y también las características de los sistemas CRISPRs encontrados en estos genomas con CRISPROne y MetaCRT, se resumen en la Tabla 7.

Identidad del genoma (bin)	Nombre del bin (%compl)	Contigs de fagos con espaciador coincidente (posición del espaciador en el arreglo de CRISPR)	Sistemas CRISPR y CRISPR/cas encontrados en contigs del genoma (bin)
<i>Alphaproteobacteria</i>	1_Bin32 (97,47%)	Ph_1263449 (1,2)	4 genes cas tipo I y II + 2R* 6R*
<i>Alphaproteobacteria</i>	1_Bin40 (99,3%)	Ph_460655 (8,10,11,19)	3 genes cas tipo II 35R* 18R* 7 genes cas tipo I + 28R*
<i>Alphaproteobacteria</i>	1_Bin78 (99,1%)	Ph_1241379 (8) Ph_335968 (14) Ph_1671820 (18)	2 genes cas + 81R 4 genes cas tipo I 10R* - -
<i>Betaproteobacteria</i>	1_Bin95 (67,54%)	Ph_1249213 (17) Ph_1876356 (23,25)	*7 genes cas tipo I* 49R*
<i>Betaproteobacteria</i>	2_Bin43 (93,85%)	Ph_2612996 (10) Ph_5158030 (10) Ph_1306099 (27) Ph_1510711 (36, mismatch)	7 genes cas tipo I + 52R 15 genes cas tipo I y III* 50R* 3 genes cas tipo I 4R - - 25R - - 7 genes cas tipo I + 19R 5R*
<i>Betaproteobacteria</i>	2_Bin131 (68,71%)	Ph_686527 (8)	7 genes cas tipo I + 35R* 7 genes cas tipo I + 3R* MetaCRT además: 11R - -
<i>Gammaproteobacteria</i>	1_Bin93 (98,64%)	Ph_2376641 (28,63)	8R* 3R + 3R 8 genes cas tipo I + 101R 4R* - - 6 genes cas tipo I*

			* 1gen cas + 105R*
<i>Deltaproteobacteria</i>	1_Bin11 (89,76%)	Ph_668266 (128) Ph_630059 (21) Ph_198957 (20, mismatch) Ph_1097811 (14) Ph_1561673 (7, mismatch)	5R -- 21R* 7 genes cas tipo I + 144R + 8R 4R -- 3R -- 3R(3),4R(3),5R(1)6R(1)8R(1)10R(1)
<i>Deltaproteobacteria</i>	1_Bin7 (94,84%)	Ph_2009184	9 genes cas tipo I + 76R* 2 genes cas tipo II y II 7 genes cas tipo I + 366R* *8 genes cas tipo III + 38R .121R(1) 6R + 5R 3R(3),4R(3),5R(6),14R(1),20R(1),24R(1),56R(1) ...
<i>Bacteroidetes, Sphingobacteriia, Haliscomenobacter</i>	2_Bin25 (97,93%)	Ph_1036195 (37)	3 genes cas tipo I 9 genes cas tipo III + 49R 10 genes cas tipo I + 45R
	1_Bin98 (76,73%)	Ph_2331691 (9)	*4 genes cas tipo I* 4R 9R* 5 genes cas tipo III *7 genes cas tipo I + 14R* 33R 4R
	1_Bin171 (81,2%)	Ph_753841 (21)	15R 6 genes cas tipo I 1 gen cas 55R
	1_Bin134 (85,26%)	Ph_626559 (79)	4R* -- 3R -- 86R* 4R
<i>Bacteroidetes, Sphingobacteriia</i>	1_Bin69 (99,51%)	Ph_2346234 (1)	23R* 24R* 2 genes cas 7 genes cas tipo III 3R* 35R* 4R 50R* 5 genes cas tipo III 10R --
	1_Bin147 (98,57%)	Ph_1057414 (1)	4 genes cas tipo II + 3R* 4R*
<i>Bacteroidetes, Cytophagia</i>	2_Bin63 (65,44%)	Ph_2734096 (14) 69R Ph_453499 (27) 60R	3R* 6R* 69R 41R* 4 genes cas tipo I* *60R* 4R -- +3 genes cas tipo I + 18R* 10 genes cas tipo III + 30R* 5R -- 9R
<i>Bacteroidetes, Cytophagia, Runella</i>	2_Bin14 (83,65%)	Ph_1761110 (23)	3 genes cas tipo I 9 genes cas tipo I + 10R* 10 genes cas tipo I + 178R 3 genes cas tipo III 8R -- 5R + 8R
	1_Bin18 (82,97%)	Ph_2076737 (5)	2 genes cas 11 genes cas tipo I+7R* 2 genes cas tipo I
<i>Actinobacteria, Acidimicrobium</i>	1_Bin119 (92,5%)	Ph_85503 (8,13,23) Ph_195104 (12) Ph_711047 (25,34) Ph213222 (6)	3 genes cas tipo I* 2 genes cas tipo I 6 genes cas tipo I + 4R* *24R* (posición 6) *37R* (posiciones 8,12,13,23,25) *5 genes cas tipo I + 5R*
<i>Actinobacteria, Intrasporangium</i>	3_Bin144 (81,45%)	Ph_2271477 (3) Ph_405962 (4)	*5 genes cas tipo I* *67R* 6 genes cas tipo I + 87R
<i>Actinobacteria,</i>	1_Bin41	Ph_48112 (Siphoviridae)	8 genes cas Tipo III + 8R

Gordonia	(98,58%)	(posiciones 1y2)	3 genes cas tipo I 10R - - 5 genes cas tipo I + 11R* 13 R + 15R
<i>Planctomycetes</i> , Planctomyces	1_Bin56 (69,38%)	Ph_82233 (1)	3R*
<i>Verrucomicrobia</i> , Akkermansia	1_Bin66 (47,34%)	Ph_1633714 (10) Ph_1523141 (16)	1 gen cas* *32R* 1 gen cas * 1 gen cas * 36R*

*Indica extremo del *contig* del bin, - - indica que las repeticiones no idénticas dentro del arreglo de CRISPR, R indica el número de repeticiones en el arreglo

Tabla 7. Resumen de resultados de la búsqueda de espaciadores CRISPR coincidentes en genomas y en *contigs* de fagos. En cada línea de la última columna se detallan todos los sistemas CRISPRs encontrados en *contigs* del genoma. Se resalta en gris el loci que contiene el espaciador que se halló en el fago. Al haber más de uno se diferencia por el color de la fuente que también indica repeticiones idénticas entre arreglos CRISPR en *contigs* dentro de un mismo bin.

Los gráficos de abundancia en el tiempo de los pares de interacción encontrados con la búsqueda de espaciadores CRISPR se muestran en el ANEXO C.

Se han podido asignar más de dos decenas de hospederos a fagos recuperados a partir del sobrenadante. Estos hospederos pertenecieron a grupos taxonómicos bacterianos diversos (*Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, *Actinobacteria*, *Planctomycetes* y *Verrucomicrobia*). Algunos ejemplos mostraron hospederos con más de un fago asociado (1_Bin78-figura 3c, 1_Bin95-figura 4c, 2_Bin43-figura 5c, 1_Bin11-figura 8c, 2_Bin63-figura16c, 3_Bin144-figura 20c) y en otros, se observaron fagos que posiblemente cambien de hospedero (1_Bin40-figura2c, 1_Bin93-figura 7c y 1_Bin7-figura9c, 1_Bin66-figura23c). Comparando perfiles de abundancia y taxonomía, se hallaron, en dos casos, las posibles bacterias a las que estos fagos pueden haber encontrado como nuevos hospederos (figura 7c y 9c).

Algunas dinámicas de interacción fago-hospedero observadas muestran modelos similares a los de Lotka-Volterra (1_Bin32-figura 1c, 1_Bin78-figura 3c, 1_Bin11-figura 8c, 1_Bin69-figura 14c, 2_Bin63-figura 16c, 2_Bin14-figura 17c).

Una particularidad se encontró en cuatro bacterias del género, *Haliscomenobacter* (dos de ellas clasificadas como *Haliscomenobacter hydrosis*), cada una con un fago diferente asociado (2_Bin25, 1_Bin98, 1_Bin171 y 1_Bin134). Se ha estudiado el uso de bacteriófagos como herramientas de biocontrol de esta especie (Kotay *et al.*, 2011). Esta bacteria es una de las bacterias filamentosas más comúnmente encontradas en sistemas de lodos activados, que al crecer desmedidamente causa importantes problemas en el proceso. La resistencia que muestran estas especies contra sus bacteriófagos (figuras 10c-13c) pone en duda que el uso de un bacteriófago como herramienta de biocontrol sea exitoso en el tiempo.

Otra particularidad se halla en el fago de *Gordonia* (1_Bin41), Ph_84112, un *Shipoviridae* que posee un arreglo de CRISPR similar al de la bacteria (Figura 41).

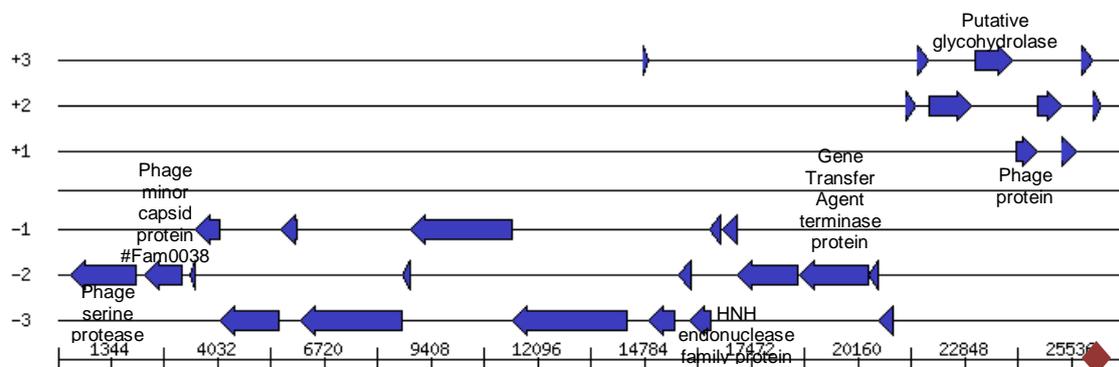


Figura 41. Mapa genómico del fago Ph_48112 de una longitud de 26899 pb. El rombo (entre las posiciones 26272-26482) representa el arreglo de CRISPR (4 repeticiones y 3 espaciadores) compartidos con la bacteria del género *Gordonia*.

La estimación de la abundancia de bacterias con o sin CRISPR mostró que en el tiempo co-existen poblaciones bacterianas resistentes y sensibles al fago (con el protoespaciador).

Búsqueda de similitud de secuencias

El otro enfoque que se ha considerado apropiado para establecer interacciones certeras fago-bacteria fue el de búsqueda de similitud de secuencias nucleotídicas. Algunos mecanismos moleculares dan como resultado la presencia de secuencias idénticas en los genomas de un fago y su hospedero. Por un lado, la homología de genes bacterianos y fagos puede indicar interacciones recientes, por ejemplo si el fago adquirió algún gen durante un evento de infección reciente del hospedero. Tanto los fagos líticos como los temperados son capaces de incorporar material genético del hospedero en su genoma. Por otro lado, otra fuente de coincidencia exacta de secuencias está dada por fagos integrados en el genoma de su hospedero (profagos).

Para explorar este tipo de interacciones se realizó un alineamiento de las secuencias nucleotídicas (BLASTn) entre los genomas bacterianos reconstruidos y los *contigs* ensamblados a partir del metaviroma. Mediante este análisis se encontraron 16 de posibles interacciones fago-hospedero que se analizaron en detalle. La siguiente tabla muestra los resultados del alineamiento realizado.

Contig del fago – (L: longitud del contig en pb)	Nombre y contig del Bin L: longitud del contig en pb Filo o clase del bin (%de completitud)	% de identidad	Longitud del alineamiento (pb)	Mismatch	Gapopen	query inicio (fago)	query final (fago)	subject inicio (bin)	subject final (bin)
Ph_1082460 L: 5339 Ph_1082460 L: 5339 Ph_1881008 L:7634 Ph_1881008 L:7634	1_Bin102 Contig81 L: 9489 1_Bin102 Contig81 L: 9489 1_Bin102 Contig59 L: 17369 1_Bin102 Contig59 L: 17369 Bacteroidetes (96,55%)	100 100 100 100	2928 1704 6097 1211	0 0 0 0	0 0 0 0	2412 670 1538 1	5339 2373 7634 1211	3209 1467 1697 160	6136 3170 7793 1370
Ph_2397247 L: 9061 Ph_2397247 L: 9061 Ph_2397247* L: 9061 Ph_2397247 L: 9061	1_Bin128 Contig530 L: 2288 1_Bin128 Contig707 L: 1586 1_Bin128 Contig54 L: 9507 1_Bin128 Contig696 L: 1636 Bacteroidetes (83,03%)	100 99,75 100 100	1323 1580 2169 1636	0 4 0 0	0 0 0 0	2913 5134 6691 700	4235 6713 8859 2335	2288 1580 9507 1636	966 1 7339 1
Ph_1537740* L: 7637	1_Bin15 Contig517 L: 3871 Betaproteobacteria (91,88%)	99,91	2162	1	1	5476	7637	36	2196
Ph_384352 L:5720	2_Bin145 Contig158 L:5858 Actinobacteria (79,46%)	99,98	5720	1	0	1	5720	54	5773
Ph_33231 L:17296 Ph_156953 L: 16325 Ph_156953 L: 16325 Ph_1368580 L: 8396 Ph_2205832 L: 6251	2_Bin185 Contig184 L:2426 2_Bin185 Contig245 L: 2185 2_Bin185 Contig437 L:1720 2_Bin185 Contig362 L: 1875 2_Bin185 Contig387 L: 1822 Bacteroidetes (39,8%)	99,75 99,95 100 99,95 99,95	2426 2185 1720 1875 1822	6 1 0 1 1	0 0 0 0 0	5048 12353 5778 1409 1848	7473 14537 7497 3283 3669	1 2185 1720 1 1822	2426 1 1 1875 1
Ph_884857 L: 7434 Ph_908964 L: 7391 Ph_908964 L: 7391 Ph_2368182 L:8533	2_Bin20 Contig635 L: 3713 2_Bin20 Contig257 L: 7361 2_Bin20 Contig257 L: 7361 2_Bin20 Contig204 L: 8162 Acidobacteria (85,90%)	100 100 99,96 100	3713 4038 2727 4830	0 0 1 0	0 0 0 0	1826 2827 1 3704	5538 6864 2727 8533	3713 4038 6864 7410	1 1 4138 2581
Ph_36540 L:15957 Ph_36540 L:15957 Ph_36540 L:15957	2_Bin219 Contig4 L: 18855 2_Bin219 Contig71 L: 6713 2_Bin219 Contig3 L: 20882	99,94 99,97 99,95	3579 6713 5625	2 2 3	0 0 0	12379 5617 1	15957 12329 5625	1 6713 5625	3579 1 1

Ph_94248 L: 16641 Ph_177642 L: 5517 Ph_250347 L: 5241 Ph_250347 L: 5241	2_Bin219 Contig3 L: 20882 2_Bin219 Contig4 L: 18855 2_Bin219 Contig4 L: 18855 2_Bin219 Contig4 L: 18855 Actinobacteria (62,97%)	99,97 99,96 99,97 99,94	15349 5517 3427 1751	4 2 1 1	0 0 0 0	1 1 1 3491	15349 5517 3427 5241	5527 9187 12107 15597	20875 3671 15533 177347
Ph_241990 L:5381	2_Bin21 Contig265 L: 7949 Chloroflexi (89,39%)	99,85	5381	8	0	1	5381	1773	7153
Ph_1685817 L: 9776 Ph_1685817 L: 9776	2_Bin88 Contig10 L: 20787 2_Bin88 Contig10 L: 20787 Alphaproteobacteria (76,03%)	100 100	5289 4462	0 0	0 0	40 5315	5328 9776	5943 10476	655 6015
Ph_1420619 L: 10220	1_Bin258 Contig176 L:3766 Actinobacteria C: 57,99%	99,73	3753	8	2	1	3752	15	3766
Ph_85503 L: 8600	1_Bin119 Contig63 L:15546 Actinobacteria (92,50%)	100	8600	0	0	1	8600	11160	2561
Ph_1864595* L: 23344 Ph_1864595 L: 23344 Ph_1864595 L: 23344 Ph_1864595 L: 23344	1_Bin18 Contig205 L:8774 1_Bin18 Contig441 L:3341 1_Bin18 Contig441 L:3341 1_Bin18 Contig510 L:2842 Bacteroidetes (82,97%)	100 99,66 99,36 100	3141 2052 1242 2037	0 7 8 0	0 0 0 0	866 12645 14744 7010	4006 14696 15985 9046	5151 1 2100 806	2011 2052 3341 2842
Ph_1235975 L: 38632 Ph_1235975 L: 38632	2_Bin27 Contig913 L:1693 2_Bin27 Contig769 L:2312 Gammaroteobacteria (88,65%)	100 100	1693 2312	0 0	0 0	9519 23647	11211 25958	1 1	1693 2312
Ph_1235975 L: 38632	1_Bin170 Contig510 L:1926 Alphaproteobacteria (72,87%)	100	1905	0	0	239	5143	22	1926
Ph_800005 L:12460	1_Bin176 Contig29 L:8104 Bacteroidetes (97,88%)	99,81	2566	5	0	7033	9598	2807	5372
Ph_907715 L:25328	1_Bin32 Contig4 L:358132 Alphaproteobacteria (97,47%)	100	4481	0	0	1952	6435	301547	306030

* Integrasas encontradas

Tabla 8. Resumen de resultados del alineamiento de secuencias nucleotídicas (BLASTn) entre genomas bacterianos y contigs del metaviroma. Los profagos se resaltan en gris.

Ciclos lisis-lisogenia

A partir de este análisis se identificaron doce interacciones fago-bacteria en las cuales el fago integró su genoma en el de la bacteria, es decir se encontró como profago. El hecho de haber recuperado a las partículas virales a partir del sobrenadante indica que también existió lisis. Estas interacciones son claros ejemplos de ciclos de lisis-lisogenia. En las figuras del anexo D (figuras 1d-12d), que muestran la dinámica de interacción, se puede observar que en todos los casos permanecen a lo largo del tiempo tanto poblaciones bacterianas como del fago, éste encontrado en el sobrenadante. Es decir, habría una coexistencia de poblaciones de bacteriófagos llevando a cabo un ciclo de lisis y, otras, un ciclo de lisogenia. También se observan, en algunos ejemplos, intervalos de tiempo en donde los bacteriófagos estarían atravesando sólo ciclos de lisogenia (Figuras 1d, 2d, 6d, 9d, 12d).

Es interesante analizar aparte dos casos especiales. El fago Ph_85503, del 1_Bin_119, además de estar integrado en el genoma de su hospedero también comparte 3 espaciadores CRISPR idénticos. El mapa genético del fago se muestra en la siguiente figura.

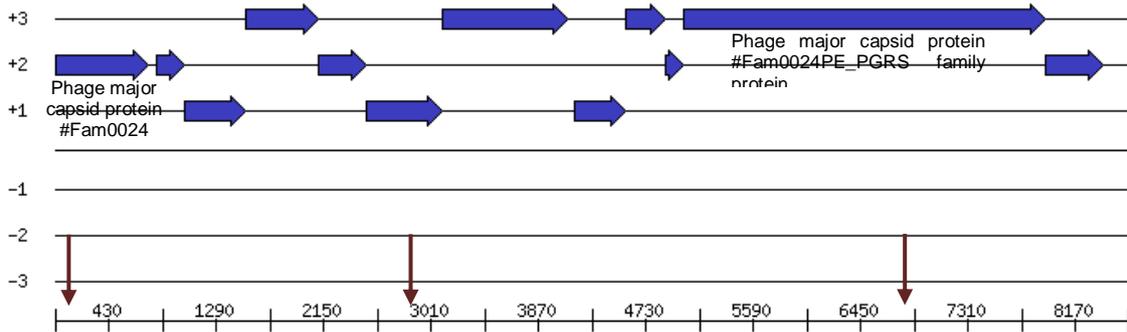


Figura 42. Mapa genético del fago Ph_85503. Con flechas se marcan los *protospacers* que comparte identidad con su hospedero, 1_Bin119. Proteínas sin rótulos se encuentran anotadas como proteínas hipotéticas.

La escasa longitud de los *contigs* obtenidos en el genoma de la bacteria no permitió explorar los alrededores del profago. El genoma del fago estaría compuesto por este *contig* (Ph_85503) y por *contigs* de longitud menores a 5000pb, que se excluyeron de este análisis, pero se vio que también forman parte del genoma de la bacteria y compartieron espaciadores CRISPR con ella. No fue posible reconstruir el fago completo a partir de estos *contigs* y obtener un *contig* del fago de mayor longitud ya que las secuencias no se superponían.

Otro caso interesante es el del 1_Bin18, que contiene un fago integrado en su genoma (Ph_1864595) y otro, diferente, que no es un profago, pero comparte un espaciador CRISPR con la bacteria (Ph_2076737). La dinámica de ambos bacteriófagos y de la bacteria son muy similares en el tiempo. (Figuras 12d y 18c). En la siguiente figura se muestra el mapa genético del fago que se encuentra como profago (Ph1864595). No se muestra el del fago que comparte espaciadores CRISPRs (Ph_2076737) ya que ninguna proteína logró anotarse.

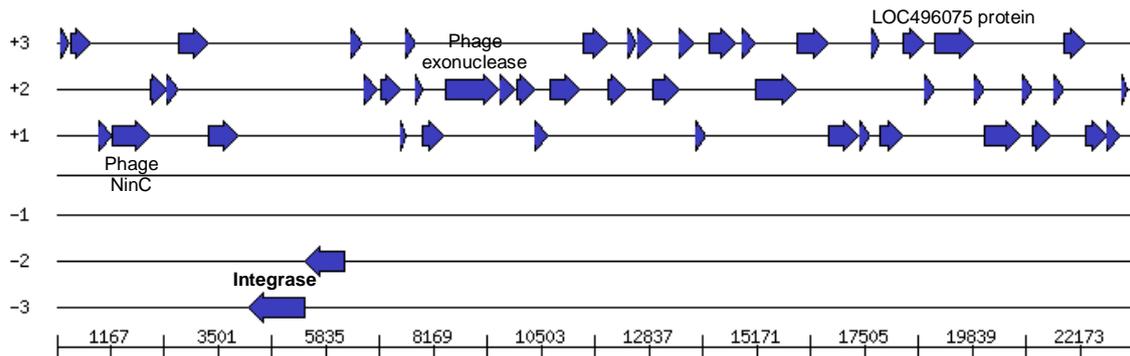


Figura 43. Mapa genético del fago Ph_1864595. Proteínas sin rótulos se encuentran anotadas como proteínas hipotéticas.

