



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Programa de Posgrado en Ciencias de la Producción y de la Salud Animal

**VIROMA DE PORCINOS DE TRASPATIO Y QUIRÓPTEROS NO
HEMATÓFAGOS**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA

Rodrigo Jesús Barrón Rodríguez

Tutora : Elizabeth Loza Rubio

INIFAP

Comité Tutor

Gary García Espinosa

Humberto Ramírez Mendoza

FMVZ - UNAM

Ciudad de México, SEPTIEMBRE 2021



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

DEDICATORIAS

Mi trabajo lo ofrezco a Dios por mi vida y mi salud, porque así puedo hacer, experimentar y disfrutar este mundo tan maravilloso. En el que se nos da lo necesario y más, y en el cual he aprendido y sigo aprendiendo a ser mejor persona.

A mi manada, Vero, mi Diego y mi Patricio, mi familia (MI GRAN TESORO). Porque de una u otra manera siempre me inspiran, me motivan y me impulsan para ser mejor, superarme, lograr mis metas y me ayudan a ser la mejor versión de mí; por siempre apoyarme en mis decisiones y locuras; y por amarme.

A mis padres por permitirme estar en esta vida y contribuir, en mucho, a hacer de mí el ser y profesionalista que soy; por darme un hogar ideal, y formarme en una familia amorosa. Por enseñarme a lograr mis metas esforzándome y haciendo las cosas lo mejor posible, de manera honesta y con empatía para con los demás. Los amo.

A mis hermanos Leslie y Julio por estar ahí siempre, por su amor y apoyo incondicionales. Por todo lo que hacen por mí.

A Jairo Betancourt, por su amistad, consejos y apoyo.

A mi amigo Said por su ayuda y apoyo.

Y a todos los familiares, amigos, profesores que no menciono de forma particular, pero que han contribuido a mi formación y me han apoyado.

AGRADECIMIENTOS

A mi tutora, Elizabeth Loza Rubio por su apoyo y enseñanzas durante más de 4 años. Y por impulsarme a realizar esta investigación.

A la doctora Edith Rojas Anaya, una gran amiga, por su ayuda y aportaciones.

Al Instituto Nacional de Investigaciones forestales agrícolas y pecuarias por abrirnos las puertas de su laboratorio y por el financiamiento de este proyecto.

Al CONACYT por otorgarme una beca durante mis estudios.

A Omar Ríos Bello, Omar Maldonado, y José López por su gran apoyo y contribución en la logística y las actividades de campo de este trabajo.

A mis tutores Gary García y Humberto Ramírez por abrirme las puertas de sus laboratorios, aportar ideas y consejos, por todo su apoyo y su ayuda.

Agradezco en especial al p.M.V.Z. Julio J. Barrón Rodríguez, mi hermano, y a Francisco A. Rosas Rodríguez, mi primo, por haberme apoyado en el trabajo de campo.

A Dr. Joel A. Pérez Vázquez, Dr. Iván Romero Espinoza y al laboratorio de virología del INER por su colaboración en la realización de este trabajo.

Al Dr. Jorge T. Ayala Sumuano por sus enseñanzas y ayuda en el trabajo bioinformático de este estudio.

RESUMEN

Enfermedades emergentes y reemergentes a menudo se asocian con entornos y actividades humanas como la agricultura y la ganadería. Una de las especies de ganado más importantes para la alimentación y la economía es el cerdo doméstico, el cual se ha introducido y criado en una amplia variedad de entornos y sistemas, incluso en libre pastoreo. Existen reportes de virus de murciélagos que han infectado a los porcinos causando importantes enfermedades como la rabia, Nipah, Menangle, Ebola y Mapuera, adaptándose y afectando también a humanos. Este trabajo obtuvo el viroma de cerdos de traspatio rural y murciélagos (*Artibeus* spp, *Macrotus waterhousii* y *Pteronotus parnellii*) en entornos superpuestos en Morelos, México. A partir de hisopos nasales o faríngeos y rectales, se utilizó secuenciación masiva y bioinformática para conocer las comunidades virales de cada especie animal. La mayor proporción de las secuencias virales obtenidas en las cuatro especies (91.1 - 98 %) estaban relacionadas con fagos, y una baja proporción (2 - 8.9%) estaba asociada a virus animales. Se encontraron secuencias de circovirus en las cuatro especies, y se ensambló un genoma parcial a partir de los datos de los porcinos. También se encontraron secuencias de saporovirus, bocavirus, astrovirus y otros virus de DNA circular en los cerdos. Secuencias de la familia *Coronaviridae* se encontraron en *P. parnellii*. Este estudio describe el viroma de cerdos sanos y tres especies de murciélagos no estudiados anteriormente, que conviven en el traspatio rural de México, encontrando una presencia abundante de bacteriófagos, mayor que lo informado anteriormente; una menor diversidad de virus animales diferentes a los reportes anteriores; y evidencia de virus no reportados previamente en el país. Los hallazgos de este estudio contribuyen con una valiosa pieza más del viroma de estas especies, proporcionando información de referencia sobre sus comunidades virales.

Palabras clave: bacteriófagos, comunidad viral, secuenciación de próxima generación, cerdos, *Phyllostomidae*, *Mormopidae*.

ABSTRACT

Emerging and reemerging diseases are often associated with human environments and activities such as agriculture and livestock. One of the most important livestock species for food and economy is the domestic swine, which has been introduced and raised in a wide variety of environments and systems, including free-roaming. There are reports of bat viruses that have infected pigs causing important diseases such as rabies, Nipah, Menangle, Ebola and Mapuera, also adapting and affecting humans. This study obtained the virome of rural-backyard swine and bats (*Artibeus* spp, *Macrotus waterhousii* and *Pteronotus parnellii*) with overlapping environments in Morelos, Mexico. Starting from nasal or pharyngeal and rectal swabs, next-generation sequencing and bioinformatic analysis were used to know the viral communities of each animal species. The major proportion of the viral sequences obtained from the four species (91.1 - 98 %) were related to phages, and a low proportion (2 – 8.9%) was associated to animal viruses. Circoviruses sequences were found in the four species, and a partial genome was assembled from the swine data. Also sequences of saporovirus, bocavirus, astrovirus and other circulad DNA viruses were found in swine. Sequences of *Coronaviridae* family were found in *P. parnellii*. This study describes the virome of healthy swine and three species of bats non studied before, that cohabit in rural backyard from Mexico, finding an abundant presence of bacteriophages, higher than previous reported (especially in young pigs); a less diversity of animal viruses different from those before reported; and evidence of viruses non reported previously in the country. Findings of this study contribute with one more valuable piece in the virome of these species, providing baseline information about their viral communities.

Keywords: bacteriophages, viral community, next generation sequencing, pigs, Phylostomidae, Mormopidae

CONTENIDO (ÍNDICE)

<i>DEDICATORIAS</i>	I
AGRADECIMIENTOS.....	II
RESUMEN.....	III
ABSTRACT.....	IV
<i>INTRODUCCIÓN</i>	1
El cerdo doméstico, infecciones ínter-especie y zoonosis.....	1
Agentes infecciosos emergentes y quirópteros.....	3
Adaptaciones virales interespecies murciélagos – porcinos en ambientes perturbados por el hombre.....	6
La secuenciación de siguiente generación: una herramienta auxiliar para el estudio de las comunidades de agentes infecciosos.....	8
Justificación.....	11
HIPÓTESIS.....	12
OBJETIVO GENERAL.....	12
MATERIALES Y MÉTODOS.....	14
Área de estudio y muestras.....	14
Muestras biológicas.....	16
Obtención de ácidos nucleicos (RNA/DNA).....	18
Construcción de las librerías.....	18
Secuenciación.....	20
Análisis Bioinformático.....	21
Análisis estadísticos y diversidad.....	27
Identificación y confirmación de virus por PCR/RT-PCR.....	27
Filogenias.....	29
<i>RESULTADOS</i>	31
Datos de secuenciación.....	31
Virus identificados por secuenciación masiva.....	31
Ensamblajes y análisis filogenético.....	37
<i>DISCUSIÓN</i>	49
<i>CONCLUSIONES</i>	61
<i>REFERENCIAS</i>	62
<i>ANEXOS</i>	77
ARTÍCULOS.....	102

Lista de cuadros

Cuadro1. Frecuencias de identificación de algunos agentes virales en quirópteros.....	5
Cuadro2. Información de las muestras y sitios de muestreo.....	17
Cuadro 3. Información de lo iniciadores (primers) utilizados para la identificación/confirmación de virus animales mediante ensayos de PCR o RT-PCR.....	30
Cuadro 4. Virus más abundantes identificados por análisis de homología (BLASTn y BLASTx) de los conjuntos de datos (lecturas) de cada especie estudiada.	34
Cuadro 5. Contigs (secuencias) de los virus más abundantes en porcinos y murciélagos.....	39

Lista de Figuras

Figura 1. Ubicación geográfica de refugios de murciélagos (B) y granjas porcinas de estudio.....	15
Figura 2. Resumen de la clasificación taxonómica de las secuencias obtenidas por secuenciación masiva a partir de muestras de las tres especies de murciélagos y los porcinos de traspatio.....	33
Figura 3. Diversidad de especies de virus estimada por rarefacción para los datos de cada especie de murciélago y porcinos.....	36
Figura 4. Árbol filogenético basado en secuencias nucleotídicas del gen de la proteína-1 (cápside) de especies del género <i>Circovirus</i> para mostrar la relación del contig de los porcinos estudiados.....	40
Figura 5. Árbol filogenético basado en el gen Cap de especies de la familia <i>Circoviridae</i> (géneros <i>Circovirus</i> y <i>Ciclovirus</i>) para mostrar la relación del contig de los porcinos estudiados.....	41
Figura 6. Árbol filogenético basado en secuencias de nucleótidos del gen Cap de la familia <i>Caliciviridae</i> para mostrar la relación del contig de los porcinos estudiados.....	42
Figura 7. Árbol filogenético basado en el gen NS1 de las especies de la subfamilia <i>Parvovirinae</i> para mostrar la relación del contig de los porcinos estudiados.....	43
Figura 8. Filogenia basada en el gen de la replicasa (Rep) del género <i>Circovirus</i> que muestra la relación del contig de <i>Artibeus</i>	44
Figura 9. Filogenia basada en el gen de la replicasa (Rep) del género <i>Circovirus</i> que muestra la relación del contig de <i>M. waterhousii</i>	45
Figura 10. Filogenia basada en el gen de la replicasa (Rep) del género <i>Circovirus</i> que muestra la relación del contig de <i>P. parnellii</i>	46
Figura 11. Filogenia basada en el gen de la RNA polimerasa dependiente de RNA (ORF1ab) de especies del género <i>Coronavirus</i> que muestra la relación del contig de los murciélagos <i>P. parnellii</i> estudiados.....	47
Figura 12. Filogenia basada en el gen de la replicasa (Rep) de especies del género <i>Circovirus</i> mostrando la relación de los contigs encontrados en los cerdos y murciélagos estudiados.....	48

INTRODUCCIÓN.

Enfermedades emergentes y reemergentes a menudo se asocian con entornos y actividades humanas que generan cambios drásticos en el uso del suelo, como la urbanización, la agricultura y la ganadería (FAO, 2016; Lambin y col., 2001). En estos escenarios se establecen condiciones en las que algunos microorganismos sufren adaptaciones que confieren la capacidad de producir enfermedades, (como los Alfairinfluzavirus) o se generan interacciones entre especies (incluidos los seres humanos) y agentes infecciosos que en la naturaleza nunca ocurrirían, permitiendo la adaptación de los microorganismos a nuevos hospederos (Patz y col., 2004). Como ejemplos de estas adaptaciones están el virus Nipah, el virus Hendra y los coronavirus similares a los relacionados con los síndromes respiratorios agudo severo (SARSr-CoV) y del Medio oriente (MERSr-CoV) (Chua, 2012; Drosten y col., 2003).

Lo anterior se ve reflejado en un incremento en la frecuencia actual de las enfermedades infecciosas emergentes y re-emergentes, que son favorecidas por factores que permiten una más rápida diseminación de agentes y que facilitan la exposición a los mismos (Cohen, 1998).

El cerdo doméstico, infecciones íter-especie y zoonosis

El cerdo doméstico (*Sus scrofa domesticus*) es una de las principales especies de ganado, debido a su adaptabilidad, y su importancia en la alimentación y la economía humanas. Esta especie ha sido introducida por el ser humano en una amplia variedad de ambientes y criado en diferentes tipos de sistemas de producción, ya sea de ciclo completo, para producción y venta de lechones o enfocado solo en la engorda de animales. Los porcinos pueden ser mantenidos en una variedad de sistemas desde los intensivos tecnificados que mantienen altas densidades poblacionales de grupos etarios con estrecha relación entre individuos (donde se facilita la transmisión, permanencia y mutación de agentes infecciosos),

hasta los productores de traspatio donde los animales pueden ser alojados en diferentes variantes de instalaciones, generalmente más rústicas, con un menor número de individuos, mayor espacio por animal, mayor exposición al ambiente o en libre pastoreo. En el sistema de traspatio no existe una modalidad de producción única y precisa, y con frecuencia los porcinos coexisten en estrecho contacto con otras especies de animales domésticos, humanos e incluso con especies silvestres, estableciendo interfaces de riesgo para la adaptación y la presencia de enfermedades (Wiethoelter y col., 2015).

Las enfermedades emergentes y reemergentes con frecuencia afectan a los porcinos (ej. influenza porcina, diarrea epidémica, gastroenteritis transmisible, circovirus, PRRS, rabia (SAGARPA, 2011) y tomando en cuenta los factores y condiciones de crianza y producción de la especie antes mencionados, son latentes los riesgos de adaptaciones y recombinaciones virales, que resultan en nuevos agentes, incluso zoonóticos, en condiciones modernas que facilitan la transmisión y la diseminación (Cohen, 1998). Como podemos ver, los suinos reúnen varias características interesantes que lo hacen un buen modelo para estudiar las comunidades virales que lo habitan, sus modificaciones e interacciones en los diferentes ambientes de producción.

El ganado porcino se infecta y padece enfermedades comunes a otras especies animales, incluyendo al hombre (zoonosis). En ocasiones el suido se ha desarrollado como hospedero accidental adaptativo de ciertos agentes, que al infectar a ganado porcino, sufren modificaciones que les confieren la capacidad de afectar nuevas especies, provocando enfermedades importantes, como el caso del virus Nipah. Cabe mencionar que los humanos también son capaces de generar patógenos para los animales (antropozoonosis). Este es el caso del coxsakievirus B5, un Enterovirus (Familia *Picornaviridae*) humano que al parecer se adaptó solo

a los porcinos, y dió origen al virus de la enfermedad vesicular del cerdo (Graves, 1973; Verdaguer y col., 2003; Zhang y col., 1993).

Actualmente encontramos al menos 868 agentes infecciosos que causan zoonosis, de los cuales 199 son virus. Así estos agentes infecciosos, además de tener interés en salud pública, afectan a varias especies de hospederos perjudicando parámetros productivos en diferentes tipos de ganado, y causando manifestaciones clínicas graves. Esto se ve reforzado por la hipótesis de que los patógenos zoonóticos tienen más probabilidades de estar asociados con enfermedades emergentes (Taylor y col., 2001).

En fechas recientes se ha descubierto que algunas enfermedades emergentes, principalmente de tipo viral, se han adaptado de los quirópteros al ganado porcino. Incluso esta última especie ha desempeñado un papel importante para que ciertos agentes infecciosos, en escenarios antropogénicos, logren causar enfermedades a otros animales, incluyendo al hombre. Ejemplos de estos virus son los paramixovirus (Nipah, Menangle), lyssavirus, coronavirus, filovirus y circovirus (Lau y col., 2010; Li y col., 2011). Así se hace evidente la importancia de vigilar los agentes virales de la especie porcina.

Agentes infecciosos emergentes y quirópteros

Los quirópteros son un orden de animales que ha cobrado un gran interés debido a los agentes infecciosos causantes de enfermedades emergentes que se han identificado en algunas de sus especies. En años recientes, se han suscitado escenarios en los que nuevas enfermedades han afectado la salud animal y la pública de forma importante, algunas con presentaciones severas y letales. Al estudiar los ambientes en los que ocurrieron dichos brotes, los agentes etiológicos se aislaron e identificaron en especies de murciélagos residentes. Algunos de los microorganismos patógenos identificados en diferentes especies de murciélagos

en el mundo son paramixovirus (Nipah, Hendra, Menangle) coronavirus (Torovirus, *Alpha* y *Betacoronavirus* entre ellos SARSr-CoV y MERSr-CoV), *Influenzavirus A*, *Lyssavirus*, *Filovirus* (Marburg virus), enterobacterias (*Salmonella* spp, *Yersinia* spp, *Listeria* spp, *Clostridium* spp, *Campylobacter* spp, *Shigella* spp, *E. coli* y *Vibrio* spp), *Leptospira* spp, *Rickettsia* spp, *Borrelia*, *Bartonella* y *Babesia* (Bowden y col., 2001; Marinkelle, 1996). Las frecuencias de identificación de agentes infecciosos en murciélagos van desde 2.4 hasta más del 50 % (Cuadro 1), dependiendo del agente infeccioso (Kuzmin y col., 2011; Young y Olival, 2016). En estos animales se han reconocido varios virus pertenecientes a 28 familias virales distintas (Lihong Chen y col., 2014). A pesar de los problemas serios en salud animal y pública que algunos de estos agentes han causado, poco se conoce de su ecología, evolución y el papel que desempeñan o pueden desempeñar los murciélagos en el mantenimiento, transmisión y dispersión interespecie de los microorganismos (Plowright y col., 2017), tomando en cuenta los hábitos y costumbres de los quirópteros, como el vuelo, la migración, interacciones entre comunidades intra e interespecies, las congregaciones en grupos numerosos y las interacciones con asentamientos rurales y urbanos. En este sentido hay que resaltar la perturbación humana, un factor común en los escenarios donde ha habido enfermedades emergentes (Afelt y col., 2018).

Existen estudios que demuestran que los murciélagos pueden infectarse y enfermar con agentes patógenos que adquieren en ambientes urbanos o pecuarios (Mühldorfer, 2013), y hay que tomar en cuenta que pueden diseminar la enfermedad a poblaciones silvestres, animales domésticos y humanos dados sus estilos de vida.

Cuadro 1. Frecuencias de identificación de algunos agentes virales en quirópteros.			
FAMILIA VIRAL	FRECUENCIA DE IDENTIFICACIÓN	ESPECIE (S)	REFERENCIA
<i>Coronaviridae</i>	6 - 55 %	Varias (Familias <i>Rhinolopidae</i> , <i>Vespertilionidae</i> , <i>Phyllostomidae</i> , <i>Hipposideridae</i> , <i>Pteropodidae</i> y <i>Molossidae</i>	(Dominguez y col., 2007; Gloza-Rausch y col., 2008; Susanna K. P. Lau y col., 2010; Poon y col., 2005; Tang y col., 2006; Tong y col., 2009)
<i>Rhabdoviridae</i>	2.44 - 22.07 %	Varias (Familias: <i>Phyllostomidae</i> , <i>Vespertilionidae</i> , <i>Molossidae</i> , <i>Pteropodidae</i> , <i>Embalonuridae</i> , y otras.	(Young y Olival, 2016)
<i>Filoviridae</i>	2.8 - 23.5%	<i>Hypsignathus monstrosus</i> , <i>Epomops franqueti</i> , <i>Myonycteris torquata</i> .	(Leroy y col., 2005)
<i>Orthomyxoviridae</i>	0 – 13.3 %	<i>Sturnira lillium</i> , <i>Artibeus planirostris</i>	(Tong y col., 2012)
<i>Paramyxoviridae</i>	0.22 – 6.13 %	Varias (Familias: <i>Pteropodidae</i> , <i>Hipposideridae</i> , <i>Vespertilionidae</i> , <i>Phyllostomidae</i> , <i>Mormopidae</i> , <i>Molossidae</i> , <i>Embalonuridae</i> , y otras.	(Young, 2016)
<i>Flaviviridae</i> , <i>Herpesviridae</i> y <i>Papillomaviridae</i>	0 – 100 %	<i>Vespertilionidae</i> *, <i>Pteropodidae</i> y otras	(Young, 2016)
* Familia en la que solamente se ha identificado papillomavirus			

En la actualidad existen más de 12 443 secuencias de virus, pertenecientes a 30 familias, reportadas en la base de datos DbatVir (L. Chen y col., 2014; “DBatVir: the Database of Bat-associated Viruses,” febrero 2021). Estos virus han sido identificados en diferentes especies de quirópteros, que, al parecer, son los hospederos naturales, pero, en ambientes de coexistencia con el hombre y animales domésticos algunos de estos agentes han causado contagios interespecie y zoonosis. Así, en los porcinos se han identificado varias infecciones interespecie que se han transmitido de especies de murciélagos en ambientes antropizados en los que coexisten ambas especies. La mayoría de estos casos han culminado en la adaptación de los agentes a la especie doméstica, causando enfermedades importantes. Existen casos en los que agentes infecciosos de porcinos se adaptan a los quirópteros, como el rubulavirus (Huang y col., 2012).

Adaptaciones virales interespecies murciélagos – porcinos en ambientes perturbados por el hombre

Es importante poner atención e interés en el ambiente (sus modificaciones y alteraciones) y en las interacciones agentes infecciosos – especies animales que se suscitan en él (Kahn y col., 2012). Sobre todo en los escenarios dónde existe una convivencia de especies diferentes, en especial si hay coexistencia de especies domésticas con fauna silvestre, ya que, al no tener datos de la evolución de los agentes infecciosos, o simplemente desconocer su existencia, no es posible vigilar, predecir ni prevenir los riesgos de brotes o epidemias. El *Nipah henipavirus*, un miembro de la familia *Paramyxoviridae*, se adaptó a nuevos hospederos gracias a la introducción de cerdos en la selva de Malasia (Luby y col., 2009). En 1998 los cerdos criados en granjas establecidas en la selva de este país comenzaron a presentar una enfermedad neurológica-respiratoria que después fue transmitida a y entre humanos. El brote dejó más de 1.1 millón de cerdos muertos (por enfermedad o sacrificio) y 283 casos en humanos con 109 decesos.

Tras investigaciones el virus fue identificado en murciélagos del Género *Pteropus* spp que habitaban la selva y se alimentaban de los árboles frutales plantados en la granja para su doble propósito (Smith y Wang, 2013). Este es un claro ejemplo de un escenario dónde la introducción del cerdo como especie doméstica generó una nueva enfermedad al entrar en contacto con un agente infeccioso nativo.

Existen otros ejemplos de virus asociados a quirópteros que han causado importantes brotes en los porcinos, y se han adaptado causando manifestaciones clínicas en los humanos. Dentro de estos tenemos a los ebolavirus (Leroy y col., 2005; OMS, 2017), el virus Menangle (Bowden y col., 2001; Philbey y col., 2008, 1998). También se ha encontrado evidencia de virus que circulan en porcinos y quirópteros sin causar enfermedad. Este es el caso del virus de Tioman, el cual se identificó inicialmente en murciélagos, y posteriormente se encontró evidencia de la infección en porcinos y humanos desde 2001 (Yaiw y col., 2008, 2007).

En América se han encontrado virus de animales domésticos en quirópteros como es el caso del herpesvirus bovino y el rubulavirus porcino (Escalera-Zamudio y col., 2018; Wang y col., 2007). En el caso de los paramyxovirus esto es respaldado mediante el estudio genómico de los virus de Mapuera (Brasil) y rubulavirus (México) que develó su relación y generó probables hipótesis de su origen (Cuevas-Romero JS, Blomstrom AL, Berg, 2015; Wang y col., 2007).

Basándonos en lo anterior, es importante estudiar las comunidades de agentes infecciosos que portan los murciélagos y los animales domésticos, como los porcinos, en los probables ambientes de interfaz. De esta manera es probable vislumbrar si existen infecciones interespecie y riesgos de enfermedades para cualquier especie.

La secuenciación de siguiente generación: una herramienta auxiliar para el estudio de las comunidades de agentes infecciosos.

En la actualidad existen varias herramientas con las cuales es posible estudiar más eficaz y rápidamente a los agentes infecciosos y sus modificaciones y su nicho ecológico. La metagenómica obtiene y analiza secuencias representativas o genomas de los microorganismos que componen una comunidad en un ambiente u hospedero, por ejemplo, la microbiota presente en una muestra de suelo o en una determinada muestra del intestino de un animal (Handelsman y col., 1998). Mediante la extracción y el análisis del ácido nucleico de diferentes especies de organismos se ha podido develar causas de enfermedades, rutas de transmisión, vectores, nuevos hospederos, y agentes benéficos para determinadas especies animales, entre otros fines (Hugenholtz y Tyson 2008). En un inicio la metagenómica fue utilizada para el estudio de consorcios bacterianos a partir de los cuales es difícil aislar o cultivar ciertas bacterias (Singh y col., 2008). Recientemente son varios los informes que han utilizado a la metagenómica para describir la microbiota, con diversos fines, en seres humanos y animales tanto silvestres como domésticos (Lamendella y col., 2011; Schürch y col., 2014), siendo de gran utilidad para el estudio de las enfermedades.

Para cualquier plataforma de secuenciación masiva, existe una serie de pasos en común a seguir para la obtención de los resultados. Estos son:

I. Colección de la muestra a partir de la cual se obtendrán los ácidos nucleicos de los agentes infecciosos objetivo. Estas muestras pueden ser procesadas por varias técnicas (ej. centrifugación, filtración, etc) para obtener microorganismos específicos de interés.

II. Obtención y enriquecimiento de ácidos nucleicos, paso en el cual se purifican los ácidos nucleicos mediante métodos generales (ej. fenol - cloroformo -

isopropanol y/ó trizol) o con productos o paquetes comerciales para, posteriormente, generar un mayor número de copias del material genético de interés (ej. hexámeros aleatorios, oligodT, iniciadores M13, etc.).

III. Construcción de librerías de DNA. Las copias del material genético enriquecido son fragmentadas (mediante enzimas, sonicación o nebulización) en tamaños lo más homogéneos posible (idealmente entre 100 a 400 bases) para conformar con estos las librerías de DNA. Estos fragmentos de DNA/cDNA que conforman la librería son identificados uniéndolos a oligonucleótidos conocidos para someterlos a procesos de secuenciación,

IV. Secuenciación masiva. En este paso se determinará la secuencia de cada fragmento analizado empleando diferentes fundamentos y plataformas (ej. ilumina, Ion S5, PacBio Sequel systems, MinION, etc). Posteriormente los resultados serán emitidos como archivos de texto de gran tamaño (que ocupan gigabytes: GB) en formatos especiales (dependiendo del secuenciador).

V. Análisis bioinformático. Los archivos generados por el secuenciador empleado contendrán miles o millones de secuencias (lecturas o reads) las cuales tendrán que ser evaluadas, editadas, seleccionadas, organizadas y analizadas de forma masiva con recursos y herramientas computacionales (bioinformática), pues a una persona le tomaría demasiado tiempo realizar estas tareas, aunado a la complejidad de ordenar y organizar los resultados. Así el análisis bioinformático busca asignar homología a cada lectura (secuencia) encontrada, para develar los microorganismos presentes en la muestra.

El potencial de la secuenciación masiva puede ser visto a su vez como su desventaja, pues estas técnicas carecen de especificidad y pueden amplificar secuencias ambientales, de hospedadores y de agentes infecciosos (cualquier ácido nucleico presente en la muestra). Afortunadamente, se han desarrollado y

perfeccionado una variedad de protocolos y variantes para obtener mejores rendimientos, y así dirigir los resultados de acuerdo al grupo-tipo de organismo(s) y/o tipo de ácido nucleico de interés. La capacidad para detectar virus desconocidos complica la identificación y la interpretación de los resultados. Pero gracias a la estandarización de mejores análisis bioinformáticos y las cada vez más robustas bases de datos (debido al reporte de genomas, secuencias nucleotídicas y de aminoácidos), es cada vez más eficiente la identificación y caracterización de ácidos nucleicos obtenidos por secuenciación masiva. Otras ventajas son la gran cantidad de lecturas de secuenciación que permiten identificar agentes infecciosos que se encuentran en concentraciones relativamente bajas en la muestra; se han publicado protocolos universales de secuenciación metagenómica que son capaces de detectar tanto virus DNA como RNA de cadena simple y doble. A pesar de estas ventajas, y la constante disminución de costos de la secuenciación masiva, aunque sigue siendo más costosa, por muestra, que los métodos tradicionales (ej. Sanger, pirosecuenciación), requiere más tiempo, no sólo en el procesamiento de la muestra, sino también debido a la gran cantidad de secuencias generadas por muestra, la administración de datos es más compleja, requiere software especializado y personal capacitado y recursos informáticos con mayor capacidad. Por fortuna existen herramientas y plataformas de uso gratuito disponibles para el análisis de los datos masivos de metagenómica (ej. FastQC para evaluar calidad; Trimmomatic para edición y selección; BLAST para asignación de homología; Trinity y Velvet para ensambles; Galaxy como plataforma que condensa y pone disponibles una gran diversidad de herramientas bioinformáticas), pero es necesario aprender y profundizar en su uso para lograr buenos resultados de calidad y confiables (Ghurye y col., 2016). Sin duda la secuenciación metagenómica es una herramienta con ventajas únicas sobre otros métodos, como identificar virus en muestras donde el diagnóstico tradicional no identifica

patógenos, determinar secuencias de genoma virales directamente a partir de muestras clínicas, caracterizar completamente virus e investigar la ecología viral de complejos de enfermedades. Su uso más extendido está conduciendo al descubrimiento de nuevos agentes y está ampliando significativamente nuestro conocimiento de la epidemiología viral.

Justificación.

No existe información de las comunidades microbianas y virales de los porcinos de México, donde uno de los más importantes sistemas de producción es el traspatio (SAGARPA, 2006). En México los porcinos de traspatio aportan aproximadamente el 40 % de los más de 14 millones de cabezas producidas anualmente que generan más de 1.4 millones de toneladas de la segunda carne más consumida en el país (INEGI, 2019; SIAP, 2015; Villamar Angula y col., 2006). La granja de traspatio es un sistema de producción diverso el cual es utilizado por varias familias, principalmente, en zonas rurales y periurbanas de regiones geográficas en desarrollo. En este sistema los animales son alimentados con productos agrícolas, desperdicios de comida y por pastoreo, su manejo es variable, de acuerdo al productor y al ambiente. Los animales son mantenidos en cercos rústicos o incluso libres, donde frecuentemente coexisten en contacto directo o indirecto con varias especies domésticas, incluyendo al humano, y especies silvestres como los murciélagos. México posee 138 especies de murciélagos, ocupando el quinto lugar en diversidad de especies de este orden (Medellin y col. 2007). Las comunidades de agentes infecciosos de estas especies no han sido estudiadas en México, y poco se sabe de las comunidades virales de murciélagos americanos y de porcinos de traspatio, o si existen infecciones interespecie que estos animales pueden adquirir en ambientes de interfaz. Este trabajo toma como objetos de estudio a los murciélagos y al ganado porcino de traspatio en México. Esta información ecológica puede servir como punto de

partida para programas de prevención, vigilancia, herramientas de diagnóstico, predicción y otras aplicaciones de gran utilidad en salud animal, medicina preventiva y ecología.

HIPÓTESIS

Los murciélagos y los porcinos que convergen en ambientes de traspatio en el estado de Morelos, México son hospederos de virus, por lo que al estudiar por secuenciación masiva el material genético derivado de muestras de hisopos rectales y nasales/faríngeos de estas especies, se encontrarán las secuencias de agentes conocidos, nuevas variantes y virus que no han sido reportados en estas especies.

OBJETIVO GENERAL

Describir y, de ser posible, caracterizar (parcial o total), por secuenciación masiva, el viroma rectal y naso-faríngeo que hospedan porcinos y especies de murciélagos asociados a ambientes de traspatio rural en el estado de Morelos donde convergen estas especies. Poniendo especial interés en virus animales.

OBJETIVOS ESPECÍFICOS

Identificar áreas de muestreo de porcinos y murciélagos y coleccionar las muestras.

Procesar y obtener los ácidos nucleicos virales de las muestras.

Enriquecer los ácidos nucleicos y construir las librerías de DNA.

Llevar a cabo la secuenciación masiva de las librerías de DNA de las especies estudiadas.

Realizar la evaluación de calidad, edición, selección, ensamblaje y asignación de homología de los datos de secuenciación obtenidos.

Analizar lecturas y ensamblajes para identificar virus y obtener gráficos y estadísticas (rarefacciones, índices de Shannon-Wiener y gráficos de lecturas virales).

Reportar al GenBank las secuencias virales encontradas.

Construir los análisis filogenéticos de las secuencias y contigs virales obtenidas.

Generar publicaciones en revistas indexadas internacionales con los resultados obtenidos.

MATERIALES Y MÉTODOS

Área de estudio y muestras.

En el estado de Morelos se identificaron cuatro sitios para obtener las muestras de porcinos y quirópteros. En cada sitio se ubicó un predio o granja porcina de traspatio rural y un refugio de murciélagos cercano (la distancia entre refugios y granjas de porcinos de traspatio fue menor a diez kilómetros para garantizar la superposición (convergencia) de las actividades biológicas de los murciélagos en las granjas porcinas). Las granjas y refugios se ubicaron en los municipios de Tlaquiltenango (granjas A y D), Puente de Ixtla (granja B) y Tlaltizapan (granja C) en el estado de Morelos, México (Figura 1 y Cuadro 2). Las granjas elegidas se ubicaban en poblaciones rurales rodeadas de ambientes donde es posible encontrar especies silvestres como prociónidos, cánidos, otros carnívoros, roedores, diversas aves, además de murciélagos y albergaban otras especies de animales domésticos, como ganado vacuno, ovejas, cabras, aves, gatos y/o perros que convivían en cercanía. Los sitios de muestreo se encuentran en un clima de sabana tropical con variantes secas o semisecas (Aw1 o Aw0; basado en el sistema de clasificación climática Köppen-Geiger), con una temperatura media anual de 23.5 ° C a 25 ° C, sequía desde finales de otoño hasta principios de primavera, y precipitaciones anuales de 840 a 930 mm. En general, las granjas alojaban a los porcinos en corrales de piso de tierra o cemento, con un área mayor por animal en comparación con los estándares establecidos en sistemas intensivos (0.36-0.66 m² / animal) y, por lo tanto, con densidades de población más bajas (un mínimo de 4 m² / animal). La alimentación consistía principalmente en alimentos comerciales, maíz y restos agrícolas, y, en algunos casos, se liberaba a los cerdos (libre pastoreo). En las granjas A, C y D los animales vivían en grupos de edades mixtas. Las granjas A, C y D eran de ciclo completo y/o venta de lechones, mientras que la granja B era una engorda. No había programas de medicina preventiva. El estado de salud de los animales fue determinado a

través de una anamnesis completa y/o un examen clínico que incluyó observación del comportamiento normal, piel, anomalías en la forma o tumores en la superficie del cuerpo, heces, ausencia de signos neurológicos o respiratorios y anomalías en el movimiento. Además se observó el color de la membrana mucosa (rosa pálido), el tiempo de llenado capilar (≤ 2 segundos), y se auscultaron los nódulos linfáticos del abdomen, parotídeos, submandibulares e inguinales.

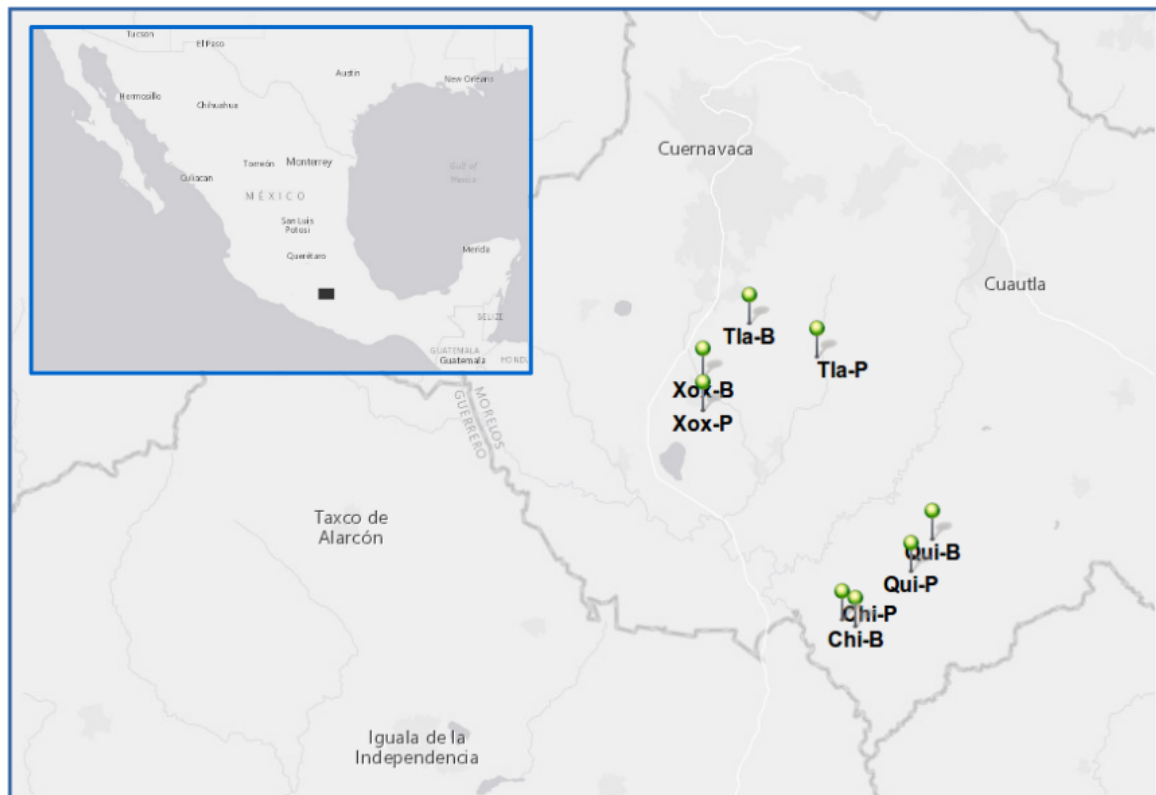


Figure 1. Ubicación geográfica de refugios de murciélagos (B) y granjas porcinas (P) de donde se obtuvieron las muestras en Xoxocotla (Xox) municipio de Puente de Ixtla, Chimalacatlán (Chi) y Quilamula (Qui) municipio de Tlaquiltenango y Tlaltizapán (Tla) en Morelos La ubicación del área de estudio (mapa grande) se representa en el mapa de la República Mexicana (mapa de la esquina superior izquierda) con el rectángulo oscuro (■). El mapa se realizó con ArcGIS en línea (<https://www.arcgis.com/home/webmap/viewer.html> fecha de acceso: 13-agosto-2017).

Los propietarios de los porcinos confirmaron avistamientos o visitas de murciélagos a las granjas, pues eran atraídos por insectos (como mosquitos y otros insectos atraídos por los focos) o por árboles frutales.

Muestras biológicas

Las muestras fueron colectadas durante los meses de febrero a julio del 2016. Se colectaron hisopos rectales y nasales de 23 porcinos de dos a cuatro meses de edad, mientras que de los murciélagos se obtuvieron hisopos faríngeos y rectales. Los murciélagos fueron capturados con redes de niebla en sus refugios. Los murciélagos se liberaron de la red de niebla, se colocaron en contenedores individuales (disminuyendo el estrés) para colectar las muestras y liberarlos. Las especies de murciélagos más abundantes capturadas fueron *Artibeus* sp (n = 17), *Pteronotus parnellii* (n = 13) y *Macrotus waterhousii* (n = 9) (Cuadro 2). Entre los especímenes de *Artibeus*, los 11 animales capturados en Tlaltzapán fueron identificados como *Artibeus phaeotis* con base en los criterios morfológicos establecidos por Medellín y col. (2007). Los otros seis ejemplares no fueron identificados a nivel de especie porque fueron liberados antes de la identificación morfológica (error en el proceso de muestreo). Las muestras se mantuvieron en 1.5 ml de medio esencial mínimo de Eagle (EMEM) a temperatura de refrigeración (2-6 °C) y fueron trasladadas al laboratorio de Biotecnología en Salud Animal del CENID-Microbiología Animal del INIFAP. El manejo y muestreo de animales para este estudio fue aprobado por el Comité Institucional para el Cuidado y Uso de Animales Experimentales (CICUAE) de la Facultad de Medicina Veterinaria y Zootecnia de la Universidad Nacional Autónoma de México (UNAM) por los protocolos DC-2016-2-2 y DC-2017-2-2.

En el laboratorio, las muestras se homogenizaron en vórtex, los hisopos fueron desechados en bolsa roja de RPBI y los viales fueron centrifugados (6000 x g/15 minutos) para separar bacterias, células y material particulado. El sobrenadante se

Cuadro 2. Información de las muestras y sitios de muestreo.

	Sitios de muestreo			
	Quilamula, Tlaquiltenango (A)	Xoxocotla, Puente de Ixtla (B)	San Miguel 30, Tlaltizapan (C)	Chimalacatlán, Tlaquiltenango (D)
N de porcinos (edad)	5 (3 meses)	7 (4 meses)	4 (4 meses)	7 (2-3 meses)
Especies de murciélagos (n)	<i>Pteronotus parnellii</i> (11)	<i>Artibeus</i> sp. (3)	<i>Artibeus</i> sp (11) <i>Pteronotus parnellii</i> (2)	<i>Artibeus</i> sp. (3) <i>Macrotus waterhousii</i> (9)
Ubicación geográfica del sitio de muestreo de porcinos	18°30'31" N, 99°1'17" O	18°39'49" N, 99°14'17" O	18°43'7" N, 99°7'9" O	18°27'46" N, 99°5'37" O
Ubicación geográfica del sitio de muestreo de murciélagos	18°32'10" N, 98°59'54' O	18°41'51" N, 99°14'05" O	18°45'0" N, 99°11'22" O	18°27'09" N, 99°04'34 O
Distancia entre sitios de muestreo de porcinos y murciélagos	4.2 km	3.18 km	8.29 km	1.57 km
Descripción del hábitat que rodea los sitios de muestreo (el tipo de vegetación en todas estas áreas era bosque bajo caducifolio de clima cálido).	Cadena de montes, cañada cubiertos por bosque, que rodean la presa de Quilamula, cerca de los campos de cultivo y el poblado de Quilamula.	Zona de hondonadas y Terreno llano, bajo el cual por se encontraba la cueva cubiertos por bosque. Algunas de las áreas circundantes más altas se utilizaban para cultivos y, con menor frecuencia, para ganado. Cercano al pueblo.	Tierras utilizadas para invernaderos florales, casas rurales y asentamientos humanos.	Zona de cerros, cubiertos por bosque, cerca de valles. Algunos de los cuales se utilizaban para cultivos.

transfirió a un tubo limpio de 1.5 ml previamente identificado. Las muestras fueron conservadas en congelación ($-75^{\circ}\text{C} \pm 5$) hasta su uso en el paso siguiente.

Obtención de ácidos nucleicos (RNA/DNA)

En un tubo limpio de 15 ml se transfirieron 200 μl del sobrenadante de cada muestra individual (nasal o faríngea y rectal) para realizar una muestra conjunto por especie. La muestra por especie se filtró a través de membranas de 0,45 μm y 0,22 μm (Millipore) para eliminar partículas del tamaño de células eucariotas y bacterianas, seguido de ultracentrifugación a $170,000 \times g$ durante 3 horas a 4°C (Beckman Coulter Optima Ty-70) para concentrar las partículas virales. El sedimento (partículas virales) se resuspendió en solución EMEM y se filtró con membrana de 0,22 μm . El filtrado se trató con DNasa (Thermo Fisher Scientific) y RNasa A (Thermo Fisher Scientific) a 37°C durante 30 minutos para digerir los ácidos nucleicos libres. Los ácidos nucleicos virales totales (tanto RNA como DNA) se extrajeron y purificaron mediante el Purelink Viral RNA/DNA Mini Kit (Invitrogen) de acuerdo con el protocolo del fabricante. Se usó una muestra con EMEM solo como testigo de extracción de RNA / DNA. La concentración de RNA / DNA se estimó por espectrofotómetro (Nanodrop - Thermo Fisher Scientific), lo que evidenció la presencia nula de ácidos nucleicos en el control. El RNA / DNA purificado se dividió en dos fracciones, una como una mezcla de RNA y la otra como una mezcla de DNA para el enriquecimiento de los ácidos nucleicos purificados.

Se cuantificó la concentración y comprobó la pureza de los ácidos nucleicos mediante espectrofotometría (Nanodrop) (relación ácido nucleico:proteínas ≥ 1.8).

Los ácidos nucleicos (DNA/RNA) fueron almacenados a -70°C hasta su utilización.

Construcción de las librerías.

Retrotranscripción y enriquecimiento de RNA. A partir del RNA se generó cDNA empleando el Transcriptor First Strand cDNA Synthesis kit V.6 (ROCHE) con oligo dT y hexámeros unidos a iniciadores M13, basados en el procedimiento de Djikeng y col. (2008). Los volúmenes y condiciones se describen en el Anexo 1.

Enriquecimiento de muestras de DNA. Para el enriquecimiento de los DNA se empleó el Faststart High Fidelity PCR System (Roche) siguiendo las instrucciones del fabricante. Los volúmenes empleados para las reacciones y el programa de termociclado se muestran en el Anexo 2.

Los productos fueron guardados a -80 °C para su uso en la generación de ácidos nucleicos enriquecidos de doble cadena.

Generación de DNA y cDNA de doble cadena. Con las muestras de cDNA y DNA enriquecidas, se generó la doble cadena empleando el fragmento largo de la polimerasa I (fragmento klenow, New England Biolabs) siguiendo las condiciones del fabricante descritas en el Anexo 3.

Amplificación de cDNA y DNA por PCR. Se realizó un enriquecimiento de los DNA/cDNA de doble cadena mediante PCR empleando el Faststart High Fidelity PCR System (Roche) siguiendo las instrucciones del fabricante (Anexo 4).

