

UNIVERSITAT POLITÈCNICA DE VALÈNCIA



**UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA**

**ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA
AGRONÒMICA I DEL MEDI NATURAL**

**Evaluación del intercambio de genes entre
micovirus y sus huéspedes: el caso de la proteína de
la envuelta viral**

Trabajo Final de Grado en Biotecnología

Curso académico: 2019-2020

Autor: Alfonso i Comos, Pau

Tutor externo: Elena Fito, Santiago Francisco

Cotutor Académico: Forment Millet, José Javier

Valencia, Julio de 2020

Convocatoria de Julio de 2020



Evaluación del intercambio de genes entre micovirus y sus huéspedes: el caso de la proteína de la envuelta viral

Valencia, Julio de 2020

Alumno: Alfonso i Comos, Pau

Tutor externo: Elena Fito, Santiago Francisco¹

Cotutor académico: Forment Millet, José Javier²

Resumen

Los micovirus son virus que infectan a hongos y son una de las clases menos conocidas de virus hasta la fecha. La gran mayoría de los micovirus identificados son virus de RNA y dentro de estos predominan aquellos de clase III, con un genoma dsRNA. El interés de conocer los micovirus radica en su uso como agentes de biocontrol de plagas fúngicas que pueden afectar a los cultivos o precisamente para mantener buenas medidas de higiene en cultivos de champiñones y otros hongos comestibles. A nivel básico, el estudio filogenético de los micovirus, como del resto de virus, nos permite hacernos una idea de su ecología y de cómo han coevolucionado con sus huéspedes.

Este estudio persigue comprobar la hipótesis de que el genoma de los micovirus es un mosaico compuesto de una región propiamente viral, donde se encuentra el gen de la RNA replicasa dependiente de RNA (RdRP), y otra de origen eucariota, que contendría el gen de la proteína de la envuelta (CP), que en algún momento de la evolución pudo haberse transferido horizontalmente desde un hospedador fúngico a un virus ancestral.

Para ello, se reconstruyeron árboles filogenéticos con las secuencias proteicas RdRP y CP de distintos grupos de micovirus de dsRNA, haciendo coincidir las especies de RdRP y CP para poder hacer una evaluación de la cofilogenia, es decir, en qué medida ambos loci explican la misma historia evolutiva. Los árboles separaban los mismos grupos, pero en el árbol de CP no se consiguen establecer relaciones entre grupos que sí se observan con RdRP.

La búsqueda PSI-BLAST de la CP de un grupo de micovirus no clasificados de dsRNA reveló una gran similitud de secuencia con una proteína hipotética de hongos presente en numerosas especies, mayor a la que presenta la CP de este grupo con la de otros micovirus. La comparación entre las estructuras simuladas *in silico* de estas proteínas fúngicas y la CP de este grupo de virus no clasificados mostró una importante homología. Estos datos, en conjunto, sugieren que este gen eucariota, de función desconocida en el hongo, se transfirió a un antepasado de estos virus, adoptando una nueva función como CP. Eventos de transferencia de este tipo pueden ser responsables de que no se puedan establecer las mismas relaciones con los árboles de CP y RdRP.

Palabras clave:

Micovirus, filogenia, evolución viral, redes filogenéticas, transferencia de genes virus-huésped, coevolución, proteína de la envuelta, alineamiento estructural de proteínas

1 Instituto de Biología Integrativa de Sistemas (I2SysBio), CSIC-UV

2 Instituto Universitario Mixto de Biología Molecular y Celular de Plantas (IBMCP), CSIC-UPV

Examining the gene transfer between mycoviruses and its hosts: a look into the case of the coat protein

Abstract

Viruses that infect fungi are known as mycovirus. Recently, more and more mycovirus species are being identified and having their genomes sequenced. Nevertheless, the majority of mycoviruses are yet to be discovered. By now, most of known mycoviruses have RNA-based genomes, being most of these dsRNA viruses or class III viruses, according to the Baltimore classification. Main applications of having an integrated knowledge of mycoviruses and their environment comprehend their use as biocontrol agents against crop's fungal pathogens and their exclusion from edible mushroom cultivations. A phylogenetic approach to mycoviruses would yield many information at a fundamental level involving mycoviral evolution and their relation to eukaryotic species ecologically related i.e. mainly the fungi that serve them as hosts.

This study's main hypothesis is that of the mycoviral genome having a mosaic structure, this means the genome having two origins, one viral and other eukaryotic, the viral part comprising the RNA-dependent RNA polymerase (RdRP) gene and the eukaryotic part being composed of the coat protein (CP) gene, which might have been transferred to an ancestral virus from its fungal host.

For this purpose, two phylogenetic trees were reconstructed for RdRP and CP, using sequences of the same virus species on each side to compare the results and address the question of both loci explaining the same evolutionary history. The CP tree was not able to find relationships between virus groups as the RdRP tree did, although the same final groups were segregated with both trees.

A PSI-BLAST search with the CP sequence of the different mycovirus groups included in the study was able to find an hypothetical protein with many orthologues in fungal genomes which showed a very significant sequence homology with the CP of a unclassified dsRNA mycovirus group. Some structures of different gruoups' CP and this hypothetical protein were predicted *in silico* to establish relationships that could not be seen working only with sequences. The CP of this unclassified mycovirus group was distinct to others' and more similar to this fungal hypothetical protein. Both results suggest that this eukaryotic gene is a strong candidate to have been transferred to this viruses' ancestor and to have replaced the native viral CP of this group. This and other plausible horizontal gene transfers between mycoviruses and their hosts that were not seen in this study could explain at least partly the high divergence of groups observed in the CP tree for viruses related in RdRP.

Keywords:

Mycovirus, phylogenetics, viral evolution, phylogenetic networks, virus-host gene transfer, coevolution, coat protein, structural alignment

ÍNDICE GENERAL

1. Introducción	1
1.1. Historia de la micovirología	1
1.2. Ecología de los micovirus	2
1.3. Taxonomía y características	4
1.4. Origen evolutivo de los micovirus	6
2. Materiales y métodos	9
2.1. Búsquedas BLAST	9
2.2. Alineamiento de secuencias	9
2.3. Análisis filogenético	9
2.3.1. Construcción de árboles filogenéticos	9
2.3.2. Construcción de redes filogenéticas	10
2.4. Análisis estructural de CP y proteínas del huésped	10
2.4.1. Predicción de estructuras y función	10
2.4.2. Alineamiento de estructuras	11
3. Resultados	11
3.1. Búsquedas BLAST	11
3.2. Análisis filogenético	16
3.2.1. Árboles filogenéticos	16
3.2.2. Red filogenética de la secuencia concatenada de RdRP y CP	20
3.2.3. Red filogenética de CP con la proteína hipotética de hongos	22
3.3. Análisis estructural de CP y la proteína de hongo similar a ThBMV1	24
3.3.1. Predicción estructural	24
3.3.2. Predicción funcional	27
3.3.3. Alineamiento de estructuras entre grupos	27
4. Discusión	29
4.1. Relaciones entre los grupos de virus estudiados	29
4.2. Secuencias en hongos similares a CP de micovirus	29
4.3. Similitud estructural de la CP de ThBMV1-like con una proteína hipotética de hongos	30
5. Conclusión	31
6. Bibliografía	32

ÍNDICE DE TABLAS

Tabla 1	4
Tabla 2	10
Tabla 3	14

ÍNDICE DE FIGURAS

Figura 1	2
Figura 2	13
Figura 3	17
Figura 4	19
Figura 5	20
Figura 6	21
Figura 7	22
Figura 8	23
Figura 9	25
Figura 10	26
Figura 11	28

1. INTRODUCCIÓN

1.1. HISTORIA DE LA MICOVIROLOGÍA

El primer indicio de la existencia de micovirus se remonta a los años 30, cuando se aisló un virus capaz de provocar la lisis de levaduras cultivadas. Sin embargo, por aquel entonces se consideró que el causante de esta muerte celular era un bacteriófago (Wieringa y Wiebols, 1936). En las décadas posteriores se hicieron otras observaciones de valor en distintas partes del mundo sin demostrar el origen viral de las patologías que se identificaban. Sinden y Hauser (1948) describieron detalladamente la enfermedad de La France en *Agaricus bisporus*, el champiñón comestible. Sus resultados sugerían por primera vez la existencia de enfermedades virales en hongos superiores capaces de reducir el crecimiento del hongo y provocar fenotipos aberrantes.

Otros estudios apuntaron la existencia en el hongo *Helminthosporium victoriae* de enfermedades de transmisión citoplasmática y, por tanto, heredables (Lindberg, 1959). Según Marcou (1961), la senescencia en *Podospora anserina* se transmitía por un “principio infectivo extracromosómico”. A pesar de no conocer su naturaleza, se demostró que podía propagarse a células sanas a través de las hifas al comunicarse el citoplasma de distintos individuos de la misma especie y producir el fenotipo enfermo. A finales de los años 50 se demostró en *Olpidium brassicae* que los hongos podían ser vectores de virus de plantas en el caso de la enfermedad de las venas ensanchadas de la lechuga, aunque no se determinó si el virus podía multiplicarse en el hongo (Grogan y Campbell, 1966).

Por otra parte, se había comprobado que filtrados de *Penicillium funiculosum* eran capaces de inducir el interferón, provocando así una acción antiviral en ratones inoculados con el virus de la gripe porcina (Shope, 1953a). Esta sustancia inductora del interferón en animales que Shope llamó “helenina” perdía su actividad si el filtrado era sometido unos minutos a elevadas temperaturas o después de años de almacenamiento en el laboratorio (Shope, 1953b). Sin embargo, no pudo demostrar cuál era la naturaleza química de la helenina. Paralelamente, se habían obtenido resultados muy similares en *Penicillium stoloniferum* (Powell et al., 1952). Tanto es así que Shope indica en su estudio que no puede estar seguro de que la helenina no fuera la misma sustancia inductora del interferón que se había aislado de *P. stoloniferum*. Cabe destacar que los interferones no se descubrirían hasta varios años más tarde (Isaacs y Lindenmann, 1957), por lo que Powell y Shope solo concluyeron que esta sustancia era capaz de hacer que los ratones sobreviviesen a dosis de virus notablemente superiores a las reconocidas como letales.

Poco más de una década más tarde se sabría que esa acción antiviral se debía a la presencia de virus de doble cadena de RNA (dsRNA) en el citoplasma de mohos del género *Penicillium*. El sistema inmune reconoce en estos virus una serie de patrones moleculares asociados a patógenos (PAMPs) como el dsRNA o la propia partícula viral para inducir el interferón y disminuir la capacidad replicativa de los virus (Ellis y Kleinschmidt, 1967; Buck et al., 1971).

Pero no fue hasta 1962 que nació definitivamente la micovirología con el trabajo de Hollings (1962). Hollings diferenció por la morfología de sus viriones tres virus relacionados con la enfermedad de La France descrita unos años antes en *A. bisporus*.

Hasta la fecha, algo más de 2.000 artículos sobre micovirus se han publicado en PubMed (<https://pubmed.ncbi.nlm.nih.gov/>). Hay una gran diferencia entre la atención que han recibido los micovirus y otros tipos de virus, como los virus de plantas (Figura 1). Los primeros estudios estuvieron motivados por su impacto económico como patógeno de levaduras y hongos comestibles como *A. bisporus* y su actividad inductora del interferón, pero poco a poco se han ido identificando nuevos campos donde la micovirología resulta interesante. Entre ellos emerge su uso como agentes de biocontrol de hongos patógenos de plantas y animales, algo que ha sido objeto de estudio de otros tantos artículos (García-Pedrajas et al., 2019). Con la llegada de las plataformas de secuenciación masiva ‘NGS’ se ha agilizado la caracterización de nuevas especies de micovirus, sobre todo con el

uso de estrategias de metagenómica (Marzano et al. 2016). Esto supone una fuente de información clave para el análisis de la historia evolutiva de los micovirus, objeto del presente estudio.

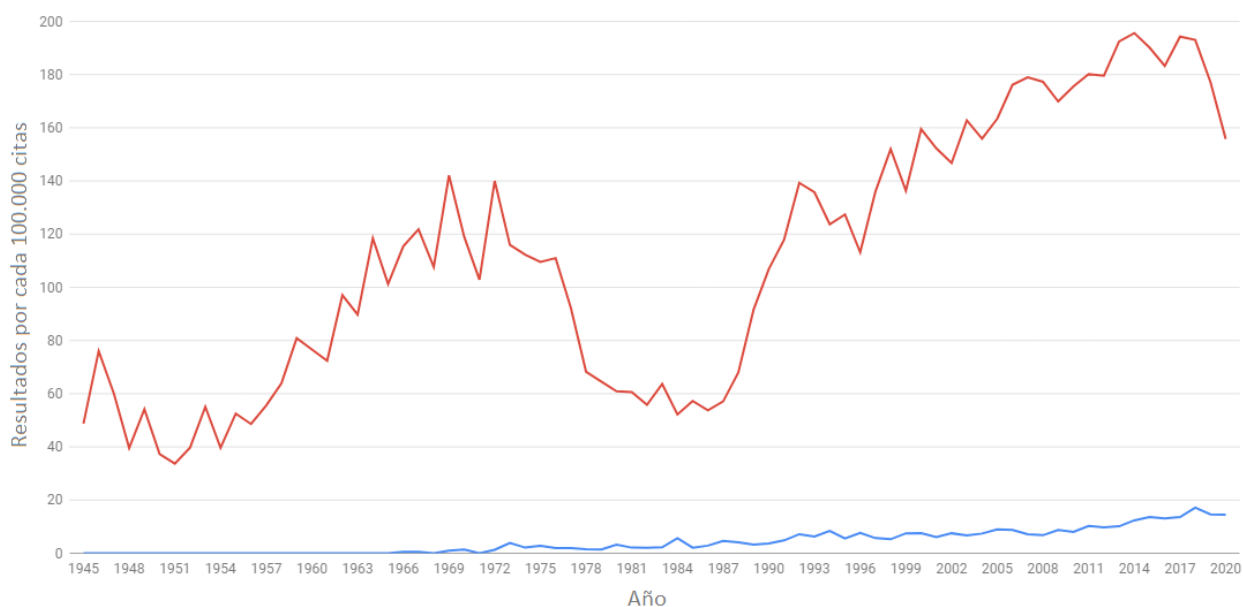


Figura 1. Evolución de la proporción de citas en los artículos publicados anualmente en PubMed con el término “Mycovirus” (azul) o “Plant Virus” (rojo) desde 1945 hasta 2020. Gráfico generado con PubMed by Year (disponible en <https://esperr.github.io/pubmed-by-year/>).

En cuanto a la terminología, los micovirus también pueden llamarse virus micófagos, pero es importante mencionar que no deben confundirse los micovirus con los micobacteriófagos, estos últimos encuentran en las bacterias del género *Mycobacterium* su huésped (Hatfull, 2018). De todas formas, el término micófago debe evitarse, ya que solo una minoría de los micovirus son capaces de producir la lisis de sus huéspedes (Bozarth, 1972).

1.2. ECOLOGÍA DE LOS MICOVIRUS

En lo que respecta a la ecología de los micovirus, están extendidos en los cinco phyla clásicos de los hongos, es decir, ascomicetos, basidiomicetos, quitridios, zigomicetos y glomeromicetos. En los glomeromicetos, hongos simbiotes obligados de plantas que forman micorrizas arbusculares, la existencia de micovirus se demostró hace tan solo cinco años (Ezawa et al., 2015). Esta distribución sugiere que la existencia de micovirus se remonta a estados muy tempranos de la historia evolutiva de sus huéspedes (Ghabrial, 1998).

Los micovirus, por lo general, carecen de una fase extracelular en su ciclo viral (Ghabrial, 1998). Solo se ha registrado que la infección de viriones libres en laboratorio es posible en levaduras y, más recientemente, se ha descubierto que el ascomiceto *Sclerotinia sclerotiorum* puede infectarse artificialmente con un micovirus de ADN, sugiriendo su uso como un fungicida natural (Yu et al., 2013). No se descarta que en estos casos concretos exista en la naturaleza una fase extracelular, si bien no se ha demostrado. La dificultad para infectar hongos *in vitro* con micovirus de RNA se considera un factor que ha limitado los avances de la micovirología (Pearson et al., 2009). Actualmente hay estrategias para superar esta barrera como la transfección de protoplastos con cDNA viral o el uso de vectores de transformación con el cDNA viral clonado, aunque requieren un conocimiento avanzado de la biología molecular tanto del virus como del huésped (Nuss, 2011).

De todas formas, esto no ha evitado que se hayan extendido ampliamente, favorecidos por los frecuentes eventos de comunicación citoplasmática o plasmogamia que se producen en el ciclo vital de los hongos. La incidencia varía de un 10 a un 100% de individuos infectados dependiendo de la especie que se esté considerando (Pearson et al., 2009).

Un punto de complejidad en su ecología supone el hecho de que un mismo micelio esté coinfectado por distintos micovirus (Chu et al., 2004; Thapa y Roossinck, 2019). No solo eso, es posible que ciertos micovirus interactúen dentro del huésped, modulando los síntomas el uno del otro o favoreciendo recíprocamente su multiplicación (Romaine y Schlagnhauser, 1995). Por otra parte, no se ha documentado ningún caso de encapsidación heteróloga entre micovirus (Ghabrial, 1994).

Los mecanismos de propagación de los micovirus incluyen la transmisión vertical y horizontal, siempre por vía citoplasmática:

En el primer caso tenemos la formación de esporas sexuales y asexuales. Hay evidencias de que en ambos escenarios se puede transmitir el virus, pero algunos autores apuntan a que las esporas sexuales de los ascomicetos, las ascosporas, son más proclives a generar nuevos individuos libres de virus (Hillman y Suzuki, 2004). Sin embargo, en distintos estudios se ha observado que la transmisión en esporas sexuales, como las ascosporas de *Fusarium graminearum*, puede llegar a ser del 100% (Chu et al., 2004). Además, puede ser importante qué individuo es el que originalmente estaba infectado en la reproducción sexual. La tasa de transmisión del Botrytis virus X de *Botrytis cinerea* es superior cuando el parental femenino, el micelio, es el portador del virus (Tan et al., 2007). Todos estos resultados apuntan a que no se puede tomar una conclusión aplicable para todos los hongos y todos los micovirus (Ghabrial, 1998).

La transmisión horizontal se basa principalmente en la anastomosis hifal. Este fenómeno consiste en la fusión de los citoplasmas de distintos micelios a través de sus hifas, favoreciendo el intercambio de determinantes citoplasmáticos, así como de material genético. La anastomosis es importante en el proceso de formación de micorrizas y suponen en los ecosistemas la comunicación de hongos a lo largo de grandes superficies (Novais et al., 2017). Esta es la vía de transmisión más utilizada por los equipos de investigación en las infecciones artificiales (Suzaki et al., 2005). La propagación del virus por las hifas puede producirse intraespecíficamente, así como entre distintas especies siempre que mantengan una compatibilidad vegetativa, determinada, entre otros, por factores nucleares. Al ser la transmisión estrictamente intracelular, estos grupos de compatibilidad vegetativa limitan el rango infectivo de los micovirus (Pearson et al., 2009). Algunos autores sugieren que la transmisión de micovirus a través de las hifas involucra mecanismos de incompatibilidad entre los genotipos del hongo y del micovirus de forma análoga a lo que ocurre en plantas (Biella et al., 2002). Los hongos cuentan además con un sistema de defensa antiviral basado en silenciamiento por RNA, aunque su eficiencia no es total, entre otros motivos, porque los micovirus podrían codificar supresores del silenciamiento, proteínas para interferir en este mecanismo de defensa (Nuss, 2011).

Estos mecanismos de transmisión aprovechando la elevada plasmogamia que tiene lugar en los hongos y la falta de fases extracelulares hacen que los micovirus estén generalmente involucrados en infecciones persistentes a lo largo de toda la vida de un hongo. A día de hoy, se cree que la mayoría de los micovirus causan infecciones latentes o crípticas, en las que el hongo presenta un fenotipo libre de síntomas (Roossinck, 2015). Aunque la dificultad de infectar hongos con micovirus en laboratorio y la coinfección simultánea y generalizada por varios micovirus distintos también ha dificultado el establecimiento de relaciones entre ciertos micovirus y el fenotipo de sus huéspedes.

Algunos virus han evolucionado para ir más allá de limitarse a que su huésped los retenga de forma avirulenta en su interior y confieren al hongo ventajas selectivas frente a los individuos no infectados (Ghabrial, 1998). Un ejemplo de esto son las conocidas levaduras Killer, infectadas con el virus L-A (Bevan et al., 1973). Fuera de las levaduras, también se ha documentado una producción de toxinas killer inducida por micovirus en especies como *Ustilago maydis* (Park et al., 1996). Asimismo, hay casos en los que hongos patógenos de plantas o animales infectados con micovirus muestran hipervirulencia (Shah et al., 2019; Özkan y Coutts, 2015).

Resulta problemático para abordar la cuestión de los efectos en el huésped a nivel ecológico que los micovirus mejor estudiados sean aquellos con consecuencias económicas directas como el virus causante de la enfermedad de La France en champiñones o algunos micovirus asociados a la hipovirulencia de hongos fitopatógenos como *Cryphonectria parasitica* o *Fusarium spp.* (Bryner et al., 2014; Sharma et al., 2018) cuando estos síntomas, en realidad, no son representativos de los micovirus (Roossinck, 2015; Pearson et al., 2009).

En resumen, el modo de vida generalmente críptico de los micovirus sugiere que los efectos sobre su huésped son resultado de una coevolución en un estado relativamente maduro, donde los virus con efectos deletéreos sobre el huésped son una minoría (Ghabrial, 1998). Esta coadaptación huésped-virus se hace patente en casos como el descrito por Hyder et al. (2013), en el que se demostró que los efectos de un micovirus latente avirulento de *Heterobasidion* podían cambiar y causar efectos detrimentales en ambientes distintos al natural. La posibilidad de infección podría estar determinada por la presencia de nutrientes y ciertos compuestos químicos en el entorno del hongo (Thapa y Roossinck, 2019). El mutualismo virus-hongo podría derivar en ocasiones en una simbiosis a tres vías incluyendo también a las plantas, ya que se ha comprobado que tanto el hongo endófito *Curvularia protuberata* como la monocotiledónea *Dichantheium lanuginosum*, uno de sus huéspedes, son más tolerantes a elevadas temperaturas cuando el hongo está infectado con un micovirus específico (Márquez et al., 2007).

