

Metavirómica en masas de agua: pasado, presente y perspectivas futuras

Metaviromics in water bodies: past, present and future perspectives

Marcos López-Pérez¹ y Francisco José Fernández²

¹Departamento de Ciencias Ambientales. Universidad Autónoma Metropolitana-Lerma

²Departamento de Biotecnología. Universidad Autónoma Metropolitana-Iztapalapa.
Apdo. Postal 55-535. Col. Vicentina. México, D. F., 09340. México
e-mail fjfp@xanum.uam.mx

López-Pérez M. y F. J. Fernández. 2013. Metavirómica en masas de agua: pasado, presente y perspectivas futuras. *Hidrobiológica* 23 (3): 287-302.

RESUMEN

Los virus son los microorganismos más abundantes en la biosfera. Debido a su gran versatilidad para infectar a todo tipo de organismos y a su condición de parásitos intracelulares obligados, tienen efectos muy importantes en la dinámica de los ecosistemas. La identificación y análisis de las interacciones en las comunidades víricas es un trabajo complicado, debido a la dificultad intrínseca en el cultivo de estos microorganismos. Además, la carencia de secuencias de ADN conservadas entre grupos imposibilita su identificación con una tecnología del tipo ADN_r, la común en la actualidad para la identificación de los organismos vivos. La metagenómica surge como una alternativa para la caracterización molecular de los virus en sus ambientes naturales, una alternativa que arroja información relevante en cuanto a la composición y estructura en dicho ambiente. En este artículo de revisión se hace un repaso del papel de los virus en los ecosistemas acuáticos, además de enfatizar en la importancia de una adecuada implementación de las técnicas de metagenómica. Finalmente, se revisan las principales herramientas informáticas para el tratamiento de los datos metagenómicos obtenidos.

Palabras clave: Acervo genético, metagenomas víricos, papel ecológico, virus acuáticos.

ABSTRACT

Viruses are the most abundant organisms in the biosphere. Because of their versatility to infect all types of organisms and their status as obligate intracellular parasites, they have important effects on ecosystem dynamics. Identification and analysis of interactions in viral communities is a hard work, due to the inherent difficulty in culturing these microorganisms. Furthermore, the lack of conserved DNA sequences among groups precludes identification of the viruses by using rDNA technology, as the current molecular standard for the identification of living organisms. Metagenomics arises as an alternative to the molecular characterization of the virus in its natural environments, which yields information relevant to its composition and structure. In this article, we review the role of viruses in aquatic ecosystems, in addition to emphasizing the importance of proper implementation of these metagenomic techniques. Finally, we review the main tools for the treatment of the obtained metagenomic data.

Key words: Genetic stock, metaviromics, ecological role, aquatic viruses.

INTRODUCCIÓN

Los virus son partículas pequeñas, normalmente con tamaños de entre 20 y 200 nm. Básicamente están constituidos de material

genético (ADN o ARN, de cadena sencilla o doble), rodeado por una envoltura de proteínas. Algunos virus cuentan, además, con una capa externa de naturaleza lipídica. Se caracterizan fundamentalmente por carecer de metabolismo propio, lo que les obli-

ga a interactuar con otros organismos y utilizar su maquinaria bioquímica y metabólica; de tal manera que son parásitos intracelulares obligados.

Tradicionalmente, los organismos procariotas han sido considerados la parte más relevante de los ecosistemas acuáticos, en cuanto a su impacto en la dinámica de estos ambientes. Sin embargo, en los últimos 15 años, y con base en las características generales de los ciclos de vida víricos, las investigaciones han demostrado que los virus influyen de forma muy importante en los procesos bioquímicos, ciclos de nutrientes, distribución de tamaños de partículas, biodiversidad relativa y distribución de algas y bacterias (Fuhrman, 1999). Además, existen reportes que indican que son los organismos más abundantes en ambientes acuáticos y están implicados en el control de las poblaciones, tanto de fitoplancton como de bacterias (Sieburth *et al.*, 1988; Bratbak *et al.*, 1990). El estudio del ADN de comunidades completas (metagenómica) aplicada a los virus, lo que también es conocido como metavirómica, inició en el año 2002, con la caracterización de comunidades víricas marinas (Breitbart *et al.*, 2002). La mayor parte de las publicaciones sobre metagenómica vírica en cuerpos y masas de agua se refiere a zonas oceánicas, sin embargo existen reportes donde este análisis se extiende a masas de agua dulce, como aguas residuales (Cantalupo *et al.*, 2011), lagos (López-Bueno *et al.*, 2009) y humedales (Jackson & Jackson, 2008). En la tabla 1 se muestran varias familias de virus de ADN y ARN que han sido identificados en algunos de los principales estudios de metavirómica en ecosistemas acuáticos.

LA TAXONOMÍA DE VIRUS EN EL CONTEXTO DE LA METAGENÓMICA

La taxonomía de los virus tradicionalmente ha resultado ser problemática cuando se equipara con la taxonomía del resto de los organismos que componen la biosfera. El problema fundamental deriva de que estos organismos se encuentran en la interfase conceptual entre la materia viva y la inerte, siendo estructural y metabólicamente mucho menos complejos, lo que dificulta la búsqueda de características especiales que diferencien a nivel específico una estirpe de otras. A pesar de estos inconvenientes, el comité internacional para la taxonomía de virus (*International Committee Taxonomy of Viruses*, ICTV), ha definido una serie de criterios para clasificar taxonómicamente a estos organismos. Los tres criterios que se tienen en cuenta son: el tipo de organismo al que parasitan (es decir, si éste pertenece al dominio Eukarya, Bacteria o Archaea), la morfología del virión o partícula vírica (si son virus envueltos o desnudos y consideraciones acerca de la forma de sus cápsidas) y la constitución de su genoma (es decir, si son virus de ADN de doble cadena o cadena sencilla, o bien si son virus de ARN de una o dos cadenas y otras características, como si las moléculas de ácidos nucleicos están fraccionadas o no). Para virus de especial interés también se tienen en cuenta

otros criterios, como los síntomas que genera su patogénesis, antigenicidad, perfil de proteínas y rango de hospedadores (Medina & García-Sastre, 2011). En metavirómica es importante resaltar que por una parte las muestras ambientales contienen una enorme diversidad viral y, por otra, la mayoría de los virus en el planeta son aún desconocidos (Kuzmin *et al.*, 2006). En los análisis metavirómicos frecuentemente se analiza a través de microscopía de epifluorescencia la fracción vírica de la muestra (Suttle *et al.*, 1990; Marie *et al.*, 1999), identificando la morfología de los virus de un modo superficial, debido a la gran diversidad viral de las muestras. Adicionalmente, es posible inferir el tipo de hospedadores que tienen los virus de una muestra haciendo una exploración de la diversidad de los organismos presentes en la misma. Sin embargo, el análisis metavirómico suele perder información referente a la morfología y los huéspedes a los que parasitan, al centrarse fundamentalmente en el análisis de las secuencias de ácidos nucleicos recuperadas. Por ello, es frecuente que en los análisis de metavirómica la clasificación se quede en niveles altos en la jerarquía taxonómica, como las familias (Middelboe *et al.*, 2003), y no se busque como objetivo principal establecer criterios de diversidad viral, sino profundizar en el análisis de la dinámica de estas comunidades virales y su interacción con el ecosistema donde viven. Un ejemplo novedoso del análisis metavirómico, centrado en elementos funcionales y de dinámica de interacción entre la comunidad viral y el ecosistema, es el que se ha basado en el papel de los genes implicados en fotosíntesis aislados de genomas de los virus (Lindell *et al.*, 2005).

VIRUS EN ECOSISTEMAS ACUÁTICOS; IMPORTANCIA, ABUNDANCIA Y ACTIVIDAD

La abundancia relativa de estos microorganismos en el agua ha sido medida, como ya se comentó, en masas de agua oceánica y de agua dulce. Con respecto a las masas de agua oceánicas, algunas investigaciones indican que los virus son las entidades biológicas más abundantes en el mar, tanto en zonas cercanas a la costa como en zonas alejadas, en aguas boreales y tropicales (Wommack *et al.*, 1992; Boehme *et al.*, 1993; Børshiem, 1993; Cochlan *et al.*, 1993; Paul *et al.*, 1993; Maranger *et al.*, 1994; Hara *et al.*, 1996; Noble & Fuhrman, 1998), llegando a superar entre 5 y 25 veces en número de partículas-individuos a las poblaciones bacterianas y constituyendo el factor más importante para el control de dichas poblaciones (Paul *et al.*, 1993; Peduzzi & Schiemer, 2004), lo que repercute directamente en la dinámica del ecosistema. De este mismo efecto sobre las bacterias se deriva el impacto de su actividad en el ciclo de varios nutrientes, como en el caso del carbono (Fig. 1). Para profundizar en este sentido, es importante aclarar que en los ciclos víricos se pueden describir tres tipos básicos de reproducción: a) infección o ciclo lítico, en el que el virus ataca a la célula hospedadora; a continuación se produce una abundante progenie en el interior de la célula que termina lisándose y consecuentemente mueren-

TABLA 1: Familias de virus ADN y ARN identificados mediante análisis metavirómicos en ecosistemas acuáticos. TABLA 1: Principales estudios y familias de virus ADN y ARN identificados mediante análisis metavirómicos en ecosistemas acuáticos. Las iniciales P y S se refieren a la tecnología de secuenciación (pirosecuenciación y *shotgun*, respectivamente) utilizada en cada referencia. Las iniciales que aparecen en esta columna se refieren a la técnica utilizada en la referencia. El orden se corresponde respectivamente con el orden de las referencias.

