

# UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA  
AGRONÒMICA I DEL MEDI NATURAL



## ***IDENTIFICACIÓN Y ANÁLISIS FILOGENÉTICO DE LOS DOMINIOS ALKB DE LAS REPLICASAS DE VIRUS DE PLANTAS***

TRABAJO DE FIN DE GRADO

AUTOR: D. Pau Rubio Costa

TUTOR EXPERIMENTAL: Dr. Vicente Pallás Benet  
TUTOR UPV: Prof. Frederic Aparicio Herrero

València, 26 de Julio de 2021



## **Título**

Identificación y análisis filogenéticos de los dominios AlkB de las replicasas de virus de plantas.

## **Resumen**

La metilación del nitrógeno en posición 6 de la adenosina (m6A) es la modificación interna más abundante de mRNAs en eucariotas. Las proteínas conocidas como m6A *writers* y *erasers* (*escritoras* y *borradoras*) definen el estado de metilación de un mRNA, mientras que las m6A *readers* (*lectoras*) reconocen este nucleótido modificado y controlan el destino de los mRNAs metilados. Aunque el estudio de la modificación del ARNm está todavía en sus inicios, el intenso interés que despierta su participación en importantes procesos biológicos, así como en patologías como el cáncer, ha llevado a rápidos avances en la comprensión de cómo las modificaciones del ARNm controlan la expresión génica en los sistemas de mamíferos. En mamíferos, las *erasers* pertenecen a la familia de proteínas AlkB (oxigenasas dependientes de 2-oxoglutarato y Fe (II) ), que en *Escherichia coli* se ha demostrado que protege los ácidos nucleicos contra el daño por metilación, mientras que las *readers* son proteínas de unión al RNA que contienen un dominio llamado YT521-B homolog (YTH). Tal y como ocurre en animales, recientemente se ha demostrado que la modificación m6A de los mRNAs en plantas es esencial para su desarrollo, así como en la respuesta a estreses abióticos y bióticos. Algunos virus de plantas, principalmente los pertenecientes a las familias *Alphaflexiviridae*, *Betaflexiviridae* and *Closteroviridae*, contienen un dominio AlkB en la secuencia de sus correspondientes replicasas. Su función en el ciclo de infección de estos virus es todavía desconocida pero su conservación en los virus que usualmente infectan huéspedes leñosos hace sugerir un papel relevante en las interacciones virus- huésped. En este trabajo se tratará de identificar este dominio en los virus caracterizados molecularmente en los últimos años y establecer una posible correlación entre su presencia y su tropismo de huésped. Por otra parte, se estudiarán las relaciones filogenéticas de los distintos géneros de virus que presentan dicho dominio y se compararán con las más clásicas basadas en las relaciones filogenéticas de la proteína de cubierta.

**Palabras clave:** virus de plantas; proteínas de cubierta; desmetilasas de RNA; modificación de RNA; m6A; Betaflexiviridae; familia AlkB.

**Alumno/a:** D. Pau Rubio Costa

**Tutor experimental:** Dr. Vicente Pallás Benet

**Tutor UPV:** Prof. Frederic Aparicio Herrero

València, 26 de Julio de 2021

**Title**

Identification and phylogenetic analysis of AlkB domains of plant virus replicases.

**Abstract**

Methylation of the nitrogen at position 6 of adenosine (m<sup>6</sup>A) is the most abundant internal modification of mRNAs in eukaryotes. Proteins known as m<sup>6</sup>A writers and erasers define the methylation state of an mRNA, while m<sup>6</sup>A readers recognise this modified nucleotide and control the fate of methylated mRNAs. Although the field of mRNA modification is young, the intense interest in its involvement in important biological processes, as well as in pathologies such as cancer, has led to rapid advances in the understanding how mRNA modifications control gene expression in mammalian systems. In mammals, erasers belong to the AlkB (2-oxoglutarate and Fe (II)-dependent oxygenase) family of proteins, which in *Escherichia coli* protects nucleic acids against methylation damage, while readers are RNA-binding proteins containing a domain called YT521-B homolog (YTH). As in animals, it has recently been shown that m<sup>6</sup>A modification of mRNAs in plants is essential for their development as well as in the response to abiotic and biotic stresses. Some plant viruses, mainly those belonging to the families *Alphaflexiviridae*, *Betaflexiviridae* and *Closteroviridae*, contain an AlkB domain in the sequence of their corresponding replicases. Their function in the infection cycle of these viruses is still unknown, but their conservation in viruses that usually infect woody hosts suggests a relevant role in virus-host interactions. In this work, we will try to identify this domain in viruses molecularly characterised in recent years and to establish a possible correlation between its presence and host tropism. On the other hand, the phylogenetic relationships of the different virus genera that present this domain will be studied and compared with the more classical ones based on the phylogenetic relationships of the coat protein.

**Key words:** plant viruses; coat proteins; demethylases; RNA modification; *Betaflexiviridae*; AlkB family.

**Student:** Mr. Pau Rubio Costa

**Experimental tutor:** Dr. Vicente Pallás Benet

**UPV tutor:** Prof. Frederic Aparicio Herrero

València, 26 de Julio de 2021

## **Dedicatorias y agradecimientos**

Me gustaría darle las gracias a mi tutor del TFG, el doctor Vicente Pallás, por su ayuda, por su trato siempre amable, por el tiempo que me ha dedicado y por darme la oportunidad de realizar este trabajo. También me gustaría mostrar mi agradecimiento al profesor Javier Forment, por toda la ayuda que me ha dado, por su paciencia, por haberme enseñado cual es mi vocación y por haber estado siempre cuando lo necesité.

Después de estos últimos años de carrera, no podía acabar sin mostrar también mi agradecimiento a todas aquellas personas que me han acompañado en este apasionante pero duro camino. Me gustaría sobre todo mostrar mi agradecimiento a mi familia, a mis amigos, y a mi pareja, los cuales han estado siempre a mi lado y me han dado la fuerza y la motivación para siempre seguir adelante. Sin su ayuda nada de esto hubiera sido posible.



# ÍNDICE

<b>1. Introducción</b> .....	1
1.1. Modificaciones de los ácidos nucleicos.....	1
1.2. N6-metiladenosina (m6A).....	2
1.2.1. Writers (Escritoras).....	3
1.2.2. Readers (Lectoras).....	4
1.2.3. Erasers (Borradoras).....	6
1.3. N6-metiladenosina en el ciclo de replicación viral.....	6
1.4. AlkB.....	8
1.4.1. La familia AlkB: contexto, historia y mecanismo molecular.....	8
1.4.2. AlkB en bacterias.....	10
1.4.3. AlkB en mamíferos.....	11
1.4.4. AlkB en plantas.....	12
1.5 Dominios AlkB en virus ssRNA+.....	13
1.5.1. Conocimiento actual.....	13
1.5.2. Familias de virus que contienen el dominio AlkB.....	14
1.5.2.1. <i>Closteroviridae</i> .....	14
1.5.2.2. <i>Alphaflexiviridae</i> .....	14
1.5.2.3. <i>Betaflexiviridae</i> .....	15
1.5.2.4. Otras familias.....	15
1.5.3. Hipótesis sobre su función.....	15
<b>2. Objetivos</b> .....	18

<b>3. Materiales y métodos</b> .....	20
3.1. Identificación y preparación de secuencias de proteínas virales que contienen el dominio AlkB.....	20
3.2. Alineamientos múltiple.....	21
3.3. Filogenia.....	22
<b>4. Resultados y discusión</b> .....	24
4.1. Identificación y recuento de especies virales con el dominio AlkB.....	24
4.2. Alineamiento múltiple de especies virales con el dominio AlkB.....	24
4.2.1. Residuos críticos para la funcionalidad de AlkB.....	24
4.2.1.1. Residuos que participan en la unión con el ion metálico.....	25
4.2.1.2. Residuos que participan en la unión del cofactor $\alpha$ -ketoglutarato.....	26
4.2.2. Otros residuos conservados dentro de los dominios AlkB virales.....	27
4.3. Alineamientos múltiples de aislados de virus que infectan árboles frutales.....	29
4.3.1. Apple stem pitting virus.....	29
4.3.2. Apple chlorotic leaf spot virus.....	30
4.3.3. Citrus leaf blotch virus .....	30
4.3.4. Cherry necrotic rusty mottle virus .....	30
4.3.5. Plum bark necrosis stem pitting associated virus.....	30
4.4. Filogenias.....	31
<b>5. Conclusión</b> .....	37
<b>6. Bibliografía</b> .....	39
<b>7. Anexo</b> .....	47

## Índice de figuras

**Figura 1.** Proteínas de regulación de la modificación m<sup>6</sup>A en humano.

**Figura 2.** Filogenia general de la familia AlkB.

**Figura 3.** Eliminación del grupo metilo en el DNA y RNA catalizada por los Homólogos AlkB humanos: ALKBH1, ALKBH2, ALKBH3, ALKBH5 y FTO.

**Figura 4.** Estructura multidominio de la ALKBH9B de *Arabidopsis thaliana*.

**Figura 5.** Valores de los parámetros seleccionados para la construcción de los alineamientos múltiples en el programa Mega-X.

**Figura 6.** Valores de los parámetros seleccionados para la construcción de los árboles filogenéticos en el programa Mega-X.

**Figura 7.** Alineamiento múltiple de los dominios AlkB de virus ssRNA<sup>+</sup> que infectan a plantas.

**Figura 8.** Árbol filogenético de los dominios AlkB virales.

**Figura 9.** Árbol filogenético de la proteína de cubierta.

**Tabla 1 del Anexo.** Tabla de virus de ssRNA de polaridad positiva encontrados que poseen el dominio AlkB, junto con su acrónimo, género, familia, y características del huésped.

**Figura 1 del Anexo.** Alineamiento múltiple del dominio AlkB de los aislados de Apple stem pitting virus.

**Figura 2 del Anexo.** Alineamiento múltiple del dominio AlkB de los aislados de Apple chlorotic leaf spot virus.

**Figura 3 del Anexo.** Alineamiento múltiple del dominio AlkB de los aislados de Citrus leaf blotch virus.

**Figura 4 del Anexo.** Alineamiento múltiple del dominio AlkB de los aislados de Cherry necrotic rusty mottle virus.

**Figura 5 del Anexo.** Alineamiento múltiple del dominio AlkB de los aislados de Plum bark necrosis stem pitting associated virus.



# 1. Introducción

## 1.1. Modificaciones de los ácidos nucleicos

Los ácidos nucleicos, tanto el RNA como el DNA, son polímeros clave en los organismos celulares (Grosjean, 2009). Estos codifican la información biológica que determina la composición de las proteínas y participan, gracias a sus distintas formas, en su ensamblaje con otras macromoléculas. (Belmont *et al.*, 2001). Desde el descubrimiento del DNA por Friedrich Miescher en el año 1869 (Weaver, 2012) hasta nuestros días, el estudio de los ácidos nucleicos ha sido esencial en la comprensión de los sistemas biológicos. Sin embargo, durante el periodo de 1920 a 1945, se creía que los ácidos nucleicos estaban formados únicamente por los 4 nucleósidos canónicos, fueran derivados ribonucleicos o desoxirribonucleicos (Grosjean, 2009). Hoy sabemos que las modificaciones de los ácidos nucleicos son omnipresentes en los tres reinos de la vida, ampliando significativamente la complejidad de los genomas y actuando como un importante medio de regulación en la transcripción de genes (Ontiveros *et al.*, 2019). Las alteraciones químicas de los nucleósidos se han observado en todos los tipos de DNA y de RNA, siendo particularmente abundantes en los RNAs no codificantes (Grosjean, 2009). Estas constituyen las piezas clave para la compleja regulación multinivel de la expresión genética (Duechler *et al.*, 2016).

La primera modificación descubierta, la N<sup>5</sup>-metilcitosina (m5dC), aunque hallada por primera vez en 1925 por Johnson y Coghill (1925) en el genoma de *Tubercle bacillus* (Grosjean, 2009), no fue confirmada hasta 1948 (Hotchkiss, 1948), donde Hotchkiss identificó en un cromatograma un quinto pico, denominando a este quinto nucleósido animal epicitosina, anticipando, sin saberlo, la definición moderna de la epigenética (Rausch *et al.*, 2020). Sin embargo, se han descrito relativamente pocas modificaciones en el DNA genómico, siendo las más comunes metilaciones simples, concretamente la nombrada anteriormente (m5dC) y la N<sup>6</sup>-metiladenosina (m6dA) y N<sup>4</sup>-metilcitosina (m4dC) (Grosjean, 2009). Las metilaciones del DNA tienen importantes funciones tanto en procariontes, donde principalmente protegen al genoma del huésped de las endonucleasas de restricción, como en eucariotes, donde están implicadas en procesos como la regulación génica, la impronta genómica y el desarrollo embrionario. (Cheng, 1995).

En el caso del RNA, la historia ha sido diferente y, desde luego mucho más compleja. A diferencia del DNA, hoy en día sabemos que todas las posiciones de una pirimidina o un anillo de pirimidina pueden ser modificadas post-transcripcionalmente, no únicamente por metilación o hidroximetilación, sino por desaminación, acetilación, transglicosilación, reducción, tiolación, oxidación, ribosilación, formilación, isomerización, selenización o la adición o transferencia de múltiples grupos (Grosjean, 2009). Desde el descubrimiento de la primera modificación química del RNA, la pseudouridina por Davis y Allen (1957), se han descrito más de 170 alteraciones distintas, conocidas colectivamente como el epitranscriptoma (Esteve-Puig *et al.*, 2020). Este tipo de modificaciones suceden mayoritariamente a nivel post-transcripcional en el RNA de transferencia (tRNA) y ribosomal (rRNA), aunque unas diez suceden también en el RNA mensajero (mRNA) (Baquero-Perez *et al.*, 2021). Las modificaciones en el RNA poseen un amplio rango de funciones en el metabolismo de este, como su procesamiento, *splicing*, poliadenilación, edición, estructura, estabilidad, localización, iniciación de la traducción y regulación de su expresión (Esteve-Puig *et al.*, 2020).

## 1.2. N<sup>6</sup>-metiladenosina (m<sup>6</sup>A)

La N<sup>6</sup>-metiladenosina (m<sup>6</sup>A) es un tipo de metilación reversible (Esteve-Puig *et al.*, 2020) que sucede en la posición N<sup>6</sup> de la adenosina. (He *et al.*, 2019). Fue descrita por primera vez en los años 70 por Desrosiers *et al.* (1974) los cuales la hallaron en el mRNA de células de hepatoma de Novikoff. Más tarde, también fue descrita en virus (Lavi y Shatkin, 1975; Krug *et al.*, 1976) y en plantas (Kennedy *et al.*, 1979). Sin embargo, debido a la falta de métodos de detección eficientes, el interés en esta modificación decayó a finales de esta década. Aunque, en el año 2012, este renació gracias al avance de las tecnologías de secuenciación, donde dos grupos de investigación (Dominissini *et al.*, 2012; Meyer *et al.*, 2012) combinando la inmunoprecipitación con anticuerpos con la secuenciación de alto rendimiento, consiguieron que el mapeo de m<sup>6</sup>A a través del transcriptoma se convirtiera en una realidad (Meyer y Jaffrey, 2017; Yang *et al.*, 2018). Este tipo de modificación es la más común en los mRNA de los organismos eucariotas y en ciertos tipos de RNA virales (Deng *et al.*, 2015), y ocurre de media en un total de tres a cinco posiciones en cada transcrito (Zhu *et al.*, 2019). Esta modificación aparece por norma general en solo los mRNAs nucleares, ya que ocurre de forma co-transcripcional en el núcleo (Zaccara *et al.*, 2019), con algunas excepciones como es el caso de los virus de la familia *Flavoviridae*, que presentan la modificación, pero únicamente se replican en el citosol (Baquero-Perez *et al.*, 2021).

La modificación m<sup>6</sup>A se produce principalmente sobre la secuencia DRACH (donde la D puede ser una adenosina, una guanina o un uracilo, la R puede ser una guanina o una adenina, y la H una citosina, una adenina o un uracilo) (Linder *et al.*, 2015). La frecuencia de este motivo, sin embargo, excede en gran número el número de m<sup>6</sup>A que se pueden encontrar en los transcritos, indicando que es necesario algo más que el motivo para que se produzca la metilación (Batista, 2017). Sin embargo, esta modificación no siempre ocurre en este motivo, apareciendo fuera de este en el 20% de los casos (Baquero-Perez *et al.*, 2021). Esta modificación también se produce en otros motivos específicos de los distintos linajes, como el motivo URUAY (siendo R una guanina o una adenina y la Y un uracilo o una adenina), un motivo preferente de plantas (Arribas-Hernández and Brodersen, 2020). Cuando se introducen mutaciones en estos motivos característicos no se produce la modificación (Kim *et al.*, 2020). Se ha visto que esta modificación no se localiza de forma aleatoria en los transcritos, sino que se ve enriquecida en diferentes zonas, como lo son el último exón, cerca del codón de parada, en las regiones 3' no codificantes (3' UTR) o algunos exones largos (Batista, 2017; Yang *et al.*, 2018). En plantas, además, también se encuentra en codones de inicio, aunque en una frecuencia menor (Kim *et al.*, 2020).

Esta alteración es de vital importancia en múltiples procesos que participan en la regulación y procesamiento del RNA. m<sup>6</sup>A actúa de diversas maneras sobre el mRNA, sea aportándole inestabilidad, participando en la traducción de transcritos, en el procesamiento de estos mediante *splicing*, promoviendo la exportación del núcleo y en la regulación mediante la separación fase líquido-líquido. (Zaccara *et al.*, 2019). Recientes trabajos han vinculado a m<sup>6</sup>A con el desarrollo en animales y a enfermedades humanas (Shi *et al.*, 2019), entre ellas el cáncer, donde se ha visto que, dependiendo del tipo celular y de la clase de tumor, esta puede promover o suprimir la tumorigénesis (Esteve-Puig *et al.*, 2020). En plantas, m<sup>6</sup>A regula un amplio rango de procesos biológicos, como lo son el desarrollo embrionario, el desarrollo de tricomas, del meristemo apical caulinar, de las transiciones entre las distintas fases de crecimiento y de la respuesta a los estreses tanto bióticos como abióticos que sufren las plantas. (Martínez-Pérez *et al.*, 2017; Arribas-Hernández and Brodersen, 2020; Kim *et al.*, 2020).

La regulación de la modificación m<sup>6</sup>A en los mRNA es un proceso dinámico que involucra a m<sup>6</sup>A metiltransferasas para añadir la modificación (*writers*), proteínas de unión a m<sup>6</sup>A para reconocer y unírsele (*readers*) y m<sup>6</sup>A desmetilasas para eliminarla (*erasers*) (Shi *et al.*, 2019) (Figura 1).

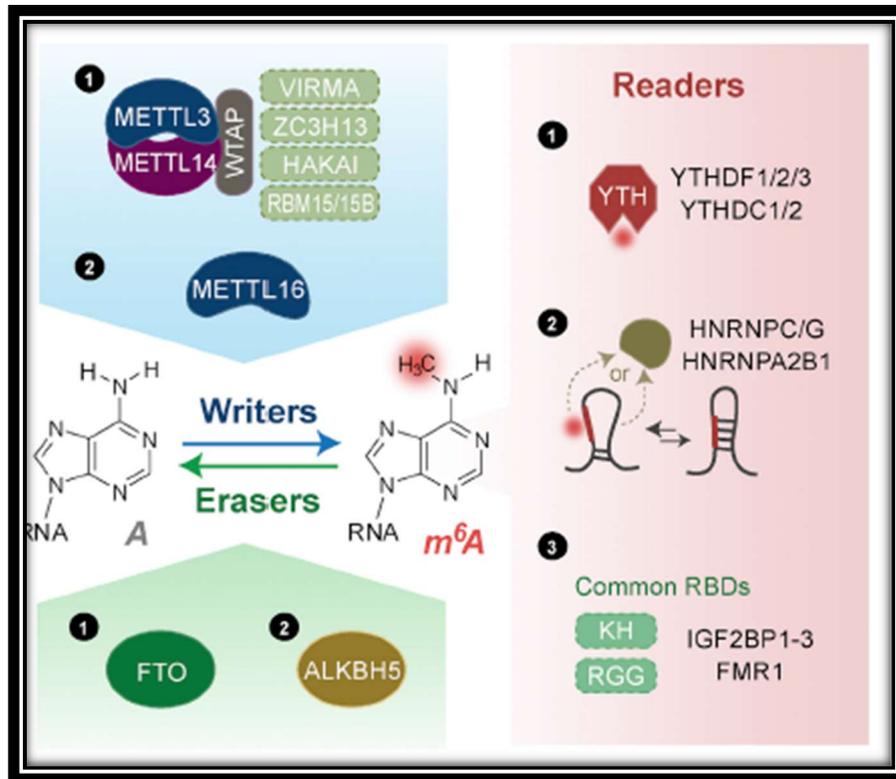


Figura 1. Proteínas de regulación de la modificación m<sup>6</sup>A en humanos. Shi *et al.*, 2019.