En la Figura 44 se muestra el locus de la bacteria correspondiente al sistema CRISPR cuyo arreglo de CRISPR contiene el espaciador encontrado en el fago (Ph_2076737). No ha sido posible explorar los alrededores del profago en el genoma bacteriano debido a la escasa longitud de los *contigs* obtenidos.

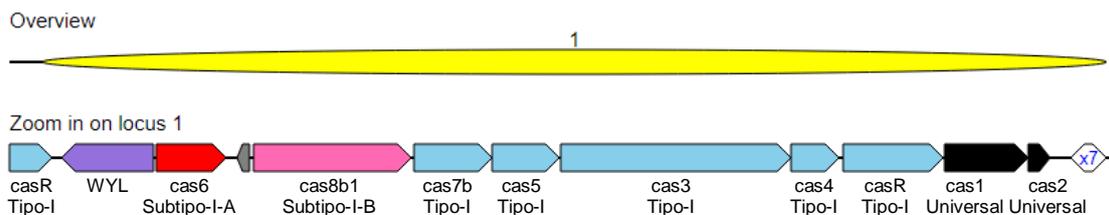


Figura 44. Locus encontrado con CRISPROne en uno de los *contigs* del 1_Bin18 correspondiente al sistema CRISPR que contiene el espaciador coincidente al del fago Ph-2076737 en la posición 5. En amarillo se muestra la visión general del sistema CRISPR en el *contig* del bin.

Los restantes cuatro pares de interacción detectados con el método de búsqueda de similitud de secuencias que no corresponderían, en principio, a integración de fagos en el genoma de la bacteria se analizaron individualmente.

Por un lado los genomas bacterianos 2_Bin27 (*Gammaproteobacteria*), y 1_Bin170 (*Alphaproteobacteria*), comparten el mismo fago (Ph_1235975). El mapa genético del fago se muestra en la Figura 45. Este fago podría ser un

Por último, un ejemplo similar al anterior es el del genoma de 1_Bin32 que comparte una secuencia similar con el fago Ph_907715.

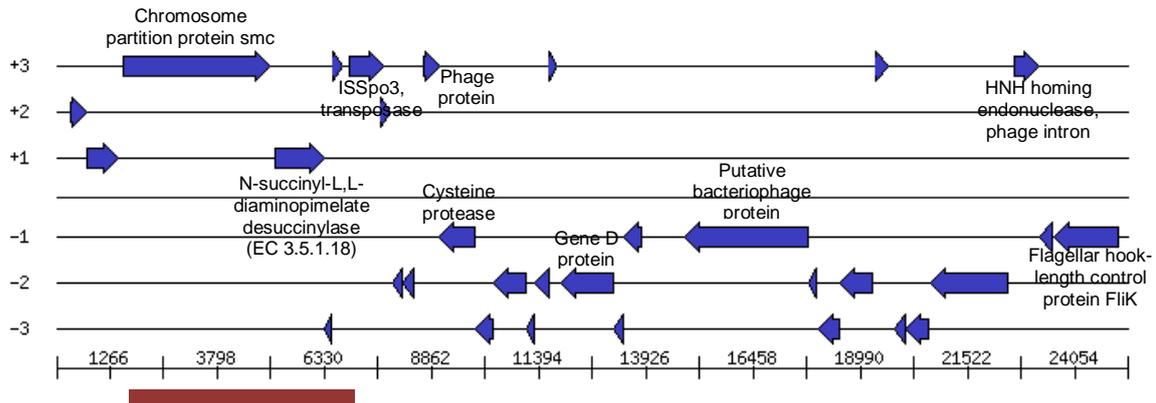


Figura 47. Mapa genético del fago Ph_907715. Proteínas sin rótulos se encuentran anotadas como proteínas hipotéticas. La barra de color bordó esquematiza el alineamiento compartido con la bacteria.

Lisogenia

También existen herramientas computacionales desarrolladas que permiten detectar profagos integrados en genomas bacterianos (Fouts, 2006; Lima-Mendez *et al.*, 2008; Akhter *et al.*, 2012; Roux *et al.*, 2015). En este trabajo se encontraron con VirSorter secuencias virales dentro de 156 de los 257, es decir en el 60% de los genomas. Algunos de ellos corresponden a los ejemplos anteriormente analizados. En la mayoría de ellos no se ha podido recuperar el fago en el sobrenadante, lo que indicaría que, si bien puede ser una cuestión de profundidad de la secuenciación del metaviroma, existiría una gran proporción de bacterias en este ambiente con fagos integrados en su genoma por largos períodos de tiempo, es decir largos períodos de lisogenia, sin ciclos de lisis.

Aún se cuenta con otro método para encontrar pares de interacción fago-bacteria por similitud de secuencias más cortas que los espaciadores CRISPRs. Existen secuencias cortas coincidentes que provienen de los sitios de integración en el genoma del fago (attP) y en el genoma de la bacteria (attB). Estos sitios de integración comparten homología de secuencia sólo en unas pocas bases (entre 2 y 15 aproximadamente, dependiendo del fago, la bacteria

y la recombinasa que utilizan). Si se llegaran a encontrar secuencias cortas idénticas, para descartar que no sean coincidencias al azar, es posible buscar si esas bases se ubican cerca de integrasas (en el fago) y cerca o dentro de ARNt (en la bacteria), ya que suelen encontrarse en esas zonas. En este análisis no se encontraron alineamientos mayores a 99% de similitud, que sean de longitudes menores a 28 pb (correspondientes a espaciadores CRISPR).

DISCUSIÓN

Influencia de factores abióticos sobre la estructura de las comunidades bacterianas de sistemas de barros activados y su relación con las estrategias de crecimiento

Para garantizar la máxima eficiencia en el tratamiento de efluentes con sistemas de barros activados es crucial que las bacterias en el reactor, principales degradadoras de la materia orgánica de los efluentes, utilicen los sustratos lo más eficientemente posible y, a su vez, que formen un floc denso y estable que pueda ser luego fácilmente separado por gravedad del líquido tratado. Los manuales de operación de sistemas de barros activados disponibles (Gott *et al.*, 2007; Boyajian and Glueckstein, 1987; Jenkins *et al.*, 2004) han documentado extensamente cuáles son las mejores condiciones empíricas de operación que permiten que los flocs microbianos posean estas propiedades. Las razones por las cuales determinados parámetros operativos resultan óptimos, han sido relacionadas con el metabolismo bacteriano y sus formas de crecimiento. En la Figura 48 se muestra una curva típica de crecimiento bacteriano y la curva de sustrato disponible correspondiente (DBO en este caso) en función del tiempo, en un sistema de barros activados.

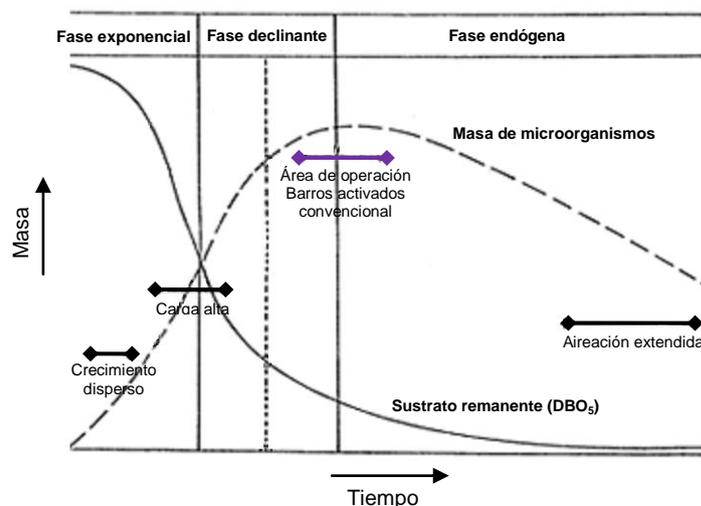


Figura 48. Curva de crecimiento bacteriano y sustrato remanente a lo largo del tiempo en barros activados. (Adaptado de: Jay Bevlins, Activated sludge process, <http://web.deu.edu.tr/atiksu/ana52/acti6.html>)

El rango óptimo de operación en un sistema de barros activados convencional se produce cuando la biomasa en el reactor crece entre las fases declinante y endógena, es decir cuando el alimento en relación a la cantidad de microorganismos es limitante, lo que en términos de parámetros operativos representa una determinada relación F/M y un determinado TRS. A TRS muy bajos, es decir cuando las concentraciones de sustrato son elevadas, las bacterias crecen en fase exponencial tendiendo a un crecimiento disperso, no floculento, y por consiguiente ocurre una pérdida de sólidos que resulta en un tratamiento deficiente. Cuando el TRS es excesivamente alto, y la materia orgánica muy escasa, las células bacterianas, que se encuentran en fase endógena avanzada, mueren, los flocs se rompen, formándose lo que se conoce como “pin floc”, que son pequeños flocs con escasa capacidad de sedimentar, por lo que se obtiene un efluente turbio.

Los resultados presentados en este trabajo indican que los factores abióticos, en particular el TRS y las condiciones inestables de operación (período perturbado) seleccionan diferentes comunidades bacterianas según sus estrategias de crecimiento. Vuono y colaboradores manipularon el TRS de una PTE a escala y observaron que los filos considerados estrategias-r (Fierer *et al.*, 2007) incrementaron su abundancia en el período de bajo TRS, mientras que en el período de alto TRS predominaron filos relacionados a estrategias-K (Vuono *et al.*, 2015). Los resultados de este trabajo proveen evidencia directa de que las bacterias seleccionadas en cada período poseen diferencias en el número de operones de ARN ribosomal, un rasgo que permite predecir dos importantes componentes de la reproducción: la velocidad y la eficiencia de crecimiento (Roller *et al.*, 2016).

Estas observaciones, permiten establecer una relación entre las estrategias de crecimiento bacterianas y su capacidad de floculación en sistemas de barros activados (Figura 49).

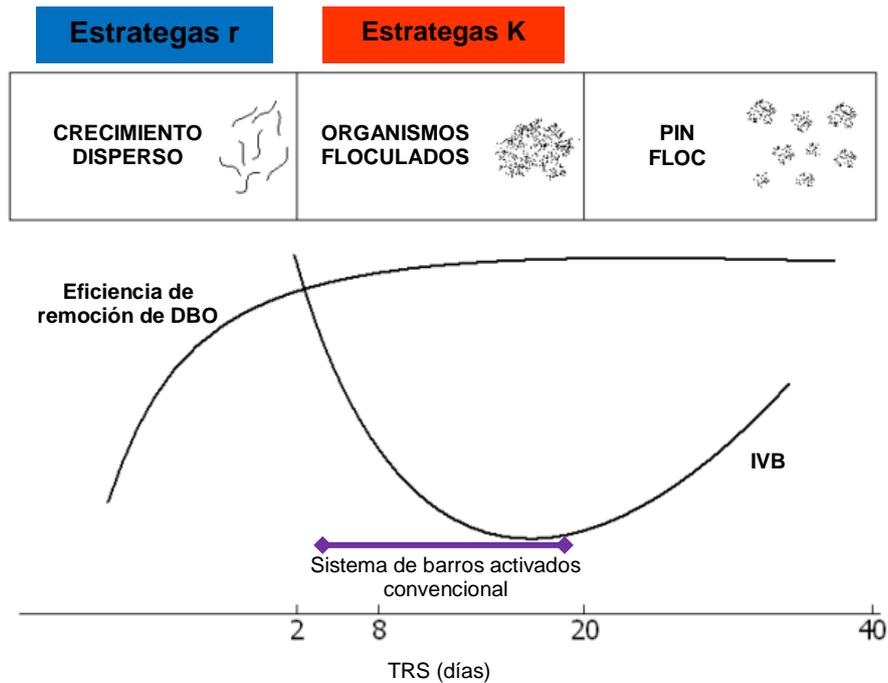


Figura 49. Selección de poblaciones con diferentes estrategias de crecimiento. (Adaptado de: http://www.maine.gov/dep/water/wwtreatment/activated_sludge_process_control.pdf)

El hecho de que la velocidad y la eficiencia de crecimiento sean favorecidas bajo regímenes contrastantes de disponibilidad de recursos (Roller 2016), permite entender por qué se seleccionan bacterias de bajo número de operones *rnn* en un proceso estable, cuando los recursos son escasos (es decir, un sistema produciendo efluente con baja demanda de oxígeno). Esto explica asimismo el motivo por el cual el uso de bacterias seleccionadas en el laboratorio (generalmente bacterias de crecimiento rápido) no resulta eficiente para solucionar problemas operativos relacionados con baja actividad biológica en procesos de barros activados (Stephenson and Stephenson, 1992).

La clasificación de organismos en estrategias r y estrategias K provee una importante base conceptual para realizar predicciones específicas acerca de los atributos ecológicos y preferencia de nichos de organismos superiores (Pianka, 1970). En general es sabido que los estrategias-r, copiótrofos, están adaptados a maximizar su crecimiento cuando los recursos son abundantes, en tanto que los estrategias-K, oligótrofos, son capaces de competir en ambientes con recursos limitados (Fierer *et al.*, 2007). Si bien el concepto ha sido

generalizado a los microorganismos, las bacterias han desarrollado un amplio rango de estrategias de crecimiento y sobrevivencia para maximizar su éxito reproductivo. Es importante comprender los mecanismos que permiten estas adaptaciones en las bacterias y cómo están reflejadas en sus genomas.

Bases genómicas de estrategias de crecimiento bacterianas

Las bases moleculares de adaptación a las estrategias tróficas bacterianas han sido escasamente investigadas. Los estudios existentes se basan en la exploración de genomas de microorganismos marinos (Haggerty and Dinsdale, 2017; Lauro *et al.*, 2009). Lauro y colaboradores compararon los genomas de dos bacterias modelo de estilos de vida opuestos, oligotrófico y copiotrófico, a fin de relacionar las características genómicas de cada una de sus estrategias tróficas, comprender sus mecanismos de adaptación, y poder establecer predicciones de estilos de vida bacterianos a partir de datos genómicos.

En este trabajo, a partir del estudio de los metagenomas, se mostró que el contenido genómico de bacterias de barros activados refleja sus estrategias de crecimiento tróficas. Se ha observado que los copiótrofos tienen un mayor potencial genético para la replicación y transcripción, mientras que los oligótrofos poseen incrementado el metabolismo secundario, similar a lo observado en bacterias marinas (Lauro *et al.*, 2009). Esto muestra cómo las bacterias copiotróficas, promoviendo sus mecanismos de crecimiento y reproducción conseguirían obtener un crecimiento rápido, mientras que las estrategias K, favoreciendo sus funciones complementarias lograrían subsistir mejor en un ambiente de bajos recursos.

Se sabe que bacterias copiotróficas además de tener una abundante maquinaria de expresión génica tienden a tener transportadores de baja afinidad, que les permite un rápido crecimiento (Koch, 1971) mientras que, los oligótrofos de crecimiento lento tienen transportadores de alta afinidad que les permite crecer favorablemente en concentraciones bajas de nutrientes (Button,

1991). Estas diferencias, que han sido observadas principalmente en bacterias marinas, no pudieron ser confirmadas en este trabajo.

La presencia de múltiples copias de genes ARNr que sucede en las bacterias copiotróficas les permite lograr un aumento en el contenido de ribosomas requeridos para su crecimiento rápido cuando hay amplia disponibilidad de nutrientes (Roller *et al.*, 2016; Klappenbach *et al.*, 2000). Como se ha mencionado, se ha demostrado que el número de copias del operón *rrn* permite predecir la tasa y la eficiencia de crecimiento, siendo el número de operones ARN ribosomales directamente proporcional a las tasas de crecimiento e inversamente proporcional a la eficiencia de uso de carbono (Roller *et al.*, 2016).

Además del número copias de ARN ribosomales, existen otras características relacionadas con diferentes estrategias reproductivas que se han utilizado como estimadores de las tasas de crecimiento bacteriana, una de las cuales se ha explorado en este trabajo. Por un lado, se ha reportado que el sesgo en el uso de codones, es decir la utilización preferente de ciertos codones sinónimos sobre otros, es uno de los rasgos bacterianos, relacionados con sus estrategias de crecimiento, que mejor correlaciona con la velocidad de crecimiento (Vieira-Silva and Rocha, 2010). Este sesgo les permitiría maximizar el reclutamiento de los ARNt que se encuentran abundantes en los ribosomas de bacterias de crecimiento rápido y así lograr una traducción rápida y precisa (Bulmer, 1991). Por otro lado, se han propuesto otros estimadores, basados en la diferencia de cobertura en genomas bacterianos que resulta de la replicación bidireccional desde un único origen de replicación (Brown *et al.*, 2016; Vinekar *et al.*, 2016). El algoritmo utilizado en este trabajo, iRep permite estimar las tasas de replicación bacterianas a partir de genomas ensamblados que no necesariamente deben encontrarse como una única secuencia ensamblada (Brown *et al.*, 2016). Si bien en este trabajo no se ha podido establecer una relación entre el iRep y las observaciones que mostraron selección de bacterias según sus estrategias de crecimiento en los diferentes períodos, el número de bacterias a las que se les pudo calcular el iRep y además, el número de copias de operones *rrn*, no sería suficiente para establecer una conclusión.

Los resultados presentados proveen un aporte a la comprensión de los mecanismos de las adaptaciones de bacterias al ambiente (Lauro *et al.*, 2009; Roller *et al.*, 2016; Haggerty and Dinsdale, 2017; Vieira-Silva and Rocha, 2010). Mediante un análisis genómico sería posible inferir el estilo de vida trófico a partir de datos del genoma o del metagenoma, lo que representa una gran herramienta de estudio, considerando especialmente la dificultad del cultivo de microorganismos, sobre todo de los oligótrofos.

Rol de los protozoos en el sistema de barros activados y su uso como indicadores de calidad de tratamiento

Se ha postulado que los protozoos participan en los sistemas de barros activados reduciendo los sólidos en suspensión y por consiguiente la turbidez del efluente final, gracias al consumo de bacterias dispersas en el licor mezcla (Nicolau *et al.*, 2001).

En una típica sucesión de protozoos en una planta de barros activados (Figura 50), los flagelados son los primeros organismos dominantes, debido a su tipo de alimentación, predominantemente materia orgánica y bacterias dispersas. A medida que la materia orgánica comienza a ser degradada, los flagelados son reemplazados por los ciliados de nado libre, que luego de alcanzar un pico de abundancia son reemplazados por ciliados caminadores (*Hypotrichia*) y finalmente por los ciliados sésiles (*Peritichia*). Esta sucesión se explica por los estilos de vida y de alimentación de los protozoarios y por la selección ejercida en el sedimentador.

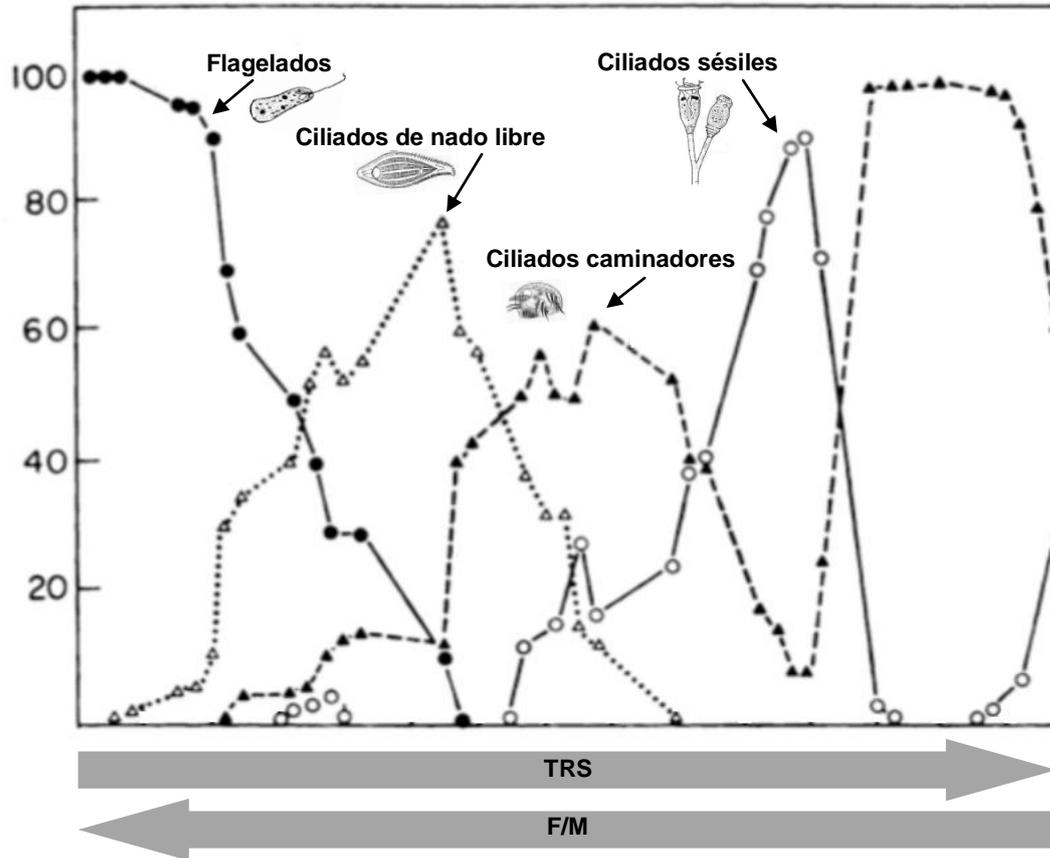


Figura 50. Sucesión de protozoos en una planta de barros activados (Adaptada de: Wastewater Biology: The microlife.)

Las condiciones óptimas de operación permiten obtener un barro con buena sedimentación. En un barro con buenas propiedades de sedimentación los protozoos mayoritarios suelen ser los ciliados sésiles y los caminadores (Madoni, 2011).

