



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE CIENCIAS

ANÁLISIS FILOGENÉTICO Y MOLECULAR DEL GEN  
*ARGONAUTA9* EN *Marchantia polymorpha*.

TESIS QUE PARA OBTENER EL TÍTULO DE

BIÓLOGA

PRESENTA

CECILIA MONSERRAT LARA MONDRAGÓN

TUTORES

DR. MARIO A. ARTEAGA VÁZQUEZ

DRA. PATRICIA LEÓN MEJÍA



CIUDAD UNIVERSITARIA, 2016



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

1. DATOS DEL ALUMNO

LARA

MONDRAGÓN

CECILIA MONSERRAT

44 37 13 23

UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO

FACULTAD DE CIENCIAS

BIOLOGÍA

308266605

2. DATOS DEL TUTOR

DR.

MARIO ALBERTO

ARTEAGA

VÁZQUEZ

DRA.

PATRICIA

LEÓN

MEJÍA

3. DATOS DEL SINODAL 1

DR.

VÍCTOR MANUEL

VALDÉS

LÓPEZ

4. DATOS DEL SINODAL 2

DRA.

LAURA

KAWASAKI

WATANABE

5. DATOS DEL SINODAL 3

DR.

JOSÉ LUIS

REYES

TABOADA

6. DATOS DEL TRABAJO ESCRITO

ANÁLISIS FILOGENÉTICO Y MOLECULAR

DEL GEN *ARGONAUTA9* EN

*MARCHANTIA POLYMORPHA*

86 P.

2016

PROTEVS ACTVS ADAPTVS



# ÍNDICE



<b>RESUMEN</b>	<b>1</b>
<b>AGRADECIMIENTOS</b>	<b>2</b>
<b>INTRODUCCIÓN</b>	<b>3</b>
<b>1. Antecedentes históricos del silenciamiento mediado por RNAs pequeños</b>	<b>3</b>
<b>2. Silenciamiento mediado por RNAs pequeños en plantas</b>	<b>5</b>
2.1 Clasificación de los RNAs pequeños regulatorios en plantas.....	6
2.2 La maquinaria central del silenciamiento mediado por RNAs pequeños .....	6
<b>3. Las proteínas efectoras del silenciamiento: ARGONAUTAS</b>	<b>8</b>
3.1 Características estructurales y funcionales de las proteínas AGO.....	9
3.1.1 Características bioquímicas de las proteínas AGO .....	9
3.1.2 Características estructurales de las proteínas AGO.....	9
3.1.2.1 El dominio N-terminal .....	9
3.1.2.2 El dominio MID .....	11
3.1.2.3 El dominio PAZ .....	11
3.1.2.4 El dominio PIWI .....	11
3.1.2.5 El dominio L2 .....	13
3.2 Diversidad funcional de las proteínas AGO de <i>A. thaliana</i> .....	13
<b>4. El clado AGO4/AGO6/AGO9 participa en el TGS: Vía de Metilación del DNA dependiente de RNA</b>	<b>15</b>
4.1 La vía de Metilación del DNA dependiente de RNA.....	16
4.1.1 La vía <i>RdDM</i> canónica.....	16
4.1.1.1 Biogénesis de siRNAs dependiente de RNA polimerasa IV.....	16

4.1.1.2	Metilación <i>de novo</i> dependiente de la RNA polimerasa V .....	16
4.1.2	AGO6 en la vía <i>RdDM</i> .....	17
4.1.3	AGO9 en la vía <i>RdDM</i> .....	19
4.2	Importancia biológica de la <i>RdDM</i> .....	20
4.2.1	Control de Elementos Transponibles y Respuesta a Estrés .....	20
4.2.2	Defensa contra patógenos .....	21
4.3	Importancia evolutiva de la <i>RdDM</i> en las plantas terrestres .....	21
4.3.1	Conservación evolutiva de los componentes de la <i>RdDM</i> .....	23
4.3.2	Las RNA polimerasas IV y V como importantes innovaciones evolutivas en la <i>RdDM</i> .....	24
<b>5.</b>	<b>El origen de las embriofitas y <i>Marchantia polymorpha</i> como modelo emergente</b>	<b>26</b>
	<b>PLANTEAMIENTO DEL PROBLEMA</b>	<b>31</b>
	<b>HIPÓTESIS</b>	<b>32</b>
	<b>OBJETIVO GENERAL</b>	<b>32</b>
	<b>OBJETIVOS PARTICULARES</b>	<b>32</b>
	<b>MATERIALES Y MÉTODO</b>	<b>33</b>
I.	Análisis <i>in silico</i> de la familia AGO en <i>M. polymorpha</i> e identificación de <i>MpAGO9</i>	33
	Identificación de la familia ARGONAUTA en <i>M. polymorpha</i> .....	33
	Identificación del gen ortólogo a ARGONAUTA9 en <i>M. polymorpha</i> .....	33
	Reconstrucción filogenética .....	34
II.	Análisis <i>in silico</i> de la familia AGO en <i>M. polymorpha</i> e identificación de <i>MpAGO9</i>	35
III.	Monitoreo de la expresión de <i>MpAGO9</i> durante el ciclo de vida de <i>M. polymorpha</i> .	35
	Diseño de oligonucleótidos para el extremo 3'UTR .....	35
	Siembra, mantenimiento y colecta de tejidos vegetales. ....	35
	Extracción de RNA total . ....	36
	Síntesis de cDNA y Reacción en Cadena de la Polimerasa (PCR) .....	36
III.	Ensayo de Complementación por Fluorescencia Bimolecular ( <i>BiFC</i> )	36
	Diseño de oligonucleótidos para amplificar los interactores .....	37
	Obtención de DNA genómico. ....	37
	Obtención de los fragmentos de PCR-attB. ....	37
	Reacción BP: Clonación en el vector de entrada pDONR221 .....	38
	Transformación de <i>Escherichia coli</i> por electroporación. ....	39
	Escrutinio de clonas .....	39

Clonación en los vectores de destino pSPYNE y pSPYCE. ....	39
Linearización del inserto-attL por PCR. ....	40
Reacción de recombinación LR ....	40
Transformación de <i>E. coli</i> por electroporación. ....	41
Escrutinio de clonas transformantes ....	41
Transformación por electroporación de <i>Agrobacterium</i> ....	41
Agroinfiltración en <i>Nicotiana benthamiana</i> ....	42
Preparación de protoplastos de <i>N. benthamiana</i> ....	43
Confirmación de la expresión de los transgenes y acumulación de proteína. ....	43
Búsqueda e identificación de posibles interactores. ....	44
Caracterización del gen <i>MpNRPE1</i> ....	45
Identificación de dominios WG/GW. ....	45
<b>RESULTADOS</b>	<b>46</b>
I. Conservación evolutiva de la secuencia: Identificación del gen homólogo a <i>AGO9</i> en <i>M. polymorpha</i> . ....	46
II. Conservación de la estructura proteica: Distribución de dominios y estructura tridimensional de la proteína MpAGO9. ....	48
III. Determinación del perfil de expresión: Dominios de expresión de <i>MpAGO9</i> durante el ciclo de vida de <i>M. polymorpha</i> . ....	50
IV. Interacción proteína-proteína: Ensayo de Complementación por Fluorescencia Bimolecular ( <i>BiFC</i> ) ....	51
<b>DISCUSIÓN</b>	<b>57</b>
<b>CONCLUSIONES Y PERSPECTIVAS</b>	<b>64</b>
<b>LITERATURA CITADA</b>	<b>66</b>
<b>APÉNDICES</b>	<b>80</b>
<b>ANEXOS</b>	<b>85</b>



## RESUMEN



El silenciamiento génico mediado por RNAs es un proceso conservado en prácticamente todos los eucariontes. En las plantas existen mecanismos de silenciamiento génico mediados por RNAs altamente diversificados. Todas las vías de silenciamiento mediadas por RNAs necesitan de la participación de proteínas de la familia ARGONAUTA (AGO) para ser efectuadas. En *Arabidopsis thaliana* existen 10 miembros de la familia AGO, siendo el clado III (clado AGO4/6/8/9) el efector de la vía de Metilación del DNA dependiente de DNA (*RdDM* por sus siglas en inglés), una vía especializada en el silenciamiento a nivel transcripcional de transposones y elementos repetidos en el genoma. AGO9 desempeña un papel importante en la especificación de la línea germinal en las angiospermas y su actividad representa un punto de conexión entre procesos de regulación epigenética y programas de desarrollo. Debido a lo anterior, consideramos que es muy interesante determinar, desde una perspectiva evolutiva, su nivel de conservación en las embriofitas y estudiar su composición molecular en *Marchantia polymorpha*, una especie representante del clado basal del linaje de las embriofitas. En el presente estudio, empleamos herramientas bioinformáticas que nos permitieron identificar al gen ortólogo de AGO9 de *A. thaliana* en el genoma de *M. polymorpha*, *MpAGO9*. Asimismo, se determinó su estructura génica, se modeló la estructura tridimensional de la proteína y se determinó la arquitectura de sus dominios. Nuestros resultados muestran que existe un alto grado de conservación en residuos esenciales determinantes para la adecuada función de la proteína *MpAGO9*. Determinamos que *MpAGO9* se expresa a niveles comparables durante el desarrollo gametofítico de la planta, al igual que su ortólogo en *A. thaliana*. Por otra parte, mediante ensayos tipo *BiFC* se probó si *MpAGO9* es capaz de establecer interacciones *in vivo* con una proteína interactora conocida (NRPE1). Los resultados no fueron los esperados debido a propiedades intrínsecas de las proteínas evaluadas. Sin embargo, lo anterior no descarta el hecho de que estas proteínas interactúen como la evidencia bioinformática indica. Finalmente, este estudio muestra que AGO9 es un gen que se encuentra altamente conservado en todos los grupos de las plantas terrestres, lo que sugiere que éste se originó en un momento previo a su divergencia. De la misma manera, es posible que la función de este gen se encuentre conservada y debido a su papel en la regulación epigenética durante el desarrollo -como se ha observado en traqueofitas-, es probable que haya sido fundamental en la evolución de las plantas terrestres.

## AGRADECIMIENTOS



Este trabajo fue concebido y realizado en el Laboratorio de Epigenética y Biología del Desarrollo del Instituto de Biotecnología y Ecología Aplicada (epilab), UV bajo la tutoría del Dr. Mario Alberto Arteaga Vázquez. La segunda fase experimental de este trabajo se realizó en el Instituto de Biotecnología, UNAM en el Laboratorio de la Dra. Patricia León Mejía.

Agradezco infinitamente al Dr. Mario Arteaga Vázquez por haberme dado la oportunidad de trabajar en su laboratorio y por haberme ayudado a encontrar mi verdadera pasión. Le agradezco todo el apoyo que me brindó desde que llegué por primera vez al laboratorio, con una gran curiosidad pero completamente inexperta.

A la Dra. Patricia León Mejía por su apoyo durante mi estancia en el IBt y por la oportunidad de continuar aprendiendo Biología Molecular de Plantas.

Agradezco profundamente al M. en C. Luis de Luna Valdez y al Biol. Carlos Villaseñor por su apoyo durante la realización del ensayo de interacción BiFC. Asimismo, agradezco a los miembros del jurado por sus valiosas aportaciones. Agradezco también a todos los miembros del epilab y compañeros del IBt que de alguna manera estuvieron involucrados en la realización de este trabajo.

A Ared Nayeli Mendoza Mejia por nuestras largas discusiones, valiosas críticas, apoyo en situaciones complicadas y por su amistad.

Y finalmente, quiero agradecer a mis padres, José Lara Tenorio y Ma. Elena Mondragón Olascoaga por ser ejemplo de fortaleza y perseverancia. Les agradezco por todo el cariño y apoyo que me han brindado a lo largo de mi vida, sin ustedes, nada de esto habría sido posible.

# INTRODUCCIÓN



## 1. Antecedentes históricos del silenciamiento mediado por RNAs pequeños

El término interferencia por RNA (RNAi por el inglés *RNA interference*) comprende aquellos fenómenos de regulación génica en donde intervienen RNAs pequeños que generalmente conllevan al silenciamiento, tanto a nivel transcripcional como a nivel post-transcripcional (Castel & Martienssen, 2013; Matzke & Moshier, 2014). El fenómeno de RNAi fue inicialmente descrito en plantas (Napoli *et al.*, 1990; van der Krol *et al.*, 1990). Sin embargo, algunos años después se encontró que está altamente conservado en prácticamente todos los eucariontes (Borges & Martienssen, 2015; Kim *et al.*, 2009).

El primer reporte de un fenómeno similar al RNAi fue el de Napoli y Jorgensen en 1990. En este estudio se pretendía conocer si la Chalcona Sintasa (CHS) era la enzima limitante en la biosíntesis de antocianinas, los pigmentos que dan la coloración violeta a los pétalos de las petunias. El estudio consistió en la introducción de una copia adicional del gen de la *CHS* en petunias, los resultados fueron completamente inesperados ya que además de flores sin pigmentación, observaron flores con patrones diversos de coloración en los pétalos. Posteriormente, al analizar los niveles del mRNA de la *CHS* en las plantas transgénicas, observaron que había una disminución notable en los niveles de dicho mRNA en comparación con las plantas silvestres. Con base en estos datos, propusieron que la introducción de una copia adicional del gen *CHS* suprimía no sólo la expresión del gen endógeno sino también la del gen introducido. Es decir, se co-suprimía la expresión del gen de ambas copias de *CHS* (Napoli *et al.*, 1990).

En el año de 1992, Romano y Macino reportaron un fenómeno similar en el hongo filamentoso *Neurospora crassa*<sup>1</sup>, en donde la introducción del RNA homólogo al gen *albino-1*, que codifica una enzima clave para la biosíntesis de carotenoides, provocaba el silenciamiento de éste generando un fenotipo albino (Romano & Macino, 1992). Para el año de 1995, este fenómeno se había reportado en el nemátodo

<sup>1</sup> Este fenómeno se nombró *Quelling*, cuya traducción al español es “detener (algo) utilizando la fuerza”.

*Caenorhabditis elegans*. En este estudio observaron que la introducción de RNA sentido o antisentido al mensajero del gen *par-1*, provocaba su degradación (Guo & Kemphues, 1995). Este estudio fue clave para determinar la naturaleza del RNA que producía el efecto de silenciamiento pues mostraron que tanto el RNA sentido como el RNA antisentido al mRNA del gen *par-1* tenían el mismo efecto de silenciamiento (Guo & Kemphues, 1995). Con base en lo anterior, en el año de 1998 Fire y Mello elucidaron el papel del dsRNA en el silenciamiento y lograron relacionar el fenómeno que observaron con lo que antes se había descrito en plantas (Co-supresión) y en hongos (*Quelling*) (Fire *et al.*, 1998).

dsRNA: RNA de doble cadena.

DICER: Enzima ribonucleasa que corta precursores de dsRNA en RNAs pequeños de 20-30 nt.

ARGONAUTA: Familia de proteínas efectoras del silenciamiento mediado por RNAs pequeños que se unen a RNAs pequeños a través de su dominio PAZ y MID, en algunos casos presentan actividad de corte mediante su dominio PIWI.

La hipótesis inicial proponía que el dsRNA se separaba y la hebra antisentido formaba un dúplex con el mRNA con lo que se producía su degradación, sin embargo, nunca se logró comprobar la existencia de moléculas de RNA de un tamaño similar al mRNA blanco. Bajo la premisa anterior, en el año de 1999 Hamilton y Baulcombe se dieron a la tarea de identificar formas más pequeñas del RNA. En los mecanismos de silenciamiento que analizaron (co-supresión y silenciamiento inducido por virus) encontraron que estaban involucrados RNAs de alrededor de 25 nt y que el dsRNA era en realidad un precursor de éstos (Hamilton & Baulcombe, 1999).

En años subsecuentes, diferentes grupos de investigación lograron caracterizar la maquinaria capaz de ejercer el silenciamiento mediado por RNAs pequeños en diferentes modelos. El grupo de Hannon y colaboradores logró, mediante aproximaciones bioquímicas, identificar dos fases en el proceso del RNAi en *Drosophila*: 1) la fase de iniciación, en donde el dsRNA se procesaba en el RNA de ~25 nt y 2) la fase efectora, en donde se llevaba a cabo el corte del mRNA. Asimismo, lograron identificar a la enzima responsable de la fase de iniciación, una RNasa tipo III a la que se nombró DICER (Bernstein *et al.*, 2001). Durante el transcurso de ese mismo año se logró identificar a un miembro de la familia ARGONAUTA como enzima responsable de efectuar el silenciamiento (Hammond *et al.*, 2001) y se determinó que era el componente esencial del Complejo de Silenciamiento Inducido por RNA (RISC por sus siglas en inglés *RNA Induced Silencing Complex*) debido a que establecía una interacción directa con el RNA pequeño (Martinez *et al.*, 2002). Para este momento, miembros de la familia ARGONAUTA ya se habían

identificado en diferentes organismos, incluyendo *Arabidopsis thaliana* (Bohmert *et al.*, 1998), *N. crassa* (Cogoni & Macino, 1997) y *C. elegans* (Tabara *et al.*, 1999).

El mecanismo de silenciamiento mediado por RNA a nivel Post-transcripcional (PTGS por sus siglas en inglés *Post Transcriptional Gene Silencing*) comenzaba a ser ampliamente comprendido, muchos de sus componentes y su papel en el proceso ya habían sido caracterizados. Sin embargo, un número creciente de reportes indicaban que el RNAi no sólo actuaba a nivel post-transcripcional sino que podía tener un papel importante en la regulación génica a nivel transcripcional (TGS por sus siglas en inglés *Transcriptional Gene Silencing*). El TGS involucra modificaciones del DNA como metilación de las citosinas y modificaciones de la cromatina (Wassenegger, 2005; Castel & Martienssen, 2013). El primer reporte de un fenómeno de represión génica a nivel transcripcional mediado por RNA se publicó en 1994, mucho tiempo antes de que se elucidara el mecanismo de RNAi (Wassenegger *et al.*, 1994). En este estudio observaron que al generar plantas transgénicas de tabaco con cDNA del Viroide del Tubérculo Fusiforme de la Papa (PSTVd del inglés *Potato Spindle Tuber Viroid*), los genes insertados mostraban metilación *de novo* y debido a que la metilación se detectaba una vez que se llevaba a cabo la replicación RNA-RNA del cDNA del PSTVd y además, ésta era específica de la región correspondiente al transgén, se concluyó que la molécula responsable de este fenómeno de silenciamiento era su propio RNA mensajero (Wassenegger *et al.*, 1994). Algunos años después, el papel de los RNAs pequeños y el RNAi en la Metilación del DNA (se nombró como *RdDM* por sus siglas en inglés *RNA-directed DNA Methylation*) se demostró en estudios con *A. thaliana* (Mette *et al.*, 2000; Onodera *et al.*, 2005). Un proceso de silenciamiento similar a la *RdDM* se identificó en *Schizosaccharomyces pombe* (Verde *et al.*, 2004) y en la línea germinal de metazoarios (Brennecke *et al.*, 2007) por lo que se propuso que el TGS también se encuentra conservado evolutivamente a pesar de que presenta algunas variaciones (Castel & Martienssen, 2013).

Metilación del DNA:  
Adición covalente de un grupo metilo (CH<sub>3</sub>) en el carbono 5 de la Citosina. La metilación conlleva frecuentemente a la represión transcripcional.

Cromatina: Es un complejo de DNA y proteínas (histonas) que forma a los cromosomas dentro del núcleo de las células eucariontes.

cDNA: DNA complementario.

Viroide: Son moléculas circulares de RNA no codificante de cadena sencilla, carecen de cápsida y se replican de manera autónoma. Son altamente patógenicas para plantas.

## 2. Silenciamiento mediado por RNAs pequeños en plantas

En los eucariontes se han descrito diferentes mecanismos de silenciamiento mediado por RNAs pequeños, sin embargo, en las plantas se observan vías

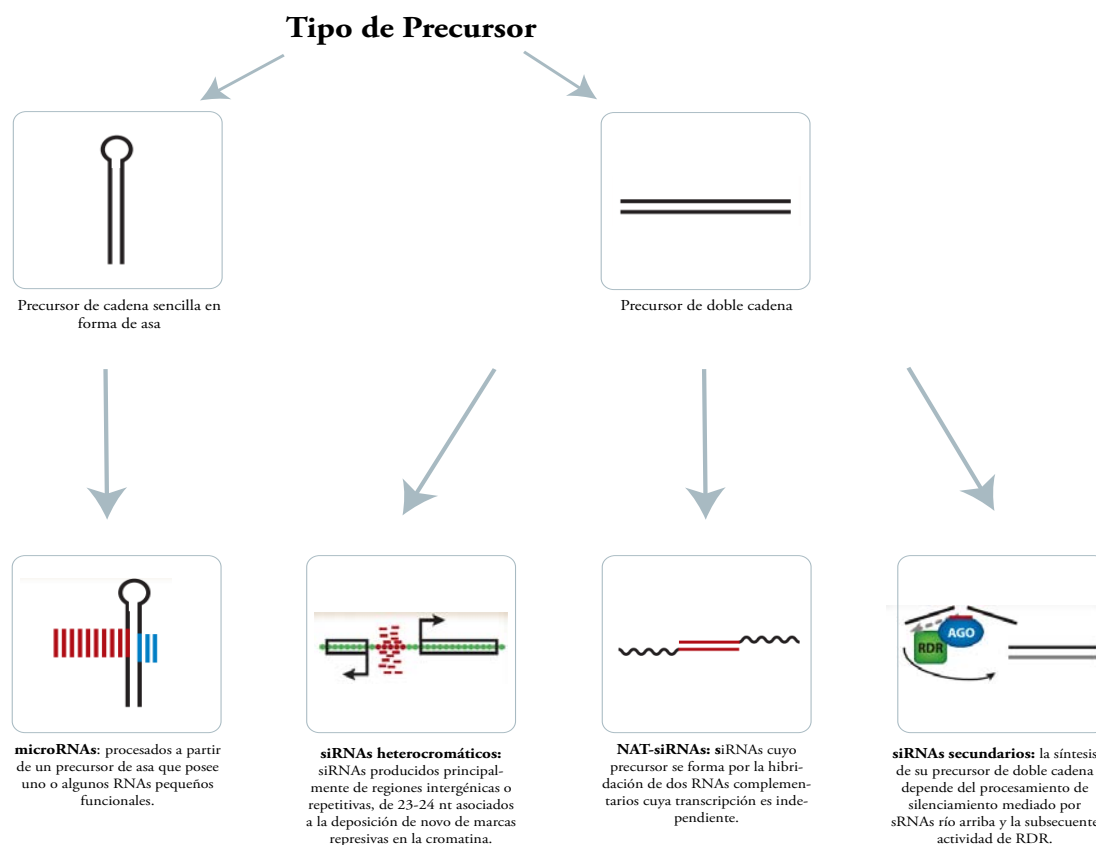
sumamente diversificadas, mientras que en otros organismos de manera natural, como es el caso de *Saccharomyces cerevisiae*, no se observa este tipo de regulación (Bonnet *et al.*, 2006; Großhans & Filipowicz, 2008).