### 1.2.1. Writers (Escritoras)

Las modificaciones químicas del RNA están catalizadas por complejos multiproteínicos *writers* (Esteve-Puig *et al.*, 2020) Así, en el genoma de los mamíferos existen cuatro metiltransferasas para la formación de m<sup>6</sup>A en los distintos RNAs (Zaccara *et al.*, 2019). Casi la totalidad de m<sup>6</sup>A se instala en los mRNA co-transcripcionalmente por un complejo formado por múltiples subunidades, con un núcleo estable que está formado por la methyltransferasa-like 3 (METTL3) (MTA en *Arabidopsis*) y la methyltransferase-like 14 (METTL14), siendo el primero la subunidad catalítica y el segundo un componente esencial para facilitar la unión del RNA (Shi *et al.*, 2019). METTL3-METTL14 es el complejo responsable de modificar la mayor parte de los RNA sintetizados por la RNA polimerasa II además de los mRNAs, como los RNAs largos no codificantes (lncRNAs), los precursores de los microRNAs (pri-miRNAs) y los RNAs circulares (circRNAs) (Berlivet *et al.*, 2019). Este complejo también está formado por otras subunidades caracterizadas en distintos estudios como son: la proteína asociada al tumor de Wils 1 (WTAP), (Shi *et al.*, 2019) que actúa como un andamiaje (Visvanathan y Somasundaran, 2018) que se une a METTL3-METTL14 y se necesita para la óptima localización del complejo y el correcto reclutamiento del sustrato; Vir-like m<sup>6</sup>A metiltransferasa associated (VIRMA) que es crítica para la deposición del m<sup>6</sup>A específicamente en la 3' UTR; la proteína con dedo de zinc tipo CCCH

13 (ZC3H13), que facilita la localización nuclear del complejo *writer*; la proteína de unión al RNA 15/15B (RBM15/15B) que se une a las regiones ricas en uracilos y facilita la metilación de ciertos RNAs (Shi *et al.*, 2019); y CBL1, también conocida como HAKAI, una ubiquitin ligasa que interacciona con el complejo de la E-cadherina (Zaccara *et al.*, 2019). Este complejo instala m6A de forma preferente sobre el motivo DRACH (He y He, 2021), aunque es necesario algo más que este para que se produzca la metilación. La selección de transcritos sobre los cuáles se efectúa la modificación se produce, al menos, a través de dos vías: mediante el reclutamiento por parte de promotores del complejo METT3-METTL14 a través de factores de transcripción, y también a través de las modificaciones de las histonas. Por otro lado, la elección de los motivos DRACH en los cuales se produce la metilación se produce mediante RBPs (proteínas de unión al RNA) que se unen a lugares cercanos al motivo y reclutan al complejo *writer* para que efectúe su acción en los motivos DRACH cercanos a este, y también por el reclutamiento a través de la RNA polimerasa 2 (RNA pol II), la cual une el complejo *writer* para inducir la metilación y de la cual se ha visto que su ralentización e inhibición provoca un aumento en los niveles de m6A (Zaccara *et al.*, 2019). En *Arabidopsis* también existe este complejo, formado por MTA (homólogo de METTL3) asociado a MTB (homólogo de METTL14) junto con otras subunidades como FIP37 (homólogo de WTAP), VIR (homólogo de VIRMA), HAKAI y posiblemente otras subunidades que aún no hayan sido caracterizadas (Bhat *et al.*, 2018). De igual forma, este complejo se encuentra conservado en todos los linajes de los eucariotas superiores, sugiriendo que la modificación m6A de los transcritos formados por la RNA pol II es un mecanismo ancestral que evolucionó en el primer ancestro común de estos (Berlivet *et al.*, 2019).

Otra metiltransferasa, conocida como la metiltransferasa-like 16 (METTL16), cataliza la formación de m6A en un reducido grupo de mRNAs y de RNAs no codificantes (Zaccara *et al.*, 2019). Mientras que el complejo METTL3-METTL14 preferiblemente instala m6A en el motivo RRACH, METTL16 instala m6A en una secuencia y estructura distintas. Los ensayos realizados sobre esta proteína sugieren que METTL16 prefiere una secuencia UAC(m6A)GAGAA ubicada en la protuberancia de un RNA estructurado en forma de bucle-tallo. Dos sustratos validados de esta metiltransferasa son el U6 RNA pequeño nuclear (snRNA) (Shi *et al.*, 2019), el cual está implicado en el proceso de *splicing* (Zaccara *et al.*, 2019) y una horquilla en la región 3' UTR de la metionina adenosiltransferasa 2A mRNA (MAT2A) que codifica la S-adenosilmetionina sintetasa (SAM). Una de sus implicaciones funcionales es participar en la regulación de la homeostasis de SAM. (Shi *et al.*, 2019). En *Arabidopsis*, su homólogo es FIONA1 (FIO1), el cuál es esencial en el mantenimiento de los ritmos circadianos de la planta. (Pendleton *et al.*, 2017)

Otros complejos *writers* de mamíferos que se encargan de formar m6A en el RNA son la  $\alpha$ -adenosil-metiltransferasa ZCCHC4, que introduce el único m6A que aparece en el rRNA 28s, y el complejo formado por METTL5-TRMT112, donde METTL5 es la subunidad catalítica y TRMT112 el adaptador alostérico, que se encarga del sitio único m6A de la subunidad rRNA 18s. (Zaccara *et al.*, 2019)

### **1.2.2. Readers (Lectoras)**

A nivel molecular, el papel más común de m6A es influenciar en la unión de las proteínas a su RNA diana. Esta metilación puede actuar repeliendo RBPs o bien actuar como una señal de unión para las denominadas proteínas *readers*. Estas proteínas son un tipo de RBP que, una vez unida a su RNA diana pueden controlar su destino o reclutar de forma directa o indirecta a proteínas efectoras (Berlivet *et al.*, 2019)

De entre estas, las proteínas con los dominios YTH son la familia de *readers* más ampliamente estudiada y caracterizada (Esteve-Puig *et al.*, 2020). Este dominio está presente en más de 170 proteínas y se encuentra conservado evolutivamente en las especies eucariotas. Aunque se conocía desde hace tiempo que estas eran proteínas de unión al RNA, su ligando se mantuvo desconocido hasta la aparición de dos estudios (Dominissini *et al.*, 2012; Wang *et al.*, 2014) en los que se descubrió que los miembros de la familia YTH en mamíferos son candidatos a *readers* de m6A (Liao *et al.*, 2018). Esta familia de dominios de proteínas fue la primera en ser descubierta como *readers* de m6A y propició las bases mecánicas para entender los efectos de m6A en mRNA (Zaccara *et al.*, 2019). En vertebrados, las proteínas que contenían el dominio YTH fueron clasificadas inicialmente como las *YTH domain family* (YTHDF) y como *YTH domain containing proteins* (YTHDC) acorde a su distribución subcelular (Berlivet *et al.*, 2019) encontrándose las primeras en el citoplasma y las segundas en el núcleo (Shi *et al.*, 2019). Más tarde se descubrió que estos dos grupos pertenecen a dos subclados evolutivos, uno conteniendo YTHDF y el otro YTHDC. Los miembros de estos subclados se etiquetan como motivos tipo DF y DC (Berlivet *et al.*, 2019) Los humanos poseen 2 proteínas YTHDC (YTHDC1 y YTHDC2) y 3 proteínas YTHDF (YTHDF1-3) (Reichel *et al.*, 2019) altamente similares (Zaccara *et al.*, 2019) Todos los dominios YTH adoptan un plegamiento canónico conservado que consiste en 3  $\alpha$ -hélices y 6  $\beta$ -láminas. Estas últimas están organizadas en una atípica forma de barril que forma, junto con las  $\alpha$ -hélices, un núcleo hidrofóbico. De esta forma los dominios YTH usan un mecanismo conservado para unirse a m6A consistente en acomodar esta modificación en un bolsillo aromático hidrofóbico formado por 3 triptófanos o 2 triptófanos y una tirosina/leucina (Berlivet *et al.*, 2019). Dentro de los motivos DC, DC1 se relaciona con el *splicing* del mRNA, el silenciamiento epigenético y la exportación nuclear de mRNAs. Se localiza en unas estructuras moteadas en el núcleo, bautizadas como cuerpos YT, que se cree que son motas nucleares. DC2, por otro lado, es, de las proteínas humanas con dominio YTH, la que más levemente se une a m6A, y se encuentra principalmente en los testículos (Zaccara *et al.*, 2019), encargándose de mediar procesos como la espermiogénesis y la traducción y estabilidad de los mRNAs. (Shi *et al.*, 2019). En cuanto a las proteínas DF, estas comparten una alta identidad de aminoácidos en toda su secuencia. Sin embargo, diferentes estudios han reportado que cada una de las 3 proteínas DF tiene un efecto distinto sobre los mRNAs: DF1 incentiva la traducción de los mRNAs modificados; DF2 promueve su degradación y DF3 posee ambas funciones. (Zaccara *et al.*, 2019). Pese a encontrarse normalmente en el citoplasma, ciertos ensayos han demostrado que estas proteínas pueden translocarse al núcleo en determinadas condiciones de estrés como el shock térmico o una infección viral (Shi *et al.*, 2019). En *Arabidopsis*, una búsqueda sistemática de homología mostró que esta especie contiene 13 proteínas con el dominio YTH, que han surgido a partir de múltiples eventos de multiplicación. Estas fueron nombradas como Evolutionary Conserved C-Terminal Region (ECT). ECT1-ECT11 pertenecen al grupo de DF, mientras que ECT12 Y 13 al DC. (Reichel *et al.*, 2019)

Otras proteínas *readers* identificadas son la Fragile X mental retardation 1 (FMR1), que tiene un efecto tanto en la traducción como en la estabilidad de los mRNAs, y las insulin-like growth factor 2 mRNA-binding proteins 1–3 (IGF2BP1- 3), que aportan estabilidad a los mRNA. (Shi *et al.*, 2019). También se cree que los ribosomas podrían actuar como *readers*. Por otro lado, m6A puede reclutar RBPs de forma indirecta, ya que la presencia de un solo m6A en una hélice corta de RNA puede desestabilizarlo, dando a las RBPs que se unen a motivos cercanos, o directamente a m6A, un acceso más fácil al RNA. Este concepto es conocido como el “m6A structural switch” y ha sido demostrado con proteínas como HNRNPC, HNRNPG y HNRNPA2B1 (Zaccara *et al.*, 2019)

### 1.2.3. Erasers (Borradoras)

Los m6A *erasers* son desmetilasas que transforman m6A en A mediante desmetilación activa (Zaccara *et al.*, 2019; Shi *et al.*, 2019). Aunque METTL3 fue descubierto como *writer* hace varias décadas, la identidad de las desmetilasas permaneció como un misterio hasta la pasada, donde Jia *et al.* (2011) demostraron que la proteína conocida como “fat mass and obesity associated protein” (FTO), exhibía una eficiente actividad desmetilasa. Poco después ALKBH5 fue descubierta por Zheng *et al.* (2013). La demostración de su actividad desmetilasa proporcionó la primera evidencia de una modificación post-transcripcional reversible en el RNA. (Yang *et al.*, 2018). Las proteínas que se encargan por tanto de la eliminación de m6A pertenecen a la familia de las proteínas homólogas de alkylation B (AlkB) (Reichel *et al.*, 2019), en las que nos centraremos más adelante (Veáse “AlkB”).

## 1.3. N6-metiladenosina en el ciclo de replicación viral

Las alteraciones del RNA tienen papeles clave en los ciclos de infección de los virus. Durante la infección, los RNA virales se ven modificados químicamente y estos afectan a los ciclos de replicación. Además, cambian el paisaje epitranscriptómico de sus huéspedes y dichos cambios también les afectan. (Baquero-Perez *et al.*, 2021). La modificación m6A fue descrita en ciertos virus de RNA largo tiempo atrás, como en el caso del virus de la influenza (IAV) (Krug *et al.*, 1976), el virus de la leucemia felina (Thomason *et al.*, 1976), el virus del sarcoma aviar B77 (Stoltzfus y Dimock, 1976; Dimock y Stoltzfus, 1977), o el virus sincitial respiratorio (RSV) (Kane y Beemon, 1985) (Dang *et al.*, 2019). Si bien es cierto que la existencia de esta modificación se conoce desde hace más de 40 años, como hemos visto anteriormente, la falta de métodos para mapear el m6A hizo que el interés por esta modificación decayera, hasta el año 2012, donde dos grupos distintos de investigación desarrollaron una tecnología NGS que permitía la secuenciación de esta modificación en el transcriptoma. (Dominissini *et al.*, 2012, Meyer *et al.*, 2012). Desde entonces, m6A se ha convertido en un foco de investigación con el fin de elucidar su papel dentro de las infecciones virales (Dang *et al.*, 2019). El enfoque para la caracterización del papel de m6A en los RNA virales se ha basado principalmente en la manipulación de la maquinaria celular encargada de m6A mediante sobre- o infra- expresión de componentes clave, o bien mutaciones en ciertos motivos DRACH en los RNA virales con el fin de suprimir ciertos sitios de modificación. Se ha visto que la mayor parte de los pasos del ciclo viral se ven afectados por las modificaciones m6A, siendo estos efectos pro- o anti- virales dependiendo del virus en cuestión. (Baquero-Perez *et al.*, 2021).

Se ha estudiado un amplio espectro de virus que afectan a animales para esta modificación. El primer virus de RNA donde se descubrió la alteración en sus mRNAs, el IAV, se estudió en el año 2017 por Courtney *et al.* (2017). Estos demostraron que, inhibiendo la adición de m6A de forma global se producía una disminución de la expresión génica y la replicación virales, mientras que la sobreexpresión de la proteína lectora YTHDF2 producía un efecto contrario (Courtney *et al.*, 2017). En el caso del VIH, diferentes grupos han investigado el papel de m6A en la infección por VIH-1. (Kennedy *et al.*, 2016; Tirumuru *et al.*, 2016; Lichinchi *et al.*, 2016; Lu *et al.*, 2018). Mientras que algunos asignaron a m6A un papel positivo, el resto le asignaron uno negativo, aunque las contradicciones que se observaron podrían explicarse por las diferencias de metodología o por las enzimas diferenciales de las rutas de señalización distintas estudiadas (Baquero-Perez *et al.*, 2021). Los virus de la familia *Flaviviridae*, que se replican exclusivamente en el citosol y poseen genomas estructurados, también han sido objeto de estudio en la modulación

del estado de metilación de su genoma. Así, se ha detectado m6A en diversos miembros de esta familia como DENV, ZIKA, WNV, YFV, y HCV. En HCV, el silenciamiento del METTL3-METTL14 o de los *readers* YTHDF1-3 incrementó la producción de virus, mientras que el de la *eraser* FTO produjo un efecto contrario. Lo mismo sucede con ZIKA, sobre el cual las modificaciones m6A regulan negativamente su infección. (Baquero-Perez *et al.*, 2021). Otros virus de RNA estudiados han sido el virus de stomatitis vesicular (VSV), para él se ha visto que, al igual que en el caso de los flavivirus, m6A afecta negativamente a su ciclo vital (Dang *et al.*, 2019); los coronavirus, concretamente el virus de diarrea epidémica porcina (PEDV) y el Sars-CoV-2, (para los cuales también resulta perjudicial m6A, y otros como el Enterovirus 71 (EV71), que mostró fuertes alteraciones de la maquinaria reguladora de m6A en sus huéspedes, indicando una poderosa manipulación epitranscriptómica en estos, y mostró divergencias en la acción de m6A sobre su ciclo de replicación viral según el tipo de células afectadas, y el virus respiratorio sincitial (RSV), que mostró una producción aumentada cuando se producía una sobreexpresión de *readers* y *writers* y disminuida en el caso de los *erasers*, viéndose positivamente afectado por m6A (Baquero-Perez *et al.*, 2021).

Se ha demostrado que la metilación m6A está involucrada en la resistencia de las plantas hacia los patógenos (Zhang *et al.*, 2021a). El papel de esta modificación sobre la infección viral se ha estudiado en plantas, concretamente en *Arabidopsis* (Martínez-Pérez *et al.*, 2017), en tabaco (Li *et al.*, 2018), en trigo (Zhang *et al.*, 2021a) y en arroz (Zhang *et al.*, 2021b). En *Arabidopsis*, Martínez-Pérez *et al.* (2017) identificaron m6A en los genomas de dos virus de la familia *Bromoviridae*: virus del mosaico de la alfalfa (AMV) y el virus del mosaico del pepino (CMV). Ambos virus se replican en el citoplasma, como los virus de la familia *Flavoviridae*, indicando que la presencia de m6A es un fenómeno común en este tipo de virus. Se piensa que en plantas el fenómeno m6A está ligado al silenciamiento y la pérdida de estabilidad de los mRNAs. En AMV los autores vieron que, al igual que sucedía con los Flavivirus, la presencia de m6A regulaba negativamente su infección viral. Se cree que este virus es capaz de utilizar maquinaria desmetilasa de *Arabidopsis* (ALKBH9B) con el fin de eliminar esta modificación. CMV, por otro lado, no interactuó con la ALKBH9B in vivo, aunque esto podría explicarse porque su proteína de cubierta (CP), a diferencia de la CP de AMV, que es una multiproteína que interactúa con múltiples factores del huésped y necesaria para la infección y replicación del virus, carece de funciones tan esenciales, y posiblemente la interacción de esta con la maquinaria *eraser* sea necesaria para la actividad ALKBH9B en los vRNAs. (Martínez-Pérez *et al.*, 2017). En tabaco, Li *et al.* (2018) indagaron sobre el efecto de la infección con el virus del mosaico del tabaco (TMV). Los resultados mostraron una reducción de los niveles de m6A y un particular aumento de la expresión del homólogo de ALKBH5, así como una disminución de los niveles de *writers* y *erasers*, y sugirieron que esto podía ser debido a una interacción del virus con la maquinaria m6A. (Li *et al.*, 2018). En trigo, Zhang *et al.* (2021a) caracterizaron la modificación durante la infección por WYMV. Como en casos anteriores, diferentes genes mostraron una expresión y niveles de m6A anormales, la mayoría de los cuales involucrados con las respuestas de defensa de la planta. (Zhang *et al.*, 2021a). Por último, en arroz, Zhang *et al.* (2021b) caracterizaron la presencia de m6A en el transcriptoma durante las infecciones virales causadas por dos de los más devastadores patógenos de este cultivo: Rice stripe virus (RSV) y Rice black-stripe dwarf virus (RBSDV). Estos vieron que m6A se encontraba enriquecido en las plantas infectadas y que estaba asociada fuertemente a genes que se expresaban en niveles bajos durante la infección viral, teniendo un papel esencial en la regulación de la expresión de muchos genes de vías antivirales (Zhang *et al.*, 2021b).

## 1.4. AlkB

### 1.4.1. La familia AlkB: contexto, historia y mecanismo molecular

Las proteínas implicadas en la eliminación de grupo metilo de m<sup>6</sup>A pertenecen a la familia de los homólogos de Alkylation B (AlkB), la cual forma parte de la superfamilia de oxigenasas (2OG-oxigenasas) 2-oxoglutarato y Fe (II) dependientes (Reichel *et al.*, 2019). La 2OG-oxigenasas constituyen una enorme superfamilia ampliamente extendida en bacterias, hongos, plantas y animales. (Jia *et al.*, 2017). Estas catalizan diversas reacciones que típicamente involucran la oxidación de un sustrato orgánico utilizando una molécula de dióxígeno (Aravind y Koovin, 2001). Estas proteínas comparten un plegamiento característico con una estructura conservada formada por un núcleo *jelly roll* (en remolino) de β-láminas con α-hélices flanqueantes (Bratlie y Drabløs, 2005). Las 2-OG oxigenasas dependen de un hierro ferroso como cofactor para lograr la unión de un oxígeno molecular y que se produzca las consecuentes reacciones de oxidación. En prácticamente todos los casos las 2OG-oxigenasas acoplan la oxidación de dos electrones del sustrato primario a la descarboxilación oxidativa del cosustrato 2-oxoglutarato dando lugar a succinato y un dióxido de carbono. De esta forma se da lugar a la formación de productos específicos (Farrow y Facchini, 2014). Entre las proteínas que conforman esta superfamilia, son muy pocos los residuos que se encuentran totalmente conservados, únicamente aquellos involucrados en la coordinación del ion Fe (II) y el 2-oxoglutarato (Bratlie y Drabløs, 2005).

Dentro de esta, los homólogos de AlkB constituyen una omnipresente familia de proteínas que está representada en casi todos los organismos celulares con la excepción de arqueas y levaduras (**Figura 2**). La función de las AlkB es revertir el daño por metilación en los ácidos nucleicos, de tal forma que aseguran la integridad genómica frente a los efectos deletéreos que resultan de las alquilaciones por agentes químicos (Moore y Meng, 2019). Los agentes alquilantes se encuentran tanto en el ambiente como dentro de las células y pueden causar la alquilación de la mayoría de los átomos de nitrógeno y oxígeno en las bases de DNA y RNA. Algunas de estas interfieren en el emparejamiento de bases bloqueando la transcripción, la replicación y la traducción, e incluso pueden provocar emparejamientos erróneos dando lugar a mutaciones. Como el daño por alquilación representa una seria amenaza a la transmisión e integridad de la información genética todos los organismos están equipados con varios sistemas de reparación para contrarrestar los efectos, como las AlkB (van den Born *et al.*, 2008). Por ello, la delección de AlkB da lugar normalmente a una elevada tasa de mutaciones, a la activación de la respuesta SOS e incluso a la muerte de los organismos (Moore y Meng, 2019). Sin embargo, las funciones de las proteínas AlkB también van más allá de la reparación del material genético, participando en otros procesos como por ejemplo la modificación del tRNA y la regulación genética (Bjørnstad *et al.*, 2011). Generalmente, las bacterias portan una única copia de AlkB, mientras que los eucariotas poseen múltiples homólogos, los cuales poseen una mayor especialización que aquellos encontrados en bacterias (Moore y Meng, 2019). Entre ellas, se han encontrado AlkB que utilizan como sustrato DNA, RNA e incluso proteínas. (Moore, 2018). Interesantemente, se ha visto que ciertos virus de RNA de cadena simple de polaridad positiva (ssRNA<sup>+</sup>) que infectan a plantas portan un homólogo AlkB en su material genético que repara DNA y, preferiblemente, RNA (Fedeles *et al.*, 2015).