Tipo de ecosistema/organismo	Familia vírica	Hospedero	Tipo de técnica de secuenciación	Referencias
Virus de ADN				
Virus de ADN doble hebra				
Lagos antárticos, agua reciclada, aguas termales, océano Ártico	Herpesviridae	Mamíferos	P, S, P, P	Angly <i>et al.</i> (2006); Schoenfeld <i>et al.</i> (2008); Vega Thurber <i>et al.</i> (2008); López-Bueno <i>et al.</i> (2009)
Aguas termales	Fuselloviridae	Arqueas	S	Schoenfeld <i>et al.</i> (2008)
Lagos antárticos	Phycodnaviridae	Algas	P	López-Bueno <i>et al.</i> (2009)
Lagos antárticos	Mimiviridae	Algas	P	López-Bueno <i>et al.</i> 2009
Arrecifes de coral	Adenoviridae	Mamíferos	S	Marhaver <i>et al.</i> (2008)
Lagos antárticos, aguas termales	Rudiviridae	Archaea	S, P	Schoenfeld <i>et al.</i> (2008); López-Bueno <i>et al.</i> (2009)
Océano Ártico	Papillomaviridae	Mamíferos	P	Angly <i>et al.</i> (2006)
Agua reciclada	Siphoviridae	Bacterias	P	Rosario <i>et al.</i> (2009)
Virus de ADN de cadena sencilla				
Lagos antárticos, arrecifes de coral	Nanoviridae	Plantas	P, P	McDaniel <i>et al.</i> (2008); López-Bueno <i>et al.</i> (2009)
Tortugas de mar	Gyrovirus genus	Aves	S	Fei Fan <i>et al.</i> (2009b)
Lagos antárticos	Circoviridae	Mamíferos/Aves	P	López-Bueno <i>et al.</i> (2009)
Lagos antárticos, océano Ártico	Inoviridae	Bacteria	P, P	Angly <i>et al.</i> (2006); López-Bueno <i>et al.</i> (2009)
Lagos antárticos, agua reciclada, arrecifes de coral	Geminiviridae	Plantas	S, S, P, P, S	Breitbart <i>et al.</i> (2004); Kim <i>et al.</i> (2008); Vega Thurber <i>et al.</i> (2008); López-Bueno <i>et al.</i> (2009); Blinkova <i>et al.</i> (2010)
<i>Phoca vitulina</i> , <i>Eumetopias jubatus</i>	Anelloviridae	Mamíferos	S, S	Fei Fan <i>et al.</i> (2009a); Fei Fan <i>et al.</i> (2011b)
Lagos antárticos	Microviridae	Bacteria	P	López-Bueno <i>et al.</i> (2009)
Virus de ARN				
Virus ARN de cadena sencilla, polaridad positiva o negativa				
Agua reciclada, lagos artificiales	Picornaviridae	Aves/Mamíferos	P, P, P	Djikeng <i>et al.</i> (2009); Kapoor <i>et al.</i> (2008); Li <i>et al.</i> (2010)
Océano Ártico	Retroviridae	Aves/Mamíferos	P, P	Angly <i>et al.</i> (2006); Park <i>et al.</i> (2011)
Lagos artificiales	Flaviviridae	Mamíferos	P	Djikeng <i>et al.</i> (2009)
Lagos artificiales	Flexiviridae	Plantas	P	Djikeng <i>et al.</i> (2009)
Océano Ártico	Coronaviridae	Aves/Mamíferos	P	Angly <i>et al.</i> (2006)
Virus de ARN cadena doble				
Lagos artificiales	Picobirnavirus	Mamíferos	P	Djikeng <i>et al.</i> (2009)
Lagos artificiales	Reoviridae	Aves/Mamíferos/ Insectos/ Plantas/Hongos	P	Djikeng <i>et al.</i> (2009)
Lagos artificiales	Partiviridae	Hongos	P	Djikeng <i>et al.</i> (2009)

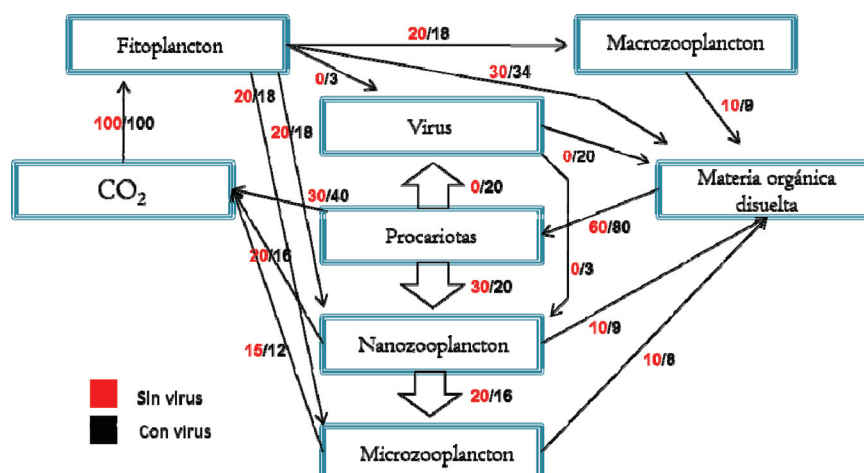


Figura 1. Modelo del flujo de carbono en ecosistemas marinos, en el que se teoriza la ausencia y presencia de virus. Este modelo presupone que la lisis debida a actividad lítica supone un 50% de la mortalidad bacteriana y un 7% de la mortalidad del fitoplancton. Los números asociados a las flechas se refieren a unidades de carbono transferidas entre los diferentes macro-componentes del ecosistema. Modificado de Fuhrman, 1992 (ver dicho trabajo para profundizar en el análisis y en las cifras).

do. Posteriormente, los viriones son liberados al ambiente donde comienzan un nuevo ciclo. b) infección crónica, donde, a diferencia de la infección lítica, la célula hospedadora no muere al ser liberada la progenie, sirviendo así como célula hospedadora durante sucesivas generaciones. c) ciclo lisogénico, en donde el ácido nucleico del virus, una vez integrado en el ADN genómico de la célula hospedadora, se mantiene latente, reproduciendo su material genético con el de la célula que lo contiene (en lo que se denomina estado de *provirus*) (Fuhrman, 1999). En este último caso, cambios en la fisiología celular del hospedador, como por ejemplo condiciones de estrés, pueden derivar en la activación de una infección o ciclo lítico (Fuhrman, 1999). Es importante entender, por lo tanto, cómo esta dinámica reproductiva puede repercutir en la redistribución de nutrientes en el ecosistema, ya que en función de ella será mayor o menor la mortalidad bacteriana y, en consecuencia, la fracción de biomasa que está disponible para la captación por otros estratos de organismos de la pirámide trófica en el ecosistema. En el momento en que las células se lisan, se liberan al medio los virus y toda la biomasa celular queda en forma de materia orgánica (MO) disuelta. Esta MO disuelta, en su mayoría, impacta de nuevo en la generación de nuevas células bacterianas (Bratbak *et al.*, 1990; Proctor *et al.*, 1990). De forma independiente, es interesante destacar que la lisis bacteriana y la subsiguiente liberación del contenido celular repercute en la concentración y en las condiciones fisicoquímicas de esa zona de la columna de agua. Un ejemplo de este efecto es la capacidad de agregación de componentes celulares debida a la presencia de polímeros (Proctor *et al.*, 1991). De forma alternativa, es posible pensar en agregados celulares que tienden a descender a aguas más profundas, pero por efectos de la lisis vírica se disgregan permaneciendo en la zona eufótica (Proctor *et al.*, 1991).

Con respecto a los virus que afectan a algas y cianobacterias, éstos permiten que sus tasas declinen durante las explosiones de población de estos organismos y se restituya la distribución de nutrientes en el ecosistema (Yau *et al.*, 2011). Adicionalmente, un aspecto interesante que debe ser resaltado es el papel que pueden jugar los virus en la dinámica del clima. En este sentido, se ha publicado que la lisis celular (la cual es derivada en gran medida de la actividad lítica de los virus) libera DMS (dimetil sulfuro), un gas que se ha comprobado que interviene en los procesos de generación de las nubes. Existen estudios que han demostrado que la infección vírica causa la liberación intracelular del precursor del DMS (el dimetildisulfopropionato, DMSP) en organismos del fitoplancton como *Micromonas pusilla* (R. W. Butcher) I. Manton *et al.* Parke. Este gas puede ser incorporado por otros microorganismos o bien puede ser liberado a la atmósfera (Kiene & Bates, 1990). Los virus, además, afectan a la diversidad y estructura de las comunidades microbianas, interviniendo en algunos casos (como ocurre con los bacteriófagos), en el intercambio de material genético entre diferentes miembros de las comunidades bacterianas, algo que también contribuye a la dinámica evolutiva en estos ecosistemas (Fuhrman, 1992; Wilhelm & Suttle, 1999; Wommack & Colwell, 2000; Brussaard, 2004; Weinbauer, 2004; Weinbauer & Rassoulzadegan, 2004; Suttle, 2005, 2007; Munn, 2006; Ogunsein, 2008).

DIVERSIDAD VÍRICA EN AGUA DULCE

Los estudios de diversidad vírica realizados en masas de agua dulce han sido establecidos con base en las características fundamentales que diferencian cada uno de estos ambientes. En este sentido hay estudios realizados sobre humedales artificiales y naturales, ríos, lagos, ambientes extremos y aguas residuales.

Lagos. Los estudios de metagenómica viral en lagos se han realizado en gran variedad de ecosistemas. Uno de los ejemplos que aporta datos más interesantes es el trabajo de López-Bueno y colaboradores (2009), realizado en lagos antárticos. En dicho trabajo, los análisis metagenómicos revelaron que las comunidades de virus son enormemente diversas y abundantes. A diferencia de lo hallado en otros viomas lacustres, donde normalmente hay una abundancia dominante de bacteriófagos, en este ambiente se encontró una gran presencia de virus de eucariotas, incluyendo ficolnavirus y otros virus de ADN de cadena sencilla. Clasen y colaboradores (2008) realizaron investigaciones sobre la abundancia viral en 16 lagos. Los resultados indicaron que existe una enorme variabilidad entre las comunidades víricas en función del ecosistema. Adicionalmente, establecieron un interesante paralelismo entre las poblaciones de bacterias y cianobacterias y la abundancia relativa de los virus en diferentes ambientes climáticos, algo que ha sido reportado también en otros trabajos (Djikeng *et al.*, 2009; Rodríguez-Brito *et al.*, 2010). Debe destacarse la posible incidencia de los virus en el ambiente, por su repercusión en las concentraciones relativas de gases atmosféricos. Esto es debido a la fuerte presión de selección que ejercen sobre las cianobacterias y, por lo tanto, a las alteraciones del metabolismo fotosintético de estos microorganismos (López-Bueno *et al.*, 2009).

Humedales. Es interesante resaltar que los estudios morfológicos y moleculares en humedales indican que una gran mayoría de los virus encontrados en estos ambientes son bacteriófagos. Sin embargo, es común identificar también virus que parasitan algas, protozoos e incluso plantas vasculares (Wommack & Cowell, 2000). En cuanto a los principales trabajos que analizan la diversidad y dinámica de las poblaciones víricas en humedales, destacan los realizados en humedales naturales, aunque también hay estudios sobre la diversidad viral en humedales artificiales (Gersberg *et al.*, 1989; Nokes *et al.*, 2003; Vidales *et al.*, 2003; Vymazal, 2005). En los naturales, las investigaciones se han enfocado recurrentemente en el análisis de las condiciones ambientales para explicar las fluctuaciones en la diversidad vírica (Fischer *et al.*, 2003; Filippini *et al.*, 2006; Filippini & Middelboe, 2007). Un ejemplo de dichos estudios fue el realizado en el humedal Talladega, en Alabama, EUA (Farnell-Jackson & Ward, 2003), en el que se evaluó el patrón de diversidad vírica asociado con las diferentes estaciones del año. Los autores hallaron que las diferencias estacionales estaban asociadas con variaciones en la abundancia de las poblaciones bacterianas, pero también del carbono orgánico disuelto y de otros nutrientes de naturaleza inorgánica (Farnell-Jackson & Ward, 2003). La diversidad en los resultados obtenidos fue explicada por los autores debido a la alta heterogeneidad de estos ecosistemas. Es posible argumentar, por ejemplo, que la profundidad será un factor muy relevante, de manera que si hay poca profundidad, la comunidad que habita el bentos o los virus asociados a plantas en el sedimento podrán constituir una frac-

ción muy importante de la muestra sometida a análisis. De igual forma, la profundidad afectará directamente a la cantidad de luz que penetra en el agua, lo que afecta a los microorganismos con metabolismo fotosintético y a su vez repercute en la dinámica de las poblaciones bacterianas (y consecuentemente, en los virus que hospedan).