### 2.1 Clasificación de los RNAs pequeños regulatorios en plantas

Los RNAs pequeños en las plantas varían en longitud desde los 20 hasta los 30 nucleótidos (nt). Existen numerosas clasificaciones según su función y/o su biogénesis (Mallory & Vaucheret, 2006; Bonnet *et al.*, 2006; Axtell, 2013). En las plantas, la clasificación más aceptada actualmente consta de dos grupos principales: 1) los microRNAs, cuyos precursores son moléculas de RNA de hebra sencilla que muestran un alto grado de auto-complementariedad intramolecular, formando estructuras tallo y asa y 2) los siRNAs, cuyos precursores son RNA de doble cadena con complementariedad perfecta o al menos mayor a los miRNAs (Aalto & Pasquinelli, 2012; Axtell, 2013). A su vez, los siRNAs pueden subdividirse en tres grupos: 1) los siRNA heterocromáticos (hc-siRNAs), producidos principalmente en regiones intergénicas o repetidas y que están asociados a la formación de heterocromatina, 2) los siRNAs de transcrito natural antisentido (nat-siRNAs por *natural antisense transcript small interfering RNA*), cuyo precursor se forma a partir de la hibridación de dos transcritos complementarios sintetizados de manera independiente y 3) los siRNAs secundarios (también conocidos como *phased-siRNAs*, phasiRNAs), que son siRNAs que se producen a partir de un precursor de doble cadena generado por otro proceso de silenciamiento mediado por sRNAs río arriba y por RNA polimerasas dependientes de RNA (RDR) (Axtell, 2013). Esta última categoría se encuentra ejemplificada en principal medida por los *trans-acting* siRNAs o ta-siRNAs, éstos se generan a partir de transcritos no codificantes de los genes *TAS* gracias a la acción de microRNAs, RDR6 y DCL4 y tienen la capacidad de actuar en *trans* para silenciar a nivel post-transcripcional otros blancos (Fei *et al.*, 2013; Yoshikawa *et al.*, 2005) (Figura 1).

### 2.2 La maquinaria central del silenciamiento mediado por RNAs pequeños

Todos los procesos de silenciamiento mediados por RNAs pequeños, independientemente del tipo de RNA pequeño involucrado, requieren de la participación de enzimas clave para poder llevarse a cabo. El proceso de silenciamiento de manera general se divide en dos etapas: 1) la generación de los RNAs pequeños a partir de un precursor más grande y 2) el silenciamiento dirigido por complementariedad de bases a un blanco específico.



**Figura 1. Clasificación de los RNAs pequeños según su biogénesis y función.** Se muestra un diagrama que clasifica a los RNAs pequeños según su tipo de precursor y la forma en que llevan a cabo el silenciamiento. Basado en Axtell (2013).

La enzima responsable de la primera etapa es una endorribonucleasa de tipo III que corta específicamente RNA de doble cadena, en las plantas esta familia de enzimas se conoce como DICER LIKE (DCL) (Doyle *et al.*, 2012). El primer miembro de la familia DCL fue caracterizado en *A. thaliana* (Jacobsen *et al.*, 1999; Shauer *et al.*, 2002). Actualmente se sabe que existen 4 miembros de esta familia en *A. thaliana*: DCL1, involucrada en la producción de microRNAs de 21 nt (Park *et al.*, 2002; Kurihara & Watanabe, 2004); DCL2, que genera siRNAs a partir de RNAs de doble cadena derivados de RNA viral y transgenes (Mlotshwa *et al.*, 2008); DCL3, que produce siRNAs de 24 nt usando precursores de doble cadena generados a partir de la transcripción de regiones repetidas, transposones o derivados de transgenes por acción de la RNA polimerasa dependiente de DNA IV (Pol IV) (Wei *et al.*, 2014) y DCL4, que produce *trans acting*-siRNAs (un tipo específico de siRNAs secundarios) de 21 nt a partir de los productos de degradación mediada por miRNAs de los transcritos de los genes *TAS* (Xie *et al.*, 2004) (Molnar *et al.*, 2011; Doyle *et al.*, 2012).

Por otra parte, la segunda fase consiste en el ensamblaje del RISC, cuyos componentes mínimos necesarios son: el RNA pequeño que fungirá como guía y una proteína de la familia ARGONAUTA, que efectuará el silenciamiento (Kuhn & Joshua-Tor, 2013). La proteína ARGONAUTA, guiada por el RNA

pequeño, efectuará el silenciamiento a un blanco que muestre complementariedad de bases. Si se trata de PTGS, el complejo RISC atenúa la traducción de un mRNA blanco o lo degrada mediante su actividad de corte o *slicer*, mientras que si se trata de TGS, inducirá la deposición de marcas epigenéticas represivas al asociarse con modificadores de la cromatina (Vaucheret, 2008). A continuación se hablará más en detalle sobre esta familia de proteínas.

### 3. Las proteínas efectoras del silenciamiento: ARGONAUTAS

Pleiotrópico: Del griego *πλείων* (pleio)= más numeroso o abundante y *τρόπος* (tropos) = cambio o giro. Se refiere a un gen que ejerce múltiples características en el fenotipo

Las proteínas ARGONAUTAS (AGO) conforman una familia con alto grado de conservación evolutiva en todos los dominios de la vida, se expresan en un amplio rango de tejidos y en altos niveles durante diferentes etapas de la vida de los organismos (Mallory & Vaucheret, 2010). El miembro fundador de esta familia de proteínas se descubrió en el año de 1998 mediante escrutinios de mutantes de genes involucrados en el desarrollo en *A. thaliana*: los mutantes nulos de *AGO1* presentaban fenotipos pleiotrópicos, incluyendo hojas en forma tubular que asemejaban los tentáculos de un molusco cefalópodo del género *Argonauta* (Bohmert *et al.*, 1998). El número de genes que codifican a proteínas AGO varía enormemente, incluso entre especies cercanas (Höck & Meister, 2008; Mallory & Vaucheret, 2010). En la Tabla 1 se enlista el número de genes que codifican a proteínas AGO en diferentes especies.

Tabla 1. Número de genes que codifican a proteínas ARGONAUTA en diferentes especies

Especie	Número de genes	Referencia
<i>Saccharomyces cerevisiae</i>	0	[1,2]
<i>Schizosaccharomyces pombe</i>	1	[1,2]
<i>Caenorhabditis elegans</i>	27	[1,2]
<i>Drosophila melanogaster</i>	5	[1,2]
<i>Homo sapiens</i>	8	[1,2]
<i>Rattus norvegicus</i>	8	[1,2]
<i>Mus musculus</i>	8	[1,2]
<i>Arabidopsis thaliana</i>	10	[1,2]
<i>Zea mays</i>	18	[3]
<i>Oryza sativa</i>	19	[4]
<i>Physcomitrella patens</i>	6	[5]

[1]: Höck & Meister, 2008; [2]: Mallory & Vaucheret, 2010; [3]: Qian *et al.*, 2013; [4]: Kapoor *et al.*, 2008; [5]: Arif *et al.*, 2013.



### 3.1 Características estructurales y funcionales de las proteínas AGO

#### 3.1.1 Características bioquímicas de las proteínas AGO

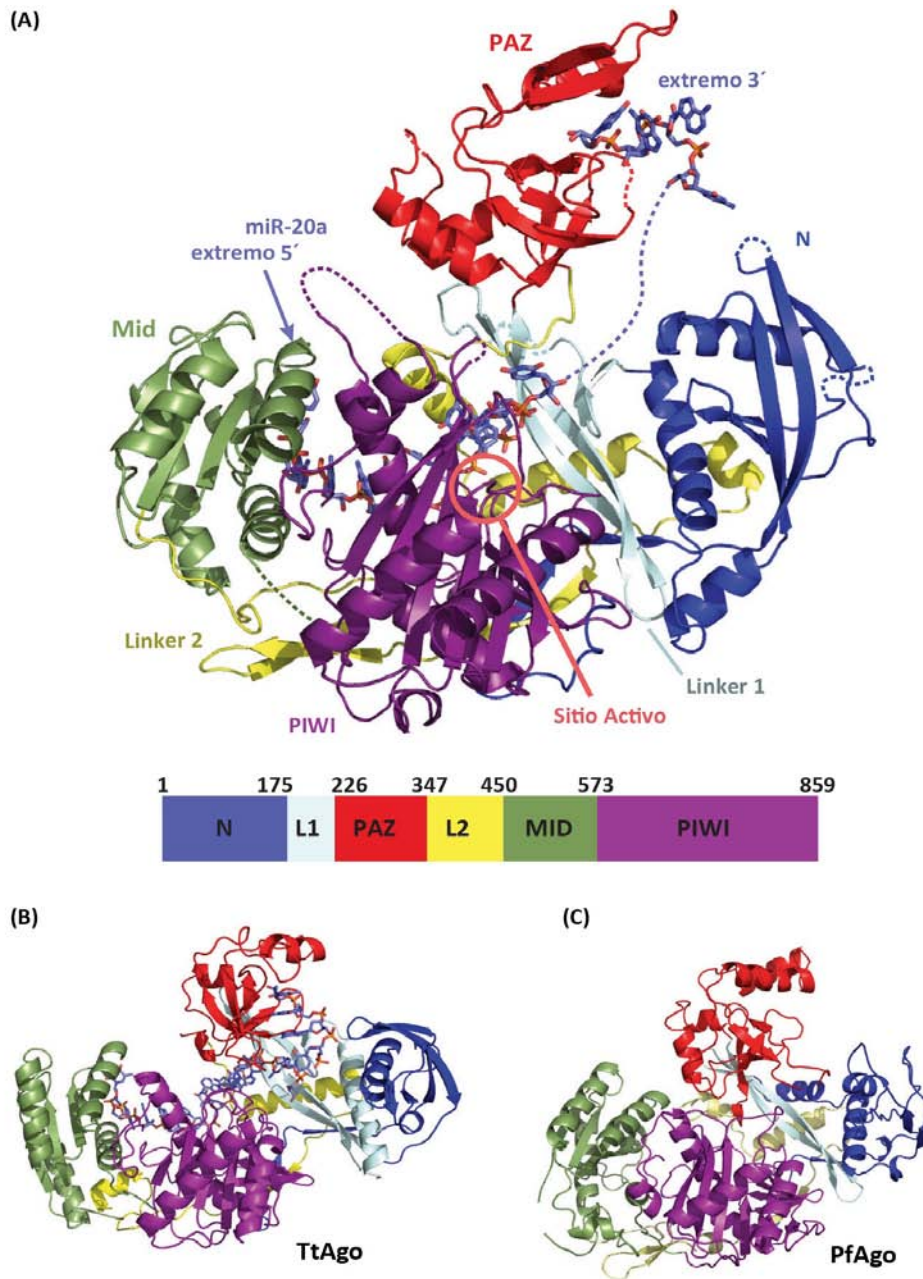
Las proteínas AGO poseen propiedades moleculares específicas que les permiten llevar a cabo sus funciones durante el silenciamiento mediado por RNAs pequeños: 1) Tienen la capacidad de establecer uniones fuertes con RNA de cadena sencilla para ensamblar el RISC, 2) Poseen un dominio de endonucleasa tipo RNasa H que las provee de actividad de corte (*slicer*); este dominio está involucrado en al menos cuatro diferentes procesos: a) efectúa el corte en el mRNA al que el RNA pequeño muestra complementariedad durante el PTGS; b) pueden amplificar la respuesta de silenciamiento por la producción de fragmentos de RNA que sirven como sustrato para RNA polimerasas dependientes de RNA convirtiéndose en fuentes de siRNAs secundarios; c) efectúan cortes para lograr la separación de la hebra complementaria y la hebra guía en el dúplex de RNA pequeño y d) actúan como plataforma de anclaje para otros cofactores durante el silenciamiento (Poulsen *et al.*, 2013; Meister, 2013).

#### 3.1.2 Características estructurales de las proteínas AGO

Las propiedades funcionales de las proteínas AGO residen en cuatro dominios: el dominio N-terminal, el dominio PIWI-AGO-ZWILLE (PAZ), el dominio MID y el dominio C-terminal o PIWI. Las proteínas AGO adoptan una estructura bilobular con una "hendidura" central delimitada por residuos de aminoácidos cargados positivamente que permiten la unión de los RNAs pequeños cargados negativamente. Los dominios MID y PIWI conforman un lóbulo, mientras que el otro lóbulo está conformado por los dominios N-terminal, L1 o DUF1785 y PAZ. Los dominios PAZ y MID están interconectados por una larga sección conocida como dominio L2 (Figura 2) (Vaucheret, 2008; Höck & Meister, 2008; Kuhn & Joshua-Tor, 2013; Poulsen *et al.*, 2013).

##### 3.1.2.1 El dominio N-terminal

El dominio N-terminal está involucrado en la disociación de la hebra pasajera de la hebra guía y se ha observado que podría tener un papel importante en la escisión del mRNA durante el PTGS (Swarts *et al.*, 2014). Después de que el dúplex de RNA se ensambla, la hebra guía permanece asociada a la proteína AGO, mientras que la hebra complementaria o pasajera se disocia. La proteína AGO reconoce a la hebra con el extremo 5' menos estable. La remoción de la hebra pasajera puede ser dependiente de corte o no. En ambos casos, el dominio N-terminal funciona como una especie de cuña, interrumpiendo los



**Figura 2. Estructura bilobular de las proteínas ARGONAUTA.** A) En la parte superior: Se muestra una representación de la proteína AGO2 de humano unida a miR20a. El dominio N-terminal se muestra en color azul, el L1 en azul claro, el PAZ en color rojo, el L2 en color amarillo, el MID en color verde y el PIWI en púrpura. miR20a se muestra en representación de varillas. El extremo 5' está unido al dominio MID, los nucleótidos 2-10 se encuentran a lo largo del surco de unión a RNA y los 17-20 se unen al dominio PAZ. En la parte inferior: Se muestra un diagrama de barra mostrando la distribución de los dominios en hAGO2. En B) se muestra la representación de la estructura de la proteína AGO de *Thermus thermophilus* (TtAGO) unido a hebra guía de DNA y en C) la estructura de la proteína AGO de *Pyrococcus furiosus* (PfAGO). Los dominios y ácidos nucleicos en B) y C) muestran los mismos colores que en A). Tomado y modificado de Kuhn & Joshua-Tor, 2013.

apareamientos entre el dúplex de RNA en su extremo 3' (durante el ensamble del RISC) o interrumpiendo el apareamiento entre el RNA pequeño guía y su blanco (durante el PTGS) (Kwak & Tomari, 2012; Faehnle *et al.*, 2013;).

### 3.1.2.2 El dominio MID

El dominio MID adquiere una forma de "bolsillo" en donde varios aminoácidos conservados establecen el sitio de interacción con el grupo fosfato del extremo 5' del RNA pequeño guía (Swarts *et al.*, 2014). Debido a lo anterior, el dominio MID ha sido implicado en la clasificación o *sorting* de los RNAs pequeños, dando una posible explicación a la preferencia de las proteínas AGO por determinado tipo de RNAs pequeños (Wei *et al.*, 2012). Asimismo, se ha observado que este dominio mantiene al RNA guía en una conformación helicoidal dentro de su *región semilla* (comprendida entre los nucleótidos 2 al 7 ó 2 al 8 a partir del extremo 5' del microRNA maduro), lo que facilita la unión con su blanco (Frank *et al.*, 2012).

Región semilla:  
Es una región heptamérica conservada que está generalmente situada en las posiciones 2-7 del extremo 5' en los miRNAs. El apareamiento entre la región semilla y el blanco es perfecto.

### 3.1.2.3 El dominio PAZ

El dominio PAZ, además de estar presente en las proteínas AGO, se encuentra también en las proteínas de la familia DICER. El dominio PAZ forma un surco central en donde se distribuyen residuos aromáticos altamente conservados que pueden unirse, de manera específica, al extremo 3' de RNAs pequeños de cadena sencilla (Ma *et al.*, 2004; Lingel *et al.*, 2004). En este sitio solamente pueden unirse RNAs pequeños que presenten dos nucleótidos sobresalientes en el extremo 3', por lo que actúa como una especie de filtro selectivo para el RISC, en donde sólo se permite la unión de RNAs pequeños generados por el procesamiento de endonucleasas tipo DCL, excluyendo a otros RNAs presentes en la célula (Hur *et al.*, 2013). Asimismo, se ha observado que esta interacción protege al RNA guía de la degradación (Poulsen *et al.*, 2013; Swarts *et al.*, 2014).

### 3.1.2.4 El dominio PIWI

El dominio PIWI es uno de los dominios fundamentales de las proteínas AGO, pues en este reside la actividad catalítica. El dominio PIWI adopta una

estructura que se asemeja a la proteína que contiene el dominio RNasa H en *Bacillus holodurans* (Song *et al.*, 2004). La estructura terciaria del dominio es bastante rígida, requiriendo iones metálicos divalentes para mantener estable su estructura y llevar a cabo su actividad catalítica. El dominio posee una tríada de residuos de aminoácidos muy importantes para llevar a cabo la actividad catalítica y se encuentran altamente conservados en todas las proteínas AGO: dos residuos de Aspartato (D) y un tercer residuo que puede ser otro Aspartato (D), Histidina (H) o en algunos casos Lisina (K) (Song *et al.*, 2004; Vaucheret, 2008). A pesar del alto grado de conservación de esta tríada catalítica, no todas las proteínas AGO presentan la actividad de corte, debido a lo anterior, se especula que existen otros factores que pueden contribuir a esta actividad (Vaucheret, 2008; Mallory & Vaucheret, 2010; Wei *et al.*, 2012).

Como una posible explicación al fenómeno antes mencionado, los estudios de Patel y colaboradores (2012) en la levadura *Kluyveromyces* y los del grupo de Wang y colaboradores (2013) en *Thermus thermophilus*, proponen la existencia de una tétrada catalítica en lugar de una tríada. Sus estudios demuestran la existencia de un cuarto residuo de Glutamato (E) necesario para la catálisis que es reclutado cuando se efectuará actividad de corte (Nakanishi *et al.*, 2012; Sheng *et al.*, 2013). En el modelo que proponen, se habla de una configuración compatible con el corte que se da cuando el residuo de Glutamato, que reside en una región móvil del dominio PIWI (dedo de glutamato), se inserta entre los residuos DDH para completar una tétrada catalítica, formando el motivo DEDH (Nakanishi *et al.*, 2012; Sheng *et al.*, 2013; Meister, 2013; Swarts *et al.*, 2013; Poulsen *et al.*, 2013).

El dominio PIWI también se ha relacionado con el establecimiento de la interacción entre las proteínas AGO y un grupo peculiar de proteínas, las proteínas con motivo GW (Poulsen *et al.*, 2013). La cristalización de la primera proteína AGO eucarionte reveló características estructurales únicas: la presencia de sitios de unión a residuos de Triptófano (W) en dos sitios próximos dentro del dominio PIWI, por lo que proteínas con repeticiones de W o ricas en este aminoácido eran candidatas a ser cofactores de las proteínas AGO (Schirle & McRae, 2012). El grupo de Lagrange y colaboradores (2007) logró probar la interacción entre AGO4 y la subunidad grande de la RNA polimerasa V de *A. thaliana* (NRPE1) que presenta motivos GW en el dominio CTD, asimismo, mostraron que al sustituir los motivos GW de NRPE1 con los motivos GW de la proteína GW182 de humano, éstos eran capaces de establecer interacción con AGO4 mostrando que los motivos GW, posteriormente nombrados como *AGO-hooks*, son plataformas de interacción para las proteínas AGO, que están conservados evolutivamente y además este tipo de proteínas pueden actuar como cofactores de las proteínas AGO durante vías de silenciamiento mediados por RNA (El-Shami *et al.*, 2007).

### 3.1.2.5 El dominio L2

El dominio L2 contribuye de manera muy importante en el mantenimiento de la estructura de la proteína pues establece interacciones con todos los dominios de la proteína AGO, de esta manera, estabiliza la interfase entre el dominio N y L1 y en particular, entre los dominios MID y PIWI. El dominio L2, cerca del dominio PAZ, adquiere una conformación en forma de espiral que contiene residuos de aminoácidos cargados positivamente que posiblemente están involucrados con la unión a RNAs pequeños (Poulsen *et al.*, 2013).

## 3.2 Diversidad funcional de las proteínas AGO de *A. thaliana*

En las plantas existe una gran diversidad de RNAs pequeños y sus funciones son efectuadas por diferentes proteínas AGO. En *A. thaliana* se han identificado 10 genes que codifican proteínas AGO. Según su identidad, se han clasificado en tres grupos: el clado I, conformado por *AGO1/AGO5/AGO10*, el clado II, conformado por *AGO2/AGO3/AGO7* y el clado III, conformado por *AGO4/AGO6/AGO8/AGO9* (Vaucheret, 2008, Zhang *et al.*, 2015) (Figura 3).

Los miembros del clado *AGO1/AGO5/AGO10* se asocian principalmente con RNAs pequeños de 21 a 22 nt y están involucrados con el silenciamiento a nivel post-transcripcional. El miembro mejor caracterizado de este clado es *AGO1*, que es el principal efector del silenciamiento mediado por miRNAs. *AGO1* se une principalmente a miRNAs procesados por la endonucleasa DCL1, así como a siRNAs provenientes de transgenes y virus. Se ha demostrado que *AGO1* presenta actividad de corte y que se une preferencialmente a miRNAs con Uridina en el extremo 5' (Song *et al.*, 2004; Baumberger & Baulcombe, 2005; Mi *et al.*, 2008; Poulsen *et al.*, 2013). En el caso de *AGO10*, se ha observado que tiene cierta redundancia funcional con *AGO1* y se ha relacionado con diferentes procesos del desarrollo, especialmente en el mantenimiento de células indiferenciadas en el meristemo apical y en el establecimiento de la polaridad de las hojas. El mecanismo por el cual se mantiene el estado indiferenciado consiste en que *AGO10* secuestra los miR165/166 e interfiere con su interacción con *AGO1*, con lo que se evita el silenciamiento de los factores de transcripción HOMEODOMAIN-LEUCINE ZIPPER clase III (HD-ZIP III), necesarios para mantener el estado indiferenciado en el meristemo apical (Vaucheret, 2008; Zhu *et al.*, 2011). El último miembro del clado, *AGO5*, a diferencia de *AGO1* y *AGO10*, muestra dominios de expresión muy específicos confinados a tejidos reproductivos tanto masculinos como femeninos (Mallory & Vaucheret, 2010; Borges *et al.*, 2010). *AGO5* se une preferencialmente a RNAs pequeños con Citosina en el extremo 5' y recientemente se le ha asociado con la respuesta de defensa antiviral (Mi *et al.*, 2008;

Brousseau & Moffett, 2015).

En el caso específico de AGO7, se ha observado que es intermediario en la biosíntesis de ta-siRNAs. AGO7 se asocia únicamente a miR390 y en conjunto dirigen la degradación de los transcritos precursores *TAS3* (Montgomery *et al.*, 2008). Asimismo, se ha observado que AGO7 controla la transición de la fase juvenil a la fase vegetativa adulta (Hunter *et al.*, 2003). Los otros dos miembros del clado, AGO2 y AGO3 son las únicas proteínas AGO en *A. thaliana* que no presentan el motivo DDH, sino un motivo con tres residuos de Aspartato (DDD). AGO2 y AGO3 son muy similares entre sí, lo que sugiere que surgieron recientemente por duplicación (Vaucheret, 2008). En el caso de AGO2, se ha observado que se asocia principalmente a RNAs pequeños con Adenosina en el extremo 5' (Mi *et al.*, 2008). Sobre la función de AGO2, se ha reportado que participa activamente en defensas antivirales y que posiblemente esta proteína se especializó en esta respuesta (Jaubert *et al.*, 2011).