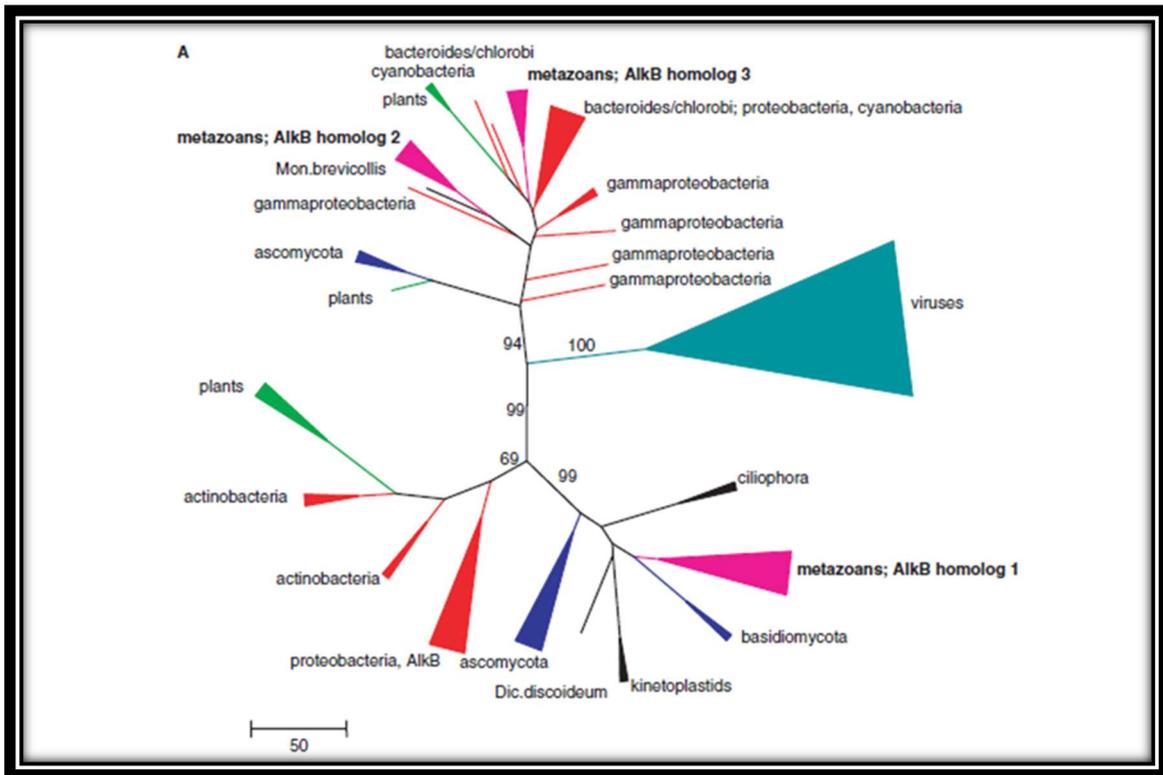


Figura 2. Filogenia general de la familia AlkB. van den Born et al., 2008

El descubrimiento de las AlkB se remonta a hace más de 40 años. En 1977, Samson y Cairns (1997) descubrieron un fenómeno en *Escherichia coli* (*E. coli*) bautizado como “Respuesta adaptativa”, por el cual se les confería a las bacterias una mayor resistencia hacia el daño por alquilación gracias a las proteínas Ada, AlkA, AlkB y AldB. (Martínez-Pérez, 2020). Para comprobarlo, hace casi 40 años Kataoka *et al.* (1983) relacionaron la deficiencia del gen de AlkB de *E. coli* con una mayor sensibilidad hacia los agentes alquilantes (Ougland *et al.*, 2004). Sin embargo, no fue hasta dos décadas más tarde cuando se produjo un importante avance al ser clasificada por Aravind y Koovin (2001) como parte de la familia 2OG-oxigenasa utilizando aproximaciones bioinformáticas, y confirmada experimentalmente como tal al año siguiente por Falnes *et al.* (2002) y Trewick *et al.* (2002) (Moore, 2018). Las investigaciones llevadas a cabo en los años que han seguido al descubrimiento de sus propiedades enzimáticas en 2002 han expandido ampliamente nuestro conocimiento sobre las funciones biológicas de las dioxigenasas AlkB. (Fedeles *et al.*, 2015).

Los homólogos AlkB son proteínas relativamente pequeñas de un tamaño de 200 a 400 aminoácidos (Aas *et al.*, 2003). La determinación de las estructuras cristalinas de la AlkB de *E. coli* y de los homólogos humanos ALKBH2, ALKBH3, ALKBH5, ALKBH7, ALKBH8 y FTO nos ha procurado una visión significativa de la desalquilación catalizada por las dioxigenasas AlkB (Fedeles *et al.*, 2015) (**Figura 3**). Aunque existe una baja conservación de secuencia entre los homólogos AlkB, las características clave de su estructura se encuentran bien conservadas. El sitio activo de estas enzimas está contenido en el característico núcleo en remolino de las 2-OG oxigenasas, un plegamiento formado por 8  $\beta$ -láminas emparejadas de dos en dos en forma helicoidal que forma el núcleo catalítico que une el cofactor y el cosustrato usando unos pocos residuos conservados: el motivo HXD conservado universalmente y el aminoácido H,

responsables de la unión de Fe (II), y el motivo RXXXXXR (donde X denota cualquier aminoácido), responsable de la unión de 2-OG. En *E. coli*, además presenta un subdominio de 90 aminoácidos que interacciona con el sustrato conocido como el Nucleotide Recognition Lid (NRL) formando una cubierta de reconocimiento de nucleótidos, aunque en otros organismos esta pobremente definida (Moore y Meng, 2019).

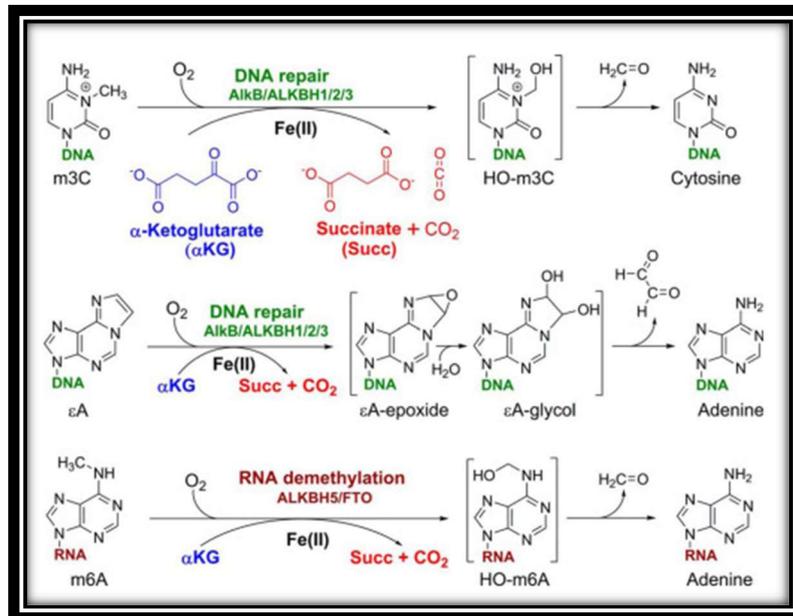


Figura 3. Ejemplos de desalquilación oxidativa en el DNA y RNA catalizada por los homólogos AlkB humanos: ALKBH1, ALKBH2, ALKBH3, ALKBH5 y FTO. Fedeles *et al.*, 2015.

#### 1.4.2. AlkB en bacterias

La mayor parte de las especies bacterianas aerobias expresan proteínas AlkB, las cuales muestran una gran diversidad de especificidad de sustrato. Debido al requerimiento de AlkB de oxígeno molecular, las bacterias anaerobias estrictas, como por ejemplo las del género *Clostridium*, no parecen poseer estas proteínas. De entre todos los homólogos de esta familia, la AlkB de *E. coli* (EcAlkB) es, de lejos, la enzima mejor estudiada de toda la familia (Fedeles *et al.*, 2015). Esta proteína realiza la desmetilación oxidativa de las bases de RNA/DNA, recuperando su estructura nativa (Marcinkowski *et al.*, 2020). EcAlkB ha mostrado actividad reparadora sobre N<sup>3</sup>-metilcitosina (N<sup>3</sup>-meC), N<sup>1</sup>-metiladenina (N<sup>1</sup>-meA), N<sup>3</sup>-metiltimina (N<sup>3</sup>-meT), N<sup>1</sup>-metilguanina (N<sup>1</sup>-meG) y otras lesiones del DNA y del RNA (Feng *et al.*, 2014). También repara aductos más voluminosos formados por grupos etilo- propilo- y los aductos exocíclicos etano- y eteno- (Marcinkowski *et al.*, 2020), aunque con menor eficiencia (van den Born *et al.*, 2008). Mediante análisis bioinformáticos, se delimitaron cuatro grupos filogenéticos distintos de entre las AlkB bacterianas anotados como 1A, 1B, 2A y 2B. Las proteínas de 1A, entre las cuales está la EcAlkB, se caracterizan por una actividad desalquilante robusta y por poseer amplias preferencias de sustrato. Este grupo incluye las proteínas AlkB de *Streptomyces*. Las proteínas del grupo 1B, que se encuentran principalmente en Proteobacterias y en las cianobacterias, también muestran amplias preferencias de sustrato. Están más estrechamente relacionadas con los homólogos eucariotas ALKBH2 y ALKBH3. Las proteínas del grupo 2A, que a menudo se encuentran también en las Proteobacterias (por ejemplo, *Agrobacterium*, *Rickettsia* y *Rhizobium*), comparten una fuerte homología con la ALKBH8 de animales y plantas, implicada en la modificación

postranscripcional del tARN. Por último, las proteínas 2B, que se encuentran habitualmente en bacterias del suelo como *Actinobacteria* y en otras como *Xanthomonas* y *Burkholderia*, son las que presentan una mayor especialización en sus sustratos, donde algunas tienen alta actividad reparadora en monoalquilaciones pero no reparan aductos más voluminosos y viceversa. Cabe destacar que, a diferencia de *E. coli*, muchos genomas bacterianos poseen dos e incluso tres genes de la familia AlkB en sus genomas. (Fedele *et al.*, 2015).

### 1.4.3. AlkB en mamíferos

En mamíferos, se han identificado nueve homólogos de AlkB en humanos (ALKBH1-9 y FTO) y en ratón (*Alkbh1-9* y *Fto*) y se ha demostrado que poseen distintas funciones biológicas y afinidades por los sustratos (Wakisaka *et al.*, 2019). De entre estos, ALKBH5 y FTO son las más recientes evolutivamente hablando, ya que solo pueden encontrarse en vertebrados, mientras que el resto está conservado en todos los metazoos (Fedele *et al.*, 2015). ALKBH1 es, de entre todos los homólogos AlkB humano, el que mayor similitud presenta con la AlkB de *E. coli*. (Müller *et al.*, 2009). Funciona como una desmetilasa mitocondrial de ácidos nucleicos que presenta actividad liasa apirimidínica/apurínica, y que no solo actúa sobre el DNA sino también sobre la histona H2A (Moore, 2018). ALKBH2 y ALKBH3 son las que presentan una actividad enzimática más similar a la *EcAlkB*. Ambas tienen como sustrato a N<sup>1</sup>-meA y N<sup>3</sup>-meC, y eliminan también aductos de eteno. La primera se localiza en el núcleo y tiene preferencia por el DNA de doble cadena, mientras que la segunda se localiza tanto en el núcleo como en el citoplasma y tiene el DNA de simple cadena como sustrato preferente. ALKBH4 también se ha encontrado tanto en el núcleo como en el citoplasma, presentando diferencia en su actividad según su localización. Se cree que en el núcleo participa en la regulación de la expresión génica y en el mantenimiento de la cromatina, mientras que en el citoplasma interactúa con una actina, desmetilándola y, de esta forma, participando en la formación del anillo contráctil durante la mitosis. De ALKBH6 se sabe muy poco, solo que está tanto en el núcleo como en el citoplasma y que posee una superficie cargada sugiriendo una posible interacción con los ácidos nucleicos (Marcisowski *et al.*, 2020). ALKBH7 es el miembro de la familia de los homólogos AlkB necesario para la necrosis inducida por alquilación y oxidación, teniendo a proteínas como sustrato. Esta es una desmetilasa localizada en las mitocondrias que, en contraste con el papel protector del resto de miembros, desencadena el colapso del potencial de membrana mitocondrial y promueve la muerte celular, además de poseer un papel importante en el metabolismo de las grasas (Wang *et al.*, 2014). ALKBH8, por otro lado, es una desmetilasa del tRNA, la cual es necesaria para llevar a el último paso en la biosíntesis del metiléster de la 5-carboximetiluridina (mcm5U) (Songe-Møller *et al.*, 2010).

De entre los homólogos humanos de AlkB, ALKBH5 y FTO son las encargadas de eliminar la modificación m<sup>6</sup>A en el RNA. ALKBH5 es una m<sup>6</sup>A desmetilasa y su actividad tiene importantes funciones en la exportación de los mRNA nucleares, así como en la asociación de las proteínas moteadas nucleares con el metabolismo del RNA. Su actividad m<sup>6</sup>A desmetilasa fue confirmada tras verse un aumento de los niveles de m<sup>6</sup>A en ratones mutados que carecían de ella. Además, la deficiencia de esta proteína producía en ratones una espermatogénesis aberrante y apoptosis situadas en los testículos, teniendo un papel esencial en los procesos biológicos reproductivos (Zheng *et al.*, 2013). FTO es la proteína asociada a la obesidad y a la masa de grasa. Como ALKBH5, también presenta una alta afinidad por m<sup>6</sup>A y está asociada a las proteínas moteadas nucleares (Jia *et al.*, 2011). Sin embargo, recientemente se ha visto que FTO presenta una afinidad

100 veces mayor por m<sup>6</sup>A, el primer nucleótido de la caperuza 5', sugiriendo que este puede ser su sustrato principal (Reichel *et al.*, 2019). Su sobreexpresión está asociada a la obesidad y a la diabetes (Fedele *et al.*, 2015), y su probable función es la de participar en el procesamiento de los mRNA y otro RNA nucleares. (Jia *et al.*, 2011).

#### 1.4.4. AlkB en plantas

En plantas, el papel de las proteínas *erasers* solo se ha determinado en unos pocos casos (Huong *et al.*, 2020). El genoma de *Arabidopsis* contiene trece homólogos de EcAlkB caracterizados bioinformáticamente: ALKBH1A, ALKBH1B, ALKBH1C, ALKBH1D, ALKBH2, ALKBH6, ALKBH8A, ALKBH8B, ALKBH9A, ALKBH9B, ALKBH9C, ALKBH10A, y ALKBH10B. (Marcinkowski *et al.*, 2020) Un estudio de la localización subcelular mostró que todas estas proteínas tienen una localización nucleocitoplasmática, con la excepción de ALKBH1D, que también se encontró en los cloroplastos, y ALKBH9B, que se localiza únicamente en el citoplasma (Martínez-Pérez *et al.*, 2017). Los datos que se tienen acerca de la función de estos homólogos indican que la duplicación y multiplicación de los homólogos de EcAlkB en plantas está conectada a la ejecución de varios roles en la célula. ALKBH1A es el homólogo de *Arabidopsis* del ALKBH1, y a su vez se cree que ALKBH1D puede tener un papel similar en los cloroplastos al de ALKBH1 en las mitocondrias (Marcinkowski *et al.*, 2020). atALKBH2 es un ortólogo funcional del homólogo humano ALKBH2, y como tal mostró actividad reparadora frente a los aductos de grupos metilo (1-meA y 3-meC) y eteno en el DNA de doble cadena (Meza *et al.*, 2012). La función de ALKBH8 es similar a la de su homólogo humano ALKBH8, participando en la hidroxilación del mcm5U del tRNA (Marcinkowski *et al.*, 2020). Recientemente, se vió que ALKBH6 desempeña un papel en la germinación de semillas, el crecimiento de las plántulas y la supervivencia de *Arabidopsis* bajo estreses abióticos. Muestra capacidad para unirse a m<sup>6</sup>A y a m<sup>5</sup>C, pero su función biológica sigue siendo desconocida (Huong *et al.*, 2020).

De entre estos, ALKBH9A, 9B, 9C, 10A y 10B muestran la mayor similitud a nivel de secuencia con el homólogo humano ALKBH5 (Marcinkowski *et al.*, 2020). Se ha demostrado que ALKBH9B y ALKBH10B tienen la capacidad de eliminar el grupo metilo de m<sup>6</sup>A (Martínez-Pérez *et al.*, 2017; Duan *et al.*, 2017).

ALKBH9B mostró actividad m<sup>6</sup>A desmetilasa tanto in vitro como in vivo, y causó efectos directos sobre la infectividad del virus AMV, indicando que posee un papel en la regulación frente a ciertas infecciones virales. Los resultados previos sugieren que ALKBH9B puede estar relacionada con el silenciamiento y degradación de los mRNA (Martínez-Pérez *et al.*, 2017). El mismo laboratorio se encargó después de realizar un análisis de mapeo de la proteína para identificar los dominios implicados de esta en la regulación del ciclo de infección viral (**Figura 4**). En su secuencia, la proteína posee regiones intrínsecamente desordenadas (IDRs) localizadas en la zona N-terminal, una de las cuales delimita el dominio AlkB y la zona C-terminal. Se vió que los residuos situados entre las posiciones 427 y 467 son críticos para la unión in vitro con el RNA viral. Se identificó un dominio de unión al RNA que contenía un motivo RGXXXRGG que solapaba con la zona C-terminal. También se observó que los residuos localizados entre las posiciones 387 y 427 son críticos para la interacción con la proteína de cubierta (CP) y, por tanto, deberían ser críticos para modular el proceso de infección viral. Así, la delección de tanto los 20 primeros residuos de la zona N-terminal como de los 40 últimos de la zona C-terminal impidieron

su acumulación en los cuerpos de los pequeños RNAs de interferencia (Alvarado-Marchena *et al.*, 2021).

Por otro lado, Duan *et al.* (2017) probaron que ALKBH10B también posee actividad desmetilasa, viéndose un aumento generalizado de los niveles de m6A en el RNA en las plantas con deficiencia de esta (Duan *et al.*, 2017). La actividad desmetilasa m6A de ALKBH10B participa en el proceso de transición floral actuando sobre las proteínas que se encargan del proceso de floración, controlando su vida media, así como su acumulación (Reichel *et al.*, 2019).

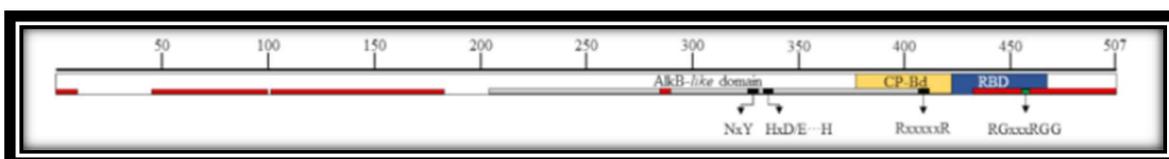


Figura 4. Estructura multidominio de la ALKBH9B de *Arabidopsis thaliana*. Las IDR predichas están marcadas en rojo. El dominio AlkB está marcado en gris. Los dominios de unión a la proteína de cubierta y al RNA aparecen en amarillo y azul, respectivamente. Alvarado-Marchena *et al.*, 2021.

## 1.5. Dominios AlkB en virus ssRNA de polaridad positiva

### 1.5.1. Conocimiento actual

Como se ha comentado anteriormente, ciertos virus de ssRNA<sup>+</sup> que infectan plantas codifican un homólogo AlkB que repara DNA y, preferiblemente, RNA (Fedeles *et al.*, 2015). Estos virus son en su mayoría miembros de las familias *Closteroviridae*, *Alphaflexiviridae* y *Betaflexiviridae*. (Moore y Meng, 2019). En estos, el dominio AlkB se encuentra codificado dentro de la poliproteína replicasa. (van den Born *et al.*, 2008). Los virus de RNA tienen genomas compactos con una capacidad limitada para la codificación, por lo que cualquier proteína o dominio que se mantenga a través de la evolución y la selección natural debe tener función que resulte importante en la supervivencia del virus. Sin embargo, la baja identidad de los dominios AlkB de virus (31%) ha hecho muy complicado sus análisis (Moore y Meng, 2019). No obstante, van den Borne *et al.* (2008) confirmaron la actividad desmetilasa de estos dominios, los cuales reactivaron eficientemente genomas de bacteriófagos metilados cuando fueron expresados en *E. coli*, indicando así que se tratan de dominios funcionales (van den Born *et al.*, 2008).

La distribución irregular de los dominios AlkB en los genomas virales hace difícil desentrañar su historia evolutiva exacta, que requeriría necesariamente múltiples eventos de adquisición o pérdida entre los diferentes taxones virales. Como AlkB está presente casi universalmente en los organismos celulares, pero solo en una minoría de virus que infectan a plantas, es probable que la adquisición de este dominio se produjera en único fenómeno de adquisición a través de transferencia horizontal de genes (Martelli *et al.*, 2007). Como los dominios AlkB de virus tienen mayor similitud con las AlkB de bacterias frente a las AlkB de plantas, es probable este gen se obtuviera a partir de una fuente bacteriana, posiblemente un patógeno o simbiote de plantas (van den Born *et al.*, 2008). Debido a la mayor prevalencia del dominio AlkB en los virus de la familia *Flexiviridae*, que más tarde se ha dividido en las familias *Alphaflexiviridae*, *Betaflexiviridae* y *Gammaflexiviridae*, es muy posible que un ancestro de esta fuera el que recibió este dominio por

primera vez (Martelli *et al.*, 2007). Posteriormente, el dominio pudo haberse diseminado a través de la especiación que se producía como resultado de la diversificación génica tras la entrada en un nuevo huésped o a través de la transferencia horizontal de genes con otros virus que coexistían en un mismo huésped (Moore y Meng, 2019).

Es de destacar que la mayor parte de los virus que contienen el dominio AlkB son importantes patógenos de plantas perennes y leñosas, tales como los árboles frutales, ocasionando cuantiosas pérdidas económicas (Moore y Meng, 2019).