Los productos fueron conservados a -80 °C, hasta su utilización para la fragmentación.

Fragmentación de cDNA y DNA de doble cadena. Los procedimientos de fragmentación, adición de índices-adaptadores, enriquecimiento, amplificación y purificación de librerías se realizaron en el laboratorio de virología del Instituto Nacional de Enfermedades Respiratorias (INER). Las muestras fueron transportadas en hielo seco para su conservación. Se midió la concentración de las muestras de cDNA y DNA de doble cadena empleando el Qubit (Thermo Fisher Scientific). Para esto se emplearon dos µl de cada muestra y 198 µl de agua desionizada, homogenizando bien la mezcla, transfiriéndola a la celda del

fluorómetro y realizando la medición. Esto se realizó para conocer la concentración de ácidos nucleicos presentes en cada muestra y mezclar el cDNA y DNA de doble cadena de cada especie (cDNA/DNA-especie) en una sola muestra (≥ 1 ng de ácidos nucleicos).

Para realizar la fragmentación de las muestras de cDNA/DNA-especie se empleó el Nextera XT DNA library prep kit (Illumina) siguiendo las instrucciones del fabricante (Anexo 5). Con este procedimiento se confirmó la longitud de 150 – 450 pb en la mayor proporción de fragmentos de ácidos nucleicos obtenidos y su concentración suficiente (1 ng) para continuar con el procedimiento.

Adición de índices-adaptadores y amplificación de librerías. Con los cDNA/DNA fragmentados de cada especie, confirmando su tamaño y concentración adecuados se realizó la adición de secuencias índice-adaptadores, para conformar como tal las librerías y amplificarlas siguiendo las indicaciones del fabricante descritas en el Anexo 6. Al terminar las librerías se purificaron para dejar únicamente los fragmentos de cDNA/DNA marcados con secuencias índices-adaptadores.

Purificación de librerías. Para purificar las librerías se empleó el Agencourt AMPure XP kit siguiendo las instrucciones del fabricante (Anexo 7). La calidad y cantidad de los fragmentos de cada librería fue evaluado y confirmado mediante los equipos Qubit y Agilent Technology 2100 Bioanalyzer (procedimientos anteriormente descritos)

Secuenciación

Las librerías purificadas fueron secuenciadas en la plataforma de secuenciación masiva MySeq sequencing system (Illumina), en la unidad de secuenciación del laboratorio de Virología del INER. De cada especie se obtuvieron millones de secuencias en formato fastq comprimidas (*.gz) en dos archivos (por especie), uno que contenía las secuencias sentido y el otro con las secuencias complemento o

antisentido (producto de la secuenciación pareada). Estos archivos fueron manipulados, procesados y analizados posteriormente por herramientas bioinformáticas.

Análisis Bioinformático

Vigilancia de la calidad de los datos. Los dos archivos de datos de secuenciación de cada especie fueron descomprimidos para evaluar sus propiedades y calidad mediante la herramienta FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Con esta herramienta se analizó cada conjunto de datos por especie monitoreando características de interés como la calidad, la longitud, la distribución de secuencias por su longitud, el contenido de bases, las secuencias sobrerrepresentadas y la presencia de secuencias de adaptadores, entre otras.

Edición y filtrado de datos. Para editar y seleccionar (filtrar) las secuencias por su calidad, además de eliminar las secuencias de iniciadores M13 y adaptadores que pudieran estar presentes en las lecturas (secuencias), se empleó la herramienta de bioinformática Trimmomatic 0.36 (Bolger y col., 2014). La herramienta funcionó a partir de la terminal (línea de comandos) utilizando el siguiente archivo de instrucciones (script), el cual se encuentra referido en el manual del software (disponible en <http://www.usadellab.org/cms/?page=trimmomatic>). A continuación se describe la instrucción empleada.

- `java -jar trimmomatic-0.36.jar PE Asp.S7.L001_R1_001.fastq Asp.S7.L001_R2_001.fastq Asp_R1_Pair.fq Asp_R1_Unpair.fq Asp_R2_Pair.fq Asp_R2_Unpair.fq ILLUMINACLIP:TrueSeq30-PE-2.fa:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:3:20 MINLEN:30`
- `java -jar trimmomatic-0.36.jar` = instrucción para activar el software.

- **PE** = se le indica a la herramienta que analizará secuencias pareadas.
- **Asp.S7.L001_R1_001.fastq** = en esta sección se le indica el nombre de los dos archivos que contienen las lecturas sentido y antisentido (dos archivos por especie) empleados en el proceso.
- **Asp_R1_Pair.fq Asp_R1_Unpair.fq** = en esta sección se indican los nombres de los archivos (*.fq = formato fasta q) en los que se guardarán las secuencias que reúnan las características requeridas.
- **ILLUMINACLIP:TrueSeq30-PE-2.fa:2:30:10** = instrucción para identificar secuencias de adaptadores e iniciadores M13 : ubicación del archivo fasta que contiene las secuencias de adaptadores e iniciadores M13 que serán buscadas y eliminadas : especifica el número de discrepancias (bases) que permitirá una coincidencia (reconocimiento de un adaptador o iniciador) : cuán precisa debe ser la coincidencia entre las dos lecturas 'ligadas por adaptador' para el alineamiento de secuencias palindrómicas pareadas (expresado en bases) : especifica cuán precisa debe ser la coincidencia entre cualquier secuencia de adaptador o iniciador contra una lectura.
- **LEADING:20** = Eliminó las bases de baja calidad del principio de las lecturas (extremo 5'). Mientras una base tenga un valor por debajo de este umbral establecido (calidad medida en phred33), la base se eliminará y se investigará la siguiente base
- **TRAILING:20** = Realizó la misma función que el comando anterior pero al final de la lectura (extremo 5').
- **SLIDINGWINDOW:3:20** = Realizó un recorte de "ventana deslizante", eliminando las bases cada vez que la calidad promedio dentro de la

ventana (número de bases establecidas, en este caso tres) fuera menor al umbral establecido (para este caso 20).

- MINLENGTH:30 = Este comando eliminó las lecturas o secuencias que resultan menores a la longitud establecida (en este caso 30 bases).

Este proceso se realizó con el conjunto de datos (lecturas o secuencias) de cada especie. Al finalizar el proceso de edición-selección los datos fueron evaluados nuevamente mediante el programa FastQC (descrita en el paso anterior) para confirmar las características adecuadas de los datos. De cada especie se obtuvieron cuatro archivos correspondientes a: 1) lecturas sentido pares; 2) lecturas antisentido pares; 3) lecturas sentido no pares y 4) lecturas antisentido no pares. Cabe mencionar que el proceso de edición-selección de datos, se llevó a cabo varias veces, variando los valores utilizados en los diferentes comandos (previamente descritos) hasta obtener los mejores resultados en la calidad y longitud de las lecturas (reflejados en esta metodología y vigiladas con el FastQC) para continuar con los pasos subsecuentes.

Búsqueda de homología (BLAST). Para realizar la búsqueda de homología del conjunto de lecturas de cada especie se empleó la herramienta MegaBLAST disponible en la plataforma Galaxy (Afgan y col., 2018) [<https://usegalaxy.org>]. Para utilizar la plataforma y las herramientas disponibles en ella se obtuvo una cuenta. Una vez realizado el registro de la cuenta se realizó el siguiente procedimiento.

- Primero se cargaron los grupos de datos de cada especie en el servidor de Galaxy seleccionando la opción [Upload File from your computer] del menú “Get Data”.
- Una vez que todos los datos estuvieron disponibles en el servidor de Galaxy, se procedió a convertirlos de formato Fastq a formato Fasta empleando la

herramienta [FASTQ to FASTA converter] disponible en el menú “Convert Formats”.

- Se concatenaron los datos correspondientes a cada especie, ordenando primero las lecturas pares, luego las sentido no pares y por último las antisentido no pares. Esto se realizó con la herramienta [Concatenate datasets tail-to-head] disponible en el menú “Text manipulation”. Así se condensó en un solo archivo todas las lecturas de una especie.
- Del menú “Convert Formats” se seleccionó la opción [FASTA to Tabular converter] para convertir todos los datos a formato tabular.
- A los archivos tabulares se les adicionó una columna con la función [Add column to an existing dataset] (disponible en el menú “Text manipulation”). Con esto se adicionó un prefijo de cada especie a su conjunto de datos correspondiente (Asp para *Artibeus* sp; Mwat para *Macrotus waterhousii*; Ppar para *Pteronotus parnelli*; y Porcinos) seguido de un número consecutivo activando la función “iterate”
- Se convirtieron los archivos tabulares a formato FASTA con la opción [Tabular to FASTA] disponible en el menú “Convert Formats”, seleccionando la columna tres como identificadores (prefijo-número consecutivo) y la columna dos como secuencia, eliminando de esta forma la columna uno que contenía los identificadores originales (códigos alfanuméricos) asignados por el secuenciador.
- Con los archivos FASTA se realizó la búsqueda de homología con las bases de datos del GenBank nt (secuencias nucleotídicas) y wgs (genomas completos) empleando la opción [Megablast] disponible en el menú “Mapping”. Para el análisis se emplearon las opciones de división de lecturas en cadenas de 28

caracteres (opción “using word size”) y un Evalue = 0.001. Los resultados se obtuvieron en archivos tabulares.

- A la par del Megablast se realizó una blastx (búsqueda de homología con proteínas reportadas a partir de la traducción de las lecturas a secuencias de aminoácidos en sus 6 posibles marcos de lecturas) empleando la herramienta [NCBI BLAST+ blastx] disponible en el menú “NCBI BLAST+”.
- Se realizó el cómputo de la longitud de cada secuencia en los archivos FASTA con la herramienta [Compute sequence length] disponible en el menú “FASTA/FASTQ” la cual realizó un cálculo de la longitud de cada secuencia reportándolo en un archivo tabular.
- Se concatenaron los archivos resultantes de los análisis del Megablast (el de bases nucleotídicas con el de genomas completos) de cada especie, con la herramienta [Concatenate datasets tail-to-head] disponible en el menú “Text manipulation”.
- Se unieron los archivos con los resultados de los dos últimos pasos (a la concatenación de los resultados del Megablast se le adicionó la columna con los resultados del cómputo de la longitud de secuencias) empleando la herramienta [Join two files] disponible en el menú “Text Manipulation” resultando un archivo tabular.
- En los archivos anteriores se seleccionaron las lecturas con similitud > 50% (con base en la longitud total de la secuencia). Esto se realizó con la herramienta [Filter] (disponible en el menú “Filter and sort”) empleando la condición: $c5/c14 > 0.5$ donde la c5 = columna con la longitud de la región alineada; c14 = longitud total de la secuencia.

- Con las secuencias del paso anterior se les asignó el nivel taxonómico más específico encontrado en la búsqueda de homología (BLAST) usando [Find lowest diagnostic rank] (disponible en Galaxy/test menú “Old metagenomics”).
- Se elaboró un resumen taxonómico con los datos encontrados utilizando la herramienta [Summarize taxonomy] (disponible en el menú “Old metagenomics” disponible en Galaxy/test menú “Old metagenomics”).

Ensamble de lecturas y búsqueda de homología de contigs. Las lecturas de cada conjunto de datos por especie fueron ensambladas *de novo* para obtener los contigs (contig: secuencia constituida por lecturas que se superponen en una misma secuencia consenso). Para esta tarea se empleó la herramienta Metavelvet versión 1.2.02 (Namiki y col., 2012) siguiendo el manual del software. El procedimiento se describe brevemente a continuación.

- Se elaboró el directorio de secuencias y archivos necesarios para realizar el ensamble (log = hashtable, Roadmaps y secuencias) de cada conjunto de secuencias con el siguiente script:

```
velveth [directorio destino para los resultados] [longitud de K-mero] -
fastq(formato de los datos de entrada) -shortPaired(orden para señalar que
se proveerá la ubicación de las lecturas pareadas) [Directorio1: ruta del
archivo 1 de lecturas pareadas empleadas en el ensamble] [Directorio 2:
ruta del archivo 2 de lecturas pareadas empleadas en el ensamble] -
short(orden para señalar que se proveerá la ubicación de las lecturas no
pareadas) [Directorio 1: ruta del archivo 1 de lecturas no pareadas]
[Directorio 2: ruta del archivo 2 de lecturas no pareadas]
```

- Se empleó velvetg para elaborar gráfica de Bruijin (construcción de nodos o sobreposición parcial o total de lecturas). Esto se realizó con los conjuntos de datos de cada especie empleando el siguiente script:

velvetg(instrucción para activar velvetg) Documentos/RoLo/Doctorado/Result-Metavelvet/Asp_Metavelvet/(directorio de la ubicación de los datos a analizar) -exp_cov auto (cobertura estimada automática) -ins_length auto (distancia entre lecturas automática). Este paso generó archivos de texto con los ensamblados.

- Los archivos con los ensamblados fueron sometidos a búsqueda de homología siguiendo los pasos descritos previamente (sección anterior “*Búsqueda de homología (BLAST)*”).

Análisis estadísticos y diversidad

Con los datos obtenidos en los resúmenes de los resultados de homología de lecturas de cada especie se realizó el cálculo de rarefacción con su interpolación y extrapolación para estimar la representatividad de la diversidad (esfuerzo de muestreo). Esto se realizó con el software iNEXT (Chao y col., 2016). Y también se obtuvo el índice de Shannon-Wiener (índice de equidad) (Moreno, 2001) para evaluar y comparar la diversidad entre muestras y contra otras reportadas.

Identificación y confirmación de virus por PCR/RT-PCR

Para confirmar la presencia de los virus animales más abundantes y la identificación de algunos virus de interés se utilizaron ensayos moleculares de PCR (virus DNA) o RT-PCR (virus RNA). Los ensayos de RT-PCR fueron realizados con el Verso 1-Step RT-PCR Hot Start kit(Thermo Fisher Scientific) y para los de PCR se empleó el Taq polymerase kit (Thermo Fisher Scientific). Los resultados fueron revelados por electroforesis (110 V / durante 60 minutos \pm 10) en geles de agarosa al 1% que fueron teñidos con solución de Midori (colorante intercalante) y visualizados en el sistema de fotodocumentación de geles Gel Doc XR+ (BIO-RAD). Los detalles de cada ensayo se describen a continuación.

RT-PCR universal para Coronavirus. Para la identificación/confirmación de coronavirus se utilizaron los iniciadores degenerados PC2S2 3'- TNA TGG GNT

GGG AYT AYC C -5' y IZS-CoV-f 3'- GAR CAR AAY TCR TGH GG -5' (De Souza Luna y col., 2007; Lelli y col., 2013) para amplificar un fragmento de 603 nucleótidos del ORF1ab (polimerasa) mediante una reacción de RT-PCR. El procedimiento de RT-PCR para la identificación/confirmación universal de coronavirus se muestra en el Anexo 8.

RT-PCR universal para paramixovirus. Para la identificación/confirmación de paramixovirus se utilizaron dos pares de iniciadores degenerados. PAR-F1 y PAR-R para los virus de la subfamilia *Paramyxovirinae* y PNE-F1 y PNE-R para los virus de la subfamilia *Pneumovirinae* (Tong y col., 2008). Las secuencias de los iniciadores, así como el gen y tamaño de la región amplificada se mencionan en el Cuadro 3. El procedimiento de RT-PCR para la identificación/confirmación universal de paramixovirus se muestra en el Anexo 9.

RT-PCR para virus de hepatitis E. La identificación/confirmación del virus de hepatitis E se realizó mediante una RT-PCR anidada, utilizando los iniciadores degenerados HEv-F-1679 y HEv-R-1680 en la primera reacción de RT-PCR. En la PCR anidada se emplearon los iniciadores Hev-F-1681 y Hev-R-1682 (La Rosa y col., 2014) (Cuadro 3). El procedimiento de RT-PCR para la identificación de virus de hepatitis E se muestra en el Anexo 10.

RT-PCR para alfainfluenzavirus. Para la identificación/confirmación de alfainfluenzavirus en las muestras de porcinos se utilizaron los iniciadores IA-M-sen e IA-M-rev (diseñados por Barrón-Rodríguez y col. (2018) mediante el análisis de 71 secuencias de referencia para la identificación de virus de aves y porcinos) (Cuadro 3). El procedimiento de RT-PCR para la identificación de alfainfluenzavirus se muestra en el Anexo 11.

PCR para Herpesvirus. Para la identificación/confirmación de herpesvirus se utilizó una PCR anidada (VanDevanter y col., 1996) utilizando los iniciadores degenerados DFA (sentido), ILK (sentido) y KGI (antisentido) en la primera

reacción. En las reacciones de la PCR anidada se emplearon los iniciadores TGV (sentido) e IYG (antisentido) (Cuadro 3). El procedimiento de PCR para la identificación /confirmación de herpesvirus se muestra en el Anexo 12.

PCR para el virus tipo Circovirus porcino 51. La identificación/confirmación del virus tipo circovirus porcino 51 se realizó mediante una PCR utilizando iniciadores específicos diseñados en primer-BLAST por Rodrigo J. Barrón Rodríguez, empleando la secuencia de referencia NC_025684 del Po-Circo-like virus 51. Los iniciadores empleados fueron F8b_Po-circo-like-virus-51_1760 y R9_Po-circo-like-virus-51_2483 (antisentido) (Cuadro 3). El procedimiento de PCR para la identificación del Circovirus porcino 51 se muestra en el Anexo 13.

Filogenias

Los contigs más largos de los virus animales más abundantes en este estudio, fueron seleccionados para realizar alineamientos múltiples con la herramienta Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/). Las secuencias de referencia usadas en los alineamientos fueron obtenidas del Gen Bank. Las secuencias de los alineamientos fueron editadas ocupando el software BioEdit v7.0.4 (disponible en <http://www.mbio.ncsu.edu/bioedit/bioedit.html>). Los alineamientos múltiples editados fueron empleados en MEGAX (Kumar y col., 2018) para construir los árboles filogenéticos empleando el método de Neighbour-joining (Saitou y Nei, 1987) (realizado con 1000 replicas de bootstrap) y las distancias evolutivas fueron calculadas utilizando el método de máxima verosimilitud compuesta (Tamura y col., 2004). Este análisis fue utilizado para identificar los grupos taxonómicos virales a los cuales pertenecen los contigs, usando criterios de demarcación del ICTV (International Committee on Taxonomy of Viruses). Las secuencias de los contigs de virus animales más abundantes fueron reportadas a la base de datos BankIt del GenBank (NCBI-NIH).

Cuadro 3. Información de los iniciadores (primers) utilizados para la identificación/confirmación de virus animales mediante ensayos de PCR o RT-PCR.

Fuente	Iniciador	Familia o Género viral	Gen objetivo	Producto
De Souza Luna y col., 2007	PC2S2-RJ_F: TNA TGG GNT GGG AYT AYC C	Coronaviridae	ORF1ab (polimerasa)	603 pb
Lelli y col., 2013	IZS-CoV-f-RJ-r: TGN GAR CAR AAY TCR TGH GG			
Tong y col., 2008	PAR-F1: GAR GGN TAN TGY CAD AAR NTN TGG AC PAR-R: GCT GAA GTT ACH GGH TCH CCD ATR TTN C	Paramyxovirinae	L	656 pb
	PNE-F1: GYG TDG GTA GDA TGT TYG CNA TGC ARC C PNE-R: DGT CCA SAA NTT YTG RCA CCA NCC YTC	Pneumovirinae	L	487 pb
La Rosa y col., 2014	HEv-F-1679: CCA YCA GTT YAT HAA GGC TCC HEv-R-1680: TAC CAV CGC TGR ACR TC	Virus Hepatitis E	ORF 1	348 pb
	Hev-F-1681: CTC CTG GCR TYA CWA CTG C Hev-R-1682: GGR TGR TTC CAI ARV ACY TC			172 pb
Barrón-Rodríguez y col. 2018	IA-M-sen: GAG TCT TCT AAC CGA GGT CG IA-M-rev: GGA TCA CTT GAA TCG CTG C	Alfainfluenzavirus	M	741 pb
Vandervanter y col., 1996	DFA: GAY TTY GCN AGY YTN TAY CC ILK: TCC TGG ACA AGC AGC ARN YSG CNM TNA A KG1: GTC TTG CTC ACC AGN TCN ACN CCY TT	Herpesvirus	DNA Polimerasa	739 pb
	TGV: TGT AAC TCG GTG TAY GGN TTY ACN GGN GT IYG: CAC AGA GTC CGT RTC NCC RTA DAT			250 pb
Barrón Rodríguez*	F8b_Po-circo-like-virus-51_1760: CGC AAC TAC TAT TCG CGT CG R9_Po-circo-like-virus-51_2483: CGA ACG AGT GAG GGG GAA TC	Virus tipo circovirus porcino 51	CAP	723 pb

*Iniciadores específicos para amplificar el Po-circo-like virus 51 (NC025684) diseñados en este estudio empleando la herramienta primer-BLAST (NCBI).

RESULTADOS

Datos de secuenciación.

El total de lecturas crudas con calidad suficiente de cada especie fue de 585,718 en *Artibeus* sp, 914,710 de *M. waterhousii*, 386,490 de *P. parnellii* y 2,736,522 en porcinos. Entre estas el porcentaje de lecturas que no mostró similitud con secuencias conocidas reportadas en las bases de datos nt y wgs del Centro Nacional de Información Biotecnológica (NCBI) fue de 95.0% (556,432) de *Artibeus* sp, 95.5% (873,548) de *M. waterhousii*, 94.1% (363,687) de *P. parnellii* y 84.8% de porcinos (2,307,018). Así 28,593 (4.9%) lecturas de *Artibeus* sp, 39,743 secuencias (4.4%) de *M. waterhousii*, 22,290 (5.9%) de *P. parnellii* y 80,863 (3%) de porcinos mostraron homología con secuencias reportadas de organismos procariontas o eucariotas. Se obtuvieron 489 secuencias (0.1%) de *Artibeus* sp, 882 (0.1%) de *M. waterhousii*, 353 (0.1%) de *P. parnellii* y 322,961 secuencias (11.9%) de porcinos asociadas a virus. Entre las secuencias virales, en *Artibeus* sp 480 se asociaron con fagos y diez con virus animales; en *M. waterhousii* 858 de las lecturas se relacionaron con fagos y 23 con virus animales; en *P. parnellii* hubo 312 asociadas con fagos, 46 con virus animales y una con virus de protozoos; y en porcinos 312,646 secuencias se relacionaron con fagos, 83 con virus de protozoarios, 10,212 con virus animales y dos con virus vegetales (Figura 2).

Virus identificados por secuenciación masiva

Las secuencias de *Artibeus* sp. representaron virus animales de tres familias (*Circoviridae*, *Retroviridae* y *Herpesviridae*) y de ocho géneros de bacteriofagos. En *M. waterhousii* se encontraron secuencias relacionadas con cinco familias de virus animales (*Circoviridae*, *Retroviridae*, *Reoviridae*, *Parvoviridae* y *Herpesviridae*) y ocho géneros de fagos. En *P. parnellii* se identificaron secuencias de cuatro familias de virus animales (*Coronaviridae*, *Retroviridae*, *Circoviridae* y *Hepeviridae*), ocho géneros de fagos y una especie de virus de alga. Mientras que las lecturas de los porcinos se asociaron con diez familias de virus animales (el

99.7% pertenecía a las familias *Circoviridae*, *Caliciviridae*, *Parvoviridae* y *Astroviridae*), una de plantas, cuatro especies de virus de protozoarios y 18 géneros de fagos (Cuadro 4).

Con respecto a las lecturas de bacteriófagos, los géneros más abundantes en *Artibeus* sp. fueron fago de *Shigella pSf1* (especie), *Jerseyvirus* y *Kaggunavirus* de la familia *Shiphoviridae* (47.7%), *Msw3virus* y *Tequatrovirus* de la familia *Myoviridae* (24.8%) y *Kp34virus*, *Teseptimavirus* y *Phikmvirus* de la familia *Podoviridae* (16%). En *M. waterhousii*, los géneros más abundante de fagos representados fueron *Tunavirus*, *Kagunavirus* y *Jerseyvirus* de la familia *Siphoviridae* (58.1%), *Tequatrovirus*, *Vequintavirus* y *Ms3virus* de la familia *Myoviridae* (25.2%) y *Teseptimavirus* y *Phi29virus* de la familia *Podoviridae* (8%). En *P. parnellii*, los géneros de virus más representados fueron *Tunavirus*, *Kagunavirus* y *Jerseyvirus* de la familia *Siphoviridae* (59.3%), *Tequatrovirus* y *Vequintavirus* de la familia *Myoviridae* (20.5%) y *Tseptimavirus* y *Phi29virus* de la familia *Podoviridae* (12.4%). Y en las lecturas de porcinos los géneros más representados (99.8% de las secuencias asignadas a fagos) fueron *Tunavirus* (23.5%), *Kagunavirus* (21.6%), *Jerseyvirus* (19.0%), *Tequatrovirus* (12.1%), *Teseptimavirus* (5.1%), *Yokohamavirus* (4.7%), *Vequintavirus* (2.8%), *Roufvirus* (1.6%), *Salasvirus* (1.1%) y *Tequintavirus* (1.0%) (Cuadro 4).

Diversidad viral

Las rarefacciones mostraron que los resultados obtenidos (lecturas/secuencias) por secuenciación masiva fueron representativos de la diversidad de la comunidad viral de cada especie muestreada (Figura 3). El índice de Shannon-Wiener se utilizó para evaluar la diversidad de especies de virus totales y virus animales en el conjunto de datos de lecturas de cada especie. Los índices de diversidad viral total para *Artibeus* sp, *M. waterhousii*, *P. parnellii* y los porcinos fueron de 3.14, 2.98, 3.19 y 2.84 respectivamente para cada especie. En el mismo orden, los valores de diversidad de virus animales fueron de 1.36, 1.17, 1.7 y 1.1, respectivamente.

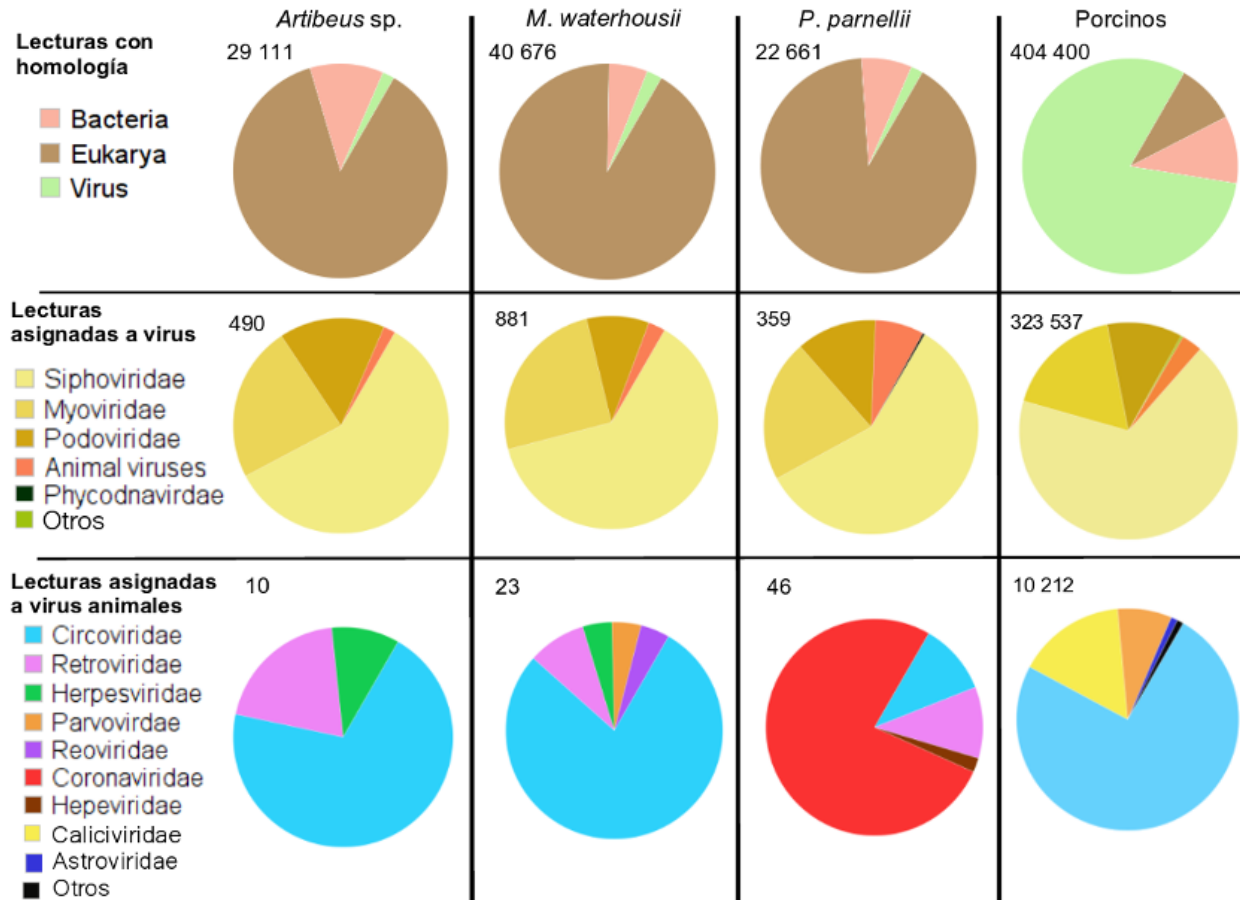


Figura 2. Resumen de la clasificación taxonómica de las secuencias obtenidas por secuenciación masiva a partir de muestras de las tres especies de murciélagos y los porcinos de traspatio. Las secuencias se asignaron a diferentes rangos taxonómicos por su homología con secuencias de referencia de las bases de datos del GenBank como resultado del BLAST.

Cuadro 4. Virus más abundantes identificados por análisis de homología (BLASTn y BLASTx) de los conjuntos de datos (lecturas) de cada especie estudiada.

Familia viral (lecturas)	Especies de Virus (lecturas)
<i>Artibeus sp</i>	
Virus animales	
<i>Circoviridae</i> (7)	Po-Circo-like-virus 51 (5); Virus de DNA circular asociado a heces de focas (1); Virus DNA circular asociado a heces de porcino (1).
<i>Herpesviridae</i> (1)	Herpesvirus de gallidos 2 (1)
<i>Retroviridae</i> (2)	Retrovirus de mono ardilla (2)
Bacteriofagos	
<i>Myoviridae</i> (131)	<i>Tequatrovirus*</i> (63); Fago de <i>Klebsiella</i> JD001 (57)
<i>Podoviridae</i> (92)	Fago de <i>Klebsiella</i> KP34 (34); <i>Teseptimavirus*</i> (26); <i>Phikmvlikevirus*</i> (17)
<i>Siphoviridae</i> (257)	Fago de <i>Shigella</i> pSf1-1 (90); <i>Jerseyvirus*</i> (68); <i>Kagunavirus*</i> (71)
<i>Macrotus waterhousii</i>	
Virus animales	
<i>Circoviridae</i> (18)	Po-Circo-like-virus 51 (15); Virus de DNA circular asociado a heces de focas (2); Po-Circo-like-virus 22 (1)
<i>Herpesviridae</i> (1)	Herpesvirus humano (1)
<i>Parvoviridae</i> (2)	Blatella germanica densovirus (2)
<i>Reoviridae</i> (1)	Aquareovirus A (1)
<i>Retroviridae</i> (1)	Retrovirus de oveja Jaagsiekte (1)
Bacteriofagos	
<i>Myoviridae</i> (220)	<i>Tequatrovirus*</i> (130); <i>Vequintavirus*</i> (63); Klebsiella phage JD001 (23)
<i>Podoviridae</i> (83)	<i>Teseptimavirus*</i> (55); <i>Phi29likevirus*</i> (14);
<i>Siphoviridae</i> (555)	Fago de <i>Shiegella</i> pSf1 (207); <i>Kagunavirus*</i> (165); <i>Jerseyvirus*</i> (127); <i>Pis4avirus*</i> (13); <i>Tequintavirus*</i> (11)
<i>Pteronotus parnellii</i>	
Animal viruses	
<i>Circoviridae</i> (5)	Po-Circo-like-virus 51 (4); Virus de ssDNA circular asociado a heces de cerdo (1)
<i>Coronaviridae</i> (35)	Coronavirus Mex_CoV-10 (6); CoV de murciélago Rousettus HKU9 (3); CoV relacionado con SARS (3); CoV de murciélago Zaria (2); Coronavirus aviar (1); coronavirus de murciélago BtCoV/133/2005 (1); Coronavirus del síndrome respiratorio de Medio Oriente (1); CoV de murciélago <i>Miniopterus</i> HKU8 (1)
<i>Retroviridae</i> (5)	Retrovirus endógeno humano (4); Pararetrovirus endógeno de cítricos (1).
<i>Hepeviridae</i> (1)	Virus de hepatitis E (1)
Bacteriofagos	
<i>Myoviridae</i> (73)	<i>Tequatrovirus*</i> (48); <i>Vequintavirus*</i> (15)
<i>Podoviridae</i> (41)	<i>Teseptimavirus*</i> (28); <i>Phi29likevirus*</i> (10)
<i>Siphoviridae</i> (198)	Fago de <i>Shigella</i> pSf1-1 (85); <i>Kagunavirus*</i> (56); <i>Jerseyvirus*</i> (41)
Virus de algas	
<i>Phycodnaviridae</i> (1)	

Cuadro 4. Continuación

Familia viral (lecturas) Especies de Virus (lecturas)

Porcinos

Virus animales

<i>Astroviridae</i> (99)	Astrovirus de cerdo salvaje/WBAstV-1/2011/HUN (17); Mamastrovirus 3 (7); astrovirus porcino 5 (1)
<i>Caliciviridae</i> (1605)	Virus de Sapporo (1461); sapovirus porcino (144);
<i>Circoviridae</i> (7612)	Po-Circo-like virus 51 (6781); Virus de DNA circular asociado a heces de focas (176); Virus de ssDNA circular asociado a heces de cerdo (119); Po-Circo-like virus 22 (46); virus circular asociado a heces de porcino (26); virus circular asociado a heces de porcino 5 (14)
<i>Herpesviridae</i> (1)	Herpesvirus bovino 4 (1)
IAS virus* (67)	Virus de heces asociado a inmunodeficiencia humana (67)
<i>Orthomyxoviridae</i> (2)	Influenza A virus (2)
<i>Parvoviridae</i> (808)	Bocavirus porcino 3 (608); Bocavirus pig/6V/China/2010 (45); virus adeno-asociado (20); Bocavirus pig/7V/China/2010 (10); bocavirus porcino (5); virus adeno-asociado-2 (4); virus adeno-asociado-Po1 (3); miniambidensovirus Acheta domesticus (2); virus adeno-asociado de primates no humanos (2); bocavirus porcino 2 (2); Bufavirus-1 humano (1); bocavirus de zorra (1); bocavirus humano (1); bocavirus porcino 4 (1); bocavirus porcino 5 (1); bocavirus porcino 6V7V (1)
<i>Picobirnaviridae</i> (1)	Human picobirnavirus (1)
<i>Picornavirus*</i> (33)	Posavirus 1 (18); sapelovirus porcino (10); Enterovirus G (2); Enterovirus B (1); Tescho virus porcino (1)
<i>Poxviridae</i> (1)	Virus de la viruela (1)
<i>Retroviridae</i> (1)	Lentivirus 1 de inmunodeficiencia humana (1)

Bacteriofagos

<i>Inoviridae</i>	Inovirus* (2)
<i>Microviridae</i> (41)	Microvirus* (2) Spiromicrovirus (1)
<i>Myoviridae</i> (56224)	Tequatrovirus* (7424), Twortlikevirus* (593), Spounalikevirus* (6), Felixounalikevirus* (2), Viunalikevirus* (5)
<i>Podoviridae</i> (36158)	Teseptimavirus* (15976), Phi29likevirus* (3321), Sp6likevirus* (1305), Ahjdlikevirus* (291), 323 N4likevirus* (46), P22likevirus* (10), Phikmvlikevirus* (10)
<i>Siphoviridae</i> (220 221)	Tequintavirs* (3240), Tunavirus* (83), Lambda likevirus* (12)

Virus de protozoarios

<i>Mimiviridae</i> (83)	Megavirus courdo (7) mimivirus de Acanthamoeba castellanii (2), Moumouvirus (2), Megavirus chiliensis (1)
-------------------------	---

* género viral

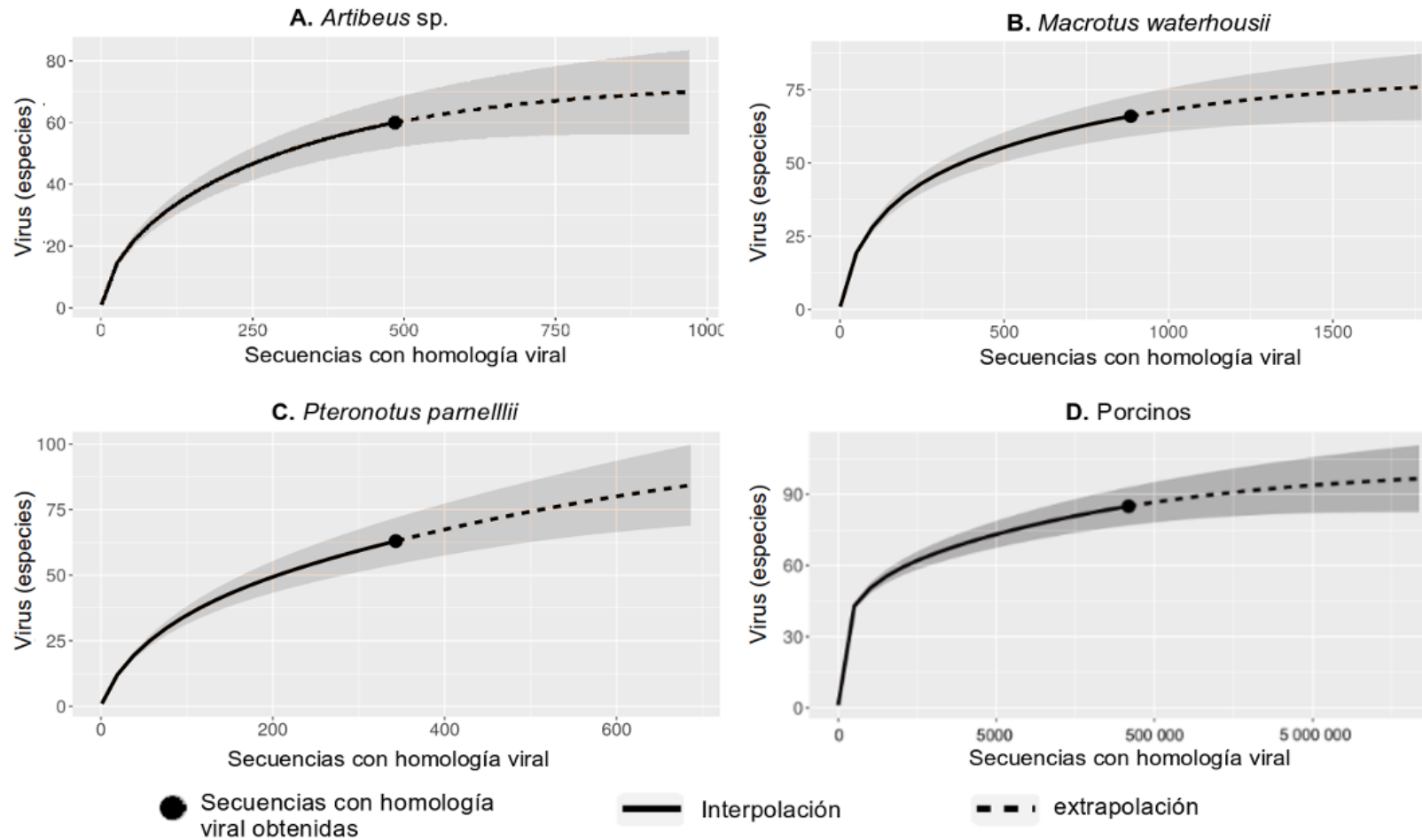


Figura 3. Diversidad de especies de virus estimada por rarefacción para los datos de cada especie de murciélago y porcinos (●). Este gráfico muestra que las especies virales identificadas por metagenómica proporcionaron una buena representación de la diversidad estimada (interpolación/extrapolación) en todas las especies.

Identificación molecular de virus animales

Las muestras fueron negativas para los ensayos de identificación molecular para herpesvirus, virus de hepatitis E, alfa-influenzavirus, paramyxovirus (subfamilias *Paramyxovirinae* y *Pneumovirinae*) y coronavirus. En la PCR de confirmación para po-circo-like virus 51, 14 muestras (ocho porcinos, dos *M. waterhousii* y cuatro *P. parnellii*) amplificaron levemente. Debido a la pérdida de estos productos durante el proceso de purificación (electroforesis), la secuenciación no se realizó. La generación de más material por PCR no fue posible debido a que el DNA de las muestras se agotó.

Ensamblaje y análisis filogenéticos

Se obtuvieron 627 contigs con homología en los datos de *Artibeus* sp, de los cuales dos mostraron identidad con virus (familias *Myoviridae* y *Siphoviridae*). En los datos de *M. waterhousii* se identificaron 608 contigs con homología, siete asignados a virus (*Siphoviridae* 4, *Myoviridae* 2, *Circoviridae* 1). En *P. parnellii* hubo 606 contigs asignados a secuencias reportadas, incluyendo cuatro asignadas a virus (*Siphoviridae* 2, *Coronaviridae* 1, *Retroviridae* 1). Con los datos de los porcinos 2665 contigs mostraron homología, entre los cuales 2083 (78,2%) estaban relacionados con bacteriofagos (97% correspondientes a caudovirales [*Podoviridae* 952, *Siphoviridae* 622 y *Myoviridae* 423] y 3% a virus animales [*Parvoviridae* 33, *Circoviridae* 14, *Caliciviridae* 6, *Astroviridae* 4, *Picornaviridae* 1]). Los resultados de BLAST de los contigs coincidieron con los resultados del BLAST de lecturas, identificando una menor proporción de virus en el de *contigs*.

Como resultado de los ensamblajes de los datos de los porcinos se logró obtener el genoma parcial de un virus tipo circovirus porcino (>60%). Otros contigs representativos de los virus animales más abundantes en porcinos fueron encontrados en este estudio (circovirus, sapovirus y parvovirus). En las tres especies de murciélagos se encontraron contigs con similitud (nt: similitud nucleotídica, aa: similitud de aminoácidos) a circovirus porcinos (nt ≤90%, aa

≤94%). Mientras que en *P. parnellii* se encontraron dos contigs similares a betacoronavirus (KC633194 BtCov/KCR155). Estas secuencias fueron registradas en la base de datos BankIt del GenBank (Cuadro 5) y fueron empleadas en la construcción de filogenias. Para la construcción de las filogenias se obtuvieron secuencias similares conocidas y de referencia de los virus de las familias *Caliciviridae* (nt y aa ≤89%), *Circoviridae* (nt ≤98%, aa ≤97.7%), *Parvoviridae* (nt ≤96%, aa ≤98.7%) y del género *Betacoronavirus* (nt ≤98.4%, aa 100%) del banco de genes (NCBI - NIH) para realizar el alineamiento de secuencias múltiples y la filogenia con su respectivo contig. El contig similar al gen de la cápside de circovirus porcinos, que se encontró en los porcinos de este trabajo, mostró mayor similitud con circovirus identificados previamente en porcinos, pero fue diferente a los circovirus porcino 1 y circovirus porcino 2 (Figura 4). Otro contig (963 pb) encontrado en los porcinos de este trabajo similar al gen CAP de circovirus mostró la mayor similitud con la secuencia de un virus encontrado en otáridos y porcinos, distante de los circovirus porcino 1 y 2 (Figura 5). El contig de los porcinos (1044 pb) similar al gen de la poliproteína de los virus de la familia *Caliciviridae* fue más similar a los sapovirus porcinos (Figura 6). En cuanto al contig empleado en la filogenia de virus de la subfamilia *Parvovirinae*, obtenido a partir de los datos de porcinos, este fue más similar a bocaparvovirus de ungulados 5 (Figura 7), virus típicamente encontrados en porcinos. Los contigs similares al gen de la replicasa de circovirus, obtenidos en las tres especies de murciélagos, se agruparon con secuencias de circovirus y virus tipo circovirus porcinos, apartados de otros circovirus (Figuras 8, 9 y 10). El contig encontrado en *P. parnellii* homólogo a betacoronavirus fue más similar a secuencias de coronavirus no clasificados del subgénero *Embecovirus* (Figura 11).

En las cuatro especies se identificaron lecturas homólogas a circovirus porcinos (gen replicasa), entre las que se encontró un fragmento de 83 bases común a

todas las especies con el cual se elaboró un árbol filogenético (Figura 12). Dichas secuencias mostraron similitudes de 82.4% a 95.7%.

Cuadro 5. Contigs (secuencias) de los virus más abundantes en porcinos y murciélagos.