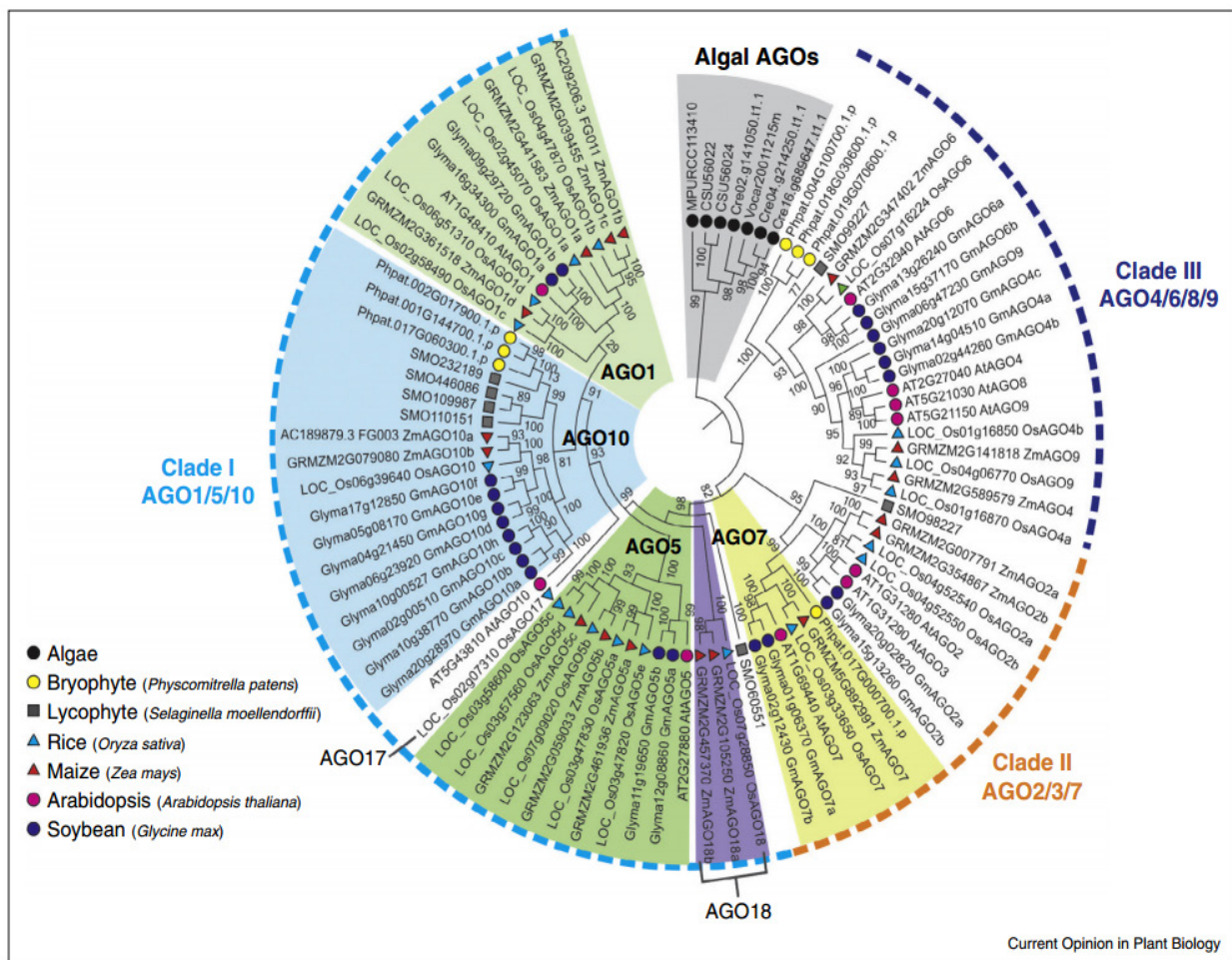


Figura 3. Reconstrucción filogenética de la familia AGO en las plantas terrestres. Se observa una reconstrucción filogenética que agrupa a la familia ARGONAUTA de diferentes especies de plantas en los tres clados principales según su identidad: El clado I o clado AGO1/5/10 se muestra de color azul claro, el clado II o clado AGO2/3/7 se muestra en color naranja y por último el clado III o clado AGO4/6/8/9 que se muestra en color azul marino. Tomada de Zhang *et al.*, 2015.

Los miembros del clado *AGO4/AGO6/AGO8/AGO9* se asocian principalmente con RNAs pequeños de 24 nt provenientes de regiones heterocromáticas y están involucrados en el silenciamiento a nivel transcripcional de transposones y regiones repetidas principalmente (Mi *et al.*, 2008; Wassenegger, 2005). *AGO4* es el miembro del clado mejor caracterizado y su participación en la vía de Metilación del DNA dependiente de RNA se demostró primero (Zilberman *et al.*, 2003; Castel & Martienssen, 2013). *AGO4* muestra dominios de expresión amplios, encontrándose desde etapas embrionarias, hasta hojas y flores, por otra parte, sus parálogos *AGO6* y *AGO9* muestran dominios de expresión restringidos: *AGO6* se encuentra únicamente en meristemas de ápice y de la raíz mientras que *AGO9* está restringida al meristemo apical de embrión y óvulos en desarrollo (Havecker *et al.*, 2010; Olmedo-Monfil *et al.*, 2010). *AGO8* es el parálogo más cercano a *AGO9*, por lo que se sugiere que se originó recientemente a causa de una duplicación. *AGO8* se expresa en bajos niveles en diferentes tejidos de *A. thaliana* y muestra cambios en el marco de lectura inducidos por *splicing*, por lo que se sugiere que es un pseudogen (Takeda *et al.*, 2008). Se ha propuesto que las diferencias en los patrones de expresión entre los miembros de este clado podrían estar relacionadas con su diversificación funcional (Havecker *et al.*, 2010).

#### 4. El clado *AGO4/AGO6/AGO9* participa en el TGS: Vía de Metilación del DNA dependiente de RNA

La vía de Metilación del DNA dependiente de RNA (*RdDM*, por las siglas en inglés *RNA-directed DNA Methylation*) es una de las vías de regulación epigenética más importantes en las plantas. Esta vía es única entre las vías de modificación de la cromatina mediada por RNAs pequeños en los eucariontes debido a que requiere de una maquinaria transcripcional especializada centrada en dos enzimas específicas de plantas: las RNA polimerasas dependientes de DNA IV y V (Matzke & Mosher, 2014; Onodera *et al.*, 2005). La *RdDM* es una vía epigenética de suma importancia en el mantenimiento de la integridad del genoma en las plantas y representa una importante extensión de sus capacidades transcripcionales (Matzke & Mosher, 2014).

Región heterocromática (heterocromatina): Cromatina altamente condensada y transcripcionalmente inactiva. Juega un importante papel en el mantenimiento de la estructura de los cromosomas, principalmente encontrada en zonas teloméricas y centroméricas

Transposón: Fragmento móvil de DNA que puede cambiar de posición en el genoma.

Parálogo (gen parálogo): Genes que muestran homología y divergieron después de un evento de duplicación. Los genes parálogos podrían o no mantener la misma función.

Pseudogen: Es una copia de un gen que generalmente carece de intrones y otras secuencias de DNA esenciales necesarias para su función. Los pseudogenes, a pesar de mostrar similitud genética con los genes funcionales no se expresan (o lo hacen en muy bajos niveles) y contienen numerosas mutaciones.

## 4.1 La vía de Metilación del DNA dependiente de RNA

La vía *RdDM* consta de dos etapas fundamentales: 1) la generación de RNAs pequeños (específicamente siRNAs de ~24 nt) dependiente de la transcripción de la RNA polimerasa IV (Pol IV) y 2) la metilación *de novo*, *locus* específica, dependiente de la transcripción de la RNA polimerasa V (Pol V). La vía canónica de la *RdDM* se describió inicialmente con AGO4. Sin embargo, se ha demostrado que sus parálogos AGO6 y AGO9 también participan en esta vía bajo ciertas condiciones. A continuación se explicará la vía *RdDM* en donde participa AGO4 y posteriormente se hará una recopilación de la información relativa a la participación de AGO6 y AGO9 en esta vía.

### 4.1.1 La vía *RdDM* canónica

#### 4.1.1.1 Biogénesis de siRNAs dependiente de RNA polimerasa IV

La biogénesis de los siRNAs inicia cuando la RNA polimerasa IV (Pol IV) se dirige a un *locus* y genera un transcrito de cadena sencilla de entre 30 a 40 nt (Blevins *et al.*, 2015; Zhai *et al.*, 2015) que servirá como molde a la RNA polimerasa dependiente de RNA 2 (RDR2) para generar dsRNA. Se ha reportado que los *loci* preferenciales de unión de la Pol IV corresponden en gran parte a transposones y numerosas regiones repetidas en donde se observan marcas epigenéticas de represión preexistentes (Mosher *et al.*, 2008). Asimismo, se ha observado que SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) funciona como un auxiliar en el reclutamiento de la Pol IV a sus *loci* (Law *et al.*, 2011).

Posteriormente, el dsRNA generado por RDR2 es procesado generando un fragmento de 24 nt por acción de la endonucleasa DICER-LIKE 3 (DCL3) (Blevins *et al.*, 2015; Zhai *et al.*, 2015). Los fragmentos de 24 nt son subsecuentemente metilados en su extremo 3' por acción de la RNA metiltransferasa HUA ENHANCER 1 (HEN1) para evitar su degradación. Finalmente, en el citoplasma, el dsRNA se incorpora a ARGONAUTA 4 (AGO4) con ayuda de HSP90 y posteriormente mediante la actividad de corte de AGO4, se escinde la hebra pasajera de la hebra guía, ensamblando así el complejo de silenciamiento (Ye *et al.*, 2012; Matzke & Mosher, 2014).

#### 4.1.1.2 Metilación *de novo* dependiente de la RNA polimerasa V

La siguiente etapa en la vía *RdDM* consiste en el establecimiento de la metilación *de novo* en *loci* específicos, proceso dependiente de la transcripción de la RNA polimerasa V (Pol V). La Pol V, con ayuda



de remodeladores de la cromatina como DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), genera un transcrito que fungirá como sustrato para guiar la unión entre el complejo AGO4/siRNA. De manera similar a Pol IV, Pol V muestra preferencia por *loci* que corresponden a transposones y regiones repetidas a lo largo del genoma, sin embargo, también se han reportado *loci* únicos correspondientes a regiones intergénicas y a transposones dentro de genes que codifican a proteínas o en promotores. Tanto la Pol IV como la Pol V muestran poca afinidad por las regiones pericentroméricas, lo que indica que estos *loci* se mantienen en su estado de heterocromatina por un mecanismo diferente a la *RdDM* (Lee *et al.*, 2012; Zheng *et al.*, 2013).

Una vez que la Pol V genera un transcrito en un *locus* específico, el complejo AGO4/siRNA establece una interacción directa con la subunidad grande de la Pol V mediante sus motivos *Ago-hooks*, presentes en el dominio CTD (El-Shami *et al.*, 2007). Asimismo, AGO4 establece interacciones con KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1) (He *et al.*, 2009) y RNA-DIRECTED DNA METHYLATION 1 (RDM1), esta última forma un punto de conexión entre AGO4 y DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), que cataliza la metilación *de novo* del DNA (Matzke & Mosher, 2014). En las plantas, la metilación del DNA puede ocurrir en Citosinas en todos los contextos de secuencia: el contexto simétrico CG y CHG (en donde H= A, T o C) y el contexto asimétrico CHH. En *A. thaliana* los patrones globales de metilación del genoma corresponden a 24% para CG, 6.7% para CHG y 1.7% para CHH. Particularmente, la metilación llevada a cabo por DRM2 corresponde a la metilación en contexto CHH que se perpetúa a través de la *RdDM* (Law & Jacobsen, 2010).

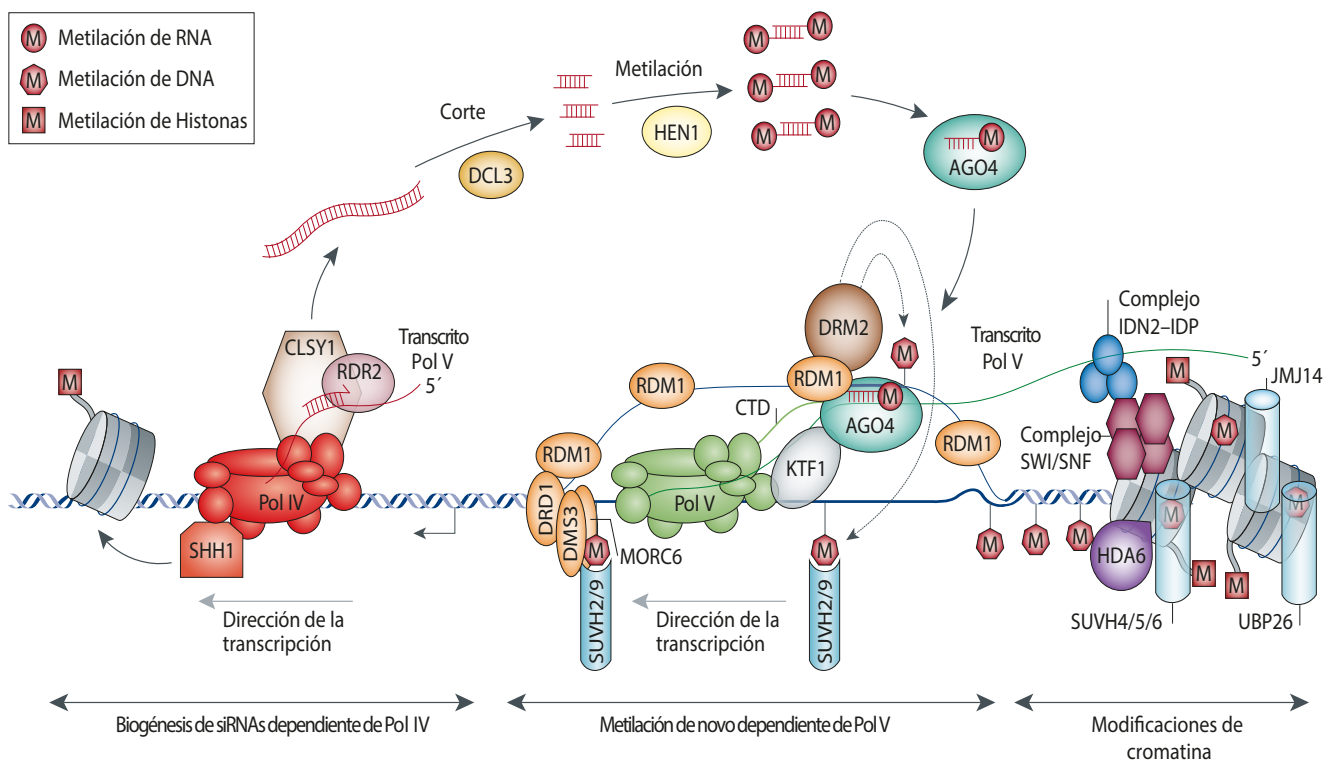
A esta última etapa de metilación del DNA se le suman un número creciente de interactores: DEFECTIVE IN MERISTEM SILENCING (DMS3) y MICRORCHIDIA 6 (MORC6) que posiblemente participan en mantener y estabilizar el estado desenrollado del DNA durante la transcripción, el remodelador de la cromatina SWI/SNF que participa en el ajuste del posicionamiento de los nucleosomas, entre otros. Adicionalmente, se ha observado que toda una maquinaria de remodeladores de la cromatina incluyendo desacetilasas de histonas como HISTONE DEACETYLASE 6 (HDA6) y metiltransferasas como SUVH4, SUVH5 y SUVH6, actúan en conjunto para el establecimiento y refuerzo de un estado de represión en los *loci* en donde se efectúa la *RdDM* (Matzke & Mosher, 2014).

#### 4.1.2 AGO6 en la vía RdDM

Mediante estudios de genética reversa se logró discernir la función de AGO6 en la *RdDM*. El primer reporte mostró que ésta tenía un papel parcialmente redundante en la acumulación de siRNAs

provenientes de *loci* heterocromáticos específicos y en la metilación de DNA (Zheng *et al.*, 2007). Estudios posteriores mostraron que la expresión de *AGO6* se encuentra restringida a ápices de raíz y brote y que mutaciones en este gen, a pesar de no mostrar un efecto en el desarrollo evidente, comprometen el silenciamiento mediado por siRNAs, por lo que se dedujo que esta proteína está relacionada con la vía *RdDM* en la que actúa de manera tejido específica (Eun *et al.*, 2011).

Recientemente se ha cambiado la visión del papel de *AGO6* durante la *RdDM*. Un par de estudios publicados en este año mostraron que *AGO6* posee una función específica y más que redundante, paralela con *AGO4* para efectuar el silenciamiento mediante metilación *de novo* (McCue *et al.*, 2015; Duan *et al.*, 2015). Por una parte, se demostró que la mayoría de los *loci* en la doble mutante *ago4 ago6* no muestran grandes cambios en los patrones de metilación en comparación con las mutantes de *ago4* o *ago6*, por ello concluyen que estas proteínas funcionan de manera paralela durante la *RdDM*. Asimismo, demostraron que *AGO4* y *AGO6* se distribuyen de manera desigual en el nucleoplasma y que éstas interactúan de manera excluyente con RNA polimerasas (*AGO4* con Pol IV mientras que *AGO6* co-localiza con Pol II), lo que sugiere que dentro de la misma célula, podrían tener funciones distintas que convergen en el silenciamiento de *loci* específicos (Duan *et al.*, 2015).



**Figura 4. Representación de la Vía canónica de Metilación del DNA dependiente del RNA (*RdDM*).** Se muestran todos los componentes reportados que participan en la vía *RdDM* canónica. Las flechas bidireccionales muestran las tres principales etapas durante esta vía: 1) Biogénesis de siRNAs dependiente de Pol IV, 2) Metilación *de novo* dependiente de Pol V y 3) Modificaciones en la cromatina. Tomado y modificado de Matzke y Mosher, 2014.

Adicionalmente, el estudio de Slotkin y colaboradores (2015) muestra que AGO6 es esencial en el silenciamiento de elementos transponibles transcripcionalmente activos mediante una vía de silenciamiento diferente e independiente de la RdDM canónica (*Pol IV-RdDM*). En este estudio se demostró que AGO6 tiene la capacidad de incorporar siRNAs de 21 a 22 nt producidos a partir de transcritos dependientes de la Pol II provenientes de elementos transponibles y procesados mediante la acción de enzimas como RDR6, DCL4, DCL2 y AGO1. Una vez que AGO6 forma el complejo con el siRNA de 21-22 nt, es capaz de dirigir la metilación *de novo* del DNA en estos *loci*. Estas observaciones demostraron el papel central de AGO6 en el establecimiento “inicial” del silenciamiento a nivel transcripcional a partir de un mecanismo de silenciamiento a nivel post-transcripcional, dicho en otras palabras, AGO6 funge como un punto de conexión siendo un intermediario entre el PTGS (ya que hace uso de siRNAs provenientes de transposones transcripcionalmente activos) y el TGS de *loci* específicos (ya que permite el establecimiento de marcas represivas en la cromatina a través de la RdDM). Finalmente, se observó que este mecanismo se restringe a células del meristemo floral debido a que AGO6 se acumula preferencialmente en estos tejidos, lo que permite el establecimiento de la metilación *de novo* en transposones antes del desarrollo de los órganos reproductivos y gametos (McCue *et al.*, 2015).

#### 4.1.3 AGO9 en la vía RdDM

Mutaciones en AGO9 en *Arabidopsis*, a diferencia de sus parálogos AGO4 y AGO6, ocasionan defectos en el desarrollo mostrando una proporción de células somáticas anormales capaces de diferenciarse en células germinales que producen numerosos sacos embrionarios, mostrando un fenotipo que asemeja a la apomixis (Olmedo-Monfil *et al.*, 2010).

AGO9 se acumula preferencialmente en las células somáticas acompañantes de la capa que se encuentra alrededor de las células reproductivas, no así en células reproductivas o sus precursores. Asimismo, AGO9 es capaz de interactuar con siRNAs de 24 nt provenientes principalmente de elementos transponibles de diferentes familias que se encuentran principalmente en regiones pericentroméricas (Olmedo-Monfil *et al.*, 2010; Durán-Figueroa & Vielle-Calzada, 2010). Debido a las observaciones anteriores, aunadas al hecho de que los fenotipos mutantes tanto

Apomixis: Es un tipo de reproducción sexual por medio de semillas. La apomixis puede dividirse en dos grandes tipos: gametofítica o esporofítica. En la apomixis gametofítica se producen gametos no reducidos (por meiosis), mientras que en la esporofítica, células somáticas dan origen directamente a un embrión.

de *rdm6* como de *sgs3* —ambos relacionados con el silenciamiento mediado por siRNAs— muestran características similares a la mutante *ago9*, se concluye que *AGO9* es necesaria para la inactivación de transposones en el óvulo a través de la propagación de señales de silenciamiento (siRNAs) desde las células somáticas circundantes, teniendo así un papel fundamental en la determinación de la línea germinal (Olmedo-Monfil *et al.*, 2010).

Se ha observado que la disrupción en la función de *AGO4* y *AGO6* si bien, disminuye drásticamente los niveles de metilación CHH en determinados *loci*, ésta no se pierde por completo. Debido a lo anterior se deduce que existen *loci* cuya metilación no depende de *AGO4* y *AGO6*. En muchos de estos *loci*, mutantes de la Pol IV y Pol V muestran una más notoria disminución en los niveles de metilación en comparación con la doble mutante *ago4 ago6*, lo que sugiere que otro miembro de la familia AGO, posiblemente *AGO9*, podría estar contribuyendo en la metilación de estos sitios (Duan *et al.*, 2015).

### 4.2 Importancia biológica de la *RdDM*

La principal función biológica de la *RdDM* se ha asociado con el mantenimiento de la estructura e integridad genómica debido a su importante papel en el silenciamiento de elementos transponibles y regiones repetidas (Fedoroff, 2012). Sin embargo, un número creciente de reportes indica que la *RdDM* posiblemente está relacionada con diferentes procesos con relevancia biológica en las plantas. A continuación se hablará de algunos procesos en los que se ha visto implicada.

#### 4.2.1 Control de Elementos Transponibles y Respuesta a Estrés

Los transposones comprenden un vasto grupo de fragmentos de DNA que poseen la capacidad de moverse e insertarse en nuevos sitios del genoma mediante un mecanismo de “corte y pegado” o a través de un intermediario de RNA. Los transposones fueron descubiertos por Barbara McClintock en maíz y los describió en el año de 1950 (McClintock, 1950). Actualmente se sabe que los transposones constituyen una gran parte del genoma en muchos eucariontes mostrando alta abundancia en las plantas en donde, por ejemplo, en el maíz alrededor del 85% del genoma son transposones (Fedoroff, 2012; Slotkin & Martienssen, 2007).

La vía *RdDM* está involucrada en el mantenimiento del estado reprimido de transposones y elementos repetitivos, evitando su activación y por ende, su inserción potencial en el genoma. Asimismo, esta vía está involucrada en el establecimiento de marcas epigenéticas en determinados *loci* manteniendo

un estado dinámico de silenciamiento y activación de genes que puede modificarse debido a la influencia de estímulos ambientales (Matzke & Moshier, 2014). El supuesto de que el ambiente puede modificar la expresión de los genes y la organización del genoma fue producto de años de investigación en plantas de maíz por la brillante genetista Barbara McClintock durante los años 50, logro que le valió el premio Nobel en el año de 1983 (McClintock, 1950; Ravindran, 2012).