La diversidad de los efectos de los micovirus sobre el fenotipo y el metabolismo de los hongos impulsa a imaginar sus potenciales aplicaciones en la medicina, la industria o el saneamiento y preservación de ecosistemas. Hay que considerar que se conocen muy pocos micovirus en relación a los que se estima que hay (Sahin y Akata, 2018) y los proyectos de identificación de nuevas especies basados en NGS no siempre incluyen un análisis detallado de los síntomas de sus infecciones.

1.3. TAXONOMÍA Y CARACTERÍSTICAS

Los micovirus son una agrupación de virus heterogénea en cuanto a su tipo de genoma se refiere (Tabla 1). La mayoría de micovirus son de dsRNA, por lo que uno de los principales métodos para determinar la presencia de micovirus en hongos ha sido la identificación de RNA bicatenario. Aun así, se estima que alrededor de un 30% de los virus de hongos son virus de RNA monocatenario positivo ((+)ssRNA) (Son et al., 2015), entre los que se encuentra el *Cryphonectria parasitica hypovirus 1* (CHV1). Mucho más minoritarios parece que son los micovirus de RNA monocatenarios negativos ((-)ssRNA) y de DNA monocatenario circular (ssDNA), con solo una especie documentada de cada clase (Liu et al., 2014; Yu et al., 2010). Por último, hay dos familias de retrotransposones LTR o virus de RNA de transcripción reversa con miembros capaces de replicarse en hongos e integrarse en su genoma (Daboussi y Capy, 2003). Es probable que existan micovirus de DNA bicatenario no identificados, ya que se han identificado en protistas cercanos a los hongos (Ghabrial et al., 2015). A pesar de los esfuerzos por clasificar en familias los nuevos micovirus descubiertos, a una parte considerable de los virus, también de dsRNA, no se les ha asignado una familia (Lin et al., 2014; Liu et al., 2019).

Tabla 1. Distribución de las familias de virus para las que se ha descubierto algún micovirus en los distintos grupos de la clasificación de Baltimore.	
*Familia propuesta por Liu et al. (2014), pero no reconocida actualmente por el Comité Internacional de Taxonomía de Virus (ICTV).	
Grupo de la clasificación de Baltimore	Familias
Grupo II: ssDNA	<i>Genomoviridae</i> (Krupovic et al., 2016)
Grupo III: dsRNA	<i>Chrysoviridae</i> <i>Endornaviridae</i>

	<i>Megabirnaviridae</i> <i>Quadriviridae</i> <i>Partitiviridae</i> <i>Reoviridae</i> <i>Totiviridae</i>
Grupo IV: (+)ssRNA	<i>Alphaflexiviridae</i> <i>Barnaviridae</i> <i>Gammaflexiviridae</i> <i>Hypoviridae</i> <i>Narnaviridae</i>
Grupo V: (-)ssRNA	<i>Mycomononegaviridae</i> *
Grupo VI: ssRNA-RT	<i>Metaviridae</i> <i>Pseudoviridae</i>

Tabla 1. Distribución de las familias para las que se ha descubierto algún micovirus en los distintos grupos de la clasificación de Baltimore.

Como cabe esperar por su distribución a lo largo de distintas clases de virus, los micovirus son bastante variables en cuanto a las características de los viriones y la organización del genoma. Una característica estructural común es el hecho de tener partículas virales desprovistas de envoltura lipídica, con la excepción de los micovirus de (-)ssRNA (Ghabrial et al., 2015). También existen micovirus que ni siquiera están encapsidados, como los hipovirus (Ghabrial y Suzuki, 2008). Dentro de los micovirus de dsRNA hay virus que tampoco forman auténticos viriones encapsidados, como los endornavirus, que se pueden considerar virus de pseudo-dsRNA (Zhang et al., 2014). Dentro de los virus de dsRNA no encapsidados se ha podido demostrar que el dsRNA purificado del *Aspergillus fumigatus* tetramycovirus-1 es infeccioso, al contrario de lo observado para el resto de micovirus de dsRNA (Kanhayuwa et al., 2015).

Considerando solo los virus de dsRNA, si bien se puede decir que los viriones son generalmente rígidos y de morfología icosaédrica, la distribución del genoma dentro de estas partículas virales es variable. Los micovirus de la familia *Totiviridae*, una de las más estudiadas, son muy sencillos en este aspecto, ya que tienen un genoma no segmentado. Sin embargo, en familias como *Chrysoviridae*, *Partitiviridae* y *Reoviridae*, el genoma puede estar dividido en hasta 12 fragmentos, codificando un mayor número de proteínas (Liu et al., 2012). Además, los virus con genomas no monopartitos pueden tener los segmentos encapsidados individualmente, como los chrysovirus o los partitivirus, lo que hace que la unidad infectiva mínima esté compuesta de más de un virión (Ghabrial et al., 2018; Nibert et al., 2013).

Muchas de las proteínas codificadas en el genoma de los micovirus aún son de función desconocida, sobre todo aquellas ORFs que no están directamente relacionadas con la fisiología básica del ciclo viral. Recientemente, se ha identificado una ORF asociada a la hipovirulencia del hongo fitopatógeno *S. sclerotiorum* en uno de sus micovirus. La transformación del hongo con este gen demostró que era capaz de reducir su virulencia y modular su metabolismo de proteínas y lípidos (Gao et al., 2020). La caracterización de más proteínas de micovirus es clave para entender los mecanismos moleculares de interacción con el huésped, así como la historia evolutiva de los micovirus.

El genoma de los micovirus de dsRNA, por lo general, tiene unos componentes mínimos, concretamente, una RNA-polimerasa dependiente de RNA (RdRP) o replicasa y al menos una proteína de cubierta (CP). En los micototivirus, como el virus L-A de *S. cerevisiae*, ambas están codificadas en la misma molécula de RNA, ya que son virus monopartitos. El gen de la replicasa suele estar en 3' y se expresa en una proteína de fusión CP-RdRP minoritaria en relación al número de CP traducidas, quedando la replicasa en el interior de la cápsida (Liu et al., 2012). En otros micovirus de dsRNA bipartitos o multipartitos, la RdRP también se empaqueta dentro de los viriones, a pesar de que se exprese por separado.

La cápsida de los virus de dsRNA no tiene solo una función protectora del genoma frente a los enzimas presentes en el citoplasma del huésped. Al permanecer inalterada durante la replicación, constituye un filtro que permite al virus importar nucleótidos para la transcripción de sus ORFs y exportar los transcritos de RNA al citoplasma para su traducción (Castón et al., 1997). De esta forma se limita la exposición del dsRNA a los componentes citoplasmáticos celulares, disminuyendo la respuesta eucariota frente al dsRNA (Luque et al., 2016). Sus cápsidas, por lo general, son $T = 1$, formadas por 60 subunidades diméricas (Luque et al., 2018).

Las funciones de CP podrían ser más complejas que la formación de estructuras pasivas, ya que se ha observado en el virus L-A de *S. cerevisiae* la presencia de dominios externos en la cápsida capaces de procesar el RNA mensajero celular eliminando su caperuza 5' (Naitow et al., 2002). Otros dominios similares se han observado en las CP de micovirus de otras familias, como *Chrysoviridae* y *Quadriviridae*, aunque no se les ha podido confirmar una función catalítica. Más estudios en este sentido podrían aclarar si el movimiento por el citoplasma es pasivo o hay una interacción con el citoesqueleto (Mata et al., 2017). Estos datos sugieren que CP es una proteína con múltiples funciones que ha ido adquiriendo a partir de pequeñas inserciones en puntos específicos de la secuencia a modo de “hotspots” que no perturben la función estructural (Luque et al., 2018). La adquisición de nuevas funciones con un aumento mínimo del genoma como las anteriores presentan una estrategia evolutiva más favorable a generar nuevas ORFs en virus de RNA (Belshaw et al., 2008).

Los virus de RNA se caracterizan por formar cuasiespecies, es decir, nubes de secuencias mutantes en torno a una secuencia maestra común. Esto se debe a la elevada tasa de mutación que tiene la replicación de un genoma basado en RNA, del orden de 10^{-3} a 10^{-5} mutaciones por cada posición en cada ciclo de replicación (Domingo y Holland, 1997). La variabilidad de estas poblaciones depende de la interacción con el huésped y del entorno, aparte de características de la infección como la tasa de multiplicación y el tamaño de la cuasiespecie. Se ha visto que estos parámetros son más determinantes para el tamaño de la nube de secuencias que la posición taxonómica del virus (Schneider y Roossinck, 2001). Esta dinámica poblacional aumenta la posibilidad de supervivencia de la cuasiespecie ante una situación de cuello de botella siempre que haya alguna secuencia que esté mejor adaptada. En micovirus esta situación puede producirse cuando cambia la especie que hace de huésped o cuando se produce algún otro cambio importante en el entorno. Sin embargo, este tipo de poblaciones suponen el coste adicional de formar un mayor número de partículas defectuosas, que son eliminadas por selección natural, pero las mutaciones deletéreas con efectos ligeros pueden ir acumulándose y suponer un peligro para la supervivencia de la especie (Elena et al., 2000).

Aparte de las mutaciones puntuales que se dan en la replicación de los virus de RNA, existen otros cambios genómicos más inusuales como la recombinación con otros micovirus y la reorganización del genoma propio, que representan una fuente de variación adicional e importantes saltos macroevolutivos. La CP de los chrysovirus es el resultado de la duplicación de una CP ancestral común a los micototivirus, lo que hace que su cápsida esté formada por 60 monómeros en vez de 60 dímeros (Luque et al., 2010).

1.4. ORIGEN EVOLUTIVO DE LOS MICOVIRUS

A día de hoy, resolver el origen evolutivo de los virus de hongos sigue siendo uno de los principales retos de la micovirología. Si bien se siguen descubriendo micovirus que aumentan la diversidad de este grupo heterogéneo, no hay duda de que los que mayor éxito han tenido y más se han extendido son aquellos de dsRNA. Los virus de dsRNA encuentran sus huéspedes a lo largo de todo tipo de organismos eucariotas, desde protozoos a vertebrados y plantas. Fuera de los eucariotas, también existen virus de dsRNA que pueden infectar bacterias (Yang et al., 2016).

Se cree que los virus de RNA podrían ser descendientes directos de los organismos primitivos que habitaban el mundo del RNA. En la hipótesis del mundo del RNA, la función de la proteína RdRP tiene una gran importancia, puesto que el RNA debía replicarse a partir de RNA. En organismos

celulares no se encuentra una proteína homóloga a RdRP, pero sí está conservada en todos los virus de RNA, por lo que es muy útil para reconstruir la historia evolutiva de los virus de RNA (Koonin et al., 2015).

El origen de los virus de dsRNA es objeto de discusión entre distintos autores. Por una parte, se ha demostrado que los virus de DNA y (+)ssRNA pueden producir dsRNA y que de esta forma podrían haberse formado distintos linajes (Weber et al., 2006). El análisis filogenético de la proteína RdRP sugiere que la replicasa de los virus de dsRNA está relacionada con la de los virus de (+)ssRNA (Koonin et al., 1989) y que podría haber habido varios eventos de formación de virus de dsRNA a partir de distintas familias de virus de (+)ssRNA, por lo que no serían monofiléticos (Ghabrial, 1998). Por otra parte, el estudio de la estructura terciaria y cuaternaria de CP mostró que los virus de dsRNA, excepto la familia *Birnaviridae* y algunos virus relacionados, muestran características muy similares, sugiriendo un origen común (Abrescia et al., 2012). Cabe destacar que el análisis estructural de la cápsida no es posible para la familia *Endornaviridae*, que se ha visto que está estrechamente emparentada con la superfamilia de los alfavirus, virus de (+)ssRNA (Koonin y Dolja, 2014).

Por esto, los micovirus de dsRNA encapsidados podrían tener un origen común, ya que no parecen estar relacionados con la familia *Birnaviridae*, tampoco en el análisis de RdRP (Koonin et al., 2015). Luque et al. (2014) afirmaron que los micovirus de dsRNA encapsidados pertenecen a un linaje único en base al estudio estructural de CP y propusieron la relación de estos virus con la familia *Cystoviridae* de bacteriófagos de dsRNA, que tienen también tienen cápsidas icosaédricas y CP con dominios con orientaciones y posiciones similares, aunque con ningún parecido apreciable a nivel de secuencia. Esta relación se extendería a lo largo de las familias *Totiviridae*, *Partitiviridae*, *Chrysoviridae* y *Reoviridae*. Otros estudios confirman la relación entre la estructura de la CP de los reovirus y los cystovirus (El Omari et al., 2013). Considerando que la familia *Quadriviridae* está relacionada con los micototivirus (Lin et al., 2012) y la familia *Megabirnaviridae* con los chrysovirus (Sato et al., 2019), todas estas familias de micovirus podrían tener un origen común según estas conclusiones. No obstante, las similitudes que se observan actualmente podrían ser el resultado de una evolución convergente, ya que las posibilidades de plegamiento de nanoestructuras como las cápsidas son limitadas (Mata et al., 2017).

En el modelo propuesto por Koonin et al. (2015) integrando la información de la replicasa y los datos estructurales de la cápsida, los reovirus serían descendientes directos de los cystovirus, mientras que los otros micovirus de dsRNA relacionados habrían adquirido solo la CP de estos bacteriófagos icosaédricos y la RdRP de virus de (+)-ssRNA de la superfamilia de virus picorna-like. Estos resultados ponen de manifiesto la importancia de combinar los datos estructurales y a nivel de secuencia de más de un locus para obtener conclusiones filogenéticas fiables.

En resumen, es probable que los micovirus de dsRNA provengan de pocos eventos de generación a partir de otros virus, seguramente (+)ssRNA. Además, algunos de estos debieron ocurrir antes de importantes diferenciaciones de los organismos celulares. Los reovirus se encuentran infectando todo tipo de eucariotas (Forrest y Dermody, 2003) y al estar emparentados con los cystovirus es posible que este linaje de virus de dsRNA tenga origen antes de la división entre eucariotas y procariotas. La aparición de los últimos antepasados comunes de dsRNA de los totivirus y las otras familias de micovirus de dsRNA relacionadas en el linaje de RdRP pudo ser posterior, ya que se encuentran infectando a eucariotas simples como los protozoos aparte de hongos y las plantas, pero no procariotas (Callanan et al., 2018). La posterior diferenciación en familias se habría producido por distintos fenómenos macroevolutivos, como las reorganizaciones intragenómicas como el caso comentado anteriormente de la duplicación de CP en los chrysovirus. Por ejemplo, se cree que los partitivirus se podrían haber originado de antepasados monopartitos de los actuales totivirus donde RdRP estaba codificada en una ORF independiente de CP, como en el micototivirus Hv190SV (Soldevila y Ghabrial, 2000). Este genoma se habría dividido en las dos moléculas monocistrónicas que se pueden observar en los partitivirus (Ghabrial, 1998). También se sospecha que ha habido eventos de transferencia horizontal de genes con plantas y hongos que podrían haber supuesto diferenciaciones importantes, aunque aún no se ha demostrado (Ghabrial y Suzuki, 2009).

Por otra parte, es posible que haya habido eventos de generación de micovirus de dsRNA más recientes a partir de virus de hongos de (+)ssRNA, ya que el micovirus de dsRNA no clasificado *Fusarium Graminearum virus-DK21* presenta similitudes muy significativas en los genes de la replicasa, la helicasa y la proteasa con los hypovirus (Kwon et al., 2009).

Se sabe que es difícil aproximar la historia evolutiva de la virosfera de RNA con árboles filogenéticos basados en la secuencia de RdRP. Aunque esta proteína esté suficientemente conservada para clasificar familias cercanas, la gran divergencia a nivel de secuencia no permite afirmar que haya un origen común de RdRP a lo largo de todos los virus de RNA (Zanotto et al., 1996). Además, se han documentado casos de recombinación homóloga entre RdRP de micovirus con virus de plantas relativamente distantes (Howitt et al., 2006). La recombinación, no solo homóloga, entre distintos virus y dentro de un mismo virus son especialmente probables en las coinfecciones que producen distintos micovirus en un mismo huésped (Onodera et al., 1993). Se sabe que algunos micovirus son capaces de inducir reorganizaciones genómicas en otros virus por medio de sus proteínas, es lo que ocurre con el factor supresor del silenciamiento de RNA p29 del hypovirus CHV1 (Sun et al., 2008).

A día de hoy, se sabe que la evolución de los micovirus no se corresponde estrictamente con la de los hongos y que se han producido saltos de huésped con las plantas, por lo que la hipótesis clásica de una coevolución continua de los micovirus con los hongos no explica toda la variación actual. Es más, hace poco se ha comprobado que hay micovirus de dsRNA extraídos directamente de hongos endófitos de la familia *Totiviridae* y *Partitiviridae* capaces de replicarse en plantas e infectarlas (Nerva et al., 2017). Anteriormente, ya se había observado que los micopartitivirus estaban mucho más relacionados filogenéticamente con los partitivirus que infectan a plantas que con el resto de micovirus. Los géneros *Alphapartitivirus* y *Betapartitivirus* están formados cada uno por dos subclados, en los que uno está compuesto de micovirus y el otro de virus de plantas (Nibert et al., 2014). Aparte de los micovirus de dsRNA, la familia de micovirus *Hypoviridae* muestra la mayor homología de secuencia aminoacídica de RdRP con el género *Bymovirus* de la familia *Potyviridae* de virus patógenos de plantas (Fauquet et al., 2005).

La interacción ecológica que se produce entre plantas, hongos fitopatógenos o endosimbiontes y los micovirus que se encuentran infectando de forma persistente a estos últimos supone un ambiente propicio para que se produzcan estos saltos de huésped, ya que las células vegetales quedan expuestas a los micovirus que lleven los hongos endófitos (Nerva et al., 2017). La frecuencia de estas adaptaciones a nuevos huéspedes, que podría parecer baja por la ausencia de una fase extracelular (Pearson et al., 2009), es desconocida, pero los saltos de huésped de planta a hongo y viceversa podrían no ser los únicos. Liu et al. (2009) identificaron en *S. sclerotiorum* un virus de RNA con similitudes significativas en la replicasa con el virus de la hepatitis E, sugiriendo la posibilidad de saltos de huésped de virus entre hongos y animales o sus antepasados.

Por último, la actual clasificación taxonómica no consigue incluir en familias muchas nuevas especies de micovirus que se identifican en proyectos de secuenciación masiva y quedan sin clasificar. Estos estudios suelen realizar un análisis filogenético de la RdRP y, si la homología con una familia conocida no es lo suficientemente elevada, los nuevos virus quedan sin clasificar. En los últimos años, la cantidad de micovirus de dsRNA no clasificados ha aumentado de unos pocos a varias decenas y es muy probable que en los próximos años se creen nuevas familias para ellos. Aun así, poco se ha profundizado en la CP de estos nuevos grupos, ya que es mucho más variable a nivel de secuencia y, por tanto, más problemática al realizar filogenias. Por ello, el estudio estructural podría aportar claves importantes.

El objetivo de este estudio es evaluar la cofilogenia de CP y RdRP en distintas familias de micovirus de dsRNA e intentar identificar la causa de la divergencia en CP analizando su secuencia aminoacídica y simulando su estructura, con la hipótesis de que la recombinación con el genoma celular del hongo ha sido un proceso importante en la historia evolutiva de la CP de al menos un grupo de virus no clasificados. El resultado de esta transferencia horizontal de genes implicaría la existencia

de genomas de micovirus de dsRNA con un genoma en forma de mosaico, con una parte propiamente viral y una parte eucariota adquirida del genoma del huésped.

2. MATERIALES Y MÉTODOS

2.1. BÚSQUEDAS BLAST

A partir de un conjunto inicial de secuencias aminoacídicas de RdRP y CP de especies de distintos grupos de micovirus se realizó una búsqueda con los algoritmos BLASTp y PSI-BLAST, respectivamente, en la base de datos de secuencias de proteínas no redundantes (nr) con los valores por defecto del servidor del NCBI (disponible en <https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Esta búsqueda en la base de datos del NCBI (Sayers et al., 2020) va enfocada igualmente a incrementar la variabilidad intragrupal y a encontrar nuevos grupos relacionados para intentar aumentar la señal filogenética del conjunto. Un total de 72 especies coincidentes de micovirus se reunieron para el análisis de RdRP y CP (Tabla S1). Los conjuntos incluyen secuencias de micovirus de las familias *Partitiviridae*, *Chrysoviridae* y micovirus de dsRNA no clasificados.

Además de la búsqueda de secuencias virales, se usó PSI-BLAST con al menos tres iteraciones y PHI-BLAST con el objetivo de encontrar regiones comunes conservadas en proteínas del huésped similares a la CP de los distintos grupos de micovirus. Así se recogieron 40 secuencias de hongos ortólogos entre sí (Tabla S2).

2.2. ALINEAMIENTO DE SECUENCIAS

El alineamiento múltiple del conjunto de secuencias de RdRP se llevó a cabo con el algoritmo MUSCLE (Edgar, 2004) integrado en el software MEGA versión X (Kumar et al., 2018). En el caso de CP, el alineamiento se realizó con distintos algoritmos para probar que las separaciones en los análisis filogenéticos son consistentes independientemente del método que se use y para encontrar aquel que permita construir un árbol filogenético con la mayor fiabilidad posible. Aparte de MUSCLE en MEGA X, se realizaron alineamientos múltiples de CP con T-Coffee (Notredame et al., 2000) versión 11.00.d625267 y el algoritmo L-INS-i de MAFFT versión 7.467 (Katoh y Standley, 2013), adecuado para conjuntos de menos de 200 secuencias, en sus servidores correspondientes (<http://tcoffee.crg.cat/>; Di Tommaso et al., 2011 y <https://mafft.cbrc.jp/alignment/server/>; Katoh et al., 2018). Los distintos algoritmos fueron empleados con los valores predeterminados de sus respectivas plataformas.