La observación microscópica de microorganismos se ha utilizado durante muchos años para relacionar la comunidad de protozoos con la calidad del efluente (Al-Shahwani and Horan, 1991; Madoni, 1994; Salvado *et al.*, 1995; Pérez-Uz *et al.*, 2010). Muchos de estos estudios se enfocaron en la búsqueda de indicadores biológicos tempranos para ser utilizados como herramienta para predecir el rendimiento de los sistemas de barros activados, así como para contribuir con el ajuste del proceso. Debido a que son sensibles a los cambios operacionales, los protozoos pueden utilizarse como indicadores biológicos de las condiciones existentes dentro del reactor, e indirectamente del rendimiento de remoción de materia orgánica (Araújo dos Santos *et al.*, 2014).

El índice biótico de barros, basado en la presencia y abundancia de ciertos grupos claves de protozoos (Madoni, 1994), es actualmente uno de los métodos usados para monitorear el desempeño de las plantas de barros activados (Arévalo et al, 2009). Algunos estudios han evaluado este índice en diversos sistemas (Arévalo *et al.*, 2009; Drzewicki and Kulikowska, 2011; Leal *et al.*, 2013; Araújo dos Santos *et al.*, 2014; Pedrazzani *et al.*, 2016). Si bien su aplicabilidad se ha considerado limitada, ya que no en todos los casos se ha encontrado una asociación entre el SBI y la calidad de los efluentes, la mayoría de los estudios publicados coincide en que este índice está relacionado con la estabilización del proceso. Sin embargo, en algunos casos, se observó una disminución en el rendimiento de la PTE, aunque los valores del SBI se mantuvieron altos (Drzewicki and Kulikowska, 2011). En otros, por el contrario, se observaron algunos valores muy bajos de SBI mientras que la eficiencia de los sistemas de tratamiento se mantuvo elevada (Arévalo *et al.*, 2009). Por esa razón algunos autores recomiendan que el uso de indicadores debería evaluarse empíricamente en cada planta en particular (Yiannakopoulou and Kaimakamidou, 2009).

Los resultados de este trabajo muestran consistencia entre los valores de SBI y los diferentes períodos operacionales que definen las condiciones del proceso. Hay una clara correspondencia entre los valores más bajos de SBI y el período operacionalmente perturbado, en los que se detectaron menor concentración de oxígeno disuelto y valores mayores de DQO en la salida del tratamiento. Inversamente, los valores más altos de SBI se corresponden con las etapas de mayor estabilidad y mejor calidad de efluente tratado. Por su parte, la correlación positiva significativa obtenida entre el SBI y la abundancia total de bacterias oligotróficas brinda un apoyo adicional al uso del SBI como indicador para el control de calidad del tratamiento.

La identificación y cuantificación de organismos eucariotas requiere de personal entrenado y un tiempo considerable de análisis. Si bien se trata de una tarea dificultosa, existen guías adecuadas para el reconocimiento de la identidad de los microorganismos. Por lo que, si bien incorporar la microscopía y cuantificación de eucariotas como herramienta de análisis de rutina diaria en las PTE significa un desafío, estos resultados apoyan la idea de que esta

herramienta podría ser de utilidad en plantas con problemas operativos frecuentes o en situaciones particulares de malfuncionamiento como indicadores tempranos de calidad del efluente o de la presencia de algún contaminante en el ingreso en plantas generalmente estables.

Desde la perspectiva del operador de planta, es importante destacar que la información que provee la observación microscópica complementa, pero no sustituye, las medidas como el OD, pH, potencial redox, turbidez, SSLM y SS30', que se utilizan habitualmente para el diagnóstico y toma de decisiones en la operación del sistema de tratamiento.

Interacciones fago-hospedero en ecosistemas de barros activados

La presencia de cualquier virus en un ambiente depende exclusivamente de la presencia de su hospedero. Por lo tanto, la comunidad de bacteriófagos en sistemas de barros activados se encuentra supeditada por la comunidad bacteriana co-existente. Existen escasos estudios sobre el metaviroma de barros activados (Parsley *et al.*, 2010; Tamaki *et al.*, 2012), y ninguno de ellos se ha enfocado en la dinámica de interacción fago-hospedero.

En este trabajo, el estudio del metaviroma permitió determinar la comunidad de bacteriófagos en un intervalo relativamente prolongado de tiempo, mostrando que los cambios en su composición resultaron ser el reflejo del patrón revelado por las poblaciones bacterianas, con respecto a su asociación con estrategias de crecimiento.

No obstante, si bien el estudio de metaviromas, en general, permite identificar secuencias virales sin necesidad de aislamiento previo, lo permite el descubrimiento de nuevos virus, en este trabajo se buscó identificar al hospedero de cada virus y analizar luego sus interacciones. Las principales estrategias utilizadas en este trabajo para la detección de interacciones fago-hospedero fueron la búsqueda de coincidencias en las secuencias de espaciadores CRISPR y la de similitud de secuencias. Estudios previos han utilizado alguna de estas estrategias para explorar el comportamiento de fagos e identificar a sus hospederos en distintos ecosistemas como el intestino

humano, ambientes marinos, entre otros (Anderson *et al.*, 2011; Cassman *et al.*, 2012; Stern *et al.*, 2012; Andersson and Banfield, 2008; Sanguino *et al.*, 2015; Berg Miller *et al.*, 2012; Minot *et al.*, 2013).

La búsqueda de espaciadores CRISPRs coincidentes entre el fago y la bacteria nos ha permitido asignar más de dos decenas de hospederos a fagos recuperados a partir del sobrenadante. Hemos determinado que los hospederos pertenecían a grupos bacterianos taxonómicos diversos. También se observaron ejemplos tanto de hospederos con más de un fago asociado, como de fagos que posiblemente cambien de hospedero.

Las estrategias que poseen los fagos para evadir la resistencia de bacterias a través de sus sistemas CRISPR/cas incluyen: recombinación, mutación, ya sea del protoespaciador o de la secuencia PAM, (motivo adyacente al protoespaciador) (Han and Deem, 2017; Weinberger *et al.*, 2012), y las recientemente descubiertas proteínas anti-CRISPR (Pawluk *et al.*, 2014). También se ha reportado que los fagos son capaces de codificar en su genoma sistemas CRISPR/cas similares a los de la bacteria, posiblemente a fin de evadir la inmunidad de su hospedero (Array *et al.*, 2016; Seed *et al.*, 2013). Es poco frecuente encontrar sistemas CRISPR/cas en plásmidos y profagos, sin embargo, se ha sugerido que podrían mediar el intercambio de CRISPR entre organismos (Godde and Bickerton, 2006). Se han encontrado en profagos de *Clostridium difficile* sistemas CRISPR similares al de su hospedero (Hargreaves *et al.*, 2014). En este trabajo se encontró en un profago de *Acidimicrobium* espaciadores coincidentes con los del sistema CRISPR del mismo hospedero en donde se encuentra integrado, y también se identificó un fago de *Gordonia*, un *Siphoviridae*, que contenía espaciadores y repeticiones similares a las de su hospedero. Si bien estos hallazgos resultan poco frecuentes, sugieren que las interacciones fago-bacteria serían aún más complejas de lo que se estima.

Por otro lado, la búsqueda de secuencias nucleotídicas similares entre los genomas bacterianos y los *contigs* de fagos, posibilitó la identificación de pares de interacción fago-bacteria en donde existen poblaciones del fago en ciclos de lisogenia y otras en ciclos líticos. También se observaron etapas en las que

sólo establecían lisogenia. VirSorter permitió detectar que el 60% de los genomas bacterianos contienen secuencias de fagos en sus genomas, lo que indicaría una gran proporción de bacterias con profagos en estado constante de lisogenia, ya que la gran mayoría de sus fagos no se han podido recuperar del sobrenadante. Estos resultados representan un interesante aporte, ya que los estudios de la lisogenia en ecosistemas son escasos en comparación con los de las infecciones líticas (Howard-Varona *et al.*, 2017). También hemos detectado al menos tres ejemplos de fagos que habrían incorporado material genético de su hospedero, en su mayoría proteínas relacionadas con la metilación. Esto podría consistir en una herramienta del fago en respuesta a la defensa de la bacteria, adquiriendo las proteínas de la bacteria capaces de introducir similares modificaciones en su ADN (metilaciones), y de este modo dejando de ser susceptibles a endonucleasas de restricción del hospedero. Los sistemas de defensa bacterianos contra sus fagos hasta el momento identificados son las modificaciones de superficie, sistemas de restricción/modificación (R/M), sistemas de infección abortiva, sistemas CRISPR/cas, sistema argonauta procarionta, BREX y DISARM (Ofir *et al.*, 2017). En muchos de ellos tanto las bacterias como los fagos, utilizan metilasas a fin de modificar su ADN. Recientemente se ha demostrado que metilasas identificadas en un fago metilaban la misma secuencia que la metiltransferasa regulada por su hospedero, y se ha deducido que estarían ocupando un papel importante en el ciclo del virus (Decewicz *et al.*, 2017).

Las múltiples interacciones fago-hospedero identificadas en un sistema de barros activados a lo largo de un período prolongado de tiempo muestran que los fagos estarían ocupando un rol importante en la modulación fina de bacterias en el sistema. Y por otro lado, muestra que el sistema de barros activados es un excelente modelo de estudio de estas interacciones.

CONCLUSIONES

Los resultados mostrados en este trabajo permiten ampliar nuestra comprensión acerca de la influencia de factores abióticos y bióticos sobre la estructura de comunidades microbianas de ecosistemas complejos y, a su vez, proveen información que contribuye al avance del conocimiento de interacciones multitróficas complejas.

Se determinó que el tiempo de retención de sólidos (TRS) y la existencia de perturbaciones en el sistema, seleccionaron a las comunidades microbianas según sus estrategias de crecimiento. Dicha conclusión se confirmó al determinarse que el número de operones de ARN ribosomal de las bacterias se condecía con sus estrategias tróficas. Estas observaciones relacionadas con las evidencias empíricas, permitieron establecer una asociación entre las estrategias de crecimiento bacterianas y su capacidad de floculación en sistemas de barros activados. Lo que indicó que las condiciones óptimas de operación seleccionan organismos floculentos y serían estrategias-K.

El análisis de los metagenomas, permitió investigar las bases genómicas de las diferentes estrategias de crecimiento bacterianas. La comparación de funciones en los metagenomas nos permitió determinar que las bacterias copiotróficas, promueven sus mecanismos de crecimiento y reproducción, lo que les facilita un crecimiento rápido, mientras que los organismos oligótrofos, favorecen su metabolismo secundario y funciones que les permiten ahorrar energía, a fin de subsistir mejor en un ambiente de bajos recursos. Estos resultados permiten aumentar nuestra comprensión sobre los mecanismos de adaptación de bacterias, y, asimismo, apoyan la hipótesis de que un análisis del genoma o del metagenoma puede ser utilizado como estimador del estilo de vida trófico bacteriano.

Por su parte, la comunidad de organismos eucariotas también respondió a las condiciones ambientales. Además fue posible evaluar su capacidad para ser empleados como microorganismos indicadores. La correlación positiva observada entre bacterias de bajo número de operones *rnm*, seleccionadas en un proceso estable, y el índice biótico de barros (SBI) apoya el uso del SBI como indicador para el control de calidad del tratamiento.

El estudio del metaviroma permitió mostrar que los cambios en la composición de bacteriófagos resultaron ser el reflejo del patrón revelado por las poblaciones bacterianas, con respecto a su asociación con estrategias de crecimiento. Las interacciones fago-hospedero identificadas en un sistema de barros activados a lo largo de un período prolongado de tiempo muestran que los fagos estarían ocupando un rol importante en la modulación fina de bacterias en el sistema. Estos resultados también representan un importante aporte a los escasos estudios existentes sobre la dinámica de interacción fago-hospedero en ecosistemas reales. Además, las observaciones puntuales realizadas sobre las interacciones de a pares fago-hospedero, en donde por ejemplo, se ha encontrado un profago con protoespaciadores, un fago con un arreglo de CRISPR y una bacteria que posee un profago además de su sistema CRISPR/cas, sugieren que, no solo los mecanismos de defensa, sino las interacciones en general entre bacteriófagos y bacterias, son aún más complejas de lo que se estima.

GLOSARIO

CRISPR: Repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas, por sus siglas en inglés (clustered regularly interspaced short palindromic repeats). Son loci de ADN en el genoma de procariotas, que contienen repeticiones cortas de secuencias (20-50pb). Entre las repeticiones se ubican espaciadores, que son segmentos cortos de ADN (20-70pb) que provienen de exposiciones previas a un virus. Asociadas a genes cas forman lo que se conoce como sistemas CRISPR/cas, el sistema inmune procariótico que otorga resistencia a fagos.

DBO: Demanda biológica de oxígeno.

DGGE: Electroforesis en gel con gradiente desnaturalizante.

DQO: Demanda química de oxígeno.

FISH: Hibridación fluorescente in situ.

IVB: Índice volumétrico de barros. El IVB es un índice utilizado periódicamente en las PTE como indicador de capacidad de sedimentación del barro. Indica el volumen que ocupa un gramo de barro.

OD: Oxígeno disuelto.

OTU: Unidad taxonómica operacional por sus siglas en inglés (operational taxonomic unit).

PAO: Organismos acumuladores de polifosfatos por sus siglas en inglés.

PTE: Planta de tratamiento de efluentes.

pVOG: Grupos ortólogos de virus procariotas.

SBI: Índice biótico de barros por sus siglas en inglés (Sludge Biotic Index). Es un indicador biológico de la calidad del efluente de salida de una PTE, establecido por Pablo Madoni en 1994. Se calcula a partir de la cuantificación de protozoos en muestras de barros activados.

SPE: Sustancias poliméricas extracelulares.

SSML: Sólidos suspendidos del licor mezcla. Concentración de sólidos suspendidos totales en el licor mezcla de la cámara de aireación de barros activados.

TRS: Tiempo de retención de sólidos. También llamado tiempo de residencia celular o edad de barros, es el tiempo promedio de permanencia de la biomasa en el sistema de barros activados.

ANEXO A. Características de genomas bacterianos reconstruidos a partir de los metagenomas de barros activados.

Nombre del BIN	Identidad	Tamaño del genoma (bp)	Número de contigs	% GC	N50	Complejidad (%)	Contaminación (%)	Heterogeneidad de especies	rrn operones
1_Bin10	<i>Proteobacteria;Deltaproteobacteria;Myxococcales;Kofferiaceae;Haliangium</i>	7905179	1574	74,02	6448	81,65	1,29	0	
1_Bin101	<i>Proteobacteria;Betaproteobacteria;Candidatus Accumulibacter;Candidatus Accumulibacter phosphatis(CAP);CAPfamily</i>	3758348	153	59,95	36256	99,05	0	0	3
1_Bin102	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	3847345	144	33,14	54797	96,55	1,64	25	2
1_Bin103	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3763787	355	67,31	16619	90,37	1,21	40	
1_Bin104	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	3639612	583	37,87	8536	87,03	0,74	0	
1_Bin105	<i>Bacteroidetes</i>	3735276	407	60,73	15214	88,17	1,23	50	
1_Bin108	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	2801543	640	71,58	5043	57,41	0,43	100	
1_Bin111	<i>Proteobacteria;Deltaproteobacteria;Myxococcales;Polyangiaceae;Sorangium</i>	7967773	492	69,43	36197	89,76	3	8,33	2
1_Bin110	<i>Bacteroidetes;Flavobacteria;Flavobacteriales</i>	3801240	321	64,56	20758	88,17	2,69	60	
1_Bin111	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Isosphaera</i>	3543311	1107	60,56	3473	58,68	0,58	0	
1_Bin112	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Hyphomicrobiaceae;Hyphomicrobium</i>	3734690	708	64,51	6503	74,36	0,4	0	
1_Bin114	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	3108510	933	72	3729	59,27	0	0	
1_Bin118	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	3587330	337	35,31	18352	91,3	0,74	50	1
1_Bin119	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	3716657	425	67,06	15320	92,5	3,42	25	1
1_Bin12	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	6709626	660	54,48	15749	80,55	0	0	1
1_Bin121	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	2990410	376	62,02	11934	80,82	0,3	0	
1_Bin123	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae</i>	2915873	431	64,82	9616	91,88	2,04	20	
1_Bin125	<i>Bacteroidetes</i>	3427084	50	37,58	156197	100	0	0	1
1_Bin126	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3267519	362	65,57	13018	89,42	0,66	25	
1_Bin128	<i>Bacteroidetes</i>	3182753	726	36,09	5156	83,03	0,79	33,33	1
1_Bin129	<i>Proteobacteria;Betaproteobacteria</i>	3294665	743	70,98	5520	78,27	2,04	75	
1_Bin134	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	3525267	416	42,29	12304	85,26	1,98	75	3
1_Bin135	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	3229177	712	36,5	5826	73,75	0,62	0	1
1_Bin136	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	3293204	800	53,29	4799	67,85	1,77	50	
1_Bin141	<i>Proteobacteria;Alphaproteobacteria;Rhodospirillales;Acetobacteraceae</i>	3079045	857	70,74	4103	63,57	0,41	100	
1_Bin142	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	3137507	771	72,63	4853	67,49	2,48	60	7
1_Bin147	<i>Bacteroidetes</i>	3161810	37	36,11	245594	98,57	0,48	0	1
1_Bin15	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	4480915	525	65,53	13985	91,88	2,66	30	
1_Bin150	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	3217468	204	70,64	37799	97,75	0,78	0	2
1_Bin16	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Rhodopirellula</i>	7230235	138	51,94	122036	98,75	1,18	0	1
1_Bin162	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	2997686	253	67,57	24441	94,13	3,38	57,69	
1_Bin163	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	2989035	19	64,84	285503	99,01	0,73	33,33	
1_Bin164	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beij</i>	1976818	913	68,2	2089	46,49	9,42	9,09	

	<i>erinckiacae</i>								
1_Bin165	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Hyphomicrobiaceae;Hyphomicrobium</i>	2921135	396	60,3	11287	90,28	0,4	50	
1_Bin166	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	2845193	326	61,39	14291	91,2	0,85	0	
1_Bin169	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	2965799	447	65,69	8962	85,04	1,3	50	
1_Bin17	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	6464719	2107	53,02	3307	55,65	0	0	
1_Bin170	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiacae</i>	2508746	615	57,88	4940	72,87	0,94	33,33	2
1_Bin171	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	3101100	458	37,26	11734	81,2	1,54	80	3
1_Bin172	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	2724079	635	66,34	5229	80,96	1,48	36,36	
1_Bin173	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae</i>	2879005	164	58,06	32412	96,13	0,06	0	1
1_Bin175	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	3085891	440	46,25	10587	64,45	1,02	66,67	1
1_Bin176	<i>Bacteroidetes;Flavobacteria;Flavobacteriales;Flavobacteriaceae;Flavobacterium</i>	2824733	49	32,61	120961	97,88	0	0	4
1_Bin178	<i>Bacteroidetes</i>	3141567	809	59,83	4756	67,3	7,44	73,33	1
1_Bin18	<i>Bacteroidetes;Cytophaga;Cytophagales;Cytophagaceae;Runella</i>	6048558	400	43,4	24463	82,97	0,6	100	2
1_Bin180	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	2859416	205	68,07	30776	90,9	1,71	0	
1_Bin181	<i>Bacteroidetes</i>	2941727	410	61,64	10467	82,21	1,97	50	
1_Bin184	<i>Proteobacteria;Alphaproteobacteria;Caulobacterales;Caulobacteraceae;Phenylobacterium</i>	2010339	659	69,15	3329	51,02	0	0	
1_Bin189	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	2485192	623	68,13	4658	64,65	0,68	50	
1_Bin19	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	3106552	285	66,73	16502	85,71	1,88	33,33	
1_Bin191	<i>Actinobacteria;Actinobacteria;Solirubrobacteriales;Conexibacteraceae;Conexibacter</i>	2548631	358	63,61	12176	83,07	9,48	80	
1_Bin192	<i>Proteobacteria;Betaproteobacteria;Rhodocyclales;Rhodocyclaceae</i>	1644216	588	64,94	2843	40,59	4,9	75,76	
1_Bin193	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	2880248	647	41,52	5914	58,58	1,49	66,67	1
1_Bin197	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	2558251	182	68,52	30228	86,51	4,27	40	
1_Bin198	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	2515157	223	58,89	18084	88,89	0,57	50	
1_Bin199	<i>Bacteroidetes;Flavobacteria;Flavobacteriales;Flavobacteriaceae;Flavobacterium</i>	2218410	627	35,13	3998	69,93	0,96	40	
1_Bin201	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiacae</i>	2774623	548	66,77	6530	76,77	2,14	33,33	
1_Bin202	<i>Proteobacteria;Betaproteobacteria</i>	2253515	722	65,1	3383	56,51	1	0	
1_Bin205	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	2460759	432	36,44	7870	76,81	3,45	85,71	3
1_Bin206	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	1828156	699	65,96	2694	45,45	0	0	
1_Bin207	<i>Spirochaetes;Spirochaetia;Spirochaetales</i>	2251275	670	53,76	3872	61,5	2,94	16,67	2
1_Bin208	<i>Actinobacteria;Actinobacteria;Actinomycetales;Nocardiodaceae;Nocardioidea</i>	1876468	564	64,59	3605	64,1	0,26	100	
1_Bin209	<i>Actinobacteria;Actinobacteria;Solirubrobacteriales;Conexibacteraceae;Conexibacter</i>	2311049	360	62,33	8420	88,75	1,72	54,55	
1_Bin210	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	2021380	615	66,45	3678	52,93	0,86	100	
1_Bin212	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	2284083	623	38,94	4100	61,2	4,68	69,23	2
1_Bin214	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	1543575	481	65,87	3454	49,64	0,47	0	
1_Bin217	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Phyllobacteriaceae</i>	2135742	711	65,64	3243	48,75	1,72	100	
1_Bin218	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	1716622	184	67,37	12928	53,95	0,27	100	
1_Bin220	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae</i>	1156178	251	66,83	5159	35,34	0	0	