Un ejemplo de lo anterior mencionado es el estudio de Pazskowski y colaboradores (2012), en donde demuestran los efectos del estrés por calor en la activación transcripcional del retrotransposón ONSEN cuando se encuentra comprometida la *RdDM*. Este grupo reportó que cuando se afecta esta vía la reactivación de ONSEN es mucho más evidente que en plantas silvestres, así mismo, observaron que nuevas inserciones de este retrotransposón en generaciones subsecuentes sólo tienen lugar en plantas mutantes en donde se interrumpe la función de Pol IV o RDR2. Estas observaciones sugieren que la *RdDM* es de suma importancia en la prevención de la movilidad transgeneracional de ONSEN. Asimismo, se observó que cuando llegaban a haber inserciones nuevas, éstas conferían sensibilidad de respuesta al estrés por calor a los genes cercanos, lo que sugiere que los eventos de movilidad pueden generar nuevos patrones de expresión génica que responden a estrés (Ito *et al.*, 2011).

#### 4.2.2 Defensa contra patógenos

La *RdDM* además de estar involucrada en respuesta a estrés abiótico, se ha relacionado con la respuesta de defensa de las plantas a patógenos bacterianos. Por una parte, similar a lo observado con el retrotransposón ONSEN, estudios globales de metilación en plantas sometidas a la infección por *Pseudomonas syringae* DC3000 mostraron que cuando se da la infección hay cambios en los patrones de metilación de transposones y que este fenómeno provoca a su vez cambios en los patrones de metilación de genes asociados a ellos, afectando sus niveles de expresión (Dowen *et al.*, 2012). Por otra parte, el grupo de Navarro y colaboradores (2013) mostraron que el péptido flg22, derivado del elicitor bacteriano flagelina, tiene un efecto de regulación negativa en diferentes componentes de la vía *RdDM* mientras que promueve la desmetilación de elementos transponibles y algunos genes de defensa, por tanto, proponen que durante la defensa antibacteriana los mecanismos de desmetilación del DNA dependiente de ROS-1 (una DNA glicosilasa) y la represión de la vía *RdDM* establecen un estado dinámico que permite la activación de determinados elementos transponibles que afectan de manera directa la expresión de genes cercanos, entre ellos genes de defensa (Yu *et al.*, 2013).

### 4.3 Importancia evolutiva de la *RdDM* en las plantas terrestres

Las plantas, como organismos sésiles, están expuestas a un gran número de estímulos ambientales a lo largo de su ciclo de vida. Dichos estímulos son percibidos y transmitidos por una miríada de vías de señales de transducción que activan determinados factores transcripcionales en el núcleo y finalmente conllevan a la activación de genes que codifican a proteínas efectoras que permiten a las plantas contender contra el estrés (Mirouze & Paszkowski, 2011).

La influencia de los estímulos ambientales en los procesos epigenéticos y sus efectos en la respuesta de las plantas al estrés, así como su carácter heredable han acrecentado la curiosidad sobre su potencial adaptativo. Si bien se sabe que estos procesos son de suma importancia para las respuestas apropiadas de las plantas al estrés a corto plazo, actualmente se desconoce si estos caracteres adquiridos podrían tener alguna relevancia en la evolución. Por otra parte, debido al estilo de vida de las plantas y el hecho de que su línea germinal se desarrolla de manera tardía, los estímulos percibidos durante el crecimiento vegetativo de la planta podrían generar una memoria epigenética que permanezca en la línea germinal y finalmente sea heredada a su progenie. Por tanto, si las modificaciones epigenéticas tienen consecuencias fenotípicas para el organismo y éstas son heredadas de manera estable, constituyen una forma de variación que podría ser sujeta a presiones selectivas y de esta manera tener relevancia para la adaptación de las siguientes generaciones si se presentara un estrés similar (Mirouze & Paszkowski, 2011; Hirsch *et al.*, 2012; Matzke & Mosher, 2014).

Memoria Epigenética: Modificaciones en el DNA de una célula que no afectan sus secuencia y fueron heredadas -ya sea a través de divisiones mitóticas o meióticas- de la célula de la que desciende.

Algunos estudios han implicado la vía *RdDM* en respuesta a estrés por ejemplo: de manera directa, en el silenciamiento de genes involucrados en el desarrollo de estomas durante condiciones de baja humedad (Tricker *et al.*, 2012) y de manera circunstancial, en la resistencia basal al estrés por calor al activar transposones y sus genes asociados (Ito *et al.*, 2012; Popova *et al.*, 2013), así como en la resistencia a estrés inducida por agentes genotóxicos (Yao *et al.*, 2010). Los efectos fenotípicos visibles en estos casos podrían incrementar la variabilidad interespecífica y por tanto, tener efecto en su adaptación a un ambiente cambiante (Matzke & Mosher, 2014).

Un fenómeno con efecto transgeneracional en donde la *RdDM* podría estar implicada es la *paramutación*. La *paramutación* fue descrita en el año de 1956 por Alexander Brink (Brink, 1956) en maíz en el gen *red1* y subsecuentemente descrita para el gen *booster1* por Ed Coe en 1959 (Coe, 1959), ambos genes que codifican a factores transcripcionales que activan la vía de biosíntesis de antocianinas. La *paramutación* es un fenómeno en donde existe interacción entre alelos y esta interacción conlleva a un cambio heredable en la expresión génica, este fenómeno ocurre sin alterar las secuencias de DNA, por tanto los estados alterados de la expresión son mediados por mecanismos epigenéticos (Stam, 2009; Arteaga-Vázquez & Chandler, 2010; Hollick, 2010). Análisis de genética reversa mostraron que los genes necesarios para poder llevar a cabo la *paramutación* en maíz muestran homología con genes involucrados en la vía *RdDM* en *A. thaliana* (Alleman *et al.*, 2006; Woodhouse *et al.*, 2006; Hale *et al.*, 2007; Erhard *et al.*, 2009). La producción de siRNAs y la naturaleza molecular de la *paramutación*, sugieren que la *RdDM* podría estar participando en este fenómeno, mostrando un escenario en donde siRNAs producidos en un alelo silenciado ocasionan una producción adicional de siRNAs en el alelo homólogo inicialmente activo (Arteaga-Vázquez & Chandler, 2010). A pesar de lo anterior, se ha mostrado que existe una diferencia entre la metilación y producción de siRNAs entre los alelos activos e inactivos, asimismo, la mutación de los componentes de la *RdDM* no muestran los mismos efectos en la *paramutación*, por tanto si bien la *RdDM* está involucrada, podrían estar actuando otros mecanismos para que este fenómeno se lleve a cabo (Stam, 2009; Arteaga-Vázquez & Chandler, 2010; Hollick, 2010; Ehrhard & Hollick, 2011). Finalmente, la relevancia evolutiva de la *paramutación* subyace en el hecho de que en ausencia de variabilidad genética, la *paramutación* podría fungir como una fuente de variación a través de la transferencia de estados transcripcionalmente favorables para la progenie (Chandler, 2007).

Paramutación:  
Interacción alélica  
que conlleva a  
cambios heredables  
en la expresión  
génica.

#### 4.3.1 Conservación evolutiva de los componentes de la *RdDM*

Estudios de genómica comparativa han demostrado que muchos de los componentes de la vía *RdDM* están conservados evolutivamente entre las plantas terrestres. Se han identificado que las proteínas AGO, miembros de la familia DCL, así como miembros de la familia RDR están presentes en todos los grupos de embriofitas. Asimismo, algunos de las subunidades de las RNA polimerasas IV

y V, componentes esenciales de la *RdDM*, se encuentran presentes en especies pertenecientes a los clados basales, como es el caso de *Physcomitrella patens* y *Marchantia paleacea* (Huang *et al.*, 2015) (Figura 5).

Asimismo, se ha observado que los siRNAs de 24 nt están presentes en todas las embriofitas, mostrando una mayor abundancia en monocotiledóneas y dicotiledóneas, seguido de una menor abundancia en angiospermas basales, gimnospermas y helechos (Matzke & Mosher, 2014). En el caso de las gimnospermas, la distribución de los siRNAs parece estar restringida a tejidos reproductivos. En briofitas como *Physcomitrella patens*, se ha encontrado bajos niveles de siRNAs que se producen principalmente en transposones (Coruh *et al.*, 2015) Por el contrario, en *Chlamydomonas reinhardtii* no se ha reportado la presencia de siRNAs de 24 nt (Arif *et al.*, 2013; Matzke & Mosher, 2014).

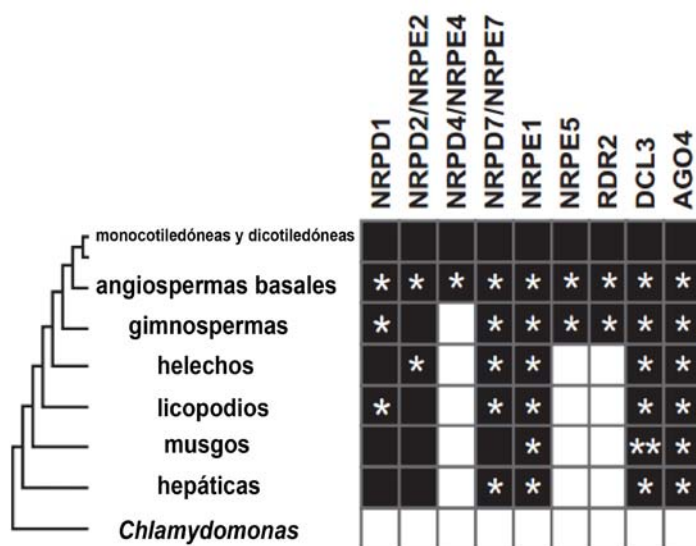


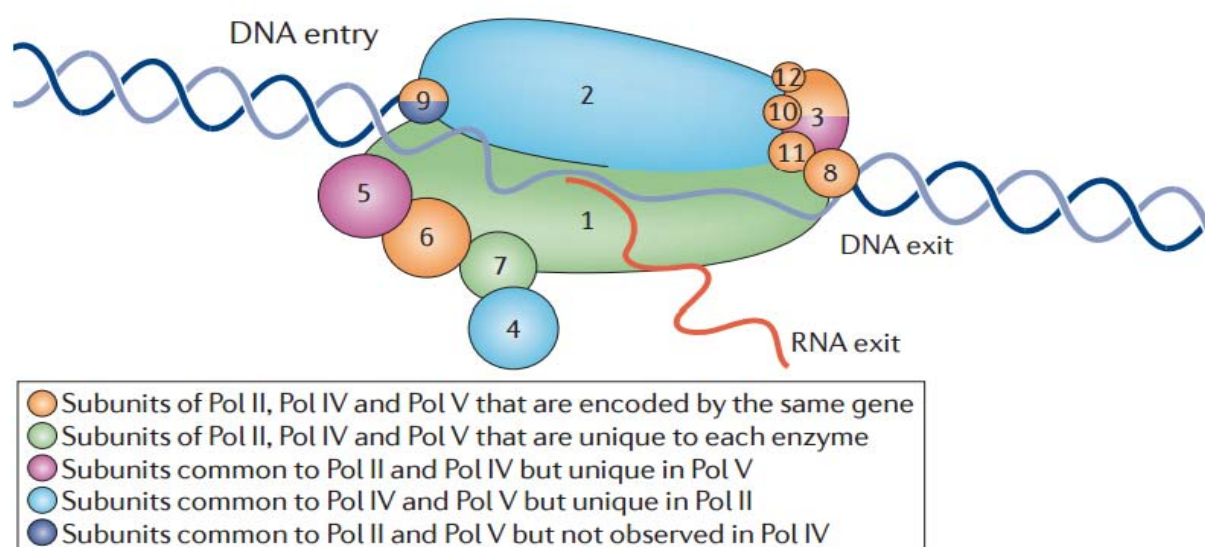
Figura 5. Conservación evolutiva de componentes de la vía *RdDM* en las plantas terrestres. En negro se muestran los genes identificados previamente en los diferentes grupos de embriofitas, los asteriscos muestran los genes identificados en el estudio de Huang y colaboradores (2015). Tomado y modificado de Huang *et al.*, 2015.

En conjunto, las observaciones anteriores sugieren que tanto los componentes de la vía *RdDM* como los siRNAs necesarios para llevarla a cabo se encuentran conservados evolutivamente entre todas las plantas terrestres y que posiblemente sean vías de silenciamiento funcionales en los diferentes grupos. Sin embargo, existen diferencias en los mecanismos a nivel específico —como subunidades específicas de las RNA polimerasas, dominios de expresión tejido-específicos, enzimas especializadas— que pudieron haber contribuido con su éxito adaptativo (Huang *et al.*, 2015).



#### 4.3.2 Las RNA polimerasas IV y V como importantes innovaciones evolutivas en la *RdDM*

En todos los eucariontes existen tres RNA polimerasas dependientes de DNA necesarias para la producción de la miríada de RNAs esenciales para la vida. De manera general, la RNA polimerasa I (Pol I) se encarga de la producción de los RNAs catalíticos que constituyen a los ribosomas, la RNA polimerasa II (Pol II) transcribe numerosos genes que codifican a proteínas, RNAs largos no codificantes y RNAs pequeños que modifican o procesan RNAs mensajeros o ribosomales y por último, la RNA polimerasa III (Pol III) transcribe varias clases de RNAs pequeños (usualmente de un tamaño menor a 500 nt) como la subunidad ribosomal 5S, RNAs de transferencia y RNAs regulatorios pequeños (Haag & Pikaard, 2011). En el caso de las plantas, se encuentran dos RNA polimerasas adicionales que están estrechamente relacionadas con la Pol II: la RNA polimerasa IV (Pol IV) y la RNA polimerasa V (Pol V), enzimas fundamentales para la vía *RdDM*. Debido a su alto grado de similitud con la Pol II, se piensa que estas polimerasas se originaron a causa de duplicación y posterior subfuncionalización, sin embargo, a pesar de su origen común, las Pol IV y V posiblemente muestren diferencias a nivel funcional debido a su Dominio Carboxilo Terminal (CTD, por sus siglas en inglés *C-terminal Domain*) que posiblemente surgió a causa de recombinaciones ilegítimas (Luo & Hall, 2007; Haag & Pikaard, 2011).



**Figura 6. Composición de las subunidades de las RNA polimerasas II, IV y V en *A. thaliana*.** Se esquematizan las posiciones relativas de las subunidades de la RNA polimerasa, mostrando las relaciones entre la Pol II, IV y V de *A. thaliana* con respecto a la composición de subunidades. Las subunidades se encuentran numeradas y coloreadas según la clave mostrada en la parte inferior. Las subunidades tres y nueve se muestran en dos colores representando formas alternativas. Tomada de Haag & Pikaard (2011).

Las Pol IV y V, al igual que el resto de las RNA polimerasas eucariontes, son holoenzimas compuestas por 12 subunidades. Algunas de las subunidades están compartidas por las cinco polimerasas, mientras que existen algunas que se incorporan de manera específica, este hecho muestra el alto grado de especialización a nivel funcional de estas enzimas. La subunidad más grande y la segunda más grande conforman el centro catalítico (Haag & Pikaard, 2011) (Figura 6). En las plantas con flor es en el grupo en el que se observa una mayor especialización de las RNA polimerasas IV y V, presentando subunidades específicas de este clado, lo que indica que esta especialización pudo haber tenido un papel muy importante en su rápida adaptación y extraordinaria radiación durante el Cretácico medio (Huang *et al.*, 2015).

Monofilético:  
Grupo o clado  
que incluye a un  
ancestro y todos sus  
descendientes

## 5. El origen de las embriofitas y *Marchantia polymorpha* como modelo emergente

El origen de las plantas terrestres fue un evento clave en la historia de la vida en la Tierra. Las plantas terrestres o embriofitas comprenden un grupo *monofilético* cuyos orígenes se remontan al período Ordovícico hace aproximadamente 480 millones de años (Bowman *et al.*, 2007, Graham & Wilcox, 2000). Las plantas heredaron muchas características bioquímicas, ultraestructurales y fisiológicas de sus ancestros acuáticos, sin embargo, durante la evolución temprana de las embriofitas surgieron numerosas innovaciones que permitieron la colonización del ambiente terrestre: la protección parental del embrión en desarrollo, el desarrollo de estructuras multicelulares para la producción de gametos, la alternancia de generaciones en donde una fase esporofítica diploide da origen a una fase gametofítica haploide multicelular, la percepción de factores ambientales como luz y gravedad, el origen de matrices extracelulares (esporopolenina, lignina y pectina), el establecimiento de redes de comunicación intercelular (plasmodesmos, fitohormonas, receptores y sus ligandos) y la diversificación de redes de regulación génica para promover la diferenciación celular (Bowman *et al.*, 2007; Wickett *et al.*, 2014).

Los eventos posteriores a la colonización del ambiente terrestre consistieron en el surgimiento de innovaciones que expandieron la diversidad

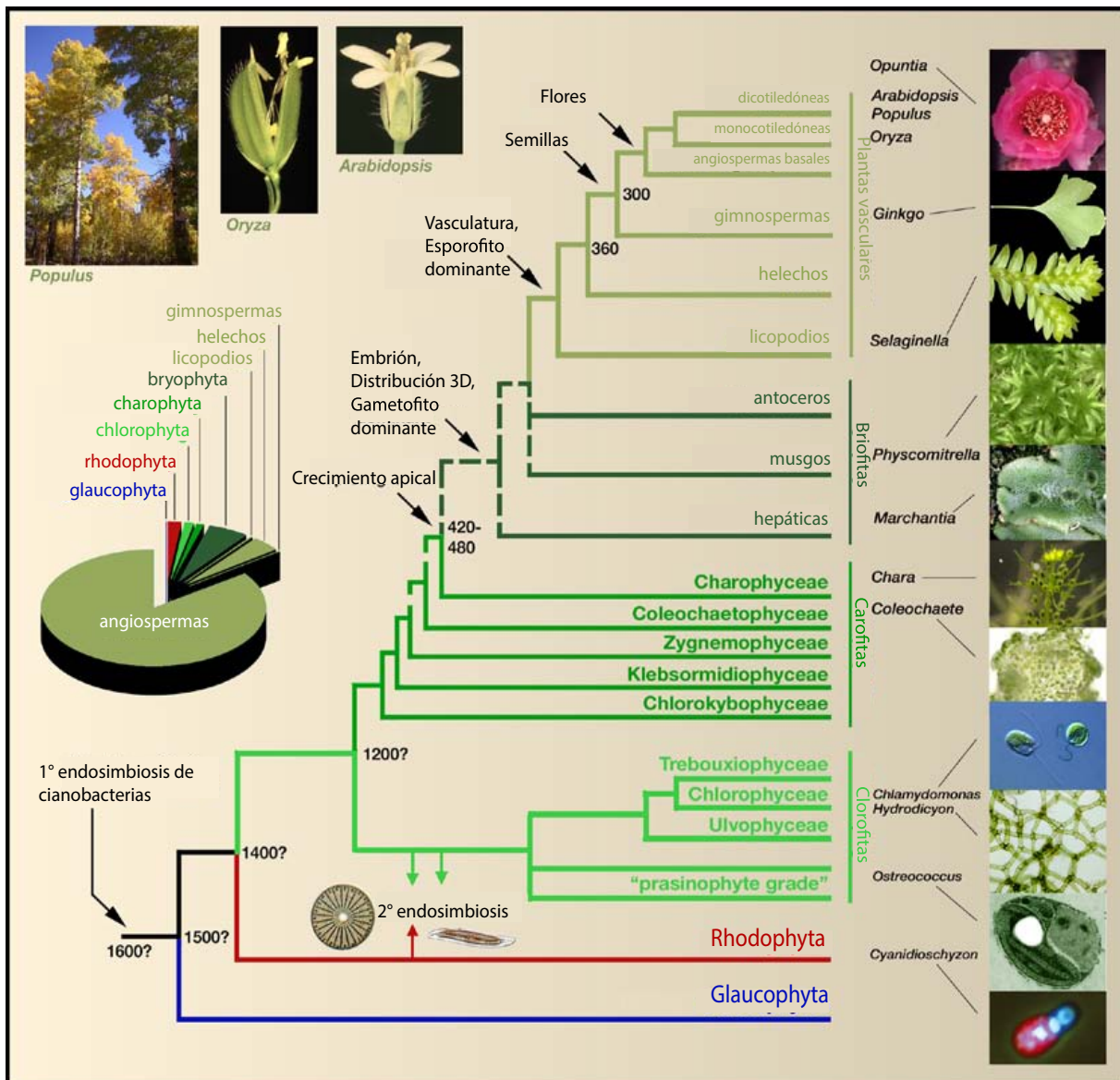
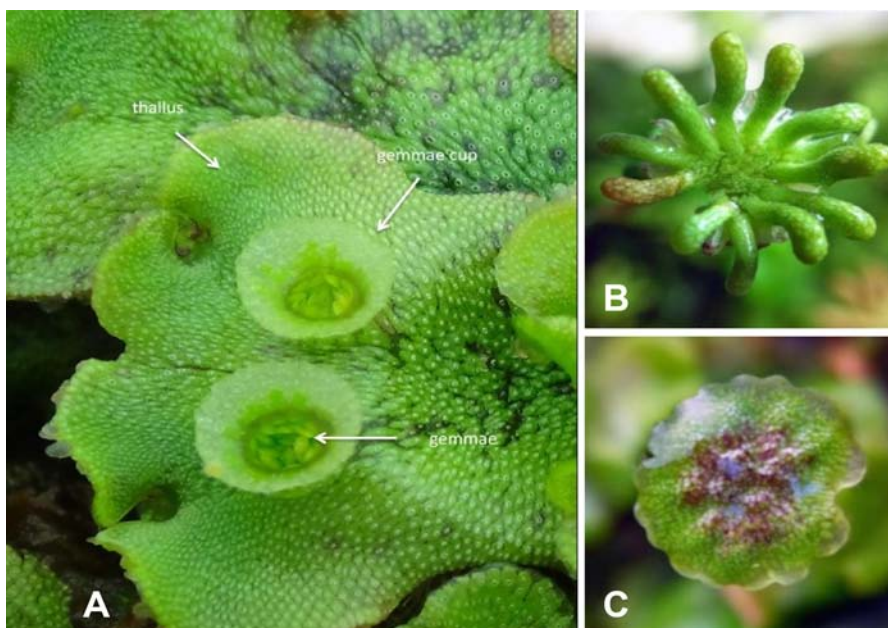


Figura 7. Relaciones filogenéticas entre las plantas terrestres. Se muestra una reconstrucción filogenética de los diferentes grupos de plantas. Las flechas indican el surgimiento de innovaciones evolutivas clave para la divergencia de las embriofitas. En la gráfica de barras se muestran los diferentes grupos de embriofitas actuales según su abundancia. Se incluyen también diferentes grupos de algas. Tomado y modificado de Bowman y colaboradores (2009).

de la flora terrestre (Bowman *et al.*, 2007). Por ejemplo, el surgimiento de un sistema vascular, así como características para prevenir la desecación en las traqueofitas, permitió el establecimiento de un sistema más eficiente para el transporte y retención del agua, fotosintatos y otros nutrientes. Las innovaciones fisiológicas fueron acompañadas de un cambio en el ciclo de vida: de un ciclo de vida gametofítico a un ciclo de vida esporofítico predominante. El origen del linaje de las espermatofitas o plantas con semillas incrementó de manera importante la provisión parental al embrión y el origen de la flor en el linaje de las angiospermas promovieron una serie de radiaciones rápidas, convirtiéndose actualmente en el grupo más diverso de plantas en el planeta (Qiu & Palmer, 1999; Bowman *et al.*, 2007; Wickett *et al.*, 2014; Graham & Wilcox, 2000).