## **1.5.2. Familias de virus que contienen el dominio AlkB**

Como hemos visto anteriormente, los virus de ssRNA<sup>+</sup> que infectan a plantas que poseen un dominio AlkB se encuentran principalmente en las familias *Closteroviridae*, *Alphaflexiviridae* y *Betaflexiviridae*. (Moore y Meng, 2019)

### **1.5.2.1. *Closteroviridae***

La familia *Closteroviridae* está formada por una serie de especies virales que infectan a plantas, principalmente dicotiledóneas, y que poseen unas partículas largas, filamentosas y de construcción helicoidal (650–2,200 nm en longitud y 12 nm de diámetro), así como un genoma de RNA de cadena simple que puede ser mono, di o tripartito. Se caracterizan por poseer uno de los genomas más largos dentro de los virus que infectan a plantas (13-19Kb). Actualmente la familia está conformada por los géneros *Closterovirus*, *Ampelovirus*, *Crinivirus* y *Velarivirus*. Su transmisión se produce principalmente a través de insectos, esencialmente áfidos, moscas blancas, cochinillas y cocoideos. (Fuchs *et al.*, 2020). De esta familia, los homólogos de AlkB han sido identificados únicamente en miembros del género *Ampelovirus*. En su genoma, las pautas de lectura abierta (ORF) 1A y 1B codifican la poliproteína replicasa que contiene los tres dominios universales necesarios para la replicación viral: metiltransferasa (MTR), helicasa (HEL) y polimerasa (POL). El dominio AlkB se encuentra insertado dentro de esta proteína, entre los dominios MTR y HEL, característica compartida con el resto de las AlkB virales (Moore, 2018). La mayoría de los miembros de los *Ampelovirus* infectan huéspedes leñosos, en los que, según su hospedador, provocan distintos efectos tales como el moteado, amarilleo y enrojecimiento de hojas, el picado del tallo, el marchitamiento o infecciones asintomáticas (Fuchs *et al.*, 2020)

### **1.5.2.2. *Alphaflexiviridae***

La familia *Alphaflexiviridae* está conformada por una serie de especies virales que infectan a plantas y hongos. Como su nombre indica, los virus de esta familia poseen un virión flexible y filamentosos (470-800 nm de longitud y 10-15 nm de diámetro) y un genoma formado por una única hebra positiva simple de RNA de 5.4 a 9 Kb. Actualmente, esta familia pertenece al orden de los *Tymovirales* y está conformada por los géneros *Allexivirus*, *Lolavirus*, *Mandarivirus*, *Platypovirus*, *Potexvirus*, *Botrexvirus* y *Sclerodarnavirus*, siendo estos dos últimos los que infectan a hongos (Kreuze *et al.*, 2020). Enfermedades como el mosaico y manchas anulares han sido asociadas a esta familia. Similarmente al resto de AlkB virales, el dominio AlkB se encuentra codificado en todos los casos conocidos hasta la fecha dentro de la replicasa, entre los dominios MTR y HEL (Moore, 2018).

### 1.5.2.3. *Betaflexiviridae*

La mayor parte de los virus que contienen el dominio AlkB en su genoma pertenecen a la familia *Betaflexiviridae*. (Moore, 2018). Similarmente a los virus de *Alphaflexiviridae* (anteriormente ambas familias, junto con *Gammaflexiviridae*, eran una única con el nombre de *Flexiviridae*) los miembros poseen un virión flexible y filamentoso (600-1000nm de longitud y 10-15 nm de diámetro.), así como un genoma formado por una sola molécula linear de ssRNA<sup>+</sup> (con un genoma de 5,9-9 Kb). (King *et al.*, 2011). Esta familia es, como *Alphaflexiviridae*, parte del orden de los *Tymovirales*, y está conformada por los géneros *Capillovirus*, *Carlavirus*, *Chordovirus*, *Citrivirus*, *Divavirus*, *Foveavirus*, *Prunevirus*, *Ravavirus*, *Robigovirus*, *Tepovirus*, *Trichovirus*, *Vitivirus* y *Wamavirus*. Con la excepción de *Carlavirus*, las infecciones naturales son en su mayor parte en huéspedes leñosos. (INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES, 2020). Prácticamente todos los géneros poseen virus que contienen el dominio AlkB. La familia posee muchos virus que producen enfermedades en plantas, incluyendo varios que se cree son los agentes etiológicos para la enfermedad de la madera rugosa. Similarmente al resto de familias, la mayoría de los virus de *Betaflexiviridae* portan el dominio AlkB dentro de la poliproteína replicasa, entre los dominios MTR y HEL (Moore, 2018). Dentro de estos, *Carlavirus* es el género donde más frecuentemente aparece el dominio AlkB. (Martelli *et al.*, 2007).

### 1.5.2.4. Otras familias

El dominio AlkB también se ha encontrado en las familias *Secoviridae* y *Potyviridae*, aunque, hasta la fecha solo un virus de cada una ha sido identificado como portador del dominio AlkB (Moore, 2018). Ambas familias están conformadas por virus de RNA + de simple cadena sin envuelta con genomas mono o bipartitos de longitudes similares, (8-11 Kb en *Potyviridae* y 9-13.7Kb en *Secoviridae*). Algunos de sus miembros constituyen importantes patógenos de plantas (Wylie *et al.*, 2017; Thompson *et al.*, 2017).

### 1.5.3. Hipótesis sobre su función

Hasta la fecha, diversos equipos de investigación han hipotetizado acerca de las posibles funciones que puede tener este dominio viral, aunque su naturaleza exacta sigue siendo desconocida (Moore, 2018).

Una de las teorías es que los virus incorporaron este dominio con el fin de contrarrestar uno de los mecanismos de defensas frente a patógenos ampliamente extendido en plantas, el silenciamiento de genes a nivel post-transcripcional (PTGS) (Bratlie y Drabløs, 2005). A lo largo de la historia, las plantas desarrollaron este sistema de defensa frente a virus y elementos génicos transponibles. Sin embargo, esta teoría choca con la reciente adquisición del dominio AlkB, al ser este un mecanismo ancestral (Moore, 2018).

Otra teoría es que los virus adquirieron este dominio con tal de contrarrestar el estrés en plantas causado por los agentes pesticidas. Diversos pesticidas causan un aumento de los niveles de metilación del DNA y del RNA, causando un ambiente de hipermetilación (Moore, 2018). Sin embargo, esta teoría no parece tampoco muy probable debido a que la mayoría de los virus que

contienen AlkB no infectan plantas tratadas con estos pesticidas y, además, no habría dado tiempo a que se produjese tanto la difusión como la divergencia filogenética que muestran las AlkB virales con sus ancestros bacterianos. (van den Born *et al.*, 2008).

Como se ha comentado anteriormente, la mayoría de virus que contienen el dominio AlkB infectan a plantas perennes leñosas. En estas, los virus deben sobrevivir durante años en el floema, un ambiente hostil. van den Born *et al.* (2008) apoyan la idea de que estos virus codifican estas AlkB con el fin de remediar el posible daño por metilación continuado que sufren en este tipo de ambientes, de tal forma que incrementen su propagación mediante la reparación de estas lesiones al entrar a nuevas células (van den Born *et al.*, 2008). Actualmente esta es la teoría más favorecida de la presencia de AlkB en los virus de ssRNA+ que infectan a plantas. (Moore, 2018).



## **2. Objetivos**

El objetivo de este trabajo ha sido la identificación de los dominios AlkB en el máximo número de especies virales posibles con el fin de caracterizar aquellas especies descritas recientemente que poseen el dominio, y posteriormente realizar un análisis filogenético para tratar de buscar patrones y arrojar más luz a su intrigante origen evolutivo, así como realizar alineamientos múltiples y estudiar la conservación de los motivos principales entre las especies virales. Se considera que este objetivo es esencial antes de abordar la demostración de la funcionalidad in vivo de dicho motivo en los genomas de los virus de RNA de plantas.



### 3. Materiales y métodos

#### 3.1. Identificación y preparación de secuencias de proteínas virales que contienen el dominio AlkB

Con el objetivo de encontrar todas las secuencias de virus de ssRNA+ que poseían este dominio, se realizó un Position Search Iterated Blast (PSI-blast) contra la base de datos del NCBI (<https://www.ncbi.nlm.nih.gov/>), primero contra únicamente la RefSeq, y luego contra la Non-Redundant (nr), para obtener el total de especies y capturar secuencias de la máxima calidad. Se usó una secuencia inicial de una proteína AlkB de *E. coli* de la cepa K-12 (Referencia: P05050) de alta calidad, tomada de la base de datos Uniprot (<https://www.uniprot.org/>), restringiéndose los residuos 129 al 210. Se utilizó la extensión PSI-blast ya que esta permite incrementar enormemente la sensibilidad hacia las secuencias que presentan relaciones débiles pero biológicamente relevantes (Altschul *et al.*, 1997).

Se excluyeron de la búsqueda los taxones de *Eukarya*, *Prokarya*, y *Archaea* con el fin de reducir el número de *hits* que no nos interesaban y, para las iteraciones posteriores de la búsqueda, se seleccionaron las secuencias encontradas pertenecientes a virus de ssRNA+ que infectaban a plantas para la creación de una matriz de score para posiciones específicas. Se capturó una entrada para cada especie y se redujo el total de secuencias, eliminando algunas entradas etiquetadas como especies distintas que en realidad no lo eran. A partir de estas, la secuencia de los dominios AlkB se identificó buscando el motivo RXXXXXR y capturando 170 aminoácidos aguas arriba de este, ya que según vieron Moore y Meng (2019) mediante análisis bioinformáticos en virus los dominios AlkB poseen una longitud comprendida entre 150 y 170 aminoácidos (Moore y Meng, 2019). También, con el fin de realizar un estudio comparativo, se capturaron las secuencias de la CP de dichas especies virales, buscando en la base de datos entradas con preferencia por aquellas de mejor calidad.

Para las especies virales estudiadas, se realizó también una búsqueda bibliográfica, con el fin de caracterizar los huéspedes de dichos virus y determinar si estos eran perennes o anuales-bienales y si eran huéspedes herbáceos o leñosos. Para ello se indagó en distintas publicaciones científicas y en algunas bases de datos de plantas como PLANTS (<https://plants.sc.egov.usda.gov/home>), gestionada por el departamento de agricultura de los Estados Unidos (E.E.U.U.) y World Flora Online (<http://www.worldfloraonline.org/>), anteriormente conocida como ThePlantList.

Por último, para ciertas especies virales que son característicos patógenos de árboles frutales se capturo su dominio AlkB en todos los aislados disponibles en las bases de datos realizando un procedimiento similar: buscando el dominio RXXXXXR y capturando la secuencia 170 aminoácidos aguas arriba. Para poder trabajar con estas secuencias, se crearon ficheros en formato FASTA a partir de los ficheros CSV creados con Excel, utilizando un pequeño script de Python 3 con el editor IDLE 3.9.4 al que, según el caso, se realizaron pequeñas modificaciones para determinar la longitud de secuencia con la que quería trabajar.

### 3.2. Alineamientos múltiple

Se creó un alineamiento múltiple del total de las 146 especies virales analizadas. Al realizar un alineamiento inicial, se observó una alta variabilidad en los primeros residuos de las secuencias, por lo que la longitud se redujo a 140 aminoácidos aguas arriba del motivo RXXXXXR. Esto se hizo con el propósito de poder visualizar las partes más conservadas, principalmente el núcleo catalítico y algunos residuos conservados de la región N-terminal. Con el fin de tener una mejor visualización, se redujo a 58 el número de especies y se hizo otro alineamiento con los mismos parámetros, adjuntado en la **Figura 7**. En este alineamiento se incluyeron las secuencias de la EcAlkB utilizada para la búsqueda, y también de ALKBH5 (referencia: Q6P6C2), obtenida en la base de datos Uniprot.

También se realizaron alineamientos múltiples para los todos los aislados encontrados de ciertos virus que infectan a árboles frutales, concretamente de Apple stem pitting virus (ASPV), Apple chlorotic leaf spot virus (ACLSV), Cherry necrotic rusty mottle virus (CNRMoV), Plum bark necrosis stem pitting-associated virus (PBNSPaV), y Citrus leaf blotch virus (CLBV). Se seleccionaron estos virus ya que estos afectan a cultivos muy importantes en el sector primario y porque poseen un alto número de aislados caracterizados en las bases de datos biológicas. Estos se realizaron de la totalidad de los 170 residuos calculados con tal de observar también el nivel de conservación entre estos de la zona inicial. Para su visualización en el trabajo se escogieron 11 aislados representativos de cada especie.

Los alineamientos se crearon utilizando el programa Mega-X, versión 10.2.5 (Kumar *et al.*, 2018), con el algoritmo Muscle (Edgar, 2004), un algoritmo utilizado para el alineamiento múltiple de proteína, que el programa tiene incorporado. Se utilizó UPGMA como método de *clustering* en el total de iteraciones. Se decidió utilizar este método ya que se ha visto que para los alineamientos múltiples este presenta unos resultados ligeramente mejores que Neighbour-Joining, aunque este último es más útil para estimar los árboles evolutivos (Edgar, 2004). Para la visualización y la edición de los alineamientos se utilizó el programa Jalview, versión 2.11.4.14 (Waterhouse *et al.*, 2009). Para su edición también se utilizó el programa Paint, versión 20h2. Los valores de los parámetros utilizados en los alineamientos múltiples se muestran en la **Figura 5**.

Option	Setting
<b>GAP PENALTIES</b>	
Gap Open	-2.90
Gap Extend	0.00
Hydrophobicity Multiplier	1.20
<b>MEMORY/ITERATIONS</b>	
Max Memory in MB	2048
Max Iterations	16
<b>ADVANCED OPTIONS</b>	
Cluster Method (Iterations 1,2)	UPGMA
Cluster Method (Other Iterations)	UPGMA
Min Diag Length (Lambda)	24

Figura 5. Valores de los parámetros seleccionados para la construcción de los alineamientos múltiples en el programa Mega-X.

### 3.3. Filogenia

Para la creación de los árboles filogenéticos se realizaron alineamientos múltiples con 146 especies virales tanto para el dominio AlkB como para la proteína de cubierta (CP), una proteína viral bien caracterizada. Estos alineamientos se realizaron utilizando el programa Mega-X, versión 10.2.5 (Kumar *et al.*, 2018), con el algoritmo Muscle. Para una mejor construcción del árbol se escogió en este caso Neighbour-joining como método de *clustering*, ya que, tal y como se ha comentado anteriormente, estima mejor los árboles evolutivos (Edgar, 2004). Basándose en las predicciones de tamaño de los dominios AlkB virales, se seleccionó un tamaño de 150 aminoácidos aguas arriba desde el motivo RXXXXXR (Moore y Meng, 2019) Las filogenias fueron construidas utilizando Mega-X, versión 10.2.5 (Kumar *et al.*, 2018). Se construyeron para los dominios AlkB y para la CP. Se calculó para ambas el árbol de máxima verosimilitud (Maximum Likelihood Tree).

Como método de sustitución se eligió el modelo de Jones-Taylor-Thornton. Se escogió el *bootstrapping* como método para testear las filogenias resultantes. Se calcularon un total de 1000 réplicas de *bootstrap* para ambos árboles. A partir de los árboles consenso obtenidos tras el *bootstrapping*, se estableció un umbral del  $\geq 30\%$  en los valores de *bootstrap* para la formación del árbol condensado. Posteriormente los árboles resultantes fueron visualizados y editados utilizando la herramienta online Interactive Tree Of Life (ITOL), versión 6.3 (Letunic y Bork, 2021). Usando esta herramienta se colorearon los acrónimos según su familia y se marcaron con un círculo aquellos nodos cuyo valor de Bootstrap era  $\geq 70\%$ . Los valores de los parámetros utilizados en la construcción de las filogenias se muestran en la **Figura 6**.

Option	Setting
<b>ANALYSIS</b>	
Statistical Method →	Maximum Likelihood
<b>PHYLOGENY TEST</b>	
Test of Phylogeny →	Bootstrap method
No. of Bootstrap Replications →	1000
<b>SUBSTITUTION MODEL</b>	
Substitutions Type →	Amino acid
Model/Method →	Jones-Taylor-Thornton (JTT) model
<b>RATES AND PATTERNS</b>	
Rates among Sites →	Uniform Rates
No of Discrete Gamma Categories →	Not Applicable
<b>DATA SUBSET TO USE</b>	
Gaps/Missing Data Treatment →	Use all sites
Site Coverage Cutoff (%) →	Not Applicable
<b>TREE INFERENCE OPTIONS</b>	
ML Heuristic Method →	Nearest-Neighbor-Interchange (NNI)
Initial Tree for ML →	Make initial tree automatically (Default - NJ/BioNJ)
Initial Tree File →	Not Applicable
Branch Swap Filter →	None
<b>SYSTEM RESOURCE USAGE</b>	
Number of Threads →	4

Figura 6. Valores de los parámetros seleccionados para la construcción de los árboles filogenéticos en el programa Mega-X.



## 4. Resultados y discusión

### 4.1. Identificación y recuento de especies virales con el dominio AlkB

Uno de los principales objetivos de este trabajo ha sido identificar todos los virus de ssRNA+ que infectan a plantas que poseen este dominio AlkB con el fin de ampliar y actualizar el repertorio conocido caracterizado en anteriores trabajos. (Bratlie y Drabløs, 2005; Moore y Meng, 2019). La búsqueda de homología se realizó por tanto en marzo de 2021. A esta fecha, se utilizó una secuencia de *E. coli* como input para realizar un PSI-blast. Se decidió utilizar una secuencia de *E. coli* con el fin de no sesgar la búsqueda utilizando un dominio AlkB viral ya encontrado. Se identificaron un total de 184 especies virales que mostraron scores relevantes en la búsqueda por PSI-blast. Del total de las especies, un número de 76 especies han sido registradas en la base de datos del NCBI en los últimos 5 años, de las cuales 52 pertenecen a los últimos 3. De estas especies se realizó una búsqueda bibliográfica con tal de averiguar las características de los principales huéspedes en cuestión. Como sucede en los trabajos anteriores, la gran mayoría de las especies virales infectan especies perennes. Se creó una tabla con la totalidad de las especies virales, género, familia y las características de su huésped (**Tabla 1 de Anexo**). Aquellas que pertenecen a los últimos 5 años se marcaron con un asterisco. Con el fin de poder trabajar con las secuencias y realizar un análisis de filogenia, el total de secuencias virales se redujo a 146, eliminando del análisis ciertos virus altamente similares que no aportaban nueva información, así como otros virus que no tenían sus secuencias completas o que daban problemas. Las especies virales que se usaron para el análisis aparecen remarcadas en negrita en la **Tabla 1 del Anexo**.

### 4.2. Alineamiento múltiple de especies virales con el Dominio AlkB

Se realizó un alineamiento múltiple con el total de las 146 secuencias utilizadas para los análisis. Debido a la dificultad de visualización que suponía un alineamiento de tales dimensiones, se redujo el número de secuencias a un total de 58 especies que resultaban representativas. De los 170 aminoácidos capturados según las predicciones funcionales realizadas por Moore y Meng (2019), los primeros 30 aminoácidos mostraron un elevado grado de variación, por lo que, como hemos comentado anteriormente, no se incluyeron en el alineamiento (**Figura 7**). En este alineamiento se evaluaron los motivos principales de las secuencias, tomando como base las estructuras de otras proteínas AlkB como EcAlkB, ALKBH2, ALKBH3, ALKBH5 y las AlkB bacterianas del grupo 1B. Como hemos visto anteriormente, pese a la baja conservación que existe entre las proteínas AlkB, las características clave estructurales se mantienen de forma universal en los dominios AlkB funcionales (Moore y Meng, 2019).

#### 4.2.1. Residuos críticos para la funcionalidad de AlkB

En los dominios AlkB de los virus de ssRNA+ que infectan a plantas la zona C-terminal corresponde al núcleo catalítico. Este tiene un tamaño de 100 aminoácidos. (Moore y Meng, 2019). Dentro de este, en línea con las observaciones hechas por van den Born *et al.*, (2008) existen una serie de residuos en las secuencias de los dominios AlkB involucrados en la actividad enzimática que son críticos para la funcionalidad de estos. En los dominios AlkB virales, que se conozca, estos son el motivo HXDX<sub>n</sub>H y el motivo RXXXXXR, siendo X cualquier aminoácido. (van den Born *et al.*, 2008)

#### 4.2.1.1. Residuos que participan en la unión con el ion metálico

En la primera caracterización de EcAlkB, Aas *et al.* (2003) identificaron tres residuos localizados en el núcleo catalítico que presuntamente estaban implicados en la unión del cofactor metálico: La histidina en posición 131, la asparraguina en posición 133 y la histidina en posición 187 (HXDX<sub>n</sub>H) (Aas *et al.*, 2003). En ALKBH5 también aparece el mismo motivo intacto con la misma función en las posiciones 204 (His), 206 (Asp) y 266(His). (Xu *et al.*, 2014; Feng *et al.*, 2014). Hoy sabemos que estos residuos se encuentran conservados de forma universal entre las proteínas AlkB funcionales (Moore y Meng, 2019). El ion metálico es quelado por una molécula de agua, estos residuos y, de forma bidentada, por los residuos carboxilato C1 y cetona c2 del cofactor ketoglutarato. Los estudios anteriores de los dominios AlkB virales pusieron de manifiesto que el motivo HXDX<sub>n</sub>H se mantenía conservado a lo largo de las especies virales (Bratlie y Drabløs, 2005; Moore y Meng, 2019). van den Born *et al.* (2008) vieron que en las secuencias virales las mutaciones tanto de la primera histidina a alanina como del aspártico a alanina provocaron una pérdida de la actividad catalítica (van den Born *et al.*, 2008) Sin embargo, estos grupos caracterizaron un número de virus relativamente bajo, contándose 37 especies. Por tanto, se estudió la presencia de este motivo en las 146 especies virales analizadas con el fin de ver si estos patrones se mantenían.

El motivo HXD se mantuvo íntegro en prácticamente todas las secuencias. La única excepción fue Yam Latent Virus (YLV), del género *Carlavirus*, que contenía una tirosina en lugar de la histidina. No se conoce el efecto de una sustitución por tirosina, pero debido a la ausencia de carga positiva en el nuevo aminoácido y a la universalidad del motivo, es probable que este dominio AlkB no sea funcional. Al profundizar en su huésped, vemos que este virus ha sido extraído de *Dioscorea opposita*, una especie de ñame conocida como ñame chino. Esta es un tubérculo que se suele cosechar a los 6 meses de su plantación (Andrés *et al.*, 2016). Por tanto, este virus no debe mantenerse por largos periodos de infección, y que su dominio AlkB no sea funcional concuerda con la teoría de que estos dominios son útiles en virus que infectan plantas perennes y leñosas (Moore y Meng, 2019).