Adicionalmente, se utilizó el programa Gblocks versión 0.91b (disponible en http://molevol.cmima.csic.es/castresana/Gblocks_server.html) para reducir los alineamientos a bloques conservados y así poder visualizar mejor parecidos remotos a nivel de secuencia (Castresana, 2000).

2.3. ANÁLISIS FILOGENÉTICO

2.3.1. Construcción de árboles filogenéticos

Para cada alineamiento múltiple se evaluó por máxima verosimilitud con MEGA X cuál era el modelo de sustitución de aminoácidos más apropiado, es decir, aquel con un criterio de información bayesiano (BIC) inferior. Para éste y los posteriores análisis en MEGA X, se ignoraron las posiciones con una cobertura menor al 95%. Los árboles filogenéticos de RdRP y los distintos alineamientos de CP se construyeron en MEGA X por máxima verosimilitud con el modelo LG (Le y Gascuel, 2008) y WAG (Whelan y Goldman, 2001), respectivamente, con 1000 réplicas de bootstrap. En ambos casos, se descartó una distribución uniforme de las tasas de variación en cada posición en favor de una

distribución gamma con 5 categorías, ya que minimizaba el valor de BIC. Los alineamientos múltiples de CP se evaluaron con TREE-PUZZLE versión 5.3.rc16 (Schmidt et al., 2002) para estimar si los datos eran aptos para un análisis por máxima verosimilitud o si, por contra, la información filogenética del conjunto de secuencias era deficiente para la reconstrucción de árboles filogenéticos (Strimmer y Haseler, 1997). Este programa se ejecutó con el modelo de aminoácidos más apropiado para cada alineamiento múltiple predicho con MEGA X y el resto de parámetros por defecto. El conjunto de secuencias definitivo de CP excluye las secuencias de chrysovirus, ya que aumentaba el número de indeterminaciones en los análisis filogenéticos, quedando así reducido a 64 secuencias. La visualización y comparación de los árboles de CP y RdRP se realizó en el programa Dendroscope versión 3.6.3 (Huson y Scornavacca, 2012).

2.3.2. Construcción de redes filogenéticas

Se empleó el programa SplitsTree (Huson, 1998) versión 4.15.1 para construir redes filogenéticas con las 64 secuencias de CP y las 40 secuencias ortólogas de hongos similares a la CP de un grupo de micovirus. Se usaron los algoritmos Neighbor-Net (Bryant y Moulton, 2004), Split Decomposition (Bandelt y Dress, 1992) y Parsimony Splits (Huson et al. 2010) para evitar sesgos por el método utilizado con 1000 réplicas de bootstrap. Adicionalmente, se construyeron redes filogenéticas para el alineamiento de la secuencia concatenada de RdRP y CP, en este orden, de cada especie para observar si ha habido eventos de recombinación entre las especies de micovirus analizadas.

2.4. ANÁLISIS ESTRUCTURAL DE CP Y PROTEÍNAS DEL HUÉSPED

2.4.1. Predicción de estructuras y función

Se simularon las estructuras de CP de cada grupo de virus observado en las filogenias y proteínas del huésped con el servidor de I-TASSER (Zhang, 2009; Roy et al., 2012; disponible en <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) para encontrar similitudes más remotas a las que se pueden observar a nivel de secuencia bajo la premisa de que la estructura proteica tiene un grado de conservación mayor que los residuos que la conforman. Se eligieron secuencias lo suficientemente alejadas dentro de su grupo para evitar artefactos asociados a una subrama no representativa y encontrar un consenso. Se simularon un total de 12 secuencias (Tabla 2). La simulación estructural y funcional de I-TASSER se basa, entre otros aspectos, en la similitud de la estructura secundaria predicha a partir de la secuencia de la proteína problema a entradas del banco de datos de proteínas PDB (disponible en <https://www.rcsb.org/>). A continuación, se usan estas estructuras como plantillas para el plegamiento de la proteína problema con un enfoque iterativo para, finalmente, mostrar aquellos modelos que minimicen la energía de la estructura. Para cada proteína simulada se proponen cinco modelos.

La función de la proteína problema se puede aproximar a partir de los términos GO de las proteínas del PDB que mejor alinean con los plegamientos predichos. Se recogieron los términos GO enriquecidos en cada grupo de proteínas problema como la suma de todas las veces que aparecía un término GO entre las estructuras análogas de cada proteína.

Tabla 2. Especies seleccionadas para la predicción de estructuras proteicas y comparación de las mismas.			
Especie	Última clasificación taxonómica común	Accesión de la proteína en el NCBI	Longitud (aa)
Beauveria bassiana RNA virus 1 (BbRV1)	Micovirus de dsRNA monopartitos no clasificados	AKC57300	315
Alternaria longipes dsRNA virus 1 (AIRV1)		YP_009052468	394
Trichoderma harzianum	Micovirus de dsRNA	YP_009553331	314

bipartite mycovirus 1 (ThBMV1)	bipartitos no clasificados		
Cryphonectria parasitica bipartite mycovirus 1 (CpBMV1)		YP_007985676	331
Botryotinia fuckeliana partitivirus 1 (BfPV1)	<i>Gammapartitivirus</i>	YP_001686790	436
Fusarium solani virus 1 (FsV1)		NP_624351	413
Grifola frondosa partitivirus 1 (GfPV1)	<i>Betapartitivirus</i>	BBG56943	637
Trichoderma citrinoviride partitivirus 1 (TcPV1)		AZT88593	663
Verticillium dahliae chrysovirus 1 (VdCV1)	<i>Alphachrysovirus</i>	YP_009507946	1016
<i>Aspergillus bombycis</i>	<i>Pezizomycotina</i> (subdivisión de <i>Ascomycota</i>)	XP_022391016	270
<i>Capronia epimyces</i>		XP_007732275	253
<i>Erysiphe necator</i>		KHJ31704	285

2.4.2 Alineamiento de estructuras

Los modelos predichos de mejor calidad se sometieron a alineamientos por parejas dentro de cada grupo para determinar una estructura consenso con TM-align (Zhang y Skolnick, 2005; disponible en <https://zhanglab.ccmb.med.umich.edu/TM-align/>). Este programa devuelve una puntuación para el alineamiento, llamada *TM-score*, que informa acerca de la exactitud del alineamiento. Este valor varía de 0 a 1, considerándose valores por debajo de 0.3 alineamientos azarosos y por encima de 0.5 alineamientos entre proteínas con plegamientos similares y más idénticos según el *TM-score* se acerca a 1. Este valor se aplica para cada una de las proteínas alineadas, estando normalizado por la longitud de la secuencia.

Para tener una buena visión de las regiones con mayor identidad una vez se ha visto qué modelos son creíbles para cada grupo, se realizaron alineamientos por parejas y múltiples en el servidor MATRAS (Kawabata y Nishikawa, 2000; Kawabata 2003; disponible en <http://strcomp.protein.osaka-u.ac.jp/matras/>).

3. RESULTADOS

3.1. BÚSQUEDAS BLAST

La búsqueda PSI-BLAST con las CP de cada grupo de micovirus no consiguió encontrar secuencias de otros grupos de micovirus, a pesar de ser un algoritmo apropiado para detectar parentescos remotos. Tampoco dentro de la familia *Partitiviridae* era posible encontrar secuencias ortólogas de CP entre géneros, al contrario de lo que ocurre con RdRP usando el mismo algoritmo. Con la CP de los betapartitivirus de hongos se encuentran en la primera iteración secuencias de CP de betapartitivirus que infectan a plantas con un *E-value* significativo y comparable al de los micovirus de este género. El uso de PHI-BLAST está muy limitado por la variabilidad de secuencia de CP, siendo complicado encontrar motivos perfectamente conservados aún dentro de cada grupo.

De las 72 especies de micovirus recogidas hasta aquí a partir de RdRP y CP, 7 pertenecían a un grupo de micovirus de dsRNA no clasificados de genoma no segmentado encontrados a partir de *Beauveria bassiana* RNA virus 1 (BbRV1), y 15 a un grupo no clasificado de micovirus de dsRNA bipartitos como *Trichoderma harzianum* bipartite mycovirus 1 (ThBMV1). Otras 25 secuencias de gammapartitivirus se juntaron a partir de *Fusarium solani* virus 1 (FsV1), 16 de betapartitivirus desde *Rosellinia necatrix* partitivirus 6 (RnPV6) y se incluyó a *Botrytis cinerea* partitivirus 2 (BcPV2), que se encontró por estar relacionado en RdRP a los betapartitivirus. A pesar de no estar clasificado en un género de la familia *Partitiviridae* en el buscador de taxonomía del NCBI (disponible en <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>), tiene un parentesco mucho más estrecho con especies clasificadas como alfapartitivirus y reconocidas por el ICTV como *Rosellinia necatrix* partitivirus 2 (RnPV2), tal como se indica en su publicación original (Kamaruzzaman et al., 2019). Por último, se incluyeron 8 secuencias de alfachrysovirus a partir de *Verticillium dahliae* chrysovirus 1 (VdCV1). Todas las especies empleadas para el análisis filogenético se pueden consultar en la Tabla S1.

Debido a las diferencias entre las CP de los distintos grupos, la búsqueda de proteínas eucariotas candidatas a tener un origen común con CP debe hacerse por separado. Así se pudieron encontrar en algunos casos proteínas interesantes que se expondrán a continuación.

En primer lugar, la búsqueda PSI-BLAST a partir de la CP de BbRV1 mostraba en la primera iteración, lo que es análogo a una búsqueda BLASTp, una secuencia de una proteína hipotética de la levadura *Lipomyces starkeyi* (ODQ71866) con un $E\text{-value} = 10^{-13}$, significativo, aunque lejano al $5 \cdot 10^{-40}$ que se observa con la CP de *Trichoderma harzianum* mycovirus 1 (AYU71188), el último resultado de un virus en la lista. Esta secuencia proteica de *L. starkeyi* está predicha a partir de un *scaffold* no ubicado en un cromosoma (KV454296) y secuenciado en un proyecto de WGS (Riley et al., 2016).

El hecho de que no se encuentre ninguna secuencia de hongo relacionada al realizar una búsqueda BLASTp con esta proteína y que tenga una homología clara con este grupo de micovirus, hasta 3e-23 con *Alternaria longipes* dsRNA virus 1 (AIRV1), hace pensar que, o bien se ha secuenciado un virus que estaba infectando a ese aislado, o bien ha habido una integración del gen o parte del gen de CP de un virus de este tipo en el genoma de esta cepa. Si este *scaffold* representa verdaderamente secuencias de una sola molécula, esta última hipótesis parece más probable, ya que en otras partes de este *scaffold* hay similitud con otros hongos. Además, el hecho de que la ORF dé lugar a una proteína truncada en la región más conservada si la comparamos con estas CP (Figura 2) dificulta la opción de que se haya secuenciado directamente al virus, aparte de lo improbable que es dada la naturaleza de RNA de estos virus y el método de secuenciación, de DNA. En todo caso, no se trataría del fenómeno de transferencia del huésped al virus que se quiere encontrar.

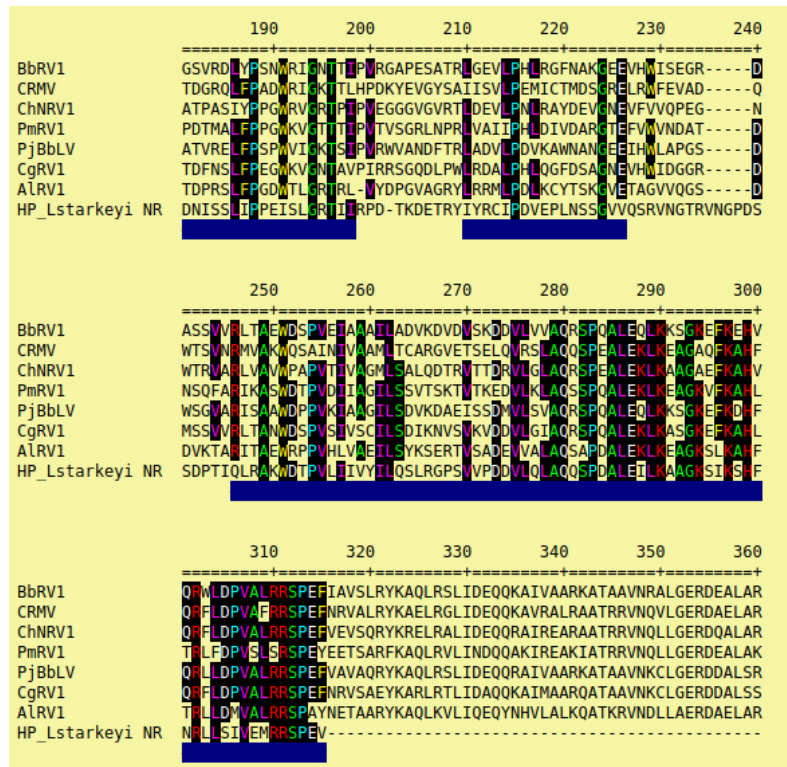


Figura 2. Extracto del alineamiento con MUSCLE de las secuencias de CP de los siete micovirus de dsRNA monopartitos no clasificados incluidos en el estudio con la proteína hipotética de *L. starkeyi* encontrada en la búsqueda PSI-BLAST visualizado en Gblocks con los parámetros por defecto para mostrar los bloques mejor conservados, marcados con rectángulos azules. Los nombres completos de los virus pueden consultarse en el anexo (Tabla S1). La secuencia de la proteína predicha de *L. starkeyi* es homóloga a estas CP, si bien hay un codón de parada prematuro que acorta la proteína por debajo de los 200 residuos. Al saltar de genoma, esta proteína cambia por completo su presión de selección y las mutaciones que son improbables en virus porque alterarían gravemente su función, pueden ocurrir en el genoma del huésped. De todas formas, no hay evidencia de que se exprese tal proteína truncada en la cepa de *L. starkeyi* NRRL Y-11557, por lo que la mutación podría ocurrir como en una región no codificante cualquiera. La elevada homología, así como la ausencia de secuencias genómicas similares en hongos al realizar una búsqueda tBLASTn con esta proteína hipotética, lleva a pensar que la integración, en todo caso, ha sido reciente.

En las siguientes iteraciones se encontraron numerosas secuencias de peroxirredoxinas bacterianas sin ninguna relación biológica aparente. En la segunda iteración se encontró una secuencia de una proteína hipotética del hongo *Cyphellophora europaea* (XP_008711212) con un *E-value* = 10^{-3} similar a otras quinasas de hongos. El parecido, sin embargo, se reduce a muy pocas posiciones y el tamaño de esta proteína es de alrededor de 900 aminoácidos, comparado con los poco más de 200 residuos de estas CP. Además, estas posiciones no están conservadas con las proteínas quinasa ortólogas, lo que hace aún más difícil imaginar que exista una relación real con estas CP.

En segundo lugar, la búsqueda PSI-BLAST con la CP de ThBMV1 rescataba en la primera iteración numerosas proteínas sin anotación funcional de hongo con un *E-value* tanto o más significativo que el de otros virus de este grupo de micovirus de dsRNA bisegmentados no clasificados o ThBMV1-like. De las 15 especies de este grupo, todas mostraban esta relación con estas proteínas de hongo ortólogas entre sí y el virus con una CP que minimizaba el *E-value* con estas proteínas de hongo fue *Cryphonectria parasitica* bipartite mycovirus 1 (CpBMV1), con un valor de $6 \cdot 10^{-82}$ con la proteína hipotética de *Ceratocystis platani* (KKF94277), más significativo que el valor que presentan otros virus con CpBMV1, siendo el más bajo $2 \cdot 10^{-65}$ entre CP de CpBMV1 y ThBMV1 (Tabla 3). El tamaño de esta proteína es variable, desde 165 residuos en *Blumeria graminis* f. sp. *hordei* hasta más de 400

en *Monosporascus* sp. La longitud de las CP de este grupo de micovirus se encuentra entre los 300 y los 400 residuos.

Tabla 3. Primeros diez resultados al realizar una búsqueda BLASTp con la secuencia aminoacídica de CP de CpBMV1. Las proteínas de hongos relacionadas con este grupo de micovirus están anotadas como proteínas hipotéticas.

Nombre del organismo	Cobertura (%)	<i>E-value</i>	Identidad (%)	Longitud (aa)	Accesión en GenBank
CpBMV1	100	0.0	100.00	331	YP_007985676
<i>Ceratocystis platani</i>	69	$6 \cdot 10^{-82}$	52.59	233	KKF94277
<i>Colletotrichum graminicola</i> M.1001	74	10^{-74}	54.88	277	XP_008094161
<i>Ceratocystis fimbriata</i> CBS 114	57	10^{-74}	57.14	199	PHH49570
<i>Monosporascus</i> sp. CRB-8-3	85	10^{-73}	42.55	415	RYP54417
<i>Golovinomyces cichoracearum</i>	80	10^{-69}	40.98	299	RKF65151
ThBMV1	94	$2 \cdot 10^{-65}$	40.63	314	YP_009553331
PvABMV1	98	$6 \cdot 10^{-64}$	40.37	308	QGZ98416
<i>Erysiphe necator</i>	82	$6 \cdot 10^{-60}$	39.71	285	KHJ31704
CthTV	82	$8 \cdot 10^{-53}$	39.27	331	YP_001976145

A pesar de que la mayoría de genomas de hongo con esta proteína estén en estado de borrador y la gran mayoría de proteínas sean de función desconocida, el ortólogo de *A. niger* (TPR01093.1) está anotado como proteína con un dominio de unión a DNA/RNA. Se seleccionaron 40 de estas proteínas hipotéticas de hongo ortólogas, que pueden consultarse en el anexo, para incluirlas en una red con las CP de este grupo de micovirus y establecer los parientes eucariotas más cercanos, ya que esta homología tan acusada hace que este gen sea un candidato muy interesante a haberse transferido al genoma viral en algún momento relativamente reciente. Las especies incluidas pertenecen, mayoritariamente, a la división de los ascomicetos, aunque también hay una pequeña parte de hongos de otros linajes, como *Smittium simulii* y *Tilletia laevis*, cuyas proteínas hipotéticas muestran un parecido menor a la CP de este grupo de micovirus. Pero, en todo caso, esta proteína parece estar relativamente conservada a lo largo de estas divisiones de hongos. Se simularon las estructuras de tres de estas proteínas, todas de ascomicetos por su mayor parecido a la proteína viral, para establecer el parentesco con la CP desde otro punto de vista e intentar resolver cuál podría ser la función de esta proteína en el hongo.

En tercer lugar, la búsqueda a partir de la CP del betapartitivirus RnPV6 (YP_009177605) encontró solo secuencias virales en la primera iteración de PSI-BLAST. Sin embargo, en la segunda se encuentran secuencias de insectos y plantas con un *E-value* significativo. La única secuencia de plantas, de *Trifolium medium*, está anotada como proteína de la cubierta (MCI33606) y las únicas proteínas similares a esta son CP de betapartitivirus, presentando una identidad aún mayor con betapartitivirus que infectan a plantas. De insectos, las secuencias encontradas de *Aphis craccivora* (KAF0690353) y *Apolygus lucorum* (KAE9436552), ortólogas entre sí, se encuentran en una situación similar, sin más secuencias homólogas eucariotas. Cabe destacar que no se han descrito

betapartitivirus que infecten insectos, pero es posible que se hayan detectado en estas especies de áfidos y chinches de las plantas, respectivamente, que pueden actuar de vectores de virus de plantas. No serían, en definitiva, genes eucariotas que se podrían haber transferido a los virus de dsRNA, ya que no están extendidos en plantas ni en insectos.

Con la CP de los gammapartitivirus, por medio de FsV1, se encontró una sola secuencia no viral a lo largo de las iteraciones de PSI-BLAST. Más concretamente, una proteína no caracterizada de *Fusarium coffeatum* (RBR24337) con un $E\text{-value} = 7 \cdot 10^{-9}$ en la tercera iteración, pero la cobertura es de solo un 13% de la secuencia de CP de FsV1. Al buscar secuencias homólogas a esta proteína se vuelve a llegar a un punto muerto, con solo una secuencia remotamente similar en *F. fasciculatum* (RFN53806). Con $E\text{-values}$ del orden de $6 \cdot 10^{-2}$ se encuentran dos secuencias de proteínas transmembrana transportadoras de aminoácidos de *Talaromyces pinophilus* (KAF3407480) y *Penicillium* sp. 'occitanis' (PCG95955), pero se pierde por completo la similitud con la CP de estos partitivirus. En este caso no parece haber una proteína fúngica candidata a haberse transferido desde el huésped al virus. No parece, pues, que la gran divergencia en la CP entre los géneros *Beta*- y *Gammapartitivirus*, que sí son similares en RdRP, pueda explicarse con alguna proteína del huésped que se haya transferido o, de ser así, ha cambiado tanto respecto a la proteína eucariota con la que guardaría una relación de xenología que no se puede rastrear un candidato.

Por último, a partir del chrysovirus VdCV1, la búsqueda PSI-BLAST dio como resultado, en la primera iteración, otros chrysovirus y virus relacionados de hongos, plantas e insectos, además de una proteína del hongo *Oidium neolycopersici* anotada como proteína de la cubierta (RKF53147), con un $E\text{-value} = 4 \cdot 10^{-32}$, un valor tan significativo por cubrir un 47% de la secuencia de CP de VdCV1, cuyo tamaño es de 1016 aminoácidos. Esta secuencia de *O. neolycopersici* de 549 residuos es, de nuevo, una proteína de traducción conceptual a partir de un *contig* de WGS con un codón de parada prematuro.