Ríos. El análisis de diversidad vírica en ríos es el menos reportado en la literatura científica. Uno de los primeros trabajos fue realizado en el río Danubio, con la finalidad de evaluar el efecto de las comunidades víricas sobre las poblaciones de bacterias (Mathias *et al.*, 1995). En otro estudio, realizado en una zona de estuario de Francia, se evaluó la distribución del viroplankton en función de las variables ambientales. La abundancia viral fue relacionada con la abundancia y actividad del bacterioplankton, con los pigmentos fotosintéticos, la concentración de nutrientes y otros parámetros físico-químicos, variables que a su vez condicionan la dinámica de las poblaciones bacterianas (Auguet *et al.*, 2005). Uno de los trabajos más recientes en el que se evaluó la diversidad del viroplankton, se realizó en dos ríos de la República Checa (Slovácková & Maršálek, 2008), pudiéndose determinar una correlación entre la abundancia de virus y bacterias con la cantidad de clorofila total y la concentración de ortofosfatos. Finalmente, otro de los estudios en esta misma línea de investigación se realizó en el sedimento de un estuario, evaluándose la distribución de virus y bacterias asociadas a la actividad diagenética (Middelboe *et al.*, 2003).

Ambientes extremos. La presencia de virus en ambientes extremos se debe fundamentalmente a que son capaces de invadir células de organismos de los tres dominios presentes en estos ecosistemas (Eukarya, Archaea y Bacteria), a pesar de que los organismos predominantes en dichos ambientes sean arqueas. En la literatura, el análisis metagenómico de estas comunidades se ha hecho en función del ambiente objeto de estudio, como es el caso del análisis que se realizó sobre fuentes geotérmicas en Yellowstone (Schoenfeld *et al.*, 2008). En este estudio se encontró que la diversidad vírica fue similar a la hallada en ambientes no extremos. Además, y a diferencia de lo que se pensaba anteriormente (que los virus sometidos a condiciones extremas de temperatura tenderían a permanecer en el interior celular, manteniendo así un ciclo de vida no lítico), las secuencias halladas en los virus aislados en este trabajo mostraron secuencias que codificaban en su mayoría para un ciclo lítico (Schoenfeld *et al.*, 2008).

Otro de los ambientes extremos donde se han realizado estos estudios son los lagos hipersalinos. En estos ecosistemas, a pesar de que son ambientes con una gran carga de partículas virales (10^9 partículas/ml), la cantidad de halovirus aislados es muy baja (Santos *et al.*, 2010). Para explicar la gran abundancia de partículas virales en estos ambientes se proponen dos hipótesis, a) la alta estabilidad física de las partículas víricas en estos ambientes y b) una alta producción de partículas virales, propiciada

por las condiciones de este entorno (Santos *et al.*, 2007; Schapira *et al.*, 2009; Bettarel *et al.*, 2010). Es muy llamativa la alta relación partícula viral por bacteria encontrada en ambientes hipersalinos (alrededor de 100), comparada con la relación en ambientes acuáticos no extremos (de 5 a 10 partículas virales) (Wommack & Colwell, 2000; Sime-Ngando *et al.*, 2010). De forma general, la cantidad de virus en estos ambientes puede ser correlacionada con la cantidad de células en una muestra (Santos *et al.*, 2010). Esta tendencia ha sido confirmada en varios estudios adicionales (Guixa-Boixareu *et al.*, 1996; Bettarel *et al.*, 2010).

DIVERSIDAD VÍRICA EN AGUA SALADA

Los estudios de metaviromica en ecosistemas marinos diferencian, en general, los ambientes costeros de los no costeros (también llamado océano profundo), así como la localización geográfica del océano objeto del estudio.

Ambientes costeros. Uno de los primeros trabajos que se enfocó al análisis de la dinámica viral en este ambiente estableció que no hay cambios apreciables en las poblaciones virales muestreadas en ambientes costeros, a pesar de que se aislaron en recipientes y se sometieron a cambios en la concentración de nutrientes, irradiación con fuertes intensidades de luz o drásticos cambios de temperatura. Además, comprobaron fuertes cambios en la abundancia viral del ambiente en función de parámetros temporales y de la profundidad (Bratbak *et al.*, 1996).

Otros estudios se han enfocado al estudio de la diversidad viral utilizando el análisis de genes relacionados con la actividad fotosintética. En este sentido, es destacable la presencia de genes fotosintéticos en poblaciones virales de diferentes tamaños (Sandaa *et al.*, 2008). Este resultado es indicativo de cómo los virus pueden estimular el metabolismo de sus hospedadores con el objetivo de desviar energía hacia la síntesis de nuevos viriones.

Finalmente, uno de los trabajos que más repercusión ha tenido en este tipo de ambientes describió el análisis metagenómico de virus de ARN (Culley *et al.*, 2006). Se encontró que los virus de ARN de cadena sencilla (ya fueran de polaridad positiva o negativa) y los de ARN de doble cadena afectan a todos los niveles tróficos (lo que generaría un fuerte impacto, tanto económico como ecológico). El análisis también reveló que contienen genomas de pequeño tamaño, una variable que facilitó el posterior análisis metagenómico. La mayoría de los virus de ARN aislados en estos ambientes marinos son evolutivamente distantes de los virus conocidos, y su diversidad genética ha sido muy poco explorada con respecto a los que afectan a la comunidad procariota. Adicionalmente, no se encontraron bacteriófagos de ARN, lo cual soporta la teoría de que la mayoría de bacteriófagos marinos tienen genomas de ADN (Weinbauer, 2004).

Ambientes no costeros. Los análisis metagenómicos y de diversidad viral realizados en océanos profundos son más escasos.

Uno de los más recientes enfatiza que la mayoría de las secuencias obtenidas no han podido ser identificadas por comparación con las bases de datos (Steward & Preston, 2011), lo cual es una señal de lo poco que se conoce de la diversidad vírica de estos ambientes. Además, en este trabajo se pone de manifiesto que la mayoría de los virus encontrados (94%), fueron bacteriófagos y sólo el 6% correspondió a virus que afectan a organismos eucariotas. En otro estudio se indica que en la zona fótica predominan los bacteriófagos que infectan cianobacterias, mientras que conforme se incrementa la profundidad descende la abundancia relativa de estos virus (Williamson *et al.*, 2008). Por otra parte, algunos autores han propuesto que en estos ambientes los virus son uno de los elementos fundamentales para la generación de diversidad microbiana, gracias al intercambio genético que favorecen en sus ciclos de infección (Hendrix *et al.*, 2000). Finalmente, uno de los trabajos que más trascendencia ha tenido en este tipo de ambiente, se realizó en cuatro regiones oceánicas diferentes (Océano Ártico, la costa de British Columbia, el Golfo de México y el Mar de los Sargazos; Angly *et al.*, 2006). Los análisis metaviromicos se caracterizaron porque los procesos de ensamblaje de las secuencias de los virus aislados fueron diferentes, lo que probablemente reflejaría la existencia de diferentes presiones de selección, indicando de forma adicional que el grado de diversidad viral también se ve afectado por procesos de migración. Asimismo, este estudio reflejó la variabilidad existente en función de la latitud y la longitud, dejando claro que a pesar de que hay especies endémicas, la gran mayoría de la especies de virus están repartidas en diferentes regiones oceánicas (Angly *et al.*, 2006).

MUESTREO EN ESTUDIOS DE METAGENÓMICA VIRAL EN MASAS DE AGUA

Con el objetivo de evaluar la variabilidad de la comunidad vírica en función de las condiciones ambientales, en la mayoría de los estudios de metaviromica publicados se realiza una toma de muestras distribuida tanto espacial como temporalmente (Djikeng *et al.*, 2009; Oh *et al.*, 2011). Este tipo de muestreo debe aplicarse en contexto con las particularidades del ecosistema objeto de estudio, como por ejemplo la zona más oxigenada o donde la actividad fotosintética sea mayor en la columna de agua. Otro de los procedimientos que sistemáticamente ha sido utilizado en la toma de muestras para estudios de metagenómica viral es el filtrado (Figura 2), y se ha demostrado que el tamaño de las partículas víricas varía en función de la época de año en la que se produzca la recolecta de las muestras (López-Bueno *et al.*, 2009). Eso implica la necesidad de variar el tamaño de poro de los filtros utilizados en el muestreo para evaluar si existen diferencias en la abundancia de las especies, en cuanto a su tamaño, a lo largo del año. Es preciso añadir en este punto que un rasgo que tienen en común los procesos de filtrado es la necesidad de eliminar las partículas en suspensión con un prelavado, de manera que se evite que en-

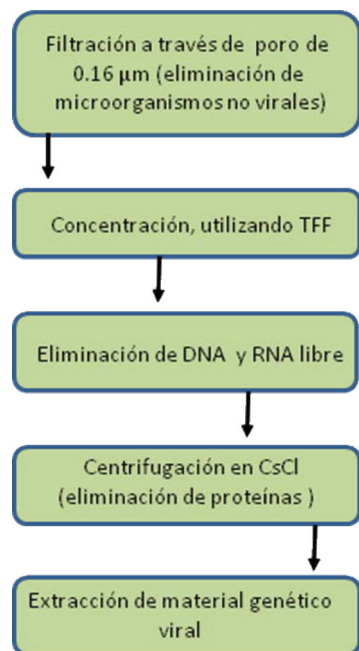


Figura 2. Ejemplo de metodología utilizada en la filtración para la obtención de la fracción vírica en masas de agua.

tren en el análisis ácidos nucleicos celulares. Además, también es habitual utilizar enzimas degradadoras de ácidos nucleicos, con el objetivo de evitar que se analicen ácidos nucleicos libres, en solución (Rebekah *et al.*, 2009). Esta manera de proceder asegura que se capten únicamente los viriones. Finalmente, el proceso de aislamiento de un virus debe ser confirmado visualmente mediante microscopía electrónica (Marie *et al.*, 1999). Como ejemplo, en el trabajo de Djikeng y colaboradores, antes de procederse al aislamiento de la fracción vírica hicieron un prelavado con filtración tangencial a 2.0 y 1.0 µm. Seguidamente, se aplicó ultracentrifugación o filtración tangencial para partículas por encima de 300.000 Da (Djikeng *et al.*, 2009). Otros trabajos hacen un filtrado preliminar de la muestra con papel Wattman de 1.6 µm de tamaño de poro y posteriormente aplican una filtración tangencial con un tamaño de poro de 0.22 µm, aplicando presión con una bomba peristáltica.