1_Bin223	<i>Actinobacteria;Actinobacteria;Actinomycetales;Microbacteriaceae</i>	2213419	332	70,43	10526	93,65	1,42	50	
1_Bin224	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	2199296	511	42,02	4912	71,98	2,96	85,71	
1_Bin229	<i>Proteobacteria;Gammaproteobacteria;Pseudomonadales;Moraxellaceae;Acinetobacter</i>	1998143	254	41,19	12324	95,22	0,65	33,33	
1_Bin23	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	6274754	496	63,9	21169	97,26	3,31	18,75	4
1_Bin233	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	1889320	531	67,05	3892	61,42	6,57	75	2
1_Bin235	<i>Unclassified Bacteria</i>	1231193	297	38,71	4873	49,87	2,09	33,33	
1_Bin237	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	2160277	217	57,96	22036	90,51	6,71	73,68	
1_Bin240	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	1905968	623	62,46	3355	53,87	0,64	12,5	
1_Bin244	<i>Actinobacteria;Actinobacteria;Actinomycetales;Nocardoidaceae;Nocardioidea</i>	2130113	263	61,84	12435	82,86	2,68	87,5	
1_Bin247	<i>Actinobacteria;Actinobacteria</i>	2037711	147	39,64	39096	92,74	11,54	64,29	
1_Bin248	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	2392063	542	38,06	5693	67,07	8,13	90	1
1_Bin252	<i>Proteobacteria;Gammaproteobacteria</i>	1965254	327	51,17	9183	83,1	6,94	85,19	
1_Bin253	<i>Proteobacteria;Alphaproteobacteria;Caulobacterales;Caulobacteraceae</i>	1830781	644	66,87	3017	52,56	0,63	25	
1_Bin254	<i>Unclassified Bacteria</i>	1871460	570	38,87	3562	57,07	0,02	100	
1_Bin257	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	1441469	326	35,07	5621	79,61	1,64	25	
1_Bin258	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	2130583	488	72,91	5329	57,99	1,79	60	
1_Bin259	<i>Unclassified Bacteria</i>	1767914	624	69,31	2844	52,92	1,3	66,67	
1_Bin261	<i>Bacteroidetes</i>	1783971	578	38,46	3100	53,92	1,11	66,67	
1_Bin262	<i>Bacteroidetes;Bacteroidia;Bacteroidales;Porphyromonadaceae;Paludibacter</i>	1417304	533	42,08	2677	50,48	1,49	66,67	7
1_Bin265	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae</i>	2016044	414	64,64	5415	61,78	0,45	0	
1_Bin266	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	1691059	481	35	3983	58,04	0,49	0	
1_Bin270	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	1667068	511	38,78	3496	58,32	0,41	50	
1_Bin271	<i>Chlamydiae;Chlamydia;Chlamydiales</i>	1528399	489	41,65	3433	68,1	0,81	0	
1_Bin274	<i>Bacteroidetes</i>	1676443	429	45,51	4480	65,63	0,25	0	3
1_Bin277	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	1953065	474	60,86	5379	62,78	5,16	100	2
1_Bin281	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	1542075	493	34,24	3435	54,08	0	0	
1_Bin287	<i>Actinobacteria;Actinobacteria;Solirubrobacterales;Conexibacteraceae;Conexibacter</i>	1547623	391	66,58	4649	55,49	7,69	80	
1_Bin29	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprosiraceae;Haliscomenobacter</i>	5770548	1180	46,62	6139	78,95	1,77	0	
1_Bin297	<i>Firmicutes;Clostridia;Clostridiales;Ruminococcaceae</i>	1253915	346	57,2	4017	51,61	1,79	100	
1_Bin306	<i>Firmicutes;Bacilli;Lactobacillales</i>	1135776	315	46,86	3914	65,22	1,37	0	
1_Bin307	<i>Candidatus_Saccaribacteria</i>	1086865	48	45,33	118096	71,3	0	0	
1_Bin308	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	1031087	418	31,01	2497	46,21	0,27	0	
1_Bin319	<i>Archaea;Euryarchaeota;Methanomicrobia;Methanosarcinales;Methanosaetaceae;Methanosaeta;Methanosaeta concilii</i>	1164473	250	53,03	5730	56,82	0,65	0	
1_Bin32	<i>Proteobacteria;Alphaproteobacteria</i>	5632013	62	64,48	178600	97,47	0	0	1
1_Bin322	<i>Bacteroidetes</i>	1035577	197	51,38	6621	52,76	0,21	100	
1_Bin33	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae</i>	5786162	822	51,5	9618	71,42	1,18	100	
1_Bin331	<i>Candidatus_Saccaribacteria</i>	1043663	60	43	76204	63,58	0	0	
1_Bin334	<i>Unclassified Bacteria</i>	972267	100	47,36	14483	63,79	3,7	25	
1_Bin336	<i>Candidatus_Saccaribacteria</i>	989902	17	50,53	166482	64,04	0	0	
1_Bin34	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	5233697	1226	69,58	5339	69,32	5,48	65,22	
1_Bin340	<i>Firmicutes;Bacilli;Lactobacillales;Streptococcaceae;Streptococcus</i>	806718	219	40,09	4037	50,99	0,37	100	
1_Bin346	<i>Unclassified Bacteria</i>	892411	16	45,89	102021	66,82	2,78	0	
1_Bin347	<i>Unclassified Bacteria</i>	790005	126	50,01	9201	77,09	2,31	66,67	
1_Bin35	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	5277815	117	70,52	91013	94,87	2,14	0	
1_Bin350	<i>Chloroflexi</i>	865983	88	51,26	15393	74,55	0	0	2
1_Bin351	<i>Candidatus_Saccaribacteria</i>	802082	26	47,3	79022	63,72	0,85	0	
1_Bin352	<i>Unclassified Bacteria</i>	820977	147	46,22	7486	56,9	1,98	100	

1_Bin355	<i>Candidatus_Saccaribacteria</i>	847582	27	39,51	42709	64,42	0	0	1
1_Bin36	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae</i>	5326293	1157	68,18	5566	82,63	5,68	33,33	
1_Bin361	<i>Unclassified Bacteria</i>	830251	70	46,62	23015	62,75	1,85	50	
1_Bin362	<i>Candidatus_Saccaribacteria</i>	801225	102	37,39	13437	64,74	0,99	0	2
1_Bin363	<i>Candidatus_Saccaribacteria</i>	845061	100	41,6	16870	59,72	1,03	50	2
1_Bin364	<i>Candidatus_Saccaribacteria</i>	811432	56	44,39	35810	66,67	0,93	100	
1_Bin365	<i>Unclassified Bacteria</i>	659030	167	40,49	4506	53,06	1,12	100	
1_Bin37	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	5418484	366	41,93	34111	95,71	0,33	0	
1_Bin370	<i>Unclassified Bacteria</i>	775515	143	45,64	7047	62,24	0	0	
1_Bin371	<i>SR1</i>	740201	185	32,56	4902	57,45	0	0	2
1_Bin377	<i>Firmicutes</i>	791102	113	47,69	13009	62,96	5,56	57,14	1
1_Bin38	<i>Proteobacteria;Deltaproteobacteria;Myxococcales;Kofleriaceae;Haliangium</i>	4725262	1691	71,27	2913	52,93	0	0	3
1_Bin380	<i>Candidatus_Saccaribacteria</i>	756713	65	36,96	26463	63,45	1,1	100	1
1_Bin39	<i>Proteobacteria;Deltaproteobacteria;Myxococcales;Polyangiaceae;Sorangium</i>	4785347	1457	72,44	3686	56,4	2,63	33,33	6
1_Bin40	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiaceae</i>	4293708	174	67,35	87763	99,3	3,74	0	4
1_Bin41	<i>Actinobacteria;Actinobacteria;Actinomycetales;Gordoniaceae;Gordonia</i>	5137547	62	66,42	152418	98,58	0,66	50	2
1_Bin418	<i>Candidatus_Saccaribacteria</i>	521659	108	38,99	6563	43,51	4,32	83,33	
1_Bin42	<i>Chloroflexi;Anaerolineae;Anaerolineales;Anaerolineaceae;Anaerolinea</i>	4926308	900	65,27	8911	89,01	14,7	79,31	
1_Bin45	<i>Acidobacteria</i>	4999488	902	69,27	7905	76,95	2,83	20	
1_Bin5	<i>Proteobacteria;Deltaproteobacteria;Myxococcales;Polyangiaceae;Sorangium</i>	12856257	1379	71,01	14333	90,5	3,94	0	2
1_Bin51	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Leptothrix;Leptospira</i>	4496455	899	69,65	6230	79,8	1,36	60	4
1_Bin52	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	4858650	148	71,47	67401	95,73	3,42	0	
1_Bin53	<i>Proteobacteria;Betaproteobacteria</i>	4793486	289	69,42	24936	90,75	1,1	42,86	
1_Bin54	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Bradyrhizobiaceae</i>	4790114	578	64,12	28891	91,74	26,34	12,9	
1_Bin55	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Paracoccus</i>	2307378	588	65,82	4408	55,97	2,2	69,23	
1_Bin56	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	4206083	1111	55,31	4346	69,38	0	0	
1_Bin57	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	4351947	1255	59,31	3771	71,17	3,45	33,33	
1_Bin58	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	4146732	818	40,9	6667	83,36	1,51	50	
1_Bin59	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	3860448	89	63,99	96513	97,48	0,8	0	
1_Bin6	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiaceae</i>	3154304	859	63,41	4200	67,98	2,98	9,09	
1_Bin61	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	4236871	603	69,31	10541	86,32	2,56	25	
1_Bin62	<i>Proteobacteria;Gammaproteobacteria</i>	4372890	165	71,69	87847	85,2	2,29	11,11	
1_Bin65	<i>Acidobacteria</i>	4415542	710	70,73	8967	81,13	3,47	0	2
1_Bin66	<i>Verrucomicrobia;Verrucomicrobiae;Verrucomicrobiales;Verrucomicrobiaceae;Akkermansia</i>	4151494	1180	59,95	3959	47,34	3,45	0	
1_Bin67	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	4049044	1281	54,67	3358	59,51	0,37	100	
1_Bin68	<i>Chloroflexi;Anaerolineae;Anaerolineales;Anaerolineaceae;Anaerolinea</i>	3993087	1099	64,01	4169	70,2	0	0	
1_Bin69	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales</i>	4329948	110	45,53	73815	99,51	0,49	0	1
1_Bin7	<i>Proteobacteria;Deltaproteobacteria;Myxococcales;Polyangiaceae;Sorangium</i>	12030210	366	70,81	62398	94,84	4,87	0	2
1_Bin70	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	4297487	52	37,31	245658	98,02	0,5	0	2
1_Bin71	<i>Bacteroidetes;Cytophagia;Cytophagales</i>	4282435	53	41,02	184558	99,73	0	0	2
1_Bin72	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	3857124	1097	56,62	3822	60,81	0	0	
1_Bin74	<i>Unclassified Bacteria</i>	3530063	1137	69,62	3330	60,55	0,85	0	
1_Bin75	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	3966697	668	68,99	8470	71,18	3,08	70	
1_Bin76	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidi</i>	4194288	93	65,55	81969	94,87	0,85	0	

	<i>microbiaceae;Acidimicrobium</i>									
1_Bin77	<i>Proteobacteria;Gammaproteobacteria;Chromatiales;Chromatiaceae;Allochrochromatium</i>	4268067	124	65,43	61199	90,47	0,93	20	2	
1_Bin78	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Bradyrhizobiaceae</i>	4108406	46	63,47	158775	99,1	0,85	50	2	
1_Bin80	<i>Actinobacteria;Actinobacteria;Actinomycetales;Mycobacteriaceae;Mycobacterium</i>	4078682	537	67,54	11112	87,62	0,88	60		
1_Bin81	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	4184935	257	65,04	29625	93,34	1,51	62,5		
1_Bin82	<i>Bacteroidetes</i>	4155102	170	39,02	57997	99,01	1,48	0	1	
1_Bin83	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Hyphomonadaceae;Hirschia</i>	3951458	349	63,17	18283	93,61	3,24	40,91	2	
1_Bin84	<i>Unclassified Bacteria</i>	4008020	433	59,14	13961	92,17	3,64	50		
1_Bin87	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales</i>	2867799	589	65,62	6361	55,64	0	0	3	
1_Bin88	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae</i>	2176557	443	63,61	6178	46,9	0	0		
1_Bin89	<i>Bacteroidetes;Flavobacteriia;Flavobacteriales</i>	3988975	298	61,77	23357	96,67	2,87	66,67	1	
1_Bin9	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Pirellula</i>	5595356	1354	63,38	4843	65,78	1,15	100		
1_Bin90	<i>Bacteroidetes;Sphingobacteriia;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	3985395	467	44,57	12519	88,14	1,23	33,33		
1_Bin91	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3979202	107	66,05	123449	98,39	0,25	0		
1_Bin93	<i>Proteobacteria;Gammaproteobacteria</i>	3974298	79	44,22	95036	98,64	0,17	0	1	
1_Bin95	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	3484239	1009	68,4	3931	67,54	3,24	45,45		
1_Bin96	<i>Bacteroidetes;Sphingobacteriia;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	3818267	881	39,46	5237	70,67	1,97	16,67		
1_Bin98	<i>Bacteroidetes;Sphingobacteriia;Sphingobacteriales;Saprosiraceae;Haliscomenobacter</i>	4047705	702	49,12	8471	76,73	0,25	100	2	
2_Bin100	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3891167	96	64,33	131652	96,63	0,76	33,33		
2_Bin113	<i>Bacteroidetes</i>	3302801	618	38,83	6864	75,74	0	0		
2_Bin115	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiaceae</i>	3597253	125	63,96	81083	93,62	14,36	95,24	3	
2_Bin116	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	3667631	662	69,53	7797	70,77	4,28	75		
2_Bin117	<i>Proteobacteria;Alphaproteobacteria;Rhodospirillales;Acetobacteraceae</i>	3337329	982	70,91	3777	49,54	0	0		
2_Bin120	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	3172205	178	64,59	30199	93,77	0,32	20		
2_Bin130	<i>Proteobacteria;Alphaproteobacteria</i>	1648381	523	62,78	3221	37,07	1,72	100		
2_Bin131	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	3644790	769	69,8	6419	68,71	1,17	33,33	1	
2_Bin133	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	3339294	405	65,53	18058	86,35	2,33	50	3	
2_Bin138	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	3213126	464	66,39	9701	94,65	0,25	100	3	
2_Bin139	<i>Bacteroidetes</i>	3250691	119	30,93	56258	97,62	0,48	0	2	
2_Bin14	<i>Bacteroidetes;Cytophagia;Cytophagales;Cytophagaceae;Runella</i>	7484387	581	42	20591	83,65	3,1	76,47	4	
2_Bin143	<i>Acidobacteria</i>	3313055	289	56,37	17379	85,99	1,76	0	1	
2_Bin145	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	3198467	530	70,82	9282	79,46	1,94	28,57	1	
2_Bin151	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	3026885	335	69,42	13603	71,6	0,5	0		
2_Bin152	<i>Acidobacteria;Actinobacteria;Actinomycetales;Nocardioideae;Nocardioidea</i>	3067694	205	68,14	28493	95,63	0,78	100	4	
2_Bin157	<i>Proteobacteria;Betaproteobacteria</i>	2744666	932	68,89	3057	44,92	2,59	50		
2_Bin160	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiaceae</i>	3054732	383	66,8	11962	86,64	2,74	41,67		
2_Bin161	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	3066818	277	66,63	19236	88,79	0,85	0	1	
2_Bin179	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	3139771	485	69,2	10510	74,2	0,72	50	1	
2_Bin183	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae</i>	2601647	159	61,43	27171	94,06	0,22	0		
2_Bin185	<i>Bacteroidetes;Flavobacteriia;Flavobacteriales;Flavobacteriaceae</i>	1138673	464	37,57	2410	39,8	2,7	50		
2_Bin187	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;</i>	1811596	600	67,18	3285	41,55	10,98	81,48		

Comamonadaceae									
2_Bin190	<i>Bacteroidetes;Flavobacteria;Flavobacteriales;Flam meovirgaceae;Flavobacterium</i>	2343333	613	48,7	4587	73,99	1,8	50	
2_Bin192	<i>Proteobacteria;Betaproteobacteria;Rhodocyclales;Rhodocyclaceae</i>	2039707	246	64,58	11994	43,34	0	0	
2_Bin20	<i>Acidobacteria</i>	5772553	963	69,42	7761	85,9	5,22	12,5	
2_Bin200	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	2444386	659	61,34	4374	53,32	1,2	55,56	6
2_Bin203	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Chitinophaga;Chitinophaga pinensis</i>	2266746	610	43,61	4205	71,35	0,69	100	3
2_Bin21	<i>Chloroflexi</i>	6157184	515	62,6	18236	89,39	1,82	0	
2_Bin215	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	2596077	652	71,24	4742	61,55	1,05	100	3
2_Bin216	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	2502336	661	65,83	4538	63,15	1,27	80	1
2_Bin218	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	1226622	478	68,55	2561	42,11	2,73	76,47	
2_Bin219	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	2145308	528	70,36	4545	62,97	0,57	50	
2_Bin24	<i>Chloroflexi</i>	6104963	980	64,48	8475	80,96	2,73	0	
2_Bin243	<i>Actinobacteria;Actinobacteria;Actinomycetales;Mycobacteriaceae</i>	1730214	577	70,74	3160	46,99	1,36	33,33	
2_Bin25	<i>Bacteroidetes;Sphingobacteria;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	6074101	215	36,38	66751	97,93	1,4	0	1
2_Bin27	<i>Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae</i>	5572021	926	64,13	8442	88,65	3,76	11,54	
2_Bin276	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	1904600	341	69,13	7279	60,67	0	0	
2_Bin293	<i>Candidatus_Saccaribacteria</i>	961643	187	45,73	7027	54,81	4,02	33,33	3
2_Bin30	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Bradyrhizobiaceae</i>	1841654	651	61,33	2959	56,38	1,72	100	
2_Bin31	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Planctomyces</i>	5703611	675	58,64	11785	88,73	2,45	25	3
2_Bin43	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Leptothrix;Leptothrix chlodnii</i>	5029507	358	68,59	28784	93,85	2,18	16,67	
2_Bin44	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	4732194	569	74,13	11480	80,64	1,23	33,33	
2_Bin46	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	1374908	380	71,47	3954	37,03	0,18	0	
2_Bin47	<i>Nitrospirae;Nitrospira;Nitrospirales;Nitrospiraceae;Nitrospira</i>	3724462	154	58,83	50445	94,63	2,47	57,14	
2_Bin6	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Hyphomicrobiaceae;Hyphomicrobium</i>	2703188	640	61,88	4912	52,87	0,88	0	
2_Bin60	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Beijerinckiaceae</i>	4582323	247	65,52	33547	93,7	0,52	0	1
2_Bin63	<i>Bacteroidetes;Cytophagia;Cytophagales;Cytophagaceae</i>	4902831	1076	42,1	6186	65,44	3,87	66,67	3
2_Bin73	<i>Bacteroidetes;Cytophagia;Cytophagales</i>	4217761	87	45,61	81103	96,47	0,99	20	1
2_Bin8	<i>Unclassified Bacteria</i>	11708180	747	55,46	24751	87,68	6,36	0	5
2_Bin86	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	4085495	117	66,65	67886	97,78	0,47	0	3
2_Bin87	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales</i>	4024885	327	65,58	19867	66,72	1,72	100	
2_Bin88	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3373900	629	63,49	6713	76,03	8,01	92,59	2
2_Bin9	<i>Planctomycetes;Planctomycetia;Planctomycetales;Planctomycetaceae;Pirellula;Pirellula staleyii</i>	2974498	1124	63,71	2658	34,5	0	0	
2_Bin92	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3396841	668	62,77	6312	80,91	1,67	33,33	
2_Bin94	<i>Actinobacteria;Actinobacteria;Actinomycetales;Mycobacteriaceae;Mycobacterium</i>	3686314	811	64,73	5534	86,63	2,26	25	
3_Bin107	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3571722	437	67,81	11289	88,77	1,06	60	
3_Bin122	<i>Proteobacteria;Betaproteobacteria;Rhodocyclales;Rhodocyclaceae;Dechloromonas</i>	2821003	266	63,01	21968	89,38	1,04	30	
3_Bin130	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales</i>	1943523	380	62,57	6017	43,1	5,17	100	
3_Bin132	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	3346140	43	65,56	159292	99,21	0,22	0	
3_Bin144	<i>Actinobacteria;Actinobacteria;Actinomycetales;Intrasporangiaceae;Intrasporangium</i>	3196582	494	68,52	10681	81,45	3,05	66,67	
3_Bin154	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales</i>	3046204	678	68,45	5399	75,01	0,91	60	

	<i>s;Rhodobacteraceae;Rhodobacter</i>								
3_Bin168	<i>Actinobacteria;Actinobacteria;Acidimicrobiales;Acidimicrobiaceae;Acidimicrobium</i>	2346951	748	69	3380	58,69	2,14	0	
3_Bin18	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae</i>	152008	52	42,57	2838	91,61	0,3	100	
3_Bin227	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	1457105	350	67,28	4789	41,61	0,69	100	
3_Bin236	<i>Actinobacteria;Actinobacteria;Actinomycetales</i>	2239751	392	70,84	7575	72,99	0	0	
3_Bin246	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	1932602	531	66,17	4187	55,26	0,58	0	
3_Bin28	<i>Proteobacteria;Alphaproteobacteria;Rhodospirillales;Acetobacteraceae</i>	5806180	202	70,66	55879	98,54	1,06	50	
3_Bin48	<i>Proteobacteria;Alphaproteobacteria;Sphingobacteriales;Sphingobacteriaceae</i>	1957969	486	61,29	4755	73,63	1,37	60	
3_Bin50	<i>Bacteroidetes;Sphingobacteriia;Sphingobacteriales;Saprospiraceae;Haliscomenobacter</i>	4936925	921	49,61	6760	74,34	0,62	0	
3_Bin6	<i>Proteobacteria;Alphaproteobacteria;Rhizobiales;Bradyrhizobiaceae</i>	2830358	130	62,73	30151	93,67	0,32	0	
3_Bin64	<i>Proteobacteria;Betaproteobacteria;Burkholderiales</i>	4340068	143	68,78	75872	97,77	0,47	0	1
4_Bin19	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	1657376	612	68,31	2755	43,79	4,31	33,33	
4_Bin227	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium</i>	1746387	546	67,2	3495	59,76	9,6	91,53	2
4_Bin48	<i>Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae</i>	2852766	201	57,92	22019	91,82	4,29	69,57	
5_Bin156	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	2061694	578	66,94	3967	46,07	10,92	65,85	
5_Bin187	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	3173513	841	66,73	4460	63,17	22,99	97,67	2
5_Bin230	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	2418492	434	66,73	8112	62,54	0	0	
5_Bin99	<i>Proteobacteria;Alphaproteobacteria;Rhodobacterales;Rhodobacteraceae;Rhodobacter</i>	3879189	97	65,17	104167	96,55	0,3	66,67	
6_Bin156	<i>Proteobacteria;Betaproteobacteria;Burkholderiales;Comamonadaceae</i>	1822207	550	66,74	3607	44,29	8,2	67,86	

ANEXO B. Comparación de taxonomía detectada con los diferentes métodos de secuenciación.