Nombre de la secuencia	Accesión asignada	Longitud del contig	RefSeq/secuencia del GenBank con mayor similitud	Similitud	Gen	Covertura en el Genoma
Po-Circo-like-virus/swine/Mex/2016/1-2	MH490914	369 pb		98 %	Rep	13.0 %
	MH490915	309 pb	Po-circo-like virus 51 NC 025684	92.2%	Prot 1 y 2	10.9%
	MH490916	422 pb		85.3%	Prot 2	14.9%
FSFA-CirV/swine/Mex/2016/Cap_partial-CDS	MH490912	963 pb	Fur seal feces associated circular DNA virus LC133373.1	95.1%	Cap	33.0 %
	MH490913	521 pb		95.4%	Rep	17.8%
Sapovirus/swine/Mex/2016/ORF1-polyprotein	MH490911	1044 pb	Porcine sapovirus strain JJ259 KT922089	89%	ORF-1	15.4 %
Bocaparvovirus_Swine/Mex/2016/NS1	MH479904	1268 pb		96.3 %	NS1	26.9 %
	MH479905	665 pb	Porcine bocavirus 3 strain 22 JF713714	96.1%	NP1	14.1%
	MH479906	762 pb		97.8%	NP1	16.2%
	MH479907	361 pb		75.6%	VP2	7.7%
Circovirus/swine/MEX/2016/Rep-partialCDS	MW584369	273 pb	Po-circo-like virus 51 NC025684	90.5 %	Rep	8.7%
Circovirus/Aspp/MEX/2016/Rep-partial-CDS	MW584366	162 pb	Circovirus PoCirV_VIRES_NX02_C7 MK377683	93 %	Rep	18.3% *
Circovirus/Mwat/MEX/2016/Rep-partial-CDS	MW584367	252 pb	Circovirus sp isolate PoCirV_VIRES_LN01_C1 MK377656	96.4 %	Rep	26.6% *
Circovirus/Ppar/MEX/2016/Rep-partial-CDS	MW584368	181 pb	Circovirus sp isolate PoCirV_VIRES_LN01_C1 MK377656	91.7 %	Rep	19.1% *
B-Coronavirus_Ppar/MEX/2016_RNApoI_CDS-1 y 2	MW356910	482 pb	Betacoronavirus BtCoV/KCR180/ Pte_par/	94 %	ORF1ab	59% *
	MW356911	201 pb	CRC/2012 KC633193	93 %		24.6% *

*Porcentaje de covertura del contig con respecto a la longitud de la secuencia referencia con mayor homología.

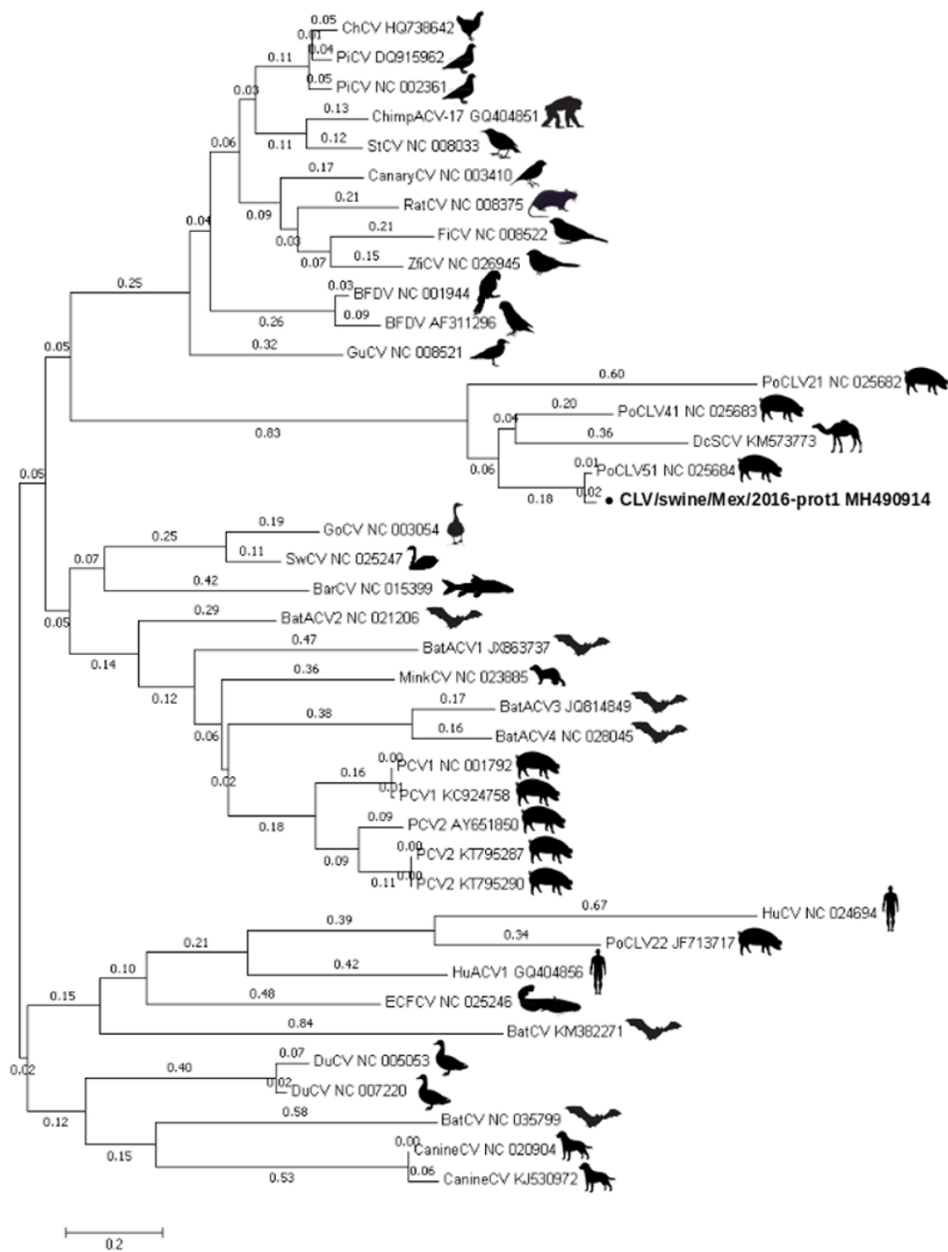


Figura 4. Árbol filogenético basado en secuencias de nucleótidos del gen de la proteína-1 (cápside) de especies del género Circovirus, que muestra la relación evolutiva de un contig de 369 bases encontrado por secuenciación masiva (negrita) en los cerdos de traspatio rural estudiados. La filogenia se infirió usando el método de neighbor-joining, y las distancias evolutivas se calcularon usando el método de máxima verosimilitud. Los análisis evolutivos se realizaron en MEGAX.

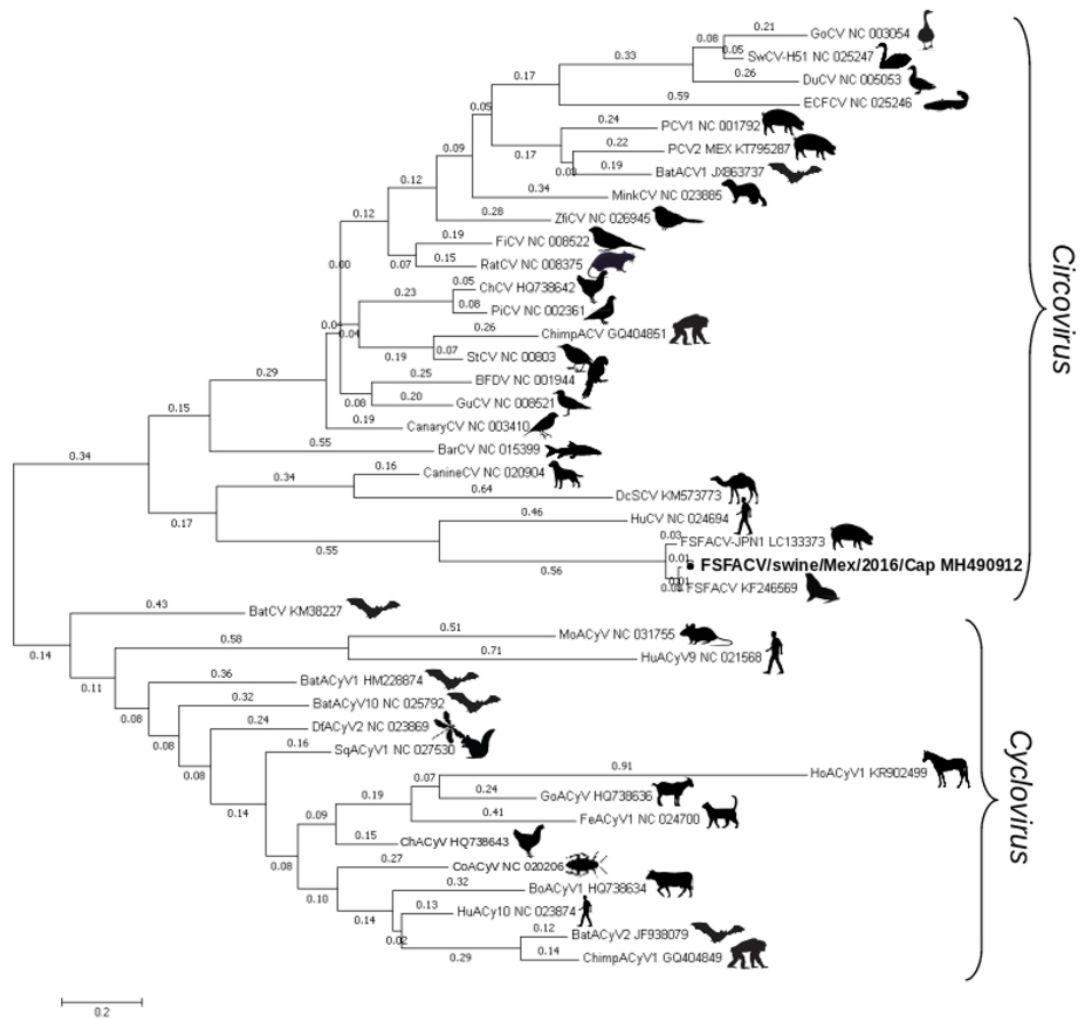


Figura 5. Árbol filogenético basado en el gen Cap de especies de la familia Circoviridae (géneros Circovirus y Cyclovirus), que muestra la relación evolutiva del contig de circovirus asociado a lobos marinos (963 bases) encontrado por secuenciación masiva (negrita) en porcinos de traspatio rural. La filogenia se infirió usando el método de neighbor-joining, y las distancias evolutivas se calcularon usando el método de máxima verosimilitud compuesta. El análisis incluyó 40 secuencias de nucleótidos de referencia del virus asignado a los géneros Circovirus y Cyclovirus. Los análisis evolutivos se realizaron en MEGAX.

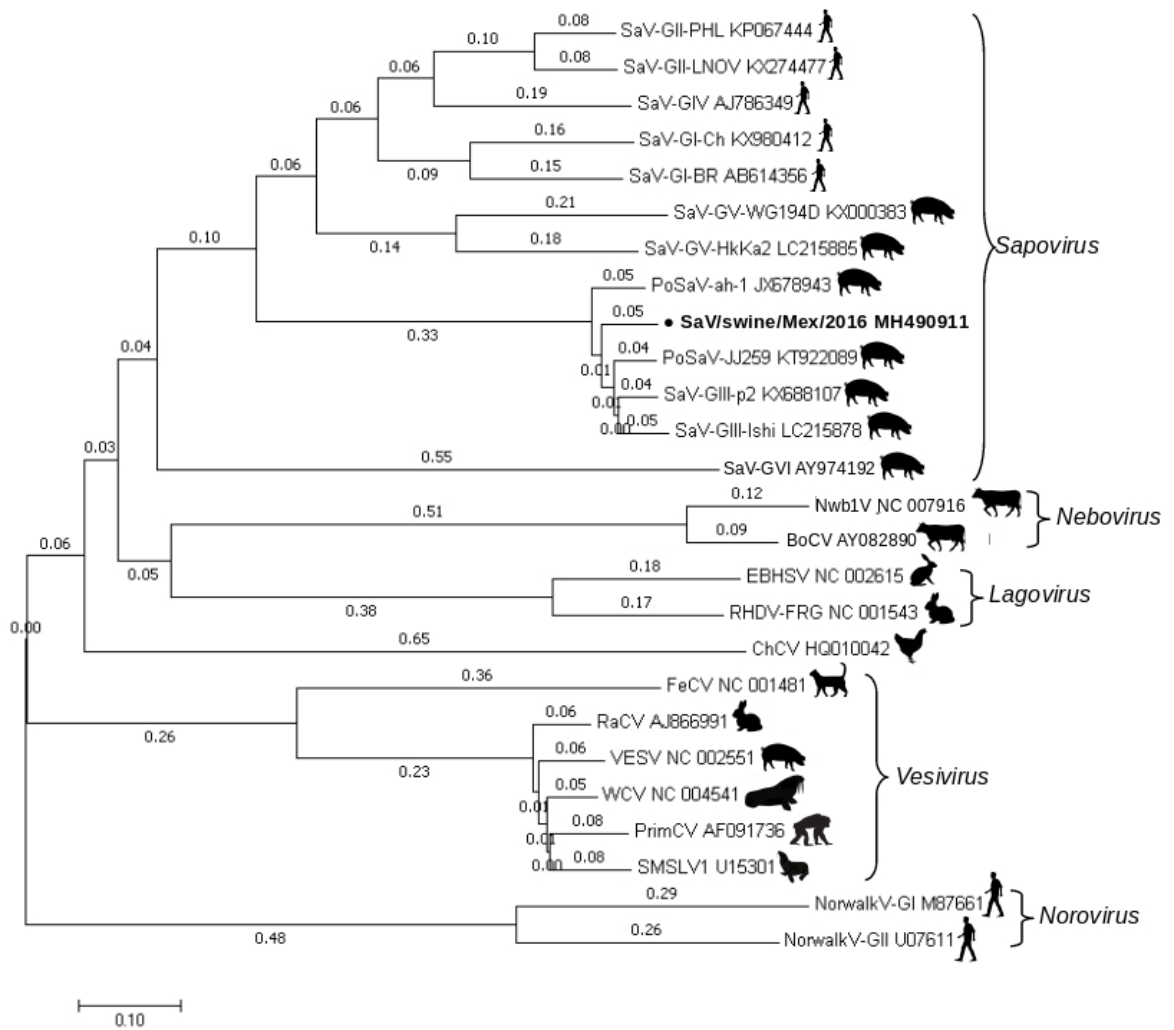


Figura 6. Árbol filogenético basado en secuencias de nucleótidos del gen Cap de la familia Caliciviridae, que muestra la relación evolutiva de un contig de 1044 bases similar al género Sapovirus encontrado por secuenciación masiva (negrita) en cerdos de traspatio rural. La filogenia se infirió usando el método de unión de vecinos (neighbor-joining), y las distancias evolutivas se calcularon usando el método de máxima probabilidad compuesta. Los análisis evolutivos se realizaron en MEGAX.

BoV(ParvoV)-NJ

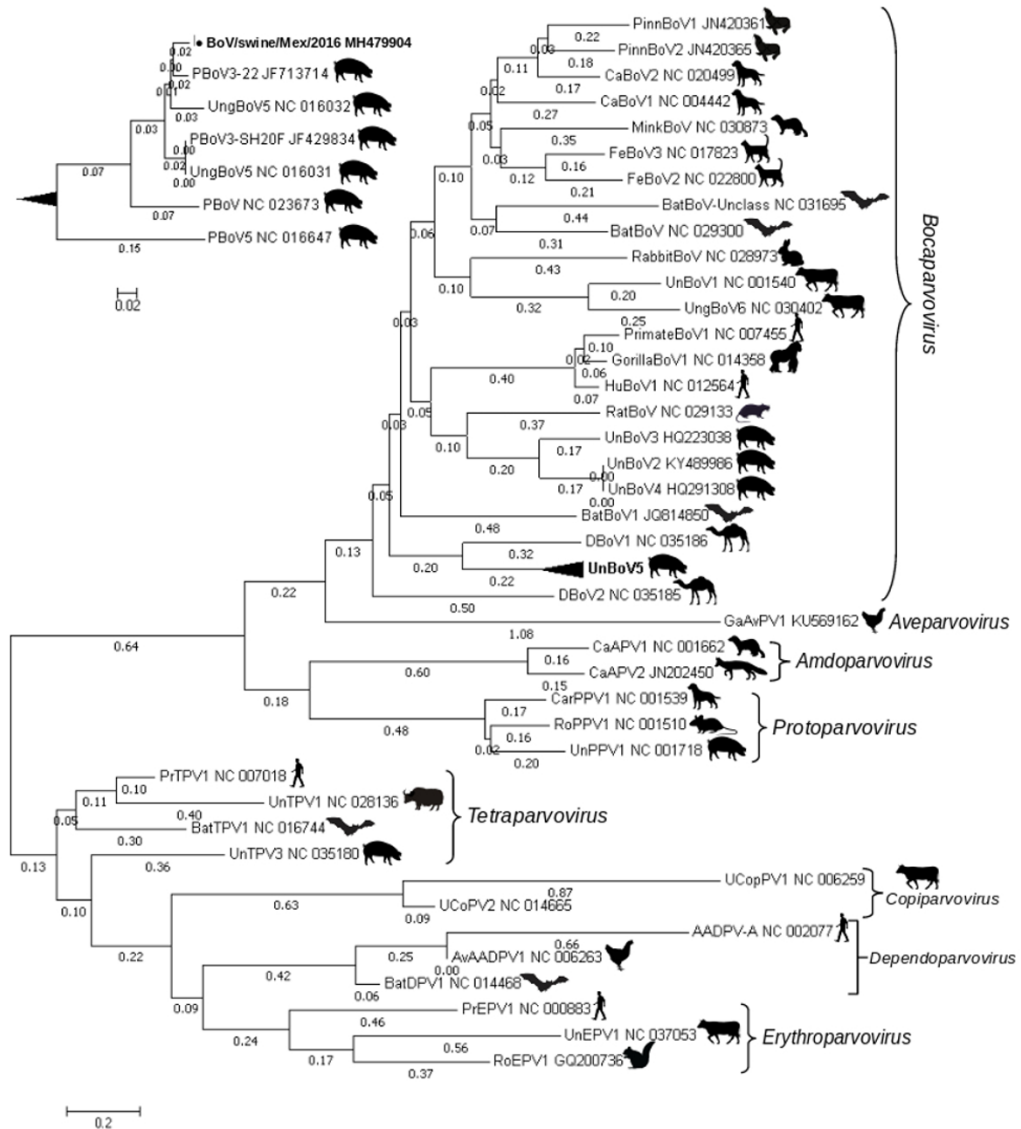


Figura 7. Árbol filogenético basado en el gen NS1 de las especies de la subfamilia Parvovirinae (géneros Amdoparvovirus, Avedparvovirus, Bocaparvovirus, Copiparvovirus, Dependoparvovirus, Erythroparvovirus, Protoparvovirus y Tetraparvovirus) que muestran la relación evolutiva con la secuencia de bocaparvovirus identificada en los porcinos de traspatio (negrita). La filogenia se infirió usando el método de unión de vecinos (neighbor-joining), y las distancias evolutivas se calcularon usando el método de máxima verosimilitud compuesta. Los análisis evolutivos se realizaron en MEGAX. La rama que involucra el bocaparvovirus 5 unglado (incluido el contig secuenciado en este estudio) se muestra en la esquina superior izquierda de la figura.

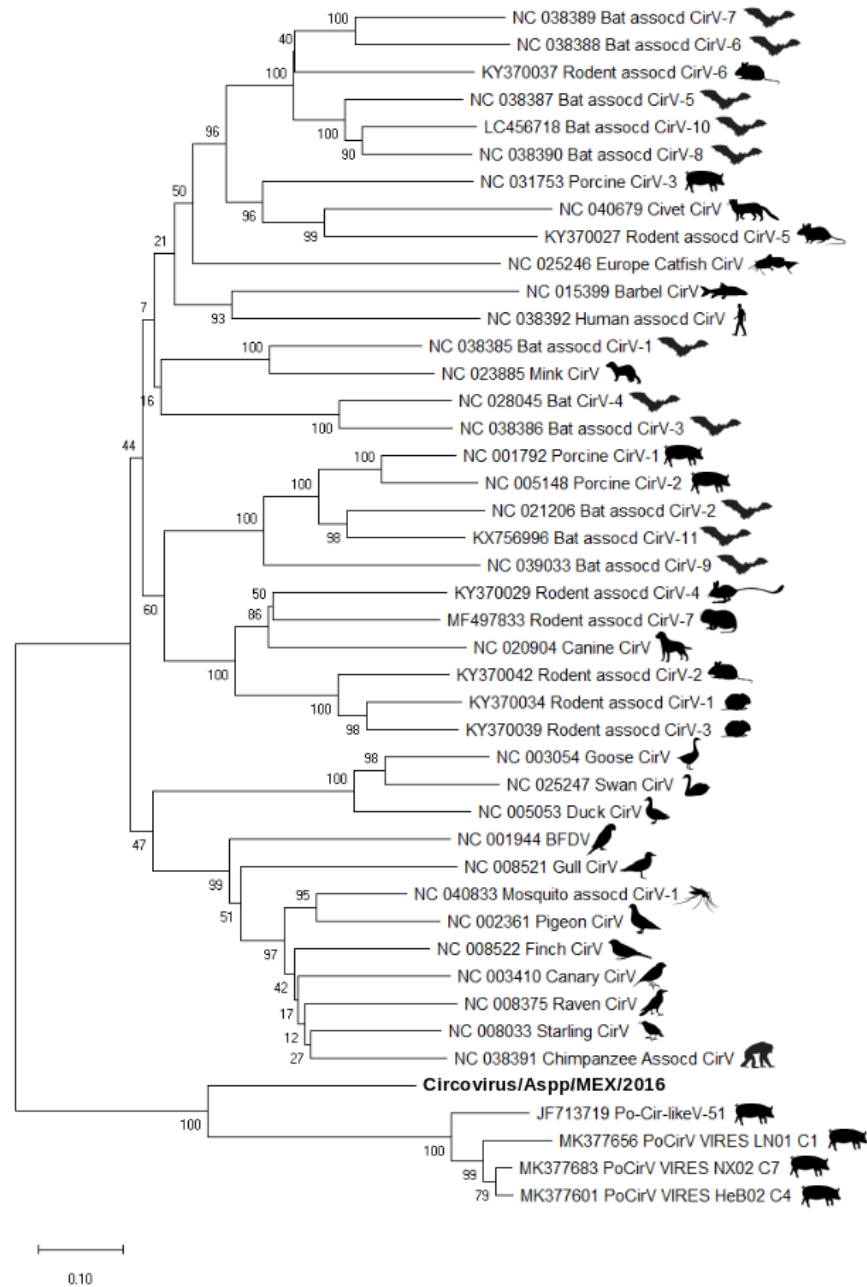


Figura 8. Filogenia basada en secuencias parciales del gen de la replicasa (Rep) de 43 especies del género *Circovirus*, que muestra la relación del contig (162 pb) encontrado por secuenciación masiva (negrita) en los datos de murciélagos *Artibeus* sp. La filogenia se infirió usando el método de neighbor-joining, y las distancias evolutivas se calcularon usando el método de máxima verosimilitud. Los análisis se realizaron en MEGAX.

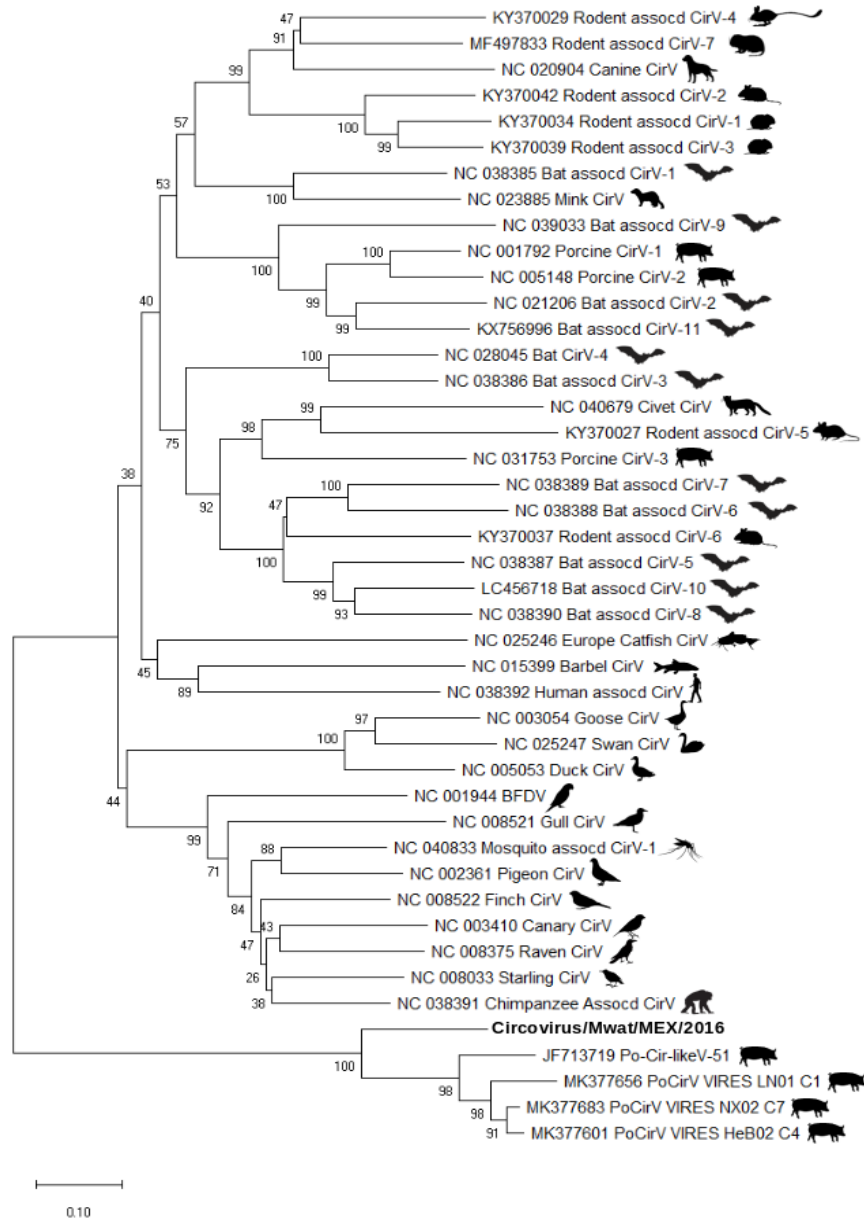


Figura 9. Filogenia basada en secuencias parciales del gen de la replicasa (Rep) de 43 especies del género *Circovirus*, que muestra la relación taxonómica del contig (252 pb) encontrado por secuenciación masiva (negrita) en los datos de murciélagos *M. waterhousii*. La filogenia se infirió usando el método de neighbor-joining, y las distancias evolutivas se calcularon usando el método de máxima verosimilitud. Los análisis se realizaron en MEGAX.

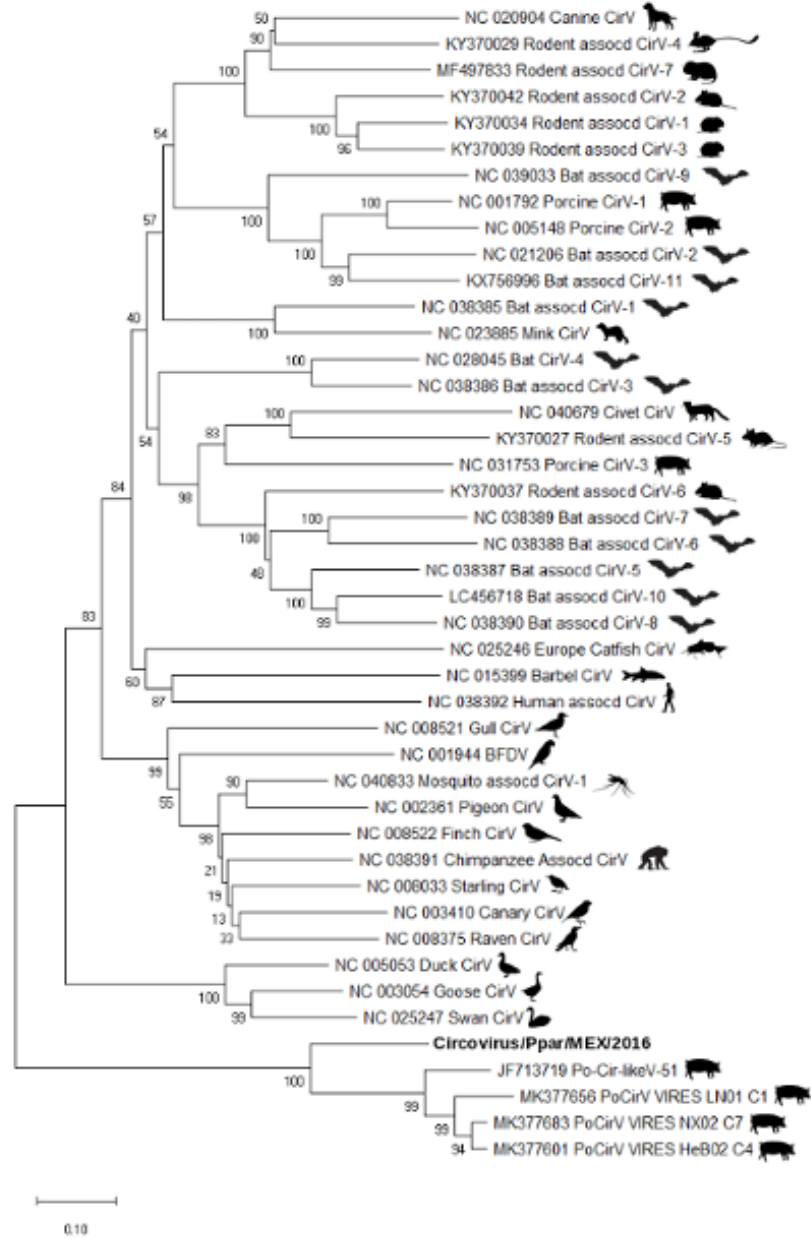


Figura 10. Filogenia basada en secuencias parciales del gen de la replicasa (Rep) de 43 especies del género *Circovirus*, que muestra la relación taxonómica del contig (181 pb) encontrado por secuenciación masiva (negrita) en los datos de murciélagos *P. parnellii*. La filogenia se infirió usando el método de neighbor-joining, y las distancias evolutivas se calcularon usando el método de máxima verosimilitud. Los análisis se realizaron en MEGAX.

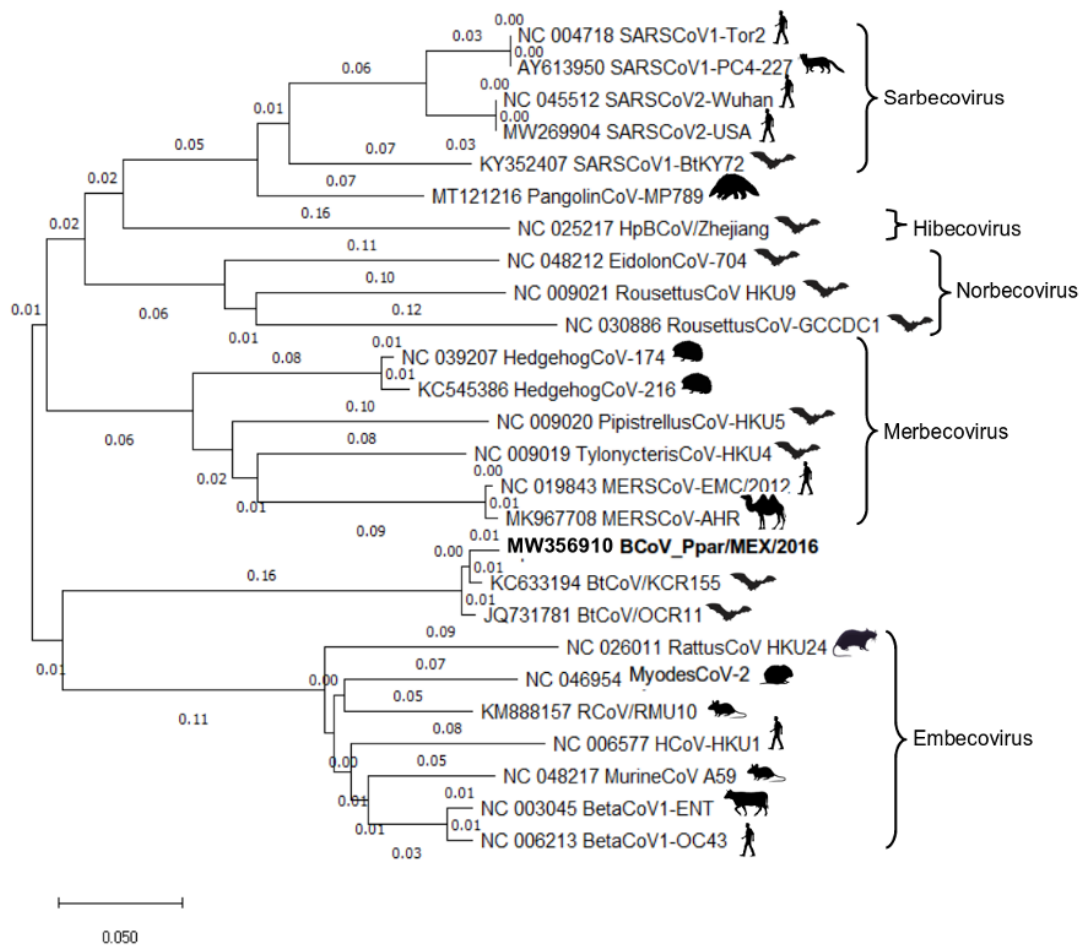


Figura 11. Filogenia sobre secuencias del gen de la RNA polimerasa dependiente de RNA (ORF1ab) de especies del género *Coronavirus*. Se infirió utilizando el método de Neighbor-Joining. Muestra la relación del contig de 482 (negrita) bases encontrado por secuenciación masiva en las lecturas de *P. parnellii*. Las distancias evolutivas se calcularon utilizando el método de máxima verosimilitud compuesta. Este análisis involucró 24 secuencias nucleotídicas de referencia. Los análisis evolutivos se realizaron en MEGA X.

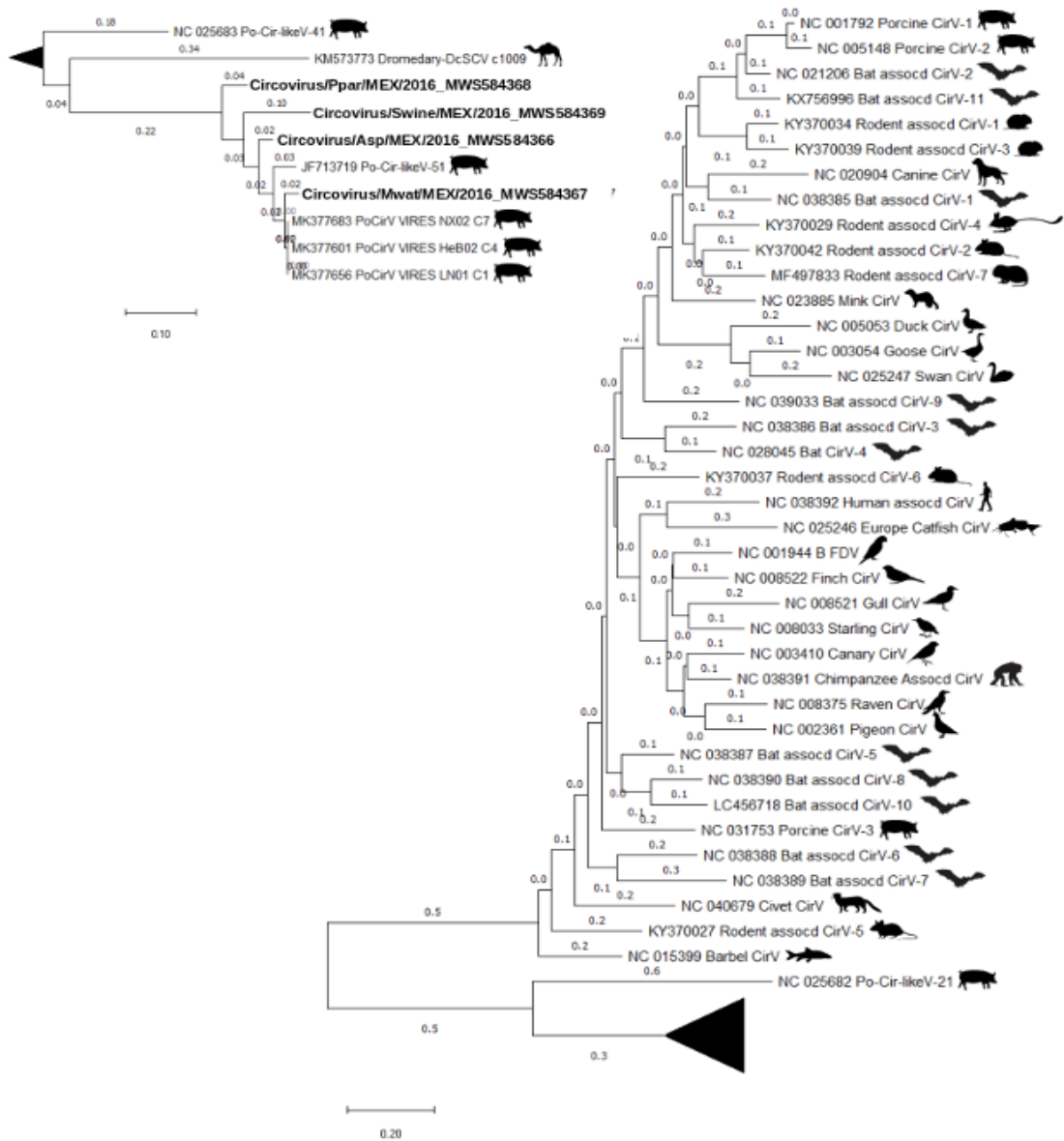


Figura 12. Filogenia basada en el gen de la replicasa (Rep) de especies del género *Circovirus*, que muestra la similitud de los contigs encontrados por secuenciación masiva (negrita) en los cerdos de traspatio rural y especies de murciélagos estudiados, usando una secuencia parcial de 83 pb común a todas las especies. La filogenia se infirió usando el método de neighbor-joining, y las distancias evolutivas se calcularon usando el método de máxima verosimilitud. Los análisis evolutivos se realizaron en MEGAX. La rama que involucra los contigs secuenciados en este estudio se muestra en la esquina superior izquierda de la figura.

DISCUSIÓN

En este estudio se analizó el viroma de murciélagos pertenecientes a la familia *Phyllostomidae* (*Artibeus* sp y *Macrotus waterhousii*), de una especie de la familia *Mormopidae* (*Pteronotus parnellii*) (Barrón-Rodríguez y col., en prensa) y de porcinos (Barrón-Rodríguez y col., 2021) que cohabitan en ambientes rurales del estado de Morelos donde el estilo de ganadería practicado es el traspatio (no intensivo). Las comunidades virales encontradas fueron diferentes a las reportadas previamente en porcinos y murciélagos (Donaldson y col., 2010; Shan y col., 2011; Sachsenröder y col., 2014; Salmier y col., 2017).

En las cuatro especies la mayor parte de la diversidad viral estuvo asociada a bacteriófagos (87 – 98% de las secuencias) presentando abundancias menores de lecturas asignadas a virus animales (12 – 2%). Este es el primer reporte de los viromas de porcinos de traspatio, ejemplares de las especies de murciélagos *Macrotus waterhousii*, *Pteronotus parnelli* y del género *Artibeus* sp. (donde la mayor proporción de individuos [n11] pertenecía a la especie *A. phaeotis*, con base en los criterios de identificación morfológica establecidos por Medellín y col. (2007).

Entre las secuencias virales, la predominancia de secuencias asociadas a bacteriofagos fue evidente en todas las especies, representando entre el 87% – 98% de las lecturas virales, siendo mayor en los datos de *Artibeus* sp. Las familias de bacteriófagos principalmente representadas fueron *Siphoviridae*, *Myoviridae* y *Podoviridae*. Las secuencias de fagos obtenidas muestran evidencia de la presencia de sus huéspedes, como las bacterias pertenecientes a los géneros *Escherichia*, *Shigella*, *Salmonella*, *Klebsiella*, *Citrobacter*, *Bacillus* y *Staphylococcus*. En estudios previos en cerdos sanos y enfermos, se describen trastornos gastrointestinales, mostrando la presencia frecuente de enfermedades a edades más tempranas en sistemas intensivos. Sachsenröder y col.

(2014) informaron que la proporción de bacteriófagos aumenta con la edad de los animales, encontrando proporciones de lecturas relacionadas con bacteriófagos de 22% a 44%, 68% a 72% y 88% a 92% para animales de doce días, 54 días y cerdas adultas, respectivamente. En los resultados de porcinos encontramos un alto porcentaje de secuencias de bacteriófagos (96,6% de las lecturas virales totales), lo que demuestra una mayor carga de estos virus a edades más tempranas que en estudios previos de viromas porcinos (Amimo y col., 2016; Sachsenröder y col., 2012; Zhang y col., 2014), incluso más que los reportados en cerdas adultas (Sachsenröder y col., 2014). Estos datos podrían estar relacionados con la salud de los porcinos estudiados, pues los fagos realizan funciones de control de bacterias. Estas funciones esenciales de los virus deben ser estudiadas en el futuro (Colbère-Garapin y col., 2007; Modi y col., 2013; Ogilvie y col., 2013). En las tres especies de murciélagos los virus con mayor abundancia correspondieron a Shigella phage pSf1-1 (género *Tunavirus*), Escherichia phage K1 (género *K1Gvirus*), Salmonella phage (género *Jerseyvirus* familia *Siphoviridae*), género *Tequatrovirus*, Klebsiella phage JD001 (género *Msw3virus*, familia *Myoviridae*), Klebsiella phage KP34 (género *Kp34virus*), y los géneros *Teseptimavirus* y *Phikmvvirus* (familia *Podoviridae*). Estas comunidades de fagos difieren de las reportadas en otros murciélagos donde las familias más representadas han sido *Podoviridae* y *Myoviridae* (Donaldson y col., 2010; He y col., 2013). Incluso estudios que trabajaron con muestras similares, reportan especies de los ordenes y familias *Caudovirales*, *Inoviridae*, *Microrviridae* (Wu y col., 2012), y *Podoviridae* (Donaldson y col., 2010). Además mostraron proporciones mayores a las reportadas en estudios anteriores (Donaldson y col., 2010; Li y col., 2010b; Salmier y col., 2017), incluso comparadas con estudios en otras especies como musaraña y demonio de Tasmania (Chen y col., 2019; Chong y col., 2019). Es fundamental prestar atención a la comunidad de fagos que alberga cualquier especie animal, ya que pueden proporcionar evidencia de ser

portadores de bacterias patógenas. Además, existe la necesidad de ahondar en la función de estos virus, ya que algunos han demostrado ser una sustitución adecuada de los antibióticos (Donaldson y col., 2010; Mokili y col., 2012; Pires y col., 2015; Herridge y col., 2020). De hecho, los fagos pueden ser una fuente de control de poblaciones de bacterias patógenas potenciales y, por lo tanto, proveer de resistencia a enfermedades a las especies que los hospedan. Por esta razón es importante realizar más estudios para comprender el papel de los fagos.

Entre los datos del viroma de los porcinos en este estudio se identificaron nuevas variantes de virus DNA de cadena sencilla circular, sapovirus, bocaparvovirus y mamastrovirus. Mientras que en los viromas de los murciélagos, los virus animales más representados fueron coronavirus (en *P. parnellii*) y circovirus.

La mayoría de los estudios disponibles sobre el viroma de porcinos parten de las heces de animales mantenidos en granjas intensivas tecnificadas con altas densidades de población o en condiciones experimentales. No hay información sobre el viroma de los cerdos en traspatio rural en este tipo de ambientes y climas de México. En este estudio, identificamos nuevas variantes de virus DNA circulares (virus tipo circovirus, virus ssDNA circular asociado al lobo marino y el género IASvirus no clasificado), sapovirus (familia *Caliciviridae*), bocaparvovirus (familia *Parvoviridae*) y mamastrovirus (familia *Astroviridae*) presentes en los cerdos de traspatio en Morelos, México, donde la agricultura de traspatio es un sistema diverso de producción de animales domésticos, que es utilizado por varias familias, principalmente en áreas rurales y periurbanas, como en muchas regiones geográficas en desarrollo del mundo. En este sistema, los animales pueden ser alimentados con alimentos comerciales, productos agrícolas y restos de comida, y mediante pastoreo; su manejo es variado, dependiendo del ambiente y de acuerdo con criterios y decisiones del productor. Esta descripción del viroma de cerdos

sanos rurales de México presenta la comunidad viral con menor diversidad de virus animales conocidos reportada en cerdos, proporcionando evidencia de nuevas variantes de virus no reportadas previamente en el país [72, 76, 89]. En informes anteriores, los virus porcinos más abundantes han sido kobuvirus, rotavirus, virus ssDNA circulares, astrovirus y enterovirus B (Sachsenröder y col., 2014). Estudios previos en porcinos (Lager y col., 2012; Sachsenröder y col., 2012; Shan y col., 2011) mostraron secuencias que representan entre nueve y 25 géneros y seis y diez familias de virus animales, mientras que los resultados de porcinos en este trabajo representaron virus de cinco géneros y cuatro familias virales. Teniendo en cuenta nuestros resultados, la mayoría de los estudios existentes sobre el viroma porcino han informado que las comunidades virales de cerdos mantenidos en granjas de alta tecnología en sistemas intensivos muestran una mayor diversidad de lecturas relacionadas con varias familias de virus animales, lo que se traduce en un mayor riesgo de enfermedades. Los sistemas intensivos en animales se han identificado como uno de los factores asociados con la aparición de enfermedades infecciosas (Patz y col., 2004). El viroma de los cerdos en entornos de producción diferentes a los sistemas intensivos y con interacciones sociales deben estudiarse, ya que estos factores ambientales probablemente influyeron en la comunidad viral de los cerdos estudiados (Biggs, 1985; Patz y col., 2004; Tung y col., 2015).