El phylum *Bryophyta sensu lato* está conformado por los musgos, antoceros y hepáticas y es considerado como el grupo de más temprana diversificación entre las embriofitas. Dentro de este grupo, diversos estudios taxonómicos y moleculares posicionan a las hepáticas como grupo hermano de las plantas vasculares o traqueofitas (Bowman, 2007; Qiu *et al.*, 2007). Sin embargo, estudios recientes muestran que los antoceros podrían ocupar esta posición (Wickett *et al.*, 2014). Las briofitas, como grupo basal, tienen características que las diferencian de las traqueofitas: carecen de un sistema vascular, no presentan lignificación en su pared celular, presentan gametos masculinos móviles y una fase gametofítica haploide dominante sobre la fase esporofítica diploide durante su ciclo de vida (Mischler, 2001). Las briofitas poseen características biológicas que representan una gran oportunidad para servir como modelos de estudio para investigación en macroevolución, genética de poblaciones y ecología (Mischler, 2001). Asimismo, las briofitas recientemente han sido adoptadas como modelo para estudios en genómica comparativa para responder preguntas fundamentales en la biología de plantas, como las bases genéticas de las innovaciones clave que permitieron a las plantas evolucionar a partir de un ancestro acuático y adaptarse a la vida terrestre, los cambios en el desarrollo responsables de la multicelularidad, así como las alteraciones en el *bauplan* dentro de las plantas terrestres (Bowman *et al.*, 2007; Ishizaki *et al.* 2015).

Bauplan: Modelo de organización, diseño o plan estructural.



**Figura 8. Gametofito de *M. polymorpha*.** A) Se muestra un talo maduro de *M. polymorpha*, se señalan con flechas los conceptáculos y dentro de ellos, los propágulos o gemas. B) Se muestra un gametangióforo femenino o arquegonióforo. C) Se muestra un gametangióforo masculino o anteridióforo. Tomado de Lara-Mondragón & Arteaga-Vázquez (2015)

*Marchantia polymorpha* (Figura 8) es una hepática taloide de amplia distribución geográfica (Steere, 1940). Esta planta tiene una larga historia como modelo experimental cuyos registros se remontan hasta mediados del siglo XV (Bowman, 2015). En los últimos siglos, se ha recabado una gran cantidad de información relativa a su anatomía, fisiología y procesos del desarrollo, adicionalmente, con el surgimiento de la era molecular y el advenimiento de la tecnología de secuenciación masiva, se ha logrado obtener la secuencia completa del cromosoma Y, cloroplasto y mitocondria, lo que ha permitido realizar estudios de genómica comparativa desde un punto de vista evolutivo (Bowman, 2015; Ishizaki *et al.*, 2015). La secuenciación del genoma de *M. polymorpha*, mediante un esfuerzo conjunto de diferentes grupos de investigación y bajo la dirección del *Joint Genome Institute* (<http://jgi.doe.gov/>), ha revelado que muchos de los genes que regulan procesos del crecimiento y desarrollo en otras plantas terrestres están conservados evolutivamente en el genoma de *M. polymorpha*, mostrando una menor redundancia (Ishizaki *et al.*, 2008; Ishizaki *et al.*, 2015).

Actualmente, el uso de *M. polymorpha* como modelo experimental en estudios de genómica comparativa es cada vez más frecuente. En años recientes se han desarrollado numerosas técnicas de biología molecular, en donde destacan diversos protocolos de transformación genética mediante *Agrobacterium* (Kubota *et al.*, 2013; Tsuboyama & Kodama, 2014; Tsuboyama & Kodama, 2015) y modificación dirigida del genoma (*i.e.* CRISPR/Cas9 (Sugano *et al.*, 2014), recombinación homóloga (Ishizaki *et al.*, 2013), uso de miRNAs artificiales (Flores-Sandoval *et al.*, 2015), entre otros), adicionalmente, sus características biológicas y su versatilidad de mantenimiento en condiciones de laboratorio, convierten a *M. polymorpha* en un modelo de estudio ideal para abordar cuestiones evolutivas, moleculares, celulares y del desarrollo (Ishizaki *et al.*, 2015; Bowman, 2015).

¿Cuáles son las características que hacen de *M. polymorpha* un modelo con gran potencial para estudios de genómica funcional? *M. polymorpha*, como otras briofitas, presenta un ciclo de vida con una fase haploide dominante, lo cual resulta muy favorable para escrutinios de mutantes en comparación con las plantas vasculares que presentan una fase predominantemente diploide. *M. polymorpha* presenta reproducción asexual mediante la producción de propágulos o gemas en estructuras denominadas conceptáculos (Figura 8.A), este tipo de reproducción resulta favorable pues permite la proliferación de biomasa isogénica en un tiempo relativamente corto. Por otra parte, también presenta reproducción sexual: las hembras producen células huevo en los arquegonios, mientras que los machos producen anterozoides en los anteridios, la fertilización depende de un medio acuoso para permitir la movilización del anterozoide hasta la célula huevo. Después de la fertilización, el embrión se divide de manera irregular para generar el esporangio que dará origen, a través de meiosis, a numerosas esporas haploides (Figura 9).

La fase diploide o esporofítica es breve y dependiente nutricionalmente de la fase gametofítica (Durand, 1908). Este tipo de reproducción y la excesiva producción de esporas permiten realizar mutagénesis en un solo evento y escrutinio de mutantes de manera rápida en comparación con especies de plantas vasculares. Adicionalmente, el ciclo de vida de *M. polymorpha* se lleva a cabo en un tiempo de alrededor de tres meses para el ciclo sexual y 2 a 3 semanas para el ciclo asexual, lo que favorece el estudio de fenómenos transgeneracionales (Ishizaki *et al.*, 2015).

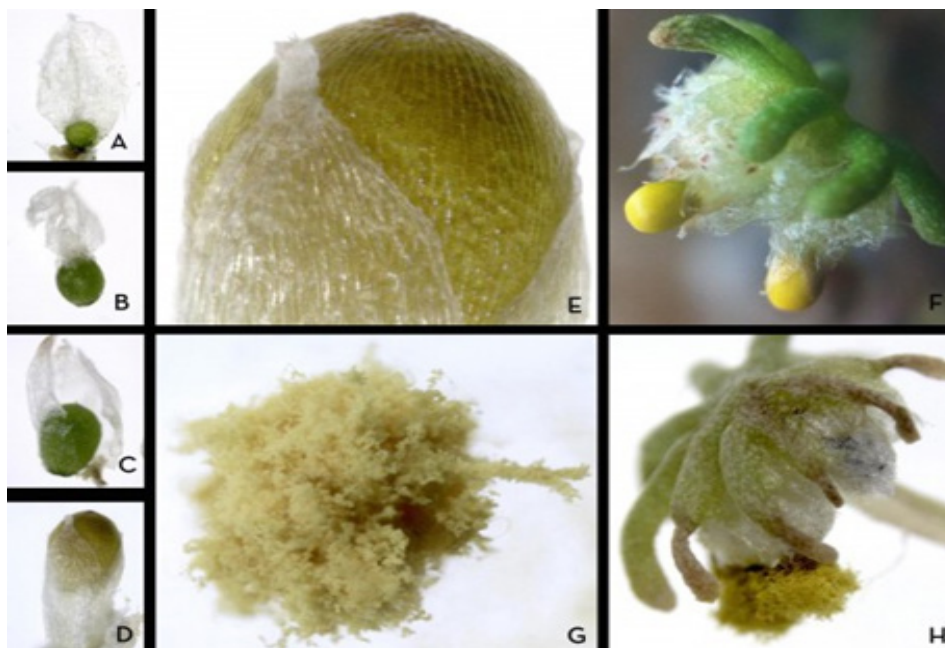


Figura 9. Desarrollo del esporangio en *M. polymorpha*. A-D) Diferentes estados de maduración del esporangio. E) Acercamiento del esporangio maduro. F) Arquegonióforo con esporangios maduros. G) Dehiscencia esporangial. H) Arquegonióforo con esporangios dehiscentes. Tomado de Lara-Mondragón & Arteaga-Vázquez (2015)

## PLANTEAMIENTO DEL PROBLEMA



La vía de metilación del DNA dependiente de RNA es una vía epigenética de suma importancia en la regulación a nivel transcripcional de elementos transponibles única en las plantas terrestres. En las angiospermas, esta vía contribuye de manera importante en el silenciamiento de transposones, elementos repetitivos y en la determinación celular. Asimismo, se le ha asociado recientemente con la respuesta ante estrés. Debido a lo anterior, se postula que esta vía constituye una fuente de variabilidad epigenética que pudo haber contribuido a la diversificación de las plantas terrestres. Por tanto, resulta interesante rastrear los orígenes de los componentes centrales de esta vía y averiguar si se encuentran conservados en los linajes basales de las embriofitas.

Los miembros del clado *AGO4/AGO6/AGO8/AGO9* son componentes centrales de la vía *RdDM*. Existen diferentes reportes en donde se evidencia el alto grado de conservación evolutiva de las proteínas de este clado en diferentes linajes de las plantas vasculares, así como en *Physcomitrella patens*, una planta no vascular. En el presente estudio emprendimos un análisis de genómica comparativa para descubrir cuáles son los miembros más ancestrales de la familia AGO en las plantas terrestres, con particular interés en la familia AGO9 debido a su papel dual en la especificación de los linajes germinales y el control de elementos transponibles.

## Hipótesis

La proteína AGO9 es necesaria para efectuar la vía de Metilación del DNA dependiente de RNA durante el control de la especificación celular del linaje germinal en *A. thaliana*, por tanto, en el genoma de *M. polymorpha* existe un homólogo a este gen y la proteína que codifica presenta características estructurales de una proteína AGO9 funcional.

## Objetivo General

Identificar el gen homólogo de AGO9 en *Marchantia polymorpha* y determinar si éste podría tener conservación evolutiva a nivel funcional.

## Objetivos particulares

Realizar un análisis bioinformático para identificar a los miembros de la familia AGO en el genoma de *M. polymorpha*.

Identificar la secuencia homóloga de AGO9 en *M. polymorpha* (MpAGO9) y realizar la anotación de este gen.

Determinar la arquitectura de dominios y modelar la estructura tridimensional de la proteína que codifica el ortólogo de AGO9 en *M. polymorpha* (MpAGO9).

Determinar los patrones de expresión del MpAGO9 en diferentes tejidos y etapas del desarrollo mediante ensayos de Transcripción Reversa acoplados a la Reacción en Cadena de la Polimerasa (RT-PCR).

Identificar mediante herramientas bioinformáticas potenciales interactores de MpAGO9 codificados en el genoma de *M. polymorpha*.

Determinar las interacciones de la proteína MpAGO9 con otros componentes de la vía RdDM mediante ensayos de Complementación por Fluorescencia Bimolecular (BiFC) en un sistema heterólogo.



## MATERIALES Y MÉTODO



### I. Análisis *in silico* de la familia AGO en *M. polymorpha* e identificación de MpAGO9

#### *Identificación de la familia ARGONAUTA en M. polymorpha*

Para identificar los miembros de la familia AGO en *M. polymorpha* se realizó una búsqueda bioinformática por medio del algoritmo tblastn, utilizando bases de datos de transcriptomas de diferentes tejidos de *M. polymorpha* generados en el Laboratorio de Epigenética y Biología del Desarrollo (INBIOTECA, Universidad Veracruzana). Como referencia, se utilizaron las secuencias de la familia AGO de *Arabidopsis thaliana* disponibles en la base de datos TAIR (*The Arabidopsis Information Resource*, [www.arabidopsis.org](http://www.arabidopsis.org)). Debido a que las proteínas AGO comparten el dominio PAZ con las proteínas de la familia DICER Like, para dicho análisis se emplearon ambos dominios (PAZ y PIWI), así como las secuencias completas de todos los miembros de la familia AGO. Una vez que se obtuvieron los transcritos correspondientes a miembros putativos de la familia AGO en *M. polymorpha*, se identificó el marco abierto de lectura de cada uno mediante la herramienta en línea *ORF Finder* (NCBI, disponible en línea <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) y se derivó su secuencia de aminoácidos. Para corroborar que los miembros putativos identificados en el análisis anterior fueran miembros de la familia AGO, se realizó un nuevo análisis mediante el programa HMMER (Finn *et al.*, 2015), que permite identificar en las secuencias de proteínas introducidas la "firma" de un dominio específico mediante comparaciones con bases de datos de proteínas. Para este análisis también se emplearon las secuencias de los dominios PAZ y PIWI.

#### *Identificación del gen ortólogo a ARGONAUTA9 en M. polymorpha*

Habiendo identificado los miembros de la familia AGO en *M. polymorpha*, se seleccionó aquella

secuencia que mostrara un mayor porcentaje de identidad con *AGO9* de *A. thaliana*. Posteriormente, se realizó un alineamiento recíproco (secuencia putativa de *AGO9* en *M. polymorpha* vs. base de datos TAIR) para confirmar que la secuencia identificada correspondiera a su contraparte en *A. thaliana*. Una vez confirmada la identidad del ortólogo de *AGO9* en *M. polymorpha* (*AGO10998* renombrado como *MpAGO9*), se determinó su estructura génica. Para lo anterior, se realizó una comparación de sus unidades transcripcionales contra el genoma de *M. polymorpha* (datos no publicados), al cual se tiene acceso gracias a convenios de colaboración con la Universidad de Kyoto, Japón. La determinación del número de intrones y exones en el gen, así como los extremos 5' y 3'UTR, se realizó con ayuda de los programas *Argo2 Genome Browser* (versión 1.0.31, *Broad Institute*) y *SnapGene* (versión 2.6, *GSL Biotech*).

### ***Reconstrucción filogenética***

Para confirmar la identidad del ortólogo de *MpAGO9*, se realizó una reconstrucción filogenética empleando las secuencias de los miembros de la familia AGO en *M. polymorpha* identificadas en el análisis anterior, así como las secuencias de las AGO de *A. thaliana* (disponibles en TAIR, [www.arabidopsis.org](http://www.arabidopsis.org)). Se realizó un alineamiento múltiple utilizando el algoritmo MUSCLE (Edgar *et al.*, 2004) y con éste se construyó un árbol filogenético de tipo *Neighbor Joining* utilizando la paquetería Phylip (versión 3.96, [evolution.genetics.washington.edu](http://evolution.genetics.washington.edu)), con 500 réplicas de bootstrap y utilizando un modelo de sustitución por bloques. Como grupo externo se utilizó la secuencia de la *AGO401* de *Drosophila melanogaster* (Apéndice 1). El árbol filogenético consenso resultante se editó con el programa *FigTree* (versión 1.2.4).

Asimismo, para determinar el grado de conservación evolutiva de este gen en las plantas terrestres se identificaron ortólogos putativos en diferentes grupos de embriofitas y se realizó una reconstrucción filogenética. Para lograr lo anterior se realizó una búsqueda de todos los miembros de la familia AGO en 11 especies representantes de los principales clados de las embriofitas (datos obtenidos en la base de datos ChromDB, [www.chromdb.org](http://www.chromdb.org)). Para identificar el ortólogo de *AGO9* en cada una de las especies seleccionadas se realizó el mismo procedimiento llevado a cabo para identificar a *MpAGO9*: identificación de cada miembro de la familia por especie, alineamiento recíproco contra la base de datos TAIR y reconstrucción filogenética utilizando los miembros putativos de la familia AGO por especie, así como los miembros de la familia en *A. thaliana*. Finalmente, las secuencias obtenidas se alinearon utilizando el algoritmo MUSCLE. El árbol filogenético se construyó bajo los parámetros descritos en el análisis anterior. Como grupo externo se utilizó la secuencia de *AGO401* de *D. melanogaster* (Apéndice 2).

## II. Identificación de dominios y modelación de la estructura proteica de MpAGO9

Empleando la herramienta en línea *Conserved Domains Database* (NCBI; Marchler-Bauer *et al.*, 2011), se identificaron los dominios característicos de la familia AGO en la secuencia de aminoácidos de MpAGO9. Asimismo, empleando el servidor SWISS-MODEL (Arnold *et al.*, 2003; <http://swissmodel.expasy.org/>) se generaron modelos tridimensionales de MpAGO9 y AtAGO9 basados en homología estructural con la proteína AGO cristalizada de *Aquifex aeolicus*, una bacteria termófila (Yuan *et al.*, 2005). Los modelos proteicos generados se visualizaron y editaron mediante el programa PyMOL *Molecular Graphics System* (Version 1.7.4 Schrödinger, LLC).

## III. Monitoreo de la expresión de MpAGO9 durante el ciclo de vida de *M. polymorpha*.

### *Diseño de oligonucleótidos para el extremo 3' UTR*

Se diseñaron oligonucleótidos específicos para amplificar una región de 127 pb dentro del extremo 3'UTR del gen para determinar el perfil de expresión de MpAGO9 en diferentes etapas del ciclo de vida de la planta. Los oligonucleótidos se diseñaron con el programa *SnapGene* (versión 2.6, *GSL Biotech*) y las propiedades de los mismos (T<sub>m</sub>, auto-complementariedad, formación de dímeros y/o estructuras secundarias) se verificaron con el programa *OligoCalc* (Kibbe, 2007; <http://www.basic.northwestern.edu/biotools/oligocalc>). Las secuencias de los oligonucleótidos sentido y antisentido se muestran en el Apéndice 3.

### *Siembra, mantenimiento y colecta de tejidos vegetales.*

Las plantas utilizadas en este estudio fueron dos variedades japonesas, Takaragaike-1 (Tak1, macho) y Kitashirakawa-2 (Kit-2, hembra), obtenidas por los convenios de colaboración mencionados anteriormente. El cultivo de las plantas *in vitro* y en sustrato se realizó con base en protocolos ya publicados (Kubota *et al.*, 2013).

Se obtuvieron gemas de plantas adultas y se sembraron en medio Gamborg B5 de fuerza media (1/2 B5) con 1% de agar. Las plantas se mantuvieron en una cámara de crecimiento a 20°C con luz blanca continua. Para la inducción de estructuras reproductivas, plantas de alrededor de 14 días se trasladaron

del medio de cultivo a un sustrato de lana mineral estéril y se sometieron a un espectro de luz roja lejana. Las condiciones de mantenimiento fueron iguales a las mencionadas anteriormente. Las colectas se realizaron de manera periódica para obtener tejidos desde gemas hasta estructuras reproductivas. Los tejidos colectados se almacenaron a -80°C hasta su uso.

### ***Extracción de RNA total.***

Los tejidos congelados se molieron por separado con mortero y pistilo estériles. La extracción de RNA total se realizó utilizando TRI Reagent (Sigma-Aldrich, Cat. No. T9424) siguiendo las indicaciones del fabricante (Apéndice 4). Posteriormente se realizó un lavado con fenol ácido (Ambion, Cat. No. AM9720) para eliminar residuos de DNA y/o proteínas (Apéndice 5). El RNA se cuantificó por espectrofotometría utilizando el equipo NanoDrop 2000 (Thermo Scientific) y se verificó su identidad mediante electroforesis en un gel de agarosa al 2% con Bromuro de Etidio (Anexo 1).

### ***Síntesis de cDNA y Reacción en Cadena de la Polimerasa (PCR)***

Con el RNA obtenido de los diferentes tejidos de *M. polymorpha*, se llevó a cabo la síntesis de cDNA por retrotranscripción mediante el kit *RevertAid RT* (Thermo Scientific, #K1691) siguiendo las instrucciones del fabricante (Apéndice 6). Se partió de 1µg de RNA total para cada reacción.

La amplificación del 3'UTR de *MpAGO9* se llevó a cabo utilizando el *GoTaq Green Master Mix* (Promega, Cat. No. M7122) que contiene la DNA polimerasa Taq, dNTPs, buffers de reacción y buffers de carga. Cada una de las reacciones de PCR se componía de: 6 µl *GoTaq Green Master Mix*, 1 µl oligonucleótido sentido, 1 µl oligonucleótido antisentido, 1 µl cDNA y 3 µl agua. Se incluyeron dos controles: como control positivo se utilizó el Factor de Elongación 1 de *M. polymorpha* (*MpEF1*) por cada muestra de una etapa específica y como control negativo una reacción sin templado. La PCR se llevó a cabo en un termociclador *My Cycler* (BioRad). Para corroborar el tamaño de los productos de PCR, éstos se corrieron en un gel de agarosa al 2% tenido con Bromuro de Etidio utilizando un marcador *O'GeneRuler ready-to-use 100 pb DNA Ladder* (Fermentas, #SM1143).

## **IV. Ensayo de Complementación por Fluorescencia Bimolecular (*BiFC*)**

Para probar si *MpAGO9* posee la capacidad de establecer interacciones con otros componentes de la vía *RdDM* se llevaron a cabo ensayos de tipo *BiFC*. Para ello, se evaluó la capacidad de interacción entre *MpAGO9* y *NRPE1*, subunidad grande de la RNA polimerasa V de *A. thaliana*, componente fundamental

de la *RdDM* y que ya se ha probado que interactúa con *AGO4* (Li *et al.*, 2006; El-Shami *et al.*, 2007).

### ***Diseño de oligonucleótidos para amplificar los interactores.***

Se diseñaron oligonucleótidos específicos para amplificar el gen *AGO9* de *M. polymorpha*. Asimismo, se diseñó un par de oligonucleótidos para *MpAGO9* adicionando los sitios de recombinación attB necesarios para la introducción al vector de entrada en el sistema de clonación Gateway y truncando el codón de paro para poder realizar fusiones traduccionales al extremo C- terminal. La interacción entre las proteínas *AGO4* y *NRPE1* de *A. thaliana* se ha probado en estudios previos mediante inmunofluorescencia y co-precipitación (Li *et al.*, 2006; El-Shami *et al.*, 2007), por lo que para este estudio se probó también la interacción de estas proteínas mediante el sistema *BiFC*, por tanto, se diseñaron oligonucleótidos para amplificar dichos genes y otro par para adicionar los sitios de recombinación attB y truncar el codón de paro (Apéndice 3).

### ***Obtención de DNA genómico.***

Se extrajo DNA genómico a partir de talos maduros de *M. polymorpha* mediante el protocolo de extracción con el buffer CTAB (Apéndice 7). El DNA se cuantificó mediante espectrofotometría y su integridad se verificó mediante electroforesis en un gel de agarosa al 0.8% teñido con Bromuro de Etidio. El DNA de *A. thaliana* fue proporcionado por la Bióloga Araceli Oropeza Aburto del Laboratorio Nacional de Genómica para la Diversidad (LANGEBIO-IPN). El DNA obtenido sirvió como templado para las reacciones de PCR subsecuentes.