MAG: genomas ensamblados a partir de las lecturas del metagenoma obtenidas por la secuenciación al azar. (MAG por sus siglas: metagenome-assembled-genome)

MA16S: genes ARNr 16S ensamblados a partir de las lecturas del metagenoma obtenidas por la secuenciación al azar (por su analogía con MAG)

Amp16S: amplicones de la región V3-V4 del gen ARNr 16S obtenidos por la secuenciación de amplicones.

Filo/Clase	MAG	MA16S	Amp16S
<i>Acidobacteria</i>	1,26	0,47	1,71
<i>Actinobacteria</i>	14,40	11,85	10,21
<i>Bacteroidetes</i>	20,36	16,06	19,00
<i>Candidatus Saccharibacteria</i>	4,34	5,74	5,55
<i>Chloroflexi</i>	0,97	2,85	4,34
<i>Firmicutes</i>	0,99	2,25	6,22
<i>Nitrospirae</i>	0,44	0,64	0,45
<i>Planctomycetes</i>	2,51	2,64	4,78
<i>Alphaproteobacteria</i>	21,40	19,82	15,62
<i>Betaproteobacteria</i>	19,86	27,21	9,92
<i>Deltaproteobacteria</i>	1,65	4,24	2,42
<i>Gammaproteobacteria</i>	8,55	6,26	5,54
<i>Proteobacteria</i> no clasificada	0,00	0,00	1,53
<i>Verrucomicrobia</i>	0,14	0,52	1,53
<i>Bacteria</i> no clasificada	2,61	0,00	8,86

Tabla 1b. Porcentajes relativos de abundancia en el total de las muestras de filos bacterianos y clases de proteobacterias determinados a partir de MAG, MA16S y Amp16S.

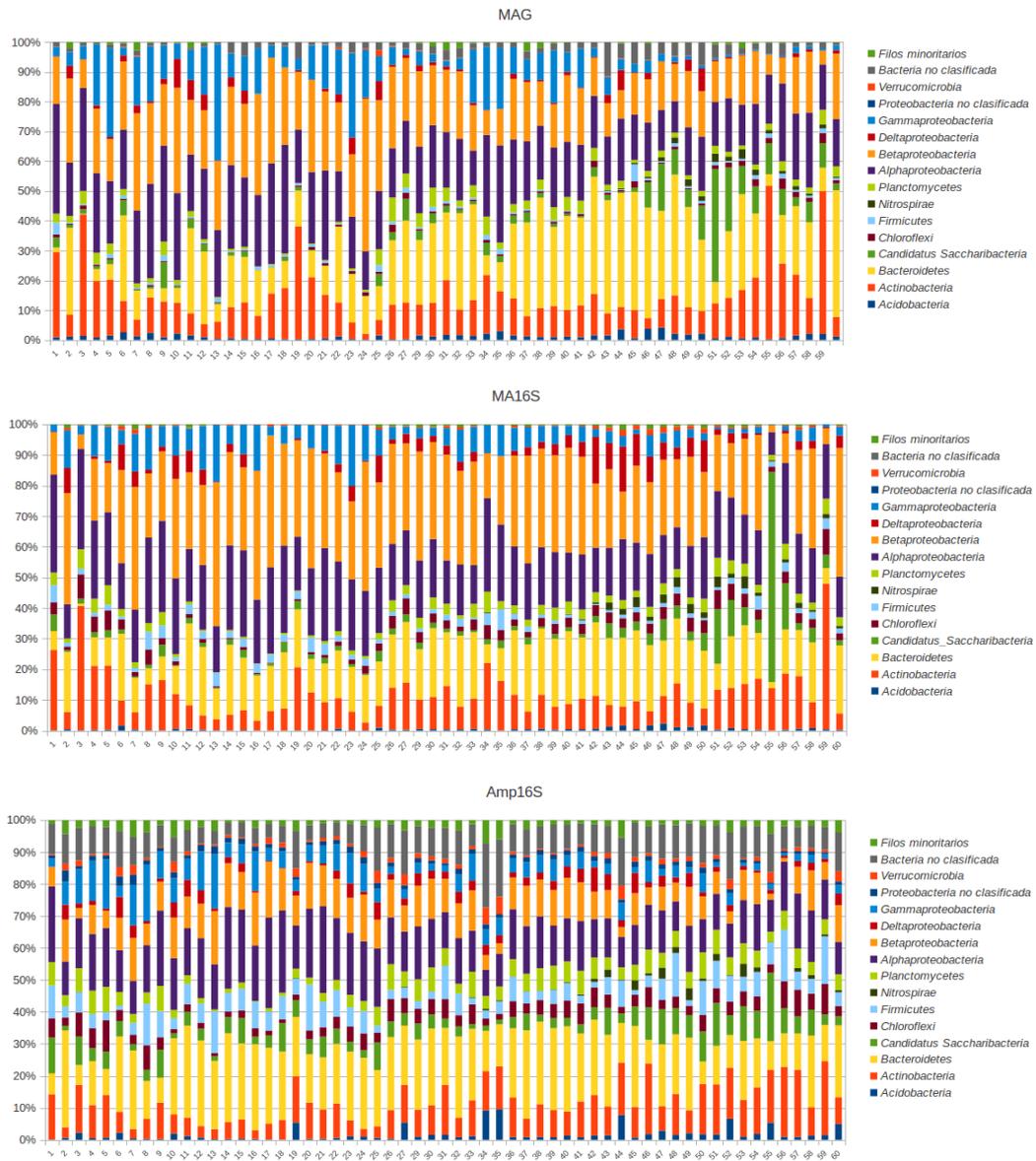


Figura 1b. Comparación de la abundancia en las 60 muestras de filos y clases de proteobacterias obtenidas a partir de Amp16S, MA16S y MAG.

ANEXO C. Figuras de abundancia en el tiempo, de pares de interacción fago-hospedero detectados a partir de la búsqueda de espaciadores CRISPRs coincidentes en el genoma de la bacteria y del fago.

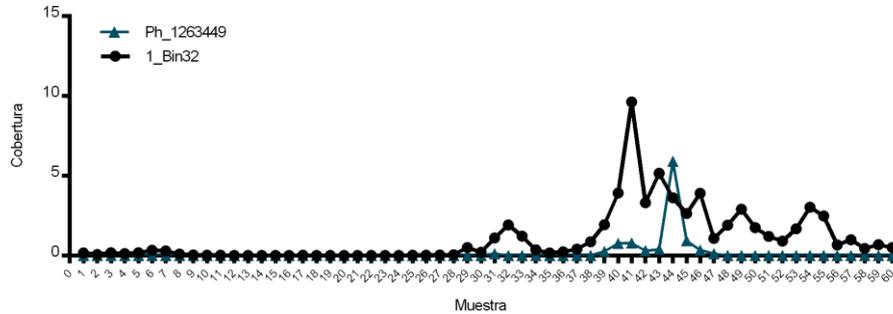


Figura 1c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin32 (*Alphaproteobacteria*) y el fago Ph_1263449

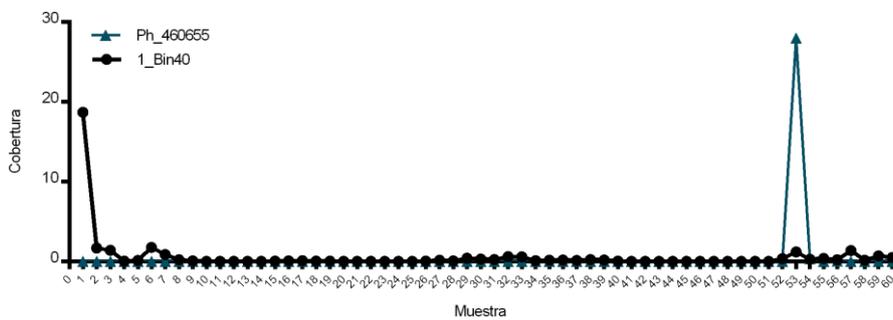


Figura 2c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin40 (*Alphaproteobacteria*, *Rhizobiales*, *Beijerinckiaceae*) y el fago Ph_460655

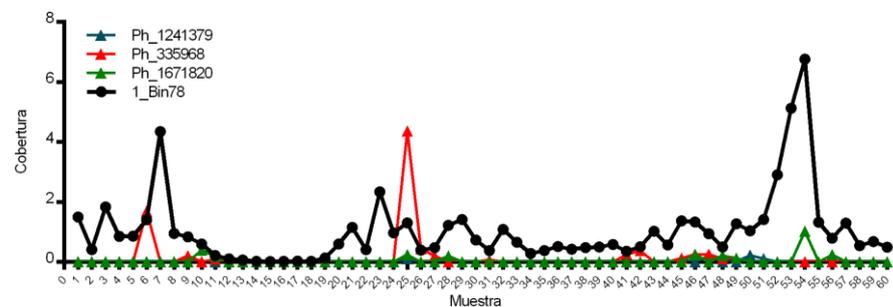


Figura 3c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin78 (*Alphaproteobacteria*, *Rhizobiales*, *Bradyrhizobiaceae*) y los fagos Ph_1241379, Ph_335968 y Ph_1671820

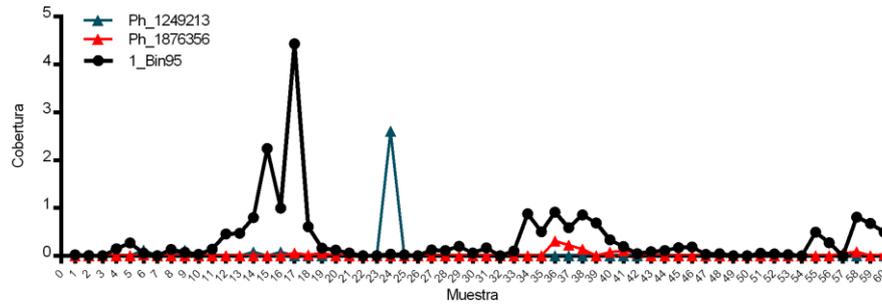


Figura 4c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin95 (*Betaproteobacteria*, *Burkholderiales*) y los fagos Ph_1249213 y Ph_1876356

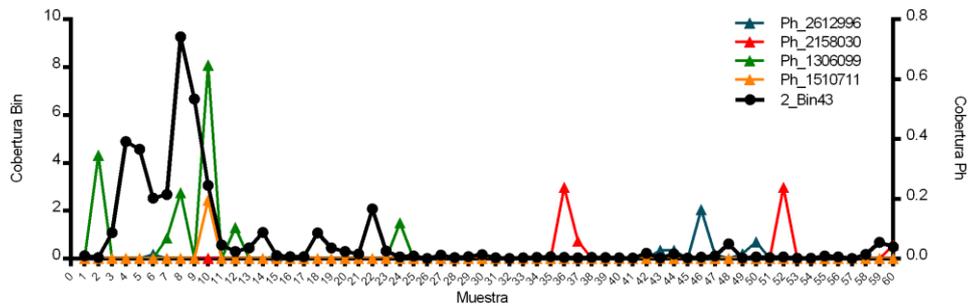


Figura 5c. Abundancia en el tiempo del par de interacción fago-hospedero: 2_Bin43 (*Betaproteobacteria*, *Burkholderiales*, *Comamonadaceae*, *Leptothrix*) y los fagos Ph_2612996, Ph_2158030, Ph_1306099 y Ph_1510711

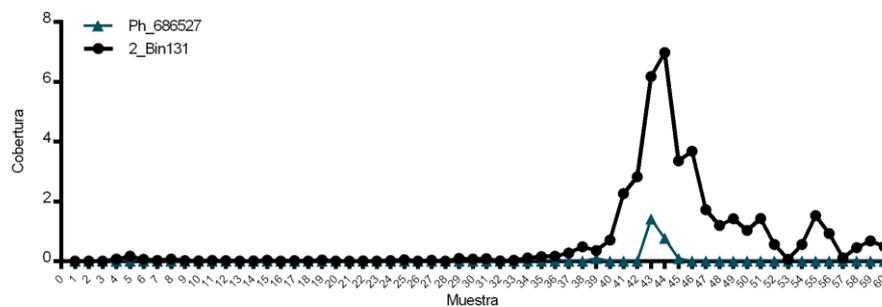


Figura 6c. Abundancia en el tiempo del par de interacción fago-hospedero: 2_Bin131 (*Betaproteobacteria*, *Burkholderiales*) y el fago Ph_686527.

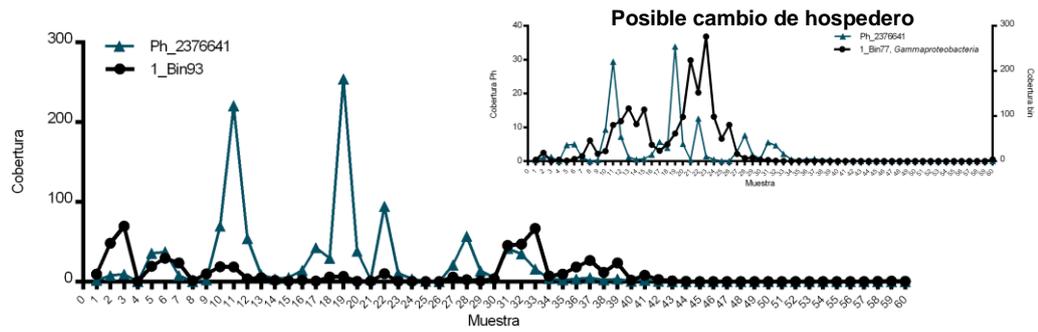


Figura 7c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin93 (*Gammaproteobacteria*) y el fago Ph_2376641

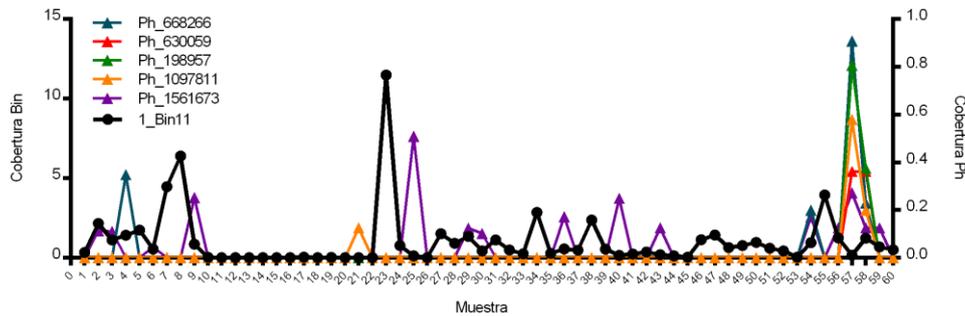


Figura 8c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin11 (*Deltaproteobacteria*, *Myxococcales*, *Polyangiaceae*, *Sorangium*) y los fagos Ph_668266, Ph_630059; Ph_198957, Ph_1097811 y Ph_1561673

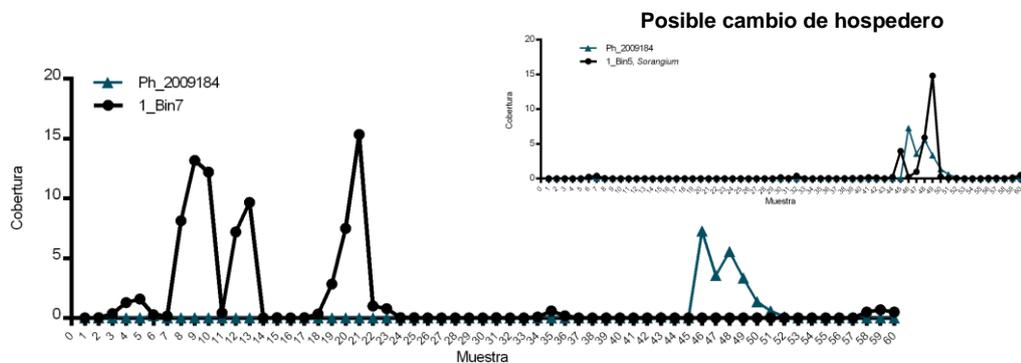


Figura 9c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin7 (*Deltaproteobacteria*, *Myxococcales*, *Polyangiaceae*, *Sorangium*) y el fago Ph_2009184

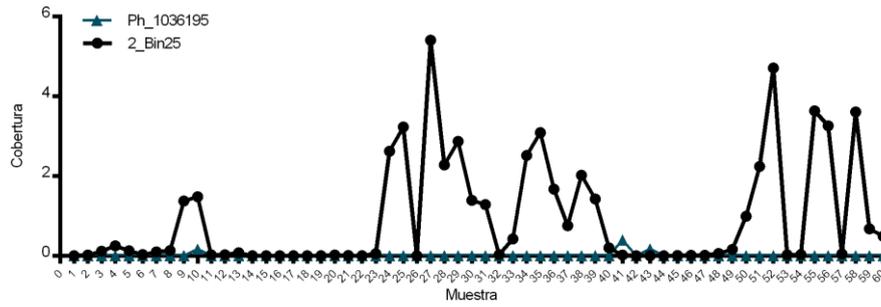


Figura 10c. Abundancia en el tiempo del par de interacción fago-hospedero: 2_Bin25 (*Bacteroidetes*, *Sphingobacteriia*, *Sphingobateriales*, *Saprospiraceae*, *Haliscomenobacter*) y el fago Ph_1036195

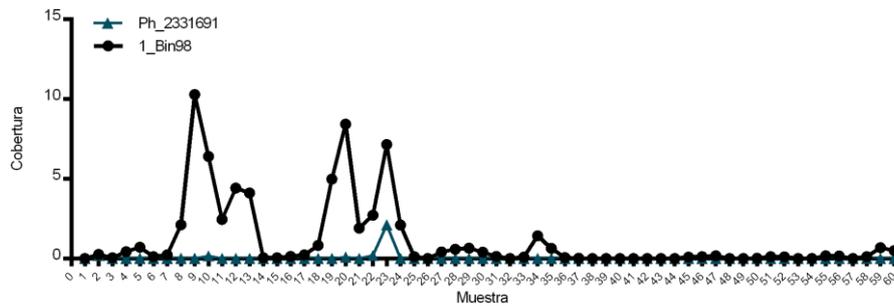


Figura 11c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin96 (*Bacteroidetes*, *Sphingobacteriia*, *Sphingobateriales*, *Saprospiraceae*, *Haliscomenobacter*) y el fago Ph_2331691

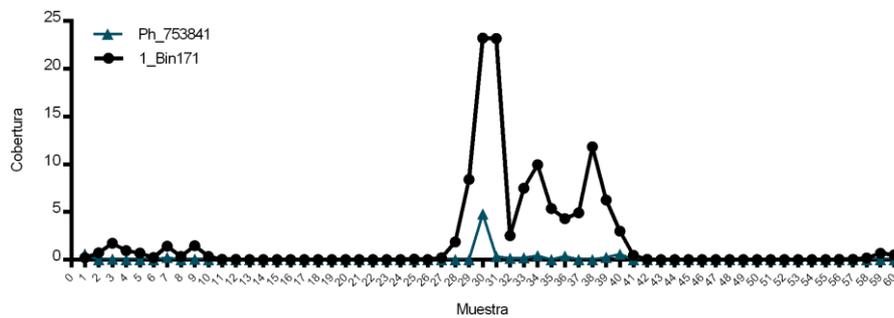


Figura 12c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin171 (*Bacteroidetes*, *Sphingobacteriia*, *Sphingobateriales*, *Saprospiraceae*, *Haliscomenobacter*) y el fago Ph_753841

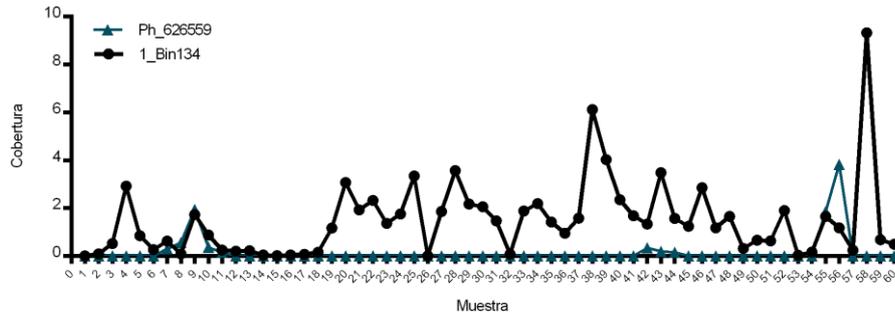


Figura 13c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin134 (*Bacteroidetes*, *Sphingobacteriia*, *Sphingobateriales*, *Saprospiraceae*, *Halicomenobacter*) y el fago Ph_626559

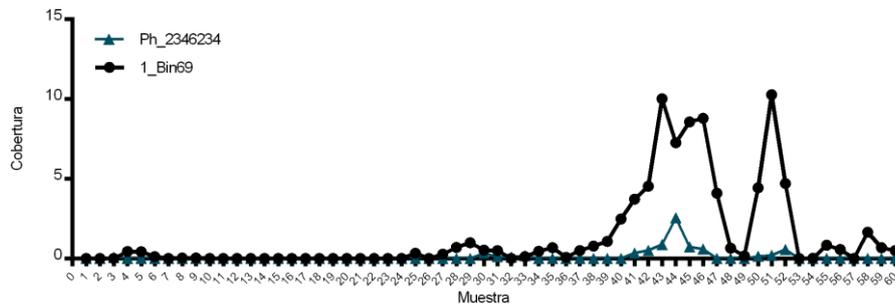


Figura 14c. Abundancia en el tiempo del par de interacción fago-hospedero: 2_Bin69 (*Bacteroidetes*, *Sphingobacteriia*, *Sphingobateriales*) y el fago Ph_2346234