Los reportes de viromas de murciélagos se han realizado en algunas especies de América, Asia, África y Europa. En los estudios que reportan el viroma de murciélagos Americanos se reportan datos de especies de las familias *Vespertilionidae* (*Eptesicus fuscus*, *Perimyotis subflavus*, *Myotis lucifugus*), *Molosidea* (*Molossus molossus*) y *Phyllostomidae* (*Desmodus rotundus*) (Donaldson y col., 2010; Salmier y col., 2017). Aquí describimos el viroma de dos especies más de murciélagos de la familia *Phyllostomidae* (*Artibeus* sp y *Macrotus*

watherhousii) y un murciélago de la familia *Mormopidae* (*Pteronotus parnellii*) encontrando diferencias entre las comunidades virales reportadas. Algunos de los estudios del viroma de murciélagos existentes describen haberse realizado en ambientes urbanizados o con especies sinantrópicas (Baker y col., 2013; Dacheux y col., 2014; Gay y col., 2014; Ge y col., 2012; Geldenhuys y col., 2018; Salmier y col., 2017) que muestran comunidades virales con un mayor número de especies de virus animales y abundancias mayores de secuencias asociadas a virus de plantas, hongos y/o protozoos, que en general reflejan diversidades diferentes a las encontradas en este trabajo. La mayoría de los estudios anteriores trabajaron con muestras de heces (Ge y col., 2012; Geldenhuys y col., 2018; Li y col., 2010b; Salmier y col., 2017), pero hay informes sobre orina (Baker y col., 2013), hisopos orales y faríngeos (Donaldson y col., 2010; Wu y col., 2012), e incluso órganos y cadáveres (Dacheux y col., 2014; He y col., 2013). Aunque algunos de estos estudios trabajaron con muestras similares, mostraron diferentes diversidades virales. Los viromas de las tres especies de murciélagos reflejaron un bajo número de secuencias asignadas a virus animales conocidos (2% - 8.9%), en comparación con las reportadas en otros estudios (> 28%) (Baker y col., 2013; He y col., 2013). Las familias virales más ampliamente identificadas en estudios anteriores de metagenómica viral de murciélagos son *Herpesviridae*, *Adenoviridae*, *Circoviridae*, *Papillomaviridae*, *Retroviridae*, *Parvoviridae*, *Coronaviridae*, *Astroviridae*, *Flaviviridae*, *Picornaviridae* y *Poxviridae* (58% - 0,07% de lecturas virales) (Geldenhuys y col., 2018, 2016; He y col., 2013; Salmier y col., 2017, Wu y col., 2012). Algunas de las familias virales mencionadas fueron representadas en las secuencias o contigs de los murciélagos estudiados. Recientemente hay más de 10 770 secuencias de virus identificadas en murciélagos, que representan 29 familias de virus animales. Al menos el 25% de estos datos se han obtenido de especies de murciélagos de América. Se han identificado/aislado rabdovirus, coronavirus, paramixovirus, ortomixovirus y virus pertenecientes a otras 18

familias virales de especies de murciélagos de América (Chen y col., 2014). Este estudio contribuye a incrementar la información disponible en las bases de datos.

La diversidad de virus animales en las especies de quirópteros estudiadas también fue baja, ya que en otros estudios se han encontrado más de trece familias virales (Geldenhuis y col., 2018; Salmier y col., 2017; Wu y col., 2016), mientras en este estudio solo estuvieron representadas entre tres a cinco familias de virus animales, con muy pocas secuencias asociadas (1-35). Además en la confirmación por PCR, no se obtuvo amplificación en ninguna muestra, lo que indica que no hubo presencia de ácidos nucleicos o variantes virales no detectadas por los iniciadores. Los resultados anteriores se traducen en un menor riesgo de virus animales conocidos o similares en los murciélagos estudiados.

Las diversidades bajas de virus animales conocidos en las tres especies de murciélagos y en los cerdos de traspatio en este estudio, podrían asociarse con las estrategias de secuenciación empleadas, como el procesamiento y agrupamiento de las muestras. La filtración, dilución, el agrupamiento de muestras, así como la purificación de ADN/ARN, el enriquecimiento, al igual que otros pasos y estrategias de secuenciación, pueden variar entre flujos de trabajo de secuenciación masiva, dependiendo de muchos factores, como el tipo de muestra, el objetivo de secuenciación, los recursos, etc. Estas variaciones pueden afectar los resultados, pero en la mayoría de los casos, su efecto no se comprende bien. Puesto que no hay estándares para la secuenciación de viomas, debemos enfrentar todos los desafíos inherentes al campo de la metagenómica orientada al estudio del vioma (Wang, 2020). En este estudio, la combinación por especie de muestras nasales/faríngeas y rectales de individuos con edades y orígenes (granjas o refugios) distintos, limitó la capacidad para evaluar la presencia de virus en individuos. Aunque se tenía la intención de identificar los virus en muestras individuales por PCR/RT-PCR, los ácidos nucleicos se

consumieron en el proceso de obtención de librerías de ácidos nucleicos de alta calidad. Sin embargo, la mayor parte de las secuencias virales encontradas en este estudio estuvieron relacionadas con fagos de enterobacterias y virus gastrointestinales de los animales, y es probable que solo una pequeña proporción haya sido de origen nasal (ej. miembros de los géneros *Rosenblumvirus*, *Twortvirus* y *Sextaevirus* y el fago Sb1 de *Staphylococcus* no asignado). En cualquier caso, el origen de algunos virus (ej. parvovirus, en cerdos) no quedó claro. Es necesario secuenciar muestras individualmente o agruparlas de acuerdo con criterios específicos para obtener información particular sobre poblaciones virales específicas, así como obtener material/muestra suficiente. En estudios anteriores, no se describe diferencia entre las comunidades de las diferentes muestras por su origen (rectal y nasal) (B M Hause y col., 2016; Ben M. Hause y col., 2016).

Entre las secuencias asociadas a virus animales, tanto en los porcinos como en las tres especies de murciélagos, encontramos secuencias asociadas a virus DNA de cadena sencilla circular. En los porcinos la mayor proporción de secuencias asociadas a virus animales (6781 lecturas) mostraron homología (nt \leq 98%; aa \leq 98,2%) con el virus tipo circovirus porcino 51, previamente identificado a partir de las heces de lechones en una granja de alta densidad en los EE. UU (Shan y col., 2011) y al virus DNA circular asociado a heces de focas (nt \leq 97.6%; aa \leq 97.7%), previamente identificado en las heces de *Arctocephalus* sp. y cerdos de Japón (Chiappetta y col., 2017; Oba y col., 2017). En las tres especies de murciélagos y en porcinos se encontraron secuencias similares (nt \leq 96.4%) al gen Rep de circovirus identificados en granjas porcinas (Shan y col., 2011). Estas secuencias mostraron en común un fragmento mayor a 80 bases el cual fue empleado para construir un árbol filogenético con similitudes de 82.4% a 95.7% (Figura 12). La secuencia de porcinos fue la más divergente (nt 82.4 a 83.3%), encontrando

mayor similitud entre las secuencias de murciélagos (87.8% a 95.7%). Los virus DNA de cadena sencilla circular se detectan comúnmente en animales, otros organismos, y muestras ambientales (Breitbart y col., 2017). Estudios previos han demostrado la transmisión de circovirus entre especies (Li y col., 2011, 2010a). Aunque que las actividades de murciélagos de este estudio se superpusieron con las granjas de porcinos estudiadas, el fenómeno de la posible transmisión de circovirus entre especies no pudo ser bien evaluado, debido al tamaño de la secuencia común a todas las especies y la carencia de material genético (consumido en librerías de ADN de calidad y ensayos de PCR) para realizar una mejor caracterización de los circovirus. Pero es importante establecer vigilancia de circovirus y otros virus, más cuando hay informes de virus adaptados de murciélagos a cerdos como Nipha (Luby et al., 2009), Menangle (Philbey y col., 1998, 2008), Ébola (Leroy y col., 2005; Organización Mundial de la Salud, 2021) y Tioman 313 (Yaiw y col., 2008). Además, se han identificado virus de cerdos en algunas especies de murciélagos (ej. Mapuera, rubulavirus) (Cuevas-Romero JS, Blomstrom AL, Berg, 2015 Wang et al., 2007). Es importante tener en cuenta que todas estas enfermedades o infecciones se han asociado con actividades antropogénicas. En cualquier caso, es importante realizar un estudio de las comunidades virales y microbianas de cualquier población de murciélagos y animales domésticos en ambientes donde estas especies coexisten y se superponen, para identificar y/o determinar la transmisión y el riesgo de patógenos (Ge y col., 2011; Wang y col., 2018).

Se encontraron 1,605 secuencias relacionadas con el género *Sapovirus* de la familia *Caliciviridae*, el cual tuvo la segunda mayor abundancia entre los virus animales en los porcinos. Estos virus tienen importancia en enfermedades gastroentéricas tanto en porcinos como en humanos (Chiba y col., 2000; Reuter y col., 2010). Todos los animales muestreados eran clínicamente sanos, a pesar de

estar infectados con sapovirus. Este hallazgo confirma los resultados encontrados por (Lager y col. (2012). Este virus se ha encontrado en otros estudios sobre viromas porcinos (Sachsenröder y col., 2012; Zhang y col., 2014). El contig del ORF-1 del sapovirus porcino encontrado en este estudio muestra una similitud del nt $\leq 89,1\%$ y aa $\leq 96.5\%$ con PoSaV-JJ259 y otros sapovirus del genogrupo 3 (GIII), que es el genogrupo porcino prototipo (Figura 6) (Green y col., 2000; Scheuer y col., 2013). Dado que hay evidencia de sapovirus humanos en México (Wang y col., 2005), la infección de los cerdos de traspatio con sapovirus necesita ser estudiada más a fondo, ya que estos virus causan enfermedades similares en animales y humanos, y entre especies la transmisión (principalmente humano-animal) y la recombinación son posibles (Martella y col., 2008; Scheuer y col., 2013; Wang y col., 2005).

La mayoría de las secuencias relacionadas con la familia *Parvoviridae* (808) mostró similitud con secuencias del bocaparvovirus de ungulados 5 (ICTV). Un contig similar al gen NS1, utilizado para caracterizar los bocaparvovirus, fue utilizado para el análisis filogenético mostrando una similitud del nt $\leq 96.3\%$ (aa $\leq 100\%$) con el bocaparvovirus 5 JF713714 (previamente clasificado como bocavirus 3) (Figura 7). Estos virus han sido reportados en América del Norte, Asia, el Reino Unido, Europa del Este y África, y han sido señalados como candidatos importantes causantes de trastornos digestivos y respiratorios (ej. síndrome de emaciación multisistémico) en cerdos (Zhou y col., 2014). Informes anteriores sobre los viromas fecales de los cerdos han descrito bocaparvovirus en animales diarreicos (Shan y col., 2011), mientras que en este estudio los animales no mostraron signos clínicos

Las 98 lecturas con similitud a las secuencias de mamastrovirus (nt $\leq 100\%$; aa $\leq 100\%$), demostraron la presencia del virus en porcinos en México confirmando que estos virus comúnmente infectan a los cerdos sin causar enfermedad, según lo

informado por Luo y col. (2011). Cabe señalar que los mamastrovirus también se han asociado con trastornos neurológicos en humanos y visones (Blomström y col., 2010; Quan y col., 2010).

También en los porcinos se encontraron secuencias relacionadas con el virus IAS. El virus IAS se reportó por primera vez en el viroma de muestras recolectadas en mataderos del mercado primario y secundario y en una estación de compra de cerdos en el sureste de los EE. UU. (B M Hause y col., 2016). Este es el segundo informe de la presencia del virus IAS en el viroma de cerdos.

Entre las secuencias de *P. parnellii* asociadas con virus animales, la mayor proporción (76,1%) estuvo relacionada con coronavirus. Entre estas secuencias, el 48,6% mostró similitud con secuencias no clasificadas de betacoronavirus de murciélago; 17.1% con una secuencia parcial de betacoronavirus mexicano (KC117198), con similitudes de nt \leq 97.9%; 8,6% con el coronavirus de murciélago Roussettus HKU9 (MG762674), nt \leq 77,8%; 8,6% con el virus relacionado con el síndrome respiratorio agudo severo (NC_004718), nt \leq 77.3%. Además, el resto de los datos mostró secuencias relacionadas con miembros de los géneros betacoronavirus (11,4%), alfacoronavirus (2,8%) y gammacoronavirus (2,8%). Los murciélagos se clasifican como reservorios de Alfa- y Betacoronavirus (Banerjee y col., 2019; Han y col., 2015), pero en este estudio solo *P. parnellii* mostró secuencias similares a miembros de esta familia viral. Las aves son consideradas reservorios de gamacoronavirus (Wille y Holmes, 2020) pero en este estudio una secuencia (58 pb) similar (nt 81%) a un gammacoronavirus aviar (HM245924 Virus de la bronquitis infecciosa CQ04) fue encontrada. Esta fue analizada individualmente mostrando una mayor similitud (nt 98%) con secuencias parciales de betacoronavirus de murciélago (JQ731781). Por lo tanto, se realizó un alineamiento múltiple de secuencias que reveló que en este caso esta secuencia de longitud corta muestra homología con una región similar entre coronavirus de

diferentes géneros (beta y gama). Este hecho llama a poner atención en los resultados del análisis bioinformático y confirmarlos. Las secuencias de este estudio que mostraron similitud ($\leq 78,6\%$) con sarbecovirus y merbecovirus fueron secuencias de corta longitud (≤ 70 pb), razón por la cual estas no fueron analizadas más a fondo. En cualquier caso, en estos entornos, una mayor investigación y vigilancia de los coronavirus podría ser útil para determinar las especies de coronavirus, las especies de murciélagos hospedadoras relacionadas y otras variables de interés

Con base en los índices de diversidad (Shanon-Wiener) obtenidos de los viromas de las tres especies de murciélagos y los porcinos, podemos afirmar que las comunidades virales entre las especies son diferentes. Estas comunidades estuvieron representadas, en su mayoría por fagos, siendo *P. parnellii* la especie con mayor diversidad viral (total 3.19; virus animales 1.7) y *M. waterhousii* la especie de murciélago con la menor diversidad viral (total 2.98; virus animales 1.17). En cuanto a la diversidad viral en los porcinos (total 2.84; virus animales 1.1) fue menor en comparación con la de los murciélagos de este estudio y la de viromas porcinos anteriormente reportados (> 4) (Sachsenröder y col., 2014, 2012). La mayoría de los estudios de viromas de murciélagos no reportaron índices de diversidad viral. Es importante obtener estos valores para tener una referencia de las diversidades virales entre especies, teniendo en cuenta otras características como los entornos y los tipos de muestras. Rico y col. (2015) describen ampliamente la diversidad de virus y murciélagos en tres sitios de muestreo con diferentes niveles de perturbación antropogénica. Pero, dado que los métodos de identificación viral utilizados (identificación molecular específica) son diferentes, no es válido comparar los valores. Cabe mencionar que al estimar la diversidad en viromas empleando un índice de equidad, como lo es el de Shanon-Wiener, hay que considerar ciertos factores que intervienen en los

resultados y al compararlos con otros índices. Entre estos se encuentran que las observaciones de cada especie viral corresponden a las lecturas/secuencias nucleotídicas obtenidas (ej. donde secuencias de diferentes genes pueden pertenecer a un mismo virus, o secuencias del mismo gen pueden pertenecer a diferentes virus); y que el número de observaciones puede ser muy variable, en este caso fueron miles en los porcinos y unidades o decenas en los murciélagos.

En este estudio se encontró una gran cantidad de lecturas sin asignación taxonómica así como en otros reportes previos sobre viomas (Dacheux y col., 2014; He y col., 2013). La mayor proporción de lecturas no asignadas estuvo en los datos de los murciélagos. Dados nuestros limitados conocimientos y recursos sobre virus, es necesario continuar trabajando con los datos obtenidos, pues es muy probable que una proporción de las lecturas sin asignación taxonómica correspondan a virus o microorganismos desconocidos (Mokili y col., 2012).

El uso de la metagenómica nos permitió identificar virus no reportados previamente en México. Los resultados sugirieron que la metagenómica fue útil para identificar las comunidades virales y su abundancia en una muestra clínica, incluso nuevas variantes y nuevos virus (B M Hause y col., 2016); sin embargo, esta herramienta aún presenta limitaciones como las asociadas a la extracción de ácidos nucleicos, la amplificación, el enriquecimiento, los procedimientos de bioinformática y la información disponible en las bases de datos de secuencias. De cualquier manera esta herramienta sigue aportando información valiosa, como la de este estudio, que sirve como base para desentrañar el complejo viroma de cualquier especie.

Es importante realizar una descripción del ambiente donde se llevan a cabo los estudios, registrando su grado de alteración antropogénica y fenómenos. Estudiando especies domésticas y silvestres que habitan entornos que actúan como interfaces de riesgo para la adaptación y la presencia de agentes

infecciosos (Patz y col., 2004; Wiethoelter y col., 2015), reportamos los primeros viromas de especies de murciélagos y porcinos que convergen en las poblaciones de granjas rurales, encontrando comunidades con altos niveles de fagos y bajos en virus animales conocidos. Nuestros resultados contribuyen con una pieza más en la valiosa información del microbioma de las diferentes especies que habitan en diferentes ambientes, sumando el de tres especies de yangoquirópteros y el de porcinos de traspatio rural.

CONCLUSIONES

Este estudio representa la primer evaluación del viroma de cerdos y murciélagos que convergen en ambientes de traspatio rural estudiados en México. Este estudio proporciona información valiosa sobre las comunidades virales de cerdos criados en condiciones diferentes a los entornos intensivos tecnificados y de las tres especies de murciélagos, no reportadas previamente. Aquí se muestran comunidades virales diferentes comparadas entre las especies estudiadas y con los viromas de porcinos y quirópteros reportados anteriormente, identificando variantes virales nuevas y nuevos virus no reportados previamente en el país, haciendo evidente la necesidad de establecer la vigilancia de sapovirus, bocaparvovirus y mamastrovirus y su relación con la presencia de enfermedades respiratorias, digestivas, reproductivas o neurológicas en porcinos, así como coronavirus en especies de murciélagos de México. Es necesario establecer la vigilancia de agentes virales como circovirus, entre otros, en especies animales domésticas y silvestres que habitan en ambientes de convergencia, para confirmar o descartar la transmisión interespecie. Estos resultados reflejan el estado de salud de los cerdos y murciélagos que viven o se asocian, respectivamente, a ambientes rurales de traspatio. Con esto se confirma que el viroma de las especies e individuos varía entre los diferentes ambientes (incluyendo los diferentes sistemas de producción), ubicaciones geográficas, así como entre las especies.

REFERENCIAS

1. Tian H, Lu C, Yang J, et al (2015) Global patterns and controls of soil organic carbon dynamics as simulated by multiple terrestrial biosphere models: Current status and future directions. *Global Biogeochem Cycles* 29:775–792. <https://doi.org/10.1002/2014GB005021>
2. Lambin EF, Turner BL, Geist HJ, et al (2001) The causes of land-use and land-cover change: moving beyond the myths. *Glob Environ Chang* 11:261–269. [https://doi.org/10.1016/S0959-3780\(01\)00007-3](https://doi.org/10.1016/S0959-3780(01)00007-3)
3. Patz JA, Daszak P, Tabor GM, et al (2004) Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. *Environ Health Perspect* 112:1092–8. <https://doi.org/10.1289/EHP.6877>
4. Chua KB (2012) Introduction: Nipah Virus—Discovery and Origin. In: *Current topics in microbiology and immunology*. pp 1–9
5. Drosten C, Günther S, Preiser W, et al (2003) Identification of a Novel Coronavirus in Patients with Severe Acute Respiratory Syndrome. *N Engl J Med* 348:1967–1976. <https://doi.org/10.1056/NEJMoa030747>
6. Cohen ML (1998) Resurgent and emergent disease in a changing world. *Br Med Bull* 54:523–532. <https://doi.org/10.1093/oxfordjournals.bmb.a011707>
7. Wiethoelter AK, Beltrán-Alcrudo D, Kock R, Mor SM (2015) Global trends in infectious diseases at the wildlife–livestock interface. *Proc Natl Acad Sci* 112:9662–9667. <https://doi.org/10.1073/pnas.1422741112>
8. SAGARPA (2011) ACUERDO mediante el cual se dan a conocer en los Estados Unidos Mexicanos las enfermedades y plagas exóticas y endémicas de notificación obligatoria de los animales terrestres y acuáticos. *D Of la Fed* 38:38–85
9. Graves JH (1973) Serological relationship of swine vesicular disease virus and Coxsackie B5 virus. *Nature* 245:314–5
10. Zhang G, Wilsden G, Knowles NJ, McCauley JW (1993) Complete nucleotide sequence of a coxsackie B5 virus and its relationship to swine vesicular disease virus. *J Gen Virol* 74 (Pt 5):845–53. <https://doi.org/10.1099/0022-1317-74-5-845>

11. Verdaguer N, Jimenez-Clavero MA, Fita I, Ley V (2003) Structure of Swine Vesicular Disease Virus: Mapping of Changes Occurring during Adaptation of Human Coxsackie B5 Virus To Infect Swine. *J Virol* 77:9780–9789. <https://doi.org/10.1128/JVI.77.18.9780-9789.2003>
12. Taylor LH, Latham SM, Woolhouse MEJ (2001) Risk factors for human disease emergence. *Philos Trans R Soc B Biol Sci* 356:983–989. <https://doi.org/10.1098/rstb.2001.0888>
13. Li L, Shan T, Soji OB, et al (2011) Possible cross-species transmission of circoviruses and cycloviruses among farm animals. *J Gen Virol* 92:768–772. <https://doi.org/10.1099/vir.0.028704-0>
14. Lau SKP, Li KSM, Huang Y, et al (2010) Ecoepidemiology and Complete Genome Comparison of Different Strains of Severe Acute Respiratory Syndrome-Related Rhinolphus Bat Coronavirus in China Reveal Bats as a Reservoir for Acute, Self-Limiting Infection That Allows Recombination Events. *J Virol* 84:2808–2819. <https://doi.org/10.1128/JVI.02219-09>
15. Bowden TR, Westenberg M, Wang LF, et al (2001) Molecular characterization of menangle virus, a novel paramyxovirus which infects pigs, fruit bats, and humans. *Virology* 283:358–373. <https://doi.org/10.1006/viro.2001.0893>
16. Marinkelle CJ (1996) *Babesia* sp. in Colombian bats (Microchiroptera). *J Wildl Dis* 32:534–535. <https://doi.org/10.7589/0090-3558-32.3.534>
17. Kuzmin I V., Bozick B, Guagliardo SA, et al (2011) Bats, emerging infectious diseases, and the rabies paradigm revisited. *Emerg Health Threats J* 4:1–17. <https://doi.org/10.3402/ehth.v4i0.7159>
18. Young CCW, Olival KJ (2016) Optimizing viral discovery in bats. *PLoS One* 11:1–18. <https://doi.org/10.1371/journal.pone.0149237>
19. Chen L, Liu B, Yang J, Jin Q (2014) DBatVir: The database of bat-associated viruses. *Database* 2014:1–7
20. Plowright RK, Parrish CR, McCallum H, et al (2017) Pathways to zoonotic spillover. *Nat Rev Microbiol* 15:502–510. <https://doi.org/10.1038/nrmicro.2017.45>

21. Afelt A, Lacroix A, Zawadzka-Pawlewska U, et al (2018) Distribution of bat-borne viruses and environment patterns. *Infect Genet Evol* 58:181–191. <https://doi.org/10.1016/j.meegid.2017.12.009>
22. Mühldorfer K (2013) Bats and Bacterial Pathogens: A Review. *Zoonoses Public Health* 60:93–103. <https://doi.org/10.1111/j.1863-2378.2012.01536.x>
23. Tang XC, Zhang JX, Zhang SY, et al (2006) Prevalence and Genetic Diversity of Coronaviruses in Bats from China. *J Virol* 80:7481–7490. <https://doi.org/10.1128/jvi.00697-06>
24. Poon LLM, Chu DKW, Chan KH, et al (2005) Identification of a Novel Coronavirus in Bats. *J Virol* 79:2001–2009. <https://doi.org/10.1128/jvi.79.4.2001-2009.2005>
25. Lau SKP, Li KSM, Huang Y, et al (2010) Ecoepidemiology and Complete Genome Comparison of Different Strains of Severe Acute Respiratory Syndrome-Related Rhinolophus Bat Coronavirus in China Reveal Bats as a Reservoir for Acute, Self-Limiting Infection That Allows Recombination Events. *J Virol* 84:2808–2819. <https://doi.org/10.1128/jvi.02219-09>
26. Gloza-Rausch F, Ipsen A, Seebens A, et al (2008) Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. *Emerg Infect Dis* 14:626–631. <https://doi.org/10.3201/eid1404.071439>
27. Dominguez SR, O’Shea TJ, Oko LM, Holmes K V. (2007) Detection of group 1 coronaviruses in bats in North America. *Emerg Infect Dis* 13:1295–1300. <https://doi.org/10.3201/eid1309.070491>
28. Tong S, Conrardy C, Ruone S, et al (2009) Detection of novel SARS-like and other coronaviruses in bats from Kenya. *Emerg Infect Dis* 15:482–485. <https://doi.org/10.3201/eid1503.081013>
29. Leroy EM, Kumulungui B, Pourrut X, et al (2005) Fruit bats as reservoirs of Ebola virus. *Nature* 438:575–6. <https://doi.org/10.1038/438575a>
30. Tong S, Li Y, Rivaller P, et al (2012) A distinct lineage of influenza A virus from bats. *Proc Natl Acad Sci* 109:4269–4274. <https://doi.org/10.1073/pnas.1116200109>

31. Chen L, Liu B, Yang J, Jin Q (2014) DBatVir: the database of bat-associated viruses. Database 2014:bau021-bau021. <https://doi.org/10.1093/database/bau021>
32. DBatVir: the Database of Bat-associated Viruses. <http://www.mgc.ac.cn/DBatVir/>. Accessed 21 Feb 2021
33. Huang K, Zhu H, Fan X, et al (2012) Establishment and lineage replacement of H6 influenza viruses in domestic ducks in southern China. J Virol 86:6075–83. <https://doi.org/10.1128/JVI.06389-11>
34. Kahn RE, Morozov I, Feldmann H, Richt JA (2012) 6th International Conference on Emerging Zoonoses. Zoonoses Public Health 59:2–31. <https://doi.org/10.1111/j.1863-2378.2012.01539.x>
35. Luby SP, Gurley ES, Hossain MJ (2009) Transmission of Human Infection with Nipah Virus. Clin Infect Dis 49:1743–1748. <https://doi.org/10.1086/647951>
36. Smith I, Wang LF (2013) Bats and their virome: An important source of emerging viruses capable of infecting humans. Curr Opin Virol 3:84–91. <https://doi.org/10.1016/j.coviro.2012.11.006>
37. Organizacion Mundial de la Salud (2021) Organizacion Mundial de la salud Enfermedad por el virus del Ebola. Oms
38. Philbey AW, Kirkland PD, Ross AD, et al (1998) An apparently new virus (family Paramyxoviridae) infectious for pigs, humans, and fruit bats. Emerg Infect Dis 4:269–271. <https://doi.org/10.3201/eid0402.980214>
39. Philbey A, Kirkland P, Ross A, et al (2008) Infection with Menangle virus in flying foxes (*Pteropus* spp.) in Australia. Aust Vet J 86:449–454. <https://doi.org/10.1111/j.1751-0813.2008.00361.x>
40. Yaiw KC, Crameri G, Wang L, et al (2007) Serological Evidence of Possible Human Infection with *Tioman virus*, a Newly Described Paramyxovirus of Bat Origin. J Infect Dis 196:884–886. <https://doi.org/10.1086/520817>

41. Yaiw KC, Bingham J, Crameri G, et al (2008) Tioman Virus, a Paramyxovirus of Bat Origin, Causes Mild Disease in Pigs and Has a Predilection for Lymphoid Tissues. *J Virol* 82:565–568. <https://doi.org/10.1128/JVI.01660-07>
42. Escalera-Zamudio M, Taboada B, Rojas-Anaya E, et al (2018) Viral Communities Among Sympatric Vampire Bats and Cattle. *Ecohealth* 15:132–142. <https://doi.org/10.1007/s10393-017-1297-y>
43. Wang LF, Hansson E, Yu M, et al (2007) Full-length genome sequence and genetic relationship of two paramyxoviruses isolated from bat and pigs in the Americas. *Arch Virol* 152:1259–1271. <https://doi.org/10.1007/s00705-007-0959-4>
44. Cuevas-Romero JS, Blomstrom AL, Berg M (2015) Molecular and epidemiological studies of Porcine rubulavirus infection - an overview. *Infect Ecol Epidemiol* 5:1–8. <https://doi.org/10.3402/iee.v5.29602> [pii]
45. Handelsman J, Rondon MR, Brady SF, et al (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5:R245-9
46. Hugenholtz P, Tyson GW (2008) Metagenomics. *Nature* 455:481–483. <https://doi.org/10.1038/455481a>
47. Singh B, Bhat TK, Kurade NP, Sharma OP (2008) Metagenomics in animal gastrointestinal ecosystem: A microbiological and biotechnological perspective. *Indian J Microbiol* 48:216–227. <https://doi.org/10.1007/s12088-008-0027-0>
48. Schürch AC, Schipper D, Bijl MA, et al (2014) Metagenomic survey for viruses in Western Arctic caribou, Alaska, through iterative assembly of taxonomic units. *PLoS One* 9:e105227. <https://doi.org/10.1371/journal.pone.0105227>
49. Lamendella R, Santo Domingo JW, Ghosh S, et al (2011) Comparative fecal metagenomics unveils unique functional capacity of the swine gut. *BMC Microbiol* 11:103. <https://doi.org/10.1186/1471-2180-11-103>
50. Ghurye JS, Cepeda-Espinoza V, Pop M (2016) Metagenomic assembly: Overview, challenges and applications. *Yale J Biol Med* 89:353–362

51. Villamar Angula L, Barrera Wadgymar MA, Gallardo Nieto JL (2006) Situación actual y perspectiva de la producción de carne de porcino en México 2006
52. SIAP (2015) Población ganadera. Secr Agric Ganad Desarro Rural Pesca y Aliment 2015
53. INEGI IN de EG e I (2019) Ganadería. In: Encuesta Nac. Agropecu. <https://www.inegi.org.mx/temas/ganaderia/#Tabulados>. Accessed 24 Mar 2021
54. Medellín RA, Arita HT, Sánchez O (2007) Identificación de los murcielagos de Mexico, clave de campo. Inst Ecol UNAM Segunda ed:1–80
55. Djikeng A, Halpin R, Kuzmickas R, et al (2008) Viral genome sequencing by random priming methods. *BMC Genomics* 9:5. <https://doi.org/10.1186/1471-2164-9-5>
56. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>
57. Afgan E, Baker D, Batut B, et al (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:W537–W544. <https://doi.org/10.1093/nar/gky379>
58. Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: An extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res* 40:. <https://doi.org/10.1093/nar/gks678>
59. Chao A, Ma KH, Hsieh TC (2016) iNEXT (iNterpolation and EXTrapolation) Online. Program and User's Guide published at http://chao.stat.nthu.edu.tw/wordpress/software_download/. 30043:1–14. <https://doi.org/10.13140/RG.2.2.25777.79200>
60. Moreno CE (2001) Métodos para medir la biodiversidad
61. De Souza Luna LK, Heiser V, Regamey N, et al (2007) Generic detection of coronaviruses and differentiation at the prototype strain level by reverse

- transcription-PCR and nonfluorescent low-density microarray. *J Clin Microbiol* 45:1049–1052. <https://doi.org/10.1128/JCM.02426-06>
62. Lelli D, Papetti A, Sabelli C, et al (2013) Detection of coronaviruses in bats of various species in Italy. *Viruses* 5:2679–2689. <https://doi.org/10.3390/v5112679>
 63. Tong S, Chern SWW, Li Y, et al (2008) Sensitive and broadly reactive reverse transcription-PCR assays to detect novel paramyxoviruses. *J Clin Microbiol* 46:2652–2658. <https://doi.org/10.1128/JCM.00192-08>
 64. La Rosa G, Fratini M, Muscillo M, et al (2014) Molecular characterisation of human hepatitis E virus from Italy: comparative analysis of five reverse transcription-PCR assays. *Virology* 472:11–22. <https://doi.org/10.1016/j.virus.2014.05.011>
 65. Barrón-rodríguez RJ (2018) Isolation and characterization from a temporary artificial pond in Mexico. *Vet México OA* 5:1–16. <https://doi.org/10.21753/vmoa.5.2.475>
 66. VanDevanter DR, Warrener P, Bennett L, et al (1996) Detection and analysis of diverse herpesviral species by consensus primer PCR. *J Clin Microbiol* 34:1666–71
 67. Kumar S, Stecher G, Li M, et al (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>
 68. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
 69. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci* 101:11030–11035. <https://doi.org/10.1073/pnas.0404206101>
 70. Barrón-Rodríguez RJ, Parra-Laca R, Rojas-Anaya E, et al (2021) Evidence of Viral Communities in three Species of Bats from Rural Environment in Mexico. *Acta Chiropterologica*, en prensa:

71. Barrón-Rodríguez RJ, Rojas-Anaya E, Ayala-Sumuano JT, et al (2021) Swine virome on rural backyard farms in Mexico: communities with different abundances of animal viruses and phages. *Arch Virol* 166:475–489. <https://doi.org/10.1007/s00705-020-04894-y>
72. Shan T, Li L, Simmonds P, et al (2011) The Fecal Virome of Pigs on a High-Density Farm. *J Virol* 85:11697–11708. <https://doi.org/10.1128/JVI.05217-11>
73. Sachsenröder J, Twardziok SO, Scheuch M, Johne R (2014) The general composition of the faecal virome of pigs depends on age, but not on feeding with a probiotic bacterium. *PLoS One* 9:15–22. <https://doi.org/10.1371/journal.pone.0088888>
74. Donaldson EF, Haskew AN, Gates JE, et al (2010) Metagenomic Analysis of the Viromes of Three North American Bat Species: Viral Diversity among Different Bat Species That Share a Common Habitat. *J Virol* 84:13004–13018. <https://doi.org/10.1128/JVI.01255-10>
75. Salmier A, Tirera S, Thoisy B De, et al (2017) Virome analysis of two - sympatric bat species (*Desmodus rotundus* and *Molossus molossus*) in French Guiana. *PLoS One* 1–25. <https://doi.org/10.1371/journal.pone.0186943>
76. Sachsenröder J, Twardziok S, Hammerl JA, et al (2012) Simultaneous identification of DNA and RNA viruses present in pig faeces using process-controlled deep sequencing. *PLoS One* 7:. <https://doi.org/10.1371/journal.pone.0034631>
77. Zhang B, Tang C, Yue H, et al (2014) Viral metagenomics analysis demonstrates the diversity of viral flora in piglet diarrhoeic faeces in China. *J Gen Virol* 95:1603–1611. <https://doi.org/10.1099/vir.0.063743-0>
78. Amimo JO, El Zowalaty ME, Githae D, et al (2016) Metagenomic analysis demonstrates the diversity of the fecal virome in asymptomatic pigs in East Africa. *Arch Virol* 161:887–897. <https://doi.org/10.1007/s00705-016-2819-6>
79. Colbère-Garapin F, Martin-Latil S, Blondel B, et al (2007) Prevention and treatment of enteric viral infections: possible benefits of probiotic bacteria. *Microbes Infect* 9:1623–31. <https://doi.org/10.1016/j.micinf.2007.09.016>

80. Modi SR, Lee HH, Spina CS, Collins JJ (2013) Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499:219–222. <https://doi.org/10.1038/nature12212>
81. Ogilvie LA, Bowler LD, Caplin J, et al (2013) Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. *Nat Commun* 4:2420. <https://doi.org/10.1038/ncomms3420>
82. He B, Li Z, Yang F, et al (2013) Virome Profiling of Bats from Myanmar by Metagenomic Analysis of Tissue Samples Reveals More Novel Mammalian Viruses. *PLoS One* 8:. <https://doi.org/10.1371/journal.pone.0061950>
83. Li L, Victoria JG, Wang C, et al (2010) Bat Guano Virome: Predominance of Dietary Viruses from Insects and Plants plus Novel Mammalian Viruses. *J Virol* 84:6955–6965. <https://doi.org/10.1128/JVI.00501-10>
84. Chen L, Gu W, Liu C, et al (2019) Characteristics of the tree shrew gut virome. *PLoS One* 14:e0212774. <https://doi.org/10.1371/journal.pone.0212774>
85. Chong R, Shi M, Grueber CE, et al (2019) Fecal Viral Diversity of Captive and Wild Tasmanian Devils Characterized Using Virion-Enriched Metagenomics and Metatranscriptomics. *J Virol* 93:. <https://doi.org/10.1128/JVI.00205-19>
86. Mokili JL, Rohwer F, Dutilh BE (2012) Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2:63–77. <https://doi.org/10.1016/j.coviro.2011.12.004>
87. Pires DP, Vilas Boas D, Sillankorva S, Azeredo J (2015) Phage Therapy: a Step Forward in the Treatment of *Pseudomonas aeruginosa* Infections. *J Virol* 89:7449–56. <https://doi.org/10.1128/JVI.00385-15>
88. Herridge WP, Shibu P, O’Shea J, et al (2020) Bacteriophages of *Klebsiella* spp., their diversity and potential therapeutic uses. *J. Med. Microbiol.* 69:176–194
89. Lager KM, Ng TF, Bayles DO, et al (2012) Diversity of viruses detected by deep sequencing in pigs from a common background. *J Vet Diagnostic Investig* 24:1177–1179. <https://doi.org/10.1177/1040638712463212>

90. Biggs PM (1985) Infectious Animal Disease and Its Control. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 310:259–274
91. Tung J, Barreiro LB, Burns MB, et al (2015) Social networks predict gut microbiome composition in wild baboons. *Elife* 2015:1–18. <https://doi.org/10.7554/eLife.05224>
92. Baker KS, Leggett RM, Bexfield NH, et al (2013) Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology* 441:95–106. <https://doi.org/10.1016/j.virol.2013.03.014>
93. Dacheux L, Cervantes-Gonzalez M, Guigon G, et al (2014) A preliminary study of viral metagenomics of french bat species in contact with humans: Identification of new mammalian viruses. *PLoS One* 9:. <https://doi.org/10.1371/journal.pone.0087194>
94. Gay N, Olival KJ, Bumrungsri S, et al (2014) Parasite and viral species richness of Southeast Asian bats: Fragmentation of area distribution matters. *Int J Parasitol Parasites Wildl* 3:161–170. <https://doi.org/10.1016/j.ijppaw.2014.06.003>
95. Ge X, Li Y, Yang X, et al (2012) Metagenomic Analysis of Viruses from Bat Fecal Samples Reveals Many Novel Viruses in Insectivorous Bats in China. *J Virol* 86:4620–4630. <https://doi.org/10.1128/jvi.06671-11>
96. Geldenhuys M, Mortlock M, Weyer J, et al (2018) A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in *Neoromicia* bats within South Africa. *PLoS One* 13:e0194527. <https://doi.org/10.1371/journal.pone.0194527>
97. Wu Z, Ren X, Yang L, et al (2012) Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. *J Virol* 86:10999–1012. <https://doi.org/10.1128/JVI.01394-12>
98. Wu Z, Yang L, Ren X, et al (2016) Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME J* 10:609–620. <https://doi.org/10.1038/ismej.2015.138>

99. Wang D (2020) 5 challenges in understanding the role of the virome in health and disease. *PLoS Pathog* 16:e1008318.
<https://doi.org/10.1371/journal.ppat.1008318>
100. Hause BM, Duff JW, Scheldt A, Anderson G (2016) Virus detection using metagenomic sequencing of swine nasal and rectal swabs. *J Swine Heal Prod* 24:304–308
101. Hause BM, Padmanabhan A, Pedersen K, Gidlewski T (2016) Feral swine virome is dominated by single-stranded DNA viruses and contains a novel Orthopneumovirus which circulates both in feral and domestic swine. *J Gen Virol* 97:2090–2095. <https://doi.org/10.1099/jgv.0.000554>
102. Chiappetta CM, Cibulski SP, Lima FES, et al (2017) Molecular Detection of Circovirus and Adenovirus in Feces of Fur Seals (*Arctocephalus* spp.). *Ecohealth* 14:69–77. <https://doi.org/10.1007/s10393-016-1195-8>
103. Oba M, Katayama Y, Naoi Y, et al (2017) Discovery of fur seal feces-associated circular DNA virus in swine feces in Japan. *J Vet Med Sci* 79:1664–1666. <https://doi.org/10.1292/jvms.16-0642>
104. Li L, Kapoor A, Slikas B, et al (2010) Multiple Diverse Circoviruses Infect Farm Animals and Are Commonly Found in Human and Chimpanzee Feces. *J Virol* 84:1674–1682. <https://doi.org/10.1128/JVI.02109-09>
105. Breitbart M, Delwart E, Rosario K, et al (2017) ICTV Virus Taxonomy Profile: Circoviridae. *J Gen Virol* 98:1997–1998. <https://doi.org/10.1099/jgv.0.000871>
106. Ge X, Li J, Peng C, et al (2011) Genetic diversity of novel circular ssDNA viruses in bats in China. *J Gen Virol* 92:2646–2653.
<https://doi.org/10.1099/vir.0.034108-0>
107. Wang H, Li S, Mahmood A, et al (2018) Plasma virome of cattle from forest region revealed diverse small circular ssDNA viral genomes. *Virol J* 15:11.
<https://doi.org/10.1186/s12985-018-0923-9>
108. Chiba S, Nakata S, Numata-Kinoshita K, Honma S (2000) Sapporo virus: history and recent findings. *J Infect Dis* 181 Suppl:S303–S308.
<https://doi.org/10.1086/315574>

109. Reuter G, Zimsek-Mijovski J, Poljsak-Prijatelj M, et al (2010) Incidence, diversity, and molecular epidemiology of sapoviruses in swine across Europe. *J Clin Microbiol* 48:363–368. <https://doi.org/10.1128/JCM.01279-09>
110. Green KY, Ando T, Balayan MS, et al (2000) Taxonomy of the caliciviruses. *J Infect Dis* 181 Suppl:S322–S330. <https://doi.org/10.1086/315591>
111. Scheuer KA, Oka T, Hoet AE, et al (2013) Prevalence of porcine Noroviruses, molecular characterization of emerging porcine sapoviruses from finisher swine in the United States, and unified classification scheme for sapoviruses. *J Clin Microbiol* 51:2344–2353. <https://doi.org/10.1128/JCM.00865-13>
112. Wang QH, Han MG, Funk JA, et al (2005) Genetic diversity and recombination of porcine sapoviruses. *J Clin Microbiol* 43:5963–5972. <https://doi.org/10.1128/JCM.43.12.5963-5972.2005>
113. Martella V, Lorusso E, Banyai K, et al (2008) Identification of a porcine calicivirus related genetically to human sapoviruses. *J Clin Microbiol* 46:1907–1913. <https://doi.org/10.1128/JCM.00341-08>
114. Zhou F, Sun H, Wang Y (2014) Porcine bocavirus: Achievements in the past five years. *Viruses* 6:4946–4960. <https://doi.org/10.3390/v6124946>
115. Luo Z, Roi S, Dastor M, et al (2011) Multiple novel and prevalent astroviruses in pigs. *Vet Microbiol* 149:316–323. <https://doi.org/10.1016/j.vetmic.2010.11.026>
116. Blomström A-L, Widén F, Hammer A-S, et al (2010) Detection of a Novel Astrovirus in Brain Tissue of Mink Suffering from Shaking Mink Syndrome by Use of Viral Metagenomics. *J Clin Microbiol* 48:4392–4396. <https://doi.org/10.1128/JCM.01040-10>
117. Quan P-L, Wagner TA, Briese T, et al (2010) Astrovirus Encephalitis in Boy with X-linked Agammaglobulinemia. *Emerg Infect Dis* 16:918–925. <https://doi.org/10.3201/eid1606.091536>
118. Wille M, Holmes EC (2020) Wild birds as reservoirs for diverse and abundant gamma- and deltacoronaviruses. *FEMS Microbiol Rev* 44:631. <https://doi.org/10.1093/FEMSRE/FUAA026>

119. Han HJ, Wen H ling, Zhou CM, et al (2015) Bats as reservoirs of severe emerging infectious diseases. *Virus Res* 205:1–6.
<https://doi.org/10.1016/j.virusres.2015.05.006>
120. Banerjee A, Kulcsar K, Misra V, et al (2019) Bats and coronaviruses. *Viruses* 11
121. Rico O, Ojeda-Flores R, Sotomayor-Bonilla J, et al (2015) Diversidad viral de comunidades de murciélagos en paisajes transformados de México. Rico, Oscar Ojeda-Flores, Rafael Sotomayor-Bonilla, Jesús Zambrana-Torrel, Carlos Loza, Elizab Aguirre, A Alonso Suzán, Gerardo 2:1–23

ANEXOS

Anexo 1. Retrotranscripción y enriquecimiento de RNA.

- Se preparó la mezcla maestra para el total de muestras con los siguientes reactivos y volúmenes:

Mezcla/RT-A

REACTIVO	Concentración	Volumen por muestra	Volumen por total de muestras (4)
Oligo dT		1 µl	4 µl
M13r1	469.02 µM	0.63 µl	2.54 µl
M13r2	660.16 µM	0.45 µl	1.8 µl
M13r3	625.27 µM	0.47 µl	1.88 µl
M13r4	507.91 µM	0.59 µl	2.36 µl
RNA (10 ng)		10 µl	

Se transfirieron tres µl de la Mezcla/RT-A a un tubo limpio previamente identificado y posteriormente se adicionó el RNA de cada muestra-especie a su tubo correspondiente homogenizando.

- Se incubaron las muestras a 65°C durante diez minutos.

Durante la incubación se preparó la mezcla empleada en el paso subsecuente con los siguientes reactivos. La mezcla se preparó para todas las muestras.

Mezcla/RT-B

REACTIVO	Volumen por muestra	Volumen por total de muestras (4)
Buffer de Transcriptor	4 µl	16 µl
Transcriptor reverse transcripatse	0.5 µl	2 µl
Inhibidor de Rnasa	0.5 µl	2 µl
DNTP Mix (10 mM)	2 µl	8 µl

- Al finalizar la incubación a 65 °C se adicionaron siete µl de la Mezcla/RT-B a cada muestra homogenizando la reacción.