PROBLEMAS FRECUENTES EN LA METAGENÓMICA DE VIRUS

A pesar de que los virus son agentes biológicos presentes prácticamente en toda diversidad de ambientes en el planeta, la comprensión de la dinámica de sus poblaciones en los diferentes ecosistemas está muy limitada, debido fundamentalmente a la imposibilidad de cultivarlos y trabajar con ellos en el laboratorio. De manera adicional, la carencia de secuencias genéticas conservadas entre las especies, como el ADN ribosómico utilizado para eucariotas y procariotas, incrementa todavía más la dificultad

del estudio de estos microorganismos (Breitbart *et al.*, 2002). Además, una vez que se han obtenido los datos metagenómicos de muestras ambientales, con mucha frecuencia se encuentran fragmentos truncados de material genético, elementos de ADN libre y marcos abiertos de lectura incompletos. Finalmente, otra de las dificultades que se plantea en la mayoría de los estudios de metagenómica viral es que la gran mayoría de los virus aislados en muestras ambientales son desconocidos y, consecuentemente, la mayoría de los genes secuenciados no tienen similitud suficiente como para poder ser comparados con las bases de datos (Lorenzi *et al.*, 2011). Debido a estos factores, es necesaria la implementación de tecnologías más sofisticadas en el análisis de secuencia, como la secuenciación *shotgun* (Venter *et al.*, 2004; Edwards & Rohwer, 2005; Delwart, 2007) o la pirosecuenciación (Eriksson *et al.*, 2008). Estas nuevas tecnologías han servido para incrementar considerablemente la información contenida en las bases de datos.

IMPORTANCIA DE LAS TECNOLOGÍAS DE SECUENCIACIÓN EN EL CONTEXTO DE LA METAVIRÓMICA

En el contexto de la metavirómica, es importante conocer las características de las tecnologías de secuenciación, con el propósito de comprender el análisis bioinformático que se requiere y a partir de ahí establecer inferencias sobre el ambiente objeto de estudio. Es necesario hacer énfasis en que los metaviromas son muestras enormemente diversas, a lo que se suma el hecho de que la mayoría de los virus no han sido descritos (lo cual hace de las bases de datos herramientas menos eficientes en el análisis de los mismos). Como ya se mencionó anteriormente, en estudios de metavirómica se han utilizado mayoritariamente dos tipos de metodología para la secuenciación y obtención de los metaviromas: la pirosecuenciación y la secuenciación *shotgun*. La pirosecuenciación es una técnica reciente enfocada a la determinación de un número elevado de secuencias de ADN (Fakhrai-Rad *et al.*, 2002; Margulies *et al.*, 2005). Algunas de las ventajas de este método son la economía y el ser procesos menos laboriosos y más rápidos que otras tecnologías existentes. Sin embargo, esta técnica también conlleva algunas desventajas, como que la longitud de las secuencias obtenidas es mucho más corta (del orden de 100-250 pb o incluso fragmentos más cortos, de 40-60 pb (Fakhrai-Rad *et al.*, 2002), con una tasa de error de entre 5 y 10 errores por kpb). Esta tasa de error tiene efecto en el proceso de ensamblaje, por lo que es preciso implementar metodologías para distinguir estos errores de los producidos por la tasa de variación natural debida a la mutación. En este punto es necesario hacer referencia a dos términos que, de forma general, se tienen en cuenta en análisis posteriores. Por una parte, el término haplotipo, que se utiliza en estudios de metavirómica para referirse a un conjunto de polimorfismos de un solo nucleótido (SNPs) en la secuencia (Purcell *et al.*, 2007). Y por otra, el concepto de cuasiespecie: en

los virus de ARN este concepto es relevante y puede entenderse como distribuciones dinámicas de genomas víricos no idénticos, pero estrechamente relacionados entre sí, sometidos a un proceso continuo de variación genética, competición y selección, que actúan como una unidad de selección. En estos casos, no se podría hablar sólo de un genoma definido y la secuencia consenso aceptada sería un promedio de variantes (Domingo *et al.*, 2000).

El ensamblaje de las secuencias obtenidas por pirosecuenciación requiere un trabajo mayor, dado que los metaviomas están en estos casos constituidos por un número de secuencias (*reads*) muy elevado. A pesar de que la longitud de las secuencias es corta, esta metodología ha sido utilizada de forma recurrente para el análisis de metaviomas (Angly *et al.*, 2006; Kapoor *et al.*, 2008; Djikeng *et al.*, 2009; Culley *et al.*, 2010; Li *et al.*, 2010; Rodríguez-Brito *et al.*, 2010; Boujelben *et al.*, 2012). Es necesario recalcar, en cualquier forma, que reconstruir genomas virales no es el objetivo principal en los análisis de metaviómica, sino más bien enfocar el análisis a un tipo de secuencia concreta de la comunidad vírica y ponerla en contexto con fenómenos de interacción con el ecosistema. Para ejemplificar esta situación, se puede mencionar el trabajo de Sharon y colaboradores, quienes identificaron genes implicados en la fotosíntesis de cianobacterias en genomas virales (Sharon *et al.*, 2009).

La otra metodología ampliamente usada en este tipo de estudios es la secuenciación *shotgun*, también conocida como secuenciación *de novo* (Fei Fan *et al.*, 2009b), una modificación del clásico proceso de secuenciación basado en el método de Sanger (Sanger & Coulson, 1975; Sanger *et al.*, 1977). En referencia a este método de secuenciación, y comparando con la tecnología de pirosecuenciación, es preciso aclarar que en este caso la longitud de los fragmentos obtenida es de entre 800 y 1000 pb, con una tasa de error de 0.01/kpb (Malet *et al.*, 2003; Huse *et al.*, 2007). Hay numerosos ejemplos de estudios donde se utilizó este método para el análisis metaviómico (Djikeng *et al.*, 2009; Oh *et al.*, 2011; Steward & Preston, 2011). La tecnología *shotgun* se desarrolló para lograr secuenciación de genomas; para ello se requiere un trabajo previo de generación de clones, donde cada uno de estos clones lleva un fragmento de la metagenoteca. Este método es más laborioso y más caro, sin embargo provoca la aparición de menos errores. En referencia al ensamblaje, esta tecnología presenta el problema de que cuando existen fragmentos de ADN de secuencia repetida es difícil ensamblar los fragmentos para generar los *contigs* (Schatz *et al.*, 2010).

Una nueva tecnología de secuenciación está adquiriendo una importancia creciente y tiene como objetivo leer grandes fragmentos de ácidos nucleicos (del orden de miles de nucleótidos), la denominada *Illumina Solexa sequencing*, con una capacidad de procesamiento similar a las tecnologías de secuenciación tradicional que producen fragmentos de longitud más corta, co-

mo la secuenciación *shotgun* o la pirosecuenciación (Korlach *et al.*, 2008).

BIOINFORMÁTICA UTILIZADA EN METAGENÓMICA VIRAL

Una de las razones fundamentales por las que la metagenómica se está utilizando es su gran utilidad en el análisis ecológico de los diferentes ambientes, debido a que permite establecer hipótesis referentes a las relaciones entre los miembros de una comunidad. Estas inferencias derivan de diferentes parámetros, siendo algunos de los más importantes la información depositada en las bases de datos y el análisis bioinformático de dichos datos, lo que será objeto de revisión del presente apartado.

Ensamblaje de secuencias. El proceso de ensamblaje en metagenómica se refiere a la unión de pequeños fragmentos de ADN previamente secuenciados (denominados en inglés *contigs*), basándose en la similitud que comparten estos segmentos de ADN en sus extremos. La unión de los *contigs* da como resultado una hebra de ADN continua denominada *scaffold* (en inglés, su traducción más literal sería andamio). En esta metodología es común la utilización de fósidos o vectores de alta capacidad, con lo que consecuentemente se limita el número de *contigs* y por lo tanto se facilita el ensamblaje. Sin embargo, el genoma de los virus se caracteriza por ser pequeño: es común hallar trabajos donde el tamaño de los *contigs* es de muy poca longitud (20-50 pb) (Angly *et al.*, 2006; Boujelben *et al.*, 2012), por lo que la estrategia de ensamblaje suele ser diferente de lo habitual. Para poder realizar este proceso de ensamblaje existen varias herramientas informáticas que han sido utilizadas en análisis metaviómicos, por ejemplo Phrap (www.phrap.org), Arachne (Willner *et al.*, 2009; Wommack *et al.*, 2009), Celera Assembler (Williamson *et al.*, 2008, 2012), PGA (<http://www.paracel.com/>) y CAP3 (Roux *et al.*, 2012). Una vez que se tienen las secuencias ensambladas, éstas son incorporadas a las bases de datos para posteriormente ser comparadas.

Para un adecuado ensamblaje de secuencias, es necesario discriminarlas en varios grupos: fagos, otros virus, elementos móviles, elementos repetidos, material genético bacteriano, de arqueas y de eucariotas. En este sentido, los elementos transponibles, secuencias de inserción, retrotransposones y plásmidos son agrupados primeramente como elementos móviles. En otro gran grupo se conjuntan las secuencias correspondientes a bacterias, arqueas y eucariotas. Para cada grupo de virus filogenéticamente cercano debe hacerse un análisis bioinformático, comparando las secuencias *query* o problema con las bases de datos y generándose subgrupos relacionados (para posteriormente conformar un agrupamiento que englobe a toda la diversidad de la muestra). Un ejemplo establecido para bacteriófagos, que se ha seguido en varios estudios, es el basado en el método de Rohwer y Edwards (2002). Este método opera de la siguiente

manera: se realiza un apareamiento del metaviroma en BLAST. De entre todas las secuencias, aquellas que tienen un *e-value* menor de 0.001 se consideran "especies" o estirpes relacionadas o cercanas a lo depositado en esa base de datos. Aquellas secuencias con menor índice de similitud no son utilizadas. Las secuencias seleccionadas son mapeadas en busca de marcadores, utilizando un método de parsimonia (utilizada para detectar el número de cambios en la secuencia debidos a mutaciones) y una función de cálculo de probabilidad. Posteriormente se utiliza un software con el objetivo de identificar los diferentes *contigs*, por ejemplo Sequencher 4.0 (Gene Codes, Ann Arbor, MI, USA). En estos paquetes informáticos se debe predeterminar el grado de solapamiento de las diferentes lecturas (*reads*), por lo general se establece una longitud de secuencia solapada que permita diferenciar fagos filogenéticamente muy cercanos, como por ejemplo el ϕ T7 y ϕ T3 en *Escherichia coli* (Migula) Castellani & Chalmers (Rohwer & Edwards 2002). Este tipo de análisis puede hacerse con otros paquetes informáticos para determinar qué tan parecida es la estructura de la población que se genera con ambos modelos. El proceso general de tratamiento de los datos se muestra en la figura 3. Este tipo de metodología ha sido seguida en varios estudios de metavirómica (Rodríguez-Brito *et al.*, 2010). En otros estudios, particularmente para la clasificación y análisis de la diversidad de fagos cercanos a los T4, se ha utilizado la proteína gp23, utilizando cebadores específicos para amplificar el

gen codificante a partir de las muestras ambientales. Los clones resultantes se organizan en un dendrograma (López-Bueno *et al.*, 2009). De forma general, algunas de las principales plataformas utilizadas para este tipo de análisis son la Virus Database del ICTV (<http://www.ncbi.nlm.nih.gov/ICTVdb/>), utilizando para ello tBLASTx. Otra de las bases de datos comúnmente utilizadas para fagos es la que se alberga en la página <http://www-rohan.sdsu.edu/~brodrigu/PPT/>.

Finalmente, uno de los procesos que se realiza con el análisis de las comparaciones es la generación de árboles filogenéticos. Para ello se puede utilizar la herramienta PHY-FI (Fredslund, 2006).