Aparte de esta secuencia de hongo aislada, una serie de proteínas hipotéticas de la familia *Brassicaceae*, siendo la más similar la secuencia de *Microthlaspi erraticum*, con una cobertura del 30% de CP de VdCV1, un $E\text{-value} = 3 \cdot 10^{-23}$ en la primera iteración y un tamaño de 499 aminoácidos. La similitud entre esta secuencia y la de *O. neolycopersici* se extiende a lo largo de la práctica totalidad de ambas secuencias. Los aparentes ortólogos en *Brassica oleracea* (VDD09411), *B. napus* (CDY45789), *B. rapa* (VDD10380) y *B. cretica* (KAF3553546), que aparecen en las siguientes iteraciones guardan la relación con la secuencia de CP y también con la proteína de *O. neolycopersici*, aunque la significación disminuya porque la secuencia se reduce a unos 270 aminoácidos. Estos 270 aminoácidos, sea como sea, son homólogos a la parte final de la proteína de *M. erraticum*, luego no sería una disrupción debida a un nuevo codón de parada prematuro, sino que la pérdida es en N-terminal.

Finalmente, cabe destacar que se han descrito chrysovirus que infectan a plantas de esta familia, como *Brassica campestris chrysovirus 1* (BcCV1; YP_009667005), otro alfachrysovirus. De hecho, al realizar una búsqueda PSI-BLAST con la secuencia de *M. erraticum*, las secuencias más relacionadas en la primera iteración eran la CP de BcCV1 con un $E\text{-value} = 10^{-122}$ y la de *Raphanus sativas chrysovirus 1* (RsCV1; YP_009667002), otro chrysovirus de una brasicácea, con $E\text{-value} = 4 \cdot 10^{-112}$. A continuación, en orden de significación, se encontraban las secuencias del género *Brassica* mencionadas anteriormente y una secuencia aún más corta de *Arabis nemorensis* (VVB06284), de 124 residuos, que empieza su alineamiento con *M. erraticum* en el mismo punto que las secuencias del género *Brassica*. Haciendo la misma búsqueda con *O. neolycopersici*, se obtuvo un resultado similar, los chrysovirus más relacionados eran de plantas, aunque por poca diferencia, de un $E\text{-value} = 8 \cdot 10^{-52}$ con BcCV1 y $4 \cdot 10^{-49}$ con RsCV1 a $2 \cdot 10^{-41}$ con *Colletotrichium gloeosporioides chrysovirus 1* (YP_009667014) y muchos otros microchrysovirus cercanos a ese valor.

La interpretación, de nuevo, no debe ser la de un gen que se transfiere horizontalmente desde el huésped al virus, pues esta secuencia apenas está conservada y extendida en eucariotas. Para que esto fuese así, este gen en eucariotas debería ser como mínimo tan antiguo como para haberse transferido al

antepasado común entre chrysovirus de hongos, plantas e insectos. Por lo que se sabe, el origen de la CP de los chrysovirus parece estar más relacionado con la duplicación de una estructura similar a la CP de los totivirus (Luque et al., 2010). Parece más probable, pues, que un transcrito de CP de un chrysovirus se haya retrotranscrito e integrado en el genoma de algún antepasado común de las especies nombradas de la familia *Brassicaceae* al que estaba infectando.

La presencia de varias secuencias homólogas en distintas brasicáceas permite ver si alguna región es menos variable y si conserva suficiente homología con el genoma de los chrysovirus. Si esto ocurre tanto a nivel de aminoácidos como de nucleótidos, podría ser un indicio de que este pseudogen se transcribe a un RNA que interfiere con el mRNA viral de CP, ya que la expresión de una proteína a partir de este inserto parece improbable por las diferencias a nivel de ORF.

Al contrario de lo que se ha pensado durante muchos años, la integración de genes de virus de dsRNA en el genoma nuclear de eucariotas es un fenómeno relativamente extendido y, en algunos casos, pueden llegar a expresarse y tener una función en el huésped (Liu et al., 2010). De todos modos, la existencia de esta secuencia en *O. neolycopersici* es difícil de explicar por su parentesco más estrecho con chrysovirus de plantas.

3.2. ANÁLISIS FILOGENÉTICO

3.2.1. Árboles filogenéticos

El cladograma construido a partir del alineamiento múltiple con MUSCLE de 72 secuencias de RdRP y 1226 posiciones de longitud conseguía separar tres grandes linajes: el de los partitivirus, el de los chrysovirus y el de los micovirus de dsRNA no clasificados monopartitos y bipartitos (Figura 3). El árbol mostrado es el que minimizaba el valor del logaritmo de la verosimilitud, que en este caso fue de -37713.99. Los distintos tipos de partitivirus quedan agrupados bajo un nodo con un valor de bootstrap de 0.84. Los otros linajes principales de RdRP presentan cada uno un último antepasado común más creíble no compartido con el resto de secuencias, con valores de *bootstrap* de 0.9510 para los micovirus no clasificados y 1 para los chrysovirus.

Las secuencias de BcPV2 y CPVRdRP-4 pertenecen muy probablemente a virus del género *Alphapartitivirus* (Kamaruzzaman et al., 2019; Ong et al., 2017) y, a pesar de tener poca información por contar solo con dos secuencias, se agrupan significativamente con los betapartitivirus. La credibilidad del nodo que separa estas dos secuencias y agrupa a BcPV2 con los betapartitivirus en vez de con CPVRdRP-4 es baja y puede deberse a diferencias de secuencia dentro de los alfapartitivirus. Sin embargo, otros estudios filogenéticos con más secuencias de cada género no consiguen establecer una relación entre alfa- y betapartitivirus que los haga monofiléticos respecto al resto de partitivirus (Vainio et al., 2018), por lo que puede ser un sesgo debido a la poca señal filogenética que se tiene con solo dos secuencias. Las diferencias que han separado estos géneros no parecen radicar en la organización genómica de los virus, ya que los tres géneros son predominantemente bipartitos, con algunas excepciones de betapartitivirus con un tercer segmento de dsRNA no codificante como AhV (Oh y Hillman, 1995). También hay casos como el de algunos deltapartitivirus, un género de virus que solo infectan a plantas relacionado con los gammapartitivirus, con tres moléculas de dsRNA monocistrónico (Kumar et al., 2017).

De manera análoga, los micovirus de dsRNA no clasificados monopartitos y bipartitos se agrupan bajo un mismo clado muy robusto con este conjunto de datos, indicando que los virus con genomas bipartitos no provienen de un evento ancestral de formación de micovirus de dsRNA con dos segmentos. Descartar que los partitivirus y los virus ThBMV1-like vienen de un mismo evento de segmentación implica rechazar la posibilidad de que los virus monopartitos no clasificados BbRV1-like se hayan originado a partir de la unión de moléculas de virus bipartitos.

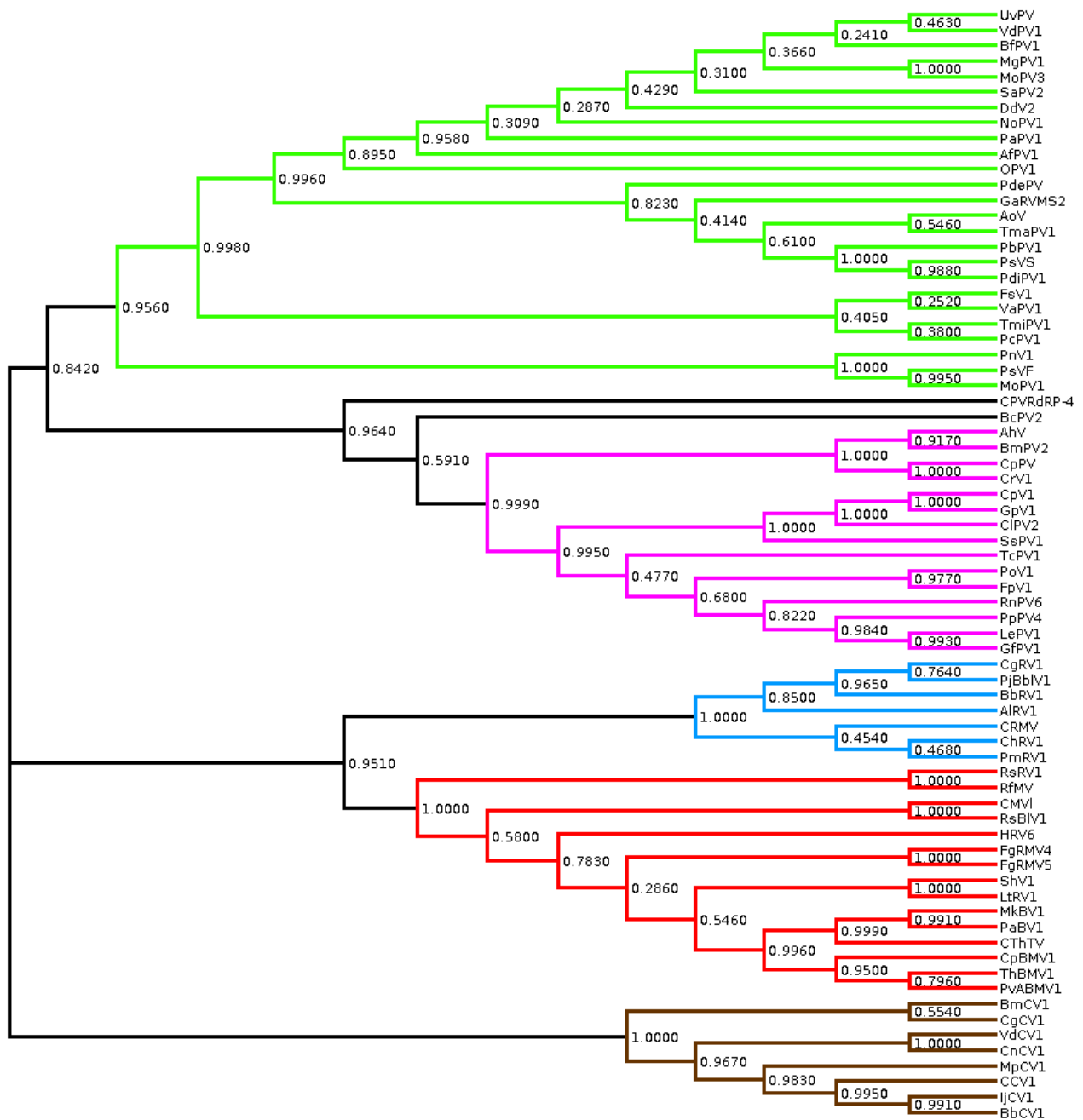


Figura 3. Árbol filogenético obtenido a partir del alineamiento múltiple de las 72 secuencias de RdRp. El valor mostrado en cada nodo se refiere al valor de *bootstrap* tras 1000 réplicas con los datos del alineamiento, donde un valor de 1 o cercano a 1 indica que la bifurcación que produce el nodo se da en la práctica totalidad de las réplicas. La longitud de las ramas está normalizada de forma que todas lleguen a la misma vertical, luego solo la topología puede interpretarse de este árbol. Los diferentes colores muestran los principales grupos que se consideran a lo largo del trabajo: gammaviruses (verde), betaviruses (morado), micovirus de dsRNA monopartitos no clasificados (azul), micovirus de dsRNA bipartitos no clasificados (rojo) y chrysovirus (marrón). Las secuencias en negro pertenecen al género *Alphavirus*.

En definitiva, el hecho de que la encapsidación de las moléculas de dsRNA sea individual, algo muy común en los micovirus de dsRNA, podría minimizar los efectos de este tipo de reorganizaciones genómicas como saltos macroevolutivos si no hay un cambio claro de las regiones codificantes. Una consecuencia de la segmentación es la capacidad de superar el espacio limitado que hay dentro de la

cápsida icosaédrica, aunque no siempre se explota esta supuesta ventaja. Si bien los chrysovirus tienen un genoma de unas 12 kb divididas en cuatro segmentos, notablemente superior a las hasta 6.7 kb que se encuentran en los totivirus, monopartitos, el genoma de los partitivirus bipartitos se mantiene en el rango de las 3-4 kb, incluso menor a virus no segmentados como los totivirus.

Para la CP se realizaron tres árboles distintos, uno para cada alineamiento múltiple de las 64 secuencias, excluyendo los chrysovirus. El primero, realizado con MUSCLE, mostraba los cuatro grupos restantes con valores de *bootstrap* bastante fiables, excepto para los gammapartitivirus, donde el último nodo común a la mayoría de ellos tiene un valor de *bootstrap* de 0.68. Las tres secuencias de gammapartitivirus MoPV1, PsVF y PnV1 quedaban fuera de este clado, sin ninguna relación fiable aparente en CP más cercana con los otros gammapartitivirus que la que comparten con el resto de virus del árbol. Los bajos valores de *bootstrap* de los nodos y la ausencia de una relación entre virus ThBMV1-like y BbRV1-like como ocurre en el árbol de RdRP, motivó el uso de otros algoritmos de alineamiento múltiple para reconstruir árboles filogenéticos.

El árbol resultante del alineamiento de MAFFT (Figura 4), con un tamaño de 1084 posiciones, conservaba las mismas grandes divisiones en grupos y consiguió aumentar el valor de *bootstrap* del nodo principal de los gammapartitivirus comentado anteriormente, pero seguía sin poder relacionarlos con el clado aislado de tres gammapartitivirus. Por otra parte, el último nodo común y único a los virus ThBMV1 y BbRV1 en el árbol de MAFFT tiene un valor de *bootstrap* 0.54. Aun así, no se puede concluir que este parentesco sea real y, en todo caso, la divergencia en CP es muy superior a la observada en RdRP. El valor del logaritmo de la verosimilitud de este árbol fue de -18602.05. Al evaluar la aptitud de los tres alineamientos múltiples mediante TREE-PUZZLE para la reconstrucción de un árbol por máxima verosimilitud, se vio que el alineamiento de MAFFT constituía el conjunto de datos con una mayor señal filogenética, con solo un 5.8% de cuartetos no resueltos y un 5% de cuartetos semiresueltos en el análisis. El árbol a partir del alineamiento de T-Coffee era, de los tres, el que presentaba un mayor porcentaje de indeterminaciones en este análisis y no conseguía agrupar los virus ThBMV1-like y los betapartitivirus bajo nodos con un valor de *bootstrap* superior a 0.3.

Los alineamientos múltiples de CP muestran una variabilidad intragrupal notable y muy poca similitud entre grupos, lo que hace muy difícil asignar un origen común a esta proteína entre grupos, ni siquiera dentro de la familia *Partitiviridae* (Figura 5). En referencia a los alfapartitivirus, representados solo por BcPV2, no se puede concluir que estén estrechamente relacionados en CP a ninguno de estos cuatro grupos. La secuencia de CPVCP-a1 pertenece a un betapartitivirus (Ong et al., 2017) y no es una molécula del mismo organismo que CPVRdRP-4.

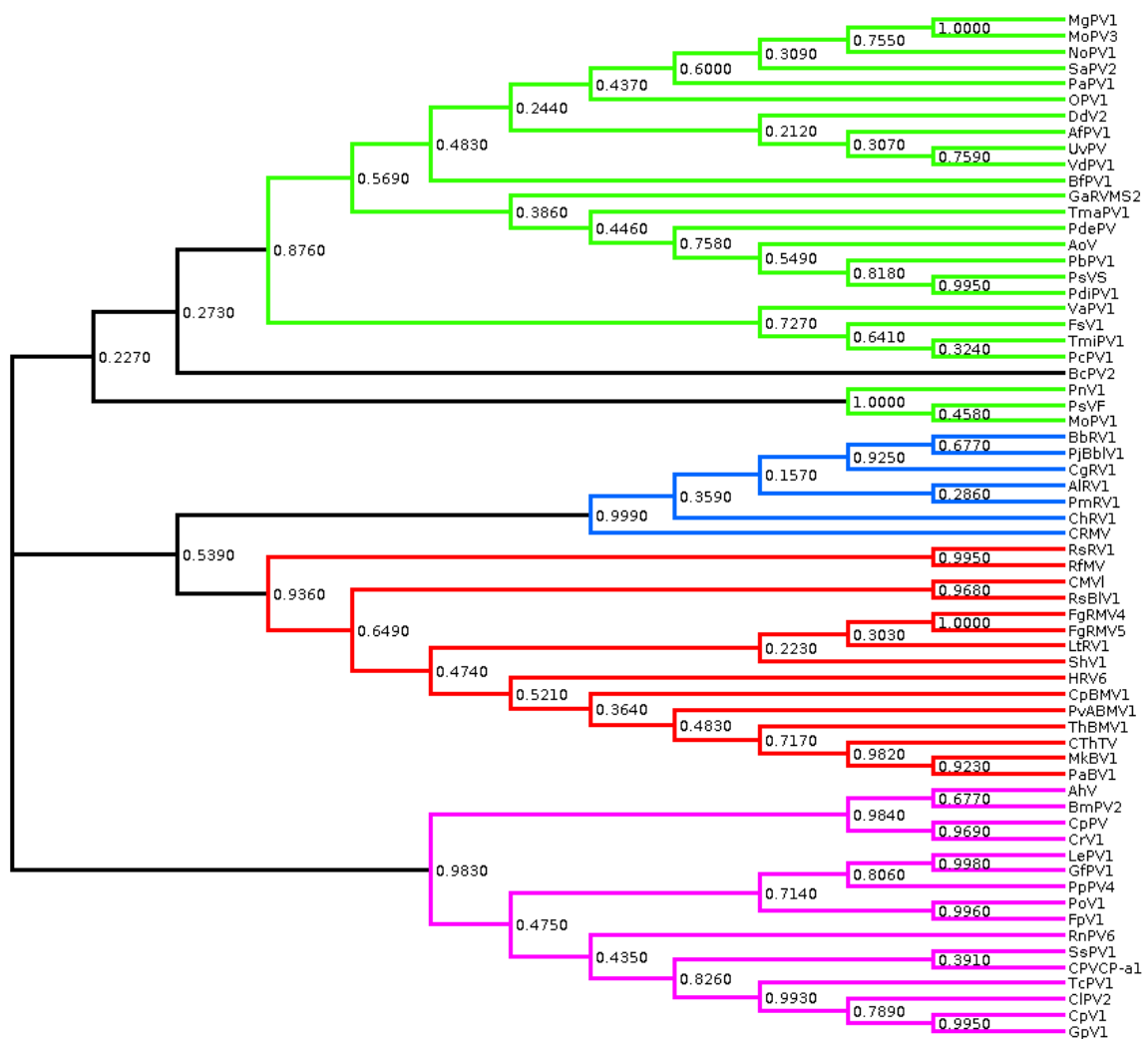


Figura 4. Árbol filogenético obtenido a partir del alineamiento múltiple mediante MAFFT de las 64 secuencias de CP, excluyendo a los chrysovirus mostrando en cada nodo los valores de bootstrap con 1000 réplicas. El árbol muestra solo la topología, la longitud de las ramas no tiene un significado biológico. Los principales grupos que se han considerado en el estudio están coloreados en verde (*Gammaviridae*), azul (micovirus de dsRNA monopartitos no clasificados o BbrV1-like), rojo (micovirus de dsRNA bipartitos no clasificados o ThBMV1-like) y morado (*Betaviridae*). Las secuencias de gammaviridae y betaviridae podrían estar sesgadas hacia las secuencias que se usaron para recuperarlas en la búsqueda BLAST y no ser representativas de sus géneros, ya que se ha secuenciado una cantidad importante de partitivirus y solo se escogió una pequeña parte. De esta manera, los virus PsVF, MoPV1 y PnV1 no tendrían por qué representar una minoría dentro de los gammaviridae, solo una minoría de las secuencias elegidas, pudiendo existir otros gammaviridae que expliquen mejor la relación entre CP de estos virus y el resto de su género. Sin embargo, sí se puede concluir que forman un subclado lejano al resto de gammaviridae en el árbol de RdRP.

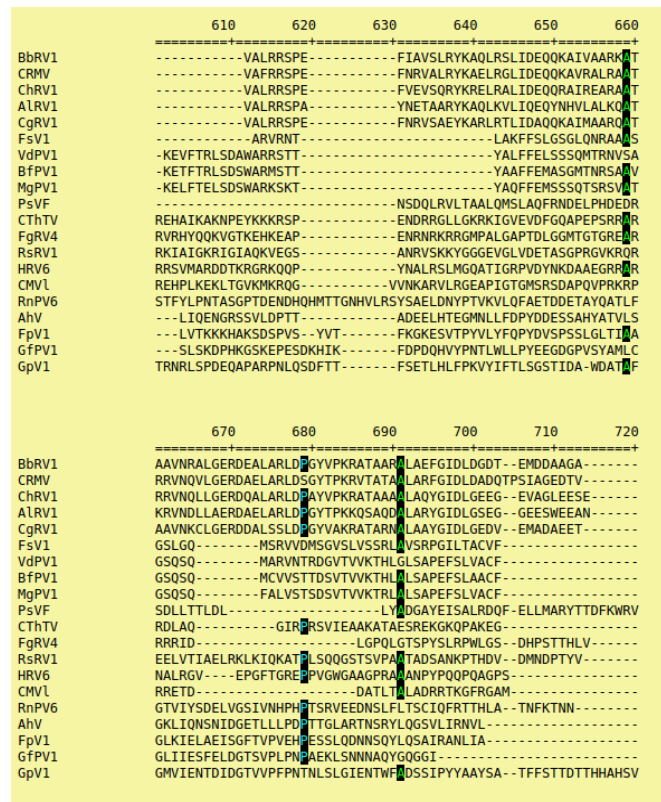


Figura 5. Extracto del alineamiento mediante MAFFT de la proteína CP visualizado en Gblocks. Se seleccionaron 20 virus distintos, cinco de cada uno de los cuatro grupos, es decir, *Beta-* (RnPV6 y los cuatro siguientes), *Gammapartitivirus* (FsV1 y los cuatro siguientes) y micovirus de dsRNA no segmentados (BbRV1 y los cuatro siguientes) y bipartitos (CthTV y los cuatro siguientes). Se eligió un número igual de virus de cada grupo para no sesgar la identificación de bloques conservados a regiones conservadas dentro de cada grupo. Para ello, la elección se basó en los árboles filogenéticos realizados para intentar representar ramas distantes de cada grupo. El programa Gblocks con los parámetros más permisivos no consiguió encontrar ninguna región conservada entre estas secuencias. La región mostrada es la que cuenta con más posiciones marcadas, es decir, que presentan el mismo aminoácido en 11 o más de las 20 secuencias.