Moore y Meng (2019) propusieron a su vez una extensión del motivo HXD a HXDDE, viéndose que después del motivo aparecían de forma generalizada estos dos aminoácidos cargados negativamente. Con el fin de comprobarlo se estudió su presencia en las secuencias AlkB. Se vio que el aspártico en posición 4 no se mantuvo intacto en todas las secuencias. En 46 de los aislados apareció una asparraguina en su lugar, apareciendo en aislados virales pertenecientes a distintos géneros y familias y con huéspedes de características similares. Curiosamente en *Closteroviridae* la mayoría de los virus estudiados tenían este cambio. A parte de estos, solo hubo otra excepción: una sustitución por un ácido glutámico en Shallot virus X (ShVX), del género *Allexivirus*. En cuanto al glutámico de la posición 5, este se mantuvo íntegramente conservado en las secuencias con la excepción de tres especies: ShVX, Narcissus common latent virus (NCLV), y Passiflora latent virus (PLV), siendo estos dos últimos *Carlavirus*. ShVX presentó una glutamina en lugar del ácido glutámico. ShVX no presentaba por tanto ninguno de los dos residuos pertenecientes a la extensión del motivo. Sin embargo, su huésped, la chalota, es un bulbo que, pese a ser perenne, se cultiva de forma anual. Es por esto por lo que, si la hipótesis de que el dominio AlkB tiene su función en las infecciones a largo plazo de plantas leñosas, no esperamos que este dominio AlkB posea una función y por tanto no es de sorprender que esté sometido a selección negativa. Por ello, no es un virus determinante en la caracterización de submotivos dentro de las AlkB virales. NCLV mostró una leucina en lugar del glutámico. Al igual que en el caso anterior, el narciso (*Narcissus tazetta*) es un bulbo que, pese a que es perenne, se cultiva anualmente, por lo que

tampoco esperamos que posea un dominio AlkB funcional. En cuanto a PLV, el aislado mostró una glicina en lugar del glutámico. La pasionaria es una planta trepadora cuyo fruto es el maracuyá. Esta planta es perenne y de tallo leñoso. Sin embargo, al indagar en su entrada, vemos que este aislado viral fue anteriormente transferido a plantas de quinoa (*Chenopodium quinoa*), una hierba anual. Por tanto, esta sustitución podría deberse a que tras su transferencia el dominio AlkB este acumulando mutaciones a un ritmo elevado. Por ello, vemos que en aquellos virus donde esperamos que el dominio AlkB tenga un papel en la supervivencia este ácido glutámico se encuentra altamente conservado.

Por otro lado, la histidina final del motivo (H 187 en EcAlkB y H 266 en ALKBH5) también se conservó en casi la totalidad de las secuencias virales. En ALKBH5, la mutación de este residuo por una alanina provoca daños en su actividad enzimática. (Zheng *et al.*, 2014). Solo se observaron dos excepciones: Grapevine berry inner necrosis virus (GBINV), del género *Trichovirus*, y NCLV. GBINV mostró en su lugar una arginina. Todo y que la arginina es un aminoácido de carga positiva como la histidina y no se conocen los efectos de dicha mutación sobre la actividad catalítica, es probable que este dominio AlkB no sea funcional, ya que este motivo se conserva de forma universal en todos los homólogos de AlkB. Esto resulta sorprendente puesto que la vid es una especie perenne y leñosa y esperaríamos encontrar un dominio AlkB funcional. Al profundizar más a fondo en la entrada del NCBI (id: YP\_004293216.1) vemos que la secuencia viral se ha obtenido tras inocular plantas de tabaco (*Nicotiana occidentalis*) con partículas virales obtenidas de una vid enferma y luego se ha transferido y mantenido en quinoa. Ambas son plantas herbáceas anuales o de corta vida. Por tanto, que este dominio AlkB no fuera funcional concordaría con la hipótesis principal de la funcionalidad viral de AlkB. El otro caso fue el de NCLV, que mostró una tirosina en lugar de una histidina. Como en casos anteriores, aunque no exista una caracterización de esta concreta mutación, es probable que se trate de un dominio AlkB no funcional. Esto resulta coherente con lo planteado anteriormente, ya que como hemos visto narciso es un bulbo que se cultiva anualmente y, por tanto, no esperamos que este dominio posea actividad. Además, en el alineamiento se ha observado que este motivo caracterizado por el motivo H parece extenderse a un motivo HXV. Esta valina apareció en un total de 129 secuencias de las 146 analizadas. En el resto de los casos, esta valina aparece sustituida por una leucina o isoleucina, siendo ambos residuos hidrofóbicos como la valina. Por tanto, el patrón de HXZ donde Z es un residuo hidrofóbico que puede ser una valina, una leucina o una isoleucina se mantuvo en todas las secuencias.

En conclusión, podemos observar que la totalidad de los aislados de virus que infectan a plantas perennes y leñosas (a excepción de GBINV, aunque como hemos visto este fue transferido a otros huéspedes) mostraron el motivo HXDX<sub>n</sub>H totalmente conservado entre sus secuencias. Así mismo, dentro de este la extensión de HXD a HXD(D/N) aparece también conservado íntegramente en aquellas secuencias AlkB de aislados que infectan a huéspedes leñosos, de los cuales se espera que sean funcionales, al igual que la extensión en la última histidina a HXZ recién caracterizada.

#### **4.2.1.2. Residuos que participan en la unión del cofactor $\alpha$ -ketoglutarato**

Entre las AlkB, además de con la quelación del ion metálico, el cofactor  $\alpha$ -ketoglutarato se estabiliza con una serie de residuos. Entre estos, un motivo altamente conservado entre las AlkB es el motivo RXXXXXR, que aparece en los residuos 206-210 de EcAlkb. (Moore y Meng, 2019). El motivo RXXXXXR, caracterizado anteriormente en los anteriores estudios, se mantuvo en su

totalidad en todas las secuencias virales. Acerca de este, Moore y Meng (2019) propusieron una corrección de este motivo a RXXSXTXR por encontrarlo en todas las secuencias analizadas. De este motivo, la treonina correspondiente a la cuarta posición se mantuvo conservada íntegramente en la totalidad de las secuencias. La serina en la posición 2 también apareció en la gran mayoría de los casos (97,8%). Las únicas excepciones fueron Actinidia seed borne latent virus (ASBLV) y Caucasus prunus virus (CPV), del género *Prunovirus*, y Camellia ringspot associated virus 1 (CRSaV1), que aún no tiene un género asignado. Estos tres aislados tienen en su lugar una asparraguina, que al igual que la serina y la treonina es un residuo polar sin carga. Además, Moore y Meng (2019) observaron que entre los residuos los aminoácidos variables designados con X correspondían con aminoácidos hidrofóbicos. Esto, aunque sucedía en la mayor parte de las secuencias estudiadas, no ocurría en todos los casos, hallándose 37 aislados que mostraban otro tipo de residuo en su lugar. Los huéspedes de los virus que poseían estos residuos distintivos no mostraron ningún patrón común, encontrándose tanto en virus de distintos géneros y familias y que infectan tanto a plantas anuales como perennes y de estas tanto leñosas como herbáceas. Por tanto, aunque en la mayor parte de los casos este patrón se mantiene, no puede generalizarse como una propiedad de las AlkB virales.

#### **4.2.2. Otros residuos conservados dentro de los dominios AlkB virales**

Además de los motivos críticos para la función enzimática de AlkB, existen algunos residuos y motivos que se encuentran altamente conservados entre las secuencias de los dominios AlkB virales.

Como se ha visto antes, en la región N-terminal de la EcAlkB se encuentra el NRL formado por 90 residuos (Moore y Meng, 2019). En este dominio, existen ciertos aminoácidos que participan en el reconocimiento específico de sustrato, como los son el triptófano de la posición 69 y la tirosina de la posición 76. Uno de los residuos de este dominio que se vio que está conservado en las secuencias de los AlkB de virus es la tirosina 76 de EcAlkB, que el alineamiento aparece en la posición 40 en los dominios virales (van den Born *et al.*, 2008). En nuestro alineamiento, este residuo apareció en todas las secuencias virales con una única excepción, Grapevine rupestris stem pitting-associated virus (GRSPaV), del género *Foveavirus*, que en su lugar mostró una histidina, el cual ya había sido caracterizado en los alineamientos de Moore y Meng (2019). Otros residuos altamente conservados que eran comunes a otras proteínas AlkB fueron caracterizados en los anteriores estudios. Algunos de estos residuos son la arginina que aparece en posición 25, que también aparece en las ALKBH2 y 3 y en las AlkB bacterias del grupo 1B, y el triptófano en posición 53, que son compartidas por todas estas proteínas AlkB además de por la EcAlkB (van den Born *et al.*, 2008). Algo similar sucede con el motivo en bacterias conformado por LXNXY, que en el alineamiento se encuentra en la posición 85. Este motivo cambia para los dominios AlkB virales, siendo en su lugar LXQXY. La fenilalanina en posición 122 también es otro residuo que se encuentra conservado y que se mantiene en EcAlkb. La metionina de la posición 165 también es un residuo con un alto porcentaje de identidad, el cual pese a no existir en la EcAlkB si aparecen en otras AlkB bacterianas, principalmente aquellas del grupo 1B. La treonina en posición 114 y la glicina en 118 también poseen un alto nivel de identidad, siendo estos residuos exclusivos de los dominios AlkB virales. Todos los residuos y motivos nombrados tuvieron un porcentaje de identidad superior al 90%.



### 4.3. Alineamientos múltiples de aislados de virus que infectan árboles frutales

La fruta producida por numerosas especies perennes leñosas es esencial para la salud y la nutrición de los seres humanos. Un gran número de virus taxonómicamente distintos son detectados en este tipo de cultivos, y son un impedimento significativo para su producción y sostenibilidad. (Moore y Meng, 2019). Como hemos comentado anteriormente, en estudios anteriores sobre patógenos que infectan estos cultivos parecen tener dominios AlkB funcionales. (Bratlie y Drabløs, 2005; van den Born *et al.* 2008; Moore y Meng, 2018). Si la hipótesis de que este dominio se utiliza para evitar los daños de metilaciones en los virus que sobreviven en huéspedes perennes y leñosos en infecciones duraderas es correcta, esperaríamos por tanto que este dominio fuera funcional en todos los casos, estando sujeto a selección positiva. Por ello, hemos analizado los residuos críticos de estos, ya que no existe conocimiento del efecto de las mutaciones en los otros residuos conservados. De esta forma, se observaron los motivos HXDX<sub>n</sub>H y RXXXXXR de las secuencias de los dominios de múltiples aislados de 5 virus (ASPV, ACLSV, CLB, CNRMoV, y PBNPaV) que son patógenos de árboles frutales, con el fin de comprobar si la funcionalidad de AlkB se mantiene en todos los aislados.

#### 4.3.1. Apple stem pitting virus

Se encontraron un total de 43 aislados de ASPV caracterizados en las principales bases de datos biológicas (NCBI y EBI). De estos se realizó un alineamiento múltiple. Entre las distintas secuencias se vio un alto nivel de conservación. Los motivos HXDX<sub>n</sub>H y RXXXXXR se mantuvieron intactos en la mayor parte de los casos. Sin embargo, hubo dos excepciones: los aislados ABV56203.1 y CDZ92711.

El aislado ABV56203.1 mostró una pérdida total del motivo HXDX<sub>n</sub>H: la histidina inicial es sustituida por una glutamina, el aspártico por una tirosina y la histidina final por una leucina. Como hemos visto anteriormente estos residuos son de vital importancia para la actividad enzimática de AlkB, por lo que puede predecirse que este dominio AlkB no posee actividad catalítica. Además, dentro de todas las secuencias, este aislado fue el único que rompió el patrón RXSXTXR dentro del motivo RXXXXXR final, todo y que si conservó las argininas inicial y final. Si nos fijamos en el resto de la secuencia y la comparamos con el resto de aislados, puede verse a simple vista que particularmente este aislado presenta un alto grado de variabilidad respecto al resto de secuencias. Investigando en su entrada, vemos que la secuencia de este aislado ha sido obtenida a partir de una translación conceptual de una muestra viral de peral (*Pyrus communis*). No hay mucha más información sobre dicha entrada. Esto resulta intrigante porque el peral es una planta perenne y leñosa.

Por otro lado, está el aislado CDZ92711.1. Este mostró una pérdida del motivo HXD dentro del motivo HXDX<sub>n</sub>H, aunque conservó la histidina final. La histidina inicial se vio sustituida por una metionina, mientras que el ácido aspártico lo fue por asparraguina. En este caso el dominio RXSXTXR se conservó íntegramente como en el resto de las secuencias. Sin embargo, la falta del motivo HXD nos lleva a pensar que el dominio AlkB de este aislado viral no debería ser funcional. Al indagar en su entrada, se vio que la secuencia fue obtenida a partir de la secuenciación de un genoma de ASPV obtenida de un ejemplar de manzano (*Malus Domestica*, variedad Gold Spur), un huésped leñoso y perenne. Por tanto, la pérdida de función de este dominio resulta también peculiar.

El alineamiento reducido se ha adjuntado en la **Figura 1** del **Anexo**.

### 4.3.2. Apple chlorotic leaf spot virus

Se encontraron 46 aislados registrados de ACLSV en las bases de datos biológicas. Entre estos, se vio un alto nivel de conservación de las secuencias. Todos estos aislados mostraron en sus secuencias los motivos  $HXD\text{X}_n\text{H}$  y  $RXXXXXR$  totalmente conservados. Dentro de este último, el patrón viral  $RXS\text{X}\text{TXR}$  se mantuvo también entre todos los aislados. Sin embargo, el patrón de residuos hidrofóbicos situados en las posiciones X se vio en muy pocas secuencias. La mayor parte de los aislados mostraron en la segunda posición una treonina. Uno de ellos, CCK73640.1, mostró además una prolina en la posición 6. Por tanto, entre las secuencias de este virus este patrón no se mantiene. El alineamiento reducido se ha plasmado en la **Figura 2 del Anexo**.

### 4.3.3. Citrus leaf blotch virus

Se encontró un total de 20 aislados de CLBVB registrados en las principales bases de datos biológicas. Todos ellos mantuvieron en su totalidad los principales motivos responsables de la actividad enzimática, tanto  $HXD\text{X}_n\text{H}$  como  $RXXXXXR$ . Dentro de este último, todos los aislados mostraron el patrón  $RXS\text{X}\text{TXR}$ , así como residuos hidrofóbicos en las posiciones X. Curiosamente, se observó una elevada conservación de prácticamente todos los residuos de la secuencia, incluyendo la región de los primeros 30 aminoácidos. Vemos por tanto que la funcionalidad del dominio AlkB parece mantenerse en todos los aislados, y, además, las secuencias poseen un alto nivel de conservación. El alineamiento reducido puede verse en la **Figura 3 del Anexo**.

### 4.3.4. Cherry necrotic rusty mottle virus

Un total de 11 aislados de CNRMoV se encontraron en las principales bases de datos biológicas. En estas, los principales motivos  $HXD\text{X}_n\text{H}$  y  $RXXXXXR$  se mantuvieron totalmente conservados. Dentro de este último, el patrón  $RXS\text{X}\text{TXR}$  se observó en todos los casos. Lo mismo ocurrió para los residuos hidrofóbicos de las posiciones X. En general, todas las secuencias mostraron un nivel bastante alto de similitud en la toda su extensión. El alineamiento reducido se ha plasmado en la **Figura 4 del Anexo**.

### 4.3.5. Plum bark necrosis stem pitting associated virus

Para PBNPaV, se encontró un total de 20 aislados registrados en las bases de datos biológicas. Entre estos, la mayor parte mostraron un alto porcentaje de conservación entre sí, curiosamente incluso la región más variable que corresponde a los primeros 30 residuos. Los principales motivos correspondientes a la actividad catalítica AlkB ( $HXD\text{X}_n\text{H}$  y  $RXXXXXR$ ) se encontraron conservados en la mayoría de las secuencias. En este caso hubo una sola excepción, CDM63857.1. Este aislado mostró una sustitución en el motivo  $HXD$ , teniendo una histidina en lugar del ácido aspártico. Es posible por tanto que este dominio AlkB sea no funcional. Al indagar en su origen, se vio que las partículas virales fueron obtenidas y secuenciadas a partir de hojas de albaricoque (*Prunus armeniaca*). Siendo este un huésped perenne y leñoso esto resulta sorprendente.

Además, pese a que el motivo RXXXXXR se mantuvo en todas las secuencias, la extensión de este a RXSXTXR fue truncada en un aislado, AIH06941.1, el cual posee una arginina en la tercera posición. Además, otros 2 aislados rompieron el patrón común del resto de poseer residuos hidrofóbicos en X. Uno fue el comentado anteriormente, CDM63857.1, que como hemos visto carecía a su vez del motivo HXD. Este mostró en la cuarta posición un triptófano. El otro fue el aislado AIH06934.1, que mostró una tirosina en la sexta posición.

El alineamiento reducido se ha adjuntado en la **Figura 5** del **Anexo**.

#### 4.4. Filogenias

El origen e historia evolutiva de la inclusión de los dominios AlkB en los virus de ssRNA+ que infectan a plantas es una cuestión intrigante. En anteriores estudios, se concluyó que la integración del dominio era un fenómeno reciente, producido en un único evento de adquisición probablemente por un ancestro de las familias *Betaflexiviridae*, posiblemente del linaje *Carlavirus* debido al alto número de especies virales de este género que lo contienen. (van den Born *et al.*, 2008; Moore y Meng, 2019). En el estudio de Moore y Meng (2019) las filogenias de los dominios AlkB mostraban patrones de agrupación distintos a los que mostraban ciertas proteínas que se usan en la replicación y transcripción de virus como es la replicasa. Mientras que esta última mostraba una filogenia discreta agrupada acorde a las clasificaciones taxonómicas de familia, el dominio AlkB mostraba patrones dispersos con miembros de las distintas familias agrupadas en diferentes clados. Esta diferencia en los patrones de agrupación es compatible con la hipótesis de que la adquisición del dominio AlkB se produjo en un único evento de integración con la posterior diseminación del dominio mediante transferencia horizontal (Moore y Meng, 2019). En este trabajo se calcularon las filogenias tanto del dominio AlkB como de la CP en un total de 146 aislados virales. La ampliación de las filogenias anteriores con las especies AlkB tanto novedosas como anteriormente no caracterizadas puede resultar útil para comprobar si estos patrones vistos anteriormente se mantienen. La filogenia de la CP se utilizó para compararse con la de AlkB por ser una proteína conservada entre las especies virales esencial para su ciclo vital.

Entre ambos árboles existen sendas diferencias. Lo primero que resulta llamativo del árbol de los dominios AlkB (**Figura 8**) son los bajos valores de *bootstrap* que presentan la mayor parte de los nodos. En contraste con la filogenia de la CP (**Figura 9**), donde todas las especies virales han sido clasificadas en algún clado superando por tanto el umbral de establecido del 30%, un total de 25 especies no se incluyeron en ningún clado. Además, vemos que hay muy pocos nodos con un alto nivel de confianza que incluyan un número elevado de aislados. El máximo número de especies que se han incluido en un clado con un *bootstrap* superior al 70% ha sido de únicamente 9, en contraste con la de la CP donde en un mismo clado de alta confianza se ha agrupado a más de 50 especies.

La filogenia de la CP mostró un agrupamiento discreto acorde a la taxonomía. Los miembros de las distintas familias se agruparon en respectivos clados comunes. Las únicas excepciones fueron dos virus: Garlic yellow virus (GYV) de *Alphaflexiviridae* y Sweet potato virus F (SPVF) de *Potyviridae*. Respecto a la familia *Alphaflexiviridae*, excepto GYV, todos los miembros del género *Allexivirus* se agruparon en un clado conjunto junto con la única especie de *Lolavirus*, Lolium latent virus (LoLV) y el aún no clasificado dentro de la misma familia Escobaria Virus (EV). Los aislados de *Potexvirus* se agruparon en dos clados distintos, donde en uno de ellos se incluyó el virus Ambrosia asymptomatic virus 1, un miembro de la familia *Flexiviridae* sin

caracterizar. Los dos miembros de *Mandarivirus* se agruparon juntos en uno de estos clados junto con los *Potexvirus*. De forma similar ocurrió con la familia *Betaflexiviridae*. Todos los miembros del género *Carlavirus* aparecieron en un mismo clado con un nodo inicial con alto valor de *bootstrap*, donde también se incluyeron los virus GYV y SPVF. Curiosamente estos virus aparecieron en subclados agrupados con especies que poseen el mismo o similares huéspedes. Algo similar ocurrió con el género *Citivirus*, sus miembros se agruparon en un único clado con la excepción de dos virus, Humulus Lupulus Citivirus (HLciV1) y Avocado citivirus 1 (AciV). Lo mismo ocurrió con los miembros de *Foveavirus*, *Prunevirus* y *Robigovirus*. De igual manera, con la excepción de Fig Latent Virus 1 (FLV1), todos los miembros de *Trichovirus* compartieron un mismo clado. El resto de los virus no clasificados de esta familia se repartieron entre los distintos clados de esta. La familia *Closteroviridae* mostró un patrón similar, todos sus miembros (del género *Ampelovirus*) compartieron un mismo clado. La única excepción fue el único miembro de la familia que no está clasificado en ningún género, Actinidia virus 1 (AV1). Respecto a *Potyviridae*, a excepción del ya mencionado SPVF, sus dos miembros restantes también compartieron un clado. Como podemos observar, la clasificación taxonómica resulta muy exacta en esta filogenia, llegando en la mayoría de los casos a clasificarse correctamente incluso por el género.