### ***Obtención de los fragmentos de PCR-attB.***

La Reacción en Cadena de la Polimerasa (PCR) se llevó a cabo utilizando la DNA polimerasa de alta fidelidad *Phusion* (New England Biolabs, Cat. No. M0530S). Las reacciones de PCR para cada uno de los genes a amplificar consistían en: 4  $\mu\text{L}$  de Buffer HF 5x, 0.4  $\mu\text{L}$  de dNTPs, 1  $\mu\text{L}$  de su respectivo oligonucleótido sentido y 1  $\mu\text{L}$  de su respectivo oligonucleótido antisentido a una concentración de 0.5 pmoles/ $\mu\text{L}$ , 0.2  $\mu\text{L}$  de *Phusion*, 12.4  $\mu\text{L}$  de agua libre de nucleasas y 1  $\mu\text{L}$  de templado (aproximadamente 20 ng).

El tamaño de los genes a amplificar (adicionando los sitios de recombinación attB y truncando el codón de paro) y las condiciones de PCR empleadas se muestran en la Tabla 2. La PCR se llevó a cabo en un termociclador *Mastercycler gradient* (Eppendorf).

**Tabla 2. Condiciones de PCR para cada uno de los genes amplificados.**

Nombre	Tamaño	Programa de PCR
MpAGO9-attB	5,528 pb	98° C por 3 minutos; 35 ciclos de 98°C por 20 segundos, 60°C por 30 segundos y 72°C por 3 minutos con 30 segundos; seguido de 5 minutos a 72°C.
AtAGO4-attB	4,767 pb	98° C por 3 minutos; 35 ciclos de 98°C por 20 segundos, 60°C por 30 segundos y 72°C por 3 minutos; seguido de 5 minutos a 72°C.
AtNRPE1-attB	8,376 pb	98° C por 3 minutos; 35 ciclos de 98°C por 20 segundos, 60°C por 30 segundos y 72°C por 4 minutos con 30 segundos; seguido de 5 minutos a 72°C.

El tamaño de los fragmentos obtenidos se verificó mediante electroforesis en geles de agarosa al 0.8% teñidos con Bromuro de Etidio. Se utilizó el marcador molecular de *O'GeneRuler 1kb DNA Ladder Plus* (Fermentas, # SM1343).

Posteriormente los fragmentos de PCR-attB se purificaron con PEG 30% y MgCl<sub>2</sub> 30 mM según el protocolo indicado en el manual *Gateway BP Clonase II Enzyme Mix* (Invitrogen) (Apéndice 8). Los productos de PCR-attB se resuspendieron en buffer TE pH 8, se cuantificaron por espectrofotometría y se verificaron nuevamente por electroforesis en gel de agarosa al 0.8% teñido con Bromuro de Etidio.

### ***Reacción BP: Clonación en el vector de entrada pDONR221***

Las reacciones de recombinación BP se realizaron utilizando la *Gateway BP Clonase II Enzyme Mix* (Invitrogen, Cat. No. 11789) siguiendo las instrucciones del fabricante. Las reacciones BP se establecieron utilizando cantidades equimolares de vector de entrada pDONR221 y producto de PCR-attB. Dicha relación equimolar se calculó mediante la siguiente fórmula:

$$ng = f\text{mol} (N) \left[ \frac{660 \text{ fg}}{f\text{mol}} \right] \left[ \frac{1 \text{ ng}}{10^6 \text{ fg}} \right]$$

Donde N es igual al tamaño del DNA (ya sea el tamaño del vector o el tamaño del inserto) en pb.

La reacción BP para cada uno de los genes se preparó de la siguiente manera: producto de PCR-

attB (hasta 150 ng), vector de entrada pDONR221 (hasta 150 ng) y buffer TE pH 8 hasta alcanzar el volumen final de 8  $\mu$ L. Finalmente, se adicionaron 2  $\mu$ L de *BP Clonase II Enzyme Mix*. Las reacciones se incubaron a 25°C durante toda la noche. Al día siguiente las reacciones se detuvieron adicionando 1  $\mu$ L de Proteinase K y se incubaron a 37°C durante 10 minutos.

### ***Transformación de Escherichia coli por electroporación***

Para la electroporación se utilizaron 50  $\mu$ L de células electrocompetentes de la cepa DH5  $\alpha$  a las que, una vez descongeladas, se les adicionaron 3-4  $\mu$ L de producto de la reacción BP previamente dializado en una membrana de celulosa. Las células se transfirieron a una celda de electroporación y al ponerse en contacto con el electrodo, se dio un pulso de 1.8 kV durante 5.8 milisegundos. Posteriormente se adicionó 1 mL de medio SOC (2% Triptona, 0.5% Extracto de Levadura, 0.05% NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 20 mM Glucosa), el contenido de la celda se transfirió a un tubo Eppendorf de 1.5 mL e incubó a 37°C durante 1 hora.

Habiendo transcurrido el tiempo de incubación, las bacterias se sembraron en placas Petri con medio de selección LB (Luria-Bertani) sólido con Kanamicina (50  $\mu$ g/mL) y se incubaron a 37°C durante 12 horas.

### ***Escrutinio de clonas***

Los plásmidos de las bacterias transformantes se extrajeron por lisis alcalina (Apéndice 9). El DNA plasmídico extraído se digirió con enzimas de restricción para corroborar si la recombinación BP fue exitosa. Las enzimas de restricción utilizadas variaban en función de la construcción a analizar. Las reacciones se incubaron a 37°C durante 2-4 horas. Los productos de la digestión se separaron mediante electroforesis en gel de agarosa al 0.8% teñido con Bromuro de Etidio y se compararon con los patrones de restricción de digestiones *in silico* de las mismas construcciones, utilizando las mismas enzimas. Las digestiones *in silico* se realizaron utilizando el software *Serial Cloner* (versión 2.6.1, Serial Basics).

### ***Clonación en los vectores de destino pSPYNE y pSPYCE.***

Los vectores de destino que se utilizaron para el ensayo *BiFC* son los pSPYNE y pSPYCE, éstos son vectores binarios que permiten la visualización de interacciones entre proteínas en células vegetales vivas. Ambos vectores permiten la expresión de los genes de interés fusionados al extremo N-terminal (pSPYNE) o al extremo C-terminal (pSPYCE) de la Proteína Amarilla Fluorescente (YFP por sus siglas en

inglés) (Walter *et al.*, 2004).

### ***Linearización del inserto-attL por PCR***

Debido a que el vector de entrada de las construcciones generadas posee el gen de resistencia a Kanamicina y los vectores de destino pSPYCE y pSPYNE presentan la misma resistencia, fue necesario eliminarla del vector de entrada para reducir el número de falsos positivos durante la recombinación LR. Para ello, se realizó PCR de cada una de las construcciones utilizando los oligonucleótidos para el M13, sitios que flanquean la región del inserto y mantienen los sitios de recombinación attL, necesarios para la recombinación LR, intactos.

Los componentes de las reacciones de PCR fueron los siguientes: 4  $\mu$ L de Buffer HF 5x, 0.4  $\mu$ L de dNTPs, 1  $\mu$ L del oligonucleótido M13 sentido y 1  $\mu$ L oligonucleótido M13 antisentido a una concentración de 0.5 pmoles/ $\mu$ L, 0.2  $\mu$ L de Phusion, 12.4  $\mu$ L de agua libre de nucleasas y 1  $\mu$ L de templado (aproximadamente 20 ng/ $\mu$ L).

El tamaño de los productos de PCR se verificó mediante electroforesis en gel de agarosa al 0.8% teñido con Bromuro de etidio. Asimismo, los productos de PCR se purificaron directamente del gel utilizando el kit *QIAquick Gel Extraction* (Qiagen, Cat. No. 28706) (Apéndice 10), se cuantificaron por espectrofotometría y se verificaron nuevamente por electroforesis.

### ***Reacción de recombinación LR***

Con los productos de PCR-attL purificados se llevó a cabo la reacción de recombinación LR utilizando la *LR Clonase II Enzyme Mix* (Invitrogen, Cat. No. 11791) siguiendo las instrucciones del fabricante. En la tabla 3 se muestra el vector de entrada y el vector de destino entre los que se llevaron a cabo las reacciones LR. La Figura 10 muestra un diagrama de las construcciones generadas.

**Tabla 3. Listado de los vectores de entrada y los vectores de destino con los que se realizó la reacción de recombinación LR**

<b>Vector de entrada</b>	<b>Vector de destino</b>	<b>Construcción final</b>
pENTR-MpAGO9	pSPYNE	MpAGO9:nYFP
pENTR-AtAGO4	pSPYNE	AtAGO4:nYFP
pENTR-AtNRPE1	pSPYCE	AtNRPE1:cYFP



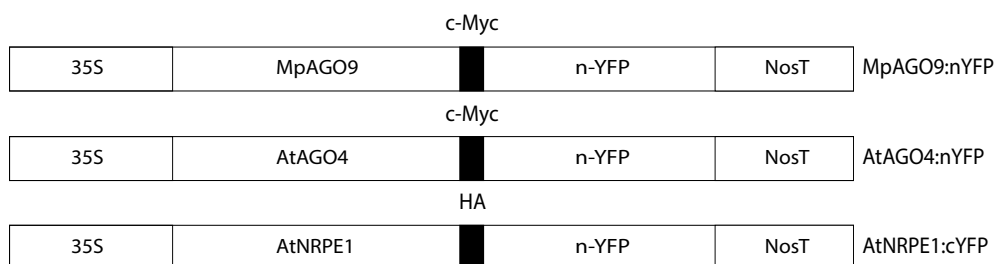


Figura 10. Construcciones generadas mediante la reacción de recombinación LR. En la parte superior MpAGO9:nYFP, seguida de AtAGO4:nYFP, en la parte inferior AtNRPE1:cYFP.

Los componentes de cada una de las reacciones de recombinación fueron: 150 ng de producto de PCR-attL, 150 ng de vector de destino y el volumen de buffer TE pH 8 necesario para alcanzar 8  $\mu$ L. Finalmente, se adicionaron 2  $\mu$ L de *LR Clonase II Enzyme Mix*. Las reacciones se incubaron a 25°C durante toda la noche. Al día siguiente las reacciones se detuvieron adicionando 1  $\mu$ L de Proteinase K y se incubaron a 37°C durante 10 minutos

### ***Transformación de E. coli por electroporación***

Para la electroporación se utilizaron 50  $\mu$ L de células electrocompetentes de la cepa DH5 $\alpha$  a las que, una vez descongeladas, se les adicionaron 3-4  $\mu$ L de producto de la reacción LR previamente dializado. El protocolo y las condiciones de electroporación fueron las mismas mencionadas anteriormente. Finalmente, las bacterias se sembraron en placas Petri con medio de selección LB (Luria-Bertani) sólido con Kanamicina (50  $\mu$ g/mL) y se incubaron a 37°C durante 12 horas.

### ***Escrutinio de clonas transformantes***

Para corroborar la eficacia de la reacción de recombinación LR se realizó PCR en colonia. La PCR se llevó a cabo utilizando el *GoTaq Green Master Mix* (Promega, Cat. No. M7122) y los oligonucleótidos sentido de cada uno de los genes clonados y el antisentido de la YFP, ya sea para el extremo N o C-terminal según fuese el vector de destino. Los tamaños de los productos de PCR se verificaron mediante electroforesis en un gel de agarosa al 0.8%. Se utilizó el marcador molecular de *O'GeneRuler 1kb DNA Ladder Plus* (Fermentas, # SM1343).

### ***Transformación por electroporación de Agrobacterium***

Una vez que se verificaron las clonas positivas en *E. coli*, se seleccionó una sola clona de cada una de las construcciones y se extrajo plásmido por lisis alcalina. De manera independiente se transformaron

bacterias con cada una de las construcciones generadas adicionando 2-3  $\mu\text{L}$  de plásmido purificado. Para la electroporación se utilizaron 50  $\mu\text{L}$  de células electrocompetentes de la cepa C58. Las células se transfirieron a una celda de electroporación y al ponerse en contacto con el electrodo, se dio un pulso de 2.2 kV durante 5 ms. Posteriormente se adicionó 1 mL de medio LB líquido y se incubó en agitación por 3 horas a 30°C. Una vez transcurrido el período de incubación, las bacterias se sembraron en medio LB sólido con Ampicilina 100  $\mu\text{g}/\text{mL}$ , Kanamicina 50  $\mu\text{g}/\text{mL}$  y Rifampicina 100  $\mu\text{g}/\text{mL}$  y se incubaron nuevamente a 30° durante ~2 días.

### ***Agroinfiltración en *Nicotiana benthamiana****

En cada caso, a partir de una sola colonia, se preparó un precultivo de cada línea transgénica (MpAGO9:nYFP, AtAGO4:nYFP, AtNRPE1:cYFP, incluyendo también las construcciones para el control positivo, pBIN19:GFP y la línea auxiliar p19) en 1 mL de medio líquido LB con Ampicilina 100  $\mu\text{g}/\text{mL}$ , Kanamicina 50  $\mu\text{g}/\text{mL}$  y Rifampicina 100  $\mu\text{g}/\text{mL}$  y se incubó toda la noche a 30°C en agitación. Al día siguiente se inocularon con el precultivo, de manera independiente, 50 mL de medio líquido LB con los mismos antibióticos y se incubaron toda la noche a 30°C en agitación.

Transcurrido el período de incubación, los cultivos se colocaron en tubos cónicos de 50 mL y se centrifugaron a 8000 rpm durante 10 minutos a temperatura ambiente, se descartó el sobrenadante y el pellet de cada línea transgénica se resuspendió en 5 mL de  $\text{MgCl}_2$  10 mM. Con la solución de bacterias se realizaron diluciones 1:20 de cada cultivo para medir por triplicado su absorbancia a 600 nm. Se ajustaron los volúmenes de cada cultivo para obtener una absorbancia final con un valor de 1 en 5 mL de  $\text{MgCl}_2$  10 mM. La cantidad de cultivo necesaria para el ajuste se determinó mediante la siguiente fórmula:

$$V_{ix} = \frac{V_2 \times A_n}{A_{20x} \times Fd}$$

Donde  $V_2$  corresponde al volumen final (5 mL),  $A_n$  corresponde a la absorbancia final (1),  $A_{20x}$  corresponde a la media aritmética de los valores obtenidos de absorbancia durante las mediciones por triplicado y  $Fd$  al factor de dilución (20). Habiendo ajustado los volúmenes de los cultivos se realizaron las mezclas mostradas en la Tabla 4. Cada mezcla se llevó a un volumen final de 5 mL con  $\text{MgCl}_2$  10 mM y se le adicionaron 5  $\mu\text{L}$  de acetosiringona 1000 x (50  $\mu\text{g}/\text{mL}$ ). Las mezclas se incubaron durante 3 horas a temperatura ambiente.

**Tabla 4. Mezclas realizadas para la agroinfiltración.**

<b>Construcciones</b>	<b>Como:</b>
SC:cYFP + EMB:nYFP + p19	Control positivo del sistema
AtNRPE1:cYFP + AtAGO4:nYFP + p19	Control positivo de interacción
AtNRPE1:cYFP + MpAGO9:nYFP + p19	Prueba
pBIN19:GFP + p19	Control de transformación
p19	Mock

Habiendo transcurrido el período de incubación, hojas de plantas de *Nicotiana benthamiana* de entre 3 a 4 semanas fueron infiltradas por el envés, utilizando una jeringa sin aguja. Las plantas se mantuvieron en una cámara de crecimiento por al menos 48 horas después de la agroinfiltración con fotoperíodo 15 horas luz/9 horas oscuridad.

### ***Preparación de protoplastos de N. benthamiana***

Las hojas de tabaco previamente agroinfiltradas (72 horas después de la agroinfiltración) se cortaron en pequeñas tiras con una navaja nueva. Las hojas se transfirieron a una placa de Petri y se le adicionaron 3 mL de solución enzimática (Manitol 0.5 M, Celulasa 1%, Macerozima 0.05%), posteriormente se incubaron por 3 horas en agitación leve, a temperatura ambiente y en oscuridad. Finalmente los protoplastos se recuperaron y se trasladaron a un tubo nuevo en donde permanecieron hasta su observación. Los protoplastos fueron visualizados con el Microscopio confocal Olympus FV1000 Multi-fotónico invertido con un láser de longitud de onda de excitación de 514 nm y un filtro de emisión de 527 nm. En el caso de la construcción de GFP, la longitud de onda de excitación fue de 490 nm y un filtro de emisión de 509 nm.

### ***Confirmación de la expresión de los transgenes y acumulación de proteína***

Se colectaron muestras de las hojas agroinfiltradas para RNA total por el método mencionado en el apartado II. Con el RNA total, se realizaron ensayos de RT-PCR para verificar la expresión de los transgenes. La síntesis de cDNA se llevó a cabo bajo las mismas condiciones mencionadas en el apartado II. La PCR se llevó a cabo con el KAPA 3G Plant PCR kit (KAPA Biosystems) siguiendo las instrucciones del fabricante, utilizando los oligonucleótidos sentido específicos para cada transgén y los oligonucleótidos antisentido para la fusión C o N-terminal de la YFP, según fuera el caso. Los productos de PCR se verificaron por electroforesis en agarosa al 0.8% teñida con Bromuro de Etidio.

Por otra parte, para confirmar la presencia de las proteínas de interés, se realizaron ensayos de tipo Western blot utilizando extractos proteicos de las hojas agroinfiltradas. Los anticuerpos utilizados

durante el Western blot corresponden a los epítomos específicos presentes en los vectores para el ensayo *BiFC*: la construcción pSPYNE presenta un epítome c-Myc, mientras que la construcción pSPYCE presenta un epítomo HA. Por tanto, las proteínas MpAGO9 y AtAGO4 tendrían el epítomo c-Myc, mientras que AtNRPE1 presentaría el epítomo HA (Figura 10).

Se realizó una extracción de proteína total a partir de las hojas agroinfiltradas (10 mM Tris-HCl pH 7.5, 1 mM MgCl<sub>2</sub>, 25 mM NaEDTA·2H<sub>2</sub>O pH 8, 0.25 mM Sacarosa, 7.5 mM DTT, 0.1% Tritón x-100), se incluyó un control negativo de plantas no agroinfiltradas, mientras que como control positivo se colectaron tejidos de hojas agroinfiltradas con las construcciones CP:nYFP + EMB:cYFP. Se cuantificó la proteína por espectrofotometría utilizando el reactivo de Bradford. Una vez cuantificada la proteína total, las muestras se separaron electroforéticamente en un gel SDS-PAGE al 10% durante 2 horas y media a 120 V a 4°C. Posteriormente, se realizó la transferencia de las proteínas del gel a una membrana de nitrocelulosa utilizando una cámara de transferencia húmeda (2 horas a 350 mA). Una vez realizada la transferencia, la membrana se tiñó con rojo de Ponceau para verificar el éxito en la transferencia. Posteriormente, se llevó a cabo el bloqueo de la membrana incubando con la solución de bloqueo (Leche 0.5%, PBS 1x, Tritón 0.5%) durante toda la noche en agitación moderada y a 4°C. Acto seguido, se lavó la membrana con PBS 1x/Tritón 0.5% y se incubó con el anticuerpo primario ( $\alpha$ -c-Myc, en ratón, dilución 1:13,000 o  $\alpha$ -HA, en conejo, dilución 1:1000, según fuera el caso) durante toda la noche a 4°C en agitación moderada. La membrana se lavó bajo las mismas condiciones mencionadas anteriormente y en seguida se incubó con el anticuerpo secundario (anti-mouse-HRP para  $\alpha$ -c-Myc y anti-rabbit-HRP para  $\alpha$ -HA, ambos a una dilución 1:5,000) durante 30 minutos. La membrana se lavó nuevamente y finalmente se reveló utilizando el Pierce ECL Western Blotting Substrate (Thermo Scientific) siguiendo las instrucciones del fabricante.

### ***Búsqueda e identificación de posibles interactores.***

Durante la vía *RdDM* en *A. thaliana*, AGO4 establece interacciones con otras proteínas como la subunidad grande RNA polimerasa V (NRPE1; AT2G40030), KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1; AT5G04290.1) y RNA-DIRECTED METHYLATION 1 (RDM1; AT3G22680) para dirigir el silenciamiento a un *locus* específico. Con base en lo anterior, se decidió buscar si existe en *M. polymorpha* algún homólogo a alguno de estos genes.

Se llevó a cabo una búsqueda bioinformática en bases de datos de transcriptomas, así como en el genoma de *M. polymorpha* utilizando el algoritmo tblastn y como secuencias de referencia los genes NRPE1, KTF1 y RDM1 de *A. thaliana* disponibles en la base de datos TAIR ([www.arabidopsis.org](http://www.arabidopsis.org)).

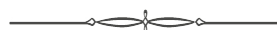
### ***Caracterización del gen MpNRPE1***

El transcrito que mostró un mayor porcentaje de identidad con *NRPE1* de *A. thaliana* se leyó en sus seis marcos abiertos de lectura, con lo que se determinó el sitio de inicio y término de la traducción. Con la secuencia de aminoácidos de *MpNRPE1* se realizó una búsqueda de dominios conservados (utilizando la base de datos CDD, mencionada anteriormente), así como la estructura cristalizada de la RNA polimerasa II de levadura para identificar los dominios conservados en la región N-terminal (Cramer *et al.*, 2001) y se determinó la distribución de éstos.

### ***Identificación de dominios WG/GW***

Se ha determinado que la presencia de reiteraciones de residuos de Glicina-Triptófano (WG/GW) en las proteínas (también conocidos como *Ago-hooks*) es crucial para que éstas puedan establecer interacciones con proteínas AGO. Debido a lo anterior, se decidió buscar si en *MpNRPE1* existen reiteraciones de WG/GW. Mediante la herramienta en línea *Agos – a web tool for WG/GW Ago – binding sites identification* (Zielezinski & Karlowski, 2011; <http://comgen.pl/agos/submit/>) se identificaron el número y la posición de dichos repetidos WG/GW en *MpNRPE1*.

# RESULTADOS



La función de las proteínas puede conocerse mediante diversas aproximaciones. El presente estudio aborda dicha cuestión desde cuatro aspectos fundamentales: I) la conservación evolutiva de la secuencia, II) la conservación evolutiva de la estructura proteica y arquitectura de dominios, III) la determinación del perfil de expresión y IV) la interacción con otras proteínas. Estos cuatro aspectos permiten inferir la función de las proteínas, particularmente, este estudio se enfocó en la búsqueda e identificación de un gen ortólogo a *AGO9* de *A. thaliana* en *Marchantia polymorpha*. A continuación se exponen los resultados obtenidos.

## **I. Conservación evolutiva de la secuencia: Identificación del gen homólogo a *AGO9* en *M. polymorpha*.**

Mediante la búsqueda bioinformática realizada se identificaron los miembros de la familia *AGO* en *M. polymorpha*: está conformada por cinco miembros, mostrando al menos un representante en cada clado (Figura 11). La familia *AGO* en *A. thaliana* se agrupa en tres clados principales según su identidad: El clado I, conformado por *AGO1/AGO5/AGO10*, el clado II, conformado por *AGO7/AGO2/AGO3* y el clado III, cuyos miembros son *AGO4/AGO6/AGO8/AGO9*, este último asociado con el silenciamiento a nivel transcripcional (Vaucheret, 2008; Zheng *et al.*, 2015). En el caso del clado III, se identificaron tres miembros, uno de ellos es un claro ortólogo de *AGO9* de *A. thaliana* mostrando un porcentaje de identidad del 43.82%. A este transcrito se le denominó *AGO10998* con base en el número de registro en el transcriptoma de *M. polymorpha*. Asimismo, con la secuencia de dicho transcrito se realizaron reconstrucciones filogenéticas que permitieron confirmar su identidad. Los resultados mostraron que, efectivamente, el gen identificado en *M. polymorpha* se agrupa con el clado *AGO4/AGO6/AGO8/AGO9* (Figura 11).