- Las muestras se incubaron a 25 °C por diez minutos.
- Posteriormente se incubaron a 50 °C durante 60 minutos.
- Al finalizar las incubaciones las muestras se mantuvieron a cuatro °C hasta su utilización en el paso siguiente.

Anexo 2. Enriquecimiento de muestras de DNA con Fast Start High Fidelity PCR System.

Para preparar la mezcla (Mezcla/DNA-en) para todas las reacciones se utilizaron los siguientes volúmenes de reactivos:

Mezcla/DNA-en

REACTIVO	Volumen por muestra	Volumen por total de muestras (4)
FastStart High Fidelity Reaction Buffer, 10x	2.5 µl	10.0 µl
M13-Hexameros (primers c/hexameros)	0.25 µl	1.0 µl
MgCl ₂	1.8 µl	7.2 µl
dNTPs	0.5 µl	2.0 µl
Agua	14.7 µl	58.8 µl
FastStart High Fidelity Enzyme Blend (5 U/µL)	0.25 µl	1.0 µl
DNA	5 µl	

Una vez preparada la mezcla/DNA-en se distribuyeron 20 µl por tubo y se añadieron cinco µl de DNA a su tubo correspondiente homogenizando adecuadamente. Los tubos cerrados se colocaron en el termociclador para llevar a cabo la PCR, el protocolo se muestra a continuación.

Programa de PCR (termociclado) para el enriquecimiento de DNA con el Fast Start High Fidelity PCR System

PASO	Temperatura	Duración
Desnaturalización inicial/activación de polimeraza	95 °C	2 minutos
Reamplificación 30 ciclos		
Desnaturalización	95 °C	30 segundos
Alineamiento	50 °C	30 segundos
Extensión	72 °C	3 minutos
Extensión final	72 °C	5 minutos
Mantenimiento	4 °C	∞

Anexo 3. Generación de DNA y cDNA de doble cadena.

Con las muestras de cDNA y DNA enriquecidas, se generó la doble cadena empleando el fragmento klenow (New England Biolabs) siguiendo las condiciones del fabricante a continuación descritas.

Se preparó la mezcla (Mezcla-dsDNA) necesaria para el número de muestras:

Mezcla-dsDNA		
REACTIVO	Volumen por muestra	Volumen por total de muestras (8)
Buffer	2.5 µl	10.0 µl
DNTPs (10 mM)	1.0 µl	4.0 µl
M13 random (20µM)	1.0 µl	4.0 µl
Agua	15.5 µl	62.0 µl
cDNA o DNA enriquecido	5.0 µl	20.0 µl
Volumen TOTAL	25.0 µl	

- Una vez preparada la mezcla se transfirieron a cada tubo de DNA o cDNA 20 µl de la Mezcla-DNA doble y se agregaron cinco µl de cDNA o DNA.
- Se incubaron a 94°C por tres minutos
- Se transfirieron a hielo durante dos minutos
- Se agregó 0.25 µl de polimerasa Klenow en cada tubo de reacción.
- Se incubaron a 37°C durante una hora
- Se desnaturalizaron 94°C por tres minutos
- Se transfirieron a hielo por dos minutos
- Se repitieron los pasos 4 y 5
- Al final se inactivó la enzima a 75°C durante 10 minutos.

Anexo 4. Amplificación de cDNA y DNA por PCR.

Se realizó un enriquecimiento de los DNA/cDNA de doble cadena mediante PCR empleando el Faststart High Fidelity PCR System (Roche). Los volúmenes empleados para las reacciones se muestran a continuación.

Mezcla AmpPCR

REACTIVO	Volumen por muestra	Volumen por total de muestras (4)
Buffer	2.5 µl	10.0 µl
M13 (primer sin hexamero)	0.25 µl	4.0 µl
M13random	0.25 µl	4.0 µl
MgCl ₂	1.8 µl	62.0 µl
dNTPs	0.5 µl	20.0 µl
Agua	14.45 µl	25.0 µl
Taq	0.25 µl	
dscDNA/DNA Marcado por M13	5µl	
	Vol. = 25 µl	

Una vez preparada y bien homogenizada la Mezcla AmpPCR se distribuyeron 20 µl por tubo y se añadieron cinco µl de dscDNA o dsDNA a su tubo correspondiente y se colocaron en el termociclador para llevar a cabo la PCR con las condiciones siguientes.

Programa de PCR (termociclado) para la amplificación de dsDNA y dscDNA

PASO	Temperatura	Duración
Desnaturalización inicial/activación de polimerasa	95 °C	2 minutos
Reamplificación 30 ciclos		
Desnaturalización	95°C	30 segundos
Alineamiento	50 °C	30 segundos
Extensión	72 °C	3 minutos
Extensión final	72 °C	5 minutos
Mantenimiento	4°C	∞

Anexo 5. Fragmentación de cDNA y DNA de doble cadena.

Para realizar la fragmentación de las muestras de cDNA/DNA-especie se empleó el Nextera XT DNA library prep kit (Illumina) siguiendo las instrucciones del fabricante a continuación descritas brevemente.

Se preparó la mezcla de reactivos (Mezcla-Frag) empleando los siguientes volúmenes:

Mezcla-Frag		
REACTIVO	Volumen por muestra	Volumen por total de muestras (4)
Tagment DNA buffer	10.0 µl	40.0 µl
H ₂ O destilada ultra pura	30 µl	100.0 µl
Volumen TOTAL por muestra	40.0 µl	

Una vez preparada y bien homogenizada la Mezcla-Frag se transfirieron 40 µl por tubo y se añadieron cinco µl de cDNA/DNA-especie a su tubo correspondiente.

A cada tubo se le añadieron cinco µl de Amplicon Tagment Mix homogenizando.


Los tubos se colocaron en el termociclador para llevar a cabo la fragmentación a 55 °C durante cinco minutos seguido de diez °C mantenimiento.

En cuanto las reacciones alcanzaron los diez °C se les añadió cinco µl de Nuetralize Tagment Buffer homogenizando por pipeteo.

La calidad y concentración de los cDNA/DNA fragmentados fueron cuantificadas y evaluadas por electroforesis de alta definición empleando el High sensitivity DNA kit con el equipo Agilent 2100 Bioanalyzer siguiendo el procedimiento del fabricante a continuación descrito brevemente.

- Se templaron los reactivos (High sensitivity DNA reagents) durante 30 minutos.
- Se programó el software 2100 expert seleccionando el tipo de ensayo “High sensitivity DNA assay”.

- Se preparó la mezcla del gel-colorante (Gel dye mix) añadiendo 15 µl del colorante (high sensitivity DNA dye concentrate) a un tubo vial del gel-matriz (high sensitivity DNA gel matrix). La mezcla se homogenizó en vórtex y se centrifugó brevemente.
 - La mezcla se transfirió a una columna/filtro (spin filter) y se centrifugó a 2240 x g ($\pm 20\%$) durante diez minutos.
 - NOTA: la mezcla fue protegida en todo momento de la exposición a la luz.
 - En la estación de preparación de chips (chip priming station) se colocó la jeringa cargada con aire (un ml), con su clip, (enroscándola firmemente).
 - En la estación se colocó un nuevo chip (high sensitivity DNA chip).
 - Se añadieron nueve µl de la mezcla del gel dye mix al pozo del chip marcado con **G**, agregando el volumen en el fondo y centro del pozo.
 - Se cerró la tapa de la estación de preparación y se presurizó (inyectó) el 1ml de aire manteniendo la presión durante 60 segundos.
 - Al transcurrir los 60 segundos, se liberó el émbolo del clip y se dejó durante 10 segundos. Posteriormente, lentamente, se continuó retrayendo el émbolo manualmente hasta la posición inicial de un ml.
 - Se abrió la estación y se agregaron nueve µl de la mezcla del gel dye mix en el fondo-centro de los pozos del chip marcados con G (pozos de la última columna filas 1,2 y 4).
 - La gel dye mix restante se almacenó en refrigeración (4 - 6 °C) después de ser utilizada.
- NOTA: Se programó el equipo Agilen 2100 Bioanalyzer introduciendo los datos de las muestras para comenzar la lectura en cuanto el chip estuviera listo.
- Se agregaron cinco µl de marcadores (High Sensitivity DNA Markers 35/10380 bp) en el fondo-centro de todos los pozos.

- Se agregó un μl de marcador de peso molecular (High Sensitivity DNA ladder) en el centro del pozo marcado con .
- Se agregó un μl de cada muestra (una muestras por pozo) en el centro de los pozos un a 11 (pozos de las columnas 1, 2 y 3 filas 1, 2 y 3, y pozos de las columnas 1 y 2 fila 4).
- A los pozos no utilizados se les agregó un μl de marcadores (High sensitivity DNA Maarkers 35/10380).
- Se colocó el chip en el vórtex de chips (Modelo MS2 S9 minishaker MARCA IKA) y se agitó a 2400 rpm durante un minuto (tiempo establecido automáticamente por el equipo).
- Se trasladó el chip al Agilen 2100 Bioanalyzer y se colocó de forma adecuada, cerrando la tapa e iniciando la lectura (electroforesis).
- Al finalizar el programa, se retiró el chip (se desechó), y se lavó el arreglo capilar del bioanalyzer con el chip de lavado lleno con un ml de agua bidestilada pura, colocándolo en el bioanalyzer (como cualquier otro chip), cerrando la tapa del equipo bioanalyzer (para que las puntas de los capilares se introduzcan en los pozos del chip de limpieza que contiene agua), esperando un minuto para luego retirar el chip de lavado. Por último se mantuvo la tapa del equipo abierta durante 20 segundos (para que el agua restante en los capilares se evaporara) y se cerró el equipo.
- Se obtuvo el informe de resultados para analizarlos.

Anexo 6. Adición de índices-adaptadores y amplificación de librerías.

Con los cDNA/DNA fragmentados de cada especie, se realizó la adición de secuencias índice-adaptadores, para conformar las librerías y amplificarlas siguiendo las indicaciones del fabricante a continuación descritas brevemente.

En un tubo nuevo de 0.2 ml previamente identificado se transfirieron 25 µl del cDNA/DNA fragmentado. Este paso se realizó con cada muestra-especie.

A cada muestra de cDNA/DNA fragmentado se le adicionaron los reactivos siguientes por separado homogenizando la mezcla:

- 15 µl de Nextera PCR Master Mix
- 5 µl del Index Adapter i5 503
- 5 µl de index Adapter i7 específico de especie, de la siguiente manera:

Index Adaptor i7	702	<i>Artibeus sp</i>
	703	<i>M. waterhousii</i>
	704	<i>P. parnellii</i>
	705	Porcinos

Una vez preparadas las reacciones se centrifugaron brevemente y se colocaron en el termociclador para realizar el siguiente programa:

72 °C durante tres minutos

95 °C durante 30 segundos.

Doce ciclos : 95 °C durante diez segundos

55 °C durante 30 segundos

72 °C durante 30 segundos

72 °C durante cinco minutos

Mantener a diez °C

Anexo 7. Purificación de librerías.

Para purificar las librerías se empleó el Agencourt AMPure XP kit siguiendo las instrucciones del fabricante a continuación descritas brevemente.

- Se homogenizó el reactivo Agencourt AMPure XP y se añadieron 50 μ l (1x volúmenes de muestra) a cada muestra homogenizando por pipeteo al menos cinco veces para promover la unión de las perlas con los fragmentos de la librería. Se incubó a temperatura ambiente durante cinco minutos.
- NOTA: Se preparó etanol al 70% (fresco) un ml/muestra. Una concentración mayor de etanol provoca lavados ineficientes (permanecen contaminantes como fragmentos cortos de DNA y enzimas de fragmentación); una concentración menor de etanol (evaporación) puede causar pérdida de la muestra (cDNA).
- Se centrifugaron las muestras un pulso y se colocaron en la gradilla magnética DynaMag 2 magnet durante tres minutos o hasta que la solución se aclaró.
- Cuidadosamente se recuperó y desecho el líquido sin tocar el concentrado de perlas adherido a la gradilla magnética.
- Sin retirar los tubos de la gradilla magnética se añadieron 300 μ l de etanol al 70% y se incubó por 30 segundos.
- Al transcurrir la incubación, se giraron los tubos 180 ° y se dejó que la solución se aclarara (30 segundos). Este paso se repitió y a continuación se retiró la solución sin tocar el concentrado de perlas.
- Se repitieron los dos pasos anteriores y posteriormente para remover el etanol residual se centrifugaron los tubos un pulso y se volvieron a colocar en la gradilla magnética permitiendo que las perlas se condensaran en la pared

del tubo para retirar el etanol absorbiéndolo con la pipeta con cuidado de no tocar ni absorber perlas.

- Se secó el tubo (evaporación del etanol) dejándolo abierto durante cinco minutos.
- Se quitó la gradilla de plástico del magneto y se agregaron 50 μ l de agua libre de nucleasas directamente al concentrado de perlas homogenizando la suspensión cinco veces (al menos) para después mezclar con vórtex durante 60 segundos y se dejó durante dos minutos a temperatura ambiente.
- Se centrifugaron los tubos un pulso de cinco segundos y se colocaron los tubos en la gradilla magnética por un minuto (suspensión aclarada).
- Con una pipeta se recuperó el líquido con el cDNA eluido (teniendo cuidado de no arrastrar perlas) y se transfirió a un nuevo tubo de 1.5 ml.

Anexo 8. Procedimiento de RT-PCR para la identificación/confirmación general de coronavirus en muestras de RNA de porcinos y murciélagos.

Los volúmenes empleados y las condiciones de termociclado se describen a continuación.

Volúmenes de reactivos del Verso 1-Step RT-PCR kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación de coronavirus (Mezcla-CoV).

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Verso enzyme Mix	0.12 µl	9.24 µl
MgCl ₂ 25 mM	0.60 µl	46.20 µl
Iniciador PC2S2-F 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador IZS-Cov-f-R 10µM (antisentido)	0.90 µl	69.30 µl
Ribolock (Bioline)	0.20 µl	15.40 µl
2x 1-step RT-PCR Hot Start Master Mix	7.50 µl	577.50 µl
Agua libre de nucleasas	1.98 µl	137.06 µl
Total Mezcla Maestra	12.00 µl	
RNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-CoV con todos los reactivos se homogenizó y se repartieron doce µl a cada uno de los 76 tubos de 0.2 ml (74 muestras y dos testigos) y se añadió el RNA correspondiente a cada tubo (en el tubo correspondiente a cada individuo muestreado se agregaron 1.5 µl de RNA de muestra nasal-cerdos o faríngea-murciélagos y 1.5 µl de RNA de muestra rectal), en el caso del testigo positivo se empleó RNA del virus de bronquitis infecciosa aviar vacunal (Nobilis IB 4-91, MSD Salud Animal) al tubo del testigo positivo y al tubo del testigo negativos se le agregó agua libre de nucleasas. Los tubos se

cerraron y se colocaron en el termociclador icycler (BIO-RAD) para llevar a cabo la RT-PCR con las siguientes condiciones.

Programa de RT-PCR (termociclado) para la identificación general de coronavirus

PASO	Temperatura	Duración
Retrotranscripción	49 °C	5 minutos
	50 °C	15 minutos
Inactivación de retortranscriptasa y activación de polimerasa	95°C	15 minutos
Reacción en cadena de la polimerasa 36 repeticiones	94°C	30 segundos
	49°C	30 segundos
	72 °C	21 segundos
Extensión final	72 °C	5 minutos
Mantenimiento	4°C	∞

Anexo 9. Procedimiento de RT-PCR para la identificación general de paramixovirus en muestras de RNA de porcinos y murciélagos.

Los volúmenes empleados para la mezcla se muestran a continuación.

Volúmenes de reactivos del Verso 1-Step RT-PCR kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación de paramixovirus (Mezcla-ParamixoV)

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Verso enzyme Mix	0.12 µl	9.24 µl
MgCl ₂ 25 mM	0.60 µl	46.20 µl
Iniciador PAR-F1 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador PAR-R 10µM (antisentido)	0.90 µl	69.30 µl
Ribolock (Bioline)	0.20 µl	15.40 µl
2x 1-step RT-PCR Hot Start Master Mix	7.50 µl	577.50 µl
Agua libre de nucleasas	1.78 µl	137.06 µl
Total Mezcla Maestra	12.00 µl	
RNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-ParamixoV con todos los reactivos bien homogenizada se repartieron doce µl a cada uno de los 76 tubos de 0.2 ml (74 muestras y dos testigos), se añadió el RNA correspondiente a cada tubo (en el tubo correspondiente a cada individuo muestreado se agregaron 1.5µl de RNA de muestra nasal o faríngea y 1.5 µl de RNA de muestra rectal). En el caso del testigo positivo se empleó RNA del rubulavirus porcino de la diarrea epidémica porcina (amablemente proporcionado por el Dr Humberto Ramírez Mendoza de la FMVZ-UNAM) y al tubo del testigo negativo se le agregó agua libre de nucleasas. Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la RT-PCR con las siguientes condiciones.

Programa de RT-PCR (termociclado) para la identificación general de paramixovirus

PASO	Temperatura	Duración
Retrotranscripción	50 °C	30 minutos
Inactivación de retortranscriptasa y activación de polimerasa	95°C	15 minutos
Reacción en cadena de la polimerasa	94°C	30 segundos
36 repeticiones	50°C	30 segundos
	72 °C	21 segundos
Extensión final	72 °C	5 minutos
Mantenimiento	4°C	∞

Anexo 10. Procedimiento de RT-PCR para la identificación del virus de hepatitis E en muestras de RNA de porcinos y murciélagos.

Los volúmenes empleados en las reacciones de RT-PCR se muestran a continuación.

Volúmenes de reactivos del Verso 1-Step RT-PCR kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación del virus de la hepatitis E (Mezcla-HepatitisEV)

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Verso enzyme Mix	0.12 µl	9.24 µl
MgCl ₂ 25 mM	0.60 µl	46.20 µl
Iniciador HEv-F-1679 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador HEv-R-1680 10µM (antisentido)	0.90 µl	69.30 µl
Ribolock (Bioline)	0.20 µl	15.40 µl
2x 1-step RT-PCR Hot Start Master Mix	7.50 µl	577.50 µl
Agua libre de nucleasas	1.78 µl	137.06 µl
Total Mezcla Maestra	12.00 µl	
RNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-HepatitisEV con todos los reactivos bien homogenizada se repartieron doce µl a cada uno de los 76 tubos de 0.2 ml (74 muestras y dos testigos) y se añadió el RNA correspondiente a cada tubo (en el tubo correspondiente a cada individuo muestreado se agregaron 1.5µl de RNA de muestra nasal o faríngea y 1.5 µl de RNA de muestra rectal), en el caso del testigo positivo se empleó DNA sintético (ORF1) replicado en bacterias transformadas y al tubo del testigo negativo se le agregó agua libre de nucleasas. Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la RT-PCR con las siguientes condiciones.

Programa de RT-PCR (termociclado) para la identificación de virus de la hepatitis E.

PASO	Temperatura	Duración
Retrotranscripción	50 °C	20 minutos
Inactivación de retortranscriptasa y activación de polimerasa	95°C	15 minutos
Reacción en cadena de la polimerasa	94°C	30 segundos
40 repeticiones	45°C	30 segundos
	72 °C	50 segundos
Extensión final	72 °C	10 minutos
Mantenimiento	4°C	∞

Empleando los productos de RT-PCR como material genético se realizó la PCR anidada con los siguientes volúmenes de reacción.

Volúmenes de reactivos del Taq polymerase kit (Thermo Fisher Scientific) empleados en las reacciones de PCR anidadas para la identificación del virus de la hepatitis E (Mezcla-HepatEV-nest).

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Verso enzyme Mix	0.12 µl	9.24 µl
MgCl ₂ 25 mM	0.60 µl	46.20 µl
Iniciador HEv-F-1681 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador HEv-R-1682 10µM (antisentido)	0.90 µl	69.30 µl
2x 1-step RT-PCR Hot Start Master Mix	7.50 µl	152.46 µl
Agua libre de nucleasas	1.98 µl	924.00 µl
Total Mezcla Maestra	12.00 µl	
RNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-HepatEV-nest con todos los reactivos bien homogenizada se repartieron doce µl a cada uno de los 76 tubos de 0.2 ml (74

muestras y dos testigos) y se añadió el DNA correspondiente a cada tubo (tres µl de la reacción de RT-PCR correspondiente). Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la PCR con las siguientes condiciones.

Programa de PCR anidado (termociclado) para la identificación de virus de la hepatitis E.

PASO	Temperatura	Duración
Desnaturalización inicial	95°C	5 minutos
Reacción en cadena de la polimerasa	94°C	30 segundos
40 repeticiones	45°C	30 segundos
	72 °C	22 segundos
Extensión final	72 °C	5 minutos
Mantenimiento	4°C	∞

Anexo 11. Procedimiento de RT-PCR para la identificación de alfaifluenzavirus en muestras de RNA de porcinos y murciélagos.

Los volúmenes empleados en la RT-PCR se muestran a continuación.

Volúmenes de reactivos del Verso 1-Step RT-PCR kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación de alfaifluenzavirus (Mezcla-influenzaV).

Reactivo	Volumen por muestra	Volumen por total de muestras (25)
Verso enzyme Mix	0.12 µl	2.76 µl
MgCl ₂ 25 mM	0.60 µl	13.80 µl
Iniciador IA-M-sen 10 µM (sentido)	0.90 µl	20.70 µl
Iniciador IA-M-rev 10µM (antisentido)	0.90 µl	20.70 µl
Ribolock (Bioline)	0.20 µl	4.60 µl
2x 1-step RT-PCR Hot Start Master Mix	7.50 µl	172.50 µl
Agua libre de nucleasas	1.78 µl	40.94 µl
Total Mezcla Maestra	12.00 µl	
RNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-InfluenzaV con todos los reactivos bien homogenizada se repartieron doce µl a cada uno de los 23 tubos de 0.2 ml (21 muestras y dos testigos) y se añadió el RNA correspondiente a cada tubo (en el tubo correspondiente a cada individuo muestreado se agregaron 1.5µl de RNA de muestra nasal y 1.5 µl de RNA de muestra rectal), en el caso del testigo positivo se empleó RNA viral de alfaifleunzavirus porcino (amablemente proporcionado por la Dra. Edith Rojas Anaya del CENID-Microbiología, INIFAP) y al tubo del testigo negativo se le agregó agua libre de nucleasas. Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la RT-PCR con las siguientes condiciones.

Programa de RT-PCR (termociclado) para la identificación de alfa-influenzavirus.

PASO	Temperatura	Duración
Retrotranscripción	50°C	30 minutos
Inactivación de RT, activación de polimerasa y desnaturalizaron	95°C	15 minutos
Reacción en cadena de la polimerasa	94°C	30 segundos
36 repeticiones	45°C	30 segundos
	72 °C	44 segundos
Extensión final	72 °C	2 minutos
Mantenimiento	4°C	∞

Anexo 12. Procedimiento de PCR para la identificación/confirmación de herpesvirus en muestras de DNA de porcinos y murciélagos.

Los volúmenes empleados en la primera reacción de PCR se muestran a continuación.

Volúmenes de reactivos del Taq polymerase kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación de herpesvirus (Mezcla-HerpesV).

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Taq polymerase	0.12 µl	9.24 µl
DNTP Mix 10 mM	0.30 µl	23.10 µl
Iniciador DFA 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador ILK 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador KGI 10µM (antisentido)	0.90 µl	69.30 µl
MgCl ₂ 25 mM	1.50 µl	115.50 µl
Buffer 10x	1.50 µl	115.50 µl
Agua libre de nucleasas	5.88 µl	446.60 µl
Total Mezcla Maestra	12.00 µl	
DNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-HerpesV con todos los reactivos bien homogenizada se repartieron doce µl a cada uno de los 76 tubos de 0.2 ml (74 muestras y dos testigos) y se añadió el DNA correspondiente a cada tubo (en el tubo correspondiente a cada individuo muestreado se agregaron 1.5µl de DNA de muestra nasal o faríngea y 1.5 µl de DNA de muestra rectal), en el caso del testigo positivo se empleó DNA viral (EHv2 proporcionado amablemente por la doctora

Edith Rojas Anaya del CENID-Microbiología, INIFAP). Al tubo del testigo negativo se le agregó agua libre de nucleasas. Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la RT-PCR con las siguientes condiciones.

Programa de PCR anidado (termociclado) para la identificación de herpesvirus.

PASO	Temperatura	Duración
Desnaturalización inicial	95°C	5 minutos
Reacción en cadena de la polimerasa 35 repeticiones	94°C 50°C 72 °C	20 segundos 20 segundos 20 segundos
Extensión final	72 °C	2 minutos
Mantenimiento	4°C	∞

Empleando los productos de la primera PCR como material genético se realizó la PCR anidada con los siguientes volúmenes de reacción.

Volúmenes de reactivos del Taq polymerase kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación de herpesvirus (Mezcla-HerpesV-nest)

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Taq polymerase	0.12 µl	9.24 µl
DNTP Mix 10 mM	0.30 µl	23.10 µl
Iniciador TGV 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador IYG 10µM (antisentido)	0.90 µl	69.30 µl
MgCl ₂ 25 mM	1.50 µl	115.50 µl
Buffer 10x	1.50 µl	115.50 µl
Agua libre de nucleasas	6.78 µl	522.06 µl
Total Mezcla Maestra	12.00 µl	
DNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-HepatEV-nest con todos los reactivos bien homogenizada se repartieron doce μl a cada uno de los 76 tubos de 0.2 ml (74 muestras y dos testigos) y se añadió el DNA correspondiente a cada tubo (tres μl de la reacción de PCR correspondiente). Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la PCR con las mismas condiciones que el pasado programa de termociclado.

Anexo 13. Procedimiento de PCR para la identificación/confirmación de virus tipo circovirus porcino 51 en muestras de DNA de porcinos y murciélagos.

Los volúmenes empleados en la reacción de PCR se muestran a continuación.

Volúmenes de reactivos del Taq polymerase kit (Thermo Fisher Scientific) empleados en las reacciones para la identificación de virus tipo-circovirus porcino 51 (Mezcla-CircoV51).

Reactivo	Volumen por muestra	Volumen por total de muestras (77)
Taq polymerase	0.12 µl	9.24 µl
DNTP Mix 10 mM	0.30 µl	23.10 µl
Iniciador F8b_Po-circo-like-virus-51_1760 10 µM (sentido)	0.90 µl	69.30 µl
Iniciador R9_Po-circo-like-virus-51_2483 10µM (antisentido)	0.90 µl	69.30 µl
MgCl ₂ 25 mM	1.50 µl	115.50 µl
Buffer 10x	1.50 µl	115.50 µl
Agua libre de nucleasas	6.78 µl	522.06 µl
Total Mezcla Maestra	12.00 µl	
DNA (por muestra)	3.00 µl	
TOTAL	15.00 µl	

Una vez preparada la mezcla-CircoV51 con todos los reactivos bien homogenizada se repartieron doce µl a cada uno de los 76 tubos de 0.2 ml (74 muestras y dos testigos) y se añadió el DNA correspondiente a cada tubo (en el tubo correspondiente a cada individuo muestreado se agregaron 1.5µl de DNA de muestra nasal o faríngea y 1.5 µl de DNA de muestra rectal), en el caso del testigo positivo se empleó DNA sintético del gen CAP replicado en bacterias transformadas y al tubo del testigo negativo se le agregó agua libre de nucleasas. Los tubos se cerraron y se colocaron en el termociclador para llevar a cabo la RT-PCR con las siguientes condiciones.

Programa de PCR (termociclado) para la identificación de virus tipo circovirus porcino 51.

PASO	Temperatura	Duración
Desnaturalización inicial	95°C	10 minutos
Reacción en cadena de la polimerasa	94°C	30 segundos
35 repeticiones	50°C	30 segundos
	72 °C	44 segundos
Extensión final	72 °C	2 minutos
Mantenimiento	4°C	∞

*Swine virome on rural backyard farms
in Mexico: communities with different
abundances of animal viruses and phages*

**Rodrigo Jesús Barrón-Rodríguez, Edith
Rojas-Anaya, Jorge Tonatiuh Ayala-
Sumuano, José Ángel Iván Romero-
Espinosa, et al.**

Archives of Virology

Official Journal of the Virology
Division of the International Union of
Microbiological Societies

ISSN 0304-8608

Arch Virol
DOI 10.1007/s00705-020-04894-y



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Austria, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



Swine virome on rural backyard farms in Mexico: communities with different abundances of animal viruses and phages

Rodrigo Jesús Barrón-Rodríguez^{1,2} · Edith Rojas-Anaya¹ · Jorge Tonatiuh Ayala-Sumano³ · José Ángel Iván Romero-Espinosa⁴ · Joel Armando Vázquez-Pérez⁴ · Moisés Cortés-Cruz⁵ · Gary García-Espinosa² · Elizabeth Loza-Rubio¹

Received: 6 May 2020 / Accepted: 6 October 2020
© Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract

Domestic swine have been introduced by humans into a wide diversity of environments and have been bred in different production systems. This has resulted in an increased risk for the occurrence and spread of diseases. Although viromes of swine in intensive farms have been described, little is known about the virus communities in backyard production systems around the world. The aim of this study was to describe the viral diversity of 23 healthy domestic swine maintained in rural backyards in Morelos, Mexico, through collection and analysis of nasal and rectal samples. Next-generation sequencing was used to identify viruses that are present in swine. Through homology search and bioinformatic analysis of reads and their assemblies, we found that rural backyard swine have a high degree of viral diversity, different from those reported in intensive production systems or under experimental conditions. There was a higher frequency of bacteriophages and lower diversity of animal viruses than reported previously. In addition, sapoviruses, bocaparvoviruses, and mamastroviruses that had not been reported previously in our country were identified. These findings were correlated with the health status of animals, their social interactions, and the breeding/rearing environment (which differed from intensive systems), providing baseline information about viral communities in backyard swine.

Introduction

The domestic swine *Sus scrofa domesticus* is an important species for the economy and food production. Swine are maintained under different conditions in diverse production systems (intensive, semi-intensive, extensive, free-roaming) and environments, where they are often in close contact with humans and other species. It is well known that agriculture, animal production, and other human activities have an impact on the interactions between domestic animals, wild species, and infectious agents. These interactions represent interfaces in which the risk for diseases is not often taken into account until outbreaks occur [1, 2].

Emerging and re-emerging diseases frequently affect swine. Introduction of swine by humans into new environments can lead to adaptation of viruses and genetic recombination between infectious agents. This can create novel viruses with zoonotic potential when transmission and spread are favored by the current conditions, such as land use changes (crops, livestock, urbanization), trade, and other anthropogenic phenomena [3]. Examples include influenza A virus [4], Nipah virus [5], Menangle virus [6, 7], Mapuera virus, rubulaviruses [8, 9], and Tioman virus [10, 11]. Most of these viruses use receptors on mucous tissue to infect their host. Thus, the study of these environments is attractive for studying the infectious agents, including the viral communities, which play an important role in animal health and, eventually, the health of consumers. Information about 1) the communities of microorganisms that inhabit swine under different production conditions, 2) the influence of the environment on these communities, and 3) the adaptations that confer pathogenicity to infectious agents and the relationships between all of these factors is still scarce.

Advances in sequencing techniques have made it possible to obtain more-precise information about the diversity of

Handling Editor: Ana Cristina Bratanich.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00705-020-04894-y>.

✉ Elizabeth Loza-Rubio
eli_rubio33@hotmail.com

Extended author information available on the last page of the article

Published online: 04 January 2021

Springer

the viral communities present in different organs and tissues of individuals of a particular species in different environments. These advances have allowed researchers to identify the viruses present in animal samples, providing a better understanding of both new and previously known viral infectious agents [12].

Some studies have been conducted on the swine virome [13]; however, the majority have focused on animals maintained on high-tech farms with large population densities or in facilities with controlled experimental conditions [14, 15]. In Mexico, due to its economic circumstances, one of the most important production systems is the backyard farm, which is an important source of food and income [16]. Backyard farming is a diverse production system that is common in rural and peri-urban areas of many regions of the country. In this system, animals can be fed with commercial feed, agricultural products and by-products, or human food leftovers or allowed to graze. Backyard breeding management varies according to the breeder and the environment. Animals are housed in rustic pens or outdoors, where they frequently coexist in close contact with other species of domestic animals and humans, and can occasionally come into contact with wild animals. This situation favors the spread and adaptation of infectious agents [17]. Because there is

little information about viral communities in swine that are bred and raised in backyards, this work aimed to study the virome of these animals in Mexico.

Materials and methods

Animals and samples

Four backyard farms in border villages and rural areas in the municipalities of Tlaquiltenango (farms A and D, Quilamula and Chimalacatlán, respectively), Puente de Ixtla (farm B, Xoxocotla), and Tlaltizapan (farm C) in the state of Morelos, Mexico, were surveyed in this study (Fig. 1, Table 1). The selection of these farms was made base on their location in rural populations surrounded by wild environments, where it is possible to find wild species such as carnivores, procyonids, canids, rodents, diverse birds, and bats. These farms also house other species of domestic animals, such as cattle, sheep, goats, poultry, cats, and dogs living in close contact. The sampled farms are located in a tropical savanna climate with dry or semi-dry variants (Aw1 or Aw0; based on the Köppen-Geiger climate classification system), with an average annual temperature of 23.5 °C to 25°C, drought from



Fig. 1 Location of swine farms (P) from which samples were obtained in Xoxocotla (Xox), Chimalacatlán (Chi), Quilamula (Qui) and Tlaltizapan (Tla) in Morelos, Mexico. The location of the study area (large map) is represented on the map of the Mexican Republic

(left corner, small map) with a dark rectangle (■). The map was made with ArcGIS online (<https://www.arcgis.com/home/webmap/viewer.html>). Date accessed: August 13, 2017.

Table 1 Information on the samples and location of the rural backyard farms in Morelos, Mexico

Farm	Farm location	Geographical location	Population (swine)	Sampled animals	Age (months)	Date
A	Quilamula, Tlaquiltenango	18°30'36" N, -99°01'17" W	15	5	3	February 2016
B	Xoxocotla, Puente de Ixtla	18°39'49" N, -99°14'17" W	10	7	4	March 2016
C	San Miguel 30, Tlaltizapan	18°43'7" N, -99°7'9" W	4	4	4	March 2016
D	Chimalacatlán, Tlaquiltenango	18°27'46" N, -99°5'37" W	14	7	2–3	July 2016

The column "Age" refers to the age of the sampled animals.

late autumn to early spring, and annual rainfall of 840 to 930 mm [18, 19]. Swine on these farms were kept in earthen- or cement-floor pens, with a larger area per animal when compared to the standard established for intensive systems (0.36–0.66 m²/animal), and therefore, the population density was lower (a minimum of 10 m²/animal) (Supplementary Fig. 1). The feeding consisted mainly of commercial feed, corn, agricultural by-products, and, in some cases, free-roaming. On farms A, B, and D, animals were living in mixed-age groups. The goals of farms A, C, and D were to breed animals for farrow-to-finish or to-nursery purposes, whereas farm B was a wean-to-finish farm. In general, there were no preventive medicine or deworming programs on any of the farms studied.

The health status of the animals was determined through anamnesis and clinical examination, which included, but was not limited to, observation of normal behavior, skin, abnormalities in body shape, or swelling on the body surface, stool, neurological or respiratory signs, and lameness. A body temperature of 38.8°C–39.2°C was considered normal, and the normal respiratory rate was 25–40 breaths per minute. In addition, mucous membrane color was monitored, capillary refill time was measured (≤ 2 s), auscultation of the abdomen was done, and the parotid, submandibular, and inguinal lymph nodes were examined.

One nasal swab and one rectal swab from 23 animals (two and four months of age) were collected between February and July 2016. These samples were placed in 1 ml of Eagle's minimum essential medium (EMEM) at 2°C to 6°C and transported to the Laboratory of Biotechnology on Animal Health (CENID Microbiología Animal - INIFAP) for the immediate RNA/DNA extraction.

RNA/DNA extraction

The samples were homogenized by vortexing and centrifuged (6000 $\times g$ for 15 min) to eliminate cells, bacteria, and debris. A pool was made by transferring a volume of 200 μ l of the supernatant of each sample to a 15-ml tube.

The nasal-rectal pool was filtered through a 0.45- μ m filter and a 0.22- μ m filter (Millipore) to remove eukaryotic and bacterial-cell-sized particles, followed by fractional ultracentrifugation at 170,000 $\times g$ for 3 hours at 4°C (Beckman Coulter Optima Ty-70) to pellet the viral particles. The pellet was resuspended in EMEM solution and filtered using a 0.22- μ m filter. The filtrate was treated with DNase (Thermo Scientific) and RNase A (Thermo Scientific) at 37°C for 30 minutes to digest non-particle-protected nucleic acids. The total viral nucleic acids (both RNA and DNA) were extracted and purified using a Purelink Viral RNA/DNA Mini Kit (Invitrogen) according to the manufacturer's protocol. A sample with EMEM alone was used as a blank for RNA/DNA extraction. The concentration of RNA/DNA was estimated using a NanoDrop spectrophotometer (Thermo Fisher Scientific). The purified RNA/DNA was divided in two fractions, one as an RNA mix and the other as a DNA mix for the enrichment of the purified nucleic acids.

DNA library construction and next-generation sequencing

Complementary DNA (cDNA) was synthesized from the RNA mix by random-primer-mediated sequence-independent single-primer amplification (RP-SISPA) using a chimeric M13 primer with a random hexamer attached to its 3' end [20]. Reverse transcription was performed using a Transcriptor First Strand cDNA Synthesis kit V.6 (Roche). Then, cDNA and DNA were denatured, followed by annealing of random primers and the use of the Klenow enzyme (New England Biolabs) to generate dscDNA and dsDNA, which were subsequently used as templates for polymerase chain reaction (PCR) amplification using M13 primers. Equimolar amounts of cDNA and DNA were pooled in one mix. The amplicon concentration was determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), and the quality of the preparation was determined by capillary electrophoresis in a 2100 Bioanalyzer (Agilent Genomics). A DNA library was prepared using a Nextera XT DNA Library

Preparation Kit and identified using a Nextera XT index kit (Illumina). Sizes and concentrations of fragmented libraries were determined using a Qubit fluorometer (Thermo Fisher Scientific) and a 2100 Bioanalyzer. The pooled library was sequenced on an Illumina MiSeq platform using a MiSeq Reagent Kit (v2). The sequencing protocol was performed according to manufacturer's instructions in a MiSeq system (Illumina) with 2×250 -bp cycles at the National Institute of Respiratory Diseases (INER).

Bioinformatics analysis

Data quality analysis was carried using the FastQC program (Braham Institute, Cambridge). FastQC reads were edited and filtered using the Trimmomatic v.0.33 command-line tool to remove chimeric M13 primers and any other remaining Illumina-specific sequences from the reads and to select the sequences with sufficient quality (Phred score ≥ 30) and length for subsequent analysis (Supplementary Fig. 2).

Taxonomic identification

Filtered reads were compared to whole genome sequences (wgs) and nucleotide (nt) databases (NCBI) using BLASTn (E-value, $<10^{-4}$). All sequences without nucleotide sequence homologs were compared to the UniProt (database of amino acid sequences) using BLASTx (E-value, 10^{-5}). Sequences corresponding to the swine genome were discarded for this analysis. Then, sequence alignments, obtained by BLASTn and BLASTx, were filtered using the "filter" command in Galaxy v1.1.0 (similarity $>50\%$). Filtered alignments were assigned to taxonomic groups using the "fetch taxonomic representation" command in Galaxy version 1.1.0. All sequences assigned to eukaryotes, protozoa, fungi, and bacteria were discarded, and the reads assigned to viruses were kept. The lowest taxonomic rank was obtained using the "find lowest diagnostic rank" command in Galaxy version 1.0.1, and the "summarize taxonomy" command in Galaxy version 1.0.0 was used to assign a taxonomic rank to the reads [21].

Assembly and phylogenetic analysis

All reads were assembled *de novo* using Velvet software [22, 23] to obtain contigs. Contigs were compared against whole genome sequences (wgs) and nucleotide (nt) databases using BLASTn and BLASTx, following the same steps described in the section on taxonomic assignment. The longer contigs of the more abundant animal viruses found in this study were selected to perform multiple sequence alignments using Clustal Omega. The alignment results were used to

infer the phylogenies in MEGA X software [24], employing the neighbor-joining method [25] (performed with a 1000 bootstrap replicates and the maximum composite likelihood method [26] to compute evolutionary distances). These analyses were used to identify the taxonomic groups to which the contigs belonged, using ICTV demarcation criteria.

Sequence accession numbers

Sequences of contigs of the more-abundant animal viruses were submitted to the GenBank (NCBI) database under accession numbers MH490914, MH490915, MH490916, MH490912, MH490913, MH490911 and MH479904.

Statistical analysis

The rarefaction of the dataset of the swine virome was estimated using iNEXT [27, 28] to show that the results obtained by massive sequencing were representative of the viral community in the sampled swine. The Shannon index (equity index) was used to assess the diversity of mammalian viruses in the reads.

Results

Overview of sequencing data

A total of 2,736,522 raw reads were generated, 2,719,920 of which were of sufficient quality (Phred score ≥ 30). Of those 84.8%, with an average length greater than 250 bases did not exhibit similarity with known sequences in genome and nucleotide databases (NCBI-NIH). BLASTn and BLASTx searches revealed that 80,863 sequences were similar to sequences from prokaryotic or eukaryotic organisms. There were 322,961 sequences related to viruses (Fig. 2a), 312,646 of which were related to phages, 10,212 to animal viruses, 83 to protozoan viruses, and the rest to plant viruses (Fig. 2b).

Identified viruses and their diversity

Reads were related to members of 40 species of animal viruses, one species of plant virus, four species of protozoan viruses, and 18 genera of phages. Of the total number of reads with similarity to viral sequences, 312,688 belonged to dsDNA viruses, 8,530 to ssDNA viruses (of which 90.5% corresponded to circular ssDNA viruses), and 1,722 to ssRNA viruses. Regarding the animal virus reads, 99.7% belonged to members of the families *Circoviridae*, *Caliciviridae*, *Parvoviridae*, and *Astroviridae*, and the unclassified immunodeficiency-associated stool (IAS) virus (Fig. 2c,

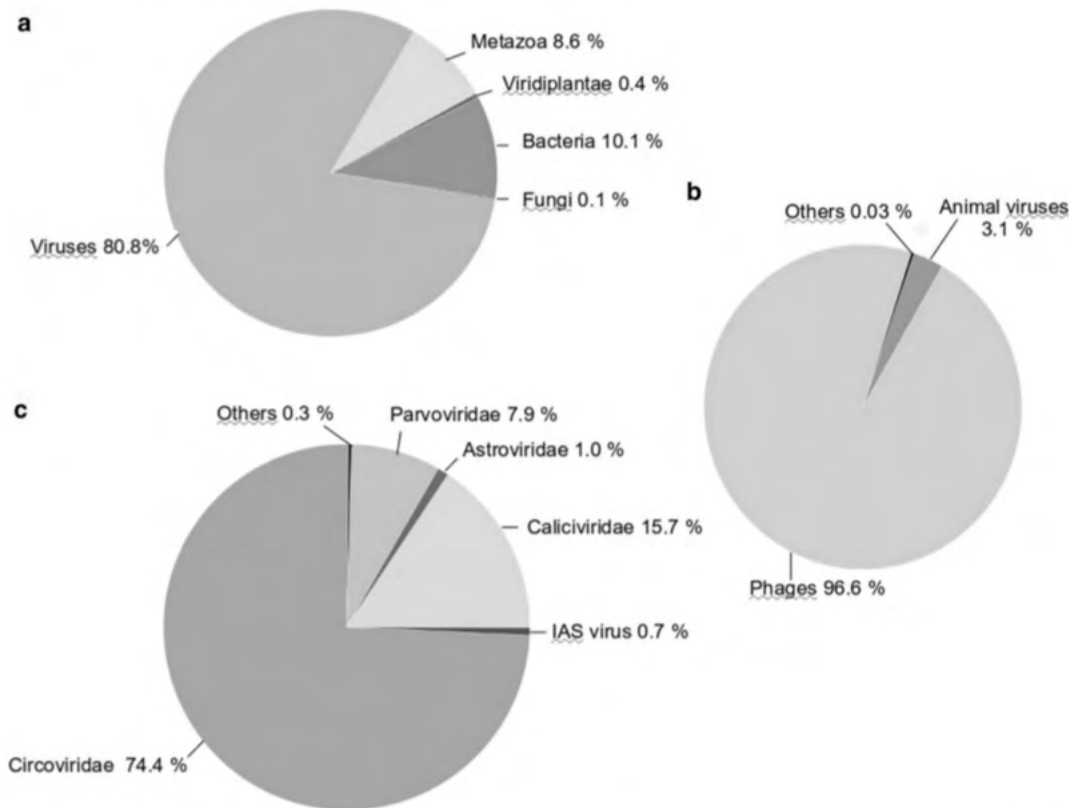


Fig. 2 Summary of the sequence classification of 46 nasal and rectal samples from backyard swine. The reads were classified at different taxonomic ranks by their assigned homology results of BLASTn and BLASTx. **(a)** Percentages of sequences classified by kingdom

taxonomic rank to those of Metazoa (animals), Viridiplantae (plants), bacteria, fungi, and viruses. **(b)** Percentages of sequences classified by virus host. **(c)** Percentages of sequences of animal viruses classified by both viral family and genus and type of nucleic acid.

Table 2 and Supplementary Table 1). Of the 312,646 reads related to bacteriophages, 213,965 (68.4%) corresponded to members of established viral genera or species, among which the most represented (99.85% of the specific reads assigned to genus or species of phages) were *Tunavirus* (23.5%), *Kagunavirus* (21.6%), *Jerseyvirus* (19.0%), *Tequatrovirus* (12.1%), *Teseptimavirus* (5.1%), *Yokohamavirus* (4.7%), *Vequintavirus* (2.8%), *Roufvirus* (1.6%), *Salasvirus* (1.1%) and *Tequintavirus* (1.0%).