Sin embargo, la filogenia de AlkB dio lugar a un paisaje distinto. A diferencia del caso anterior, los miembros de las distintas familias virales no se agruparon conjuntamente en la mayor parte de los casos, mostrando una distribución mucho más dispersa. Dentro de *Alphaflexiviridae*, los *Allexivirus* aparecieron totalmente separados en distintos clados. Papaya virus A (PaVA) apareció solo. ShVX apareció en un pequeño clado de alta confianza junto a Garlic yellow mosaic-associated virus (GYMaV), un *Carlavirus* que, curiosamente, como ShVX, es un patógeno de las especies de *Allium*. Por otro lado, Alfalfa virus S (AVS) apareció en un clado junto a dos *Ampelovirus*: Persimmon ampelovirus (PamV) y PBNSPaV. Los dos *Mandarivirus* aparecieron solos juntos en un clado. En cuanto a los *Potexvirus*, algunos de estos aparecieron en clados conjuntos, sin embargo, algunos de sus miembros aparecieron en otros clados, como *Allium virus X*, que apareció solo en el único clado de un alto número de especies; Papaya mosaic virus (PaMV), que curiosamente apareció en un clado de alta significación junto a American hop latent virus (AHLV); y GYV, que como en el caso de la filogenia de la CP apareció junto a especies de *Carlavirus* que infectan a plantas de *Allium*. El único virus del género *Lolavirus*, LoLV, apareció en un clado junto a Grapevine leafroll-associated virus 1 (GLaV1), Grapevine virus A (GVA), Actinidia virus A (AVA) y Heracleum latent virus (HeLV), siendo el primero un *Ampelovirus*, y el resto *Vitivirus*. Dentro de *Betaflexiviridae* sucedió algo similar. Los *Carlavirus* aparecieron en la mayoría de los casos, viéndose un elevado número de ellos en clados solitarios y el resto mayormente en clados donde únicamente había dos aislados. De estos destacó un clado de alto nivel de confianza donde curiosamente se agruparon los distintos *Carlavirus* que infectan a especies *Allium* junto con GYV como hemos visto anteriormente. Los miembros de *Citivirus*, con la excepción de AciV, si se agruparon conjuntamente en un clado, junto con otros virus no clasificados como Citrus leaf blotch virus Ac (CLBVAc) y Salvia Divinorum RNA virus 1 (SDV1). Por otro lado, los *Vitivirus* aparecieron esparcidos, encontrándose 3 en el clado mencionado anteriormente junto al aislado de *Lolavirus*, dos en un clado conjunto (Arracacha virus V (AVV) y Arracacha latent virus V (ALVV)) y los otros dos en otro clado de forma similar (Blueberry green mosaic associated virus (BGMAV) y Blackberry virus A (BVA)). En cuanto a *Foveavirus*, estos también aparecieron separados, viéndose cuatro de ellos en un clado conjunto, tres de ellos en otro (curiosamente los tres patógenos de la vid), dos por otro lado y el resto por separado, excepto en el caso de Rubus virus 1 (RV1) que se emparejó con AV1 de la familia

*Closteroviridae*. Los *Robigovirus*, todos patógenos del cerezo, aparecieron juntos en un clado de alta confianza. Los *Prunavirus* aparecieron por separado excepto Karelinia Prunavirus 1 (KpVA) y Apricot vein clearing associated 1 (AVCaV). Acerca de *Trichovirus*, cinco de estos (ACLSV, Apricot pseudo-chlorotic leaf spot virus (APCLSV), Peach mosaic virus (PMV), Mexico trichovirus (MtrV), y Cherry mottle leaf virus (CMoLV)) aparecieron juntos en un único clado, dos en otro (GBINV y Grapevine pinot gris virus (GPGV)) y FLV1 por separado. La familia *Closteroviridae*, apareció totalmente esparcida entre las diferentes ramas de la filogenia, con la excepción de los anteriormente comentados PamV y PBNSPaV. Lo mismo sucedió con la familia *Potyviridae*, donde sus tres aislados aparecieron totalmente separados en distintos clados. De esta destaca SPVF, que, como sucedía en la filogenia de la CP, se agrupó junto a los dos *Carlavirus* que también parasitan el boniato (*Ipomoea batatas*).

En conclusión, podemos ver que en la distribución de los distintos clados en la filogenia de las AlkB las distintas familias no se agrupan conjuntamente por norma general como en el caso de la CP. Además, curiosamente en muchos casos pueden observarse agrupaciones entre virus taxonómicamente distintos, pero con un huésped similar. Por tanto, esta distribución es coherente con la hipótesis de que el dominio AlkB sucedió por un fenómeno de adquisición único y la posterior transferencia horizontal hacia virus que infectan a los mismos huéspedes de otras familias.

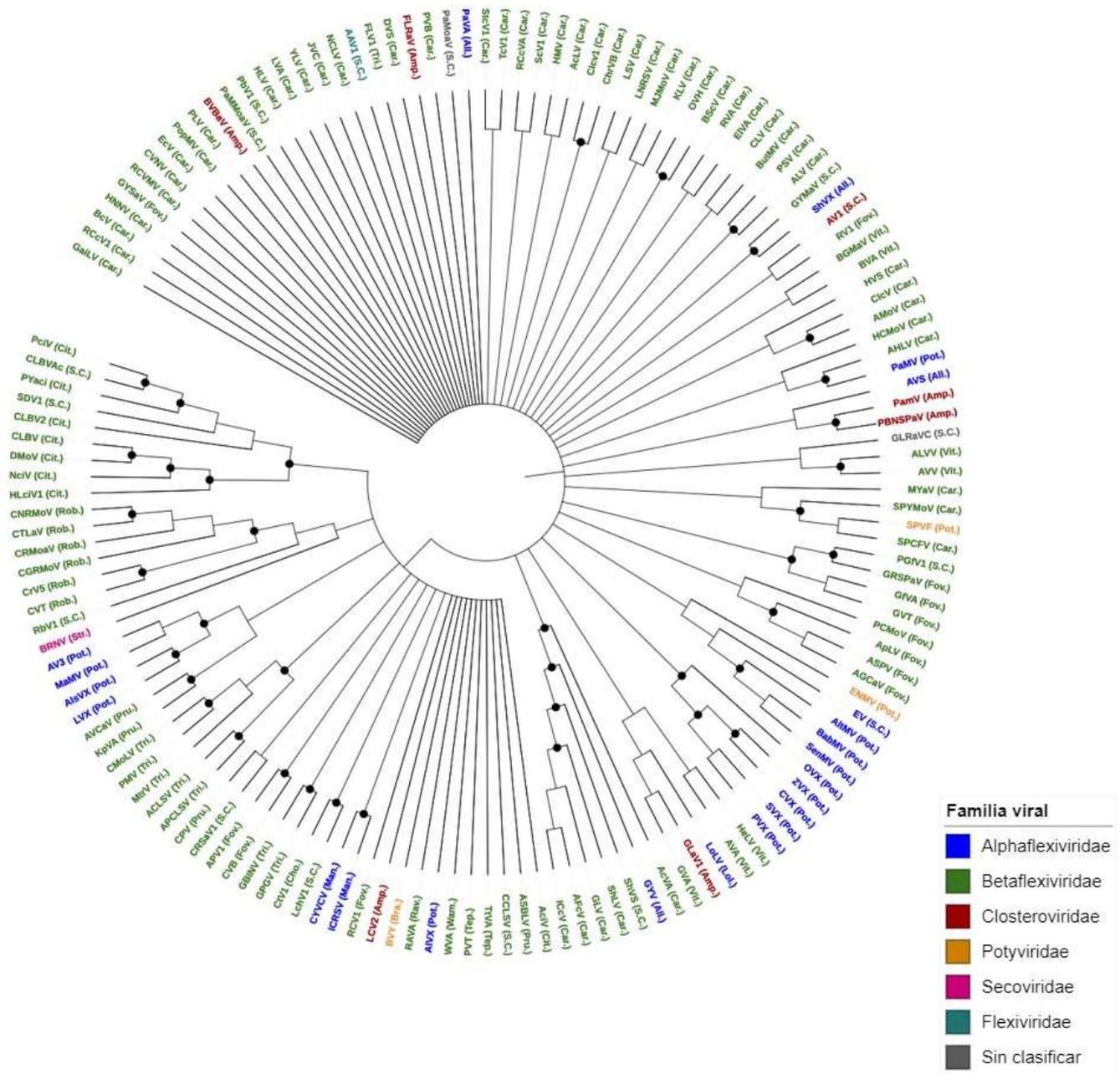


Figura 8. Árbol filogenético de los dominios AlkB virales. Las diferentes familias se representaron usando el esquema de colores indicado. Los géneros aparecen representados entre paréntesis con las tres primeras letras. Aquellos aislados cuyo género no ha sido establecido se han marcado como S.C. (Sin clasificar). Los clados que se muestran poseen un valor de bootstrap  $\geq 30\%$ . Los clados marcados poseen un valor de bootstrap  $\geq 70\%$

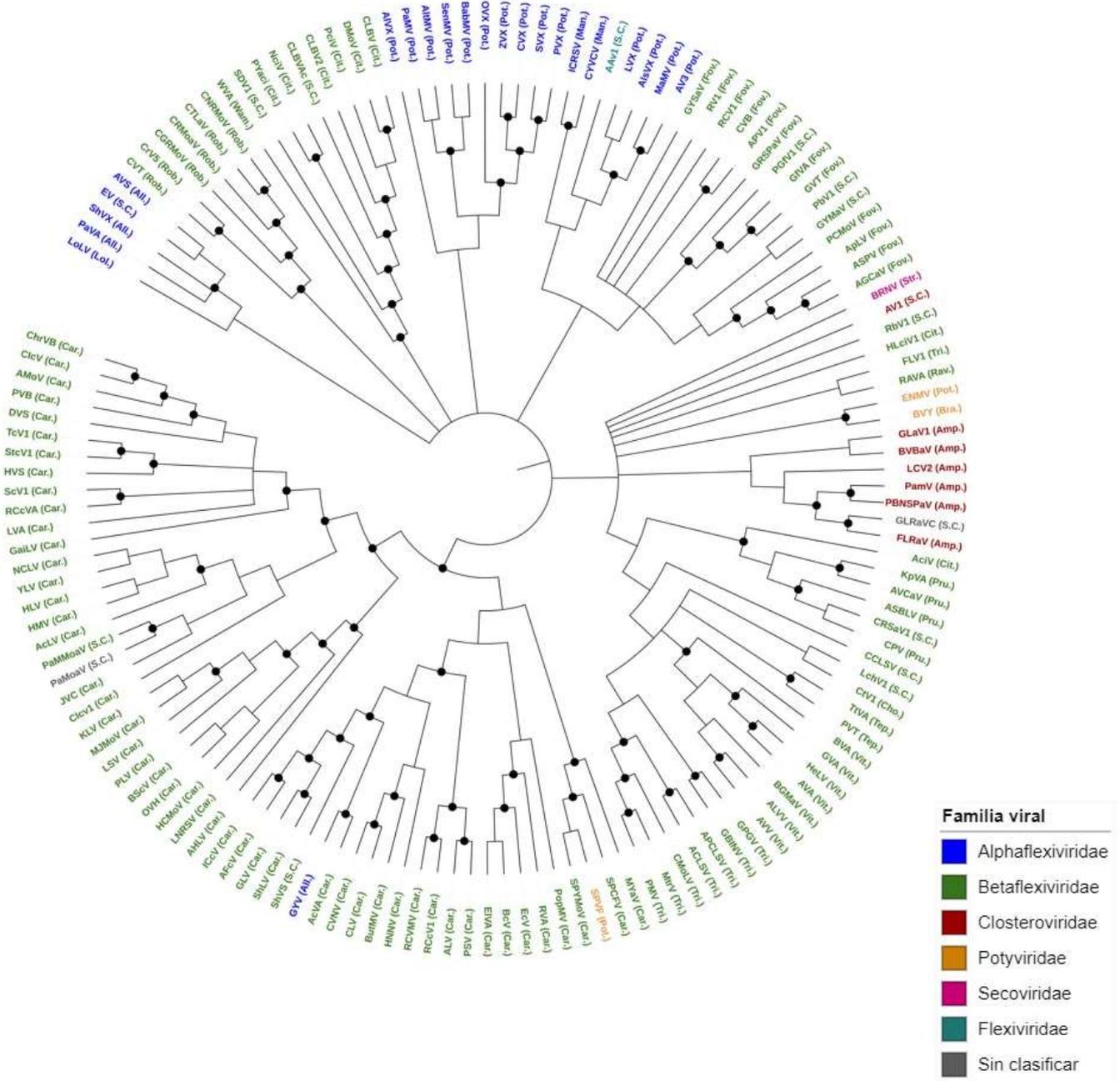


Figura 9. Árbol filogenético de la proteína de cubierta (CP). Las diferentes familias se representaron usando el esquema de colores indicado. Los géneros aparecen representados entre paréntesis con las tres primeras letras. Aquellos aislados cuyo género no ha sido establecido se han marcado como S.C. (Sin clasificar). Los clados que se muestran poseen un valor de bootstrap  $\geq 30\%$ . Los clados marcados poseen un valor de bootstrap  $\geq 70\%$



## 5. Conclusión

En este trabajo se ha identificado el dominio AlkB en un total de 194 especies de virus de ssRNA<sup>+</sup>. Estos pertenecen principalmente a las familias *Alphaflexiviridae*, *Betaflexiviridae* y *Closteroviridae*. Por tanto, se ha ampliado el catálogo de virus conocidos de poseer este patrón, lo cual puede ser útil para posteriores estudios acerca de la funcionalidad de este dominio.

Además, en los alineamientos múltiples generalizados se ha observado un elevado nivel de conservación en los residuos críticos para la funcionalidad del dominio en todos aquellos aislados de los que se esperaba encontrar un dominio AlkB funcional.

Dentro de los aislamientos múltiples de diferentes aislados de virus que infectan a árboles frutales, se ha observado un mantenimiento total de los residuos críticos en los virus CLB<sub>V</sub>, CNRM<sub>oV</sub> y ACL<sub>SV</sub>. Sin embargo, en ASP<sub>V</sub> y PBN<sub>SPaV</sub>, se observaron aislados característicos que mostraron mutaciones en estos residuos clave, y por tanto muy probablemente la posible funcionalidad de esta actividad enzimática. El uso de clones infecciosos de estos virus en los que se muten los correspondientes residuos críticos seguidos de ensayos bioquímicos y biológicos correspondientes pueden arrojar luz sobre la importancia de este dominio en el ciclo de infección de estos virus.

Por último, la filogenia de las proteínas AlkB mantuvo un patrón disperso respecto a la taxonomía en contraposición a la filogenia de las CP. Además, ciertos virus de distintas familias que poseen huéspedes similares se agruparon en clados conjuntos. Estos resultados son coherentes con la hipótesis planteada en anteriores estudios de que la obtención del dominio AlkB se produjo en un único fenómeno de adquisición reciente y este fue transferido a virus de distintas familias mediante transferencia horizontal.



## 6. Bibliografía

- Aas, P. A., Otterlei, M., Falnes, P. Ø., Vågbø, C. B., Skorpen, F., Akbari, M., Sundheim, O., Bjørås, M., Slupphaug, G., Seeberg, E., & Krokan, H. E. (2003). Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature*, 421(6925), 859–863. <https://doi.org/10.1038/nature01363>
- Altschul, S. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
- Alvarado-Marchena, L., Marquez-Molins, J., Martínez-Pérez, M., Aparicio, F., & Pallás, V. (2021). Mapping of Functional Subdomains in the atALKBH9B m6A-Demethylase Required for Its Binding to the Viral RNA and to the Coat Protein of Alfalfa Mosaic Virus. *Frontiers in Plant Science*, 12, 701683. <https://doi.org/10.3389/fpls.2021.701683>
- Andrés, C., AdeOluwa, O. O., & Bhullar, G. S. (2017). Yam (*Dioscorea* spp.). In *Encyclopedia of Applied Plant Sciences* (pp. 435–441). Elsevier. <https://doi.org/10.1016/B978-0-12-394807-6.00177-5>
- Aravind, L., & Koonin, E. V. (2001). The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biology*, 2(3), 8.
- Arribas-Hernandez, L. & Brodersen, P. (2020) Occurrence and functions of m6A and other covalent modifications in plant mRNA. *Plant Physiology*, 182, 79–96.
- Baquero-Perez, B., Geers, D., & Díez, J. (2021). From A to m6A: The Emerging Viral Epitranscriptome. *Viruses*, 13(6), 1049. <https://doi.org/10.3390/v13061049>
- Batista, P. J. (2017). The RNA Modification N<sup>6</sup>-methyladenosine and Its Implications in Human Disease. *Genomics, Proteomics & Bioinformatics*, 15(3), 154–163. <https://doi.org/10.1016/j.gpb.2017.03.002>
- Belmont, P., Constant, J. F., & Demeunynck, M. (2001). Nucleic acid conformation diversity: From structure to function and regulation. *Chemical Society Reviews*, 30(1), 70–81. <https://doi.org/10.1039/a904630e>
- Berlivet, S., Scutenaire, J., Deragon, J.-M., & Bousquet-Antonelli, C. (2019). Readers of the m6A epitranscriptomic code. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1862(3), 329–342. <https://doi.org/10.1016/j.bbagr.2018.12.008>
- Betaflexiviridae - Positive Sense RNA Viruses - Positive Sense RNA Viruses (2011) - ICTV. (2020). *Talk.Ictvonline.Org*. [https://talk.ictvonline.org/ictv-reports/ictv\\_9th\\_report/positive-sense-rna-viruses-2011/w/posrna\\_viruses/241/betaflexiviridae](https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/241/betaflexiviridae)
- Bhat, S. S., Bielewicz, D., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2018). N<sup>6</sup>-methyladenosine (m6A): Revisiting the Old with Focus on New, an *Arabidopsis thaliana* Centered Review. *Genes*, 9(12), 596. <https://doi.org/10.3390/genes9120596>
- Bjørnstad, L. G., Zoppellaro, G., Tomter, A. B., Falnes, P. Ø., & Andersson, K. K. (2011). Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe(II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4. *Biochemical Journal*, 434(3), 391–398. <https://doi.org/10.1042/BJ20101667>
- Bratlie, M. S., & Drabløs, F. (2005). Bioinformatic mapping of AlkB homology domains in viruses. *BMC Genomics*, 6(1), 1. <https://doi.org/10.1186/1471-2164-6-1>

- Cheng, X. (1995). DNA modification by methyltransferases. *Current Opinion in Structural Biology*, 5(1), 4–10. [https://doi.org/10.1016/0959-440X\(95\)80003-J](https://doi.org/10.1016/0959-440X(95)80003-J)
- Courtney, D. G., Kennedy, E. M., Dumm, R. E., Bogerd, H. P., Tsai, K., Heaton, N. S., & Cullen, B. R. (2017). Epitranscriptomic Enhancement of Influenza A Virus Gene Expression and Replication. *Cell Host & Microbe*, 22(3), 377–386.e5. <https://doi.org/10.1016/j.chom.2017.08.004>
- Dang, W., Xie, Y., Cao, P., Xin, S., Wang, J., Li, S., Li, Y., & Lu, J. (2019). N6-Methyladenosine and Viral Infection. *Frontiers in Microbiology*, 10, 417. <https://doi.org/10.3389/fmicb.2019.00417>
- Davis, F. F., & Allen, F. W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. *Journal of Biological Chemistry*, 227(2), 907–915. [https://doi.org/10.1016/S0021-9258\(18\)70770-9](https://doi.org/10.1016/S0021-9258(18)70770-9)
- Deng, X., Chen, K., Luo, G.-Z., Weng, X., Ji, Q., Zhou, T., & He, C. (2015). Widespread occurrence of N 6 -methyladenosine in bacterial mRNA. *Nucleic Acids Research*, 43(13), 6557–6567. <https://doi.org/10.1093/nar/gkv596>
- Desrosiers, R., Friderici, K., & Rottman, F. (1974). Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma Cells. *Proceedings of the National Academy of Sciences*, 71(10), 3971–3975. <https://doi.org/10.1073/pnas.71.10.3971>
- Dimock, K., & Stoltzfus, C. M. (1977). Sequence specificity of internal methylation in B77 avian sarcoma virus RNA subunits. *Biochemistry*, 16(3), 471–478. <https://doi.org/10.1021/bi00622a021>
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., & Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, 485(7397), 201–206. <https://doi.org/10.1038/nature11112>
- Duan, H.-C., Wei, L.-H., Zhang, C., Wang, Y., Chen, L., Lu, Z., Chen, P. R., He, C., & Jia, G. (2017). ALKBH10B Is an RNA N 6 -Methyladenosine Demethylase Affecting Arabidopsis Floral Transition. *The Plant Cell*, 29(12), 2995–3011. <https://doi.org/10.1105/tpc.16.00912>
- Duechler, M., Leszczyńska, G., Sochacka, E., & Nawrot, B. (2016). Nucleoside modifications in the regulation of gene expression: Focus on tRNA. *Cellular and Molecular Life Sciences*, 73(16), 3075–3095. <https://doi.org/10.1007/s00018-016-2217-y>
- 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>
- Esteve-Puig, R., Bueno-Costa, A., & Esteller, M. (2020). Writers, readers and erasers of RNA modifications in cancer. *Cancer Letters*, 474, 127–137. <https://doi.org/10.1016/j.canlet.2020.01.021>
- Falnes, P. Ø., Johansen, R. F., & Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature*, 419(6903), 178–182. <https://doi.org/10.1038/nature01048>
- Farrow, S. C., & Facchini, P. J. (2014). Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. *Frontiers in Plant Science*, 5. <https://doi.org/10.3389/fpls.2014.00524>