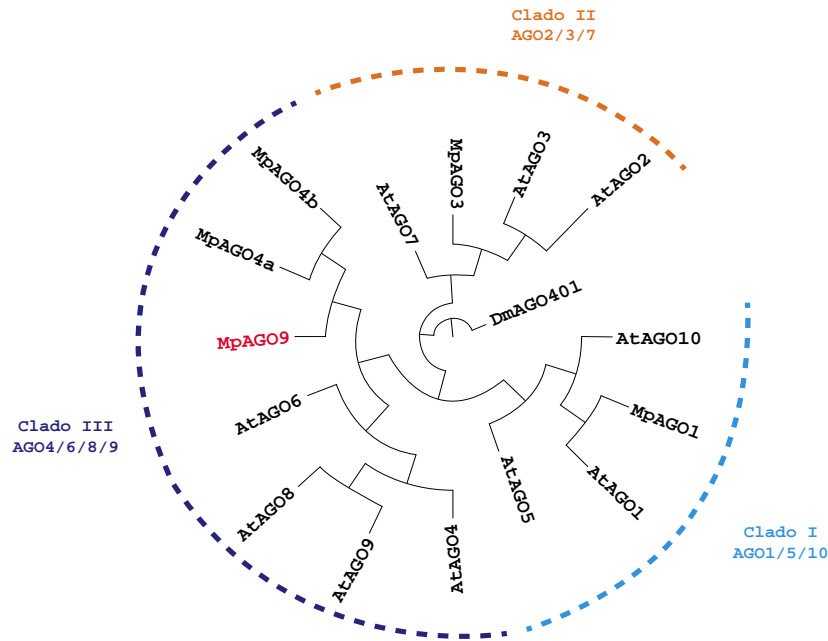


Figura 11. Reconstrucción filogenética de la familia AGO en *A. thaliana* y *M. polymorpha*. Se incluyen todos los miembros de la familia AGO de *A. thaliana* y los miembros putativos de la familia AGO en *M. polymorpha*. Se señalan los tres clados de la familia AGO y MpAGO9 se resalta en color rojo. DmAGO401 corresponde a un miembro de la familia AGO en *Drosophila melanogaster*.

Con la información obtenida se logró determinar la estructura del gen AGO10998, el que se renombró como MpAGO9 siguiendo el sistema de nomenclatura propuesto por Zhang y colaboradores (2015). El gen MpAGO9 tiene una longitud de 5470 pb, posee 19 exones y 18 intrones; la región codificante tiene un tamaño de 2433 pb y codifica una proteína de 810 aa. La longitud del gen, la composición de intrones y exones y el tamaño de la proteína que codifica MpAGO9 es similar a la del gen AGO9 de *A. thaliana* (Tabla 5, Figura 12).

Tabla 5. Comparación de las características de los genes MpAGO9 y AtAGO9.

Identificador	Tamaño del gen	No. Exones	No. Intrones	CDS	Proteína
MpAGO9	5470 pb	19	18	2433 pb	810 aa
AtAGO9	6086 pb	23	22	2691 pb	890 aa

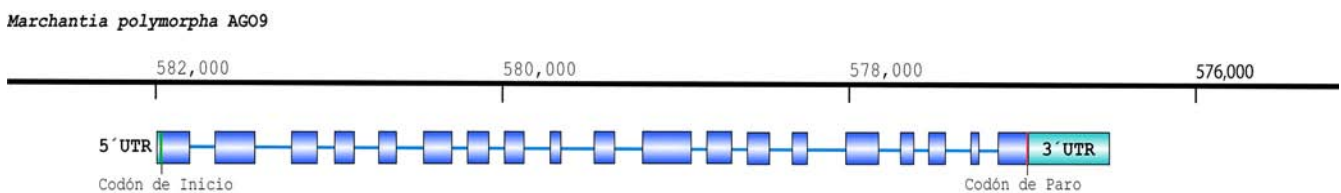


Figura 12. Estructura génica de MpAGO9. Se muestran como cajas los exones, mientras que los intrones se muestran como líneas. Los extremos 3 y 5'UTR se muestran en azul claro, asimismo, se indica el sitio de inicio (en verde) y el sitio de término (en rojo) de la traducción. La línea negra en la parte superior corresponde a la posición en el genoma de *M. polymorpha* en donde se encuentra MpAGO9.

Por otra parte, la reconstrucción filogenética de la Figura 13 ubica a *MpAGO9* en el clado de las briofitas, junto a su ortólogo en *Physcomitrella patens*. Dicho resultado sugiere que existe un alto grado de conservación evolutiva a nivel de secuencia de este gen en todas las plantas terrestres y más aún, que el origen de este gen se remonta a un momento en la historia previo a la divergencia de las embriofitas.

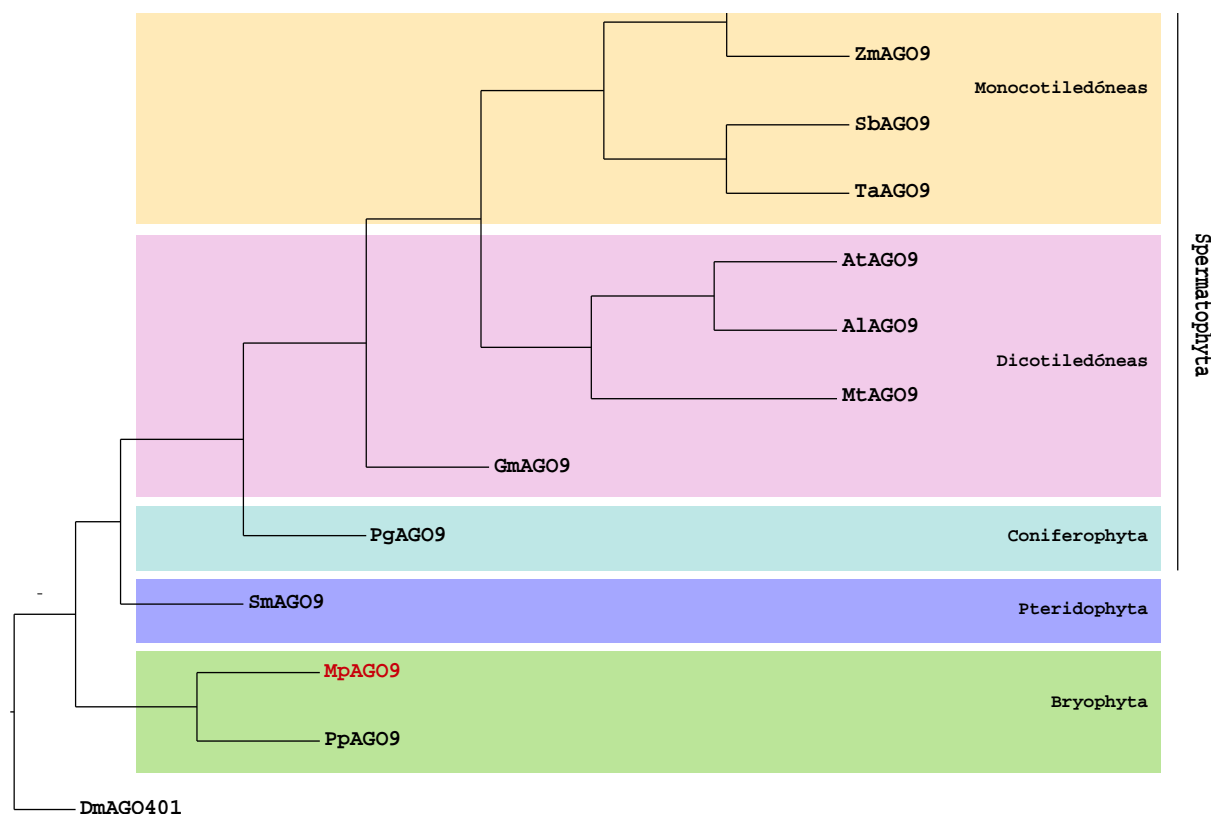


Figura 13. Reconstrucción filogenética de los ortólogos de AGO9 en las embriofitas. Se incluyen representantes de diferentes grupos de embriofitas. *MpAGO9* se resalta en color rojo, agrupado con su ortólogo en *P. patens*.

## II. Conservación de la estructura proteica: Distribución de dominios y estructura tridimensional de la proteína *MpAGO9*

En los resultados previos se mostró que *MpAGO9* está altamente conservado a nivel de secuencia. La siguiente cuestión a responder es si éste presenta conservación evolutiva a nivel estructural. Para poder responder lo anterior, se realizó una búsqueda de dominios conservados en la secuencia de aminoácidos de *MpAGO9*.

En todos los eucariontes, la familia AGO se caracteriza por la presencia de los dominios PAZ, PIWI, L1, L2 y MID (Swarts *et al.*, 2014). Los resultados del análisis de dominios muestran que *MpAGO9*



presenta un alto grado de conservación evolutiva de dichos dominios y que la arquitectura de éstos es similar a la de AtAGO9 (Figura 14).

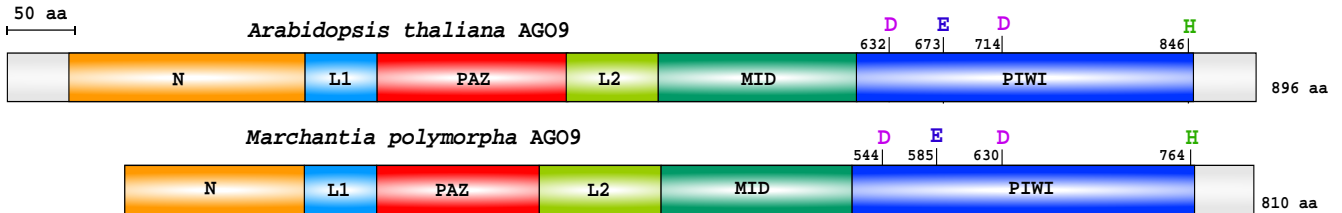
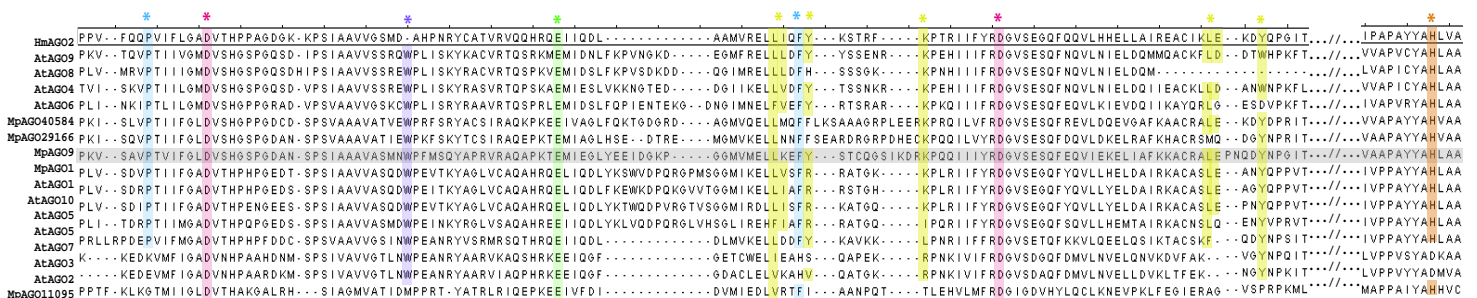


Figura 14. Comparación de la distribución de dominios entre *AtAGO9* y *MpAGO9*. En el dominio PIWI se resaltan la posición de los residuos DDH que conforman la tríada catalítica presente en todas las proteínas AGO, así como el residuo adicional de E según el modelo de Nakanishi y colaboradores (2012).

Asimismo, *MpAGO9* conserva residuos importantes para su función, como son los residuos de DDH que conforman la triada catalítica de las proteínas AGO, así como también un grupo de residuos de aminoácidos importantes para el establecimiento de interacciones con proteínas WG/GW, todos ellos distribuidos en el dominio PIWI (Figura 15). Con respecto al modelo de la tétrada catalítica propuesto por Nakanishi y colaboradores (2012), *MpAGO9*, así como todas las AGO de *A. thaliana*, conservan este residuo de Glutamato (Figura 14,15).

### Dominio PIWI



\* Trp pocket "902" \* Trp pocket "902" \* Inserción de Trp \* His Residuo catalítico \* Asp Residuo catalítico  
\* Glu Residuo catalítico

Figura 15. Alineamiento de las secuencias de aminoácidos de las proteínas AGO de *A. thaliana* y *M. polymorpha*. Se muestran con \* las posiciones de los residuos conservados necesarios para la función de las mismas. Se resalta en color gris la secuencia correspondiente a *MpAGO9*. Asimismo, se incluyó la secuencia de la proteína AGO2 humana para identificar los residuos que corresponden a los sitios de unión a Triptofano.

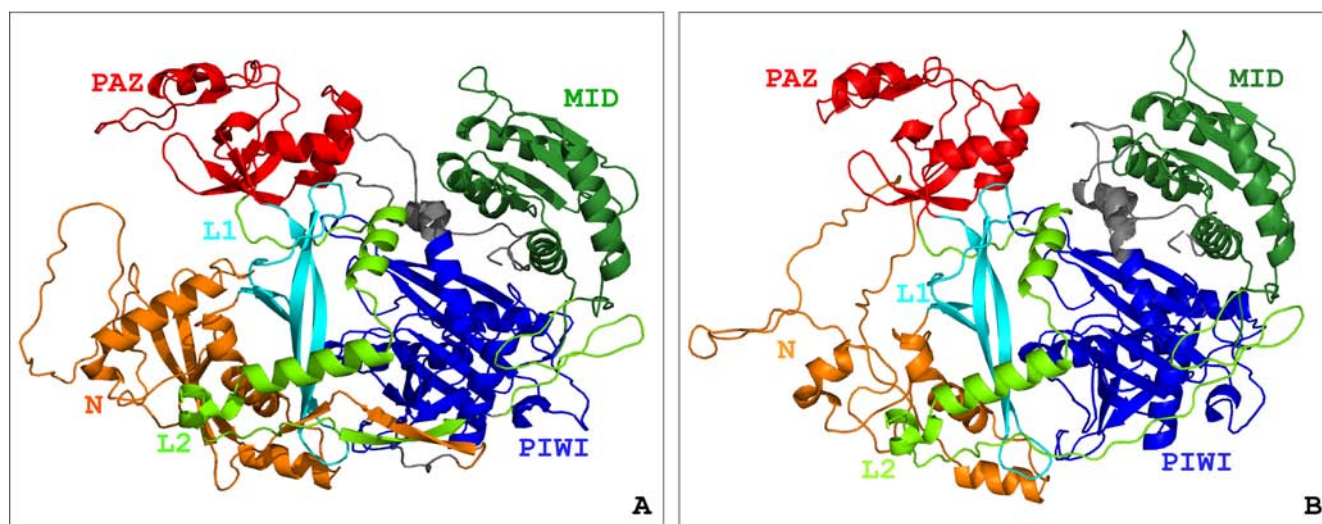


Figura 16. Estructura tridimensional de las proteínas A) AtAGO9 y B) MpAGO9. Los dominios se muestran por colores. Los modelos tridimensionales de ambas proteínas fueron generados mediante homología estructural con la proteína AGO cristalizada de *A. aeolicus*. Los colores en A y B corresponden a los dominios descritos en la Figura 14.

El modelo predicho para la estructura terciaria de la proteína MpAGO9 muestra la estructura bilobular característica de las proteínas AGO en donde los dominios N, L1, y PAZ conforman un lóbulo interconectado por el dominio L2 al lóbulo que conforman los dominios MID y PIWI (Figura 16.B). Debido a que no ha sido posible determinar hasta el momento la estructura tridimensional de alguna proteína AGO de planta, los modelos de AtAGO9 y MpAGO9 se construyeron a partir de homología estructural con la proteína AGO cristalizada de *A. eolicus* (Yuan *et al.*, 2005). Los modelos tridimensionales de las proteínas muestran cierto grado de similitud, mostrando únicamente diferencias notorias en el dominio N terminal (Figura 16).

### III. Determinación del perfil de expresión: Dominios de expresión de *MpAGO9* durante el ciclo de vida de *M. polymorpha*.

Los resultados mostrados de los análisis bioinformáticos hasta el momento sugieren que existe un alto grado de conservación evolutiva de MpAGO9 tanto a nivel de secuencia como a nivel estructural. Por tanto, la siguiente cuestión es determinar si MpAGO9 muestra conservación de la función.

Se ha reportado que en el clado *AGO4/AGO6/AGO9* de *A. thaliana*, tanto los dominios de expresión como la preferencia de unión por siRNAs provenientes de determinados *loci*, han jugado un papel esencial en su diversificación funcional (Havecker *et al.*, 2010). Bajo la premisa anterior, se decidió monitorear la expresión de *MpAGO9* durante diferentes etapas del desarrollo en la generación gametofítica de *M. polymorpha*.

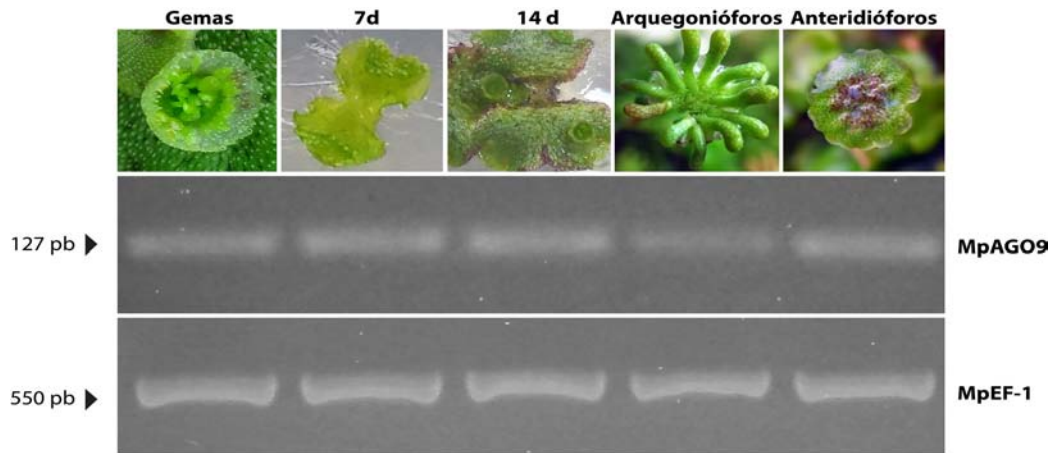


Figura 17. Perfil de expresión de *MpAGO9* durante el ciclo de vida de *M. polymorpha*. Se muestra que *MpAGO9* muestra niveles de expresión similares en gemas, talos de 7 y 14 días, arquegonióforos y anteridióforos.

Las etapas contempladas para el análisis fueron las estructuras de reproducción asexual (gemas), talos de 7 y 14 días, las estructuras reproductivas femeninas (arquegonióforos) y las masculinas (anteridióforos). Los resultados mostraron que no existe una variación evidente en los niveles de expresión en las diferentes etapas del ciclo de vida de *M. polymorpha* (Figura 17). En este caso, *MpAGO9* muestra dominios de expresión diversificados durante el desarrollo gametofítico, a diferencia de su ortólogo en *A. thaliana*, en donde su expresión se encuentra restringida en el meristemo apical del embrión y tejidos reproductivos femeninos (Havecker *et al.*, 2010; Olmedo-Monfil *et al.*, 2010).

#### IV. Interacción proteína-proteína: Ensayo de Complementación por Fluorescencia Bimolecular (*BiFC*)

Una forma de conocer la función de una proteína consiste en la identificación de sus proteínas interactoras. En el caso específico de las proteínas AGO, se ha observado que son capaces de establecer interacciones con un grupo de proteínas que presentan motivos WG/GW o *Ago hooks*. Asimismo, se ha observado que muchas de estas proteínas juegan papeles fundamentales durante la vía de Metilación del DNA dependiente de RNA (*RdDM*).

El ensayo de Complementación por Fluorescencia Bimolecular (*BiFC*) es una técnica ampliamente utilizada para la detección de interacciones de proteínas basada en la asociación de dos fragmentos de una proteína fluorescente. Los fragmentos (N y C -terminal, no fluorescentes) se fusionan a las proteínas de interés de manera que cuando éstas interaccionan, provocan la reconstitución de la fluorescencia haciendo posible la visualización de esta interacción *in vivo* (Kerppola, 2009; Kodama & Hu, 2012).

Para averiguar si MpAGO9 es capaz de establecer interacciones con proteínas con motivos WG/GW, como los análisis *in silico* sugieren, se realizaron construcciones utilizando el sistema de vectores pSPYCE/pSPYCE que permite llevar a cabo el ensayo *BiFC*. Mediante este ensayo se probó la interacción entre MpAGO9 y AtNRPE1.

Los genes de interés se clonaron en el vector de entrada pDONR221 mediante el sistema de recombinación Gateway. Las clonas positivas se identificaron mediante digestiones con enzimas de restricción, los resultados y los mapas de las construcciones se muestran en la Figura 18. Una vez verificadas las clonas, se llevó a cabo la recombinación LR entre los vectores de entrada y los vectores de destino. Las clonas positivas se verificaron mediante PCR en colonia (Figura 19). Debido a dificultades técnicas, las clonas de AtNRPE1:cYFP se verificaron utilizando el par de oligonucleótidos sentido y antisentido de este gen. Asimismo, para verificar que las construcciones después de la recombinación mantuvieran en fase la fusión traduccional, se enviaron a secuenciar en la Unidad de Síntesis y Secuenciación del Instituto de Biotecnología, UNAM. Los resultados de la secuenciación mostraron que las fusiones traduccionales se encontraban en fase (Anexo 2).

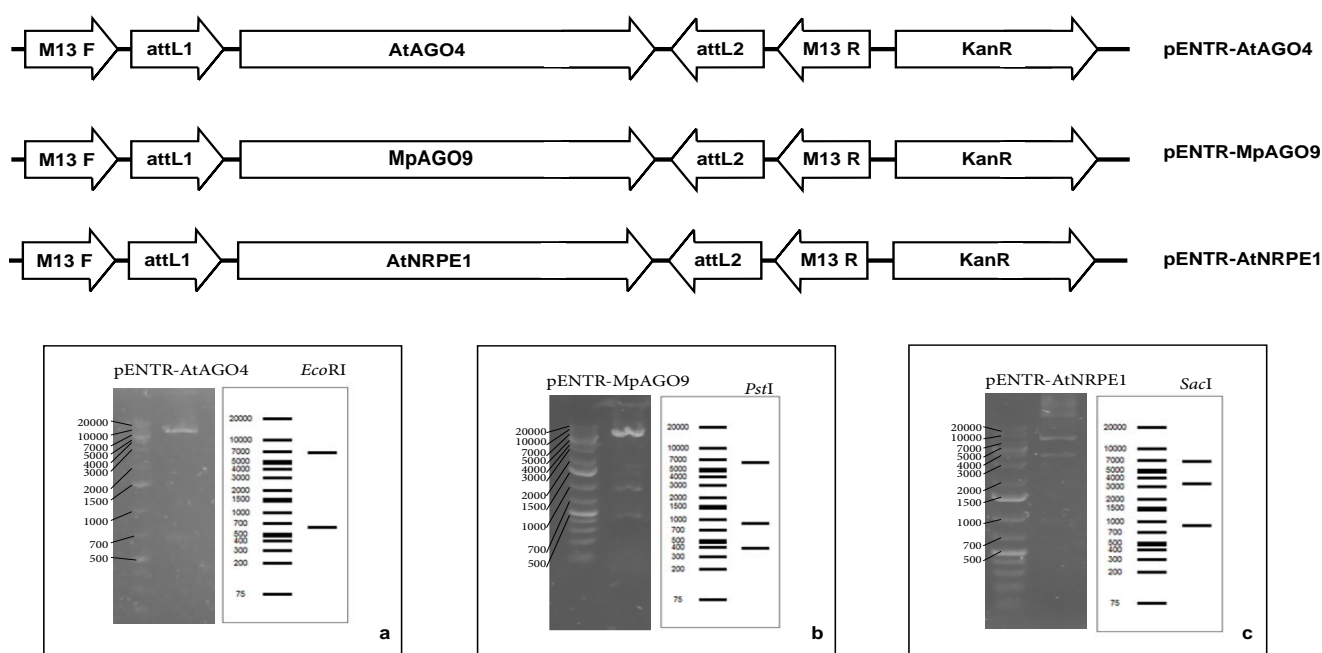


Figura 18. Digestiones de las construcciones de entrada al sistema Gateway. A) Mapa de la construcción de entrada pENTR-AtAGO4 y restricción con la enzima *EcoRI*, B) Mapa de la construcción pENTR-AtAGO4 y digestión con la enzima de restricción *PstI*, C) Mapa de la construcción pENTR-AtNRPE1 y digestión con la enzima de restricción *SacI*.