The rarefaction showed that results obtained by next-generation sequencing (NGS) were representative of the viral community in the sample (Fig. 3). The Shannon index of the diversity of mammalian viruses was 1.12.

Assembly and phylogenetic analysis

Homologs were found for 2665 contigs (4.6% of all contigs; estimated coverage, 2.86), 2083 (78.2%) of which were related to viruses (97% corresponding to members of the

order *Caudovirales* and 3% to animal viruses), 381 to bacteria (14.3%), and 48 (1.8%) to eukaryotic organisms, with the kingdom Metazoa being the most represented, with 35 reads. The results of the BLAST obtained using contigs coincided with the results obtained using individual reads, identifying the same families of bacteriophages (*Podoviridae*, *Siphoviridae*, and *Myoviridae*) and animal viruses (*Bocaparvovirus*, *Circovirus*, *Sapovirus*, *Mamastrovirus*, and others). Seven hundred thirty-one assemblies were longer than 200 bp in length (202–10,617 bp, mean, 374.9 bp).

The four longer contigs, representative of the more abundant animal viruses found in this study, were selected for phylogenetic analysis (Table 3). These contigs were submitted to the GenBank databases. Similar known and reference sequences of viruses of the families *Caliciviridae* (nt \leq 89%; aa \leq 89%), *Circoviridae* (nt \leq 98%; aa \leq 97.7%), and *Parvoviridae* (nt \leq 96%; aa \leq 98.7%) were obtained to perform multiple sequence alignment and phylogenetic analysis with their respective contig according to ICTV

Table 2 Viruses identified by BLAST analysis of the reads from the rural backyard swine dataset. Viruses represented by more than 15 reads were included in this table.

Viral family (reads)	Viral genus (reads)	Viruses (reads)
Animal viruses		
<i>Astroviridae</i> (99)	<i>Mamastrovirus</i> (98)	Astrovirus wild boar/WBAstV-1/2011/HUN (17);
<i>Caliciviridae</i> (1605)	<i>Sapovirus</i>	Sapporo virus (1461); porcine sapovirus (144);
<i>Circoviridae</i> (7612)		Po-Circo-like virus 51 (6781); fur seal faeces associated circular DNA virus (176); pig stool associated circular ssDNA virus (119); Po-Circo-like virus 22 (46); porcine associated stool circular virus (26);
	IAS virus (67)	Human immunodeficiency-associated stool virus (67)
<i>Parvoviridae</i> (808)	<i>Bocaparvovirus</i> (775), <i>Dependoparvovirus</i> (30)	Porcine bocavirus 3 (608); bocavirus pig/6V/China/2010 (45); Adeno-associated virus (20)
<i>Picornaviridae</i> (33)		Posavirus 1 (18)
Bacteriophages		
<i>Microviridae</i> (41)		
<i>Myoviridae</i> (56224)		^a T4likevirus (37824), Msw3virus (14820), V5virus (8678), Twortlikevirus (593), Kayvirus (458)
<i>Podoviridae</i> (36158)		^a T7likevirus (15976), Kp34virus (4356), Phi29likevirus (3321), Sp6likevirus (1305), Ahj-likevirus (352), N4likevirus (46)
<i>Siphoviridae</i> (220 221)		^a Tunalikevirus (73355), K1givirus (67632), Jerseyvirus (59439), Pis4avirus (5173), T5likevirs (3240), Hk578virus (2852), Nonanavirus (2053), Sextaevirus (980)
Protozoan viruses		
<i>Mimiviridae</i> (83)		

^anames included in this cell correspond to viral genera

The number in parentheses indicates the number of reads assigned to the taxonomic group.

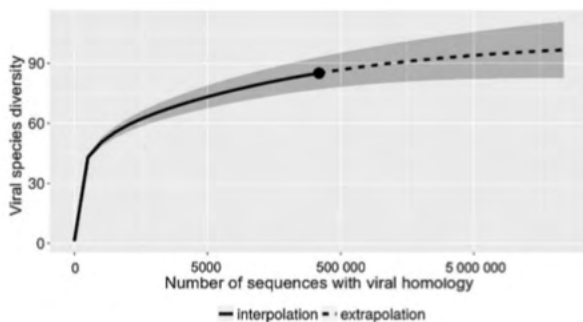


Fig. 3 Diversity of virus species estimated by sequence rarefaction (●). This graphic shows that viral species identified by metagenomics provided a good representation of the estimated real diversity (extrapolation).

demarcation criteria. The circo-like virus contig found in this work was analysed against reference sequences of the genus *Circovirus*. These viruses grouped with other porcine circo-like viruses but were different from porcine circovirus 1 and porcine circovirus 2 (Fig. 4). Another contig that was similar to the CAP gene (963 bp) of circoviruses

showed the most similarity to fur seal faeces associated circular DNA viruses found in otariids and swine. This contig was compared with reference sequences of the family *Circoviridae*. It was also distant from porcine circovirus 1 and porcine circovirus 2 (Fig. 5). The polyprotein contig of porcine sapovirus was analyzed using reference sequences of member viruses of the family *Caliciviridae*. The contig was grouped with porcine sapoviruses (Fig. 6). Figure 7 shows a phylogeny constructed with reference sequences of members of the family *Parvoviridae* in which the contig analyzed was more similar to sequences of ungulate bocaparvovirus 5, viruses typically found in swine.

Discussion

In this study, we identified new variants of circular ssDNA viruses (circo-like virus, fur seal-associated circular ssDNA virus, and IAS virus), sapoviruses (family *Caliciviridae*), bocaparvoviruses (family *Parvoviridae*), and mamastroviruses (family *Astroviridae*) present in backyard swine in Mexico. It is relevant to point out that backyard family farms

Table 3 Information about the longest contigs (sequences) of the most abundant animal viruses in this study

Name of sequence	Accession number	Length (bp)	Contig coverage	RefSeq/sequence with highest similarity in GenBank	Sequence identity	Gene(s)	Gene coverage ^a
Po-Circo-like-virus/ swine/Mex/2016/ prot1 ^b	MH490914 MH490915 MH490916	369 309 422	2.31 2.29 3.19	Po-circo-like virus 51 NC 025684	98%	Cap Cap - 2 ^c 2 ^d	70% 15 - 35% 64%
Po-Circo-like-virus/ swine/Mex/2016/1-2 Po-Circo-like-virus/ swine/Mex/2016/prot2							
FSFA-CirV/swine/ Mex/2016/Cap_ partial-CDS	MH490912 MH490913	963 521	3.88 3.61	Fur seal feces associ- ated circular DNA virus KF246569	96.15% 96.16%	Cap Rep	67% 48%
FSFA-CirV/swine/ Mex/2016/Cap_ partial-CDS							
Sapovirus/swine/ Mex/2016/ORF1- polyprotein	MH490911	1044	33.4	Porcine sapovirus strain JJ259 KT922089	89%	ORF-1	15.4%
Bocaparvovirus_Swine/ Mex/2016/NS1	MH479904	1268	4.59	Porcine bocavirus 3 strain 22 JF713714	96.34%	NS1	58.1%

^aPercentage of contig coverage with respect to the RefSeq/sequence genome with the most similarity

^bSequence used for phylogenetic analysis

^cSequence involving the final fragment of capsid gene and the start of the hypothetical protein 2 gene

^dPartial sequence of the hypothetical protein 2 gene

in Mexico often have a variety of domestic animals, mainly in rural and peri-urban areas, as is the case in many other developing regions of the world. In this system, animals can be fed with commercial feed, agricultural products, or human-food leftovers, or allowed to graze. Their management varies according to the breeder and the farm resources. This description of the virome of healthy rural backyard swine in Mexico provides evidence of new virus variants that had not been reported previously in the country [29].

Previous studies on the swine virome analyzed feces of animals maintained on intensive high-tech farms with high population densities and under experimental conditions [13–15]. There is no information about the virome of swine in a backyard system in this kind of climate zone in Mexico. The most abundant porcine viruses reported so far have been kobuviruses, rotaviruses, circular ssDNA viruses, astroviruses, and enterovirus B [30]. Previous studies [12, 14, 31] showed sequences that represented between nine and 25 genera and six to ten families of animal viruses, whereas our sequences represented viruses of five genera and four families. The backyard swine in this study had a lower diversity of known animal viruses. This might be due to the virome sequencing strategies employed, such as sample processing and pooling, among others. Sample filtration, dilution,

and pooling, as well as DNA/RNA purification and enrichment, like other sequencing steps and strategies, might vary between sequencing workflows, depending on many factors such as sample type, sequencing objective, resources, etc. These variations can affect sequencing results, but in most of cases, their effect is not well understood. Since there are no standards for virome sequencing, we must face all of the challenges inherent to virome metagenomics field [32]. Swine production environments, unlike intensive systems where animals are separated by weight and age and do not allow natural social interactions, should be studied, since these environmental factors are likely to influence which viruses are present [2, 33, 34].

The reads of phages provide evidence of the presence of their hosts, such as bacteria belonging to the genera *Escherichia* (37.1%), *Shigella* (25.3%), *Salmonella* (21.4%), *Klebsiella* (6.1%), *Citrobacter* (2.3%), *Bacillus* (0.9%) and *Staphylococcus* (0.7%). Previous studies have described the viromes of both healthy swine and those with gastrointestinal disorders, showing that diseases are frequent at younger ages in intensive systems. Sachsenröder et al. [30] reported that the proportion of bacteriophages in pigs increases with age, finding proportions of reads related to bacteriophages ranging from 22% to 44%, 68% to 72%, and 88% to

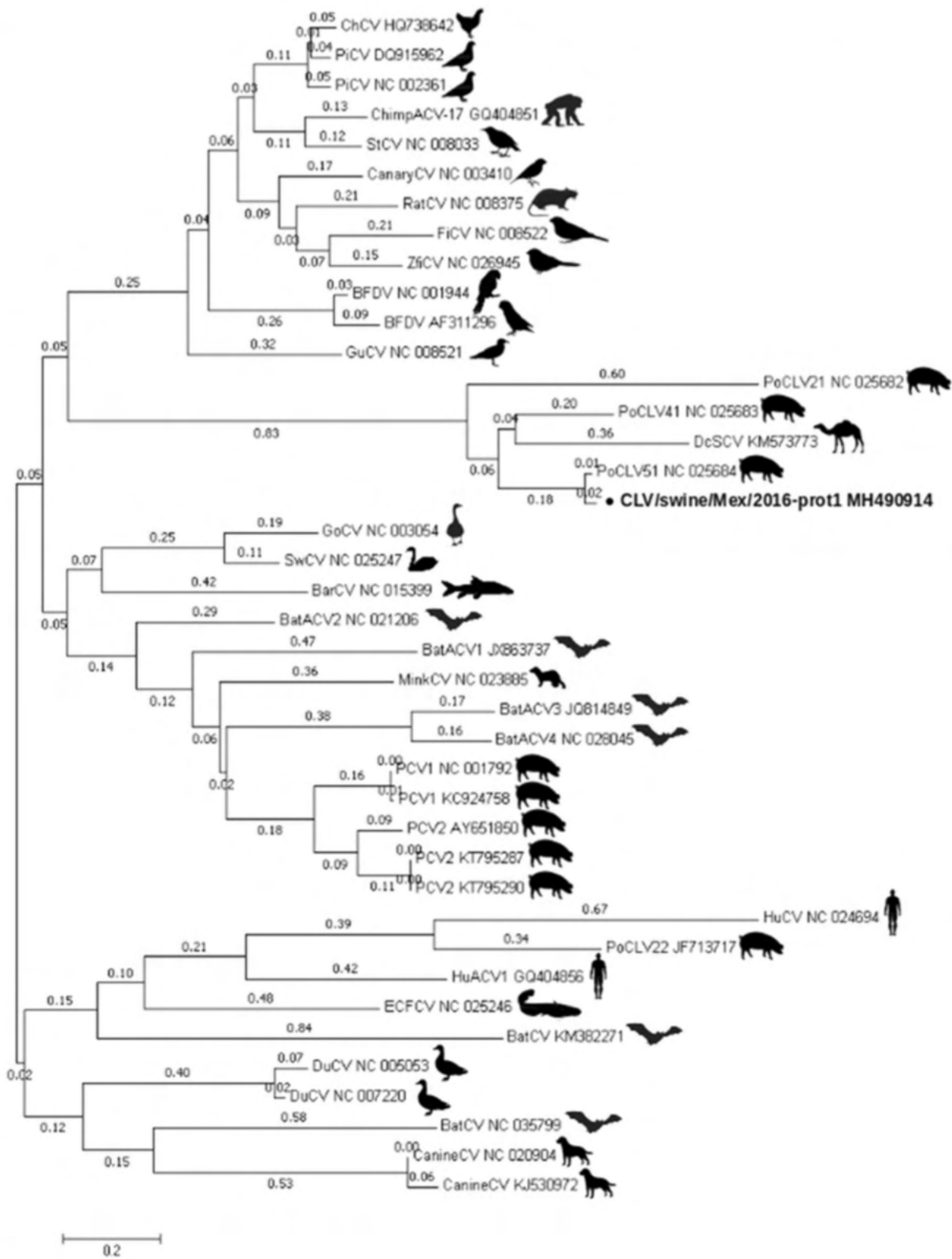


Fig. 4 Phylogenetic tree based on nucleotide sequences of the putative protein-1 (capsid) gene of members of the genus *Circovirus*. It shows the evolutionary relationship of a circo-like virus contig of 369 bases found by next-generation sequencing (bold) in rural backyard swine. The phylogeny was inferred using the neighbor-joining method, while the evolutionary distances were computed using the maximum composite likelihood method. The analysis included 40 reference nucleotide sequences of members of the genus *Circovirus*. Evolutionary analyses were carried out in MEGA7 [24–26].

92% for 12-day-old pigs, 54-day-old pigs, and adult sows, respectively. We found a high percentage of bacteriophage sequences (96.6% of total viral reads), which demonstrates a higher load of these viruses at younger ages than in previous porcine virome studies [13, 31, 35], even higher than those reported for adult sows [30]. In previous reports of swine viromes, phage communities were not well studied. Essential biological functions of these viruses, such as their role as regulators of bacterial growth, should be studied in the future [36–38].

We found numerous sequences with similarity to porcine circo-like virus 51 (nt \leq 98%; aa \leq 98.2%), which was previously identified in fecal samples from piglets on a high-density farm in the USA [14], and to fur seal feces-associated circular ssDNA virus (FSFACV) (nt \leq 97.6%; aa \leq 97.7%), previously identified in feces of *Arctocephalus* sp. and pigs from Japan [39, 40]. We found contigs related to two new variants of circoviruses, with the major proportion (6781 reads) similar to porcine circo-like virus 51. Circular ssDNA viruses are commonly detected in animals, other organisms, and environmental samples. However, further investigations of circular ssDNA viruses are needed to determine their possible pathogenicity [41–44].

We identified 1,605 reads related to members of the genus *Sapovirus* of the family *Caliciviridae*, making them second in abundance. These viruses are associated with gastroenteric illness in both swine and humans [45, 46]. All of the animals that were sampled were clinically healthy despite being infected with sapoviruses. This finding is in agreement with other studies [13, 15, 31]. The contig of porcine sapovirus ORF-1 found in this study showed \leq 89.1% nucleotide and \leq 96.5% amino acid sequence identity to PoSaV-JJ259 and other sapoviruses of genogroup 3 (GIII), which is the prototypic porcine genogroup [47, 48] (Fig. 6). Since there is evidence of human sapoviruses in Mexico [49], infection of backyard swine with sapoviruses needs to be studied further, since these viruses cause similar diseases in animals and humans, and inter-species transmission (mainly human-animal) and recombination are possible [48–50].

Most of reads related to members of the family *Parvoviridae* (808 reads) showed similarity to sequences from ungulate bocaparvovirus 5. A contig with similarity to the NS1

gene, which is used to characterize bocaparvoviruses, was used for phylogenetic analysis showed \leq 96.3% nucleotide sequence identity (aa \leq 100%) to bocaparvovirus 5 (previously classified as bocavirus 3) (Fig. 7). These viruses have been reported in North America, Asia, the UK, Eastern Europe, and Africa [51] and have been implicated in digestive and respiratory disorders (e.g., multisystemic wasting syndrome) in swine. Previous reports on the fecal viromes of pigs have described bocaparvoviruses in diarrheic animals [14], whereas in this study, the animals showed no clinical signs.

Ninety-eight reads with similarity to mamastrovirus sequences (nt \leq 100%; aa \leq 100%) were found, demonstrating their presence in swine in Mexico and confirming that these viruses commonly infect pigs without causing detectable disease, as reported by Luo et al. [52]. It should be pointed out that mamastroviruses have also been associated with neurological disorders in humans and minks [53, 54].

Reads corresponding to IAS virus were also found. IAS virus was first reported in the virome of swab pools collected from primary and secondary market slaughterhouses and a cull-swine buying station in the southeastern USA [12]. This study is the second report of the presence of IAS virus in the virome of swine.

In this study, the pooling of nasal and rectal samples from different age groups and farms limited our ability to evaluate whether they were present in mixed populations. Although we had intended to identify the viruses in individual samples by PCR/RT-PCR, the nucleic acids were consumed in the process of obtaining high-quality libraries. However, the major proportion of the viral sequences found in this study were related to enterobacterial phages and gastrointestinal animal viruses, and only a small proportion were likely to have been of nasal origin (members of the genera *Rosenblumvirus*, *Twortvirus*, and *Sextaecvirus* and the unassigned *Staphylococcus* phage Sb1). In any case, the origin of some viruses (i.e., parvovirus) was not clear. It is necessary to sequence samples individually or pool them according to specific criteria to obtain particular information about specific viral communities. In previous studies, two of which analyzed the nasal virome, one of them analyzed the nasal virome of feral swine, finding ssDNA viruses and orthopneumoviruses [55], whereas the other did not establish differences between the rectal and nasal viromes [12].

In general, the viral richness in this study was lower than in previous reports. The Shannon index obtained in this study indicated a low viral diversity (<2) in comparison to other reports (>4) [13, 30]. Including our results, most of the existing studies on the porcine virome have reported that the viral communities in swine kept on high-tech farms in intensive systems show a higher diversity of reads related

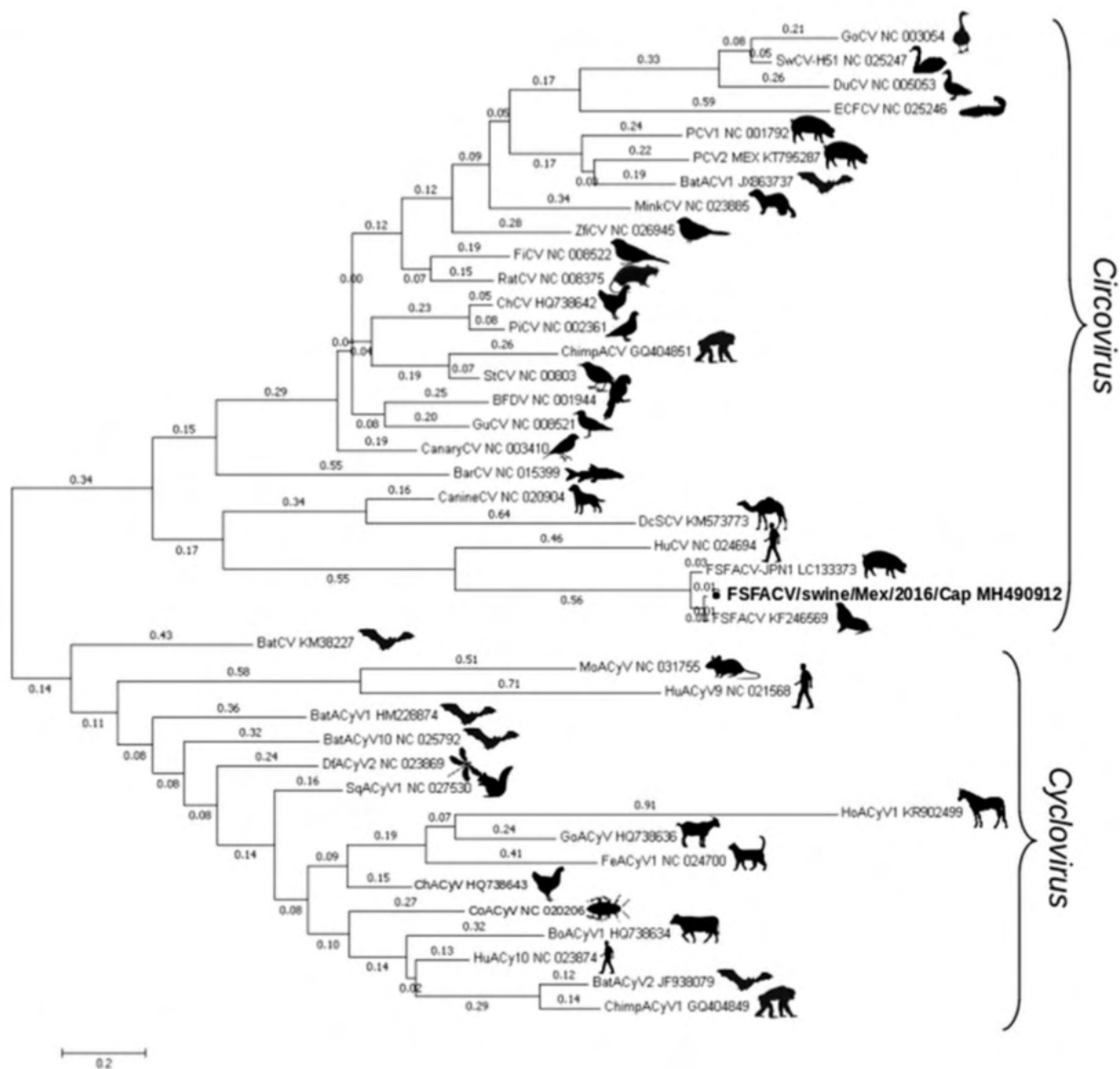


Fig. 5 Phylogenetic tree based on the Cap gene of members of the family *Circoviridae* (genera *Circovirus* and *Cyclovirus*), showing the evolutionary relationship of the swine fur seal-associated circovirus contig (963 bases) found by next-generation sequencing (bold) in rural backyard swine. The phylogeny was inferred using the neigh-

bor-joining method, while the evolutionary distances were computed using the maximum composite likelihood method. The analysis included 40 reference nucleotide sequences of members of the genera *Circovirus* and *Cyclovirus*. Evolutionary analysis was conducted in MEGA7 [24–26].

to several families of animal viruses, which translates into a higher risk of emerging diseases. The use of intensive systems for raising animals has been identified as a risk factor for the occurrence of infectious disease [2]. The backyard swine in this study showed a lower diversity of animal viruses, which could be related to their breeding environment and social interactions.

The use of metagenomics has allowed us to identify viruses not previously reported in Mexico. Not only is this technique useful for identifying viruses and determining their abundance in samples, but it also allows new variants and new viruses to be identified [12]. It also serves as a basis for investigating the complexity of the swine virome. It is necessary to establish surveillance systems for sapoviruses,

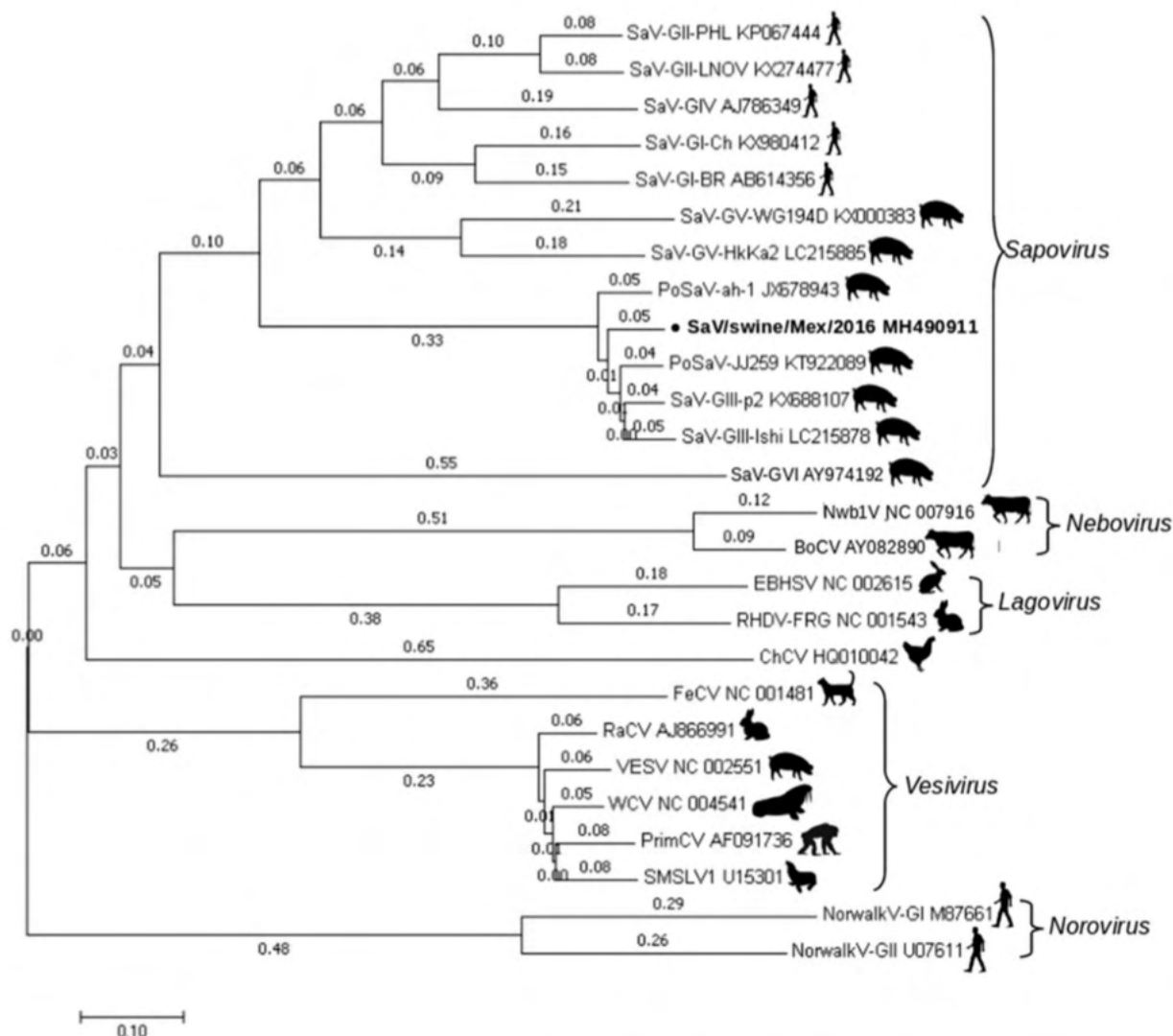


Fig. 6 Phylogenetic tree based on nucleotide sequences of the Cap gene of members of the genus *Sapovirus*, showing the evolutionary relationship of a contig of 1044 bases found by next-generation sequencing (bold) in rural backyard swine. The phylogeny was inferred using the neighbor-joining method, while the evolutionary

distances were computed using the maximum composite likelihood method. The analysis included 25 reference nucleotide sequences of members of the family *Caliciviridae*. Evolutionary analysis was carried out in MEGA7 [24–26].

bocaparvoviruses, and mamastroviruses and to determine their relationship to respiratory, digestive, reproductive, and neurologic diseases in swine in Mexico.

Conclusion

Our study is the first to evaluate the virome of rural backyard swine in Mexico, providing further information about the viral communities in swine bred under

conditions other than intensive high-tech systems. Bacteriophages may play essential roles in the swine virome, potentially controlling bacterial growth, including that of pathogenic species. We have also identified new viruses that had not previously been reported in the country and need to be monitored. This confirms that the swine virome can vary among different production systems and geographical locations.

Our findings suggest the need for a more in-depth analysis of bacteriophages and their functions in swine. These

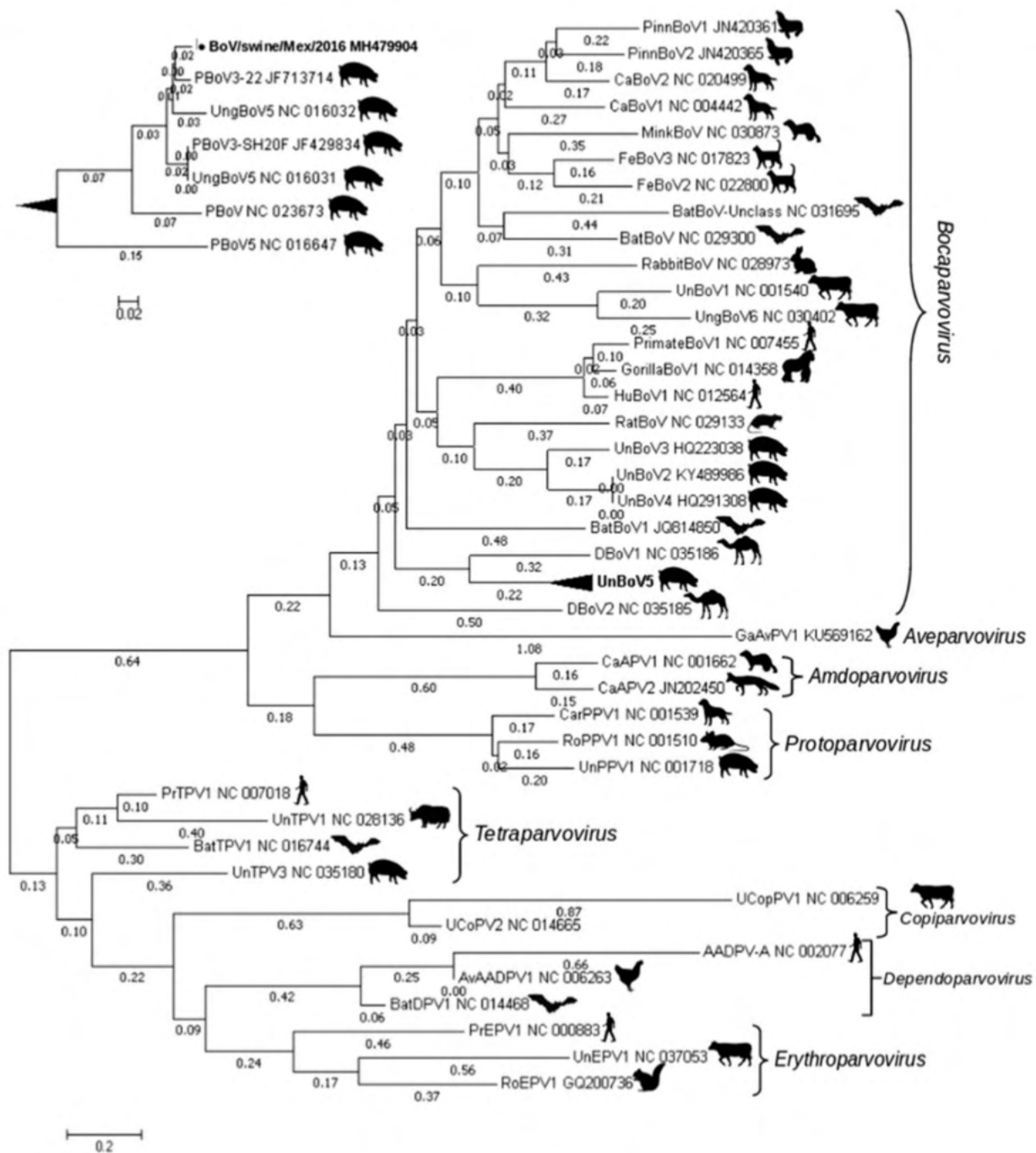


Fig. 7 Phylogenetic tree based on the NS1 gene of members of the subfamily *Parvovirinae* (genera *Amdoparvovirus*, *Aveparvovirus*, *Bocaparvovirus*, *Copiparvovirus*, *Dependoparvovirus*, *Erythroparvovirus*, *Protoparvovirus*, and *Tetraparvovirus*) showing the evolutionary relationship of the swine bocaparvovirus contig of 1198 bases (bold) found by next-generation sequencing in nasal and rectal samples of backyard swine. The phylogeny was inferred using

the neighbor-joining method, and the evolutionary distances were computed using the maximum composite likelihood method. The analysis included 46 reference nucleotide sequences of members of the subfamily *Parvovirinae*. Evolutionary analysis was conducted in MEGA7. The branch that includes ungulate bocaparvovirus 5 (including the contig sequenced in this study) are displayed in the upper left corner of the figure [24–26].

results reflect the health status of pigs under the environmental conditions of a rural backyard.

Acknowledgements We thank Omar Ríos Bello, Omar Maldonado Pineda, José López Reyes, Julio José Barrón Rodríguez, and Francisco A. Rosas Rodríguez for their help in field work. We also thank Carlos Cabello and José Eduardo Márquez García for their support during the sequencing work at the INER, and Jairo Betancourt for his help in translation of this text.

Authors' contribution Conceived by Rodrigo J. Barrón-Rodríguez and Elizabeth Loza-Rubio. Performed research: Rodrigo J. Barrón-Rodríguez, Edith Rojas-Anaya, José A.I. Romero-Espinosa, Joel A. Vázquez-Pérez and Moisés Cortés-Cruz. Analyzed the data: Rodrigo J. Barrón-Rodríguez, Jorge T. Ayala-Sumano, Gary García-Espinosa. Wrote the paper: Rodrigo J. Barrón-Rodríguez, Elizabeth Loza-Rubio, Moisés Cortés-Cruz and Gary García-Espinosa.

Funding This work was supported by the National Institute of Research in Forestry, Agriculture and Livestock (INIFAP) and the Mexican Council for Science and Technology (CONACyT) [scholarship 219623].

Availability of data and material Sequences are available in GenBank with accession numbers MH490914, MH490915, MH490916, MH490912, MH490913, MH490911 and MH479904.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval The care, handling and sampling of animals for this study were conducted in accordance with the guidelines issued by the Secretary of Agriculture and Rural Development (Spanish acronym, SADER). The experiment was approved by the Institutional Committee for Care and Use of Experimental Animals (Spanish acronym, CIC-UAE) of the Faculty of Veterinary Medicine and Animal Science at the Universidad Nacional Autónoma de México (UNAM) by protocol DC-2016/2-2.

Code availability Available upon request from the authors.

References

- Lambin EF, Turner BL, Geist HJ et al (2001) The causes of land-use and land-cover change: moving beyond the myths. *Glob Environ Chang* 11:261–269. [https://doi.org/10.1016/S0959-3780\(01\)00007-3](https://doi.org/10.1016/S0959-3780(01)00007-3)
- Patz JA, Daszak P, Tabor GM et al (2004) Unhealthy landscapes: policy recommendations on land use change and infectious disease emergence. *Environ Health Perspect* 112:1092–1098. <https://doi.org/10.1289/EHP.6877>
- Cohen ML (1998) Resurgent and emergent disease in a changing world. *Br Med Bull* 54:523–532. <https://doi.org/10.1093/oxfordjournals.bmb.a011707>
- Schultz-Cherry S, Olsen CW, Easterday BC (2011) History of Swine Influenza. In: Current topics in microbiology and immunology. *Curr Top Microbiol Immunol*, pp 21–27
- Smith I, Wang LF (2013) Bats and their virome: an important source of emerging viruses capable of infecting humans. *Curr Opin Virol* 3:84–91. <https://doi.org/10.1016/j.coviro.2012.11.006>
- Philbey AW, Kirkland PD, Ross AD et al (1998) An apparently new virus (family Paramyxoviridae) infectious for pigs, humans, and fruit bats. *Emerg Infect Dis* 4:269–271. <https://doi.org/10.3201/eid0402.980214>
- Bowden TR, Westenberg M, Wang LF et al (2001) Molecular characterization of menangle virus, a novel paramyxovirus which infects pigs, fruit bats, and humans. *Virology* 283:358–373. <https://doi.org/10.1006/viro.2001.0893>
- Wang LF, Hansson E, Yu M et al (2007) Full-length genome sequence and genetic relationship of two paramyxoviruses isolated from bat and pigs in the Americas. *Arch Virol* 152:1259–1271. <https://doi.org/10.1007/s00705-007-0959-4>
- Cuevas-Romero JS, Blomstrom AL, Berg M (2015) Molecular and epidemiological studies of Porcine rubulavirus infection—an overview. *Infect Ecol Epidemiol* 5:1–8. [https://doi.org/10.3402/iee.v5.29602r29602\[pii\]](https://doi.org/10.3402/iee.v5.29602r29602[pii])
- Yaiw KC, Cramer G, Wang L et al (2007) Serological evidence of possible human infection with tioman virus, a newly described paramyxovirus of bat origin. *J Infect Dis* 196:884–886. <https://doi.org/10.1086/520817>
- Yaiw KC, Bingham J, Cramer G et al (2008) Tioman virus, a paramyxovirus of bat origin, causes mild disease in pigs and has a predilection for lymphoid tissues. *J Virol* 82:565–568. <https://doi.org/10.1128/JVI.01660-07>
- Hause BM, Duff JW, Scheldt A, Anderson G (2016) Virus detection using metagenomic sequencing of swine nasal and rectal swabs. *J Swine Heal Prod* 24:304–308
- Sachsenröder J, Twardziok S, Hammerl JA et al (2012) Simultaneous identification of DNA and RNA viruses present in pig faeces using process-controlled deep sequencing. *PLoS One*. <https://doi.org/10.1371/journal.pone.0034631>
- Shan T, Li L, Simmonds P et al (2011) The fecal virome of pigs on a high-density farm. *J Virol* 85:11697–11708. <https://doi.org/10.1128/JVI.05217-11>
- Lager KM, Ng TF, Bayles DO et al (2012) Diversity of viruses detected by deep sequencing in pigs from a common background. *J Vet Diagn Investig* 24:1177–1179. <https://doi.org/10.1177/1040638712463212>
- Gallardo Nieto JL, Villamar Angula L, Barrera Wadgymar MA (2006) Situación actual y perspectiva de la producción de carne de porcino en México 2006
- Wiethoelter AK, Beltrán-Alcrudo D, Kock R, Mor SM (2015) Global trends in infectious diseases at the wildlife–livestock interface. *Proc Natl Acad Sci* 112:9662–9667. <https://doi.org/10.1073/pnas.1422741112>
- CEA (2014) Programa estatal hídrico de Morelos 2014 - 2018. 110
- INEGI IN de EG e I (2016) Mapa Digital de México. <http://gaia.inegi.org.mx/mdm6/?v=bGF00jE4LjYyNDAwLGxvbjotOTkuMDQ3Mjgsejo2LGw6YzQwN3xjNDE3fGMxMDJ8YzEwMHxjMzUw&layers=c401,c404,c407,c410,c417,c418>. Accessed 21 Jun 2018
- Djikeng A, Halpin R, Kuzmickas R et al (2008) Viral genome sequencing by random priming methods. *BMC Genomics* 9:5. <https://doi.org/10.1186/1471-2164-9-5>
- Afgan E, Baker D, Batut B et al (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:W537–W544. <https://doi.org/10.1093/nar/gky379>
- Zerbino DR (2011) Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinform*. <https://doi.org/10.1002/0471250953.bil105s31.Using>

23. Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>
24. Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>
25. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
26. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci* 101:11030–11035. <https://doi.org/10.1073/pnas.0404206101>
27. Hsieh TC, Ma KH, Chao A (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol* 7:1451–1456. <https://doi.org/10.1111/2041-210X.12613>
28. Chao A, Ma KH, Hsieh TC (2016) iNEXT (iNterpolation and EXTrapolation) Online. Program and User's Guide published at http://chao.stat.nthu.edu.tw/wordpress/software_download/30043:1-14. <https://doi.org/10.13140/RG.2.2.25777.79200>. Accessed 6 July 2018
29. SAGARPA (2011) ACUERDO mediante el cual se dan a conocer en los Estados Unidos Mexicanos las enfermedades y plagas exóticas y endémicas de notificación obligatoria de los animales terrestres y acuáticos. *D Of la Fed* 38:38–85
30. Sachsenröder J, Twardziok SO, Scheuch M, Johne R (2014) The general composition of the faecal virome of pigs depends on age, but not on feeding with a probiotic bacterium. *PLoS One* 9:15–22. <https://doi.org/10.1371/journal.pone.0088888>
31. Zhang B, Tang C, Yue H et al (2014) Viral metagenomics analysis demonstrates the diversity of viral flora in piglet diarrhoeic faeces in China. *J Gen Virol* 95:1603–1611. <https://doi.org/10.1099/vir.0.063743-0>
32. Wang D (2020) 5 challenges in understanding the role of the virome in health and disease. *PLoS Pathog* 16:e1008318. <https://doi.org/10.1371/journal.ppat.1008318>
33. Biggs PM (1985) Infectious animal disease and its control. *Philos Trans R Soc Lond B Biol Sci* 310:259–274
34. Tung J, Barreiro LB, Burns MB et al (2015) Social networks predict gut microbiome composition in wild baboons. *Elife* 2015:1–18. <https://doi.org/10.7554/eLife.05224>
35. Amimo JO, El Zowalaty ME, Githae D et al (2016) Metagenomic analysis demonstrates the diversity of the fecal virome in asymptomatic pigs in East Africa. *Arch Virol* 161:887–897. <https://doi.org/10.1007/s00705-016-2819-6>
36. Colbère-Garapin F, Martin-Latil S, Blondel B et al (2007) Prevention and treatment of enteric viral infections: possible benefits of probiotic bacteria. *Microbes Infect* 9:1623–1631. <https://doi.org/10.1016/j.micinf.2007.09.016>
37. Modi SR, Lee HH, Spina CS, Collins JJ (2013) Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499:219–222. <https://doi.org/10.1038/nature12212>
38. Ogilvie LA, Bowler LD, Caplin J et al (2013) Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. *Nat Commun* 4:2420. <https://doi.org/10.1038/ncomms3420>
39. Chiappetta CM, Cibulski SP, Lima FES et al (2017) Molecular detection of circovirus and adenovirus in feces of fur seals (*Arctocephalus* spp.). *EcoHealth* 14:69–77. <https://doi.org/10.1007/s10393-016-1195-8>
40. Oba M, Katayama Y, Naoi Y et al (2017) Discovery of fur seal feces-associated circular DNA virus in swine feces in Japan. *J Vet Med Sci* 79:1664–1666. <https://doi.org/10.1292/jvms.16-0642>
41. Breitbart M, Delwart E, Rosario K et al (2017) ICTV virus taxonomy profile: circoviridae. *J Gen Virol* 98:1997–1998. <https://doi.org/10.1099/jgv.0.000871>
42. Ge X, Li J, Peng C et al (2011) Genetic diversity of novel circular ssDNA viruses in bats in China. *J Gen Virol* 92:2646–2653. <https://doi.org/10.1099/vir.0.034108-0>
43. Wang H, Li S, Mahmood A et al (2018) Plasma virome of cattle from forest region revealed diverse small circular ssDNA viral genomes. *Virology* 515:5–9. <https://doi.org/10.1016/j.virus.2018.05.018>
44. Li L, Kapoor A, Slikas B et al (2010) Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J Virol* 84:1674–1682. <https://doi.org/10.1128/JVI.02109-09>
45. Chiba S, Sakuma Y, Kogasaka R et al (1979) An outbreak of gastroenteritis associated with calicivirus in an infant home. *J Med Virol* 4:249–254
46. Reuter G, Zimsek-Mijovski J, Poljsak-Prijatelj M et al (2010) Incidence, diversity, and molecular epidemiology of sapoviruses in swine across Europe. *J Clin Microbiol* 48:363–368. <https://doi.org/10.1128/JCM.01279-09>
47. Green KY, Ando T, Balayan MS et al (2000) Taxonomy of the caliciviruses. *J Infect Dis* 181(Suppl):S322–S330. <https://doi.org/10.1086/315591>
48. Scheuer KA, Oka T, Hoet AE et al (2013) Prevalence of porcine Noroviruses, molecular characterization of emerging porcine sapoviruses from finisher swine in the United States, and unified classification scheme for sapoviruses. *J Clin Microbiol* 51:2344–2353. <https://doi.org/10.1128/JCM.00865-13>
49. Wang QH, Han MG, Funk JA et al (2005) Genetic diversity and recombination of porcine sapoviruses. *J Clin Microbiol* 43:5963–5972. <https://doi.org/10.1128/JCM.43.12.5963-5972.2005>
50. Martella V, Lorusso E, Banyai K et al (2008) Identification of a porcine calicivirus related genetically to human sapoviruses. *J Clin Microbiol* 46:1907–1913. <https://doi.org/10.1128/JCM.00341-08>
51. Zhou F, Sun H, Wang Y (2014) Porcine bocavirus: Achievements in the past five years. *Viruses* 6:4946–4960. <https://doi.org/10.3390/v6124946>
52. Luo Z, Roi S, Dastor M et al (2011) Multiple novel and prevalent astroviruses in pigs. *Vet Microbiol* 149:316–323. <https://doi.org/10.1016/j.vetmic.2010.11.026>
53. Blomström A-L, Widén F, Hammer A-S et al (2010) Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics. *J Clin Microbiol* 48:4392–4396. <https://doi.org/10.1128/JCM.01040-10>
54. Quan P-L, Wagner TA, Briese T et al (2010) Astrovirus ENCEPHALITIS IN BOY WITH X-LINKED AGAMMAGLOBULINEMIA. *Emerg Infect Dis* 16:918–925. <https://doi.org/10.3201/eid1606.091536>
55. Hause BM, Padmanabhan A, Pedersen K, Gidlewski T (2016) Feral swine virome is dominated by single-stranded DNA viruses and contains a novel Orthopneumovirus which circulates both in feral and domestic swine. *J Gen Virol* 97:2090–2095. <https://doi.org/10.1099/jgv.0.000554>

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

Affiliations

Rodrigo Jesús Barrón-Rodríguez^{1,2} · Edith Rojas-Anaya¹ · Jorge Tonatiuh Ayala-Summano³ · José Ángel Iván Romero-Espinosa⁴ · Joel Armando Vázquez-Pérez⁴ · Moisés Cortés-Cruz⁵ · Gary García-Espinosa² · Elizabeth Loza-Rubio¹

Rodrigo Jesús Barrón-Rodríguez
rodjes1@hotmail.com

Edith Rojas-Anaya
edith_ra23@hotmail.com

Jorge Tonatiuh Ayala-Summano
jsummano@gmail.com

José Ángel Iván Romero-Espinosa
joseangelivan@gmail.com

Joel Armando Vázquez-Pérez
joevazpe@gmail.com

Gary García-Espinosa
gary@unam.mx

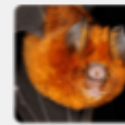
¹ Laboratorio de Biotecnología en Salud Animal, Centro Nacional de Investigación Disciplinaria en Microbiología Animal (CENID-Microbiología), Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), carretera federal México-Toluca km 15.5, colonia palo Alto, Cuajimalpa, P.C. 05110, Mexico City, Mexico

² Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México (UNAM), Avenida Universidad 3000, colonia Ciudad universitaria, Coyoacán, P.C. 04510, Mexico City, Mexico

³ Idix S.A. de C.V., Sonterra 3035 interior 26, Fraccionamiento Sonterra, P.C. 76230, Santiago de Querétaro, Querétaro, Mexico

⁴ Laboratorio de Virología, Instituto Nacional de Enfermedades Respiratorias (INER), Calzada de Tlalpan 4502, Del. Tlalpan, colonia Sección XVI, Tlalpan, P.C. 14080, Mexico City, Mexico

⁵ Centro Nacional de Recursos Genéticos (CNRG), Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), Boulevard de la biodiversidad 400, Rancho las Cruces, P.C. 47600, Tepatitlán de Morelos, Jalisco, Mexico

**Authors:**

Rodrigo Barrón Rodríguez, Rocío Parra-Laca, Edith Rojas-Anaya, José Romero-Espinoza, Jorge Ayala-Sumuano, Joel Vázquez-Pérez, Gary García-Espinosa, Elizabeth Loza-Rubio

Decision letter:

August 09, 2021

AC-00481-2020-03

Evidence of Viral Communities in three Species of Bats from Rural Environment in Mexico

Dear Dr Elizabeth Loza Rubio,

I am pleased to inform you that we shall be happy to publish a suitably revised version of your manuscript, entitled: Evidence of Viral Communities in three Species of Bats from Rural Environment in Mexico, provided you can satisfactorily address the specific points raised in the Reviewer's comments (see below).