En definitiva, es difícil aproximar la historia evolutiva de CP con árboles filogenéticos. A pesar de que separe a grandes rasgos los mismos grupos de virus actuales que el árbol de RdRP, el árbol de CP no consigue establecer relaciones ancestrales entre los grupos coloreados en las figuras. Estos resultados son consecuentes con lo que se sabe acerca de la conservación de las secuencias de ambas proteínas, donde RdRP constituye una herramienta mucho más potente para realizar estudios filogenéticos entre virus de RNA evolutivamente lejanos.

3.2.2. Red filogenética de la secuencia concatenada de RdRP y CP

De nuevo, el algoritmo de alineamiento que producía un conjunto de datos más adecuado para el análisis por máxima verosimilitud fue MAFFT, seguido de MUSCLE y T-Coffee. El alineamiento estuvo formado de 71 secuencias concatenadas, ya que se incluyeron de nuevo los chrysovirus y se eliminaron las secuencias de CPV, donde RdRP era de un alfapartitivirus y CP de un betapartitivirus, no era el mismo organismo. Los algoritmos NeighborNet, SplitDecomposition y ParsimonySplits reconstruían redes con una agrupación de ramas similar. En la red de NeighborNet se puede apreciar cómo en la base de todos los grupos hay numerosos eventos de transferencia horizontal, que se vuelven más infrecuentes según las ramas avanzan y se separan. Si bien el análisis no incorpora datos con información temporal, parece que hay pocos eventos recientes de este tipo, aunque pueden observarse en distintas parejas, como CpPV y CrV1. De todas formas, no se puede afirmar que esta recombinación entre virus suponga una fuente de variación que pueda explicar grandes divisiones de

estos grupos de micovirus. La topología de las ramas en la red muestra, principalmente, características del árbol de RdRP, como un antepasado común y único a todos los gammapartitivirus y la relación entre los dos grupos de micovirus no clasificados.

Esta red permite visualizar claramente cómo algunos virus que infectan a hongos estrechamente relacionados también guardan un parentesco notable entre ellos, como CpPV y CrV1, que infectan a hongos del género *Ceratocystis*. Luego hay secuencias prácticamente idénticas que infectan a la misma especie, como FgRMV4 y FgRMV5 o GpV1 y CpV1, algo que se corresponde con la variabilidad natural que muestran los virus de RNA. También es interesante ver cómo micovirus de hongos ecológicamente relacionados como *Magnaporthe oryzae* (MoPV3), *M. grisea* (MgPV1) y *Nigrospora oryzae* (NoPV1), tres hongos fitopatógenos del arroz, presentan una relación tan estrecha y, sorprendentemente, más significativa en el árbol de CP que en RdRP. Sin embargo, no se puede concluir con estos datos que se haya producido algún tipo de transmisión del virus entre ellos o que los tres micovirus actuales provengan de una coevolución con los antepasados comunes de estos hongos, cuyo primer rango taxonómico común es el de la clase *Sordariomycetes*.

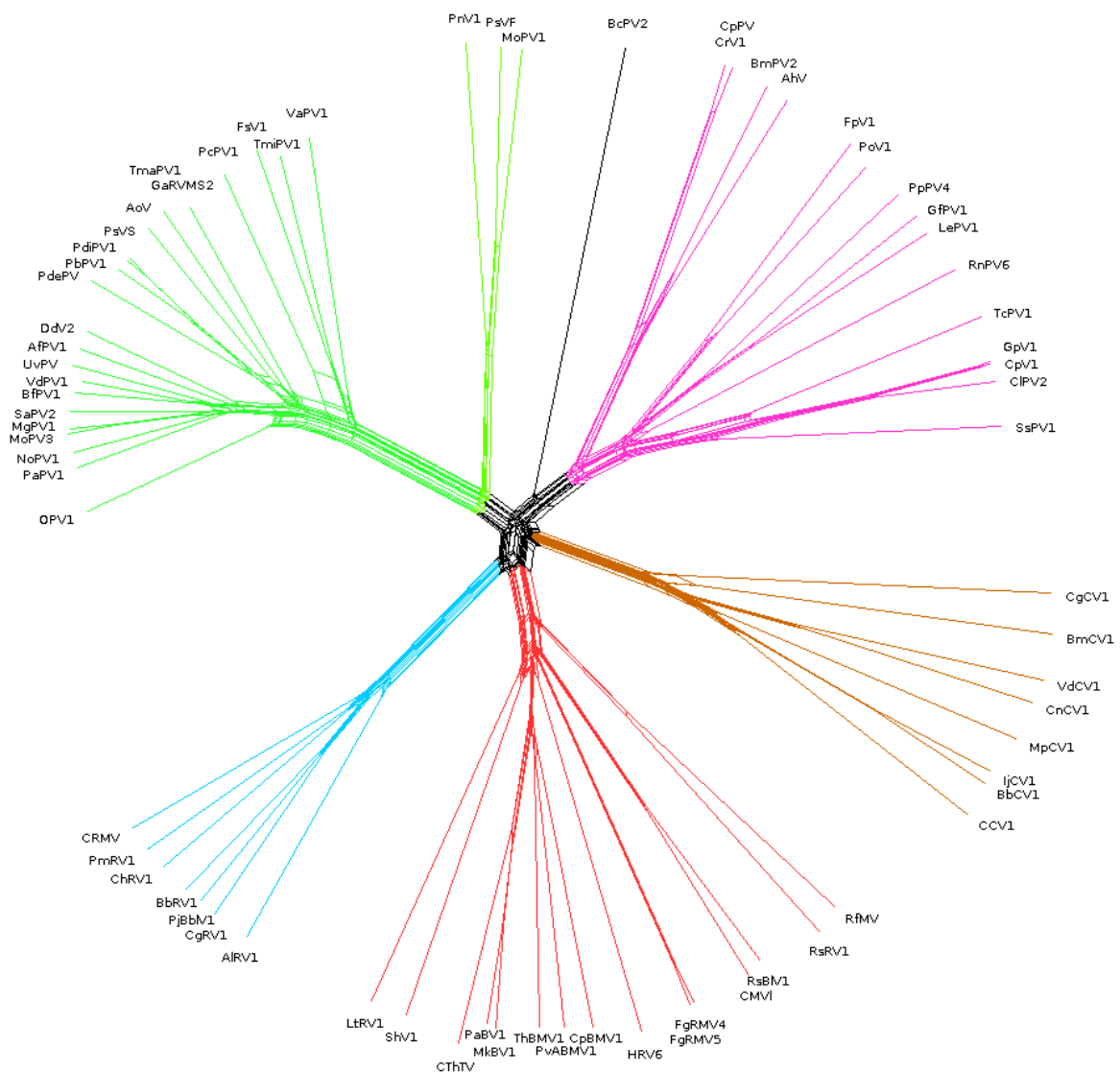


Figura 6. Red filogenética de la secuencia concatenada de RdRP y CP en 71 micovirus reconstruida mediante el algoritmo NeighborNet. Las ramas coloreadas representan los principales grupos que logra separar el árbol de RdRP: los beta- (morado), los gammapartitivirus (verde), los micovirus no clasificados monopartitos (azul) y bipartitos (rojo) y chrysovirus (marrón). No se observa ninguna conexión horizontal

entre grupos, solo eventos intragrupalos. Los valores de *bootstrap* de los nodos no se muestran en la imagen, pero los cinco grupos se separan con valores de 1, excepto los gammapartitivirus, que lo hacen con 0.97. El nodo basal no coloreado que agrupa los virus BbRV1-like y ThBMV1-like se predice con un valor de *bootstrap* de 0.93. La credibilidad de los eventos de transferencia horizontal centrales es notablemente inferior, con valores incluso por debajo de 0.05. Otros eventos, como el que une a un antepasado de PnV1 con uno de PsVF y MoPV1 se predicen con un valor de *bootstrap* de 0.53.

3.2.3. Red filogenética de CP con la proteína hipotética de hongos

Las grandes similitudes a nivel de secuencia entre la proteína hipotética de hongos y las CP de los virus ThBMV1-like pueden apreciarse en el alineamiento múltiple de las secuencias mediante MUSCLE (Figura 7). El parecido se concentra principalmente en el centro de la proteína, mientras que, en los extremos, la longitud variable en torno a este núcleo mejor conservado provoca alineamientos pobres con grandes gaps, incluso dentro de las CP. Sea como sea, la red filogenética a partir de este alineamiento no mostró que la CP sea una mezcla de origen viral y eucariota, en el sentido que en la recombinación solo se hubiera transferido la región central. Aparte de los residuos estrictamente idénticos, también es posible observar posiciones con una alta incidencia de aminoácidos similares, como el grupo de aminoácidos hidrofóbicos que forman Leu, Ile, Val, Ala y Phe o Tyr en posiciones donde también hay Trp y Phe, como en la posición 163, aminoácidos aromáticos que pueden tener un efecto similar sobre el plegamiento. De forma análoga, en las posiciones 292 y 330 se encuentran ambos aminoácidos polares con carga positiva, Lys y Arg, dentro de las secuencias de virus y hongos.

En la red (Figura S1), algunos virus ThBMV1-like como CpBMV1 colocalizan con secuencias de hongos bajo nodos con valores de *bootstrap* muy significativos, tanto como 0.998 en este caso, indicando una similitud aún mayor con estas secuencias que con el resto de virus de este grupo. Dentro de los hongos, la similitud parece ser mayor con las secuencias de ascomicetos, que eran los mejores resultados en la búsqueda BLAST.

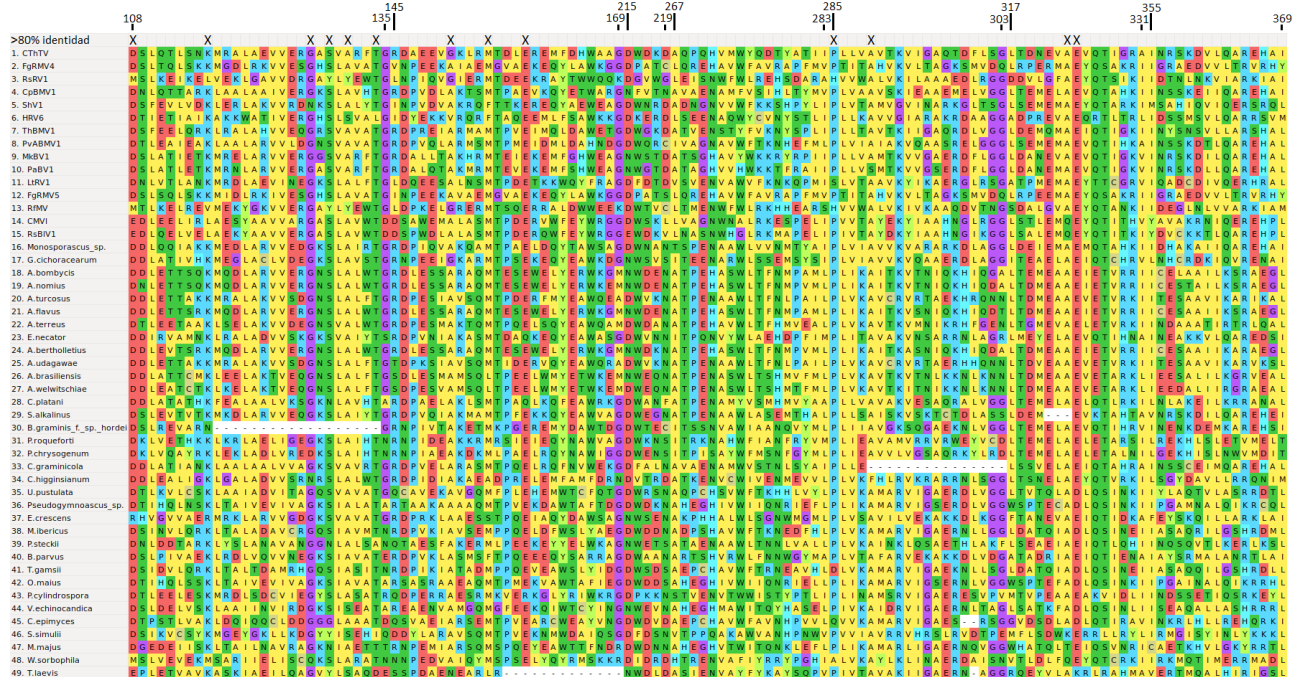


Figura 7. Concatenación de los bloques conservados en el alineamiento a lo largo de las secuencias de CP de virus ThBMV1-like y de la proteína hipotética de hongos. El alineamiento múltiple original mediante MUSCLE, de 609 posiciones de longitud, se redujo mediante el programa Gblocks con los parámetros más

permisivos a 124 posiciones con un mayor grado de identidad. Las posiciones indicadas se refieren a la situación en el alineamiento múltiple. Las posiciones con más de un 80% de incidencia del mismo aminoácido están marcadas en la parte superior con una “X”. Se eliminaron 6 secuencias del género *Aspergillus* para la toma de la imagen y no se consideraron para la determinación del porcentaje de identidad. Más residuos, principalmente Pro, fuera de estas regiones mostraban un alto grado de identidad en el alineamiento original, pero al estar aislados no se consideraron como bloques.

Como es de esperar con estos resultados, en la red a partir del alineamiento incluyendo todos los grupos de virus y estas proteínas de hongo, los virus ThBMV1-like se agrupan con las proteínas hipotéticas de hongo en la red en el lado opuesto a donde lo hacen el resto de virus (Figura 8). Esta proteína hipotética eucariota presenta, en definitiva, una similitud a nivel de secuencia con la CP de ThBMV1-like muy superior a la que presentan micovirus significativamente relacionados en RdRP como los virus BbRV1-like.

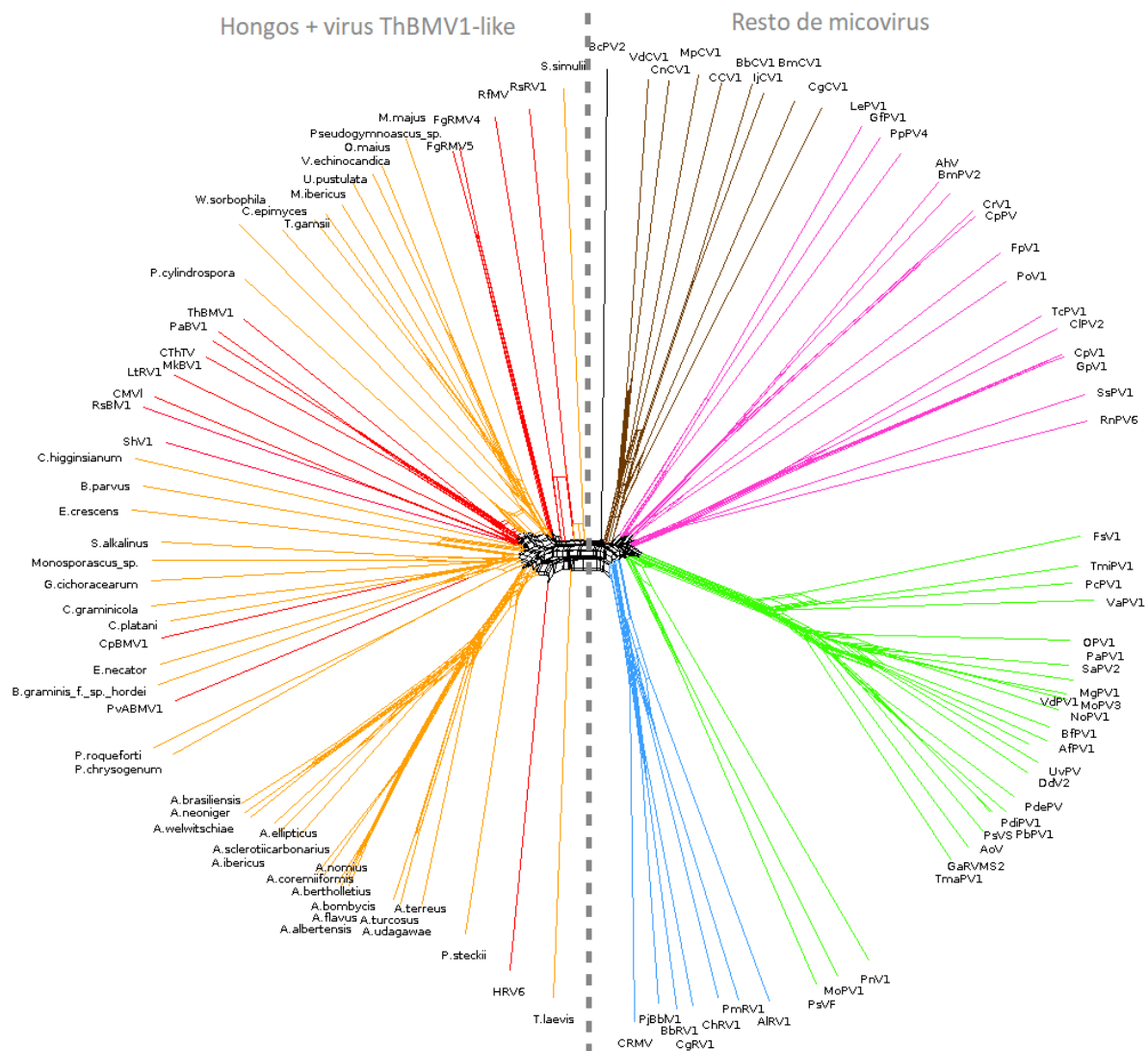


Figura 8. Red filogenética de NeighborNet del alineamiento mediante MUSCLE de todas las secuencias de CP de micovirus incluidos en el estudio y las 40 proteínas hipotéticas de hongo. El grupo de virus ThBMV1-like, al contrario del resto de grupos de virus, muestra una variabilidad intragrupal y una similitud tan elevada con la proteína fúngica tales que el grupo se dispersa dentro de las secuencias de hongos. Los hongos no ascomicetos, *T. laevis* y *S. simulii*, se muestran más alejados de las secuencias de ascomicetos que los virus ThBMV1-like. Las ramas están coloreadas según la clasificación taxonómica de las especies que comprenden: marrón (chrysovirus), morado (betapartitivirus), verde (gammapartitivirus), azul (virus BbRV1-like), rojo (virus ThBMV1-like), naranja (hongos).

3.3. ANÁLISIS ESTRUCTURAL DE CP Y LA PROTEÍNA DE HONGO SIMILAR A ThBMV1

3.3.1. Predicción estructural

En primer lugar, los betapartitivirus GfPV1 y TcPV1 presentaban primeros modelos, normalmente los más creíbles, dispares. Sin embargo, el tercer modelo de GfPV1 produce un alineamiento significativo, con un *TM-score* > 0.6, con el segundo, cuarto y quinto modelo de TcPV1. La fiabilidad de la predicción, representada por el valor *C-score* era cercana a -3 y similar, en ambos casos a la del primer modelo. Esta puntuación oscila entre -5 y 2, significando 2 una confianza máxima en el modelo. Ambas proteínas llegaron a modelos muy similares usando como plantilla distintas estructuras, lo que refuerza la probabilidad de que la estructura real se asemeje a esta. Cabe destacar que no hay una estructura disponible para la CP de los betapartitivirus en el PDB.

Aparentemente, esta proteína tiene un cuerpo formado por hélices alfa en una disposición compleja y una pequeña protuberancia en torno a los residuos 480-500 (Figura S2), que se conserva en los cuatro modelos predichos similares. El cuerpo de hélices alfa se estrecha en la región central hacia N-terminal, lo que podría suponer un sitio de interacción proteína-proteína en la dimerización de CP.

En segundo lugar, los dos gammapartitivirus, FsV1 y BfPV1, formaron estructuras muy similares en el primer modelo. El alineamiento de estos dos modelos dio lugar a un alineamiento con un *TM-score* > 0.8. Los extremos de la proteína muestran cierta variabilidad, mientras que 370 posiciones centrales se situaban a menos de 5 Å entre ambos modelos y eran responsables de la elevada identidad del alineamiento.

La estructura consenso presenta dos protuberancias en la misma dirección que emergen de una estructura principal formada principalmente por hélices alfa. Estas dos protuberancias, una al principio de la secuencia y otra al final parecen estar menos conservadas que el cuerpo principal en las proteínas analizadas, pero hay motivos comunes, como la hélice alfa en C-terminal, aunque con una orientación distinta, y el asa que se forma en la otra protuberancia.

Las principales proteínas que se han usado como plantilla para la predicción estructural han sido la CP de PsVS y PsVF (Tang et al., 2010; Pan et al., 2009), otros gammapartitivirus, mostrando la primera un valor de *TM-score* \approx 0.85 con FsV1 y BfPV1, mientras que la CP de PsVF apenas llegaba a una puntuación de 0.6. La homología estructural con estas estructuras de elucidación experimental (Figura S3) aumenta la credibilidad de estos modelos y es consistente con la segregación observada en el análisis filogenético de CP, donde PsVS compartía una relación más cercana con FsV1 y BfPV1. De todas formas, la divergencia en CP de PsVF a nivel de secuencia respecto al resto de gammapartitivirus no se traduce en grandes diferencias a nivel estructural. Una de las estructuras usadas por FsV1 como plantilla fue la CP del totivirus L-A de *S. cerevisiae* (Naitow et al., 2002), aunque el parecido es remoto y el *TM-score* entre ambas estructuras es ligeramente superior a 0.3.