Predicción de genes y anotación. La predicción de genes (*gene prediction* o *gene calling*) es el procedimiento para la identificación de proteínas y secuencias de ARN codificadas en el ADN de interés. Este procedimiento es particularmente difícil en los análisis de metavirómica, debido a que los virus usualmente no comparten la mayoría de sus genes con otras especies y esto dificulta mucho el establecimiento de relaciones filogenéticas basadas en similitud de secuencias. Este proceso puede realizarse con secuencias ensambladas o bien con secuencias no ensambladas del metagenoma. En cualquiera de los dos casos, existen dos metodologías generales que han sido utilizadas en análisis metavirómicos. En una de ellas, los análisis comparativos se basan

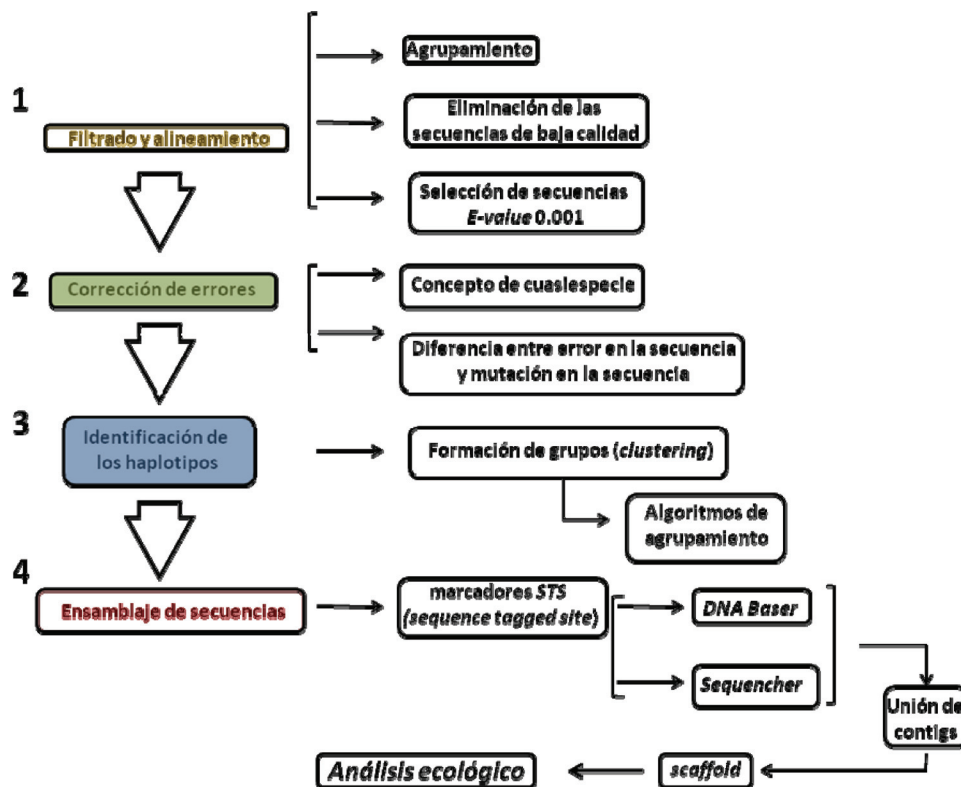


Figura 3. Esquema general de tratamiento informático de los datos de un metaviroma.

en identificar similitud con secuencias depositadas en bases de datos (mediante una comparación BLAST, por ejemplo) (Breitbart *et al.*, 2002; Williamson *et al.*, 2008; Cantalupo *et al.*, 2011). La otra alternativa consiste en utilizar, al inicio de la predicción de genes, relaciones intrínsecas entre las características del ADN a discriminar (diferenciando, por ejemplo, entre regiones codificantes y no codificantes), de manera que se puedan identificar genes sin necesidad de identificar similitud con las secuencias depositadas en las bases de datos. Hay estudios de metaviromica que utilizan con esta finalidad plataformas informáticas como Metagene (http://metagene.cb.k.u-tokyo.ac.jp/metagene/download_mga.html) (Noguchi *et al.*, 2008; Parsley *et al.*, 2010; Tanenbaum *et al.*, 2010; Lorenzi *et al.*, 2011).

Análisis de datos. Uno de los principales problemas que se presentan en el análisis de los datos en estudios de metagenómica viral es cómo detectar la presencia de secuencias virales novedosas que sean muy divergentes respecto de aquellas depositadas en bases de datos, y en las que por tanto no es posible encontrar similitudes con algoritmos de búsqueda del tipo tBLASTx (Delwart, 2007). Es particularmente importante destacar, a este respecto, que entre el 5 y el 30% de las secuencias de nucleótidos o aminoácidos obtenidas en análisis de muestras ambientales no muestran similitud significativa con las secuencias contenidas en Genbank (Delwart, 2007). Por esa razón, se han desarrollado varios métodos para poder establecer las similitudes. En virus de ARN de cadena sencilla de polaridad positiva, por ejemplo, se ha descubierto una serie de dominios proteicos implicados en la función catalítica de las enzimas (conocido en inglés como *core protein functions*), mediante el análisis de las cuales ha sido posible relacionar secuencias víricas distantes pero con funciones análogas (Koonin & Dolja, 1993). Otro ejemplo de este tipo son algunas estructuras proteicas que forman parte de las cápsidas virales y pertenecen a una familia de helicasas (las denominadas en inglés *jelly-roll capsid protein*). Estas proteínas pueden ser encontradas en partículas virales grandes y pequeñas, tanto de ADN como de ARN (Koonin *et al.*, 2006). Adicionalmente, el uso de matrices de sustitución es otra posible aplicación para establecer relaciones entre proteínas con muy bajo nivel de similitud (Dimmic *et al.*, 2002; Duhaime *et al.*, 2011). Y, finalmente, el análisis de estructura de proteínas *in silico* es otra posibilidad que aún puede ser mejorada y permitiría la detección de virus con alto grado de divergencia a través del análisis de secuencias que codifiquen estructuras conservadas en las proteínas (Delwart, 2007).

Clasificación. En un estudio metagenómico se analizan secuencias aisladas, *contigs* y genes. La asociación entre estos datos y los organismos procedentes de los ambientes de donde derivan estas muestras es siempre un objetivo deseable para la interpretación de la dinámica de un ecosistema en particular. Este proceso de asociación entre los datos de las secuencias y las especies se denomina clasificación (en inglés, *binning*). De for-

ma recurrente, algunos de los genes que han sido utilizados para realizar procedimientos de clasificación se basan en secuencias de marcadores filogenéticos. De forma similar a lo anteriormente mencionado, las comparaciones entre secuencias, junto con herramientas para su visualización como BLAST y MEGAN (Huson *et al.*, 2007), pueden también ser usadas para realizar clasificación de grandes fragmentos de ADN, de forma que se permita generar grupos filogenéticos.

Otra estrategia diferente está basada en la dinámica de procesos celulares, como la utilización de codones preferenciales, sistemas de modificación-restricción o mecanismos de reparación del ADN, los cuales causan determinados patrones en la secuencia que pueden ser identificados como marcadores moleculares (Karlin & Burge, 1995; Karlin *et al.*, 1997; Deschavanne *et al.*, 1999). En trabajos previos, Karlin *et al.* (1997) reportaron la presencia de dinucleótidos (repeticiones de mononucleótidos en la secuencia de ADN, por ejemplo AA, GG o CC) como marcadores moleculares o, como los denominaron, una "firma genética". En este sentido, los genomas víricos pueden ser modificados por la acción de mecanismos de regulación de la célula a la que parasitan, algunos de los cuales ya fueron mencionados anteriormente. Sería este efecto el que podría generar de modo reiterativo secuencias que compartan los patrones de distribución de estos dinucleótidos, elementos que servirían en ese caso para el proceso de agrupación (y servirían también como marcadores moleculares). Esta propiedad de los genomas ha sido utilizada en diversos estudios para la identificación de grupos de secuencias con cierto grado de similitud y para el establecimiento de relaciones filogenéticas (Nakashima *et al.*, 1998; Sandberg *et al.*, 2001; Teeling *et al.*, 2004; Abe *et al.*, 2005; Dalevi *et al.*, 2006). Un estudio interesante para la clasificación en metagenómica viral ha permitido el desarrollo de herramientas informáticas como ProVIDE, con la que se ha podido procesar entre un 3 y un 19% de las secuencias que permanecerían sin clasificar utilizando otras aplicaciones, como MEGAN (Shankar Ghosh *et al.*, 2011).

PERSPECTIVAS

En los últimos diez años se ha producido un importante esfuerzo por comprender el papel de los virus en los ecosistemas acuáticos. Un indicativo de esto son los numerosos trabajos desarrollados cuyo objetivo específico fue comparar diferentes variables de la ecología vírica en ecosistemas acuáticos de agua dulce y salada (Clasen *et al.*, 2008; Danovaro *et al.*, 2008). En este sentido, una de las conclusiones que se ha podido establecer es que es necesario particularizar en las condiciones ecológicas de cada ambiente para poder definir las relaciones complejas entre diferentes formas de materia orgánica y los grupos de bacterias presentes. Otra cuestión que es posible extrapolar en cuanto a la dinámica del funcionamiento de las comunidades víricas es que la mayoría de las investigaciones indican una gran complejidad y variabilidad de las interacciones de estos microorganismos en

los ecosistemas (Middelboe *et al.*, 2008). Desde este particular punto de vista, las herramientas metagenómicas son idóneas para el análisis de los virus.

En sus inicios, la aplicación de la metagenómica se enfocaba al estudio de biomoléculas de interés, fundamentalmente proteínas en comunidades bacterianas, lo que también se ha denominado metagenómica funcional (Uchiyama & Miyazaki, 2009). Sin embargo, el rápido desarrollo de nuevos métodos en el aislamiento de ácidos nucleicos, estrategias de clonación y técnicas de rastreo han permitido la exploración y análisis de las secuencias de las comunidades microbianas, lo que a su vez ha facilitado su comparación con base en criterios taxonómicos y metabólicos (Simon & Daniel, 2011). Esto está siendo utilizado para investigar no sólo la diversidad en cuanto a especies, sino incluso la propia dinámica de los ecosistemas (Vega Thurber *et al.*, 2008).

Una de las perspectivas futuras más importantes para la metagenómica (en general, y la metavirómica en particular), será el desarrollo de programas de cómputo que permitan detectar bajos niveles de similitud entre secuencias, los cuales serán de gran ayuda para el análisis de este tipo de datos y el establecimiento de posibles relaciones entre organismos (Altschul *et al.*, 1997; Zhang *et al.*, 1998). Adicionalmente, el desarrollo de nuevas matrices BLOSUM (*Blocks of Amino Acid Substitution Matrix*) (Henikoff & Henikoff, 1992) permitiría medir con mayor precisión la distancia entre secuencias de proteínas relacionadas, lo que podría ser utilizado para la detección de virus con mayor grado de divergencia (Delwart, 2007).