- Fedeles, B. I., Singh, V., Delaney, J. C., Li, D., & Essigmann, J. M. (2015). The AlkB Family of Fe(II)/ $\alpha$ -Ketoglutarate-dependent Dioxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond. *Journal of Biological Chemistry*, 290(34), 20734–20742. <https://doi.org/10.1074/jbc.R115.656462>
- Feng, C., Liu, Y., Wang, G., Deng, Z., Zhang, Q., Wu, W., Tong, Y., Cheng, C., & Chen, Z. (2014). Crystal Structures of the Human RNA Demethylase Alkbh5 Reveal Basis for Substrate Recognition. *Journal of Biological Chemistry*, 289(17), 11571–11583. <https://doi.org/10.1074/jbc.M113.546168>
- Fuchs, M., Bar-Joseph, M., Candresse, T., Maree, H. J., Martelli, G. P., Melzer, M. J., Menzel, W., Minafra, A., Sabanadzovic, S., & Report Consortium, I. (2020). ICTV Virus Taxonomy Profile: Closteroviridae. *Journal of General Virology*, 101(4), 364–365. <https://doi.org/10.1099/jgv.0.001397>
- Grosjean, H. (2009). *DNA and RNA Modification Enzymes* (0 ed.). CRC Press. <https://doi.org/10.1201/9781498713153>
- He, L., Li, H., Wu, A., Peng, Y., Shu, G., & Yin, G. (2019). Functions of N6-methyladenosine and its role in cancer. *Molecular Cancer*, 18(1), 176. <https://doi.org/10.1186/s12943-019-1109-9>
- He, P. C., & He, C. (2021). m<sup>6</sup>A RNA methylation: From mechanisms to therapeutic potential. *The EMBO Journal*, 40(3). <https://doi.org/10.15252/embj.2020105977>
- Hotchkiss, R. D. (1948). The Quantitative Separation of Purines, Pyrimidines, and Nucleosides by Paper Chromatography. *Journal of Biological Chemistry*, 175(1), 315–332. [https://doi.org/10.1016/S0021-9258\(18\)57261-6](https://doi.org/10.1016/S0021-9258(18)57261-6)
- Huong, T. T., Ngoc, L. N. T., & Kang, H. (2020). Functional Characterization of a Putative RNA Demethylase ALKBH6 in Arabidopsis Growth and Abiotic Stress Responses. *International Journal of Molecular Sciences*, 21(18), 6707. <https://doi.org/10.3390/ijms21186707>
- Jia, B., Tang, K., Chun, B. H., & Jeon, C. O. (2017). Large-scale examination of functional and sequence diversity of 2-oxoglutarate/Fe(II)-dependent oxygenases in Metazoa. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1861(11), 2922–2933. <https://doi.org/10.1016/j.bbagen.2017.08.019>
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., & He, C. (2011). N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chemical Biology*, 7(12), 885–887. <https://doi.org/10.1038/nchembio.687>
- Johnson, T. B., & Coghill, R. D. (1925). The discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus. *J Am Chem Soc.* 47, 7.
- Kane, S. E., & Beemon, K. (1985). Precise localization of m<sup>6</sup>A in Rous sarcoma virus RNA reveals clustering of methylation sites: Implications for RNA processing. *MOL. CELL. BIOL.*, 5, 9.
- Kataoka, H., Yamamoto, Y., & Sekiguchi, M. (1983). A new gene (alkB) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. *Journal of Bacteriology*, 153(3), 1301–1307. <https://doi.org/10.1128/jb.153.3.1301-1307.1983>

- Kennedy, T. D., & Lane, B. G. (1979). Wheat embryo ribonucleates. XIII. Methyl-substituted nucleoside constituents and 5'-terminal dinucleotide sequences in bulk poly(A)-rich RNA from imbibing wheat embryos. *Canadian Journal of Biochemistry*, 57(6), 927–931. <https://doi.org/10.1139/o79-112>
- Kennedy, E. M., Bogerd, H. P., Kornepati, A. V. R., Kang, D., Ghoshal, D., Marshall, J. B., Poling, B. C., Tsai, K., Gokhale, N. S., Horner, S. M., & Cullen, B. R. (2017). Posttranscriptional m6A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host & Microbe*, 22(6), 830. <https://doi.org/10.1016/j.chom.2017.11.010>
- Kim, J., Shim, S., Lee, H., & Seo, P. J. (2020). M6A mRNA Modification as a New Layer of Gene Regulation in Plants. *Journal of Plant Biology*, 63(2), 97–106. <https://doi.org/10.1007/s12374-020-09239-5>
- King, A. M. Q., Lefkowitz, E., Adams, M. J., & Carstens, E. B. (2011). *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses (English Edition) (1.a ed.)*. Elsevier.
- Kreuze, J. F., Vaira, A. M., Menzel, W., Candresse, T., Zavriev, S. K., Hammond, J., Hyun Ryu, K., & Report Consortium, I. (2020). ICTV Virus Taxonomy Profile: Alphaflexiviridae. *Journal of General Virology*, 101(7), 699–700. <https://doi.org/10.1099/jgv.0.001436>
- Krug, R. M., Morgan, M. A., & Shatkin, A. J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *Journal of Virology*, 20(1), 45–53. <https://doi.org/10.1128/jvi.20.1.45-53.1976>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lavi, S., & Shatkin, A. J. (1975). Methylated simian virus 40-specific RNA from nuclei and cytoplasm of infected BSC-1 cells. *Proceedings of the National Academy of Sciences*, 72(6), 2012–2016. <https://doi.org/10.1073/pnas.72.6.2012>
- Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293–W296. <https://doi.org/10.1093/nar/gkab301>
- Li, Z., Shi, J., Yu, L., Zhao, X., Ran, L., Hu, D., & Song, B. (2018). N 6 -methyl-adenosine level in *Nicotiana tabacum* is associated with tobacco mosaic virus. *Virology Journal*, 15(1), 87. <https://doi.org/10.1186/s12985-018-0997-4>
- Liao, S., Sun, H., & Xu, C. (2018). YTH Domain: A Family of N 6 -methyladenosine (m 6 A) Readers. *Genomics, Proteomics & Bioinformatics*, 16(2), 99–107. <https://doi.org/10.1016/j.gpb.2018.04.002>
- Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G. M., Bansal, V., Wang, Y., Mason, C. E., & Rana, T. M. (2016). Dynamics of the human and viral m6A RNA methylomes during HIV-1 infection of T cells. *Nature Microbiology*, 1(4), 16011. <https://doi.org/10.1038/nmicrobiol.2016.11>
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., & Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods*, 12(8), 767–772. <https://doi.org/10.1038/nmeth.3453>

- Lu, W., Tirumuru, N., St. Gelais, C., Koneru, P. C., Liu, C., Kvaratskhelia, M., He, C., & Wu, L. (2018). N6-Methyladenosine-binding proteins suppress HIV-1 infectivity and viral production. *Journal of Biological Chemistry*, 293(34), 12992–13005. <https://doi.org/10.1074/jbc.RA118.004215>
- Marcinkowski, M., Pilżys, T., Garbicz, D., Steciuk, J., Zugaj, D., Mielecki, D., Sarnowski, T. J., & Grzesiuk, E. (2020). Human and Arabidopsis alpha-ketoglutarate-dependent dioxygenase homolog proteins—New players in important regulatory processes. *IUBMB Life*, 72(6), 1126–1144. <https://doi.org/10.1002/iub.2276>
- Martelli, G. P., Adams, M. J., Kreuze, J. F., & Dolja, V. V. (2007). Family Flexiviridae: A Case Study in Virion and Genome Plasticity. *Annual Review of Phytopathology*, 45(1), 73–100. <https://doi.org/10.1146/annurev.phyto.45.062806.094401>
- Martínez-Pérez, M., Aparicio, F., López-Gresa, M. P., Bellés, J. M., Sánchez-Navarro, J. A., & Pallás, V. (2017). Arabidopsis m6A demethylase activity modulates viral infection of a plant virus and the m6A abundance in its genomic RNAs. *Proceedings of the National Academy of Sciences*, 114(40), 10755–10760. <https://doi.org/10.1073/pnas.1703139114>
- Martínez-Pérez, M. (2020). Una proteína de Arabidopsis thaliana perteneciente a la familia AlkB de dioxigenasas dependientes de hierro (II) y 2-oxoglutarato interfiere en el ciclo infeccioso del virus del mosaico de la alfalfa. (Trabajo Final de Máster). Instituto de Biología Celular y Molecular de Plantas (IBMCP). Valencia.
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., & Jaffrey, S. R. (2012). Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and near Stop Codons. *Cell*, 149(7), 1635–1646. <https://doi.org/10.1016/j.cell.2012.05.003>
- Meyer, K. D., & Jaffrey, S. R. (2017). Rethinking m6A Readers, Writers, and Erasers. *Annual Review of Cell and Developmental Biology*, 33(1), 319–342. <https://doi.org/10.1146/annurev-cellbio-100616-060758>
- Moore, C. (2018). Prediction of the Molecular Boundary and Evolutionary History of Novel Viral AlkB Domains Using Homology Modeling and Principal Component Analysis. 105.
- Moore, C., & Meng, B. (2019). Prediction of the molecular boundary and functionality of novel viral AlkB domains using homology modelling and principal component analysis. *Journal of General Virology*, 100(4), 691–703. <https://doi.org/10.1099/jgv.0.001237>
- Müller, T. A., Meek, K., & Hausinger, R. P. (2010). Human AlkB homologue 1 (ABH1) exhibits DNA lyase activity at abasic sites. *DNA Repair*, 9(1), 58–65. <https://doi.org/10.1016/j.dnarep.2009.10.011>
- Ontiveros, R. J., Stoute, J., & Liu, K. F. (2019). The chemical diversity of RNA modifications. *Biochemical Journal*, 476(8), 1227–1245. <https://doi.org/10.1042/BCJ20180445>
- Ougland, R., Zhang, C.-M., Liiv, A., Johansen, R. F., Seeberg, E., Hou, Y.-M., Remme, J., & Falnes, P. Ø. (2004). AlkB Restores the Biological Function of mRNA and tRNA Inactivated by Chemical Methylation. *Molecular Cell*, 16(1), 107–116. <https://doi.org/10.1016/j.molcel.2004.09.002>
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., & Conrad, N. K. (2017). The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell*, 169(5), 824–835.e14. <https://doi.org/10.1016/j.cell.2017.05.003>

- Rausch, C., Hastert, F. D., & Cardoso, M. C. (2020). DNA Modification Readers and Writers and Their Interplay. *Journal of Molecular Biology*, 432(6), 1731–1746. <https://doi.org/10.1016/j.jmb.2019.12.018>
- Reichel, M., Köster, T., & Staiger, D. (2019). Marking RNA: M6A writers, readers, and functions in Arabidopsis. *Journal of Molecular Cell Biology*, 11(10), 899–910. <https://doi.org/10.1093/jmcb/mjz085>
- Samson, L., & Cairns, J. (1977). A new pathway for DNA repair in Escherichia coli. *Nature*, 267(5608), 281–283. <https://doi.org/10.1038/267281a0>
- Shi, H., Wei, J., & He, C. (2019). Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. *Molecular Cell*, 74(4), 640–650. <https://doi.org/10.1016/j.molcel.2019.04.025>
- Songe-Møller, L., van den Born, E., Leihne, V., Vågbø, C. B., Kristoffersen, T., Krokan, H. E., Kirpekar, F., Falnes, P. Ø., & Klungland, A. (2010). Mammalian ALKBH8 Possesses tRNA Methyltransferase Activity Required for the Biogenesis of Multiple Wobble Uridine Modifications Implicated in Translational Decoding. *Molecular and Cellular Biology*, 30(7), 1814–1827. <https://doi.org/10.1128/MCB.01602-09>
- Stoltzfus, C. M., & Dimock, K. (1976). Evidence for Methylation of B77 Avian Sarcoma Virus Genome RNA Subunits. *J. VIROL.*, 18, 10.
- Thomason, A. R., Brian, D. A., Velicer, L. F., & Rottman, F. M. (1976). Methylation of high-molecular-weight subunit RNA of feline leukemia virus. *Journal of Virology*, 20(1), 123–132. <https://doi.org/10.1128/jvi.20.1.123-132.1976>
- Thompson, J. R., Dasgupta, I., Fuchs, M., Iwanami, T., Karasev, A. V., Petrzik, K., Sanfaçon, H., Tzanetakis, I., van der Vlugt, R., Wetzel, T., Yoshikawa, N., & ICTV Report Consortium. (2017). ICTV Virus Taxonomy Profile: Secoviridae. *Journal of General Virology*, 98(4), 529–531. <https://doi.org/10.1099/jgv.0.000779>
- Tirumuru, N., Zhao, B. S., Lu, W., Lu, Z., He, C., & Wu, L. (2016). N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *ELife*, 5, e15528. <https://doi.org/10.7554/eLife.15528>
- Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., & Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature*, 419(6903), 174–178. <https://doi.org/10.1038/nature00908>
- Visvanathan, A., & Somasundaram, K. (2018). mRNA Traffic Control Reviewed: N6-Methyladenosine (m<sup>6</sup>A) Takes the Driver's Seat. *BioEssays*, 40(1), 1700093. <https://doi.org/10.1002/bies.201700093>
- Wakisaka, K. T., Muraoka, Y., Shimizu, J., Yamaguchi, M., Ueoka, I., Mizuta, I., Yoshida, H., & Yamaguchi, M. (2019). Drosophila Alpha-ketoglutarate-dependent dioxygenase AlkB is involved in repair from neuronal disorders induced by ultraviolet damage. *NeuroReport*, 30(15), 1039–1047. <https://doi.org/10.1097/WNR.0000000000001323>
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., Ren, B., Pan, T., & He, C. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, 505(7481), 117–120. <https://doi.org/10.1038/nature12730>

- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9), 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>
- Weaver, R. F. (2012). *Molecular biology* (5th ed). McGraw-Hill.
- Wei, C.-M., Gershowitz, A., & Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell*, 4(4), 379–386. [https://doi.org/10.1016/0092-8674\(75\)90158-0](https://doi.org/10.1016/0092-8674(75)90158-0)
- Wylie, S. J., Adams, M., Chalam, C., Kreuze, J., López-Moya, J. J., Ohshima, K., Praveen, S., Rabenstein, F., Stenger, D., Wang, A., Zerbini, F. M., & ICTV Report Consortium. (2017). ICTV Virus Taxonomy Profile: Potyviridae. *Journal of General Virology*, 98(3), 352–354. <https://doi.org/10.1099/jgv.0.000740>
- Xu, C., Liu, K., Tempel, W., Demetriades, M., Aik, W., Schofield, C. J., & Min, J. (2014). Structures of Human ALKBH5 Demethylase Reveal a Unique Binding Mode for Specific Single-stranded N6-Methyladenosine RNA Demethylation. *Journal of Biological Chemistry*, 289(25), 17299–17311. <https://doi.org/10.1074/jbc.M114.550350>
- Yang, Y., Hsu, P. J., Chen, Y.-S., & Yang, Y.-G. (2018). Dynamic transcriptomic m6A decoration: Writers, erasers, readers and functions in RNA metabolism. *Cell Research*, 28(6), 616–624. <https://doi.org/10.1038/s41422-018-0040-8>
- Zaccara, S., Ries, R. J., & Jaffrey, S. R. (2019). Reading, writing and erasing mRNA methylation. *Nature Reviews Molecular Cell Biology*, 20(10), 608–624. <https://doi.org/10.1038/s41580-019-0168-5>
- Zhang, K., Zhuang, X., Dong, Z., Xu, K., Chen, X., Liu, F., & He, Z. (2021). The dynamics of N6-methyladenine RNA modification in interactions between rice and plant viruses. *Genome Biology*, 22(1), 189. <https://doi.org/10.1186/s13059-021-02410-2>
- Zhang, T., Wang, Z., Hu, H., Chen, Z., Liu, P., Gao, S., Zhang, F., He, L., Jin, P., Xu, M., Chen, J., & Yang, J. (2021). Transcriptome-Wide N6-Methyladenosine (m6A) Profiling of Susceptible and Resistant Wheat Varieties Reveals the Involvement of Variety-Specific m6A Modification Involved in Virus-Host Interaction Pathways. *Frontiers in Microbiology*, 12, 656302. <https://doi.org/10.3389/fmicb.2021.656302>
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C. J., Vågbø, C. B., Shi, Y., Wang, W.-L., Song, S.-H., Lu, Z., Bosmans, R. P. G., Dai, Q., Hao, Y.-J., Yang, X., Zhao, W.-M., Tong, W.-M., Wang, X.-J., Bogdan, F., ... He, C. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*, 49(1), 18–29. <https://doi.org/10.1016/j.molcel.2012.10.015>
- Zheng, H., Li, S., Zhang, X., & Sui, N. (2020). Functional Implications of Active N6-Methyladenosine in Plants. *Frontiers in Cell and Developmental Biology*, 8, 291. <https://doi.org/10.3389/fcell.2020.00291>
- Zhu, W., Wang, J., Xu, Z., Cao, M., Hu, Q., Pan, C., Guo, M., Wei, J., & Yang, H. (2019). Detection of N6-methyladenosine modification residues (Review). *International Journal of Molecular Medicine*. <https://doi.org/10.3892/ijmm.2019.4169>



## 7. Anexo

Tabla 1. Tabla de virus de ssRNA de polaridad positiva encontrados que poseen el dominio AlkB, junto con su acrónimo, género, familia, y características del huésped. En negrita aparecen aquellos virus usados para el análisis posterior. Los aislados de virus registrados en los últimos 5 años se han marcado con un asterisco.

Familia	Género	Especie	Acrónimo	Huésped Perenne		Huésped leñoso	
				Sí	No	Sí	No
<i>Alphaflexiviridae</i>	<i>Allexivirus</i>	<b>Alfalfa virus S*</b>	AVS	x			x
		<b>Shallot virus X</b>	ShVX	x			x
		<b>Papaya virus A</b>	PaVA	x		x	
		<b>Garlic yellow virus*</b>	GYV	x			x
	<i>Lolavirus</i>	<b>Lolium latent virus</b>	LoLV	x			x
	<i>Mandarivirus</i>	<b>Citrus yellow vein clearing virus</b>	CYVCV	x		x	
		<b>Indian citrus ringspot virus</b>	ICRSV	x		x	
	<i>Potexvirus</i>	<b>Allium virus X</b>	AIVX	x			x
		<b>Alstroemeria virus X</b>	AlsVX	x			x
		<b>Alternanthera mosaic virus</b>	AltMV	x			x
		<b>Asparagus virus 3</b>	AV3	x			x
		<b>Babaco mosaic virus*</b>	BabMV	x		x	
		<b>Cactus virus X</b>	CVX	x			x
		<b>Clover yellow mosaic virus</b>	CYMV	x			x
		<b>Lettuce virus X</b>	LVX		Anual		x
		<b>Malva mosaic virus</b>	MaMV	x			x
		<b>Opuntia virus X</b>	OVX	x			x
		<b>Papaya mosaic virus</b>	PaMV	x		x	
		<b>Pitaya virus X</b>	PVX	x			x
		Potexvirus sp.*	PtV	?	?	?	?
		<b>Schlumbergera virus X</b>	SVX	x			x
		<b>Senna mosaic virus</b>	SenMV	x			x
		<b>Zygocactus virus X</b>	ZVX	x			x
Sin asignar	<b>Escobaria virus</b>	EV	x			x	
<i>Betaflexiviridae</i>	<i>Capillovirus</i>	<b>Currant virus A</b>	CVA	x		x	
	<i>Carlavirus</i>	<b>Aconitum latent virus</b>	AcLV	x			x
		<b>Alfalfa latent virus</b>	ALV	x			x

Familia	Género	Especie	Acrónimo	Huésped Perenne		Huésped leñoso	
				Sí	No	Sí	No
<i>Betaflexiviridae</i>	<i>Carlavirus</i>	American hop latent virus	AHLV	x			x
		Atractylodes mottle virus	AMoV	x			x
		Blueberry scorch virus	BScV	x		x	
		Butterbur mosaic virus	ButMV	x			x
		Chrysanthemum virus B	ChrVB	x			x
		Chrysanthemum virus R*	ChrVR	x			x
		Coleus vein necrosis virus	CVNV	x			x
		Daphne virus S	DVS	x			x
		Elderberry carlavirus A	EIVA	x		x	
		Elderberry carlavirus B	EIVB	x		x	
		Elderberry carlavirus C	EIVC	x		x	
		Elderberry carlavirus D	EIVD	x		x	
		Elderberry carlavirus E	EIVE	x		x	
		Gaillardia latent virus	GaiLV	x			x
		Garlic latent virus	GLV	x			x
		Helleborus net necrosis virus	HNNV	x			x
		Hop latent virus	HLV	x			x
		Hop mosaic virus	HMV	x			x
		Hydrangea chlorotic mottle virus	HCMoV	x		x	
		Jasmine virus C	JVC	x		x	
		Kalanchoe latent virus	KLV	x			x
		Ligustrum necrotic ringspot virus	LNRSV	x		x	
		Ligustrum virus A	LVA	x		x	
		Lily symptomless virus	LSV	x			x
		Melon yellowing-associated virus	MYaV		Anual		x
		Mirabilis jalapa mottle virus	MJMoV	x			x
		Narcissus common latent virus	NCLV	x			x
		Passiflora latent virus	PLV	x		x	
		Pea streak virus	PSV		Anual		x
		Phlox virus B	PVB	x			x
Phlox virus S	PVS	x			x		

Familia	Género	Especie	Acronimo	Huésped Perenne		Huésped leñoso	
				Sí	No	Sí	No
<i>Betaflexiviridae</i>	<i>Carlavirus</i>	Poplar mosaic virus	PopMV	x		x	
		Red clover vein mosaic virus	RCVMV	x			x
		Sweet potato chlorotic fleck virus	SPCFV	x			x
		Yam latent virus	YLV	x			x
		Shallot latent virus*	ShLV	x			x
		Opuntia virus H	OVH	x			x
		Soybean carlavirus 1*	ScV1		Anual		x
		Ilex cornuta carlavirus*	ICcV	x		x	
		Rose virus A*	RVA	x		x	
		Rose virus B*	RVB	x		x	
		Red clover carlavirus 1*	RCcV1	x			x
		Allium fistulosum carlavirus*	AFcV	x			x
		Allium carlavirus A*	AcVA	x			x
		Carya illinoensis carlavirus 1*	Cicv1	x		x	
		Elm carlavirus*	EcV	x		x	
		Tagetes carlavirus 1*	TcV1	x			x
		Stevia carlavirus 1*	StcV1	x			x
		Red clover carlavirus A*	RCcVA	x			x
		Helenium virus S*	HVS	x			x
		Sweet potato yellow mottle virus	SPYMoV	x			x
		Carnation latent virus*	CLV	x			x
		Chrysanthemum indicum carlavirus*	CicV	x			x
		Birch carlavirus*	BcV	x		x	
	<i>Chordovirus</i>	Carrot Ch virus 1	CtV1		Bienal		x
		Carrot Ch virus 2	CtV2		Bienal		x
	<i>Citivirus</i>	Citrus leaf blotch virus	CLBV	x		x	