En tercer lugar, la estructura predicha de VdCV1 muestra una muy buena homología con la estructura de la CP del chrysovirus *Penicillium chrysogenum* Virus (PcV), elucidada por criomicroscopía electrónica por Luque et al. (2014), que se usó como plantilla con un *TM-score* \approx 0.9. También se usó como plantilla la estructura del virus L-A con un *TM-score* de 0.342, siempre normalizado para VdCV1. El primer modelo de VdCV1 permitió apreciar la naturaleza duplicada de CP en los chrysovirus con dos subunidades similares unidas por una secuencia *linker* (Figura S4). Las subunidades están formadas principalmente por hélices alfa, aunque hay regiones de láminas beta, como también se observa en PcV.

En el grupo de los micovirus de dsRNA no segmentados no se pudo determinar una estructura consenso entre BbRV1 y A1RV1. Ninguno de los cinco modelos simulados para cada proteína era capaz de alinear con alguno de los modelos de la otra con un *TM-score* > 0.36. Aparte de que

predominan las hélices alfa en ambas proteínas, poco más se puede decir a nivel estructural. Las diferencias a nivel de secuencia y longitud no facilitan que se simulen orientaciones similares de estos motivos de estructura secundaria.

En quinto lugar, los primeros modelos de CP de ThBMV1 y CpBMV1 alineaban con un *TM-score* de 0.916 normalizado para ThBMV1, la puntuación más elevada entre parejas de un mismo grupo observada en el análisis. Hay pequeñas diferencias entre ambos modelos, como la sexta hélice alfa desde C-terminal, que no se predice como tal en ThBMV1, y las regiones de los extremos (Figura 9). La estructura consenso consta de 13 hélices alfa en forma de empalizada. Cerca de N-terminal, un pequeño bucle entre hélices está conservado en estructura, pero no en posición entre las dos especies analizadas, en CpBV1 está más cerca del extremo N-terminal, acortando la longitud de la hélice alfa inicial de la proteína. La proteína que más ha influido en la simulación estructural de ambos modelos fue una proteína de interacción con la proteína quinasa ATR humana (abr. ATRIP), elucidada por criomicroscopía electrónica por Rao et al. (2018). Hay evidencias de que ATRIP es capaz de formar homodímeros e interactuar con el ssDNA para dirigir la actividad reparadora del complejo ATR-ATRIP (Itakura et al., 2005; Ball et al., 2005).

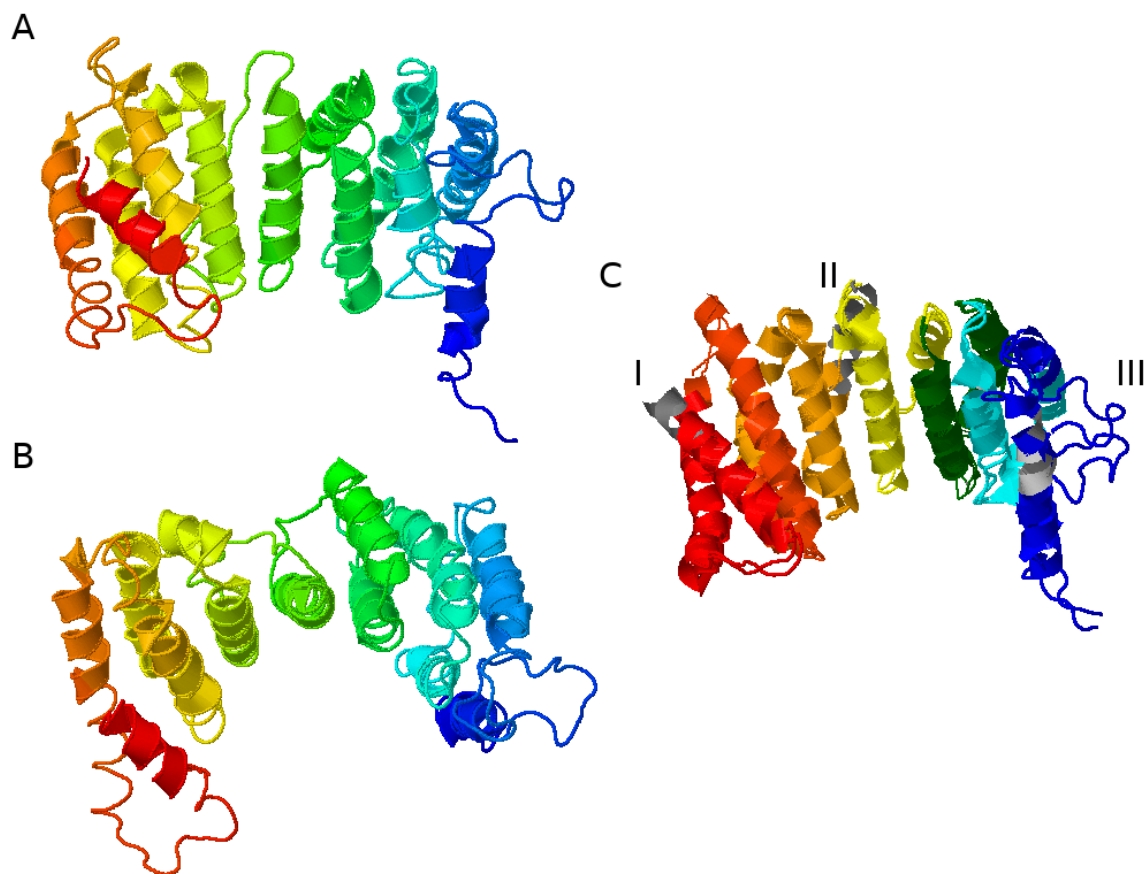


Figura 9. El esquema de colores en el rango del azul (N-terminal) al rojo (C-terminal) indica la posición en la cadena polipeptídica. (A) Visión frontal del primer modelo predicho para la CP de ThBMV1. La estructura se compone de una empalizada de 13 hélices alfa consecutivas. El plano vertical de esta empalizada, sin embargo, está torcido en los extremos en sentidos opuestos, tomando como referencia las hélices centrales. (B) Plano cenital del modelo 1 de ThBMV1. (C) Alineamiento de los primeros modelos predichos de ThBMV1 (blanco) y CpBMV1 (gris). Las regiones coloreadas presentan elevada identidad estructural. Las principales diferencias entre ambos modelos son: (I) la hélice alfa en C-terminal de CpBMV1 es más larga, (II) la octava hélice desde N-terminal o sexta desde C-terminal no se predice como tal en ThBMV1 y (III) la posición del bucle cercano al extremo N-terminal.

En último lugar, los mejores alineamientos se conseguían con el modelo 1 de *A. bombycis*, el modelo 1 de *C. epimyces* y el modelo 3 de *E. necator*. Las puntuaciones de *C-score* de los distintos modelos de estas proteínas ortólogas de hongos se situaron por debajo de -3, indicando una baja fiabilidad. Los tres alineamientos posibles dieron valores de *TM-score* entre 0.35 y 0.45, un intervalo en el que el alineamiento no se considera significativo ni aleatorio. A simple vista, la estructura sí parece ser similar, formando una empalizada más o menos recta de 9 hélices alfa consecutivas y alternadas en el caso de *A. bombycis* y *C. epimyces* y de 8 en el caso de *E. necator*, aunque en esta especie las hélices adyacentes no siempre son contiguas en la secuencia (Figura 10). Es probable que la elevada variabilidad observada no sea real y se deba a que las proteínas problema hayan utilizado distintas estructuras como plantilla en la simulación del plegamiento. Esto puede afectar a la orientación de los bucles entre estructuras secundarias como las hélices alfa.

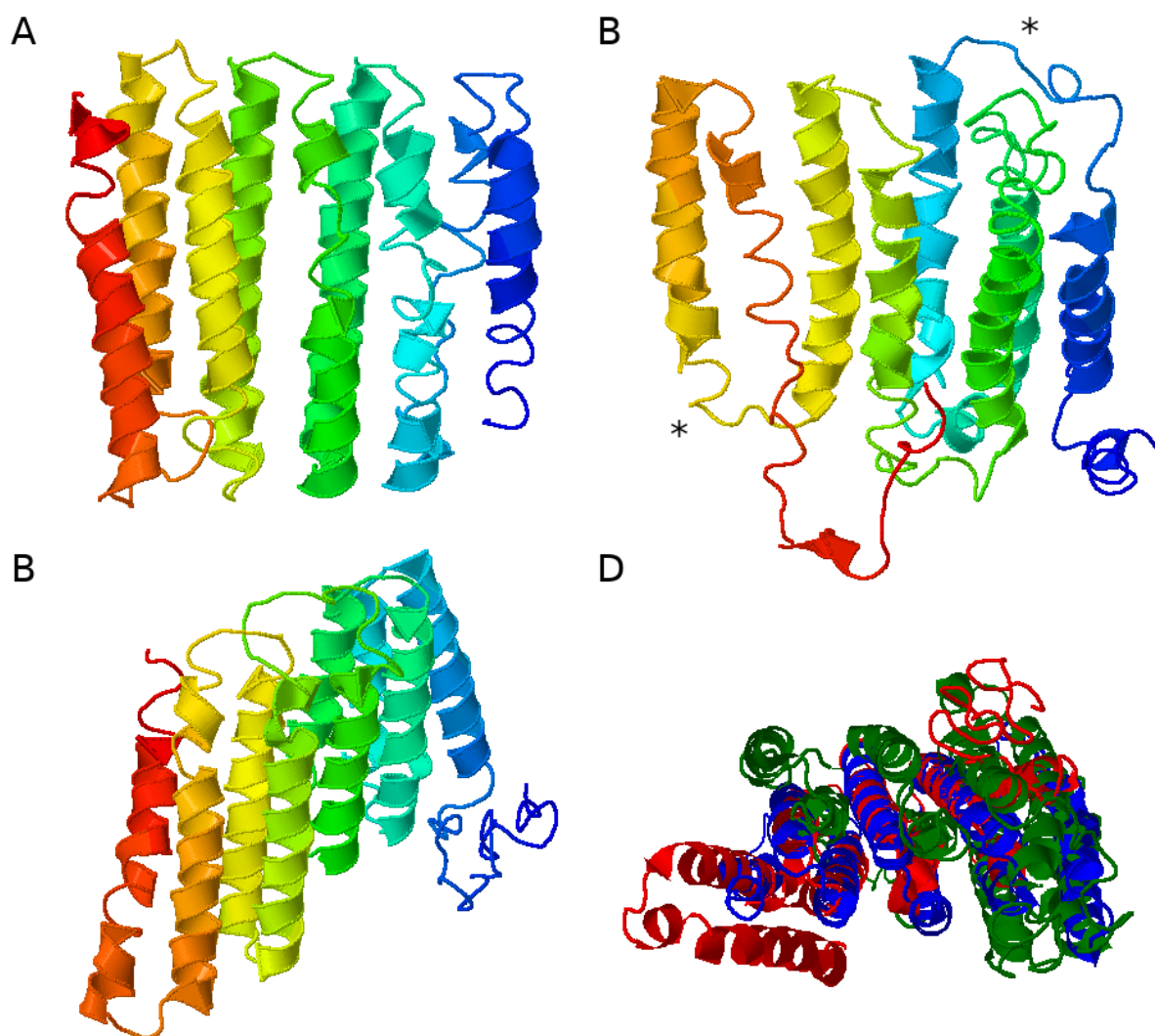


Figura 10. Plano vertical de los modelos predichos de las proteínas hipotéticas de *A. bombycis* (A), *C. epimyces* (B) y *E. necator* (C). La estructura de las tres proteínas está formada, predominantemente, por hélices alfa. Los asteriscos marcan bucles que unen hélices alfa contiguas a nivel de secuencia, pero no adyacentes en cuanto a la estructura. (D) Alineamiento múltiple de los modelos de *A. bombycis* (azul), *C. epimyces* (rojo) y *E. necator* (verde). A pesar del buen alineamiento de las hélices alfa centrales de *A. bombycis* y *C. epimyces*, no se observa una buena homología, ya que los ángulos entre hélices alfa difieren entre modelos e impiden una identidad más extensa.

3.3.2. Predicción funcional

Las funciones predichas más extendidas entre grupos son la de unión de proteínas y transporte de proteínas (Tabla S3). Para los micovirus ThBMV1-like y BbRV1-like se predijo la función de unión a RNA. En cuanto al componente celular, estructuras similares a las simuladas se localizan en el citoplasma, en el núcleo y en la envoltura nuclear, principalmente. Algunas funciones predichas repetidamente en ciertos grupos parecen no tener sentido en el contexto de una proteína de la cápsida, como los numerosos términos GO relacionados con el metabolismo predichos para los gammapartitivirus. Pocas veces se predicen términos GO relacionados con el ciclo viral, lo que se puede explicar con las escasas estructuras de CP elucidadas en el PDB, quedando así la predicción de términos GO sesgada hacia otros tipos de proteínas estructuralmente mejor estudiadas.

No hay un consenso acerca de la función de la proteína hipotética de *A. bombycis*, *C. epimyces* y *E. necator*. Las tres están enriquecidas en términos GO de unión a proteínas y *A. bombycis* y *C. epimyces* involucran términos GO relacionados con el ciclo viral, aunque con una ocurrencia muy baja. *E. necator* es responsable de que en la suma aparezcan tantos términos GO relacionados con el metabolismo.

3.3.3. Alineamiento de estructuras entre grupos

El parecido estructural de CP entre los grupos de virus es muy bajo, dándose alineamientos con valores de *TM-score* ≈ 0.3 . Tampoco la CP de los betapartitivirus puede formar un buen alineamiento con la de los gammapartitivirus, indicando grandes diferencias en la estructura de CP dentro de la familia *Partitiviridae* y asumiendo que la estructura predicha para los primeros se asemeja al plegamiento real, que no ha sido elucidado. A grandes rasgos, las estructuras predichas de CP de partitivirus y virus no clasificados predichas presentan una elevada proporción de hélices alfa y ninguna o muy pocas regiones de láminas beta.

El modelo 1 de BbRV1 presenta una estructura a modo de torsión de una empalizada de hélices alfa similar a la observada en el grupo ThBMV1-like, ya que también utilizó la proteína ATRIP humana como plantilla principal. El valor de *TM-score* en el alineamiento del modelo 1 de CP de BbV1 y el modelo 1 de ThBV1 es tan significativo como 0.89. Esta relación es coherente con los resultados de la filogenia con RdRP, donde ambos grupos de virus parecen estar relacionados. Sin embargo, no podemos concluir que este modelo simulado se acerque a la estructura real de la CP de este grupo, ya que no encontramos en AIRV1 una estructura similar.

Las tres proteínas hipotéticas de hongo simuladas alineaban con los modelos de CP de virus ThBMV1-like con valores de *TM-score* incluso superiores a los que se observaron en los alineamientos intragrupal, aunque no muy significativos en términos absolutos. El alineamiento que maximizaba este valor fue CpBMV1 con *A. bombycis*, con una puntuación de 0.51 normalizada para *A. bombycis*, la secuencia más corta. Con *C. epimyces*, CpBMV1 alineó con un *TM-score* de 0.42 y ThBMV1 con 0.41. El alineamiento de estas dos proteínas con *E. necator* fue peor, con valores de *TM-score* de 0.33 y 0.34, respectivamente. El alineamiento múltiple de la proteína hipotética de *A. bombycis* y *C. epimyces* con las CP de CpBMV1 y ThBMV1 muestra una elevada homología en 6 hélices alfa centrales (Figura 11). Al sacar la proteína de *C. epimyces* del alineamiento, la homología se extiende a lo largo de la práctica totalidad de la proteína hipotética de *A. bombycis*, si bien sus hélices son más largas que las que se observan en la CP.

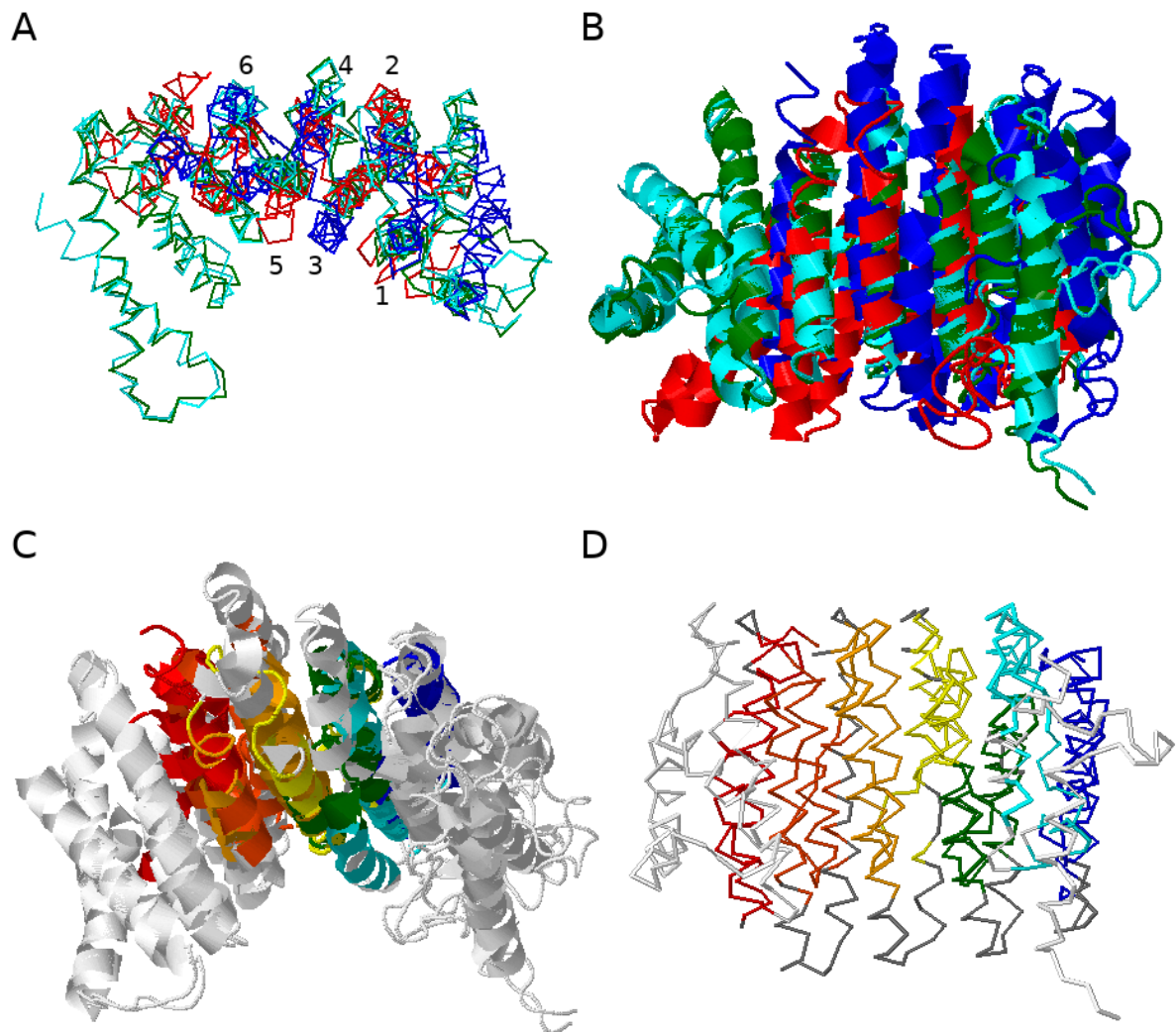


Figura 11. (A, B) Distintas perspectivas del alineamiento múltiple de las proteínas hipotéticas de *A. bombycis* (azul), *C. epimyces* (rojo), CP de ThBMV1 (verde) y CpBMV1 (turquesa). Las seis hélices alfa que colocalizan aproximadamente entre las cuatro estructuras están numeradas según su orden de aparición desde el extremo N-terminal. (C) Visión del alineamiento múltiple con las regiones posicionadas de forma similar coloreadas según su posición en la secuencia: de azul (N-terminal) a rojo (C-terminal). (D) Alineamiento de CP de ThBMV1 con la proteína hipotética de *A. bombycis* con el código de colores anterior. Las regiones grises pertenecen únicamente a *A. bombycis* y las blancas a ThBMV1. El alineamiento con este modelo de hongo en particular alarga el alineamiento a las nueve hélices alfa, si bien las hélices de la proteína de *A. bombycis* son más largas. La estructura de ThBMV1 tiene hélices adicionales que flanquean la estructura de la proteína de *A. bombycis* en ambos extremos y tomando otras direcciones, sobre todo en C-terminal.

4. DISCUSIÓN

4.1. RELACIONES ENTRE LOS GRUPOS DE VIRUS ESTUDIADOS

Como era de esperar, el análisis basado en secuencia presentó dificultades importantes para reconstruir árboles filogenéticos, de una forma más acusada en CP. A la hora de realizar búsquedas de PSI-BLAST, las secuencias de RdRP dentro de la familia *Partitiviridae* eran lo suficientemente similares para encontrar los ortólogos entre géneros. Así mismo, con una búsqueda simple de las RdRP, se pueden encontrar virus BbRV1-like con virus ThBMV1-like y a la inversa. Esta homología de secuencia indica un origen común en este locus y ha llevado a que se proponga una clasificación taxonómica común, más concretamente, la de amalgamycovirus, dentro de la familia *Amalgaviridae*, pero formando un linaje distinto del resto de amalgavirus, comprendiendo solo aquellos que infectan a hongos (Koloniuk et al., 2015).

El nombre de esta familia se debe a que estos virus presentan una serie de características en común con los totivirus y los partitivirus. Con los primeros guardan similitud a nivel de organización genómica, ya que los amalgavirus reconocidos tienen un genoma no segmentado y muy probablemente expresen proteínas de fusión RdRP-CP, mientras que con los segundos a nivel de secuencia con RdRP (Sabanadzovic et al., 2009). Con las secuencias y métodos seleccionados, no se ha podido observar una similitud significativa con los partitivirus, ya que los chrysovirus no han servido para hacer una raíz del subárbol formado por partitivirus y virus no clasificados.