Es importante tener en cuenta que la enorme cantidad de datos producidos en los estudios de metagenómica viral puede trascender el propio marco disciplinario de la ecología microbiana, lo que en algún momento podría hacer necesario plantear nuevas hipótesis que permitan la inclusión de los datos metagenómicos en disciplinas con un marco conceptual cercano, como un mecanismo de aproximación complementario (Ward, 2006). Un ejemplo de este tipo de aproximaciones se podría dar en el ámbito sanitario, en el que muchos de los patógenos de naturaleza vírica permanecen aún sin caracterizar. Los estudios de metavirómica podrían ayudar en dicha caracterización y, lo que es más importante, podrían ayudar en la prevención de patógenos emergentes (Hendrix *et al.*, 2000; Fei Fan *et al.*, 2011a).

Finalmente, otra perspectiva referida a la metagenómica viral afectará a los modelos matemáticos aplicables a los estudios que manejen datos de metagenomas de poblaciones y dinámica de comunidades (Edwards & Rohwer, 2005). El desarrollo de dichos modelos mejorará la obtención de conclusiones de tipo ecológico a partir del análisis de los datos experimentales obtenidos.

REFERENCIAS

ABE, T., H. SUGAWARA, M. KINOCHI, S. KANAYA & T. IKEMURA. 2005. Novel phylogenetic studies of genomic sequence fragments derived from

uncultured microbe mixtures in environmental and clinical samples. *DNA Research* 12: 281-290.

ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHÄFFER, J. ZHANG, Z. ZHANG, W. MILLER & D. J. LIPMAN. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389-3402.

ANGLY, F. E., B. FELTS, M. BREITBART, P. SALAMON, R. A. EDWARDS, C. CARLSON, A. M. CHAN, M. HAYNES, S. KELLEY, H. LIU, J. M. MAHAFFY, J. E. MUELLER, J. NULTON, R. OLSON, R. PARSONS, S. RAYHAWK, C. A. SUTTLE & F. ROHWER. 2006. The marine viromes of four oceanic regions. *PLoS Biology* 4 (11): e368. doi:10.1371/journal.pbio.0040368.

AUGUET, J. C., H. MONTANIÉ, D. DELMAS, H. J. HARTMANN & V. HUET. 2005. Dynamic of virioplankton abundance and its environmental control in the Charente estuary (France). *Microbial Ecology* 50 (3): 337-349.

BETTAREL, Y., A. DESNUES & E. ROCHELLE-NEWALL. 2010. Lytic-failure in cross inoculation assays between phages and prokaryotes from three aquatic sites of contrasting salinity. *FEMS Microbiology Letters* 311: 113-118.

BLINKOVA, O., J. VICTORIA, Y. LI, B. F. KEELE, C. SANZ, J.-BN. NDJANGO, M. PEETERS, D. TRAVIS, E. V., M. L. LONSDORF, A. E. WILSON, B. H. PUSEY & E. L. DELWART. 2010. Novel circular DNA viruses in stool samples of wild-living chimpanzees. *Journal of General Virology* 91: 74-86.

BOEHME, J., E. FRISCHER, S. C. JIANG, C. A. KELLOGG, S. PICHARD, J. B. ROSE, C. STEINWAY & J. H. PAUL. 1993. Viruses, bacterioplankton, and phytoplankton in the southeastern Gulf of Mexico: distribution and contribution to oceanic DNA pools. *Marine Ecology Progress Series* 97: 1-10.

BØRSHEIM, K. Y. 1993. Native marine bacteriophages. *FEMS Microbiology Ecology* 102: 141-159.

BOUJELBEN, I., P. YARZA, C. ALMANSA, J. VILLAMOR, S. MAALEJ, J. ANTÓN & F. SANTOS. 2012. Virioplankton community structure in Tunisian solar salterns. *Applied and Environmental Microbiology* 78 (20): 7429-7437.

BRATBAK, G., M. HELDAL, S. NORLAND, & T. F. THINGSTAD. 1990. Viruses as partners in spring bloom microbial trophodynamics. *Applied and Environmental Microbiology* 56: 1400-1405.

BRATBAK, G., M. HELDAL, T. F. THINGSTAD & P. TUOMI. 1996. Dynamics of virus abundance in coastal seawater. *FEMS Microbiology Ecology* 19 (4): 263-269.

BREITBART, M., P. SALAMON, B. ANDRESEN, J. M. MAHAFFY, A. M. SEGALL, A. MEAD & D. ROHWER. 2002. Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences USA* 99 (22): 14250-14255.

BREITBART, M., B. FELTS, S. KELLEY, J. M. MAHAFFY, J. NULTON, P. SALAMON & F. ROHWER. 2004. Diversity and population structure of a near-shore marine-sediment viral community. *Proceedings of the Royal Society of London Series B-Biological Sciences* 271: 565-574.

- BRUSSAARD, C. P. D. 2004. Viral control of phytoplankton populations, a review. *The Journal of Eukaryotic Microbiology* 51: 125-138.
- CANTALUPO, P. G., B. CALGUA, G. ZHAO, A. HUNDESA, A. D. WIER, J. P. KATZ, M. GRABE, R. W. HENDRIX, R. GIRONES, D. WANG, & J. M. PIPAS. 2011. Raw sewage harbors diverse viral populations. *mBio* 2 (5): e00180-11.
- CLASEN, J. L., S. M. BRIGDEN, J. P. PAYET & C. A. SUTTLE. 2008. Evidence that viral abundance across oceans and lakes is driven by different biological factors. *Freshwater Biology* 53: 1090-1100.
- COCHLAN, W. P., J. WIKNER, G. F. STEWARD, D. C. SMITH & F. AZAM. 1993. Spatial distribution of viruses, bacteria and chlorophyll *a* in neritic, oceanic and estuarine environments. *Marine Ecology Progress Series* 92: 77-87.
- CULLEY, A. I., A. S. LANG & C. A. SUTTLE. 2006. Metagenomic analysis of coastal RNA virus communities. *Science* 312: 1795-1798.
- CULLEY, A. I., C. A. SUTTLE & G. F. STEWARD. 2010. Characterization of the diversity of marine RNA viruses. In: Wilhelm, W., M. G. Weinbauer & C. A. Suttle (Eds.). *Manual of aquatic viral ecology*. American Society of Limnology and Oceanography, pp. 193-201.
- DALEVI, D., D. DUBHASHI & M. HERMANSSON. 2006. Bayesian classifiers for detecting HGT using fixed and variable order Markov models of genomic signatures. *Bioinformatics* 22: 517-522.
- DANOVARO, R., C. CORINALDESI, M. FILIPPINI, U. R. FISCHER, M. O. GESSNER, S. JA-CQUET, M. MAGAGNINI & B. VELIMIROV. 2008. Viriobenthos in freshwater and marine sediments: a review. *Freshwater Biology* 53: 1186-1213.
- DELWART, E. L. 2007. Viral metagenomics. *Reviews in Medical Virology* 17: 115-131.
- DESCHAVANNE P. J., A. GIRON, J. VILAIN, G. FAGOT & B. FERTIL. 1999. Genomic signature: characterization and classification of species assessed by chaos game representation of sequences. *Molecular Biology and Evolution* 16: 1391-1399.
- DIMMIC, M. W., J. S. REST, D. P. MINDELL & R. A. GOLDSTEIN. 2002. rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. *Journal of Molecular Evolution* 55 (1): 65-73.
- DJIKENG, A., R. KUZMICKAS, N. G. ANDERSON & D. J. SPIRO. 2009. Metagenomic analysis of RNA viruses in a fresh water lake. *PLoS One* 4 (9): e7264. doi:10.1371/journal.pone.0007264.
- DOMINGO E., E. BARANOWSKI, J. I. NUÑEZ, C. M. RUIZ-JARABO, S. SIERRA, N. MOLINA & F. SOBRINO. 2000. Cuasiespecies y evolución molecular de virus. *Revue Scientifique et Technique (International Office of Epizootics)* 19 (1): 55-63.
- DUHAIME, M. B., A. WICHELS, J. WALDMANN, H. TEELING & F. O. GLÖCKNER. 2011. Ecogenomics and genome landscapes of marine *Pseudoalteromonas* phage H105/1 *The ISME Journal* 5: 107-121.
- EDWARDS, R. A. & F. ROHWER. 2005. Viral metagenomics. *Nature Reviews Microbiology* 3 (6): 504-510.
- ERIKSSON, N., L. PACTER, Y. MITSUYA, S. RHEE, C. WANG, B. GHARIZADEH, M. RONAGHI, R. W. SHAFER & N. BEERENWINKEL. 2008. Viral population estimation using pyrosequencing. *PLoS Computational Biology* 4 (5): e1000074.
- FAKHRAI-RAD, H., N. POURMAND & M. RONAGHI. 2002. Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. *Human Mutation* 19: 479-485.
- FARNELL-JACKSON, E. A. & A. K. WARD. 2003. Seasonal patterns of viruses, bacteria and dissolved organic carbon in a riverine wetland. *Freshwater Biology* 48: 841-851.
- FEI FAN, T., C. MANIRE, K. BORROWMAN, T. LANGER, L. EHRHART & M. BREITBART. 2009a. Discovery of a novel single-stranded DNA virus from a sea turtle fibropapilloma by using viral metagenomics. *Journal of Virology* 83: 2500-2509.
- FEI FAN, T., W. K. SUEDEMEYER, E. WHEELER, F. GULLAND & M. BREITBART. 2009b. Novel anellovirus discovered from a mortality event of captive California sea lions. *Journal of General Virology* 90: 1256-1261.
- FEI FAN, T., D. L. WILLNER, Y. W. LIM, R. SCHMIEDER, B. CHAU, C. NILSSON, S. ANTHONY, Y. RUAN, F. ROHWER & M. BREITBART. 2011a. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One* 6(6): e20579. doi:10.1371/journal.pone.0020579.
- FEI FAN, T., E. WHEELER, D. GREIG, T. B. WALTZEK, F. GULLAND & M. BREITBART. 2011b. Metagenomic identification of a novel anellovirus in Pacific harbor seal (*Phoca vitulina richardsii*) lung samples and its detection in samples from multiple years. *Journal of General Virology* 92: 1318-1323.
- FILIPPINI, M. & M. MIDDELBOE. 2007. Viral abundance and genome size distribution in the sediment and water column of marine and freshwater ecosystems. *FEMS Microbiology Ecology* 60: 397-410.
- FILIPPINI, M., N. BUESING, Y. BETTAREL, T. SIME-NGANDO & M. O. GESSNER. 2006. Infection paradox: high abundance but low impact of freshwater benthic viruses. *Applied and Environmental Microbiology* 72: 4893-4898.
- FISCHER, U. R., C. WIELTSCHNIG, A. K. T. KIRSCHNER & B. VELIMIROV. 2003. Does virus-induced lysis contribute to bacterial mortality in the oxygenated sediment layer of shallow oxbow lakes? *Applied and Environmental Microbiology* 69: 5281-5289.
- FREDSLUND, J. 2006. PHY*FI: fast and easy online creation and manipulation of phylogeny color figures *BMC Bioinformatics* 7: 315.
- FUHRMAN, J. A. 1992. Bacterioplankton roles in cycling of organic matter: The microbial food web. In: Falkowski, P. G. & A. D. Woodhead (Eds.). *Primary productivity and biogeochemical cycles in the sea*. Plenum, pp. 361-383.
- FUHRMAN, J. A. 1999. Marine viruses and their biogeochemical and ecological effects. *Nature* 399: 541-548.
- GERSBERG, R. M., R. A. GEARHART & M. IVES. 1989. Pathogen removal in constructed wetlands. In: Hammer D. (Ed.). *Constructed wetlands for*