Familia	Género	Especie	Acrónimo	Huésped Perenne		Huésped leñoso		
				Sí	No	Sí	No	
<i>Betaflexiviridae</i>	<i>Citivirus</i>	Citrus leaf blotch virus 2*	CLBV2	x		x		
		Dweet mottle virus	DMoV	x		x		
		Avocado citrivirus 1*	AciV	x		x		
		Nandina citrivirus*	NciV	x		x		
		Humulus lupulus citrivirus 1*	HLeiV1	x			x	
		Peony yellowing associated citrivirus*	PYaci	x			x	
		Paeonia citrivirus A*	PciV	x			x	
	<i>Foveavirus</i>	Apple green crinkle associated virus	AGCaV	x		x		
		Apple stem pitting virus	ASPV	x		x		
		Apricot latent virus	ApLV	x		x		
		Asian prunus virus 1	APV1	x		x		
		Asian prunus virus 2	APV2	x		x		
		Asian prunus virus 3	APV3	x		x		
		Grapevine rupestris stem pitting-associated virus	GRSPaV	x		x		
		Peach chlorotic mottle virus	PCMoV	x		x		
		Rubus canadensis virus 1	RCV1	x		x		
		Grapevine virus T*	GVT	x		x		
		Grapevine foveavirus A*	GfVA	x		x		
		Cherry virus B*	CVB	x		x		
		Garlic yellow stripe associated virus*	GYSaV	x			x	
		Camellia ringspot associated virus 4*	CRSaV4	x		x		
		Rubus virus 1*	RV1	x			x	
		<i>Prunevirus</i>	Actinidia seed borne latent virus*	ASBLV	x		x	
			Apricot vein clearing associated virus	AVCaV	x		x	
	Caucasus prunus virus		CPV	x		x		
	Karelinia prunevirus A*		KpVA	x			x	
	<i>Ravavirus</i>	Ribes americanum virus A*	RAVA	x		x		
	<i>Robigovirus</i>	Cherry green ring mottle virus	CGRMoV	x		x		

Familia	Género	Especie	Acrónimo	Huésped Perenne		Huésped leñoso	
				Sí	No	Sí	No
<i>Betaflexiviridae</i>	<i>Robigovirus</i>	Cherry necrotic rusty mottle virus	CNRMoV	x		x	
		Cherry rusty mottle associated virus	CRMoaV	x		x	
		Cherry twisted leaf associated virus	CTLaV	x		x	
		Cherry robigovirus 5*	CrV5	x		x	
		Cherry virus Turkey*	CVT	x		x	
	<i>Tepovirus</i>	Prunus virus T	PVT	x		x	
		Trichosanthes tepovirus A*	TtVA	x		x	
	<i>Trichovirus</i>	Apple chlorotic leaf spot virus	ACLSV	x		x	
		Apricot pseudo-chlorotic leaf spot virus	APCLSV	x		x	
		Cherry mottle leaf virus	CMoLV	x		x	
		Grapevine berry inner necrosis virus	GBINV	x		x	
		Grapevine Pinot gris virus	GPGV	x		x	
		Peach mosaic virus	PMV	x		x	
		Mexico trichovirus*	MtrV	x		x	
	<i>Vitivirus</i>	Fig latent virus 1	FLV1	x		x	
		Actinidia virus A	AVA	x		x	
		Actinidia virus B	AVB	x		x	
		Arracacha virus V*	AVV	x			x
		Blackberry virus A*	BVA	x		x	
		Grapevine virus A	GVA	x		x	
		Grapevine virus B	GVB	x		x	
		Grapevine virus E	GVE	x		x	
		Grapevine virus F	GVF	x		x	
		Grapevine virus G*	GVG	x		x	
		Grapevine virus H*	GVH	x		x	
		Grapevine virus I*	GVI	x		x	
	Grapevine virus J*	GVJ	x		x		

Familia	Género	Especie	Acrónimo	Huésped Perenne		Huésped leñoso		
				Sí	No	Sí	No	
<i>Betaflexiviridae</i>	<i>Vitivirus</i>	Grapevine virus K*	GVK	x		x		
		Grapevine virus L*	GVL	x		x		
		Grapevine virus M*	GVM	x		x		
		<b>Heracleum latent virus*</b>	HeLV		Anual-Bienal		x	
		<b>Blueberry green mosaic associated virus*</b>	BGMaV	x		x		
		<b>Arracacha latent virus V*</b>	ALVV	x			x	
	<i>Wamavirus</i>	<b>Watermelon virus A*</b>	WVA		Anual		x	
	Sin asignar	<b>Lettuce chardovirus 1*</b>	LchV1		Anual		x	
		<b>Panax ginseng flexivirus 1*</b>	PGfV1	x		x		
		<b>Salvia divinorum RNA virus 1*</b>	SDV1	x			x	
		<b>Citrus chlorotic leaf spot virus*</b>	CCLSV	x		x		
		<b>Actinidia virus C*</b>	AVC	x		x		
		<b>Shallot virus S*</b>	ShVS	x			x	
		<b>Papaya mild mottle associated virus*</b>	PaMMoaV	x		x		
		<b>Garlic yellow mosaic-associated virus*</b>	GYMaV	x			x	
		<b>Peony betaflexivirus 1*</b>	PbV1	x			x	
		<b>Camellia ringspot associated virus 1*</b>	CRSaV1	x		x		
		Camellia ringspot associated virus 2*	CRSaV2	x		x		
		Camellia japonica associated betaflexivirus 1*	CJabV1	x		x		
		Camellia japonica associated betaflexivirus 3*	CJabV2	x		x		
		Camellia japonica associated betaflexivirus 2*	CJabV3	x		x		
	<b>Rhododendron betaflexivirus 1*</b>	RbV1	x		x			
	<b>Citrus leaf blotch virus Ac*</b>	CLBVAc	x		x			
	<i>Closteroviridae</i>	<i>Ampelovirus</i>	<b>Blackberry vein banding-associated virus</b>	BVBaV	x		x	
			<b>Grapevine leafroll-associated virus 1</b>	GLaV1	x		x	
			Grapevine leafroll-associated virus 3	GLaV3	x		x	
Grapevine leafroll-associated virus 4			GLaV4	x		x		

Familia	Género	Especie	Acrónimo	Huésped Perenne		Huésped leñoso	
				Sí	No	Sí	No
<i>Closteroviridae</i>	<i>Ampelovirus</i>	Grapevine leafroll-associated virus 5	GLaV5	x		x	
		Grapevine leafroll-associated virus 6	GLaV6	x		x	
		Grapevine leafroll-associated virus 10	GLaV10	x		x	
		Grapevine leafroll-associated virus 13	GLaV13	x		x	
		<b>Little cherry virus 2</b>	LCV2	x		x	
		<b>Plum bark necrosis stem pitting-associated virus</b>	PBNPaV	x		x	
		<b>Firespike leaf roll-associated virus*</b>	FLRaV	x		x	
		<b>Persimmon ampelovirus*</b>	PamV	x		x	
		Maize associated ampelovirus*	MaaV		Anual		x
		<b>Actinidia virus 1*</b>	AV1	x		x	
	Sin asignar	<b>Ambrosia asymptomatic virus 1</b>	AAV1	x			x
<i>Potyviridae</i>	Sin asignar	<b>Blackberry virus Y</b>	BVY	x		x	
	<i>Potyvirus</i>	<b>Endive necrotic mosaic virus</b>	ENMV		Anual		x
		<b>Sweet potato virus F*</b>	SPVF	x			x
<i>Secoviridae</i>	<i>Stramovirus</i>	<b>Black raspberry necrosis virus</b>	BRNV	x		x	
Sin asignar	Sin asignar	<b>Papaya mottle associated virus*</b>	PaMoaV	x		x	
		<b>Grapevine leafroll-associated virus Carn</b>	GLRaVC	x		x	
		Narcissus mottling-associated virus	NmotaV	x			x

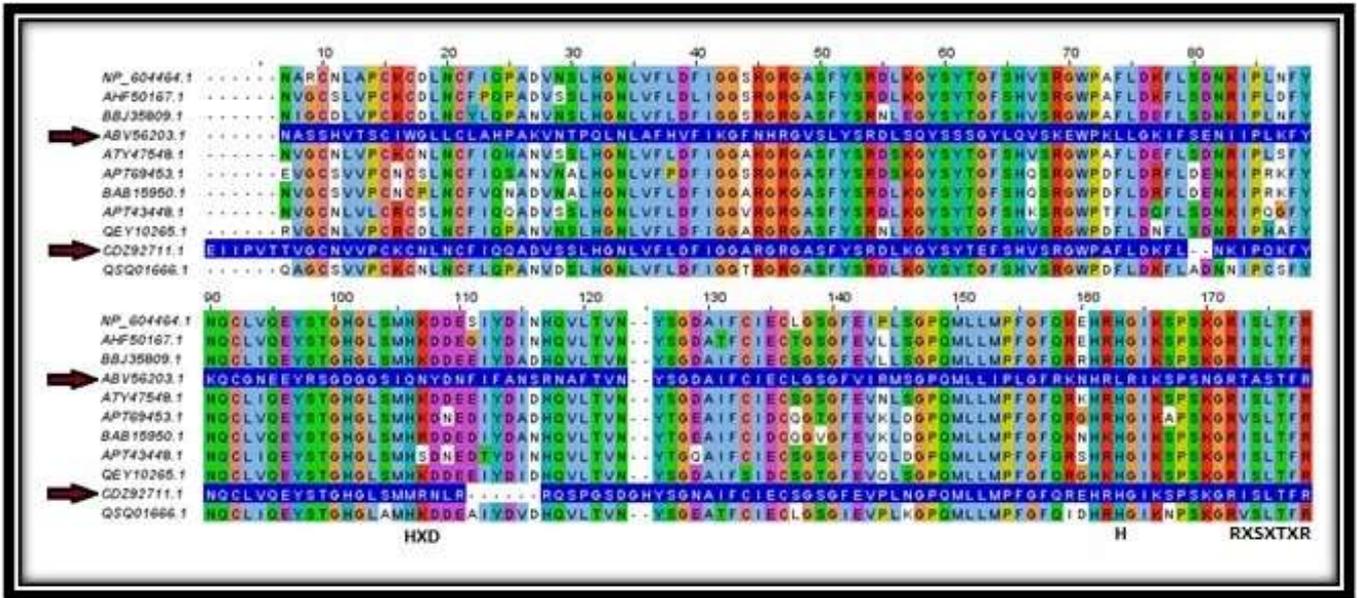


Figura 1. Alineamiento múltiple del dominio AlkB de los aislados de Apple stem pitting virus. Los aislados que poseen mutaciones en los residuos críticos han sido destacados.

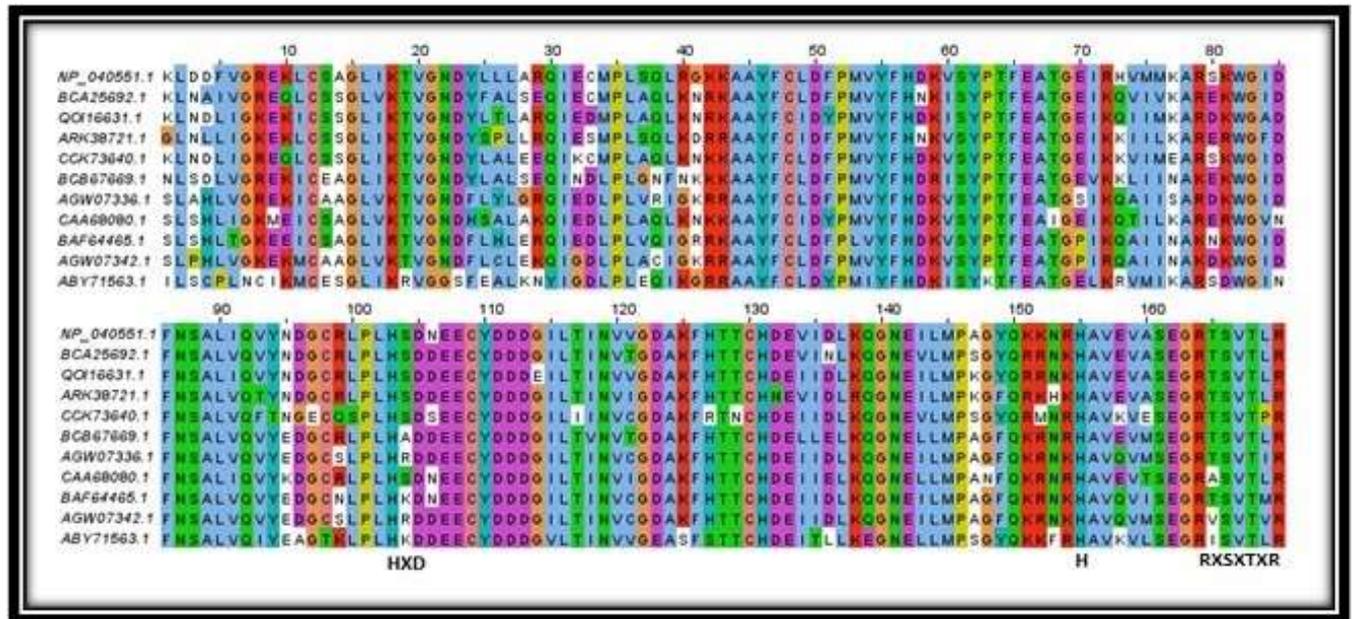


Figura 2. Alineamiento múltiple del dominio AlkB de los aislados de Apple chlorotic leaf spot virus.

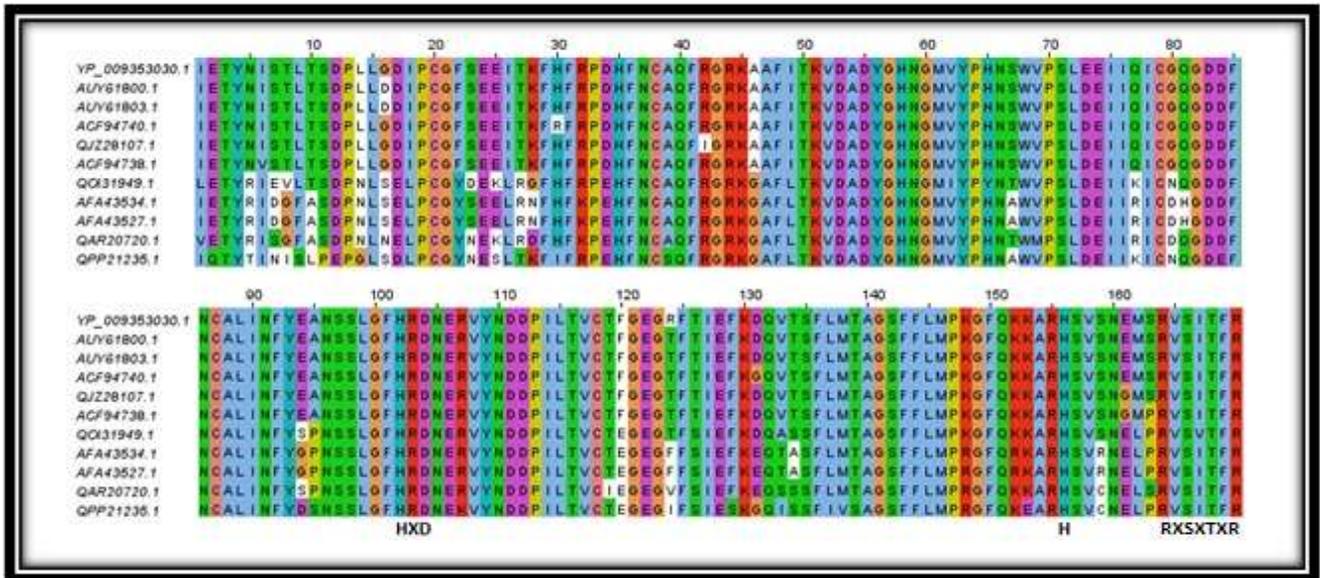


Figura 3. Alineamiento múltiple del dominio AlkB de los aislados de Citrus leaf blotch virus.

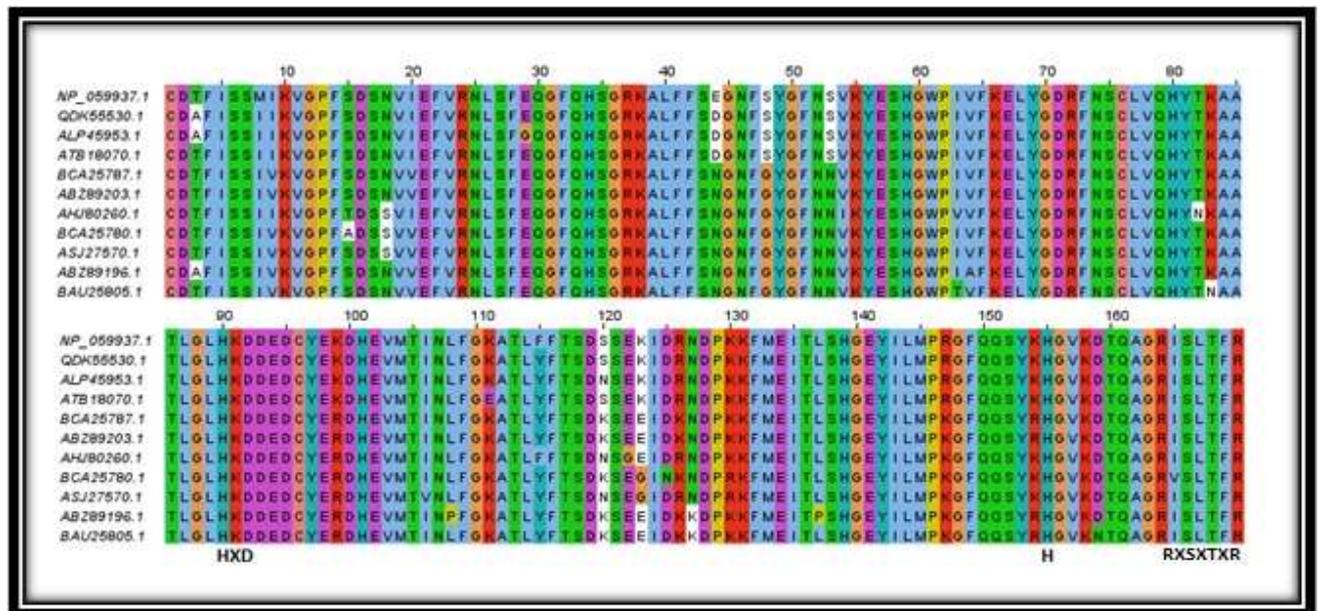


Figura 4. Alineamiento múltiple del dominio AlkB de los aislados de Cherry necrotic rusty mottle virus.

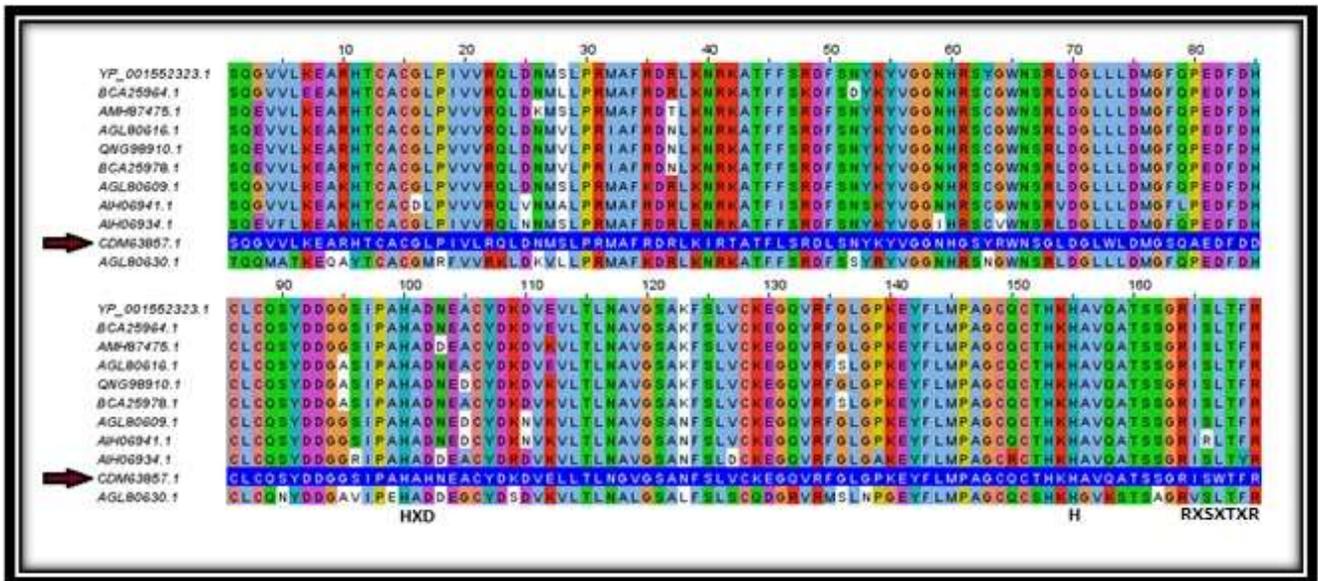


Figura 5. Alineamiento múltiple del dominio AlkB de los aislados de Plum bark necrosis stem pitting-associated virus. El aislado que posee mutaciones en los residuos críticos ha sido destacado.