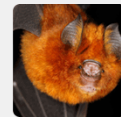
Please revise your paper according to the attached Reviewer's comments. Your manuscript will not be taken into consideration without the revisions made according to the recommendations and a point by point response to each comment.

Authors of our journal are requested to prepare a revised version of their manuscript within 30 days (please, contact us shall you require more time). This may ensure fast publication if an article is finally accepted.

Thank you for submitting your work to us.

Yours sincerely,
Wiesław Bogdanowicz
Editor-in-Chief
Acta Chiropterologica

El artículo arriba mencionado fue revisado, corregido con base en las observaciones del revisor y reenviado (23-agosto-2021) a la revista Acta Chiropterologica para su publicación. A continuación se incluye la última versión.



Evidence of Viral Communities in three Species of Bats from Rural Environment in Mexico

Type:

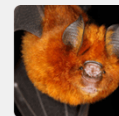
Original paper

Abstract:

Some emerging and reemerging diseases have been associated with certain species of bats. These diseases have emerged in anthropogenic environments where the conditions for spillover of infectious agents between bats, domestic animals, and humans are present. Mexico is the country with the fourth highest bat diversity in the world, and some of these bat species live in anthropogenic environments such as a backyard production system. The objective of this study was to analyze the virome of three species of bats (*Artibeus* spp., *Macrotus waterhousii* and *Pteronotus parnellii*) that inhabit roosts near rural backyard farms and have large geographic distributions. Rectal swabs were taken and analyzed by next-generation sequencing (NGS). The detected viral communities were abundant in bacteriophages with low numbers of animal virus sequences. Thus, it was possible to study the virome of these three species of bats which have not been previously reported. Viral diversities varied between each studied species and differed from previous studies. In one of these species, *P. parnellii*, sequences of Coronaviridae family were found. These results indicate that viral diversity of these three bat species includes mostly bacteriophages while showing low numbers for known animal viruses. The findings of this study contribute to our knowledge of the virome of these bat species which inhabit anthropogenic habitats and have large geographical distributions.

Keywords:

next-generation sequencing, bacteriophages, rural backyard, *Artibeus* spp, *Macrotus waterhousii*, *Pteronotus parnellii*



Evidence of Viral Communities in three Species of Bats from Rural Environment in Mexico

Virome of bats from rural environment in Mexico

Rodrigo J. Barrón-Rodríguez,^{1,2} Rocío Parra-Laca,^{1,2} Edith Rojas-Anaya,² José Á.I. Romero-Espinoza,³ Jorge T. Ayala-Sumuano,⁴ Joel A. Vázquez-Pérez,³ Gary García-Espinosa², Elizabeth Loza-Rubio^{1,*}

¹Centro Nacional de Investigación Disciplinaria en Microbiología Animal (CENID-Microbiología), Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), carretera federal México-Toluca km 15.5, colonia palo Alto, Cuajimalpa, Ciudad de México, México; C.P. 05110.

²Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México (UNAM), Avenida Universidad 3000, colonia Ciudad universitaria, Coyoacán, Ciudad de México, México. CP 04510

³Laboratorio de Virología, Instituto Nacional de Enfermedades Respiratorias (INER), Calzada de Tlalpan 4502, Del. Tlalpan, colonia Sección XVI, Tlalpan, Ciudad de México, México. C.P. 14080.

⁴Idix S.A. de C.V., Sonterra 3035 interior 26, Fraccionamiento Sonterra, Santiago de Querétaro, Querétaro México. C.P. 76230.

Correspondence Elizabeth Loza-Rubio, Centro Nacional de Investigación Disciplinaria en Microbiología Animal (CENID-Microbiología), Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), carretera federal México-Toluca km 15.5, colonia palo Alto, Cuajimalpa, Ciudad de México, México; C.P. 05110.

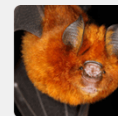
Email: eli_rubio33@hotmail.com; telephone: (52) 55-3871-8700 ext. 80324.

KEYWORDS

Next generation sequencing, bacteriophages, rural-backyard, *Artibeus*, *Macrotus waterhousii*, *Pteronotus parnellii*

Short Running title:

Virome of bats from rural environment in Mexico

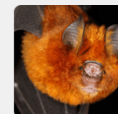


Abstract

Some emerging and reemerging diseases have been associated with certain species of bats. These diseases have emerged in anthropogenic environments where the conditions for spillover of infectious agents between bats, domestic animals, and humans are present. Mexico is the country with the fourth highest bat diversity in the world, and some of these bat species live in anthropogenic environments such as a backyard production system. The objective of this study was to analyze the virome of three species of bats (*Artibeus* spp., *Macrotus waterhousii* and *Pteronotus parnellii*) that inhabit roosts near rural backyard farms and have large geographic distributions. Rectal swabs were taken and analyzed by next-generation sequencing (NGS). The detected viral communities were abundant in bacteriophages with low numbers of animal virus sequences. Thus, it was possible to study the virome of these three species of bats which have not been previously reported. Viral diversities varied between each studied species and differed from previous studies. In one of these species, *P. parnellii*, sequences of *Coronaviridae* family were found. These results indicate that viral diversity of these three bat species includes mostly bacteriophages while showing low numbers for known animal viruses. The findings of this study contribute to our knowledge of the virome of these bat species which inhabit anthropogenic habitats and have large geographical distributions.

KEYWORDS

Next generation sequencing, bacteriophages, rural-backyard, Artibeus, Macrotus waterhousii, Pteronotus parnellii



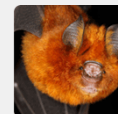
INTRODUCTION

Bats (*Chiroptera*) have gained attention as carriers or potential reservoirs of diverse, high impact pathogenic animal and human viruses (Racey *et al.*, 2015). Due to the viral richness found in some of these species, bats have created interest of the scientific community, since they may be one of the most crucial sources of mammal viruses (Luis *et al.*, 2013; Smith and Wang, 2013). Among these, lyssaviruses, coronaviruses, ebolaviruses, and paramyxoviruses (Wang *et al.*, 2006; Kuzmin *et al.*, 2011; Wang and Cowled, 2015) have been identified in bats not showing any sign of disease. These facts illustrate the importance of bats research. Emerging outbreaks take place in environments where there is change in use of land by human activities such as agriculture and animal production (Patz *et al.*, 2000, 2004). These environments represent interfaces for infectious agent adaptations and disease occurrence risk (Lambin *et al.*, 2001; Patz *et al.*, 2004; Tian *et al.*, 2015).

There is evidence of viruses of potential concern to humans, or similar to (such as SARS-CoV), infecting *Chiroptera* species naturally. Once these viruses come in close contact with domestic species they can infect, adapt and cause disease in new hosts. Examples of these are the Nipah virus (Luby *et al.*, 2009), Menangle virus, Hendra virus, Marburg and Ebola viruses and SARS-CoV (called by Smith and Wang, 2013). Although these viruses do not have the same spillover dynamics, or even unknown for some of these viruses, at the end they cause disease in domestic animals and humans. Thus, it makes the study of bat viral communities in these environments very important. One of the first steps in this process, is the identification of species present in the microbial community of the animals found at the human-wildlife-domestic animal interface (Patz *et al.*, 2000; Lambin *et al.*, 2001).

One environment where different species are present is the backyard farms. The production of diverse domestic animals in these farms, which are mainly seen in rural and peri-urban areas of developing regions, represents an essential source of food and income. In this system, domestic animals frequently coexist with humans, other domestic species and even have occasional contact with wild animals.

These interactions represent interfaces in which the risk for diseases is not often taken into account until outbreaks occur (Wiethoelter *et al.*, 2015). Backyard farms are common in Mexico, the country with the fourth highest bat diversity, harboring 138 species (Medellin *et al.*, 2007). As little information on the viral community of American bat species is available, it is relevant to undertake research on this subject, especially for those species cohabiting in environments with human-domestic animals-bats interfaces. This work aims to study the virome of three bat species that inhabiting roosts near rural backyard farms to obtain information on their viral communities.



76 MATERIALS AND METHODS

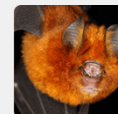
77 *Sampling*

78 Four bat roosts were identified in the state of Morelos, Mexico, near rural environments. These roosts
79 harbor a high diversity of bats as reported by personal of State Livestock Promotion and Protection
80 Committee (CFPPEM, the acronym in Spanish) (Fig. 1). These roosts were chosen because of their
81 location (<10 km, from rural settlements), close enough for bats to forage or roost within the
82 anthropized geographical areas (Table 1). These roosts were located near rural populations (range 1.6
83 km – 8.2 km) that housed species of domestic animals, such as cattle, sheep, goats, poultry, cats, and
84 dogs. Samples were obtained between February and July 2016. Bats were caught with mist nets and
85 placed in individual cages to be sampled and released. Two phyllostomid species (*Artibeus* spp. n=17;
86 and *Macrotus waterhousii* n=9) and one mormoopid species (*Pteronotus parnellii*, n=13) were
87 captured. Among specimens of *Artibeus*, the 11 animals captured in Tlaltizapan were identified as
88 *Artibeus phaeotis* based on the morphological criteria established by Medellín et al. (2007). The other 6
89 specimens were not identified to species level because they were released before morphological
90 identification (error in sampling process). Individual rectal swabs were collected using 1 ml of Eagle's
91 minimum essential medium (EMEM). Swabs were sent to the Laboratory of Biotechnology on Animal
92 Health (CENID Salud Animal e Inocuidad - INIFAP) for the immediate RNA (ribonucleic acid) and
93 DNA (desoxyribonucleic acid) extraction, at 2-6°C. The time of transportation was ≤ 4 hours.

94 The collection, care, handling and sampling of animals for this study were conducted in accordance
95 with the guidelines issued by the Secretary of Environment and Natural Resources (SEMARNAT for its
96 Spanish acronym) (authorization SGPA/DGVS/08483/15). This study was approved by the Institutional
97 Committee for Care and Use of Experimental Animals (CICUAE for its Spanish acronym) of the
98 Faculty of Veterinary Medicine and Animal Science at the Universidad Nacional Autónoma de México
99 (UNAM) by protocol DC-2016-2-2 and was conducted in strict accordance with its recommendations.

100 *RNA/DNA extraction*

101 Before the extraction step, samples were homogenized by vortexing to release the biological agents
102 present. The swabs were discarded and vials were centrifuged (6,000 x g/15 min) to separate cells,
103 bacteria, and debris. A pool of supernatants by species was made by transferring a 200 μ l from each
104 sample. The pool of rectal samples of each species was filtered through a 0.45 μ m filter (Millipore) to
105 remove eukaryotic and bacterial cell-sized particles, followed by fraction ultracentrifugation at 170,000
106 x g/3 hours at 4 °C (Beckman Coulter Optima Ty-70) to pellet viral particles. Pellets were resuspended
107 in EMEM solution and filtered with a 0.22 μ m filter. These filtrates were treated with DNase (Thermo



108 Scientific) and RNase A (Thermo Scientific) at 37°C for 30 minutes. Total viral nucleic acids (both
109 RNA and DNA) from each sample were extracted and purified by Purelink Viral RNA/DNA mini kit
110 (Invitrogen) according to the manufacturer's protocol. The concentration of RNA/DNA was estimated
111 by spectrophotometer (Nanodrop - Thermo Fisher Scientific). The purified RNA/DNA of each species
112 sample was divided into two fractions, one as RNA mix and the other as DNA mix for the enrichment
113 of the purified nucleic acids.

114 *DNA library construction and Next-generation sequencing*

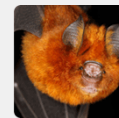
115 Complementary DNA (cDNA) was synthesized from RNA by random primer-mediated sequence-
116 independent single primer amplification (RP-SISPA) using a chimeric M13 primer with a random
117 hexamer attached to its 3'-end (Djikeng and Spiro, 2009). The reverse transcription was performed with
118 a Transcriptor First Strand cDNA Synthesis kit V.6 (ROCHE). Then, cDNA and DNA were denatured,
119 followed by annealing of random primers and the use of the Klenow enzyme (New England Biolabs),
120 generating dscDNA and dsDNA. These products were subsequently used as a template for polymerase
121 chain reaction (PCR) amplification using M13 primers. The amplified genetic mixes of cDNA and
122 DNA were equimolar pooled and their concentration was estimated by spectrophotometer (NanoDrop
123 2000 – Thermo Fisher Scientific). The quality of samples was determined by capillary electrophoresis
124 (2100 Bioanalyzer – Agilent Genomics). The DNA library was prepared with a Nextera XT DNA
125 Library Preparation Kit and identified by a Nextera XT index kit (Illumina) following the
126 manufacturer's instructions. The pooled library was sequenced with a Miseq Illumina platform using
127 MiSeq Reagent Kit (v2). The sequencing protocol was performed according to the manufacturer's
128 instructions in a Miseq system (Illumina) using 2x250bp cycles at the National Institute of Respiratory
129 Diseases (by its Spanish acronym INER).

130 *Bioinformatic analysis*

131 FastQC program (Braham Institute, Cambridge) was used for data quality analysis for each species
132 dataset. After quality analysis, reads were edited and filtered with the Trimmomatic v.0.33 command-
133 line tool (Bolger *et al.*, 2014) to remove M13 primers and any other remaining Illumina-specific
134 sequences from the reads to select the sequences with sufficient quality (Phred score ≥ 20) and length
135 for the taxonomic assignation.

136 *Taxonomic assignment and virus identification*

137 Filtered reads, for each species dataset, were compared against whole-genome sequences (wgs) and
138 nucleotide (nt) databases from the National Center for Biotechnology Information (NCBI) using the
139 tool BLASTn (E value $< 10^{-4}$). All sequences without nucleotide homology were compared against a



140 protein sequence bank (UniProt) using the tool BLASTx (E value 10^{-5}). Sequence alignments were
141 filtered by "Filter" tool Galaxy v1.1.0 (similarity >50%). Filtered alignments were assigned to
142 taxonomic representation by the "Fetch taxonomic representation" tool (Galaxy Version 1.1.0). The
143 lowest taxonomic classification was made using the "Search for lowest diagnostic classification" tool
144 (Galaxy Version 1.0.1). The "Taxonomy Summary" tool (Galaxy Version 1.0.0) was used to classify the
145 sequences. Previous analyzes were performed using the Galaxy platform (Afgan *et al.*, 2018).

146 *Assembly and phylogenetic analysis*

147 All reads were de novo assembled using Velvet software (Zerbino and Birney, 2008; Zerbino, 2010) to
148 obtain the contigs. Contigs were compared against whole genome sequences (wgs) and nucleotide (nt)
149 databases using BLASTn and BLASTx, following the same steps described in the section on
150 taxonomic assignment. This study used the longer contigs and sequences of the more abundant animal
151 viruses found to perform multiple sequence alignments (using Clustal Omega). The alignment results
152 were used to infer the phylogenies in MEGA X software (Kumar *et al.*, 2018), employing the
153 Neighbor-Joining method (Saitou and Nei, 1987) (performed with a 1000 replicates bootstrap). In this
154 study also used the maximum composite likelihood method (Tamura *et al.*, 2004) to compute the
155 evolutionary distances. These analyses showed the contigs nucleotide relationship with more similar
156 viral species and their taxonomic group members, according to ICTV demarcation criteria.

157 *Sequence accession numbers*

158 Sequences and contigs of more abundant animal viruses (coronavirus and circovirus) were submitted to
159 the Gen Bank (NCBI) to obtain their accession numbers.

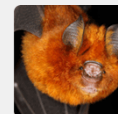
160 *Diversity analysis*

161 Rarefaction with their interpolation and extrapolation for each data set were estimated using iNEXT, an
162 R package modified by Chao *et al.* (2016) (Hsieh *et al.*, 2016), to show that diversity results obtained
163 by massive sequencing were representative. The Shannon-Wiener index (equity index) was used to
164 assess the diversity of viruses in the reads per bat-species obtained and to compare them.

165 RESULTS

166 *Overview of sequencing data*

167 There were a total of 585 718 raw reads of *Artibeus* spp., 914 710 raw reads of *M. waterhousii* and 386
168 490 of *P. parnellii* (with an average length >250 bases). The reads that did not show similarity with
169 known reported sequences were 556 432 reads (95.1%) from *Artibeus* spp., 873 548 reads (95.5%)
170 from *M. waterhousii* and 363 687 (94.1%) from *P. parnellii*. The analysis performed with BLASTn and



BLASTx with both nt and wgs databases showed homology with prokaryotic and eukaryotic organisms in 28,593 sequences (4.9 %) of *Artibeus* spp., 39 743 (4.4 %) of *M. waterhousii* and 22 290 (5.9 %) of *P. parnellii*. Within taxonomic assigned sequences, there were 518 sequences (0.1 %) of *Artibeus* spp., 933 (0.1 %) of *M. waterhousii*, and 371 (0.1 %) of *P. parnellii* related to viruses. Among viral sequences, in *Artibeus* spp. 480 reads were associated with phages and nine with animal viruses, in *M. waterhousii*, 858 were related to phages and 24 to animal viruses, and in *P. parnellii* 307 were associated with phages, 45 with animal viruses and 1 with protozoan virus.

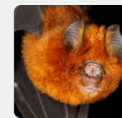
Identified viruses and their diversity (based on the reads)

Animal viruses families found in sequences of *Artibeus* spp. were *Circoviridae*, *Retroviridae*, and *Herpesviridae*. For *M. waterhousii* the families obtained were *Circoviridae*, *Retroviridae*, *Reoviridae*, *Parvoviridae* and *Herpesviridae*. For *P. parnellii* the families found were *Coronaviridae*, *Retroviridae*, *Circoviridae*, and *Hepeviridae* (Table 2 and Fig 2).

Regarding bacteriophages reads, the most abundant genus in *Artibeus* spp. were unclassified *Shigella phage Sf1* (species), *Jerseyvirus* and *K1gvirus* of *Siphoviridae* family (47.7% of the phages sequences), *Msw3virus* and *T4virus* of *Myoviridae* family (24.8% of the phages sequences), and *Kp34virus*, *T7virus* and *Phikmvirus* of *Podoviridae* family (16 % of the reads assigned to phages). In the case of *M. waterhousii* the most abundant genus of phages represented were unclassified *Tunavirus*, *K1gvirus* and *Jerseyvirus* of the *Siphoviridae* family (58.1 % of the reads of phages), *T4virus*, *V5virus* and *Ms3virus* of the *Myoviridae* family (25.2 % of the phages sequences) and *T7virus* and *Phi29virus* of *Podoviridae* family (8 % of the phages reads). Finally, for *P. parnellii* the most frequently identified genus of viruses were *Tunavirus*, *K1gvirus* and *Jerseyvirus* of *Siphoviridae* family (59.3% of the phages reads), *T4virus* and *V5virus* of the *Myoviridae* family (20.5 % of the phages reads) and *T7virus* and *Phi29virus* of the *Podoviridae* family (12.4 % of reads assigned to phages).

Assembly and phylogenetic analysis

In the BLAST, less viruses were identified based on contigs compared to reads. In *Artibeus* spp. identified bacteriophages belongs to families *Siphoviridae* and *Myoviridae*. However animal viruses were not identified in BLAST/contigs. Therefore the reads assigned to animal viruses from each bat species were aligned to map them with respective reference viral genome/sequence using Clustal Omega multiple alignment tool. Thus, it found overlapped reads of coronavirus to obtain two contigs (482 and 201 bp lengths and contig coverage of 3.6 and 2 respectively), from *P. parnellii*.



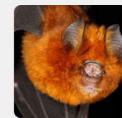
201 This study used the longer contig of coronavirus obtained from *P. parnellii* data, and the three longest
202 sequences related to circoviruses (one of each bat species data) for phylogeny analyses (Table 3). These
203 sequences are representative of the more abundant animal viruses and were submitted to the GenBank
204 databases. This study obtained similar known and reference sequences of the genus *Betacoronavirus*
205 members (similarities: nt \leq 98.4 – 76.16 %; aa \leq 100 – 80 %) and genus *Circovirus* (similarities: nt \leq 90
206 %; aa \leq 94 %). They were used to perform multiple sequence alignment and phylogeny with its
207 respective contig/sequence according to ICTV demarcation criteria. The betacoronavirus contig found
208 in this study was grouped with other unclassified bat coronavirus, similar to members of the subgenus
209 *Embecovirus* (Fig. 3). In the phylogenies of circovirus, the sequences for each bat species (Figures 4, 5
210 and 6 respectively) were grouped with sequences of porcine circoviruses and porcine circo-like viruses,
211 in a clade apart from the other circoviruses, included porcine circovirus 1, 2 and 3. The circovirus
212 sequences found in this study do not overlap (with exception of a 83 pb short region in which the
213 circoviruses sequences of the three bat species show similarity 87.8% to 95.7%). It is for this reason
214 they were analyzed in separate phylogenies. A supplementary phylogeny based on 83 pb length
215 sequences of each three bat species was added to show the relation between them and with
216 circoviruses (Supplementary figure 1).

217 *Statistical analysis and viral diversity*

218 Despite the fact that rarefaction curves did not reach asymptotic level (specially for *P. parnellii*),
219 results obtained by next-generation sequencing showed quite a representative viral community for each
220 sampled bat species (Fig. 7). The Shannon-Wiener index was used to assess the diversity of total virus
221 species in the dataset of reads of each bat species, obtaining values of 3.14 for *Artibeus* spp., 2.98 for
222 *M. waterhousii*, and 3.19 for *P. parnellii*. The values of animal viruses were 1.36 for *Artibeus* spp., 1.17
223 for *M. waterhousii* and 1.7 for *P. parnellii*.

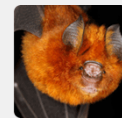
224 DISCUSSION

225 This study analyzed the viromes of bat specimens from *Artibeus* spp., *Macrotus waterhousii* and
226 *Pteronotus parnellii* species present in rural environments. The results showed that the viral diversity is
227 largely composed of abundant bacteriophages and to a lesser extent animal viruses. Among the viral
228 sequences identified in this study, the predominance of bacteriophage-associated sequences was evident
229 (87% - 98%), which are higher proportions than those reported in previous studies (0.1 - >50%)
230 (Donaldson *et al.*, 2010; Li *et al.*, 2010b; Salmier *et al.*, 2017), even when compared to other species



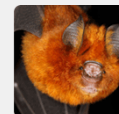
231 such as pigs, shrew, and the Tasmanian devil (Sachsenröder *et al.*, 2014; Chen *et al.*, 2019; Chong *et*
232 *al.*, 2019). Reports of bat viromes are available for some species in America, Asia, Africa, and Europe.
233 However, these studies were carried out in urban environments or on synanthropic species (Ge *et al.*,
234 2012; Baker *et al.*, 2013; Dacheux *et al.*, 2014; Gay *et al.*, 2014; Salmier *et al.*, 2017; Geldenhuys *et*
235 *al.*, 2018) and showed a different composition. Most of the viral diversity in these studies was the result
236 of animal virus species and sequences associated with plant, fungal, and protozoan viruses. Most
237 previous studies worked on fecal samples (Li *et al.*, 2010b; Ge *et al.*, 2012; Salmier *et al.*, 2017;
238 Geldenhuys *et al.*, 2018), but there are reports on urine (Baker *et al.*, 2013), oral and pharyngeal swabs
239 (Donaldson *et al.*, 2010; Wu *et al.*, 2012) and organs and carcasses (He *et al.*, 2013; Dacheux *et al.*,
240 2014) that showed variations in the viral communities. Although there are other studies that worked
241 with similar samples, these studies showed viral species of *Caudovirales*, *Inoviridae*, *Microviridae* (Wu
242 *et al.*, 2012) and *Podoviridae* (Donaldson *et al.*, 2010). This suggests bat virome varies among species
243 and their life histories (foraging guild, environment, habits, etc), indicating a complex viral ecology in
244 the bat populations (He *et al.*, 2013). But there are other factors, such as the maintenance of the
245 samples, methods of nucleic acids extraction, amplification and enrichment of nucleic acids and the
246 bioinformatics strategies used. All of these factors could affect the results of the studies on virome and
247 their diversities.

248 The sequences associated with bacteriophages represented the viral families of *Siphoviridae*,
249 *Myoviridae*, and *Podoviridae* in the three bat species. The viruses with the highest abundance
250 corresponded to Shigella phage pSf1-1 [genus *Tunalikeyvirus*], Escherichia phage K1 [genus *K1Gvirus*],
251 Salmonella phage [genus *Jerseyvirus*] (family *Siphoviridae*), genus *T4likevirus*, Klebsiella phage
252 JD001 [genus *Msw3virus*], and the genera *T7virus* and *Phikmvlikevirus* (family *Podoviridae*). The phage
253 communities reported in other bats species (Donaldson *et al.*, 2010; He *et al.*, 2013), have been
254 represented by members of *Podoviridae* and *Myoviridae* families. It is essential to pay attention to the
255 phage community in bats, and in general, in all species, as they can provide evidence of being
256 pathogenic bacteria carriers. In addition, there is the need to delve into the function of these viruses,
257 since some have shown to be an adequate substitution to antibiotics (Donaldson *et al.*, 2010; Mokili *et*
258 *al.*, 2012; Pires *et al.*, 2015; Herridge *et al.*, 2020). Indeed, these phages could be a source of bacteria
259 population control, including potential pathogens. This could make these bats more resistant to enteric
260 bacterial diseases (Herridge *et al.*, 2020; Sausset *et al.*, 2020). This suggests that further studies are
261 important to understand the role of the phages on this and other related viral ecology issues.



262 Overall, the virome of these three species of bats contained a low number of sequences assigned to
263 known animal viruses (2% - 8.9%), compared to those reported in other studies (> 28%) (Baker *et al.*,
264 2013; He *et al.*, 2013). The viral families more widely identified in previous bat viral metagenomics
265 studies are Herpesviridae, Adenoviridae, Circoviridae, Papillomaviridae, Retroviridae, Parvoviridae,
266 Coronaviridae, Astroviridae, Flaviviridae, Picornaviridae and Poxviridae (58% – 0.07% of viral reads)
267 (Wu *et al.*, 2012, 2016; He *et al.*, 2013; Salmier *et al.*, 2017; Geldenhuys *et al.*, 2018). Some of the
268 viral families mentioned were identified from the sequences or contigs of the bat species studied. There
269 are more than 10,770 virus sequences identified in bats reported in databases (Chen *et al.*, 2014), from
270 which 29 animal-viral families are represented. From these sequences, at least 25 % came from bat
271 species from America (Chen *et al.*, 2014). This study is important because it contributes to increase the
272 available information in databases.

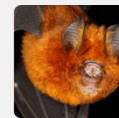
273 The diversity of animal viruses was also low compared to previous studies reported ≥ 13 viral families
274 (Wu *et al.*, 2016; Salmier *et al.*, 2017; Geldenhuys *et al.*, 2018) in contrast with 3 to 5 families found in
275 this study, with few related sequences (1-35). Our results showed a lower diversity of known or similar
276 viral sequences in the bats studied. It is important to mention that many viromes reported belong to
277 bats associated to high degree of anthropogenic environments such as cities (Ge *et al.*, 2012; Baker *et al.*,
278 2013; Afgan *et al.*, 2018; Geldenhuys *et al.*, 2018) while others do not specify the type of
279 environment (Wu *et al.*, 2012). The Shannon-Wiener diversity index supports the fact that phages
280 represented the most abundant viral community, with *P. parnellii* and *Artibeus* spp. showing the highest
281 (3.22) and lowest (2.98) total viral diversity. In relation to animal-viruses, *P. parnellii* show the highest
282 value of viral diversity (1.7) in this study and *M. waterhousii* the lowest (1.17). Taking into
283 consideration that the diversity index values lower than two (Moreno, 2001) are considered low. Most
284 of the virome in bats studies did not report viral diversity indexes so it is important to have a baseline
285 information to compare viral diversities, considering other characteristics such as environments and
286 sample types. Rico *et al.* (2015) describe in depth the viral and bat diversities in three sample sites with
287 different anthropogenic perturbation levels. But, since they used another method of viral identification
288 (specific molecular identification), it is impossible to compare values. Despite having obtained
289 diversity results close to asymptote in rarefaction curves, not all arrived a good level, specially *P.*
290 *parnellii*. This fact could be related with the large number of sequences without taxonomic assignation
291 found in the bat species of this study, unlike other reports (He *et al.*, 2013; Dacheux *et al.*, 2014). These
292 calls for a continuing analysis of these sequences, even short length reads and contigs, since it is likely



293 that a number of reads correspond to unknown viruses or limited sequences present (Mokili *et al.*,
294 2012).

295 Among sequences of *P. parnelli* associated with animal viruses, most (76.1%) were related to
296 coronaviruses. Among these coronavirus related sequences 48.6 % show similarity with unclassified
297 sequences of bat betacoronaviruses; 17.1 % with a Mexican betacoronavirus partial sequence
298 (KC117198), with similarities from $\leq 97.9\%$; 8.6 % with Roussettus bat coronavirus HKU9
299 (MG762674), $\leq 77.8\%$; 8.6 % with Severe acute respiratory related coronavirus 1 (NC_004718), \leq
300 77.3%. Additionally, the rest of the data shows sequences related to members of betacoronavirus
301 (11.4%), alphacoronavirus (2.8 %) and gammacoronavirus (2.8 %). Bats are classified as alpha- and
302 betacoronavirus reservoirs (Han *et al.*, 2015; Banerjee *et al.*, 2019), but in this study only *P. parnellii*
303 shows sequences similar to members of these viruses. Birds are considered as reservoirs of
304 gammacoronaviruses (Wille and Holmes, 2020) but in this study one sequence (58 bp) similar (nt 81%)
305 to an avian virus (HM245924 Infectious bronchitis virus CQ04) was found. The sequence was
306 individually analyzed showing also higher similarity (nt 98%) with partial sequences of bat
307 betacoronaviruses (JQ731781). Therefore a multiple sequence alignment was performed revealing that
308 in this case this short length sequence shows homology with a region similar to different genera of
309 coronaviruses. This fact calls for more attention to the bioinformatic analysis results to confirm them.
310 The sequences of this study that showed similarity ($\leq 78.6\%$) with sarbecoviruses and merbecoviruses
311 were short length (≤ 70 pb) sequences, and therefore were not further analyzed. In any case, in these
312 environments further investigation on coronaviruses surveillance could be useful to determine
313 coronavirus species and related host bat species.

314 We found sequences with similarity to *Rep* gene of porcine circo and a circo-like viruses which was
315 previously identified in fecal samples from piglets on a high-density farm in the USA (Shan *et al.*,
316 2011). The sequences found in this study showed evidence to belong to new variants of circoviruses,
317 although deeper characterization is need to determine their relationship among them and with
318 circoviruses currently reported. Previous studies have demonstrated the cross-species transmission of
319 circoviruses and cycloviruses (Li *et al.*, 2010a, 2011). Circular ssDNA viruses are commonly detected
320 in animals, other organisms, and environmental samples. However, further investigations of these
321 viruses is needed in domestic and wild animals overlapping environments to determine the interspecies
322 transmission and risks. Since bat activities of this study overlapped within the swine livestock farming
323 rural areas, it is important to establish circovirus and other viruses surveillance, more when there are
324 reports of viruses adapted from bats to swine such as Nipha (Luby *et al.*, 2009), Menangle (Philbey *et*

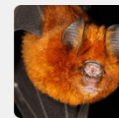


325 *al.*, 1998, 2008), Ebola (Leroy *et al.*, 2005; Organizacion Mundial de la Salud, 2021) and Tioman
326 (Yaiw *et al.*, 2008). Additionally viruses from swine have been identified in some species of bats (i.e.
327 Mapuera, rubulavirus) (Wang *et al.*, 2007; Cuevas-Romero, Blomstrom, Berg, 2015). It is important to
328 take into account that all of these diseases or infections have been associated with anthropogenic
329 activities. In any case, it is important to survey the viral and microbial communities of any bat and
330 domestic animal populations in environments where these species coexist and overlap. Surveillance
331 program must be established using the best accurate assays to identify and confirm the most viruses
332 present.

333 Metagenomics does not always seem to be the best option to identify all viral species, such as the case
334 of paramyxoviruses (Drexler *et al.*, 2012). In this study there were no reads associated to
335 paramyxoviruses, coinciding with previous studies. Here the RT-PCR would be used to identify
336 paramyxoviruses from RNA samples, but these were consumed in the nucleic acids, emphasizing the
337 importance of obtaining sufficient samples, even more with little animals such as bats.

338 In contributing with valuable virome information of bat species, here we describe the viromes of two
339 additional species belonging to the *Phyllostomidae* family (*Artibeus* spp. and *Macrotus watherhousii*)
340 and one of the *Mormopidae* family (*Pteronotus parnellii*). Previous studies on the virome of American
341 bats included species of the families *Vespertilionidae* (*Eptesicus fuscus*, *Perimyotis subflavus*, *Myotis*
342 *lucifugus*), *Molossidae* (*Molossus molossus*) and *Phyllostomidae* (*Desmodus rotundus*) (Donaldson *et*
343 *al.*, 2010; Salmier *et al.*, 2017). These studies have been done on some bat species with different life
344 histories (diet, roost environment, activities, etc) inhabiting locations of America with distinct
345 environments. There is a clear need for more studies on the bat virome, including reports of the same
346 bat species inhabiting different environments in a wide geographical distribution. One important issue
347 is the detailed description of the bats environment to identify factors associated with virome differences
348 or risk.

349 Most of the studies on the bat viromes are conducted in human disturbed environments (Ge *et al.*,
350 2012; Baker *et al.*, 2013) or include synanthropic bat species (Geldenhuys *et al.*, 2018). In this study,
351 we report different bat species in different environments located close (<10 km) to rural backyard farms
352 housing domestic animal species, such as swine, cattle, sheep, goat, poultry, cats, and dogs. In these
353 farms, animals are fed with commercial food, agricultural products, human food leftovers, grazing or
354 combinations of them. Animals are kept in rustic pens or outdoors, where they frequently coexist with
355 other domestic species, humans and wild species such as bats that feed on insects (attracted by



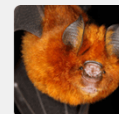
356 spotlights, leftovers or feces), fruit trees or livestock blood. This is the first report of these species
357 virome associated with rural-backyard farms.

358 In this study, the pooling of samples individually or from different sampling sites, limited our ability to
359 evaluate whether these viruses were present (i.e. viral anatomical origin, infected individuals/sampling
360 sites, different viral variants). Although we intended to identify the viruses in individual samples
361 (PCR/RT-PCR), the nucleic acids were consumed in the process of obtaining high-quality libraries.
362 Therefore it is necessary to sequence samples individually or pool them according to specific criteria.
363 This to obtain specific information about viral communities present in different bat populations or
364 subgroups. Always obtaining a sufficient sample to confirm any result.

365 One important issue is to consider analysing other viral communities from different origin samples
366 (pharyngeal, nasal, urine). The previous studies on bats have analyzed the virome from rectal or fecal
367 samples mainly, and there are important viruses that are excreted by other routes (Chua *et al.*, 2002;
368 Wacharapluesadee *et al.*, 2005). Paramyxoviruses (i.e. Nipah, Hendra) have been identified in urine and
369 saliva; coronaviruses in anal, pharyngeal and other bat organs. Han *et al.* (2015) report that humans
370 were infected by Marburg virus after visiting or working in caves in Africa, probably by aerosol
371 transmission (Timen *et al.*, 2009). Capturing animals involves hard work and resources (such as
372 preparing materials, moving to the study site, entering roosts, setting traps and capturing/obtaining and
373 restraint animals, collecting samples, etc.) that must be used to the maximum. Because capturing and
374 sampling is logistical difficult it is important to obtain the most useful samples and materials for more
375 complete and future projects.

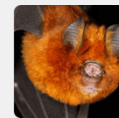
376 Although the use of metagenomics has allowed the investigation of the complexity of the virome in bat
377 species, this tool still present limitations such as those associated with nucleic acid extraction,
378 amplification, enrichment, bioinformatics, database sequences and others. In this work, the major
379 proportion of virome data of the three bat species remained unclassified, making evident the need to
380 increase the knowledge about bat viruses, as well as other microorganisms.

381 These results indicate that viral diversity includes mostly bacteriophages while showing low numbers
382 for known animal viruses. It also shows differences between viral communities of investigated bats
383 which are different from previously reported by other authors. Findings of this study contribute with
384 one more valuable piece in the virome information of these bat species with large geographical
385 distributions.

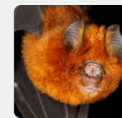


LITERATURE CITED

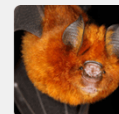
- AFGAN, E., D. BAKER, B. BATUT, M. VAN DEN BEEK, D. BOUVIER, M. ECH, J. CHILTON, D. CLEMENTS, N. CORAOR, B. A. GRÜNING, A. GUERLER, J. HILLMAN-JACKSON, S. HILTEMANN, V. JALILI, H. RASCHE, N. SORANZO, J. GOECKS, J. TAYLOR, A. NEKRUTENKO, and D. BLANKENBERG. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*. 46: W537–W544.
- BAKER, K. S., R. M. LEGGETT, N. H. BEXFIELD, M. ALSTON, G. DALY, S. TODD, M. TACHEDJIAN, C. E. G. HOLMES, S. CRAMERI, L. F. WANG, J. L. HEENEY, R. SUU-IRE, P. KELLAM, A. A. CUNNINGHAM, J. L. N. WOOD, M. CACCAMO, and P. R. MURCIA. 2013. Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology*. 441: 95–106.
- BANERJEE, A., K. KULCSAR, V. MISRA, M. FRIEMAN, and K. MOSSMAN. 2019. Bats and coronaviruses. *Viruses*. 11.
- BOLGER, A. M., M. LOHSE, and B. USADEL. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*. 30: 2114–20.
- CHAO, A., K. H. MA, and T. C. HSIEH. 2016. iNEXT (iNterpolation and EXTrapolation) Online. Program and User's Guide published at http://chao.stat.nthu.edu.tw/wordpress/software_download/. 30043: 1–14.
- CHEN, L., W. GU, C. LIU, W. WANG, N. LI, Y. CHEN, C. LU, X. SUN, Y. HAN, D. KUANG, P. TONG, and J. DAI. 2019. Characteristics of the tree shrew gut virome. *PLOS ONE*. 14: e0212774.
- CHEN, L., B. LIU, J. YANG, and Q. JIN. 2014. DBatVir: The database of bat-associated viruses. *Database*. 2014: 1–7.
- CHONG, R., M. SHI, C. E. GRUEBER, E. C. HOLMES, C. J. HOGG, K. BELOV, and V. R. BARRS. 2019. Fecal Viral Diversity of Captive and Wild Tasmanian Devils Characterized Using Virion-Enriched Metagenomics and Metatranscriptomics. *Journal of Virology*. 93.
- CHUA, K. B., C. LEK KOH, P. S. HOOI, K. F. WEE, J. H. KHONG, B. H. CHUA, Y. P. CHAN, M. E. LIM, and S. K. LAM. 2002. Isolation of Nipah virus from Malaysian Island flying-foxes. *Microbes and Infection*. 4: 145–151.
- CUEVAS-ROMERO JS, BLOMSTROM AL, BERG, M. 2015. Molecular and epidemiological studies of Porcine rubulavirus infection - an overview. *Infection Ecology & Epidemiology*. 5: 1–8.
- DACHEUX, L., M. CERVANTES-GONZALEZ, G. GUIGON, J. M. THIBERGE, M. VANDENBOGAERT, C. MAUFRAIS, V. CARO, and H. BOURHY. 2014. A preliminary study of viral metagenomics of french bat species in contact with humans: Identification of new mammalian viruses. *PLoS ONE*. 9.
- DJIKENG, A., and D. J. SPIRO. 2009. Advancing full length genome sequencing for human RNA viral pathogens. *Future Virology*. 4: 47–53.



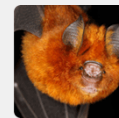
- 421 DONALDSON, E. F., A. N. HASKEW, J. E. GATES, J. HUYNH, C. J. MOORE, and M. B. FRIEMAN. 2010.
422 Metagenomic Analysis of the Viromes of Three North American Bat Species: Viral Diversity
423 among Different Bat Species That Share a Common Habitat. *Journal of Virology*. 84: 13004–
424 13018.
- 425 DREXLER, J. F., V. M. CORMAN, M. A. MÜLLER, G. D. MAGANGA, P. VALLO, T. BINGER, F. GLOZA-
426 RAUSCH, A. RASCHE, S. YORDANOV, A. SEEBENS, S. OPPONG, Y. A. SARKODIE, C. PONGOMBO, A. N.
427 LUKASHEV, J. SCHMIDT-CHANASIT, A. STÖCKER, A. J. B. CARNEIRO, S. ERBAR, A. MAISNER, F.
428 FRONHOFFS, R. BUETTNER, E. K. V. KALKO, T. KRUPPA, C. R. FRANKE, R. KALLIES, E. R. N.
429 YANDOKO, G. HERRLER, C. REUSKEN, A. HASSANIN, D. H. KRÜGER, S. MATTHEE, R. G. ULRICH, E.
430 M. LEROY, and C. DROSTEN. 2012. Bats host major mammalian paramyxoviruses. *Nature*
431 *Communications*. 3: 796.
- 432 FELSENSTEIN, J. 1985. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE
433 BOOTSTRAP. *Evolution*. 39: 783–791.
- 434 GAY, N., K. J. OLIVAL, S. BUMRUNGSRI, B. SIRIAROONRAT, M. BOURGAREL, and S. MORAND. 2014.
435 Parasite and viral species richness of Southeast Asian bats: Fragmentation of area distribution
436 matters. *International Journal for Parasitology: Parasites and Wildlife*. 3: 161–170.
- 437 GE, X., Y. LI, X. YANG, H. ZHANG, P. ZHOU, Y. ZHANG, and Z. SHI. 2012. Metagenomic Analysis of
438 Viruses from Bat Fecal Samples Reveals Many Novel Viruses in Insectivorous Bats in China.
439 *Journal of Virology*. 86: 4620–4630.
- 440 GELDENHUYS, M., M. MORTLOCK, J. WEYER, O. BEZUIDT, E. C. J. SEAMARK, T. KEARNEY, C. GLEASNER,
441 T. H. ERKKILA, H. CUI, and W. MARKOTTER. 2018. A metagenomic viral discovery approach
442 identifies potential zoonotic and novel mammalian viruses in *Neoromicia* bats within South
443 Africa. *PLOS ONE*. 13: e0194527.
- 444 HAN, H. J., H. ling WEN, C. M. ZHOU, F. F. CHEN, L. M. LUO, J. wei LIU, and X. J. YU. 2015. Bats as
445 reservoirs of severe emerging infectious diseases. *Virus Research*. 205: 1–6.
- 446 HE, B., Z. LI, F. YANG, J. ZHENG, Y. FENG, H. GUO, Y. LI, Y. WANG, N. SU, F. ZHANG, Q. FAN, and C.
447 TU. 2013. Virome Profiling of Bats from Myanmar by Metagenomic Analysis of Tissue Samples
448 Reveals More Novel Mammalian Viruses. *PLoS ONE*. 8.
- 449 HERRIDGE, W. P., P. SHIBU, J. O'SHEA, T. C. BROOK, and L. HOYLES. 2020. Bacteriophages of *Klebsiella*
450 spp., their diversity and potential therapeutic uses. *Journal of Medical Microbiology*. 69: 176–194.
- 451 HSIEH, T. C., K. H. MA, and A. CHAO. 2016. iNEXT: an R package for rarefaction and extrapolation of
452 species diversity (Hill numbers). *Methods in Ecology and Evolution*. 7: 1451–1456.
- 453 KUMAR, S., G. STECHER, M. LI, C. KNYAZ, and K. TAMURA. 2018. MEGA X: Molecular Evolutionary
454 Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*. 35: 1547–
455 1549.