- wastewater treatment municipal, industrial, and agricultural. Lewis Publishers, pp. 431-445.
- GUIXA-BOIXAREU, N., J. I. CALDERÓN-PAZ, M. HELDAL, G. BRATBAK & C. PEDRÓS-ALIÓ. 1996. Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. *Aquatic Microbial Ecology* 11: 215-227.
- HARA, S., I. KOIKE, K. TERAUCHI, H. KAMIYA & E. TANQUE. 1996. Abundance of viruses in deep oceanic waters. *Marine Ecology Progress Series* 145: 269-277.
- HENDRIX, R. W., J. G. LAWRENCE, G. F. HATFULL & S. CASJENS. 2000. The origins and ongoing evolution of viruses. *Trends in Microbiology* 8: 499-500.
- HENIKOFF, S. & J. G. HENIKOFF. 1992. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences USA* 89(22): 10915-10919.
- HUSE, S., J. HUBER, H. MORRISON, M. SOGIN & D. WELCH. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology* 8: R143.
- HUSON, D. H., A. F. AUCH, J. QI & S. C. SCHUSTER. 2007. MEGAN analysis of metagenomic data. *Genome Research* 17: 377-386.
- JACKSON, E. F. & C. R. JACKSON. 2008. Viruses in wetland ecosystems. *Freshwater Biology* 53: 1214-1227.
- KAPOOR, A., J. VICTORIA, P. SIMMONDS, C. WANG, R. W. SHAFER, R. NIMS, O. NIELSEN & E. DELWART. 2008. A highly divergent picornavirus in a marine mammal. *Journal of Virology* 82 (1): 311-320.
- KARLIN, S. & C. BURGE. 1995. Dinucleotide relative abundance extremes: a genomic signature. *Trends in Genetics* 11: 283-290.
- KARLIN, S., J. MRAZEK & A. M. CAMPBELL. 1997. Compositional biases of bacterial genomes and evolutionary implications. *Journal of Bacteriology* 179: 3899-3913.
- KIM, K. H., H. W. CHANG, Y. D. NAM, S. W. ROH, M. S. KIM, Y. SUNG, C. O. JEON, H. M. OH & J. W. BAE. 2008. Amplification of uncultured single-stranded DNA viruses from rice paddy soil. *Applied and Environmental Microbiology* 74: 5975-5985.
- KIENE, R. P. & T. S. BATES. 1990. Biological removal of dimethyl sulphide from sea water. *Nature* 345: 702-705.
- KOONIN, E. V. & V. V. DOLJA. 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Critical Reviews in Biochemistry and Molecular Biology* 28: 375-430.
- KOONIN, E. V., T. G. SENKEVICH & V. V. DOLJA. 2006. The ancient virus world and evolution of cells. *Biology Direct* 1: 29. doi:10.1186/1745-6150-1-29.
- KORLACH, J., P. J. MARKS, R. L. CICERO, J. J. GRAY, D. L. MURPHY, D. B. ROITMAN, T. T. PHAM, G. A. OTTO, M. FOQUET & S. W. TURNER. 2008. Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proceedings of the National Academy of Sciences USA* 105: 1176-1181.
- KUZMIN, I. V., G. J. HUGHES & C. E. RUPPRECHT. 2006. Phylogenetic relationships of seven previously unclassified viruses within the family Rhabdoviridae using partial nucleoprotein gene sequences. *Journal of General Virology* 87: 2323-2331.
- LI, L., J. G. VICTORIA, C. WANG, M. JONES, G. M. FELLERS, T. H. KUNZ & E. DELWART. 2010. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *Journal of Virology* 84: 6955-6965.
- LINDELL, D. J. D., JAFFE, Z. I. JOHNSON, G. M. CHURCH & S. W. CHISHOLM. 2005. Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* 438 (7064): 86-89.
- LÓPEZ-BUENO, A., J. TAMAMES, D. VELÁZQUEZ, A. MOYA, A. QUESADA & A. ALCA-MÍ. 2009. High diversity of the viral community from an Antarctic lake. *Science* 326: 858-861.
- LORENZI, H. A., J. HOOVER, J. INMAN, T. SAFFORD, S. MURPHY, L. KAGAN & S. J. WILLIAMSON. 2011. The Viral MetaGenome Annotation Pipeline (VMGAP): an automated tool for the functional annotation of viral metagenomic shotgun sequencing data. *Standards in Genomic Sciences* 4: 418-429.
- MALET, I., M. BELNARD, H. AGUT & A. CAHOUR. 2003. From RNA to quasispecies: a DNA polymerase with proofreading activity is highly recommended for accurate assessment of viral diversity. *Journal of Virology Methods* 109: 161-170.
- MARANGER, R., D. F. BIRD & S. K. JUNIPER. 1994. Viral and bacterial dynamics in arctic sea ice during the spring algal bloom near Resolute, NWT, Canada. *Marine Ecology Progress Series* 111: 121-127.
- MARGULIES, M., M. EGHOLM, W. E. ALTMAN, S. ATTIIYA, J. S. BADER, L. A. BEMBEN, J. BERKA, M. S. BRAVERMAN, Y. J. CHEN, Z. CHEN, S. B. DEWELL, L. DU, J. M. FIERRO, X. V. GOMES, B. C. GODWIN, W. HE, S. HELGESEN, C. H. HO, G. P. IRZYK, S. C. JANDO, M. L. ALLENQUER, T. P. JARVIE, K. B. JIRAGE, J. B. KIM, J. R. KNIGHT, J. R. LANZA, J. H. LEAMON, S. M. LEFKOWITZ, M. LEI, J. LI, K. L. LOHMAN, H. LU, V. B. MAKHIJANI, K. E. MCDADE, M. P. MCKENNA, E. W. MYERS, E. NICKERSON, J. R. NOBILE, R. PLANT, B. P. PUC, M. T. RONAN, G. T. ROTH, G. J. SARKIS, J. F. SIMONS, J. W. SIMPSON, M. SRINIVASAN, K. R. TARTARO, A. TOMASZ, K. A. VOGT, G. A. VOLKMER, S. H. WANG, Y. WANG, M. P. WEINER, P. YU, R. F. BEGLEY & J. M. ROTHBERG. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380.
- MARHAVER, K. L., R. A. EDWARDS & F. ROHWER. 2008. Viral communities associated with healthy and bleaching corals. *Environmental Microbiology* 10: 2277-2286.
- MARIE, D., C. BRUSSAARD, R. THYRHAUG, G. BRATBAK & D. VAULOT. 1999. Enumeration of marine viruses in culture and natural samples by flow cytometry. *Applied and Environmental Microbiology* 65: 45-52.
- MATHIAS, C. B., A. K. T. KIRSCHNER & B. VELIMIROV. 1995. Seasonal variations of viral abundance and viral control of bacterial production in a

- backwater system of the Danube River. *Applied and Environmental Microbiology* 61: 3734-3740.
- McDANIEL, L., M. BREITBART, J. MOBBERLEY, A. LONG, M. HAYNES, F. ROHWER & J. H. PAUL. 2008. Metagenomic analysis of lysogeny in Tampa Bay: implications for prophage gene expression. *PLoS One* 3(9): e3263.
- MEDINA, R.A. & A. GARCÍA-SASTRE. 2011. Influenza A viruses: new research developments. *Nature Reviews Microbiology* 9: 590-603.
- MIDDELBOE, M., R.N. GLUD & K. FINSTER. 2003. Distribution of viruses and bacteria in relation to diagenetic activity in an estuarine sediment. *Limnology and Oceanography* 48: 1447-1456.
- MIDDELBOE, M., S. JACQUET & M. WEINBAUER. 2008. Viruses in freshwater ecosystems: an introduction to the exploration of viruses in new aquatic habitats. *Freshwater Biology* 53: 1069-1075.
- MUNN, C. B. 2006. Viruses as pathogens of marine organisms from bacteria to whales. *Journal of the Marine Biological Association of the United Kingdom* 86: 453-467.
- NAKASHIMA, H., M. OTA, K. NISHIKAWA, & T. OOI. 1998. Genes from nine genomes are separated into their organisms in the dinucleotide composition space. *DNA Research* 5: 251-259.
- NOBLE, R. T. & J. A. FUHRMAN. 1998. Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquatic Microbial Ecology* 14: 113-118.
- NOGUCHI, H., T. TANIGUCHI & T. ITOH. 2008. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Research* 15: 387-396.
- NOKES, R. L., C. P. GERBA & M. M. KARPISCAK. 2003. Microbial water quality improvement by small scale on-site subsurface wetland treatment. *Journal of Environmental Science and Health* 38: 1849-1855.
- OGUNSEITAN, O. A. 2008. Genetic transduction in freshwater ecosystems. *Freshwater Biology* 53: 1228-1239.
- OH, S., A. CARO-QUINTERO, D. TSEMENTZI, N. DE LEON-RODRIGUEZ, C. LUO, R. PORRETSKY & K. T. KONSTANTINIDIS. 2011. Metagenomic insights into the evolution, function, and complexity of the planktonic microbial community of Lake Lanier, a temperate freshwater ecosystem. *Applied and Environmental Microbiology* 77 (17): 6000-6011.
- PARK, E. J., K. H. KIM, G. C. ABELL, M. S. KIM, S. W. ROH & J. W. BAE. 2011. Metagenomic analysis of the viral communities in fermented foods. *Applied and Environmental Microbiology* 77: 1284-1291.
- PARSLEY, L. C., E. J. CONSUEGRA, S. J. THOMAS, J. BHAVSAR, A. M. LAND, N. N. BHUIYAN, M. A. MAZHER, R. J. WATERS, K. E. WOMMACK, W. F. HARPER, JR. & M. R. LILES. 2010. Census of the viral metagenome within an activated sludge microbial assemblage. *Applied and Environmental Microbiology* 76 (8): 2673-2677.
- PAUL, J. H., J. B. ROSE, S. C. JIANG, C. A. KELLOGG & L. DICKSON. 1993. Distribution of viral abundance in the reef environment of Key Largo, Florida. *Applied and Environmental Microbiology* 59: 718-724.
- PEDUZZI, P. & F. SCHIEMER. 2004. Bacteria and viruses in the water column of tropical freshwater reservoirs. *Environmental Microbiology* 6: 707-715.
- PROCTOR, L. M. & J. A. FUHRMAN. 1990. Viral mortality of marine bacteria and cyanobacteria. *Nature* 343: 60-62.
- PROCTOR, L. M. & J. A. FUHRMAN. 1991. Roles of viral infection in organic particle flux. *Marine Ecology Progress Series* 69: 133-142.
- PURCELL, S. J. D. MARK & P. C. SHAM. 2007. WHAP: haplotype-based association analysis. *Bioinformatics Applications Note* 23 (2): 255-256.
- REBEKAH, R. H. & K. E. WOMMACK. 2009. Seasonal dynamics and metagenomic characterization of estuarine virobenthos assemblages by Randomly Amplified Polymorphic DNA PCR. *Applied and Environmental Microbiology* 75 (8): 2259-2265.
- RODRIGUEZ-BRITO, B., L. LI, L. WEGLEY, M. FURLAN, F. ANGLY, M. BREITBART, J. BUCHANAN, C. DESNUES, E. DINSDALE, R. EDWARDS, B. FELTS, M. HAYNES, H. LIU, D. LIPSON, J. MAHAFFY, A. B. MARTIN-CUADRADO, A. MIRA, J. NULTON, L. PASIĆ, S. RAYHAWK, J. RODRIGUEZ-MUELLER, F. RODRIGUEZ-VALERA, P. SALAMON, S. SRINAGESH, T. F. THINGSTAD, T. TRAN, R. V. THURBER, D. WILLNER, M. YOULE & F. ROHWER. 2010. Viral and microbial community dynamics in four aquatic environments. *ISME Journal* 4: 739-751.
- ROHWER, F. & R. EDWARDS. 2002. The phage proteomic tree: a genome-based taxonomy for phage. *Journal of Bacteriology* 184: 4529-4535.
- ROSARIO, C., K. NILSSON, Y. W. LIM, R. YIJUN & M. BREITBART. 2009. Metagenomic analysis of viruses in reclaimed water. *Environmental Microbiology* 11: 2806-2820.
- ROUX, S., F. ENAULT, A. ROBIN, V. RAVET, S. PERSONNIC, S. THEIL, J. COLOMBET, T. SIME-NGANDO & D. DEBROAS. 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS ONE* 7 (3): e33641.
- SANDAA, R., M. CLOKIE & N.H. MANN. 2008. Photosynthetic genes in viral populations with a large genomic size range from Norwegian coastal waters. *FEMS Microbiology Ecology* 63: 2-11.
- SANDBERG, R., G. WINBERG, C. I. BRANDEN, A. KASKE, I. ERNBERG & J. COSTER. 2001. Capturing whole-genome characteristics in short sequences using a naive Bayesian classifier. *Genome Research* 11: 1404-1409.
- SANGER, F. & A. R. COULSON. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* 94: 441-448.
- SANGER, F., S. NICKLEN & A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* 74: 5463-5467.
- SANTOS, F., A. MEYERDIERKS, A. PENNA, R. ROSSELLO-MORA, RAMANN & J. ANTON. 2007. Metagenomic approach to the study of halophages: the environmental halophage. *Environmental Microbiology* 9: 1711-1723.