Los virus ThBMV1-like presentarían el primer caso de genomas bipartitos dentro de la familia *Amalgaviridae*, con los genes de RdRP y CP dispuestos en moléculas de RNA independientes como en los partitivirus. El hecho de que los virus de ThBMV1-like estén filogenéticamente más relacionados con virus no segmentados, lleva a pensar que el evento de partición del genoma sea distinto al que originó los partitivirus. El parentesco de los virus ThBMV1-like con los partitivirus es más remoto que el parentesco que presentan con estos los otros amalgavirus (Zheng et al., 2013).

La divergencia en la CP, tanto a nivel de secuencia como estructural, hace muy difícil determinar que las proteínas que vemos actualmente en los distintos grupos de virus descienden de una única CP ancestral. Las grandes diferencias intergrupales y la relativa conservación dentro de los grupos podrían sugerir que la formación de nuevos linajes está precedida de cambios radicales en CP, originando nuevas estructuras alternativas para la formación de cápsidas icosaédricas. La proteína RdRP, que interactúa con CP para su localización en el interior de la cápsida, se vería así obligada a adaptarse a la nueva CP, manteniendo más conservadas regiones posiblemente relacionadas con la replicación del genoma, como el motivo GDD, presente en todas las especies analizadas. El proceso por el que ocurren estos cambios radicales puede deberse a situaciones de cuello de botella típicas de los virus de RNA en los que la elevada tasa de mutación permite dar con una alternativa viable o a eventos de recombinación, que pueden suponer saltos macroevolutivos drásticos, con el huésped o con otros virus, como se cree que ocurrió con la CP de los amalgavirus (Krupovic et al., 2015).

4.2. SECUENCIAS EN HONGOS SIMILARES A CP DE MICOVIRUS

Para varios grupos de micovirus se observaron secuencias similares a CP en genomas eucariotas, concretamente de plantas y hongos, que suponen la mayor parte del rango de huéspedes de los virus relacionados con los micovirus. Para las secuencias observadas en eucariotas similares a la CP de los virus BbRV1-like y los chrysovirus la explicación más probable es que se haya pseudogenizado un transcrito de CP por lo poco extendidas que están estas secuencias y la naturaleza truncada de las secuencias homólogas a CP en genomas eucariotas.

La integración en el genoma del material genético de un virus de dsRNA parece improbable, pero no es imposible que se produzca en presencia de elementos retrotransponibles, que ya se han descrito en hongos (Daboussi y Capy, 2003). De esta manera, los mRNA virales, así como los celulares, son

susceptibles de retrotranscribirse e integrarse en el genoma, degenerando la mayoría de veces en un pseudogen procesado. Se estima, que del 1 al 5% de los genes de hongos tienen en su genoma un pseudogen homólogo, incluyendo aquellos procesados y no procesados. Además, la pseudogenización de ciertos genes podría ser clave en el proceso de adaptación de ciertas especies, observándose grandes diferencias entre hongos estrechamente relacionados (Van der Burgt et al., 2014). Ya que los pseudogenes son una fuente de RNA antisentido u otro tipo de RNA interferentes (Muro y Navarro, 2010; Roberts y Morris, 2013), esto podría suponer un importante punto de interacción entre los micovirus y sus huéspedes.

Las secuencias similares a la CP de ThBMV1-like, por otra parte, están extendidas entre especies de hongos muy distantes. Es muy poco probable que estas secuencias sean homólogas de un evento de pseudogenización ancestral de CP por la elevada identidad de secuencia entre hongos evolutivamente alejados, además de conservar una longitud de la proteína relativamente similar. Estos datos sugieren que esta secuencia es un gen propio de hongos, aunque de función desconocida, putativamente relacionada con la interacción con ácidos nucleicos. Esta función es de especial interés cuando se trata de la CP, ya que debe poder discriminar qué moléculas de dsRNA encapsidar, discriminando entre especies de virus en el caso de los gammapartitivirus (Pan et al., 2009). La elevada identidad de la estructura de CP de los virus ThBMV1-like con la proteína ATRIP apunta en una dirección similar, ya que esta proteína está involucrada en la interacción de ácidos nucleicos y, además, es capaz de dimerizar, algo que muy seguramente haga esta CP para la formación de la cápsida.

Es posible que la CP de algún otro grupo de micovirus de dsRNA tenga su origen en un gen eucariota, pero los cambios de secuencia necesarios para la adaptación a la nueva función y el tiempo transcurrido desde el evento de transferencia pueden dificultar el rastreo de genes eucariotas candidatos.

4.3. SIMILITUD ESTRUCTURAL DE LA CP DE ThBMV1-like CON UNA PROTEÍNA HIPOTÉTICA DE HONGOS

Salvando las diferencias entre los tres modelos de hongos predichos que dificultan asignar una estructura consenso, esta proteína está compuesta de alrededor de 8 o 9 hélices alfa consecutivas y aproximadamente paralelas formando una empalizada relativamente recta. Las estructuras predichas de CP de los virus ThBMV1-like presentan una estructura muy similar a esta proteína hipotética.

Estos resultados, junto con el parecido a nivel de secuencia, refuerzan la idea de un origen común entre ambas proteínas, sugiriendo una transferencia desde el huésped a un antepasado común de los virus ThBMV1-like. Este evento habría provocado la pérdida de la CP nativa de los amalgamicovirus en el grupo de virus bipartitos. Queda por saber si el evento de bipartición del genoma fue simultáneo al evento de transferencia o anterior. En la primera situación, todos los amalgamicovirus bipartitos deberían tener una CP homóloga, como parece ser el caso hasta la fecha. Otras proteínas ortólogas de hongo no simuladas con un mayor tamaño podrían esclarecer si el gen que se transfirió al virus ya tenía esta estructura a modo de torsión de la empalizada, sobre todo en C-terminal, o se adquirió ya en los virus.

El éxito de este evento de sustitución de la CP nativa depende de la capacidad de la proteína eucariota original de hacer las funciones de CP, concretamente la interacción con otros monómeros para dimerizar, con RdRP y, posiblemente, con el genoma de dsRNA. Además, el tamaño de esta proteína fúngica oscila entre números de residuos que comprenden a la longitud de la CP de los virus BbRV1-like. La elevada homología de secuencia con CP podría sugerir un origen relativamente reciente de este grupo de micovirus. Más estudios acerca de la función y la localización celular de esta proteína, posiblemente transmembrana por sus hélices con regiones hidrofóbicas a una misma altura, podrían esclarecer qué características pudieron ser útiles para su éxito como CP.

5. CONCLUSIÓN

La aproximación de la evolución de los micovirus de dsRNA mediante métodos tradicionales como los árboles filogenéticos es compleja, no solo por las elevadas tasas de mutación, también por eventos de transferencia horizontal de genes. El análisis estructural permite retroceder más en la historia evolutiva y rastrear parentescos más remotos. Se recogieron secuencias de RdRP y CP de distintos grupos de micovirus de dsRNA y se buscaron genes eucariotas similares a CP, encontrando para la CP del grupo de micovirus de dsRNA bipartitos no clasificados similares a ThBMV1 un gen de hongo candidato a haberse transferido desde el huésped al virus, siendo la causa de la diferenciación de este grupo de virus ThBMV1-like. Ambas proteínas conservan un parecido significativo a nivel estructural y de secuencia.

6. BIBLIOGRAFÍA

1. Abrescia, N. G. A., Bamford, D. H., Grimes, J. M., & Stuart, D. I. (2012). Structure Unifies the Viral Universe. *Annual Review of Biochemistry*, 81(1), 795-822. <https://doi.org/10.1146/annurev-biochem-060910-095130>
2. Ball, H. L., Myers, J. S., & Cortez, D. (2005). ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation. *Molecular Biology of the Cell*, 16(5), 2372-2381. <https://doi.org/10.1091/mbc.E04-11-1006>
3. Bandelt, H. J., & Dress, A. W. M. (1992). Split decomposition: A new and useful approach to phylogenetic analysis of distance data. *Molecular Phylogenetics and Evolution*, 1(3), 242-252. [https://doi.org/10.1016/1055-7903\(92\)90021-8](https://doi.org/10.1016/1055-7903(92)90021-8)
4. Belshaw, R., Gardner, A., Rambaut, A., & Pybus, O. G. (2008). Pacing a small cage: mutation and RNA viruses. *Trends in Ecology and Evolution*, 23(4), 188-193. <https://doi.org/10.1016/j.tree.2007.11.010>
5. Bevan, E. A., Herring, A. J., & Mitchell, D. J. (1973). Preliminary characterization of two species of dsRNA in yeast and their relationship to the «Killer» character. *Nature*, 245(5420), 81-86. <https://doi.org/10.1038/245081b0>
6. Biella, S., Smith, M. L., Aist, J. R., Cortesi, P., & Milgroom, M. G. (2002). Programmed cell death correlates with virus transmission in a filamentous fungus. *Proceedings of the Royal Society B: Biological Sciences*, 269(1506), 2269-2276. <https://doi.org/10.1098/rspb.2002.2148>
7. Bozarth, R. F. (1972). *Mycoviruses: a new dimension in microbiology*. En *Environmental health perspectives* (Vol. 2, pp. 23-39). National Institute of Environmental Health Sciences. <https://doi.org/10.1289/ehp.720223>
8. Bryner, S. F., Prospero, S., & Rigling, D. (2014). Dynamics of cryphonectria hypovirus infection in chestnut blight cankers. *Phytopathology*, 104(9), 918-925. <https://doi.org/10.1094/PHTO-03-13-0069-R>
9. Buck, K. W., Chain, E. B., & Himmelweit, F. (1971). Comparison of interferon induction in mice by purified *Penicillium chrysogenum* virus and derived double-stranded RNA. *The Journal of general virology*, 12(2), 131-139. <https://doi.org/10.1099/0022-1317-12-2-131>
10. Callanan, J., Stockdale, S. R., Shkoporov, A., Draper, L. A., Ross, R. P., & Hill, C. (2018). RNA phage biology in a metagenomic Era. En *Viruses* (Vol. 10, Número 7). MDPI AG. <https://doi.org/10.3390/v10070386>
11. Castón, J. R., Trus, B. L., Booy, F. P., Wickner, R. B., Wall, J. S., & Steven, A. C. (1997). Structure of L-A virus: A specialized compartment for the transcription and replication of double-stranded RNA. *Journal of Cell Biology*, 138(5), 975-985. <https://doi.org/10.1083/jcb.138.5.975>
12. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4), 540-552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
13. Chu, Y. M., Lim, W. S., Yea, S. J., Cho, J. D., Lee, Y. W., & Kim, K. H. (2004). Complexity of dsRNA Mycovirus Isolated from *Fusarium graminearum*. *Virus Genes*, 28(1), 135-143. <https://doi.org/10.1023/B:VIRU.0000012270.67302.35>
14. Daboussi, M.-J., & Capy, P. (2003). Transposable Elements in Filamentous Fungi. *Annual Review of Microbiology*, 57(1), 275-299. <https://doi.org/10.1146/annurev.micro.57.030502.091029>
15. Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., Taly, J. F., & Notredame, C. (2011). T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research*, 39(SUPPL. 2), 13-17. <https://doi.org/10.1093/nar/gkr245>
16. Domingo, E., & Holland, J. J. (1997). Rna Virus Mutations and Fitness for Survival. *Annual Review of Microbiology*, 51(1), 151-178. <https://doi.org/10.1146/annurev.micro.51.1.151>
17. Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792-1797. <https://doi.org/10.1093/nar/gkh340>
18. El Omari, K., Sutton, G., Ravantti, J. J., Zhang, H., Walter, T. S., Grimes, J. M., Bamford, D. H., Stuart, D. I., & Mancini, E. J. (2013). Plate tectonics of virus shell assembly and reorganization in phage $\phi 8$, a distant relative of mammalian reoviruses. *Structure*, 21(8), 1384-1395. <https://doi.org/10.1016/j.str.2013.06.017>
19. Elena, S., Miralles, R., Cuevas, J., Turner, P., & Moya, A. (2000). The Two Faces of Mutation: Extinction and Adaptation in RNA Viruses. *IUBMB Life*, 49(1), 5-9. <https://doi.org/10.1080/713803585>
20. Ellis, L. F., & Kleinschmidt, W. J. (1967). Virus-like particles of a fraction of statolon, a mould product. En *Nature* (Vol. 215, Número 5101, pp. 649-650). Nature Publishing Group. <https://doi.org/10.1038/215649a0>
21. Ezawa, T., Ikeda, Y., Shimura, H., & Masuta, C. (2015). Detection and characterization of mycoviruses in arbuscular mycorrhizal fungi by deep-sequencing. *Methods in Molecular Biology*, 1236, 171-180. https://doi.org/10.1007/978-1-4939-1743-3_13

22. Fauquet, C., Mayo, M. A., Maniloff, J., Desselberger, U., & Ball, L. A. (2005). Virus taxonomy: Eighth report of the international committee on taxonomy of viruses. En *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*. Academic Press.
23. Forrest, J. C., & Dermody, T. S. (2003). Reovirus Receptors and Pathogenesis. *Journal of Virology*, 77(17), 9109-9115. <https://doi.org/10.1128/jvi.77.17.9109-9115.2003>
24. Gao, Z., Wu, J., Jiang, D., Xie, J., Cheng, J., & Lin, Y. (2020). ORF i of mycovirus SsNSRV-1 is associated with debilitating symptoms of *Sclerotinia sclerotiorum*. *Viruses*, 12(4). <https://doi.org/10.3390/v12040456>
25. Ghabrial, S. A. (1994). New Developments in Fungal Virology. En *Advances in Virus Research* (Vol. 43, Número C, pp. 303-388). [https://doi.org/10.1016/S0065-3527\(08\)60052-4](https://doi.org/10.1016/S0065-3527(08)60052-4)
26. Ghabrial, S. A. (1998). Origin, adaptation and evolutionary pathways of fungal viruses. En *Virus Genes* (Vol. 16, Número 1, pp. 119-131). <https://doi.org/10.1023/A:1007966229595>
27. Ghabrial, S. A., & Suzuki, N. (2008). Fungal Viruses. En *Encyclopedia of Virology* (pp. 284-291). Elsevier Ltd. <https://doi.org/10.1016/B978-012374410-4.00563-X>
28. Ghabrial, S. A., & Suzuki, N. (2009). Viruses of Plant Pathogenic Fungi. *Annual Review of Phytopathology*, 47(1), 353-384. <https://doi.org/10.1146/annurev-phyto-080508-081932>
29. Ghabrial, S. A., Castón, J. R., Jiang, D., Nibert, M. L., & Suzuki, N. (2015). 50-Plus Years of Fungal Viruses. *Virology*, 479-480, 356-368. <https://doi.org/10.1016/j.virol.2015.02.034>
30. Ghabrial, S. A., Castón, J. R., Coutts, R. H. A., Hillman, B. I., Jiang, D., Kim, D. H., & Moriyama, H. (2018). ICTV Virus taxonomy profile: Chrysoviridae. *Journal of General Virology*, 99(1), 19-20. <https://doi.org/10.1099/jgv.0.000994>
31. Grogan, R. G., & Campbell, R. N. (1966). Fungi as Vectors and Hosts of Viruses. *Annual Review of Phytopathology*, 4(1), 29-52. <https://doi.org/10.1146/annurev.py.04.090166.000333>
32. Hatfull, G. F. (2018). Mycobacteriophages. *Microbiology Spectrum*, 6(5). <https://doi.org/10.1128/microbiolspec.GPP3-0026-2018>
33. Hillman, B. I., & Suzuki, N. (2004). Viruses of the Chestnut Blight Fungus, *Cryphonectria parasitica*. *Advances in Virus Research*, 63, 423-472. [https://doi.org/10.1016/S0065-3527\(04\)63007-7](https://doi.org/10.1016/S0065-3527(04)63007-7)
34. Hollings, M. (1962). Viruses associated with a die-back disease of cultivated mushroom. *Nature*, 196(4858), 962-965. <https://doi.org/10.1038/196962a0>
35. Howitt, R. L. J., Beever, R. E., Pearson, M. N., & Forster, R. L. S. (2006). Genome characterization of a flexuous rod-shaped mycovirus, Botrytis virus X, reveals high amino acid identity to genes from plant «potex-like» viruses. *Archives of Virology*, 151(3), 563-579. <https://doi.org/10.1007/s00705-005-0621-y>
36. Huson, D. H. (1998). SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics*, 14(1), 68-73. <https://doi.org/10.1093/bioinformatics/14.1.68>
37. Huson, D. H., Rupp, R., & Scornavacca, C. (2010). Phylogenetic networks: Concepts, algorithms and applications. En *Phylogenetic Networks: Concepts, Algorithms and Application* (pp. 218-219). Cambridge University Press, Cambridge, UK. <https://doi.org/10.1017/CBO9780511974076>
38. Huson, D. H., & Scornavacca, C. (2012). Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology*, 61(6), 1061-1067. <https://doi.org/10.1093/sysbio/sys062>
39. Hyder, R., Pennanen, T., Hamberg, L., Vainio, E. J., Piri, T., & Hantula, J. (2013). Two viruses of *Heterobasidion confer* beneficial, cryptic or detrimental effects to their hosts in different situations. *Fungal Ecology*, 6(5), 387-396. <https://doi.org/10.1016/j.funeco.2013.05.005>
40. Isaacs, A. ; Lindenmann, J. (1957). Virus interference. I. The interferon. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, 147(927), 258-267. <https://doi.org/10.1098/rspb.1957.0048>
41. Itakura, E., Sawada, I., & Matsuura, A. (2005). Dimerization of the ATRIP protein through the coiled-coil motif and its implication to the maintenance of stalled replication forks. *Molecular Biology of the Cell*, 16(12), 5551-5562. <https://doi.org/10.1091/mbc.E05-05-0427>
42. Kamaruzzaman, M., He, G., Wu, M., Zhang, J., Yang, L., Chen, W., & Li, G. (2019). A novel partitivirus in the hypovirulent isolate qt5-19 of the plant pathogenic fungus *botrytis cinerea*. *Viruses*, 11(1). <https://doi.org/10.3390/v11010024>
43. Kanhayuwa, L., Kotta-Loizou, I., Özkan, S., Gunning, A. P., & Coutts, R. H. A. (2015). A novel mycovirus from *Aspergillus fumigatus* contains four unique dsRNAs as its genome and is infectious as dsRNA. *Proceedings of the National Academy of Sciences of the United States of America*, 112(29), 9100-9105. <https://doi.org/10.1073/pnas.1419225112>
44. Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780. <https://doi.org/10.1093/molbev/mst010>
45. Katoh, K., Rozewicki, J., & Yamada, K. D. (2018). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20(4), 1160-1166. <https://doi.org/10.1093/bib/bbx108>