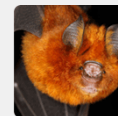
- 456 KUZMIN, I. V., B. BOZICK, S. A. GUAGLIARDO, R. KUNKEL, J. R. SHAK, S. TONG, and C. E. RUPPRECHT.
457 2011. Bats, emerging infectious diseases, and the rabies paradigm revisited. *Emerging Health*
458 *Threats Journal*. 4: 1–17.
- 459 LAMBIN, E. F., B. L. TURNER, H. J. GEIST, S. B. AGBOLA, A. ANGELSEN, C. FOLKE, J. W. BRUCE, O. T.
460 COOMES, R. DIRZO, P. S. GEORGE, K. HOMEWOOD, J. IMBERNON, R. LEEMANS, X. LI, E. F. MORAN,
461 M. MORTIMORE, P. S. RAMAKRISHNAN, J. F. RICHARDS, W. STEFFEN, G. D. STONE, U. SVEDIN, and T.
462 A. VELDKAMP. 2001. The causes of land-use and land-cover change : moving beyond the myths.
463 *Global Environmental Change*. 11: 261–269.
- 464 LEROY, E. M., B. KUMULUNGUI, X. POURRUT, P. ROUQUET, A. HASSANIN, P. YABA, A. DÉLICAT, J. T.
465 PAWESKA, J.-P. GONZALEZ, and R. SWANEPOEL. 2005. Fruit bats as reservoirs of Ebola virus.
466 *Nature*. 438: 575–6.
- 467 LI, L., A. KAPOOR, B. SLIKAS, O. S. BAMIDELE, C. WANG, S. SHAUKAT, M. A. MASROOR, M. L. WILSON,
468 J.-B. N. NDJANGO, M. PEETERS, N. D. GROSS-CAMP, M. N. MULLER, B. H. HAHN, N. D. WOLFE, H.
469 TRIKI, J. BARTKUS, S. Z. ZAIDI, and E. DELWART. 2010a. Multiple Diverse Circoviruses Infect Farm
470 Animals and Are Commonly Found in Human and Chimpanzee Feces. *Journal of Virology*. 84:
471 1674–1682.
- 472 LI, L., J. G. VICTORIA, C. WANG, M. JONES, G. M. FELLERS, T. H. KUNZ, and E. DELWART. 2010b. Bat
473 Guano Virome: Predominance of Dietary Viruses from Insects and Plants plus Novel Mammalian
474 Viruses. *Journal of Virology*. 84: 6955–6965.
- 475 LI, L., T. SHAN, O. B. SOJI, M. M. ALAM, T. H. KUNZ, S. Z. ZAIDI, and E. DELWART. 2011. Possible
476 cross-species transmission of circoviruses and cycloviruses among farm animals. *Journal of*
477 *General Virology*. 92: 768–772.
- 478 LUBY, S. P., E. S. GURLEY, and M. J. HOSSAIN. 2009. Transmission of Human Infection with Nipah
479 Virus. *Clinical Infectious Diseases*. 49: 1743–1748.
- 480 LUIS, A. D., D. T. S. HAYMAN, T. J. O'SHEA, P. M. CRYAN, A. T. GILBERT, J. R. C. PULLIAM, J. N. MILLS,
481 M. E. TIMONIN, C. K. R. WILLIS, A. A. CUNNINGHAM, A. R. FOOKS, C. E. RUPPRECHT, J. L. N.
482 WOOD, and C. T. WEBB. 2013. A comparison of bats and rodents as reservoirs of zoonotic viruses:
483 are bats special? *Proceedings of the Royal Society B: Biological Sciences*. 280: 20122753–
484 20122753.
- 485 MEDELLIN, R. A., H. T. ARITA, and O. SÁNCHEZ. 2007. Identificación de los murciélagos de México,
486 clave de campo. Instituto de Ecología, UNAM. Segunda ed: 1–80.
- 487 MOKILI, J. L., F. ROHWER, and B. E. DUTILH. 2012. Metagenomics and future perspectives in virus
488 discovery. *Current Opinion in Virology*. 2: 63–77.
- 489 MORENO, C. E. 2001. *Métodos Para Medir La Biodiversidad*. 84 p.



- 490 ORGANIZACION MUNDIAL DE LA SALUD. 2021. Organizacion Mundial de la salud Enfermedad por el
491 virus del Ebola. Oms.
- 492 PATZ, J. A., P. DASZAK, G. M. TABOR, A. A. AGUIRRE, M. PEARL, J. EPSTEIN, N. D. WOLFE, A. M.
493 KILPATRICK, J. FOUFOPOULOS, D. MOLYNEUX, D. J. BRADLEY, and M. of the W. G. on L. U. C. D.
494 WORKING GROUP ON LAND USE CHANGE AND DISEASE EMERGENCE. 2004. Unhealthy landscapes:
495 Policy recommendations on land use change and infectious disease emergence. *Environmental*
496 *Health Perspectives*. 112: 1092–8.
- 497 PATZ, J. A., T. K. GRACZYK, N. GELLER, and A. Y. VITTOR. 2000. Effects of environmental change on
498 emerging parasitic diseases. *International Journal for Parasitology*. 30: 1395–405.
- 499 PHILBEY, A., P. KIRKLAND, A. ROSS, H. FIELD, M. SRIVASTAVA, R. DAVIS, and R. LOVE. 2008. Infection
500 with Menangle virus in flying foxes (*Pteropus* spp.) in Australia. *Australian Veterinary Journal*. 86:
501 449–454.
- 502 PHILBEY, A. W., P. D. KIRKLAND, A. D. ROSS, R. J. DAVIS, A. B. GLEESON, R. J. LOVE, P. W. DANIELS, A.
503 R. GOULD, and A. D. HYATT. 1998. An apparently new virus (family Paramyxoviridae) infectious
504 for pigs, humans, and fruit bats. *Emerging Infectious Diseases*. 4: 269–271.
- 505 PIRES, D. P., D. VILAS BOAS, S. SILLANKORVA, and J. AZEREDO. 2015. Phage Therapy: a Step Forward in
506 the Treatment of *Pseudomonas aeruginosa* Infections. *Journal of Virology*. 89: 7449–56.
- 507 RACEY, P. A., C. H. CALISHER, I. V. KUZMIN, C. E. RUPPRECHT, D. E. ANDERSON, G. A. MARSH, X.-Y.
508 GE, B. HU, Z.-L. SHI, G. D. MAGANGA, V. ROUGERON, E. M. LEROY, G. TACHEDJIAN, J. A.
509 HAYWARD, J. CUI, C. KOHL, A. KURTH, K. QUEEN, M. SHI, L. J. ANDERSON, S. TONG, J. H. EPSTEIN,
510 H. E. FIELD, K. J. OLIVAL, C. C. WEEKLEY, P. DASZAK, J. A. PALLISTER, D. J. MIDDLETON, J. W.
511 WYNNE, M. TACHEDJIAN, M. L. BAKER, and P. ZHOU. 2015. Bats and Viruses A New Frontier of
512 *Emerging Infectious Diseases*(L.-F. WANG and C. COWLED, eds.). Wiley Blackwell, 387 p.
- 513 RICO, O., R. OJEDA-FLORES, J. SOTOMAYOR-BONILLA, C. ZAMBRANA-TORRELIO, E. LOZA, A. A. AGUIRRE,
514 and G. SUZÁN. 2015. Diversidad viral de comunidades de murciélagos en paisajes transformados
515 de México. Rico, Oscar Ojeda-Flores, Rafael Sotomayor-Bonilla, Jesús Zambrana-Torrelío, Carlos
516 Loza, Elizabeth Aguirre, A Alonso Suzán, Gerardo. 2: 1–23.
- 517 SACHSENRODER, J., S. O. TWARDZIOK, M. SCHEUCH, and R. JOHNE. 2014. The general composition of the
518 faecal virome of pigs depends on age, but not on feeding with a probiotic bacterium. *PLoS ONE*.
519 9: 15–22.
- 520 SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing
521 phylogenetic trees. *Molecular Biology and Evolution*. 4: 406–425.
- 522 SALMIER, A., S. TIRERA, B. De THOISY, A. FRANC, E. DARCISSAC, D. DONATO, C. BOUCHIER, V. LACOSTE,
523 and A. LAVERGNE. 2017. Virome analysis of two -sympatric bat species (*Desmodus rotundus* and
524 *Molossus molossus*) in French Guiana. *PLoS ONE*. 1–25.



- 525 SAUSSET, R., M. A. PETIT, V. GABORIAU-ROUTHIAU, and M. DE PAEPE. 2020. New insights into intestinal
526 phages. *Mucosal Immunology*. 13: 205–215.
- 527 SHAN, T., L. LI, P. SIMMONDS, C. WANG, A. MOESER, and E. DELWART. 2011. The Fecal Virome of Pigs
528 on a High-Density Farm. *Journal of Virology*. 85: 11697–11708.
- 529 SMITH, I., and L. F. WANG. 2013. Bats and their virome: An important source of emerging viruses
530 capable of infecting humans. *Current Opinion in Virology*. 3: 84–91.
- 531 TAMURA, K., M. NEI, and S. KUMAR. 2004. Prospects for inferring very large phylogenies by using the
532 neighbor-joining method. *Proceedings of the National Academy of Sciences*. 101: 11030–11035.
- 533 TIAN, H., C. LU, J. YANG, K. BANGER, D. N. HUNTZINGER, C. R. SCHWALM, A. M. MICHALAK, R. COOK,
534 P. CIAIS, D. HAYES, M. HUANG, A. ITO, A. K. JAIN, H. LEI, J. MAO, S. PAN, W. M. POST, S. PENG, B.
535 POULTER, W. REN, D. RICCIUTO, K. SCHAEFER, X. SHI, B. TAO, W. WANG, Y. WEI, Q. YANG, B.
536 ZHANG, and N. ZENG. 2015. Global patterns and controls of soil organic carbon dynamics as
537 simulated by multiple terrestrial biosphere models: Current status and future directions. *Global*
538 *Biogeochemical Cycles*. 29: 775–792.
- 539 TIMEN, A., M. P. G. KOOPMANS, A. C. T. M. VOSSEN, G. J. J. van DOORNUM, S. GÜNTHER, F. van den
540 BERKMORTEL, K. M. VERDUIN, S. DITTRICH, P. EMMERICH, A. D. M. E. OSTERHAUS, J. T. van DISSEL,
541 and R. A. COUTINHO. 2009. Response to Imported Case of Marburg Hemorrhagic Fever, the
542 Netherlands. *Emerging Infectious Diseases*. 15: 1171.
- 543 WACHARAPLUESADEE, S., B. LUMLERTDACHA, K. BOONGIRD, S. WANGHONGSA, L. CHANHOME, P. ROLLIN,
544 P. STOCKTON, C. E. RUPPRECHT, T. G. KSIAZEK, and T. HEMACHUDHA. 2005. Bat Nipah virus,
545 Thailand. *Emerging Infectious Diseases*. 11: 1949–1951.
- 546 WANG, L. F., and C. COWLED. 2015. Bats and Viruses: A New Frontier of Emerging Infectious Diseases.
547 1-368 p.
- 548 WANG, L. F., Z. SHI, S. ZHANG, H. FIELD, P. DASZAK, and B. T. EATON. 2006. Review of bats and SARS.
549 *Emerging Infectious Diseases*. 12: 1834–1840.
- 550 WANG, L. F., E. HANSSON, M. YU, K. B. CHUA, N. MATHE, G. CRAMERI, B. K. RIMA, J. MORENO-LÓPEZ,
551 and B. T. EATON. 2007. Full-length genome sequence and genetic relationship of two
552 paramyxoviruses isolated from bat and pigs in the Americas. *Archives of Virology*. 152: 1259–
553 1271.
- 554 WIETHOELTER, A. K., D. BELTRÁN-ALCRUDO, R. KOCK, and S. M. MOR. 2015. Global trends in infectious
555 diseases at the wildlife–livestock interface. *Proceedings of the National Academy of Sciences*.
556 112: 9662–9667.
- 557 WILLE, M., and E. C. HOLMES. 2020. Wild birds as reservoirs for diverse and abundant gamma- and
558 deltacoronaviruses. *FEMS Microbiology Reviews*. 44: 631.



- 559 WU, Z., L. YANG, X. REN, G. HE, J. ZHANG, J. YANG, Z. QIAN, J. DONG, L. SUN, Y. ZHU, J. DU, F. YANG,
560 S. ZHANG, and Q. JIN. 2016. Deciphering the bat virome catalog to better understand the ecological
561 diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME Journal*. 10: 609–
562 620.
- 563 WU, Z., X. REN, L. YANG, Y. HU, J. YANG, G. HE, J. ZHANG, J. DONG, L. SUN, J. DU, L. LIU, Y. XUE, J.
564 WANG, F. YANG, S. ZHANG, and Q. JIN. 2012. Virome analysis for identification of novel
565 mammalian viruses in bat species from Chinese provinces. *Journal of Virology*. 86: 10999–1012.
- 566 YAIW, K. C., J. BINGHAM, G. CRAMERI, B. MUNGALL, A. HYATT, M. YU, B. EATON, D. SHAMALA, L.-F.
567 WANG, and K. THONG WONG. 2008. Tioman Virus, a Paramyxovirus of Bat Origin, Causes Mild
568 Disease in Pigs and Has a Predilection for Lymphoid Tissues. *Journal of Virology*. 82: 565–568.
- 569 ZERBINO, D. R. 2010. Using the Velvet de novo assembler for short-read sequencing technologies. *Curr.*
570 *Protoc. Bioinforma.* 1–13.
- 571 ZERBINO, D. R., AND E. BIRNEY. 2008. VELVET: ALGORITHMS FOR DE NOVO SHORT READ ASSEMBLY USING
572 DE BRUIJN GRAPHS. *GENOME RESEARCH*. 18: 821–829.

Table 1

[Download source file \(6.83 kB\)](#)

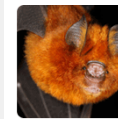
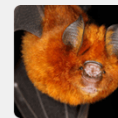


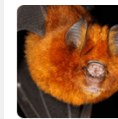
Table 1. Information on the samples and location of the sampled bat roosts in Morelos, Mexico.

	Sampled roosts sites			
	Quilamula, Tlaquiltenango (A)	Xoxocotla, Puente de Ixtla (B)	San Miguel 30, Tlaltizapan (C)	Chimalacatlán, Tlaquiltenango (D)
Geographical location	18°32'10" N, 98°59'54" O	18°41'51" N, 99°14'05" O	18°45'0" N, 99°11'22" O	18°27'09" N, 99°04'34" O
Habitat type immediately surrounding the bat roosts (the vegetation type in all of these areas was low deciduous forest of warm climate).	Located at the base of a chain of hills, that surround the Quilamula's dam, near to crops land and the Quilamula village.	Located in middle height of a slope. Some of the surrounding highest area were used to crops and, lesser frequency to livestock.	An extended forked subterranean cave, located under a plain terrain surrounded by land used for crops, floral greenhouses, rural houses, and other human settlements.	Located in the middle height of a hill (in the area of hills), near valleys some of which were used for crops.
Type of roost	Natural	Natural	Natural	Abandoned mine
Bat species identified in roost	<i>Pteronotus parnellii</i> , <i>Glossophaga</i> sp, <i>Desmodus rotundus</i>	<i>Artibeus</i> sp, <i>Balantiopteryx plicata</i>	<i>Artibeus phaeotis</i> , <i>Pteronotus parnellii</i> , <i>Natalus lanatus</i> , <i>Glossophaga</i> sp.	<i>Macrotus waterhousii</i> , <i>Artibeus</i> sp, <i>Glossophaga</i> sp, <i>Myotis</i> sp
Sampled bat species (n)	<i>Pteronotus parnellii</i> (11)	<i>Artibeus</i> sp. (3)	<i>Artibeus phaeotis</i> (11) <i>Pteronotus parnellii</i> (2)	<i>Artibeus</i> sp. (3) <i>Macrotus waterhousii</i> (9)
Distance between sampled roosts and rural settlement	4.2 km	3.18 km	8.29 km	1.57 km
Distance between sampled bat roosts	B: 30.09 km C: 30.3 km D: 12.0 km	A: 30.09 km C: 8.4 km D: 31.7 km	A: 30.3 km B: 8.4 km D: 34.5 km	A: 12.0 km B: 31.7 km C: 34.5 km

**Table 2.** More abundant viruses identified by homology (BLAST) analysis of the reads in each bat species data set.

Viral family (reads)	Virus(es) species (reads)
<i>Artibeus</i> sp	
Animal viruses	
Circoviridae (7)	Po-Circo-like-virus 51 (5); Fur seal faeces-associated circular DNA virus (1); Porcine associated stool circular DNA virus (1).
Herpesviridae (1)	Gallid herpesvirus 2 (1)
Retroviridae (2)	Squirrel monkey retrovirus (2)
Bacteriophages	
Myoviridae (131)	T4 likevirus† (63); Klebsiella phage JD001 (57)
Podoviridae (92)	Klebsiella phage KP34 (34); T7virus† (26); Phikmvlikevirus† (17)
Siphoviridae (257)	Shigella phage pSf1-1 (90); Jerseyvirus† (68); K1gvirus† (71)
<i>Macrotus waterhousii</i>	
Animal viruses	
Circoviridae (18)	Po-Circo-like-virus 51 (15); Fur seal faeces-associated circular DNA virus (2); Po-Circo-like-virus 22 (1)
Herpesviridae (1)	Human herpesvirus (1)
Parvoviridae (2)	Blatella germanica densovirus (2)
Reoviridae (1)	Aquareovirus A (1)
(Retroviridae (1)	Jaagsiekte sheep retrovirus (1)
Bacteriophages	
Myoviridae (220)	T4likevirus† (130); V5virus† (63); Klebsiella phage JD001 (23)
Podoviridae (83)	T7likevirus† (55); Phi29likevirus† (14);
Siphoviridae (555)	Shigella phage pSf1 (207); K1gvirus† (165); Jerseyvirus† (127); Pis4avirus† (13); T5likevirus† (11)
<i>Pteronotus parnellii</i>	
Animal viruses	
Circoviridae (5)	Po-Circo-like-virus 51 (4); Pig stool associated circular ssDNA virus (1)
Coronaviridae (35)	Coronavirus Mex_CoV-10 (6); Rousettus bat CoV HKU9 (3); Severe acute respiratory related CoV (3); Zaria bat CoV (2); Avian coronavirus (1); Bat coronavirus BtCoV/133/2005 (1); Middle East respiratory syndrome coronavirus (1); Miniopertus bat CoV HKU8 (1)
Retroviridae (5)	Human endogenous retrovirus (4); citrus endogenous pararetrovirus (1).
Hepeviridae (1)	Hepatitis E virus (1)
Bacteriophages	
Myoviridae (73)	T4likevirus† (48); V5virus† (15)
Podoviridae (41)	T7likevirus† (28); Phi29likevirus† (10)
Siphoviridae (198)	Shigella phage pSf1-1 (85); K1gvirus† (56); Jerseyvirus† (41)
Algae virus	
Phycodnaviridae (1)	

† viral genus

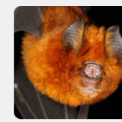
Table 3[Download source file \(6.45 kB\)](#)**Table 3.** Information of sequences of more abundant animal viruses submitted to the GenBank.

Name of sequence	Accession assigned	Length (bp)	Contig coverage	RefSeq/sequence with higher homology in GenBank	Similarity	Gene(s)	Gene(s) Coverage ^a
B-Coronavirus_Ppar/MEX/2016_RNApol_CDS-1	MW356910	482	3.6	KC633194 Betacoronavirus BtCoV/KCR155/Pte_par/CRC/2012 RNA-dependent_RNA polymerase gene, partial cds KC633194	94 %	ORF1ab	59.1 %
B-Coronavirus_Ppar/MEX/2016_RNApol_CDS-2	MW356911	201	2	KC633193 Betacoronavirus BtCoV/KCR180/Pte_par/CRC/2011 RNA-dependent_RNA polymerase gene, partial cds	93 %	ORF1ab	24.6 %
Circovirus/Aspp/MEX/2016/Rep-partial-CDS	MW584366	162	2	MK377683_Circovirus_PoCirV_VIRES_NX02_C7 replicase gene, partial cds	82 %	Rep	22.6 %
Circovirus/Mwat/MEX/2016/Rep-partial-CDS	MW584367	252	2	MK377656.1 Circovirus sp. isolate PoCirV_VIRES_LN01_C1 replicase gene, partial cds	87.7 %	Rep	26.6 %
Circovirus/Ppar/MEX/2016/Rep-partial-CDS	MW584368	181	2	NC 025684 Po-circo-like virus 51, complete genome	90 %	Rep	6.7 %

^a Percentage of contig coverage with respect to the RefSeq/sequence genome/partial cds with greater homology.^b Sequence used for phylogenetic analysis.

Figure 1

[Download source file \(1.91 MB\)](#)



*Acta
Chiropterologica*

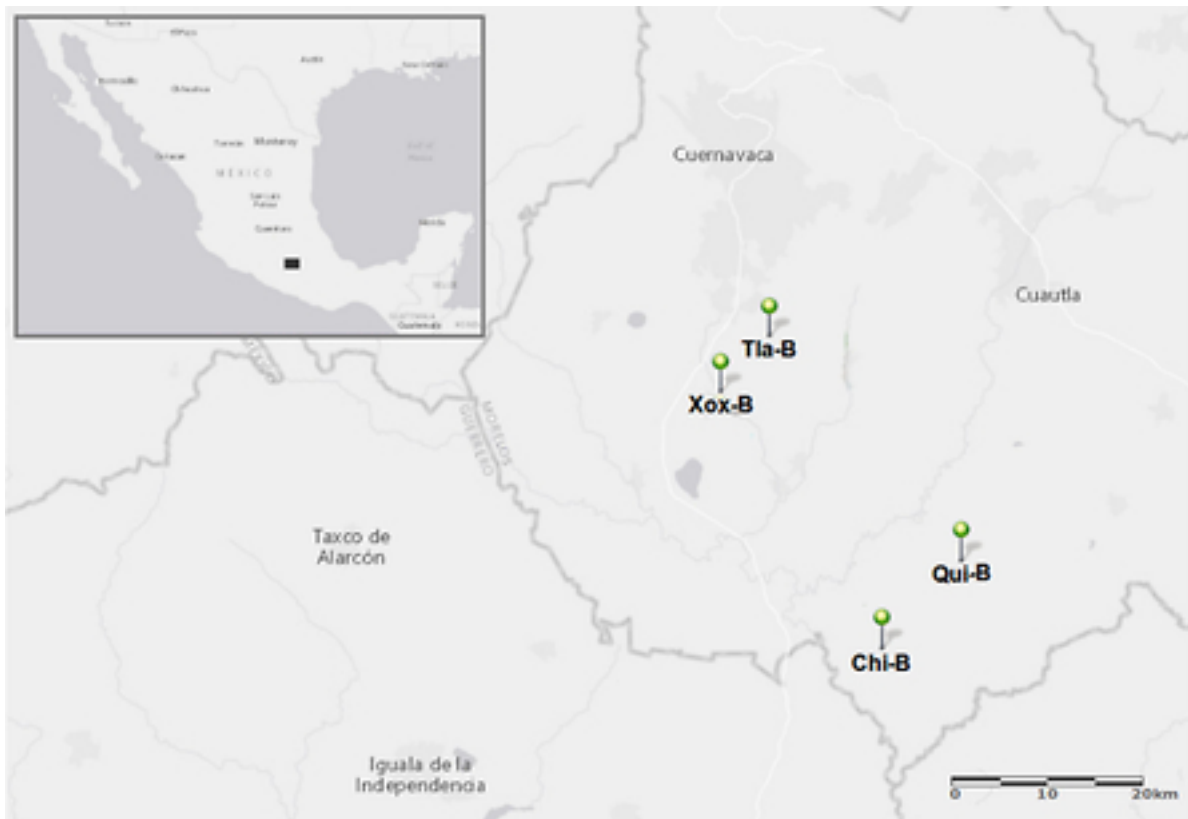


Figure 1. Geographic location of bat roosts (B) in Xoxocotla (Xox), Chimalacatlan (Chi), Quilamula (Qui) and Tlaltizapan (Tla) in Morelos. Location of the study area (big map) is represented on the map of the Mexican Republic (left corner small map) with the dark rectangle (■). The map was made with ArcGIS on line (<https://www.arcgis.com/home/webmap/viewer.html> date accessed: august-13-2017).

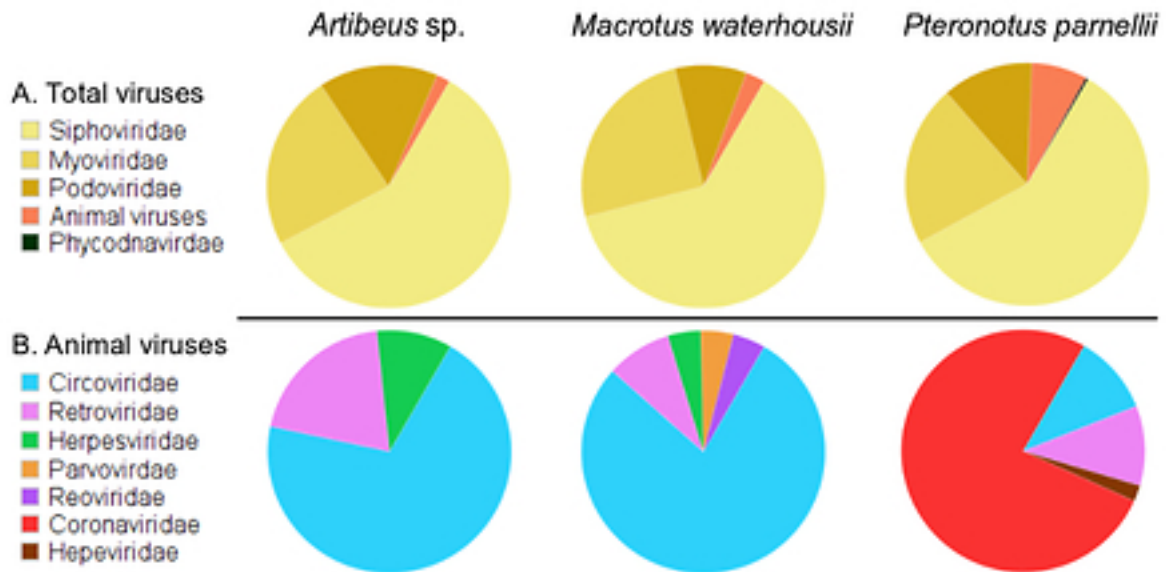
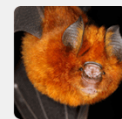


Figure 2. Summary of sequences (reads) assigned by BLASTn/BLASTx. (A) Sequences of total viruses. (B) Sequences assigned to animal viruses.

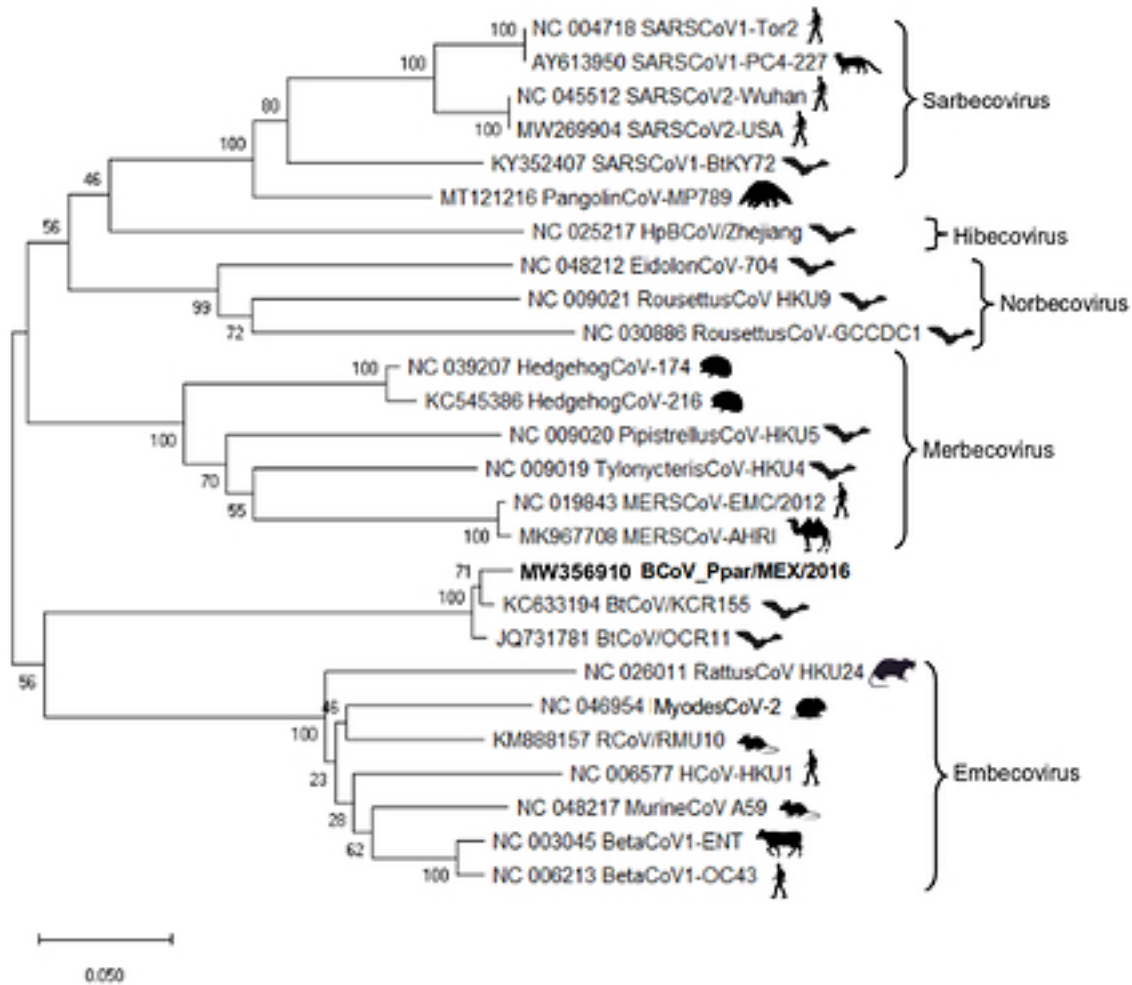
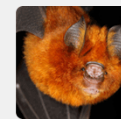


Figure 3. Phylogenetic tree on nucleotide sequences of the RNA dependent RNA polymerase gene (ORF1ab) of species of the Coronavirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the contig of 482 bases found by next generation sequencing (bold) in reads of *P. parnellii*. The optimal tree with the sum of branch length = 2.11171885 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 24 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

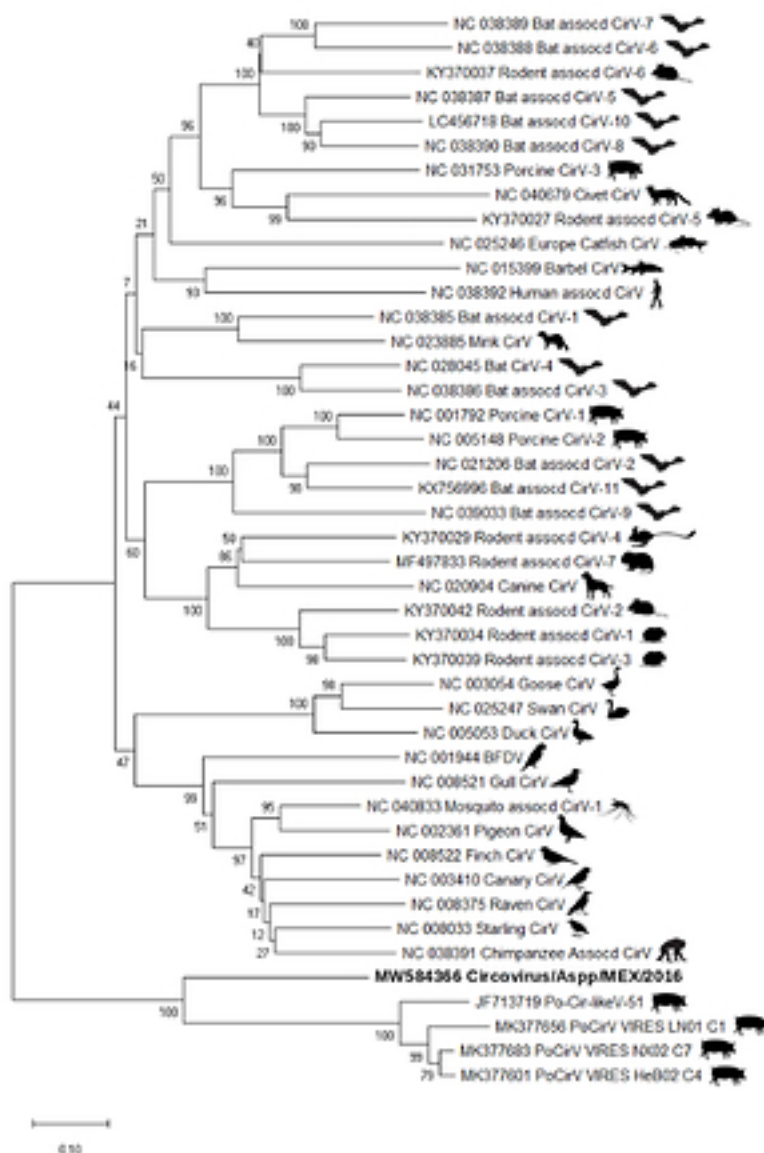
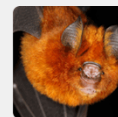


Figure 4. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequence of 252 bases found by next generation sequencing (bold) in reads of *Artibeus* spp. The optimal tree with the sum of branch length = 10.33514381 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

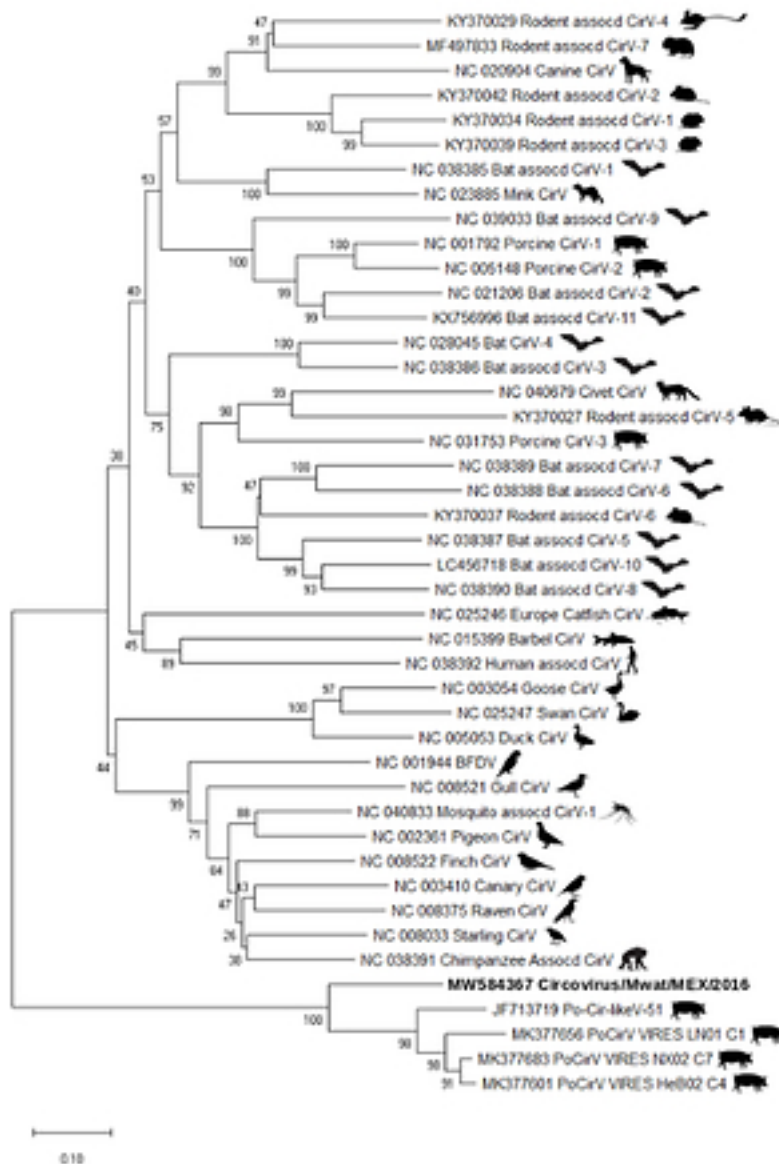
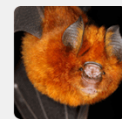


Figure 5. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequence of 252 bases found by next generation sequencing (bold) in reads of *Macrotus waterhousii*. The optimal tree with the sum of branch length = 9.93812575 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

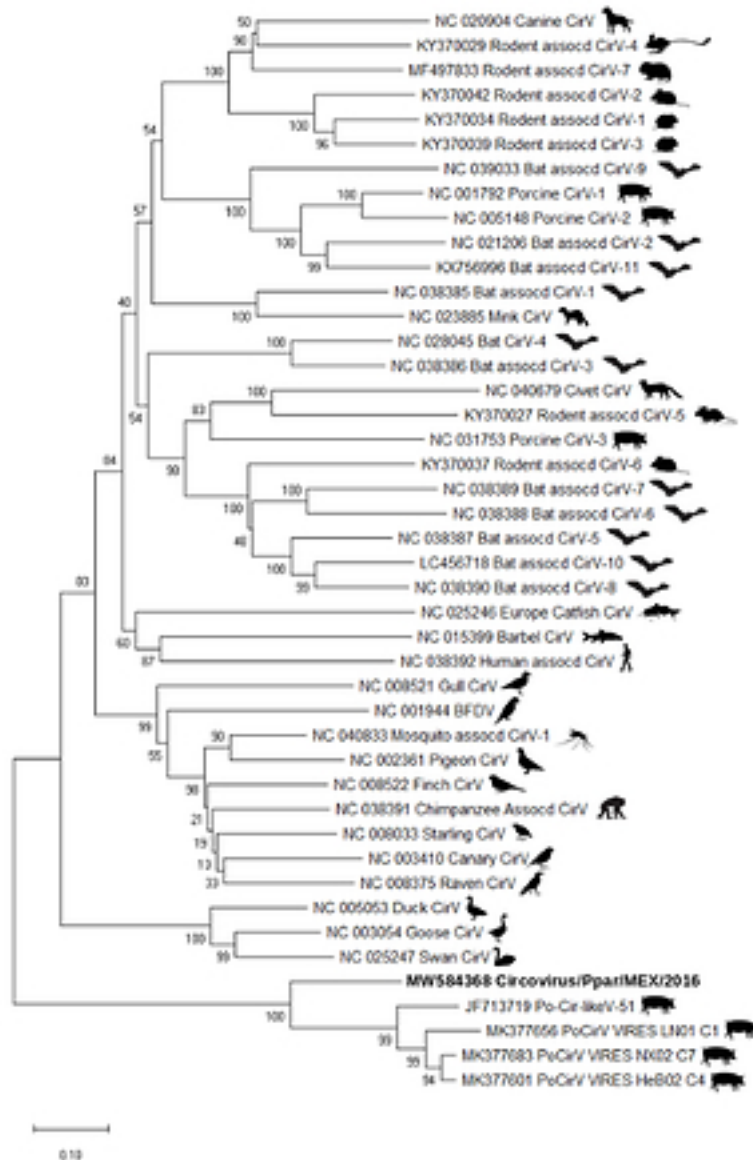
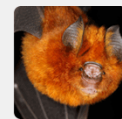


Figure 6. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequence of 250 bases found by next generation sequencing (bold) in reads of *Pteronotus parnellii*. The optimal tree with the sum of branch length = 10.27156581 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

Figure 7

[Download source file \(1.69 MB\)](#)

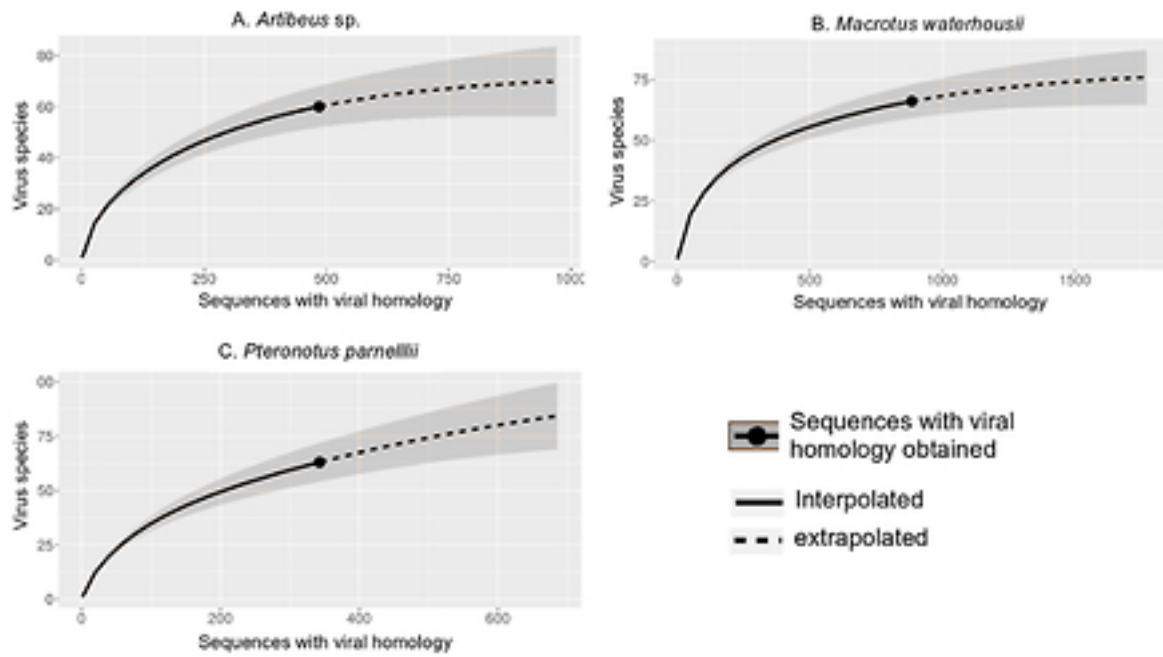
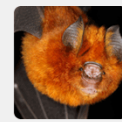
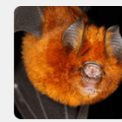


Figure 7. Diversity of virus species estimated by rarefaction for each species of bat (●). This graphic shows that viral species identified by next generation sequencing provided a good representation of the estimated diversity.



Manuscript body

Manuscript body 1 - [Download source file \(34.19 kB\)](#)

Tables

Table 1 - [Download source file \(6.83 kB\)](#)

Table 1. Information on the samples and location of the sampled bat roosts in Morelos, Mexico.

Table 2 - [Download source file \(6.75 kB\)](#)

Table 2. More abundant viruses identified by homology (BLAST) analysis of the reads in each bat species data set.

Table 3 - [Download source file \(6.45 kB\)](#)

Table 3. Information of sequences of more abundant animal viruses submitted to the GenBank.

Figures

Figure 1 - [Download source file \(1.91 MB\)](#)

Figure 1. Geographic location of bat roosts (B) in Xoxocotla (Xox), Chimalacatlan (Chi), Quilamula (Qui) and Tlaltizapan (Tla) in Morelos. Location of the study area (big map) is represented on the map of the Mexican Republic (left corner small map) with the dark rectangle (■). The map was made with ArcGIS on line (<https://www.arcgis.com/home/webmap/viewer.html> date accessed: august-13-2017).

Figure 2 - [Download source file \(1.81 MB\)](#)

Figure 2. Summary of sequences (reads) assigned by BLASTn/BLASTx. (A) Sequences of total viruses. (B) Sequences assigned to animal viruses.

Figure 3 - [Download source file \(1.8 MB\)](#)

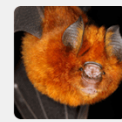
Figure 3. Phylogenetic tree on nucleotide sequences of the RNA dependent RNA polymerase gene (ORF1ab) of species of the Coronavirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the contig of 482 bases found by next generation sequencing (bold) in reads of *P. parnellii*. The optimal tree with the sum of branch length = 2.11171885 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 24 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

Figure 4 - [Download source file \(1.59 MB\)](#)

Figure 4. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequence of 252 bases found by next generation sequencing (bold) in reads of *Artibeus* spp. The optimal tree with the sum of branch length = 10.33514381 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

Figure 5 - [Download source file \(1.84 MB\)](#)

Figure 5. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequence of 252 bases found by next generation sequencing (bold) in reads of *Macrotus waterhousii*. The optimal tree with the sum of branch length = 9.93812575 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary



distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

Figure 6 - [Download source file \(1.67 MB\)](#)

Figure 6. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequence of 250 bases found by next generation sequencing (bold) in reads of *Pteronotus parnellii*. The optimal tree with the sum of branch length = 10.27156581 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).

Figure 7 - [Download source file \(1.69 MB\)](#)

Figure 7. Diversity of virus species estimated by rarefaction for each species of bat (●). This graphic shows that viral species identified by next generation sequencing provided a good representation of the estimated diversity.

Supplementary Material

File 1 - [Download source file \(1.64 MB\)](#)

Supplementary Figure 1. Phylogenetic tree on nucleotide sequences of the Rep gene of species of the Circovirus genus was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). It shows the relationship of the sequences of 85 bases found by next generation sequencing (bold) in reads of the three species of bats studied (*Artibeus* spp., *Macrotus waterhousii* and *Pteronotus parnellii*). The optimal tree with the sum of branch length = 10.45849938 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. This analysis involved 45 reference nucleotide sequences. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).