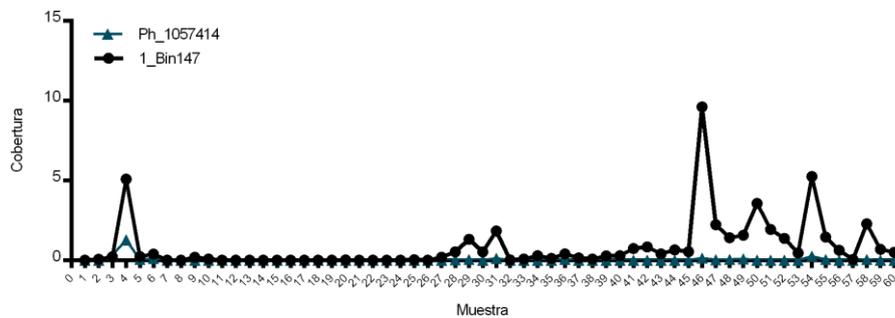


Figura 15c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin147 (*Bacteroidetes*) y el fago Ph_1057414 (*Caudovirales*, *Podoviridae*)

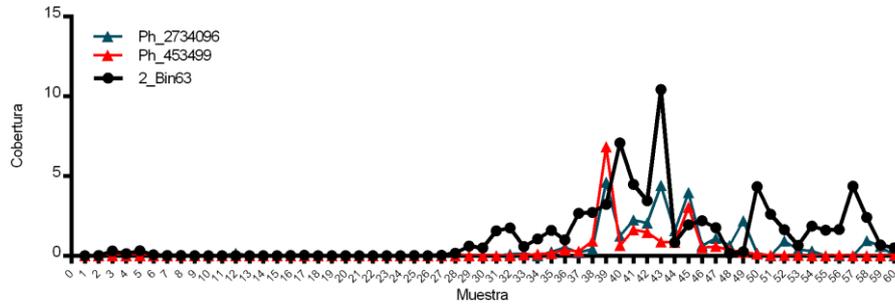


Figura 16c. Abundancia en el tiempo del par de interacción fago-hospedero: 2_Bin63 (*Bacteroidetes*, *Cytophagia*, *Cytophagales*, *Cytophageaceae*) y los fagos Ph_2734096 y Ph_453499

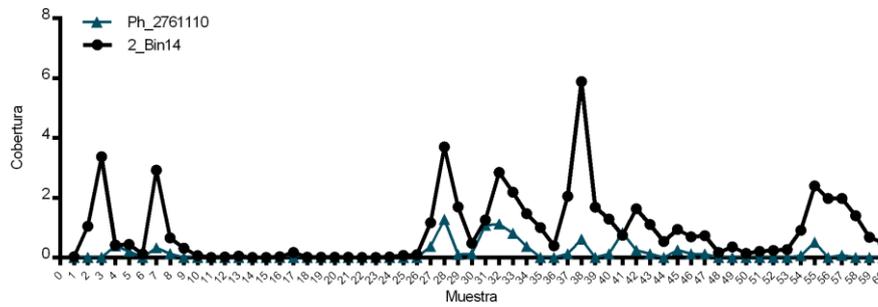


Figura 17c. Abundancia en el tiempo del par de interacción fago-hospedero: 2_Bin14 (*Bacteroidetes*, *Cytophagia*, *Cytophagales*, *Cytophageaceae*, *Runella*) y el fago Ph_1761110

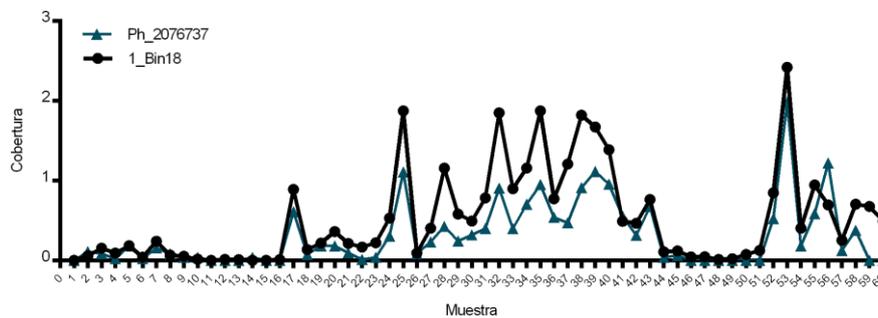


Figura 18c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin18 (*Bacteroidetes*, *Cytophagia*, *Cytophagales*, *Cytophageaceae*, *Runella*) y el fago Ph_2076737

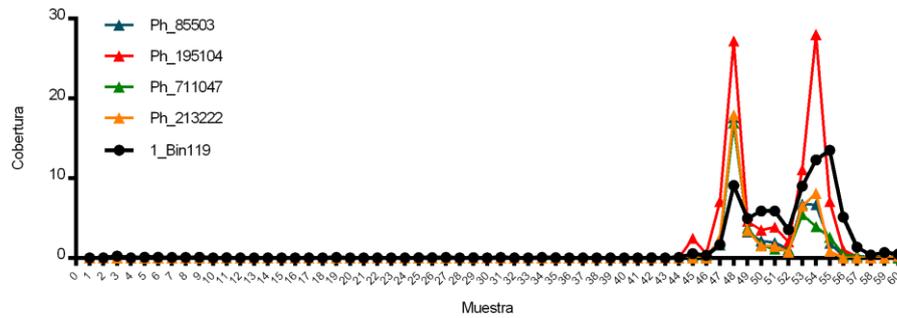


Figura 19c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin119 (*Actinobacteria*, *Actinobacteria*, *Acidimicrobiales*, *Acidimicrobiaceae*, *Acidimicrobium*) y el fago Ph_85503, Ph_195104, Ph_711047 y Ph_213222. (Los *contigs* de los fagos parecen pertenecer a un mismo fago)

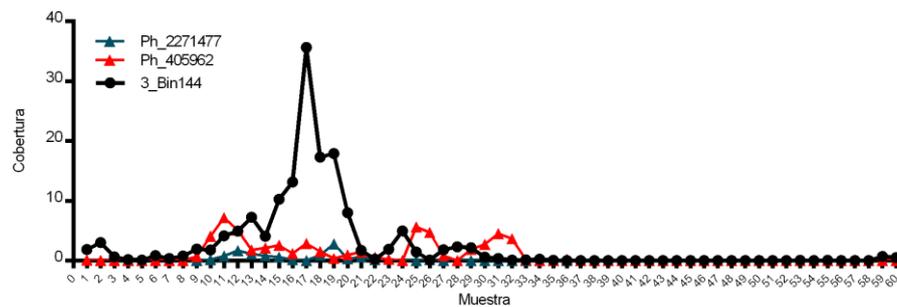


Figura 20c. Abundancia en el tiempo del par de interacción fago-hospedero: 3_Bin144 (*Actinobacteria*, *Actinobacteria*, *Actinomycetales*, *Intrasporangiaceae*, *Intrasporangium*) y los fagos Ph_2271477 y Ph_405962

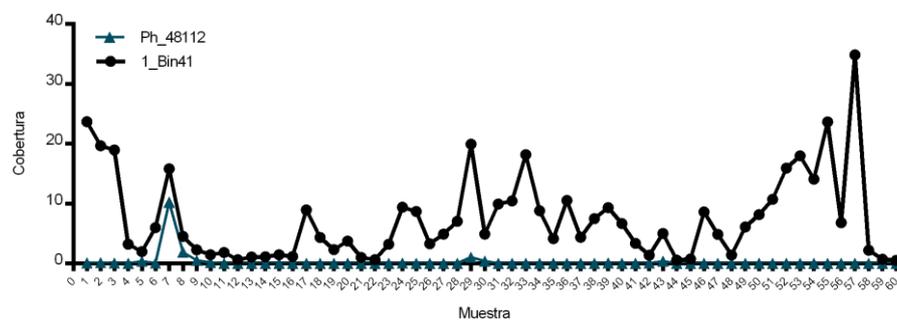


Figura 21c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin41 (*Actinobacteria*, *Actinobacteria*, *Actinomycetales*, *Gordoniaceae*, *Gordonia*) y el fago Ph_48112

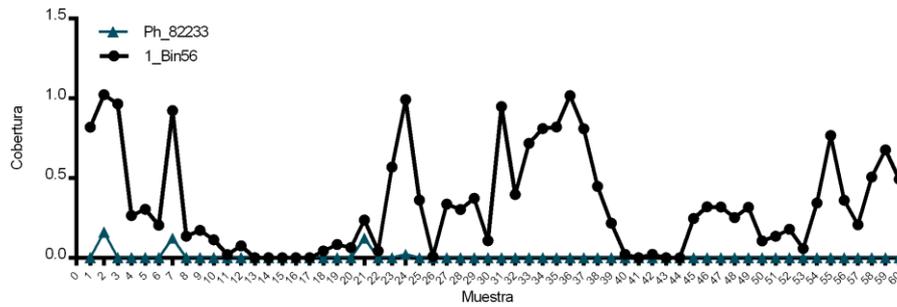


Figura 22c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin56 (*Planctomycetes*, *Planctomycetia*, *Planctomycetales*, *Planctomyces*) y el fago Ph_82233

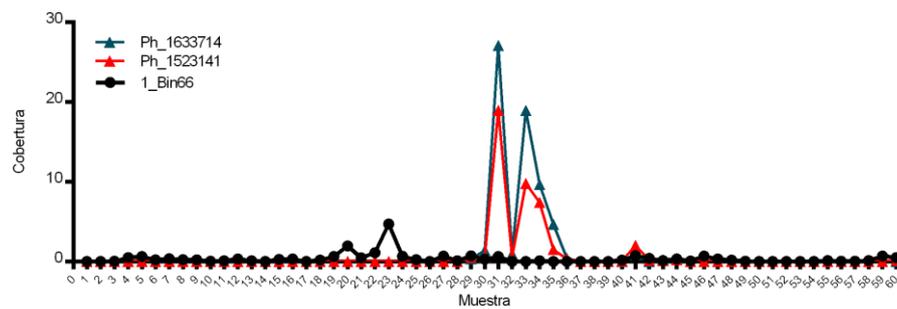


Figura 23c. Abundancia en el tiempo del par de interacción fago-hospedero: 1_Bin66 (*Verrucomicrobia*, *Verrucomicrobiae*, *Verrucomicrobiales*, *Verrucomicrobiaceae*, *Akkermansia*) y el fago Ph_1633714 y Ph_1523141 (Los *contigs* de los fagos parecen pertenecer a un mismo fago)

ANEXO D. Figuras de abundancia en el tiempo, de pares de interacción fago-hospedero detectados a partir de la búsqueda de similitud de secuencias nucleotídicas.

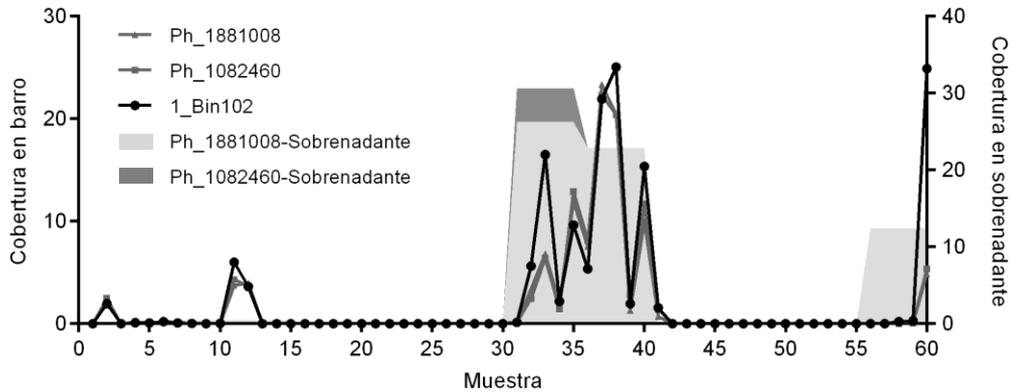


Figura 1d. Abundancia en el tiempo de 1_Bin102 (*Bacteroidetes*) y Ph_1082160 (*Caudovirales*, *Siphoviridae*) y Ph_1881008 (*Caudovirales*)

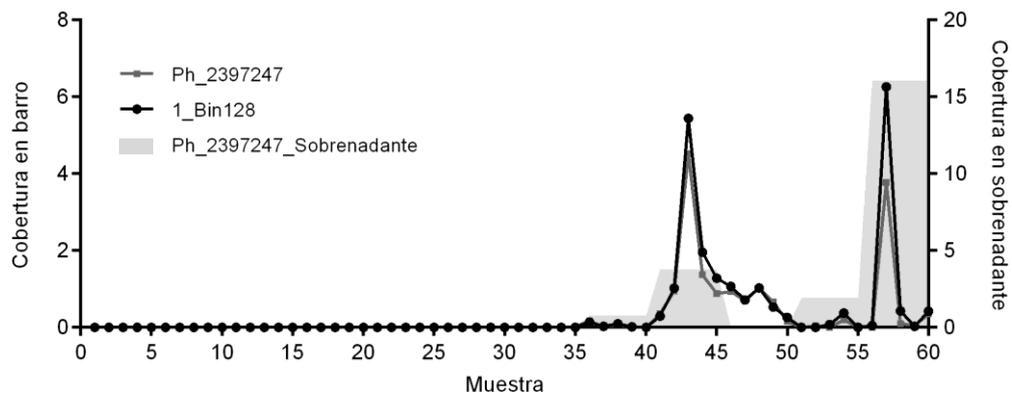


Figura 2d. Abundancia en el tiempo de 1_Bin128 (*Bacteroidetes*) y Ph_2397247 (*Caudovirales*)

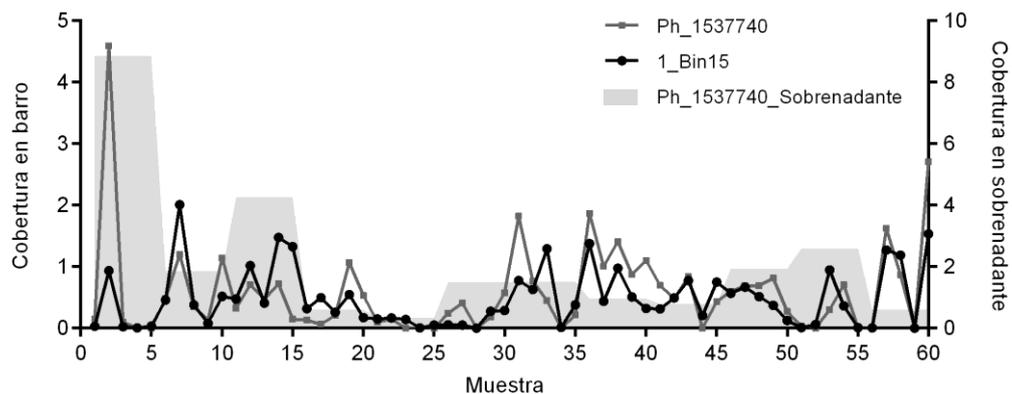


Figura 3d. Abundancia en el tiempo de 1_Bin15 (*Betaproteobacteria*) y Ph_1537740 (*Caudovirales*)

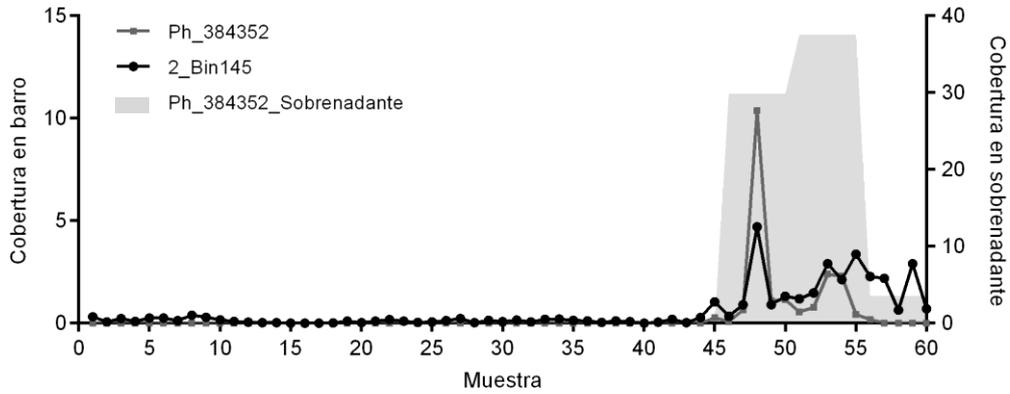


Figura 4d. Abundancia en el tiempo de 2_Bin145 (*Actinobacteria*) y Ph_384352 (*Caudovirales*, *Siphoviridae*)

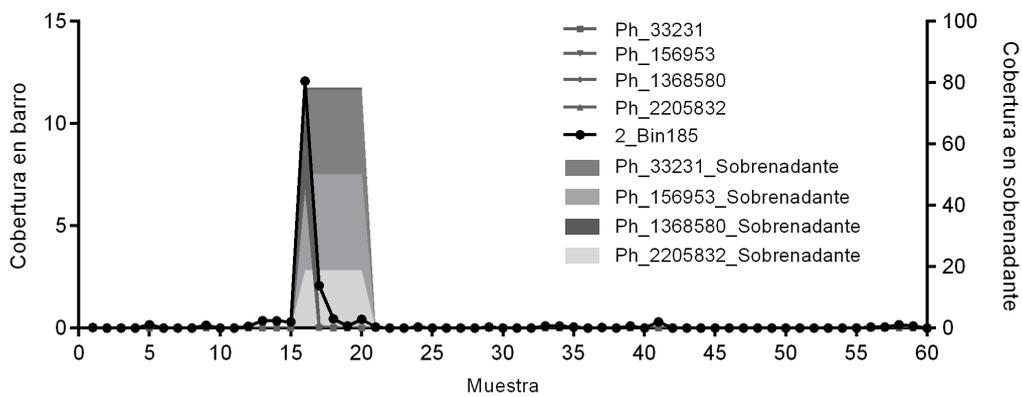


Figura 5d. Abundancia en el tiempo de 2_Bin185 (*Bacteroidetes*) y Ph_33231 (*Caudovirales*, *Myoviridae*), Ph_156953 (*Caudovirales*, *Siphoviridae*), Ph_1368580 (*Caudovirales*, *Myoviridae*), Ph_33231 (*Caudovirales*, *Siphoviridae*)

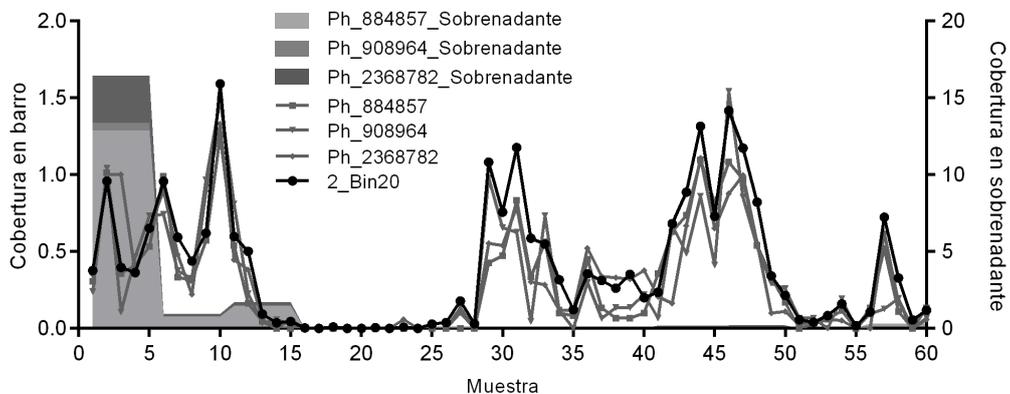


Figura 6d. Abundancia en el tiempo de 2_Bin20 (*Acidobacteria*) y Ph_884857 (*Caudovirales*, *Myoviridae*), Ph_908964 (*Caudovirales*) y Ph_2368782 (*Caudovirales*)

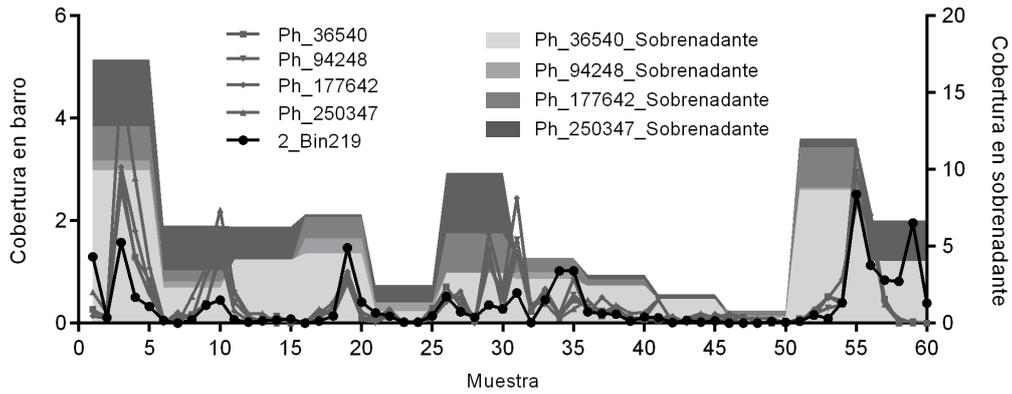


Figura 7d. Abundancia en el tiempo de 2_Bin219 (*Actinobacteria*), Ph_36540 (*Caudovirales*, *Myoviridae*), Ph_94248 (*Caudovirales*), Ph_177642 (*Caudovirales*, *Siphoviridae*), Ph_250347 (*Caudovirales*, *Myoviridae*)