46. Kawabata, T., & Nishikawa, K. (2000). Protein structure comparison using the Markov transition model of evolution. *Proteins: Structure, Function and Genetics*, 41(1), 108-122. [https://doi.org/10.1002/1097-0134\(20001001\)41:1<108::AID-PROT130>3.0.CO;2-S](https://doi.org/10.1002/1097-0134(20001001)41:1<108::AID-PROT130>3.0.CO;2-S)
47. Kawabata, T. (2003). MATRAS: A program for protein 3D structure comparison. *Nucleic Acids Research*, 31(13), 3367-3369. <https://doi.org/10.1093/nar/gkg581>
48. Koloniuk, I., Hrabáková, L., & Petrzik, K. (2015). Molecular characterization of a novel amalgavirus from the entomopathogenic fungus *Beauveria bassiana*. *Archives of Virology*, 160(6), 1585-1588. <https://doi.org/10.1007/s00705-015-2416-0>
49. Koonin, E. V., Gorbalenya, A. E., & Chumakov, K. M. (1989). Tentative identification of RNA-dependent RNA polymerases of dsRNA viruses and their relationship to positive strand RNA viral polymerases. *FEBS Letters*, 252(1-2), 42-46. [https://doi.org/10.1016/0014-5793\(89\)80886-5](https://doi.org/10.1016/0014-5793(89)80886-5)
50. Koonin, E. V., & Dolja, V. V. (2014). Virus World as an Evolutionary Network of Viruses and Capsidless Selfish Elements. *Microbiology and Molecular Biology Reviews*, 78(2), 278-303. <https://doi.org/10.1128/membr.00049-13>
51. Koonin, E. V., Dolja, V. V., & Krupovic, M. (2015). Origins and evolution of viruses of eukaryotes: The ultimate modularity. En *Virology* (Vols. 479-480, pp. 2-25). Academic Press Inc. <https://doi.org/10.1016/j.virol.2015.02.039>
52. Krupovic, M., Dolja, V. V., & Koonin, E. V. (2015). Plant viruses of the Amalgaviridae family evolved via recombination between viruses with double-stranded and negative-strand RNA genomes. *Biology Direct*, 10(1), 12. <https://doi.org/10.1186/s13062-015-0047-8>
53. Krupovic, M., Ghabrial, S. A., Jiang, D., & Varsani, A. (2016). Genomoviridae: a new family of widespread single-stranded DNA viruses. *Archives of Virology*, 161(9), 2633-2643. <https://doi.org/10.1007/s00705-016-2943-3>
54. Kumar, S., Subbarao, B. L., Kumari, R., & Hallan, V. (2017). Molecular characterization of a novel cryptic virus infecting pigeonpea plants. *PLoS ONE*, 12(8). <https://doi.org/10.1371/journal.pone.0181829>
55. Kwon, S. J., Lim, W. S., Park, S. H., Park, M. R., & Kim, K. H. (2009). Molecular characterization of a dsRNA mycovirus, *Fusarium graminearum* Virus-DK21, which is phylogenetically related to hypoviruses but has a genome organization and gene expression strategy resembling those of plant potex-like viruses. *Molecules and Cells*, 28(1), 73-74. <https://doi.org/10.1007/s10059-009-0112-1>
56. Le, S. Q., & Gascuel, O. (2008). An improved general amino acid replacement matrix. *Molecular Biology and Evolution*, 25(7), 1307-1320. <https://doi.org/10.1093/molbev/msn067>
57. Lin, Y. H., Chiba, S., Tani, A., Kondo, H., Sasaki, A., Kanematsu, S., & Suzuki, N. (2012). A novel quadripartite dsRNA virus isolated from a phytopathogenic filamentous fungus, *Rosellinia necatrix*. *Virology*, 426(1), 42-50. <https://doi.org/10.1016/j.virol.2012.01.013>
58. Lin, Y., Zhang, H., Zhao, C., Liu, S., & Guo, L. (2014). The complete genome sequence of a novel mycovirus from *Alternaria longipes* strain HN28. *Archives of Virology*, 160(2), 577-580. <https://doi.org/10.1007/s00705-014-2218-9>
59. Lindberg, G. D. (1959). A transmissible disease of *Helminthosporium victoriae*. *Phytopathology*, 49(1), 29-32.
60. Liu, H., Fu, Y., Jiang, D., Li, G., Xie, J., Peng, Y., Yi, X., & Ghabrial, S. A. (2009). A Novel Mycovirus That Is Related to the Human Pathogen Hepatitis E Virus and Rubi-Like Viruses. *Journal of Virology*, 83(4), 1981-1991. <https://doi.org/10.1128/jvi.01897-08>
61. Liu, H., Fu, Y., Jiang, D., Li, G., Xie, J., Cheng, J., Peng, Y., Ghabrial, S. A., & Yi, X. (2010). Widespread Horizontal Gene Transfer from Double-Stranded RNA Viruses to Eukaryotic Nuclear Genomes. *Journal of Virology*, 84(22), 11876-11887. <https://doi.org/10.1128/jvi.00955-10>
62. Liu, H., Fu, Y., Xie, J., Cheng, J., Ghabrial, S. A., Li, G., Peng, Y., Yi, X., & Jiang, D. (2012). Evolutionary genomics of mycovirus-related dsRNA viruses reveals cross-family horizontal gene transfer and evolution of diverse viral lineages. *BMC Evolutionary Biology*, 12(1), 91. <https://doi.org/10.1186/1471-2148-12-91>
63. Liu, L., Xie, J., Cheng, J., Fu, Y., Li, G., Yi, X., & Jiang, D. (2014). Fungal negative-stranded RNA virus that is related to bornaviruses and nyaviruses. *Proceedings of the National Academy of Sciences of the United States of America*, 111(33), 12205-12210. <https://doi.org/10.1073/pnas.1401786111>
64. Liu, C., Li, M., Redda, E. T., Mei, J., Zhang, J., Wu, B., & Jiang, X. (2019). A novel double-stranded RNA mycovirus isolated from *Trichoderma harzianum*. *Virology Journal*, 16(1), 113. <https://doi.org/10.1186/s12985-019-1213-x>
65. Luque, D., González, J. M., Garriga, D., Ghabrial, S. A., Havens, W. M., Trus, B., Verdaguer, N., Carrascosa, J. L., & Castón, J. R. (2010). The T=1 Capsid Protein of *Penicillium chrysogenum* Virus Is Formed by a Repeated Helix-Rich Core Indicative of Gene Duplication. *Journal of Virology*, 84(14), 7256-7266. <https://doi.org/10.1128/jvi.00432-10>
66. Luque, D., Gómez-Blanco, J., Garriga, D., Brilot, A. F., González, J. M., Havens, W. M., Carrascosa, J. L., Trus, B. L., Verdaguer, N., Ghabrial, S. A., & Castón, J. R. (2014). Cryo-EM near-atomic structure of a dsRNA

- fungal virus shows ancient structural motifs preserved in the dsRNA viral lineage. *Proceedings of the National Academy of Sciences of the United States of America*, 111(21), 7641-7646. <https://doi.org/10.1073/pnas.1404330111>
67. Luque, D., Mata, C. P., González-Camacho, F., González, J. M., Gómez-Blanco, J., Alfonso, C., Rivas, G., Havens, W. M., Kanematsu, S., Suzuki, N., Ghabrial, S. A., Trus, B. L., & Castón, J. R. (2016). Heterodimers as the Structural Unit of the T=1 Capsid of the Fungal Double-Stranded RNA Rosellinia necatrix Quadrivirus 1. *Journal of Virology*, 90(24), 11220-11230. <https://doi.org/10.1128/jvi.01013-16>
68. Luque, D., Mata, C. P., Suzuki, N., Ghabrial, S. A., & Castón, J. R. (2018). Capsid structure of dsRNA fungal viruses. En *Viruses* (Vol. 10, Número 9). MDPI AG. <https://doi.org/10.3390/v10090481>
69. Marcou, D. (1961). Notion de longévité et nature cytoplasmique du déterminant de sénescence chez quelques champignons. En *Ann. Sci. Natur. Bot.* (Vol. 11, pp. 653-764).
70. Márquez, L. M., Redman, R. S., Rodriguez, R. J., & Roossinck, M. J. (2007). A virus in a fungus in a plant: Three-way symbiosis required for thermal tolerance. *Science*, 315(5811), 513-515. <https://doi.org/10.1126/science.1136237>
71. Marzano, S.-Y. L., Nelson, B. D., Ajayi-Oyetunde, O., Bradley, C. A., Hughes, T. J., Hartman, G. L., Eastburn, D. M., & Domier, L. L. (2016). Identification of Diverse Mycoviruses through Metatranscriptomics Characterization of the Viromes of Five Major Fungal Plant Pathogens. *Journal of Virology*, 90(15), 6846-6863. <https://doi.org/10.1128/jvi.00357-16>
72. Mata, C. P., Luque, D., Gómez-Blanco, J., Rodríguez, J. M., González, J. M., Suzuki, N., Ghabrial, S. A., Carrascosa, J. L., Trus, B. L., & Castón, J. R. (2017). Acquisition of functions on the outer capsid surface during evolution of double-stranded RNA fungal viruses. *PLoS Pathogens*, 13(12). <https://doi.org/10.1371/journal.ppat.1006755>
73. Muro, E. M., & Andrade-Navarro, M. A. (2010). Pseudogenes as an alternative source of natural antisense transcripts. *BMC Evolutionary Biology*, 10(1), 338. <https://doi.org/10.1186/1471-2148-10-338>
74. Naitow, H., Tang, J., Canady, M., Wickner, R. B., & Johnson, J. E. (2002). L-a virus at 3.4 Å resolution reveals particle architecture and mrna decapping mechanism. *Nature Structural Biology*, 9(10), 725-728. <https://doi.org/10.1038/nsb844>
75. Nerva, L., Varese, G. C., Falk, B. W., & Turina, M. (2017). Mycoviruses of an endophytic fungus can replicate in plant cells: Evolutionary implications. *Scientific Reports*, 7(1), 1-11. <https://doi.org/10.1038/s41598-017-02017-3>
76. Nibert, M. L., Tang, J., Xie, J., Collier, A. M., Ghabrial, S. A., Baker, T. S., & Tao, Y. J. (2013). 3D Structures of Fungal Partitiviruses. En *Advances in Virus Research* (Vol. 86, pp. 59-85). Academic Press Inc. <https://doi.org/10.1016/B978-0-12-394315-6.00003-9>
77. Nibert, M. L., Ghabrial, S. A., Maiss, E., Lesker, T., Vainio, E. J., Jiang, D., & Suzuki, N. (2014). Taxonomic reorganization of family Partitiviridae and other recent progress in partitivirus research. En *Virus Research* (Vol. 188, pp. 128-141). Elsevier. <https://doi.org/10.1016/j.virusres.2014.04.007>
78. Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, 302(1), 205-217. <https://doi.org/10.1006/jmbi.2000.4042>
79. Novais, C. B. de, Pepe, A., Siqueira, J. O., Giovannetti, M., & Sbrana, C. (2017). Compatibility and incompatibility in hyphal anastomosis of arbuscular mycorrhizal fungi. En *Scientia Agricola* (Vol. 74, Número 5, pp. 411-416). Scientia Agricola. <https://doi.org/10.1590/1678-992x-2016-0243>
80. Nuss D. L. (2011). Mycoviruses, RNA silencing, and viral RNA recombination. *Advances in virus research*, 80, 25-48. <https://doi.org/10.1016/B978-0-12-385987-7.00002-6>
81. Oh, C. S., & Hillman, B. I. (1995). Genome organization of a partitivirus from the filamentous ascomycete *Atkinsonella hypoxylon*. *Journal of General Virology*, 76(6), 1461-1470. <https://doi.org/10.1099/0022-1317-76-6-1461>
82. Ong, J. W. L., Li, H., Sivasithamparam, K., Dixon, K. W., Jones, M. G. K., & Wylie, S. J. (2017). The challenges of using high-throughput sequencing to track multiple bipartite mycoviruses of wild orchid-fungus partnerships over consecutive years. *Virology*, 510, 297-304. <https://doi.org/10.1016/j.virol.2017.07.031>
83. Onodera, S., Qiao, X., Gottlieb, P., Strassman, J., Frilander, M., & Mindich, L. (1993). RNA structure and heterologous recombination in the double-stranded RNA bacteriophage phi 6. *Journal of Virology*, 67(8), 4914-4922. <https://doi.org/10.1128/jvi.67.8.4914-4922.1993>
84. Özkan, S., & Coutts, R. H. A. (2015). *Aspergillus fumigatus* mycovirus causes mild hypervirulent effect on pathogenicity when tested on *Galleria mellonella*. *Fungal Genetics and Biology*, 76, 20-26. <https://doi.org/10.1016/j.fgb.2015.01.003>
85. Pan, J., Dong, L., Lin, L., Ochoa, W. F., Sinkovits, R. S., Havens, W. M., Nibert, M. L., Baker, T. S., Ghabrial, S. A., & Tao, Y. J. (2009). Atomic structure reveals the unique capsid organization of a dsRNA virus. *Proceedings of the National Academy of Sciences of the United States of America*, 106(11), 4225-4230. <https://doi.org/10.1073/pnas.0812071106>

- 86.** Park, C. M., Banerjee, N., Koltin, Y., & Bruenn, J. A. (1996). The *Ustilago maydis* virally encoded KP1 killer toxin. *Molecular Microbiology*, *20*(5), 957-963. <https://doi.org/10.1111/j.1365-2958.1996.tb02537.x>
- 87.** Pearson, M. N., Beever, R. E., Boine, B., & Arthur, K. (2009). Mycoviruses of filamentous fungi and their relevance to plant pathology. *Molecular plant pathology*, *10*(1), 115–128. <https://doi.org/10.1111/j.1364-3703.2008.00503.x>
- 88.** Powell, H. M., Culbertson, C. G., McGuire, J. M., Hoehn, M. M., & Baker, L. A. (1952). A filtrate with chemoprophylactic and chemotherapeutic action against MM and Semliki Forest viruses in mice. *Antibiotics & Chemotherapy (Northfield, Ill.)*, *2*(8), 432-434.
- 89.** Rao, Q., Liu, M., Tian, Y., Wu, Z., Hao, Y., Song, L., Qin, Z., Ding, C., Wang, H. W., Wang, J., & Xu, Y. (2018). Cryo-EM structure of human ATR-ATRIP complex. *Cell Research*, *28*(2), 143-156. <https://doi.org/10.1038/cr.2017.158>
- 90.** Riley, R., Haridas, S., Wolfe, K. H., Lopes, M. R., Hittinger, C. T., Göker, M., Salamov, A. A., Wisecaver, J. H., Long, T. M., Calvey, C. H., Aerts, A. L., Barry, K. W., Choi, C., Clum, A., Coughlan, A. Y., Deshpande, S., Douglass, A. P., Hanson, S. J., Klenk, H. P., Jeffries, T. W. (2016). Comparative genomics of biotechnologically important yeasts. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(35), 9882-9887. <https://doi.org/10.1073/pnas.1603941113>
- 91.** Roberts, T. C., & Morris, K. V. (2013). Not so pseudo anymore: pseudogenes as therapeutic targets. *Pharmacogenomics*, *14*(16), 2023–2034. <https://doi.org/10.2217/pgs.13.172>
- 92.** Romaine, C. P., & Schlagnhauser, B. (1995). PCR analysis of the viral complex associated with La France disease of *Agaricus bisporus*. *Applied and Environmental Microbiology*, *61*(6), 2322-2325. <https://doi.org/10.1128/aem.61.6.2322-2325.1995>
- 93.** Roossinck M. J. (2015). Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. *Frontiers in microbiology*, *5*, 767. <https://doi.org/10.3389/fmicb.2014.00767>
- 94.** Roy, A., Yang, J., & Zhang, Y. (2012). COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Research*, *40*(W1), 471-477. <https://doi.org/10.1093/nar/gks372>
- 95.** Sabanadzovic, S., Valverde, R. A., Brown, J. K., Martin, R. R., & Tzanetakis, I. E. (2009). Southern tomato virus: The link between the families Totiviridae and Partitiviridae. *Virus Research*, *140*(1-2), 130-137. <https://doi.org/10.1016/j.virusres.2008.11.018>
- 96.** Sahin, E., & Akata, I. (2018). Viruses infecting macrofungi. *VirusDisease*, *29*(1), 1–18. <https://doi.org/10.1007/s13337-018-0434-8>
- 97.** Sato, Y., Miyazaki, N., Kanematsu, S., Xie, J., Ghabrial, S. A., Hillman, B. I., Suzuki, N., & Ictv Report Consortium. (2019). ICTV Virus Taxonomy Profile: Megabirnaviridae. *The Journal of general virology*, *100*(9), 1269-1270. <https://doi.org/10.1099/jgv.0.001297>
- 98.** Sayers, E. W., Beck, J., Brister, J. R., Bolton, E. E., Canese, K., Comeau, D. C., Funk, K., Ketter, A., Kim, S., Kimchi, A., Kitts, P. A., Kuznetsov, A., Lathrop, S., Lu, Z., McGarvey, K., Madden, T. L., Murphy, T. D., O’Leary, N., Phan, L., ... Ostell, J. (2020). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *48*(D1), D9-D16. <https://doi.org/10.1093/nar/gkz899>
- 99.** Schmidt, H. A., Strimmer, K., Vingron, M., & Von Haeseler, A. (2002). TREE-PUZZLE: Maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, *18*(3), 502-504. <https://doi.org/10.1093/bioinformatics/18.3.502>
- 100.** Schneider, W. L., & Roossinck, M. J. (2001). Genetic Diversity in RNA Virus Quasispecies Is Controlled by Host-Virus Interactions. *Journal of Virology*, *75*(14), 6566-6571. <https://doi.org/10.1128/jvi.75.14.6566-6571.2001>
- 101.** Shah, U., Kotta-Loizou, I., Fitt, B., & Coutts, R. (2019). Mycovirus induced hypervirulence of *Leptosphaeria biglobosa* enhances systemic acquired resistance to *Leptosphaeria maculans* in *Brassica napus*. *Access Microbiology*, *1*(1A), 787. <https://doi.org/10.1099/acmi.ac2019.po0506>
- 102.** Sharma, M., Guleria, S., Singh, K., Chauhan, A., & Kulshrestha, S. (2018). Mycovirus associated hypovirulence, a potential method for biological control of *Fusarium* species. *VirusDisease*, *29*(2), 134–140. <https://doi.org/10.1007/s13337-018-0438-4>
- 103.** Shope, R. E. (1953). An antiviral substance from *Penicillium funiculosum*. I. Effect upon infection in mice with swine influenza virus and Columbia SK encephalomyelitis virus. *The Journal of experimental medicine*, *97*(5), 601-625. <https://doi.org/10.1084/jem.97.5.601>
- 104.** Shope, R. E. (1953). An antiviral substance from *Penicillium funiculosum*. III. General properties and characteristics of helenine. *The Journal of experimental medicine*, *97*(5), 639-650. <https://doi.org/10.1084/jem.97.5.639>
- 105.** Sinden, J. W.; Hauser, E. (1948). Report on Two New Mushroom Diseases. *Mushroom Science*, *1*, 96-100.
- 106.** Soldevila, A. I., & Ghabrial, S. A. (2000). Expression of the Totivirus *Helminthosporium victoriae* 190S Virus RNA-Dependent RNA Polymerase from Its Downstream Open Reading Frame in Dicistronic Constructs. *Journal of Virology*, *74*(2), 997-1003. <https://doi.org/10.1128/jvi.74.2.997-1003.2000>

107. Son, M., Yu, J., & Kim, K. H. (2015). Five Questions about Mycoviruses. *PLoS pathogens*, 11(11), e1005172. <https://doi.org/10.1371/journal.ppat.1005172>
108. Strimmer, K., & Von Haeseler, A. (1997). Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences of the United States of America*, 94(13), 6815-6819. <https://doi.org/10.1073/pnas.94.13.6815>
109. Sun, L., & Suzuki, N. (2008). Intragenic rearrangements of a mycoreovirus induced by the multifunctional protein p29 encoded by the prototypic hypovirus CHV1-EP713. *Rna*, 14(12), 2557-2571. <https://doi.org/10.1261/rna.1125408>
110. Suzaki, K., Ikeda, K. I., Sasaki, A., Kanematsu, S., Matsumoto, N., & Yoshida, K. (2005). Horizontal transmission and host-virulence attenuation of totivirus in violet root rot fungus *Helicobasidium mompa*. *Journal of General Plant Pathology*, 71(3), 161-168. <https://doi.org/10.1007/s10327-005-0181-8>
111. Tan, C. M. C., Pearson, M. N., Beever, R. E., & Parkes, S. L. (2007). Why fungi have sex? *XIVth International Botrytis Symposium*, 26.
112. Tang, J., Pan, J., Havens, W. M., Ochoa, W. F., Guu, T. S. Y., Ghabrial, S. A., Nibert, M. L., Tao, Y. J., & Baker, T. S. (2010). Backbone trace of partitivirus capsid protein from electron cryomicroscopy and homology modeling. *Biophysical Journal*, 99(2), 685-694. <https://doi.org/10.1016/j.bpj.2010.04.058>
113. Thapa, V., & Roossinck, M. J. (2019). Determinants of coinfection in the mycoviruses. *Frontiers in Cellular and Infection Microbiology*, 9(MAY), 169. <https://doi.org/10.3389/fcimb.2019.00169>
114. Vainio, E. J., Chiba, S., Ghabrial, S. A., Maiss, E., Roossinck, M., Sabanadzovic, S., Suzuki, N., Xie, J., & Nibert, M. (2018). ICTV virus taxonomy profile: Partitiviridae. *Journal of General Virology*, 99(1), 17-18. <https://doi.org/10.1099/jgv.0.000985>
115. van der Burgt, A., Karimi Jashni, M., Bahkali, A. H., & De Wit, P. J. G. M. (2014). Pseudogenization in pathogenic fungi with different host plants and lifestyles might reflect their evolutionary past. *Molecular Plant Pathology*, 15(2), 133-144. <https://doi.org/10.1111/mpp.12072>
116. Weber, F., Wagner, V., Rasmussen, S. B., Hartmann, R., & Paludan, S. R. (2006). Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses. *Journal of Virology*, 80(10), 5059-5064. <https://doi.org/10.1128/jvi.80.10.5059-5064.2006>
117. Whelan, S., & Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution*, 18(5), 691-699. <https://doi.org/10.1093/oxfordjournals.molbev.a003851>
118. Wieringa, K. T., & Wiebols, G. L. W. (1936). De aardappelschurft en de heterolyse der schurftparasiet. *Tijdschrift Over Plantenziekten*, 42(5), 235-240. <https://doi.org/10.1007/BF02812401>
119. Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, 43(W1), W174-W181. <https://doi.org/10.1093/nar/gkv342>
120. Yang, Y., Lu, S., Shen, W., Zhao, X., Shen, M., Tan, Y., Li, G., Li, M., Wang, J., Hu, F., & Le, S. (2016). Characterization of the first double-stranded RNA bacteriophage infecting *Pseudomonas aeruginosa*. *Scientific Reports*, 6. <https://doi.org/10.1038/srep38795>
121. Yu, X., Li, B., Fu, Y., Jiang, D., Ghabrial, S. A., Li, G., Peng, Y., Xie, J., Cheng, J., Huang, J., & Yi, X. (2010). A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(18), 8387-8392. <https://doi.org/10.1073/pnas.0913535107>
122. Yu, X., Li, B., Fu, Y., Xie, J., Cheng, J., Ghabrial, S. A., Li, G., Yi, X., & Jiang, D. (2013). Extracellular transmission of a DNA mycovirus and its use as a natural fungicide. *Proceedings of the National Academy of Sciences of the United States of America*, 110(4), 1452-1457. <https://doi.org/10.1073/pnas.1213755110>
123. Zannotto, P. M., Gibbs, M. J., Gould, E. A., & Holmes, E. C. (1996). A reevaluation of the higher taxonomy of viruses based on RNA polymerases. *Journal of virology*, 70(9), 6083-6096. <https://doi.org/10.1128/jvi.70.9.6083-6096.1996>
124. Zhang, Y., & Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, 33(7), 2302-2309. <https://doi.org/10.1093/nar/gki524>
125. Zhang, Y. (2009). I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins: Structure, Function and Bioinformatics*, 77(SUPPL. 9), 100-113. <https://doi.org/10.1002/prot.22588>
126. Zhang, R., Liu, S., Chiba, S., Kondo, H., Kanematsu, S., & Suzuki, N. (2014). A novel single-stranded RNA virus isolated from a phytopathogenic filamentous fungus, *Rosellinia necatrix*, with similarity to hypo-like viruses. *Frontiers in Microbiology*, 5(JULY). <https://doi.org/10.3389/fmicb.2014.00360>
127. Zheng, L., Liu, H., Zhang, M., Cao, X., & Zhou, E. (2013). The complete genomic sequence of a novel mycovirus from *Rhizoctonia solani* AG-1 IA strain B275. *Archives of Virology*, 158(7), 1609-1612. <https://doi.org/10.1007/s00705-013-1637-3>