- SANTOS, F., P. YARZA, V. PARRO, C. BRIONES & J. ANTON. 2010. The metavirome of a hypersaline environment. *Applied and Environmental Microbiology* 12: 2965-2976.
- SCHAPIRA, M., M. J. BUSCOT, S. C. LETERME, T. POLLET, C. CHAPPERON & L. SEURONT. 2009. Distribution of heterotrophic bacteria and virus-like particles along a salinity gradient in a hypersaline coastal lagoon. *Aquatic Microbial Ecology* 54: 171-183.
- SCHATZ, M. C., A. L. DELCHER & S. L. SALZBERG. 2010. Assembly of large genomes using second-generation sequencing. *Genome Research* 20: 1165-1173.
- SCHOENFELD, T., M. PATTERSON, P. M. RICHARDSON, K. E. WOMMACK, M. YOUNG & D. MEAD. 2008. Assembly of viral metagenomes from Yellowstone hot springs. *Applied and Environmental Microbiology* 74: 4164-4174.
- SHANKAR GHOSH, T., M. H. MOHAMMED, D. KOMANDURI, & S. S. MANDE. 2011. ProViDE: A software tool for accurate estimation of viral diversity in metagenomic samples. *Bioinformatics* 6 (2): 91-94.
- SHARON, I., A. ALPEROVITCH, F. ROHWER, M. HAYNES, F. GLASER, N. ATAMNA-IS-MAEEL, R. Y. PINTER, F. PARTENSKY, E. V. KOONIN, Y. I. WOLF, N. NELSON & O. BÉJÀ. 2009. Photosystem I gene cassettes are present in marine virus genomes. *Nature* 461: 258-262.
- SIEBURTH, J. M., P. W. JOHNSON & P. E. HARGRAVES. 1988. Ultrastructure and ecology of *Aureococcus anophagefferens* gen. et sp. nov. (Chrysophyceae): the dominant picoplankton during a bloom in Narragansett Bay, Rhode Island, Summer 1985. *Journal of Phycology* 24: 416-425.
- SIME-NGANDO, T., S. LUCAS, A. ROBIN, K. P. TUCKER, J. COLOMBET, Y. BETTAREL, E. DESMOND, S. GRIBALDO, P. FORTERRE, M. BREITBART & D. PRANGISHVILI. 2010. Diversity of virus-host systems in hypersaline Lake Retba, Senegal. *Environmental Microbiology* 13 (8): 1956-1972.
- SIMON, C. & R. DANIEL. 2011. Metagenomic analyses: past and future trends. *Applied and Environmental Microbiology* 77 (4): 1153-1161.
- SLOVÁČKOVÁ, H. & B. MARŠÁLEK. 2008. Virioplankton and microbial communities in two Czech rivers (Svratka and Morava River). *Aquatic Sciences* 70 (3): 282-291.
- STEWART, G. F. & C. M. PRESTON. 2011. Analysis of a viral metagenomic library from 200m depth in Monterey Bay, California, constructed by direct shotgun cloning. *Virology Journal* 8: 287-301.
- SUTTLE, C. A., A. M. CHAN & M. T. COTTRELL. 1990. Infection of phytoplankton by viruses and reduction of primary productivity. *Nature* 347: 467-469.
- SUTTLE, C. A. 2005. Viruses in the sea. *Nature* 437: 356-361.
- SUTTLE, C. A. 2007. Marine viruses; major players in the global ecosystem. *Nature Reviews Microbiology* 5: 801-812.
- TANENBAUM, D. M., J. GOLL, S. MURPHY, P. KUMAR, N. ZA-FAR, M. THIAGARAJAN, R. MADUPU, T. DAVIDSEN, L. KAGAN & S. KRAVITZ. 2010. Metagenomic shotgun sequencing data. *Standards in Genomics Sciences* 2: 229-237.
- TEELING, H., A. MEYERDIERKS, M. BAUER, R. AMANN, & F. O. GLOCKNER. 2004. Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environmental Microbiology* 6: 938-947.
- UCHIYAMA, T. & K. MIYAZAKI. 2009. Functional metagenomics for enzyme discovery: challenges to efficient screening. *Current Opinion in Biotechnology* 20 (6): 616-622.
- VEGA THURBER, R. L., K. L. BAROTT, D. HALL, H. LIU, B. RODRIGUEZ-MUELLER, C. DESNUES, R. A. EDWARDS, M. HAYNES, F. E. ANGLY, L. WEGLEY & F. L. ROHWER. 2008. Metagenomic analysis indicates that stressors induce production of herpes-like viruses in the coral *Porites compressa*. *Proceedings of the National Academy of Sciences USA* 105: 18413-18418.
- VENTER, J. C., K. REMINGTON, J. F. HEIDELBERG, A. L. HALPERN, D. RUSCH, J. A. EISEN, D. W. I. PAULSEN, K. E. NELSON, W. NELSON, D. E. FOUTS, S. LEVY, A. H. KNAP, M. L. LOMAS, K. NEALSON, O. WHITE, J. PETERSON, J. HOFFMAN, R. PARSONS, H. BADEN-TILLSON, C. PFANNKUCH, Y. H. ROGERS & H. O. SMITH. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304 (5667): 66-74.
- VIDALES, J. A., C. P. GERBA & M. M. KARPISCAK. 2003. Virus removal from wastewater in a multispecies subsurface-flow constructed wetland water. *Environment Research* 75 (3): 238-245.
- VYMAZAL, J. 2005. Removal of enteric bacteria in constructed treatment wetlands with emergent macrophytes: a review. *Journal of Environmental Science and Health* 40: 1355-1367.
- WARD, N. 2006. New directions and interactions in metagenomics research. *FEMS Microbiology Ecology* 55: 331-338.
- WEINBAUER, M. G. 2004. Ecology of prokaryotic viruses. *FEMS Microbiology Reviews* 28: 127-181.
- WEINBAUER, M. G. & F. RASSOULZADEGAN. 2004. Are viruses driving microbial diversification and diversity? *Environmental Microbiology* 6: 1-11.
- WILHELM, S. W. & C. A. SUTTLE. 1999. Viruses and nutrient cycles in the sea. *BioScience* 49: 781-788.
- WILLIAMSON, S. J., D. B. RUSCH, S. Y. AARON, L. HALPERN, K. B. HEIDELBERG, J. I. GLASS, C. ANDREWS-PFANNKUCH, D. FADROSH, C. S. MILLER, G. SUTTON, M. FRAZIER & J. C. VENTER. 2008. The Sorcerer II global ocean sampling expedition: metagenomic characterization of viruses within aquatic microbial samples. *PLoS ONE* 3 (1): e1456. doi:10.1371/journal.pone.0001456.
- WILLIAMSON, S. J., L. ZEIGLER ALLEN, H. A. LORENZI, D. W. FADROSH, D. B. MATHANGI THIAGARAJAN, J. P. MCCROW, A. TOVCHIGRECHKO, S. YOOSEPH & J. C. VENTER. 2012. Metagenomic exploration of viruses throughout the Indian Ocean. *PLoS ONE* 7 (10): e42047. doi:10.1371/journal.pone.0042047.
- WILLNER, D. M., M. FURLAN, R. HAYNES, F. E. SCHMIEDER, J. ANGLY, S. SILVA, TAMMADONI, B. NOSRAT, D. CONRAD & F. ROHWER. 2009. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. *PLoS ONE* 4 (10): e7370. doi:10.1371/journal.pone.0007370.

- WOMMACK, K. E. & R. R. COLWELL. 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiology and Molecular Biology Reviews* 64: 69-114.
- WOMMACK, K. E., R. T. HILL, M. KESSEL, E. RUSSEK-COHEN, & R. R. COLWELL. 1992. Distribution of viruses in the Chesapeake Bay. *Applied and Environmental Microbiology* 58: 2965-2970.
- WOMMACK, K. E., S. R. BENCH, J. BHAVSAR, D. MEAD & T. HANSON. 2009. Isolation independent methods of characterizing phage communities 2: characterizing a metagenome. *Methods in Molecular Biology* 502: 279-289.
- YAU, S., F. M. LAURO, M. Z. DE MAERE, M. V. BROWN, T. THOMASA, M. J. RAFTERY, C. A. PFANNKUCH, M. LEWIS, J. M. HOFFMAN, J. A. GIBSON & R. CAVICCHIOLIA. 2011. Virophage control of antarctic algal host-virus dynamics. *Proceedings of the National Academy of Sciences USA* 108 (15): 6163-6168.
- ZHANG, Z., A. A. SCHÄFFER, W. MILLER, T. L. MADDEN, D. J. LIPMAN, E. V. KOONIN & S. F. ALTSCHUL. 1998. Protein sequence similarity searches using patterns as seeds. *Nucleic Acids Research* 26: 3986-3990.
- Recibido:* 5 de mayo de 2013.
- Aceptado:* 15 de octubre del 2013.