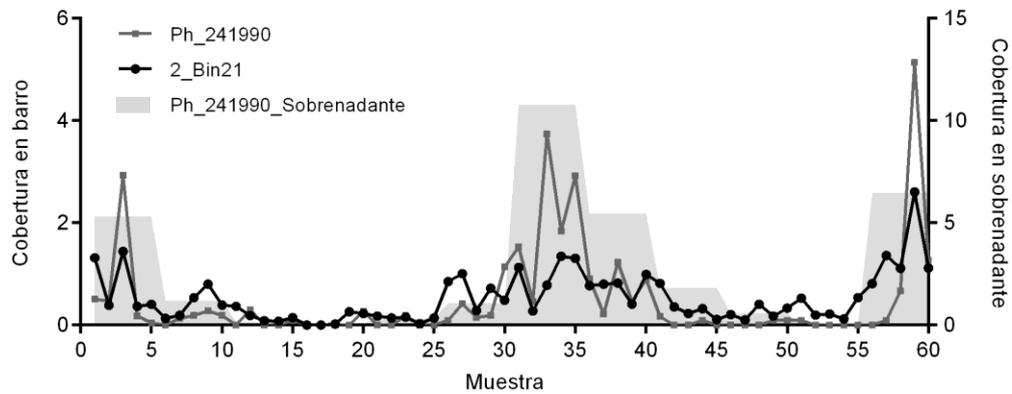


Figura 8d. Abundancia en el tiempo de 2_Bin21 (*Chloroflexi*) y Ph_241990 (*Caudovirales*)

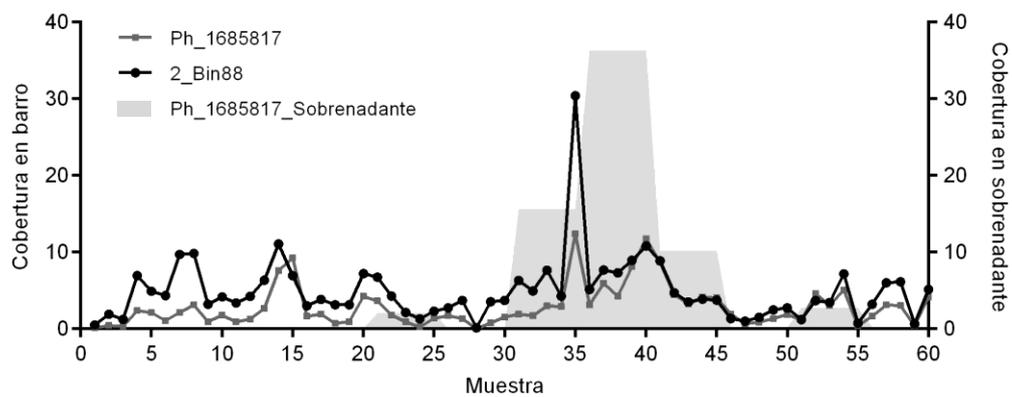


Figura 9d. Abundancia en el tiempo de 2_Bin88 (*Alphaproteobacteria*) y Ph_1685719 (*Caudovirales*)

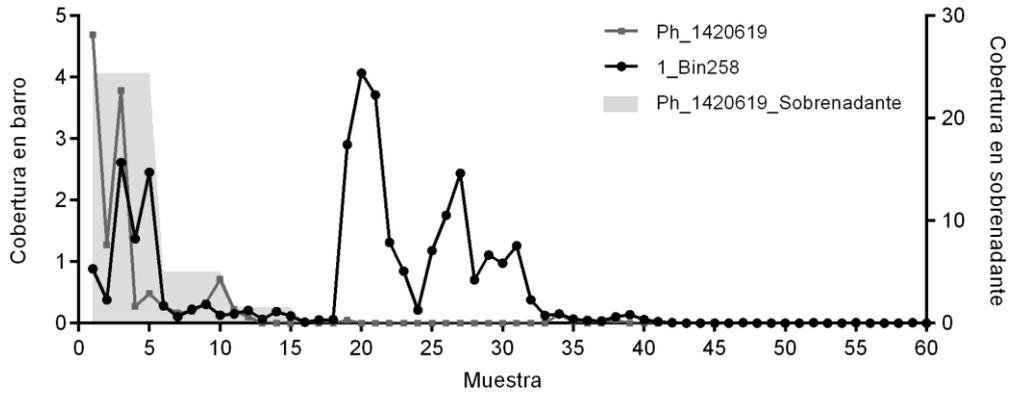


Figura 10d. Abundancia en el tiempo de 1_Bin258 (*Actinobacteria*) y Ph_1420619 (*Caudovirales*, *Siphoviridae*)

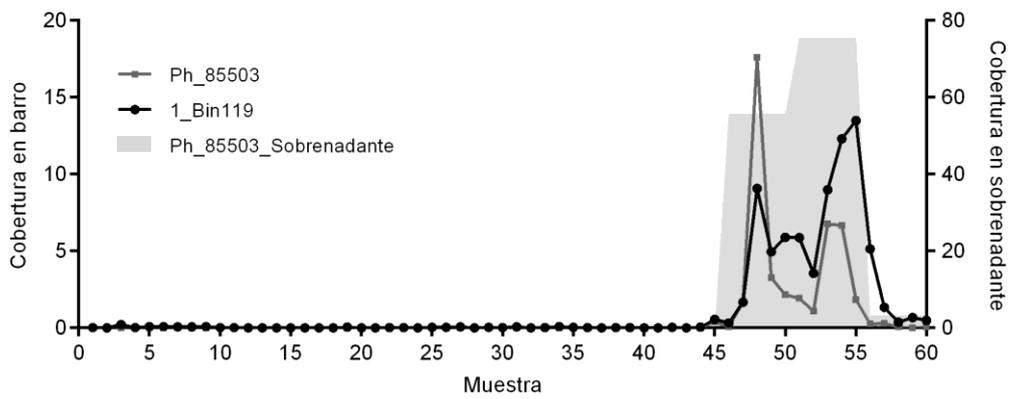


Figura 11d. Abundancia en el tiempo de 1_Bin119 (*Actinobacteria*) y Ph_85503 (*Caudovirales*)

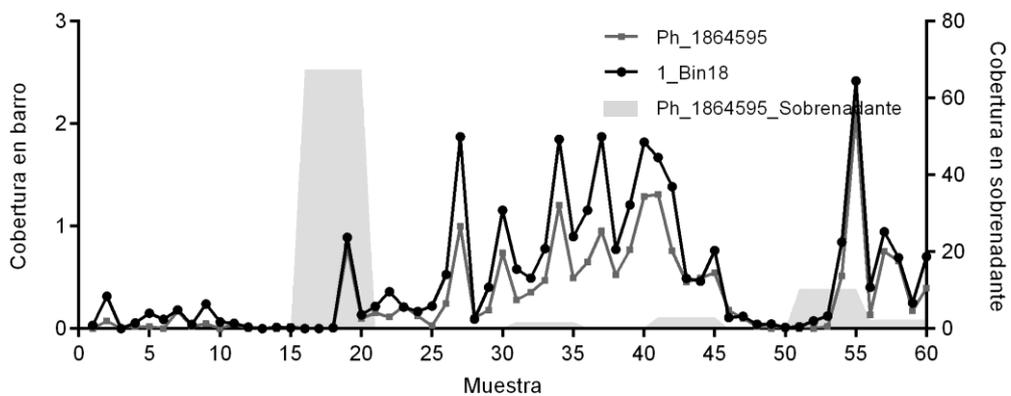


Figura 12d. Abundancia en el tiempo de 1_Bin18 (*Bacteroidetes*) y Ph_1864595 (*Caudovirales*)

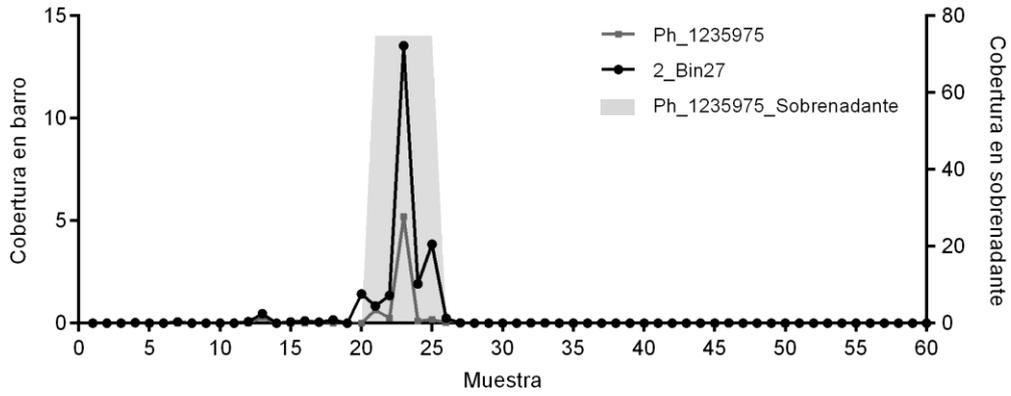


Figura 13d. Abundancia en el tiempo de 2_Bin27 (*Gammaproteobacteria*) y Ph_1235975 (*Caudovirales*)

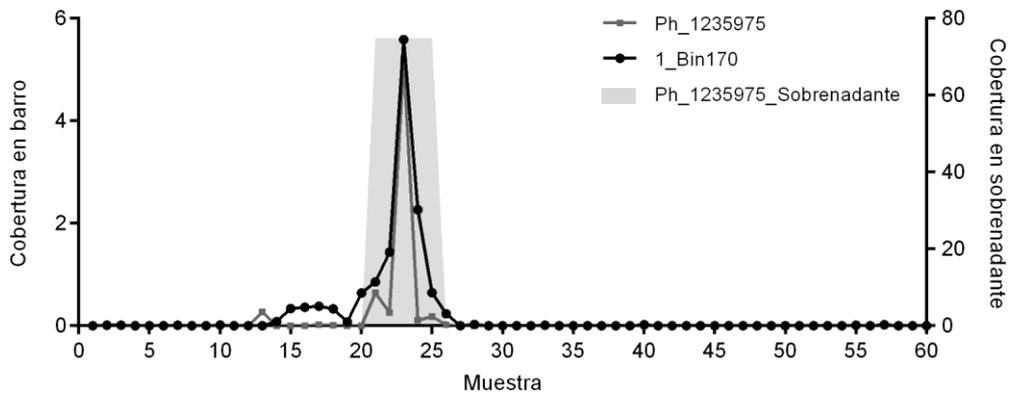


Figura 14d. Abundancia en el tiempo de 1_Bin170 (*Alphaproteobacteria*) y Ph_1235675 (*Caudovirales*)

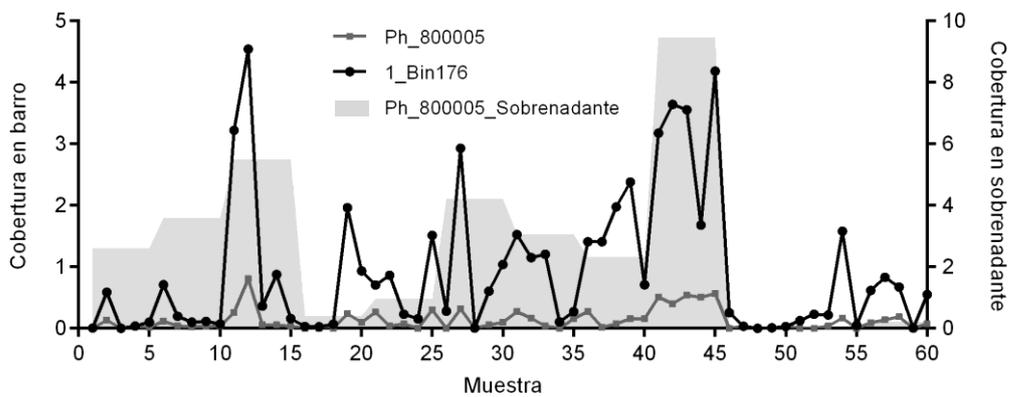


Figura 15d. Abundancia en el tiempo de 1_Bin176 (*Bacteroides*) y Ph_800005 (*Caudovirales*, *Myoviridae*)

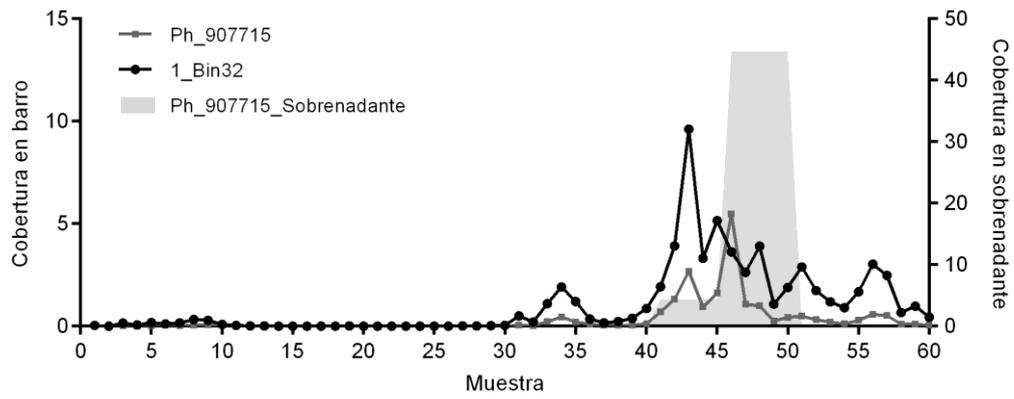


Figura 16d. Abundancia en el tiempo de 1_Bin32 (*Alphaproteobacteria*) y Ph_907715 (*Caudovirales*)

BIBLIOGRAFÍA

Abma WR, Schultz CE, Mulder JW, van der Star WRL, Strous M, Tokutomi T, *et al.* (2007). Full-scale granular sludge Anammox process. *Water Science and Technology* **55**: 27–33.

Adriaenssens EM, Cowan DA. (2014). Using signature genes as tools to assess environmental viral ecology and diversity. *Applied and Environmental Microbiology* **80**: 4470–4480.

Akhter S, Aziz RK, Edwards RA. (2012). PhiSpy: A novel algorithm for finding prophages in bacterial genomes that combines similarity-and composition-based strategies. *Nucleic Acids Research* **40**: e126.

Al-Shahwani SM, Horan NJ. (1991). The use of protozoa to indicate changes in the performance of activated sludge plants. *Water Research* **25**: 633–638.

Alhamlan FS, Ederer MM, Brown CJ, Coats ER, Crawford RL. (2013). Metagenomics-based analysis of viral communities in dairy lagoon wastewater. *Journal of Microbiological Methods* **92**: 183–188.

Allers E, Moraru C, Duhaime MB, Beneze E, Solonenko N, Barrero-Canosa J, *et al.* (2013). Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. *Environmental Microbiology* **15**: 2306–2318.

Amann RI, Ludwig W, Schleifer K-H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews* **59**: 143–169.

Anderson RE, Brazelton WJ, Baross JA. (2011). Using CRISPRs as ametagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiology Ecology* **77**: 120–133.

Andersson AF, Banfield JF. (2008). Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**: 1047–1050.

Araújo dos Santos L, Ferreira V, Pereira MO, Nicolau A. (2014). Relationship between protozoan and metazoan communities and operation and performance parameters in a textile sewage activated sludge system. *European Journal of Protistology* **50**: 319–328.

Arden E, Lockett WT. (1914). Experiments on the oxidation of sewage without the aid of filters. *Journal of Society of Chemical Industry* **33**: 523–539.

Arévalo J, Moreno B, Pérez J, Gómez MA. (2009). Applicability of the Sludge Biotic Index (SBI) for MBR activated sludge control. *Journal of Hazardous Materials* **167**: 784–789.

Array C, Dna P, Polymerase B. (2016). Viruses infecting a freshwater filamentous cyanobacterium (*Nostoc* sp.) encode a functional CRISPR array and a proteobacterial DNA polymerase B. *mBio* **7**: 1–11.

Aw TG, Howe A, Rose JB. (2014). Metagenomic approaches for direct and cell culture evaluation of the virological quality of wastewater. *Journal of Virological Methods* **210**: 15–21.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, *et al.* (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**: 75.

Baker GC, Smith JJ, Cowan DA. (2003). Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* **55**: 541–555.

Barr JJ, Slater FR, Fukushima T, Bond PL. (2010). Evidence for bacteriophage activity causing community and performance changes in a phosphorus-removal activated sludge. *FEMS Microbiology Ecology* **74**: 631–642.

Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, *et al.* (2015). metaxa2: Improved identification and taxonomic classification of small and large subunit rRNA

in metagenomic data. *Molecular Ecology Resources* **15**: 1403–1414.

Berg Miller ME, Yeoman CJ, Chia N, Tringe SG, Angly FE, Edwards RA, *et al.* (2012). Phage-bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome. *Environmental Microbiology* **14**: 207–227.

Berk SG, Gunderson JH. (1993). Wastewater organisms: A Color atlas. Lewis Publishers.

Besemer J. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Research* **29**: 2607–2618.

Blackall LL, Crocetti GR, Saunders AM, Bond PL. (2002). A review and update of the microbiology of enhanced biological phosphorus removal in wastewater treatment plants. *Antonie van Leeuwenhoek* **81**: 681–691.

Bolger AM, Lohse M, Usadel B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bouvy M, Bettarel Y, Bouvier C, Domaizon I, Jacquet S, Le Floc'h E, *et al.* (2011). Trophic interactions between viruses, bacteria and nanoflagellates under various nutrient conditions and simulated climate change. *Environmental Microbiology* **13**: 1842–1857.

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, *et al.* (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* **35**: 725–731.

Boyajian E, Glueckstein LJ. (1987). Manuals of practice for water pollution control. Activated sludge manual. Water Environmental Federation.

Breitbart M. (2012). Marine viruses: truth or dare. *Annual Review of Marine Science* **4**: 425–448.

Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, *et al.* (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports* **5**: 8365.

Brown CT, Olm MR, Thomas BC, Banfield JF. (2016). Measurement of bacterial replication rates in microbial communities. *Nature Biotechnology* **34**: 1256–1263.

Brown MR, Camézuli S, Davenport RJ, Petelenz-Kurdziel E, Øvreås L, Curtis TP. (2015). Flow cytometric quantification of viruses in activated sludge. *Water Research* **68**: 414–422.

Bulmer M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129**: 897–907.

Button DK. (1991). Biochemical basis for whole-cell uptake kinetics: Specific affinity, oligotrophic capacity, and the meaning of the Michaelis constant. *Applied and Environmental Microbiology* **57**: 2033–2038.

Cai L, Ye L, Tong AHY, Lok S, Zhang T. (2013). Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. *PLoS ONE* **8**: e53649.

Cai L, Zhang R, He Y, Feng X, Jiao N. (2016). Metagenomic analysis of Virioplankton of the subtropical Jiulong river estuary, China. *Viruses* **8**: E35.

Cassman N, Prieto-Davó A, Walsh K, Silva GGZ, Angly F, Akhter S, *et al.* (2012). Oxygen minimum zones harbour novel viral communities with low diversity. *Environmental Microbiology* **14**: 3043–3065.

Chow C-ET, Kim DY, Sachdeva R, Caron DA, Fuhrman JA. (2014). Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *The ISME Journal* **8**: 816–829.

- Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, *et al.* (2014). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Research* **42**: 633–642.
- Crocetti GR, Hugenholtz P, Bond PL, Schuler A, Keller J, Jenkins D, *et al.* (2000). Identification of polyphosphate-accumulating organisms and design of 16S rRNA-directed probes for their detection and quantitation. *Applied and Environmental Microbiology* **66**: 1175–1182.
- Curds CR, Cockburn A. (1970). Protozoa in biological sewage-treatment process-II. Protozoa as indicators in the activated-sludge process. *Water Research* **4**: 237–249.
- Cydzik-Kwiatkowska A, Zielińska M. (2016). Bacterial communities in full-scale wastewater treatment systems. *World Journal of Microbiology & Biotechnology* **32**: 66.
- Daims H. (2009). Use of fluorescence in situ hybridization and the daime image analysis program for the cultivation-independent quantification of microorganisms in environmental and medical samples. *Cold Spring Harbor Protocols* **4**: pdb.prot5253.
- Daims H, Lücker S, Wagner M. (2006a). Daime, a novel image analysis program for microbial ecology and biofilm research. *Environmental Microbiology* **8**: 200–213.
- Daims H, Purkhold U, Bjerrum L, Arnold E, Wilderer PA, Wagner M. (2001). Nitrification in sequencing biofilm batch reactors: Lessons from molecular approaches. *Water Science and Technology* **43**: 9–18.
- Daims H, Taylor MW, Wagner M. (2006b). Wastewater treatment: a model system for microbial ecology. *Trends in Biotechnology* **24**: 483–489.
- Decewicz P, Radlinska M, Dziewit L. (2017). Characterization of Sinorhizobium sp. LM21 prophages and virus-encoded DNA methyltransferases in the light of comparative genomic analyses of the sinorhizobial virome. *Viruses* **9**: 1–19.
- Drzewicki A, Kulikowska D. (2011). Limitation of Sludge Biotic Index application for control of a wastewater treatment plant working with shock organic and ammonium loadings. *European Journal of Protistology* **47**: 287–294.
- Edgar RC. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**: 996–998.
- Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. (2016). Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiology Reviews* **40**: 258–272.
- Eikelboom DH. (1975). Filamentous organisms observed in activated sludge. *Water Research* **9**: 365–388.
- Eikelboom DH, Geurkink B. (2002). Filamentous micro-organisms observed in industrial activated sludge plants. *Water Science and Technology* **46**: 535–542.
- Erijman L, Figuerola ELM, Guerrero LD, Ayarza JM. (2011). Impact of the recent advances in the analysis of microbial communities on the control of the wastewater treatment process. *Revista Argentina de microbiología* **43**: 127–135.
- Ewert DL, Paynter MJB. (1980). Enumeration of bacteriophages and host bacteria in sewage and the activated-sludge treatment process. *Applied and Environmental Microbiology* **39**: 576–583.
- Fierer N, Bradford MA, Jackson RB. (2007). Toward an ecological classification of soil bacteria. *Ecology* **88**: 1354–1364.
- Filee J, Tetart F, Suttle C a, Krisch HM, Filée J, Tétart F, *et al.* (2005). Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 12471–12476.
- Foissner W. (2016). Protists as bioindicators in activated sludge: Identification, ecology and

future needs. *European Journal of Protistology* **55**: 75–94.

Fouts DE. (2006). Phage_Finder: Automated identification and classification of prophage regions in complete bacterial genome sequences. *Nucleic Acids Research* **34**: 5839–5851.

Fredriksson NJ, Hermansson M, Wilén B-M. (2013). The choice of PCR primers has great impact on assessments of bacterial community diversity and dynamics in a wastewater treatment plant. *PLoS ONE* **8**: e76431.

Fujii T, Nakayama N, Nishida M, Sekiya H, Kato N, Asakawa S, *et al.* (2008). Novel capsid genes (g23) of T4-type bacteriophages in a Japanese paddy field. *Soil Biology and Biochemistry* **40**: 1049–1058.

Gerardi MH, Horsfall FL. (1995). Wastewater biology: The microlife. Water Environment Federation.

Giloh H, Sedat J. (1982). Fluorescence microscopy: reduced photobleaching of rhodamine and fluorescein protein conjugates by n-propyl gallate. *Science* **217**: 1252–1255.

Godde JS, Bickerton A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution* **62**: 718–729.

Gott R, George R, Garcia B, Gott N. (2007). New Mexico wastewater systems operator certification study manual. Chapter 10: activated sludge. The Water Utilities Technical Assistance Program For The Utility Operators Certification Program.

Haandel AC Van, Lubbe JGM Van Der. (2012). Handbook of biological wastewater treatment. Second Edi. IWA Publishing.

Haggerty JM, Dinsdale EA. (2017). Distinct biogeographical patterns of marine bacterial taxonomy and functional genes. *Global Ecology and Biogeography* **26**: 177–190.

Han P, Deem MW. (2017). Non-classical phase diagram for virus bacterial coevolution mediated by clustered regularly interspaced short palindromic repeats. *Journal of The Royal Society Interface* **14**: 20160905.

Hannigan GD, Meisel JS, Tyldsley AS, Zheng Q, Hodkinson BP, Sanmiguel AJ, *et al.* (2015). The human skin double-stranded DNA virome: Topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *mBio* **6**: e01578-15.

Hantula J, Kurki A, Vuoriranta P, Bamford DH. (1991). Ecology of bacteriophages infecting activated sludge bacteria. *Applied and Environmental Microbiology* **57**: 2147–2151.

Hargreaves KR, Flores CO, Lawley TD, Clokie MRJ. (2014). Abundant and diverse clustered regularly interspaced short palindromic repeat spacers in *Clostridium difficile* strains and phages target multiple phage types within this pathogen. *mBio* **5**: e01045-13.

Henze M, van Loosdrecht MCM, Ekama GA, Brdjanovic D. (2008). Biological wastewater treatment - Principles, modelling and design. IWA Publishing.

Hesselmann RPX, Werlen C, Hahn D, van der Meer JR, Zehnder AJB. (1999). Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. *Systematic and Applied Microbiology* **22**: 454–465.

Hong S, Bunge J, Leslin C, Jeon S, Epstein SS. (2009). Polymerase chain reaction primers miss half of rRNA microbial diversity. *The ISME Journal* **3**: 1365–1373.

Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. (2017). Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *The ISME Journal* **11**: 1511–1520.

Ibarbalz FM. (2016). Metagenómica de lodos activados. Factores determinantes del ensamblado de comunidades bacterianas en el tratamiento de efluentes. Universidad de

Buenos Aires. Facultad de Ciencias Exactas y Naturales.

Ibarbalz FM, Figuerola ELM, Erijman L. (2013). Industrial activated sludge exhibit unique bacterial community composition at high taxonomic ranks. *Water Research* **47**: 3854–3864.

Ibarbalz FM, Orellana E, Figuerola ELM, Erijman L. (2016). Shotgun metagenomic profiles have a high capacity to discriminate samples of activated sludge according to wastewater type. *Applied and Environmental Microbiology* **82**: 5186–5196.

Ibarbalz FM, Pérez MV, Figuerola ELM, Erijman L. (2014). The bias associated with amplicon sequencing does not affect the quantitative assessment of bacterial community dynamics. *PLoS ONE* **9**: e99722.

Jenkins D, Richard MG, Daigger GT. (2004). *Manual on the Causes and Control of Activated Sludge Bulking, Foaming, and Other Solids Separation Problems*. 3rd Editio. IWA Publishing.

Jo SJ, Kwon H, Jeong SY, Lee CH, Kim TG. (2016). Comparison of microbial communities of activated sludge and membrane biofilm in 10 full-scale membrane bioreactors. *Water Research* **101**: 214–225.

Ju F, Zhang T. (2015). Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. *The ISME Journal* **9**: 683–695.

Julia J, Yossef C, Marina DL, Ariel K, Edouard J, Antonis C. (2014). Multiple micro-predators controlling bacterial communities in the environment. *Current Opinion in Biotechnology* **27**: 185–190.

Kanagawa T, Kamagata Y, Aruga S, Kohno T, Horn M, Wagner M. (2000). Phylogenetic analysis of and oligonucleotide probe development for Eikelboom type 021N filamentous bacteria isolated from bulking activated sludge. *Applied and Environmental Microbiology* **66**: 5043–5052.

Kang DD, Froula J, Egan R, Wang Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**: e1165.

van Kessel MAHJ, Speth DR, Albertsen M, Nielsen PH, Op den Camp HJM, Kartal B, *et al.* (2015). Complete nitrification by a single microorganism. *Nature* **528**: 555–559.

Khan MA, Satoh H, Katayama H, Kurisu F, Mino T. (2002). Bacteriophages isolated from activated sludge processes and their polyvalency. *Water Research* **36**: 3364–3370.

Kim TS, Jeong JY, Wells GF, Park HD. (2013). General and rare bacterial taxa demonstrating different temporal dynamic patterns in an activated sludge bioreactor. *Applied Microbiology and Biotechnology* **97**: 1755–1765.

Klappenbach JA, Dunbar JM, Thomas M, Schmidt TM. (2000). rRNA operon copy number reflects ecological strategies of bacteria. *Applied and Environmental Microbiology* **66**: 1328–1333.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, *et al.* (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research* **41**: e1.

Koch AL. (1971). The adaptive responses of *Escherichia coli* to a feast and famine existence. *Advances in Microbial Physiology* **6**: 147–217.

Kong Y, Nielsen JL, Nielsen PH. (2004). Microautoradiographic study of *Rhodocyclus*-related polyphosphate-accumulating bacteria in full-scale enhanced biological phosphorus removal plants. *Applied and Environmental Microbiology* **70**: 5383–5390.

Kotay SM, Datta T, Choi J, Goel R. (2011). Biocontrol of biomass bulking caused by *Haliscomenobacter hydrossis* using a newly isolated lytic bacteriophage. *Water Research* **45**: 694–704.

- Kragelund C, Kong Y, van der Waarde J, Thelen K, Eikelboom D, Tandoi V, *et al.* (2006). Ecophysiology of different filamentous Alphaproteobacteria in industrial wastewater treatment plants. *Microbiology* **152**: 3003–3012.
- Kragelund C, Levantesi C, Borger A, Thelen K, Eikelboom D, Tandoi V, *et al.* (2007a). Identity, abundance and ecophysiology of filamentous Chloroflexi species present in activated sludge treatment plants. *FEMS Microbiology Ecology* **59**: 671–682.
- Kragelund C, Remesova Z, Nielsen JL, Thomsen TR, Eales K, Seviour R, *et al.* (2007b). Ecophysiology of mycolic acid-containing Actinobacteria (Mycolata) in activated sludge foams. *FEMS Microbiology Ecology* **61**: 174–184.
- Kristensen DM, Waller AS, Yamada T, Bork P, Mushegian AR, Koonin E V. (2013). Orthologous gene clusters and taxon signature genes for viruses of prokaryotes. *Journal of Bacteriology* **195**: 941–950.
- Kumar PS, Brooker MR, Dowd SE, Camerlengo T. (2011). Target region selection is a critical determinant of community fingerprints generated by 16S Pyrosequencing. *PLoS ONE* **6**: e20956.
- Kunin V, He S, Warnecke F, Peterson SB, Garcia Martin H, Haynes M, *et al.* (2008). A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Research* **18**: 293–297.
- Kwon S, Kim TS, Yu GH, Jung JH, Park HD. (2010). Bacterial community composition and diversity of a full-scale integrated fixed-film activated sludge system as investigated by pyrosequencing. *Journal of Microbiology and Biotechnology* **20**: 1717–1723.
- Labonté JM, Swan BK, Poulos B, Luo H, Koren S, Hallam SJ, *et al.* (2015). Single-cell genomics-based analysis of virus–host interactions in marine surface bacterioplankton. *The ISME Journal* **9**: 2386–2399.
- Langmead B, Salzberg SL. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357–9.
- Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S, *et al.* (2009). The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences* **106**: 15527–15533.
- Leal AL, Dalzochio MS, Flores TS, De Alves AS, Macedo JC, Valiati VH. (2013). Implementation of the sludge biotic index in a petrochemical WWTP in Brazil: Improving operational control with traditional methods. *Journal of Industrial Microbiology and Biotechnology* **40**: 1415–1422.
- Lee SH, Kang HJ, Park HD. (2015). Influence of influent wastewater communities on temporal variation of activated sludge communities. *Water Research* **73**: 132–144.
- Levantesi C, Rossetti S, Thelen K, Kragelund C, Krooneman J, Eikelboom D, *et al.* (2006). Phylogeny, physiology and distribution of 'Candidatus Microthrix calida', a new Microthrix species isolated from industrial activated sludge wastewater treatment plants. *Environmental Microbiology* **8**: 1552–1563.
- Li D, Liu CM, Luo R, Sadakane K, Lam TW. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.
- Lima-Mendez G, Van Helden J, Toussaint A, Leplae R. (2008). Prophinder: A computational tool for prophage prediction in prokaryotic genomes. *Bioinformatics* **24**: 863–865.
- Madoni P. (1994). A sludge biotic index (SBI) for the evaluation of the biological performance of activated sludge plants based on the microfauna analysis. *Water Research* **28**: 67–75.

- Madoni P. (2011). Protozoa in wastewater treatment processes: A minireview. *Italian Journal of Zoology* **78**: 3–11.
- Maixner F, Noguera DR, Anneser B, Stoecker K, Wegl G, Wagner M, *et al.* (2006). Nitrite concentration influences the population structure of Nitrospira-like bacteria. *Environmental Microbiology* **8**: 1487–1495.
- Meyer F, Paarmann D, D'Souza M, Olson R, Glass E, Kubal M, *et al.* (2008). The metagenomics RAST server - a public resource for the automatic phylo- genetic and functional analysis of metagenomes. *BMC bioinformatics* **9**: 386.
- Miller CS, Baker BJ, Thomas BC, Singer SW, Banfield JF. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biology* **12**: R44.
- Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD. (2013). Rapid evolution of the human gut virome. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 12450–5.
- Mizuno CM, Rodriguez-Valera F, Kimes NE, Ghai R. (2013). Expanding the marine virosphere using metagenomics. *PLoS Genetics* **9**: e1003987.
- Muyzer G, De Waal EC, Uitterlinden AG. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59**: 695–700.
- Nicolau A, Dias N, Mota M, Lima N. (2001). Trends in the use of protozoa in the assessment of wastewater treatment. *Research in Microbiology* **152**: 621–630.
- Nielsen PH, Aquino de Muro M, Nielsen JL. (2000). Studies on the in situ physiology of Thiothrix spp. present in activated sludge. *Environmental Microbiology* **2**: 389–398.
- Ofir G, Melamed S, Sberro H, Mukamel Z, Silverman S, Yaakov G, *et al.* (2017). DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nature Microbiology* doi: 10.1038/s41564-017-0051-0.
- Ory P, Hartmann HJ, Jude F, Dupuy C, Del Amo Y, Catala P, *et al.* (2010). Pelagic food web patterns: Do they modulate virus and nanoflagellate effects on picoplankton during the phytoplankton spring bloom? *Environmental Microbiology* **12**: 2755–2772.
- Otawa K, Lee SH, Yamazoe A, Onuki M, Satoh H, Mino T. (2007). Abundance, diversity, and dynamics of viruses on microorganisms in activated sludge processes. *Microbial Ecology* **53**: 143–152.
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, *et al.* (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research* **42**: D206-14.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research* **25**: 1043–1055.
- Parks DH, Tyson GW, Hugenholtz P, Beiko RG. (2014). STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**: 3123–3124.
- Parsley LC, Consuegra EJ, Thomas SJ, Bhavsar J, Land AM, Bhuiyan NN, *et al.* (2010). Census of the viral metagenome within an activated sludge microbial assemblage. *Applied and Environmental Microbiology* **76**: 2673–2677.
- Pawluk A, Bondy-Denomy J, Cheung VHW, Maxwell KL, Davidson AR. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* **5**: e00896.

- Pedrazzani R, Menoni L, Nembrini S, Manili L, Bertanza G. (2016). Suitability of Sludge Biotic Index (SBI), Sludge Index (SI) and filamentous bacteria analysis for assessing activated sludge process performance: the case of piggery slaughterhouse wastewater. *Journal of Industrial Microbiology and Biotechnology* **43**: 953–964.
- Pérez-Uz B, Arregui L, Calvo P, Salvadó H, Ferná Ndez N, Rodríguez E, *et al.* (2010). Assessment of plausible bioindicators for plant performance in advanced wastewater treatment systems. *Water Research* **44**: 5059–5069.
- Pericard P, Dufresne Y, Blanquart S, Touzet H. (2017). MATAM: reconstruction of phylogenetic marker genes from short sequencing reads in metagenomes. *bioRxiv* doi: 10.1093/bioinformatics/btx644.
- Petrovski S, Seviour RJ, Tillett D. (2011). Characterization of the genome of the polyvalent lytic bacteriophage GTE2, which has potential for biocontrol of *Gordonia*, *Rhodococcus*, and *Nocardia* stabilized foams in activated sludge plants. *Applied and Environmental Microbiology* **77**: 3923–3929.
- Pholchan MK, Baptista J de C, Davenport RJ, Curtis TP. (2010). Systematic study of the effect of operating variables on reactor performance and microbial diversity in laboratory-scale activated sludge reactors. *Water Research* **44**: 1341–1352.
- Pianka ER. (1970). On r- and K-Selection. *The American Naturalist* **104**: 592–597.
- Polz MF, Cavanaugh CM. (1998). Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology* **64**: 3724–3730.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, *et al.* (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research* **41**: 590–596.
- Reysenbach A-L, Giver LJ, Wickham GS, Pace NR. (1992). Differential amplification of ribosomal RNA genes by polymerase chain reaction. *Applied and Environmental Microbiology* **58**: 3417–3418.
- Richard M. (1989). Activated Sludge Microbiology. The Water Pollution Control Federation.
- Rodríguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M, *et al.* (2010). Viral and microbial community dynamics in four aquatic environments. *The ISME Journal* **4**: 739–751.
- Roller BRK, Stoddard SF, Schmidt TM. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature Microbiology* **1**: 16160.
- Rosenberg E, Bittan-Banin G, Sharon G, Shon A, Hershko G, Levy I, *et al.* (2010). The phage-driven microbial loop in petroleum bioremediation. *Microbial Biotechnology* **3**: 467–472.
- Roux S, Enault F, Hurwitz BL, Sullivan MB. (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ* **3**: e985.
- Ryan PD, Hammer Ø, Harper DA, Paul Ryan DD. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **178kb T Harper Geological Museum** **4**: 5–7.
- Saikaly PE, Stroot PG, Oerther DB. (2005). Use of 16S rRNA gene terminal restriction fragment analysis to assess the impact of solids retention time on the bacterial diversity of activated sludge. *Applied and Environmental Microbiology* **71**: 5814–5822.
- Saleem M, Moe LA. (2014). Multitrophic microbial interactions for eco- and agro-biotechnological processes: theory and practice. *Trends in Biotechnology* **32**: 529–537.
- Salvado H, Gracia MP, Amigó JM. (1995). Capability of ciliated protozoa as indicators of effluent quality in activated sludge plants. *Water Research* **29**: 1041–1050.

- Sanapareddy N, Hamp TJ, Gonzalez LC, Hilger HA, Fodor A a., Clinton SM. (2009). Molecular diversity of a north carolina wastewater treatment plant as revealed by pyrosequencing. *Applied and Environmental Microbiology* **75**: 1688–1696.
- Sanguino L, Franqueville L, Vogel TM, Larose C. (2015). Linking environmental prokaryotic viruses and their host through CRISPRs. *FEMS Microbiology Ecology* **91**: fiv046.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537–7541.
- Seed KD, Lazinski DW, Calderwood SB, Camilli A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* **494**: 489–491.
- Serrano S, Arregui L, Perez-Uz B, Calvo P, Guine A. (2008). Guidelines for the identification of ciliates in wastewater treatment plants. IWA Publishing.
- Seviour R, Nielsen PH. (2010). Microbial ecology of activated sludge. IWA Publishing.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al.* (2003). Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research* **13**: 2498–2504.
- Shapiro OH, Kushmaro A, Brenner A. (2010). Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater. *The ISME journal* **4**: 327–336.
- Sherr EB, Sherr BF. (2002). Significance of predation by protists in aquatic microbial food webs. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* **81**: 293–308.
- Sime-Ngando T. (2014). Environmental bacteriophages: Viruses of microbes in aquatic ecosystems. *Frontiers in Microbiology* **5**: 355.
- Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K, Nikolausz M. (2007). Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiology Ecology* **60**: 341–350.
- Smith S, Vigilant L, Morin PA. (2002). The effects of sequence length and oligonucleotide mismatches on 5' exonuclease assay efficiency. *Nucleic acids research* **30**: e111.
- Stephenson D, Stephenson T. (1992). Bioaugmentation for enhancing biological wastewater treatment. *Biotechnology Advances* **10**: 549–559.
- Stern A, Mick E, Tirosh I, Sagy O, Sorek R. (2012). CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Research* **22**: 1985–1994.
- Stokholm-Bjerregaard M, McIlroy SJ, Nierychlo M, Karst SM, Albertsen M, Nielsen PH. (2017). A critical assessment of the microorganisms proposed to be important to enhanced biological phosphorus removal in full-scale wastewater treatment systems. *Frontiers in Microbiology* **8**: 718.
- Strous M, Fuerst J a, Kramer EH, Logemann S, Muyzer G, van de Pas-Schoonen KT, *et al.* (1999). Missing lithotroph identified as new planctomycete. *Nature* **400**: 446–449.
- Sutherland IW. (2001). Exopolysaccharides in biofilms, flocs and related structures. *Water Science and Technology* **43**: 77–86.
- Tamaki H, Zhang R, Angly FE, Nakamura S, Hong PY, Yasunaga T, *et al.* (2012). Metagenomic analysis of DNA viruses in a wastewater treatment plant in tropical climate. *Environmental Microbiology* **14**: 441–452.

- Thingstad TF. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography* **45**: 1320–1328.
- Thomas R, Berdjeb L, Sime-Ngando T, Jacquet S. (2011). Viral abundance, production, decay rates and life strategies (lysogeny versus lysis) in Lake Bourget (France). *Environmental Microbiology* **13**: 616–630.
- Vandermeer J. (2006). Oscillating populations and biodiversity maintenance. *Bioscience* **56**: 967–975.
- Vieira-Silva S, Rocha EPC. (2010). The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genetics* **6**: e1000808.
- Vinekar K, Schaad N, Ali M, Lucien B, Leshem E, Oboho IK, *et al.* (2016). Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science* **34**: 1–15.
- Vuono DC, Benecke J, Henkel J, Navidi WC, Cath TY, Munakata-Marr J, *et al.* (2015). Disturbance and temporal partitioning of the activated sludge metacommunity. *The ISME Journal* **9**: 425–435.
- Wagner-Döbler I, Eichner CA, Erb RW, Timmis KN. (1999). Community shocks in the activated sludge microbial analysis of bioprotection from pollutant thermal gradient gel electrophoresis. *Applied and Environmental Microbiology* **65**: 102–9.
- Wang X, Hu M, Xia Y, Wen X, Ding K. (2012). Pyrosequencing analysis of bacterial diversity in 14 wastewater treatment systems in china. *Applied and Environmental Microbiology* **78**: 7042–7047.
- Weinberger AD, Sun CL, Pluciński MM, Deneff VJ, Thomas BC, Horvath P, *et al.* (2012). Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS Computational Biology* **8**: e1002475.
- Weitz JS, Hartman H, Levin SA. (2005). Coevolutionary arms races between bacteria and bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 9535–40.
- Wells GF, Park HD, Eggleston B, Francis CA, Criddle CS. (2011). Fine-scale bacterial community dynamics and the taxa-time relationship within a full-scale activated sludge bioreactor. *Water Research* **45**: 5476–5488.
- Wu M, Scott AJ. (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28**: 1033–1034.
- Wu Q, Liu W-T. (2009). Determination of virus abundance, diversity and distribution in a municipal wastewater treatment plant. *Water research* **43**: 1101–1109.
- Xia LC, Ai D, Cram J, Fuhrman JA, Sun F. (2013). Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. *Bioinformatics* **29**: 230–237.
- Xia LC, Steele JA, Cram JA, Cardon ZG, Simmons SL, Vallino JJ, *et al.* (2011). Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC Systems Biology* **5**: S15.
- Xia S, Duan L, Song Y, Li J, Piceno YM, Andersen GL, *et al.* (2010). Bacterial community structure in geographically distributed biological wastewater treatment reactors. *Environmental Science and Technology* **44**: 7391–7396.
- Ye L, Zhang T. (2013). Bacterial communities in different sections of a municipal wastewater treatment plant revealed by 16S rDNA 454 pyrosequencing. *Applied Microbiology and Biotechnology* **97**: 2681–2690.

Yi T, Lee E-H, Kang S, Shin J, Cho K-S. (2012). Structure and dynamics of microbial community in full-scale activated sludge reactors. *Journal of Industrial Microbiology & Biotechnology* **39**: 19–25.

Yiannakopoulou T V, Kaimakamidou V. (2009). Testing the reliability of protozoa as indicators of wastewater treatment plant performance. *Fresenius Environmental Bulletin* **18**: 146–157.

Yuan C, Lei J, Cole J, Sun Y. (2015). Reconstructing 16S rRNA genes in metagenomic data. *Bioinformatics* **31**: i35–i43.

Zartarian F, Mustin C, Bottero JY, Villemain G, Thomas F, Ailleres L, *et al.* (1994). Spatial Arrangement Of The Components Of Activated-Sludge Floccs. *Water Science And Technology* **30**: 243–250.

Zeng F, Wang Z, Wang Y, Zhou J, Chen T. (2017). Large-scale 16S gene assembly using metagenomics shotgun sequences. *Bioinformatics* **33**: 1447–1456.

Zhang Q, Ye Y. (2017). Not all predicted CRISPR–Cas systems are equal: isolated cas genes and classes of CRISPR like elements. *BMC Bioinformatics* **18**: 92.