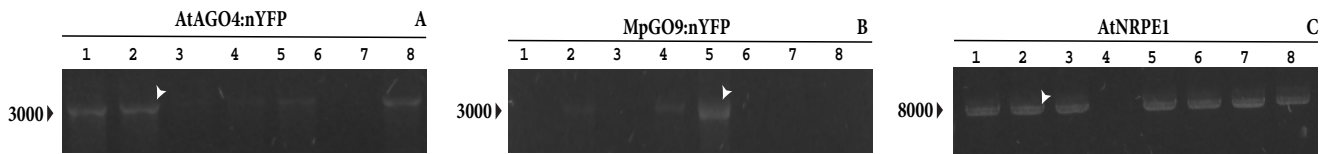


Figura 19. Verificación de la transformación con los vectores de destino mediante PCR en colonia. A) Clonas analizadas transformadas con la construcción AtAGO4:nYFP, B) Clonas transformadas con MpAGO9:nYFP, C) Clonas transformadas con la construcción AtNRPE1:cYFP. Las clonas marcadas con una flecha blanca fueron las seleccionadas para continuar el ensayo de interacción. En total se analizaron 8 clonas de cada construcción.

Una vez que se verificaron las construcciones se llevó a cabo el ensayo de interacción *BiFC* en tabaco. El control positivo del sistema fue el par de construcciones de proteínas cloroplásticas, que fueron proporcionadas y probadas con anterioridad por el M. en C. Luis de Luna y el Biol. Carlos Villaseñor (datos no publicados). Asimismo, se llevó a cabo un tratamiento *Mock* en donde se agroinfiltró únicamente con p19, un potenciador de la expresión de transgenes. La observaciones de los protoplastos a las 48, 72 y 96 horas posteriores a la agroinfiltración con las construcciones AtAGO4:nYFP + AtNRPE1:cYFP, así como MpAGO9:nYFP + AtNRPE1:cYFP no mostraron reconstitución de la fluorescencia de la YFP (Figura 20).

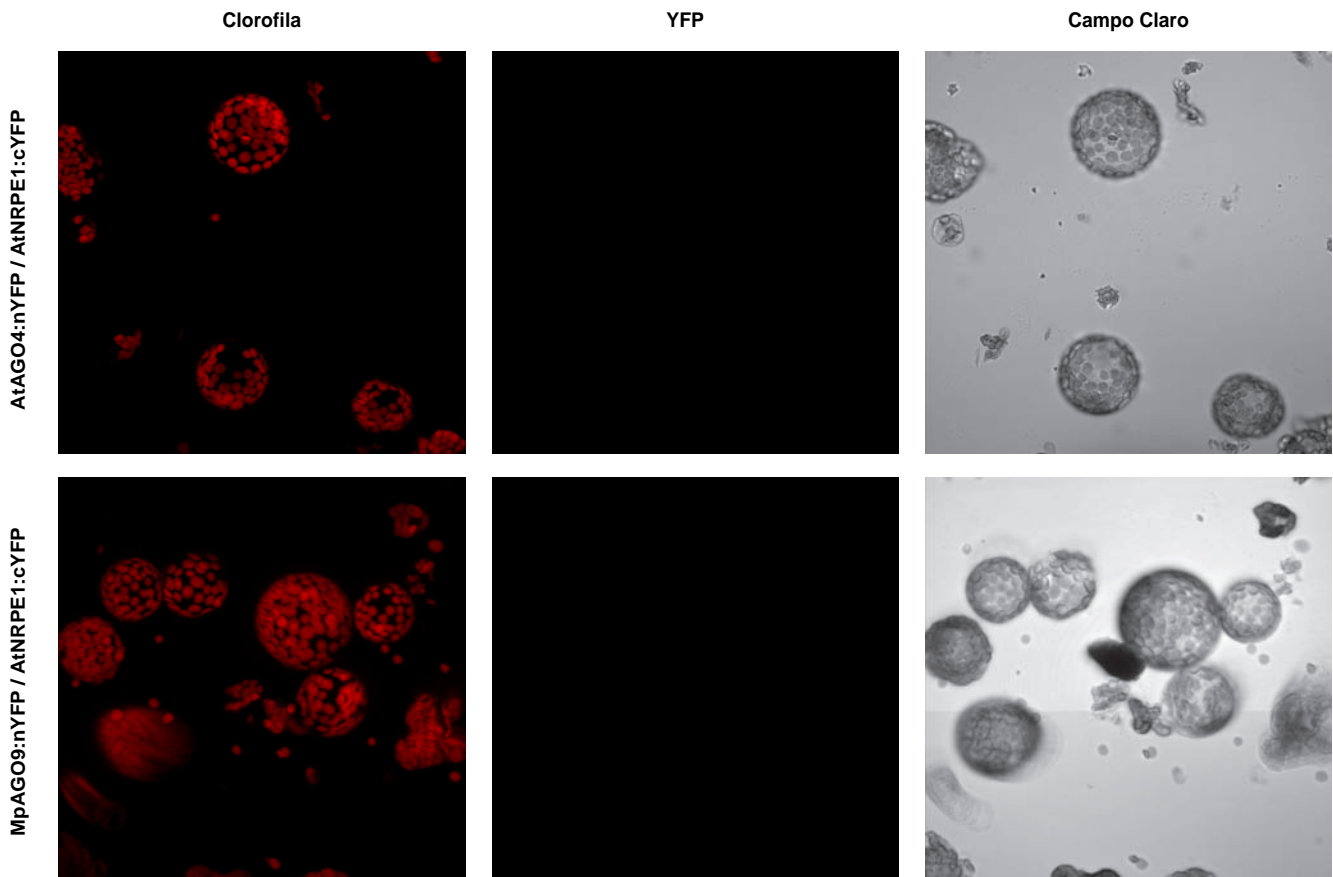


Figura 20. Protoplastos agroinfiltrados de *N. benthamiana*. Se muestran los protoplastos obtenidos de tejidos co-infiltrados con las construcciones AtAGO4:nYFP/AtNRPE1:cYFP en el panel superior y en el panel inferior los protoplastos obtenidos de la co-infiltración con las construcciones MpAGO9:nYFP/AtNRPE1:cYFP.

Para verificar que la transformación de las hojas de tabaco se hubiese efectuado de manera exitosa, se realizaron ensayos de tipo RT-PCR para evaluar los niveles de expresión de los transgenes en cada uno de los tejidos agroinfiltrados. En el caso de AtNRPE1, debido a su tamaño se diseñó un oligonucleótido interno para disminuir el tiempo de elongación durante la PCR. Asimismo, debido a dificultades técnicas, la PCR se llevó a cabo utilizando el oligonucleótido interno sentido y el oligonucleótido antisentido que se había empleado anteriormente. A pesar de que se observan inespecificidades en el caso de AtNRPE1, se observan claramente las bandas de los tamaños esperados. Por tanto, los resultados confirmaron la presencia de transcritos de cada uno de los transgenes en las plantas agroinfiltradas (Figura 21).

Asimismo, se llevaron a cabo análisis de tipo Western blot para verificar la acumulación de proteínas en las hojas que se agroinfiltraron y descartar el hecho de que la ausencia de proteína estuviese provocando que no se observara reconstitución de la YFP. Los resultados mostraron que en ninguno de los casos había proteína (Figura 22). Los resultados obtenidos durante el ensayo de interacción no permiten concluir si existe o no interacción entre las proteínas de estudio, por tanto, no se descarta el hecho de que éstas puedan establecer interacciones *in vivo*.

Debido a lo anterior, se decidió realizar una búsqueda bioinformática de potenciales interactores codificados en el genoma de *M. polymorpha*. La búsqueda se centró en la identificación de ortólogos que se ha reportado interaccionan con AGO4 en *A. thaliana*. Los resultados de la búsqueda se muestran en la Tabla 6. De los tres posibles candidatos, únicamente se identificó un ortólogo de NRPE1 en *M. polymorpha*. El transcrito con mayor porcentaje de identidad (*isotig00198*, posteriormente renombrado como *MpNRPE1a*) se leyó en sus seis marcos de lectura con los que se determinó el sitio de inicio y término de la traducción. Utilizando la secuencia de aminoácidos de MpNRPE1a se identificaron los dominios conservados, así como la característica peculiar de esta proteína: el número y posición de las reiteraciones WG/GW a lo largo de su dominio C- terminal (CTD).

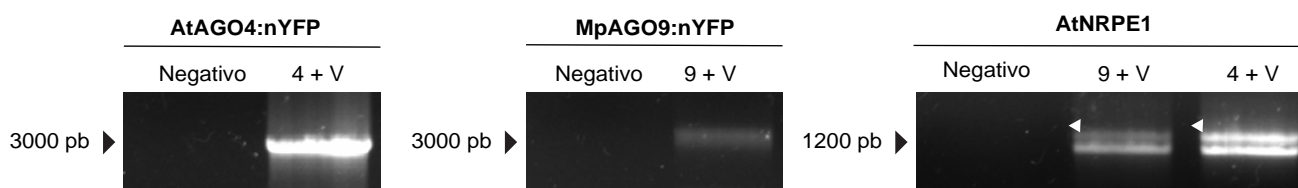


Figura 21. RT-PCR de hojas de tabaco agroinfiltradas. Se muestran los productos de RT-PCR de los tejidos agroinfiltrados. Los carriles marcados como "Negativo" corresponden a hojas sin agroinfiltrar. Los carriles marcados con "9+V" corresponden a hojas co-infiltradas con las construcciones de MpAGO9 y AtNRPE1, mientras que los carriles marcados con "4+V" corresponden a hojas co-infiltradas con AtAGO4 y AtNRPE1. En AtNRPE1 se señalan con flechas blancas los productos del tamaño esperado.

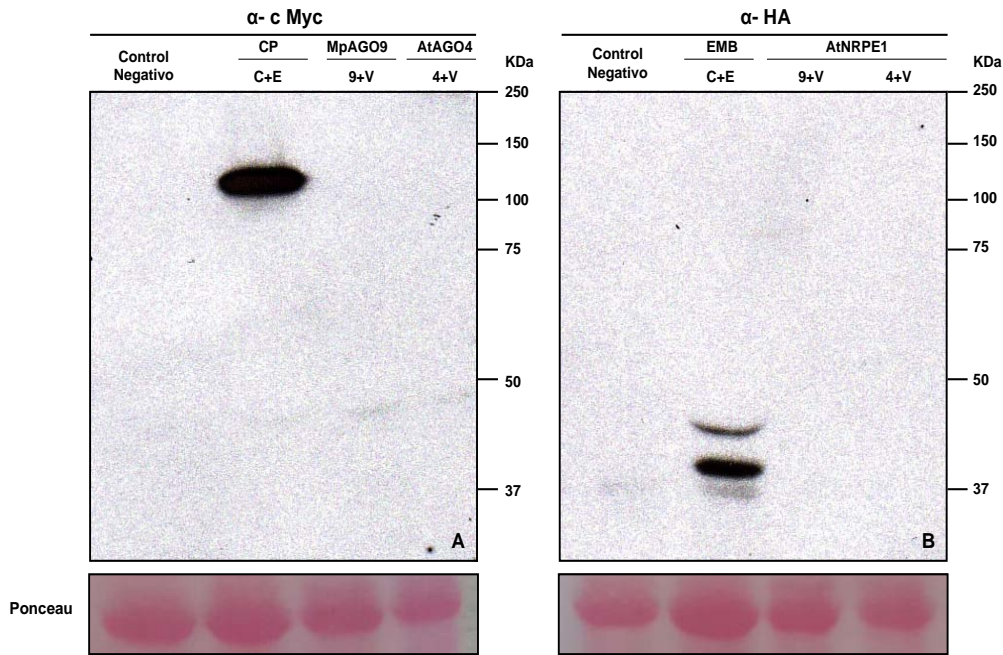


Figura 22. Inmunoblotting de extractos proteicos a partir de los tejidos agroinfiltrados. Se muestran los Western blots para detectar las proteínas de interés en los tejidos agroinfiltrados. Con el epítipo c-Myc se detectaron las proteínas AGO, mientras que con el epítipo HA se detectó NRPE1. Los controles positivos son proteínas cloroplásticas fusionadas a los mismos epítipos. "9+V" corresponde a un extracto proteico obtenido a partir de hojas co-infiltradas con las construcciones MpAGO9 y AtNRPE1; "4+V" corresponde a un extracto proteico obtenido a partir de hojas co-infiltradas con las construcciones AtAGO4 y At NRPE1. El control negativo proviene de hojas sin agroinfiltrar. Como control de carga las membranas se tiñeron con rojo de Ponceau antes de ser hibridadas.

Tabla 6. Proteínas que establecen interacciones con AGO4 durante la RdDM en *A. thaliana*.

Nombre	Identificador	Función reportada	¿Ortólogo en <i>M. polymorpha</i> ?
NUCLEAR RNA POLYMERASE E1 (NRPE1)	AT2G40030	Subunidad grande de la RNA polimerasa V, conforma parte del centro catalítico.	Sí
KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1)	AT5G04290.1	Proteína adaptadora que se une a los transcritos generados por PolV y recluta AGO4 para formar un complejo efector.	No
RNA-DIRECTED METHYLATION 1 (RDM1)	AT3G22680	Componente del complejo DDR. Establece un puente de interacción entre AGO4 y DRM1 durante la RdDM.	No

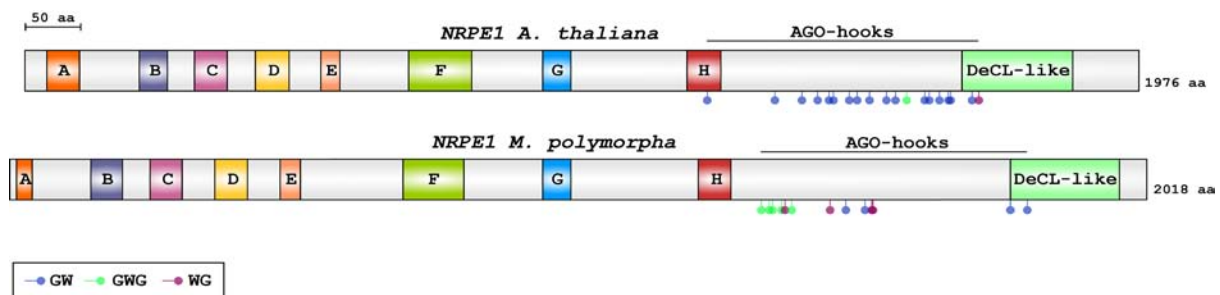


Figura 23. Distribución de los dominios de las proteínas NRPE1 de *A. thaliana* y *M. polymorpha*. Se muestran los dominios conservados entre todas las RNA polimerasas (A hasta H). Se muestra la región característica de la subunidad grande de la RNA polimerasa V, el dominio CTD donde se identifica el dominio *DeCL-like* y los dominios GW o *AgO-hooks* en donde se muestra la posición de los repetidos GW, GWG o WG.

Los resultados obtenidos muestran que MpNRPE1a presenta los dominios característicos que comparte con las demás RNA polimerasas (A-H) y el C-terminal característico de la RNA polimerasa V con el dominio *DeCL-like* y las reiteraciones de GW. MpNRPE1a posee un menor número de repetidos WG/GW y una distribución variable a lo largo del CTD en comparación con AtNRPE1 (Figura 23), teniendo una mayor abundancia de GWG, mientras que AtNRPE1 presenta un número mayor de repeticiones GW. Los resultados muestran que en el genoma de *M. polymorpha* existe un gen que codifica a una proteína que tiene el potencial de establecer interacciones con proteínas AGO.



## DISCUSIÓN



Las proteínas ARGONAUTA se encuentran altamente conservadas en todos los dominios de la vida (Swarts *et al.*, 2014). El número de genes que codifican a AGOs en los diferentes grupos muestra que esta familia de proteínas ha sido sometida a través de la evolución tanto a pérdidas como múltiples eventos de duplicación, principalmente en los genomas de las plantas en donde incluso los grupos más basales como las briofitas presentan más de un miembro —*P. patens* tiene 6 miembros de la familia AGO (Rensing *et al.*, 2008)—.

En este estudio, a través de aproximaciones bioinformáticas, se lograron identificar cinco posibles miembros de la familia AGO codificados en el genoma de *M. polymorpha*. Los tres clados se encuentran representados por al menos un miembro. El clado III o clado 4/6/8/9, relacionado directamente con la vía de Metilación del DNA dependiente de RNA (*RdDM*), fue el clado con un mayor número de representantes en *M. polymorpha*. Los resultados obtenidos concuerdan con el modelo de la evolución de la familia AGO en las plantas terrestres: la familia AGO surgió en un momento previo a la divergencia de las plantas terrestres —pues se lograron identificar miembros de esta familia en el linaje más basal de las embriofitas—, posteriormente, a través de numerosos eventos de duplicación la familia se expandió en los diferentes linajes de plantas y finalmente se especializaron en distintas funciones (Huang *et al.*, 2015; Sinhg *et al.*, 2015; Zhang *et al.*, 2015).

La importancia biológica de la *RdDM* durante la respuesta a estrés y durante la determinación celular en el desarrollo y particularmente su característica heredable, ha resaltado la posibilidad de que esta vía haya contribuido de manera importante en la evolución y diversificación de las plantas terrestres (Xu *et al.*, 2015; Olmedo-Monfil *et al.*, 2010; Matzke & Moshier, 2014). Con base en lo antes mencionado, resulta particularmente interesante el hecho de hallar tres representantes del clado III en el genoma de una especie perteneciente al considerado grupo hermano de las plantas vasculares, *M. polymorpha*.

Este estudio se concentró únicamente en el análisis del posible ortólogo del gen *AGO9* de *A. thaliana* en *M. polymorpha*, debido a que constituye un punto de conexión entre una vía de regulación epigenética y procesos de determinación celular durante el desarrollo, por lo que, pudo haber tenido un papel fundamental durante la evolución de las plantas terrestres.

Una vez que se identificó la secuencia posible ortóloga de *AGO9* en *M. polymorpha*, *AGO10998* al que se nombró como *MpAGO9*, se analizaron minuciosamente las características de la proteína que codifica. Se identificó la arquitectura de los dominios característicos de las proteínas AGO, así como residuos clave en el dominio PIWI. Entre estos residuos, se identificó la posición de los aminoácidos que conforman la tétrada catalítica en *MpAGO9*, lo que abre la posibilidad de que esta proteína posea actividad *slicer*, como muestran algunos miembros del mismo clado en *A. thaliana* (Qi *et al.*, 2006). Asimismo, se generaron modelos tridimensionales de las proteínas *AtAGO9* y *MpAGO9* y en ambos casos, la distribución espacial de los dominios L1, MID, L2, PAZ y PIWI no mostraron diferencias muy notables, no así en el dominio N-terminal. El dominio N-terminal de las proteínas AGO está estrechamente relacionado con la regulación de su actividad *slicer* (Hur *et al.*, 2013), por lo que dicha diferencia en las estructuras de este dominio podría indicar diferencias en las capacidades catalíticas de estas proteínas.

Los genes *AGO* muestran dominios de expresión que varían según la etapa de desarrollo y el tejido de la planta. Esta característica se ha relacionado estrechamente con la diversificación funcional de esta familia (Havecker *et al.*, 2010). *MpAGO9* muestra un perfil de expresión amplio, encontrándose en altos niveles en todos los tejidos evaluados de la etapa gametofítica del ciclo de vida de la planta, mientras que los otros miembros de este clado en *M. polymorpha* (*i.e.* *MpAGO4a* y *MpAGO4b*) muestran perfiles de expresión tejido específicos, expresándose principalmente en estructuras reproductivas (Flores-Martínez, 2015). Lo observado sugiere fuertemente que, como se observa con los miembros de este mismo clado en *A. thaliana*, en *M. polymorpha* los miembros del clado III podrían llevar a cabo funciones específicas según el tejido en donde se expresan (Olmedo-Monfil *et al.*, 2010; McCue *et al.*, 2015; Ziberman *et al.*, 2004).

La vía *RdDM* y particularmente la participación de *AGO9* se ha relacionado directamente con la especificación de la línea germinal tanto en *A. thaliana* como en maíz (Olmedo-Monfil *et al.*, 2010; Singh *et al.*, 2011). Asimismo, otro miembro de este clado *AGO4*, se ha relacionado con el desarrollo de células estomáticas durante condiciones de baja humedad (Tricker *et al.*, 2012). Debido a lo anterior, el hecho de que *MpAGO9* tenga un perfil de expresión amplio durante el desarrollo gametofítico sugiere que podría estar participando en la determinación de diferentes tipos celulares, pues las etapas evaluadas

corresponden a los cambios morfológicos más importantes durante el ciclo de vida de la planta: durante el estadio de gemas y talos de 7 días, se da el desarrollo de rizoides y células del talo como cámaras de aire, así como el mantenimiento de los polos meristemáticos; cuando los talos han alcanzado los 14 días de crecimiento el talo está completamente diferenciado en los polos ventral y dorsal, asimismo, durante esta etapa comienza la diferenciación celular que dará origen a las estructuras que producirán los propágulos asexuales, los conceptáculos. Finalmente, cuando el talo está maduro genera gametangióforos que darán origen a los gametos y en el caso de las hembras, portarán y sustentarán nutricionalmente al esporofito después de la fertilización (Durand, 1908; Bopp & Viktor, 1988). Todos estos eventos están marcados por procesos de diferenciación, desdiferenciación y rediferenciación celular en donde la *RdDM* y *MpAGO9* podrían estar participando en el silenciamiento de elementos transponibles y elementos repetitivos.

La historia de vida con dos generaciones multicelulares (*i.e.* esporofito y gametofito) asociadas por el desarrollo es una autapomorfía (un carácter derivado único para una especie o un grupo) de las embriofitas (Graham & Wilcox, 2000). La interacción entre estas dos formas de vida multicelulares o alternancia de generaciones ha sufrido modificaciones a través de la evolución de las plantas terrestres: en las plantas no vasculares el gametofito haploide es predominante mientras que el estadio de esporofito diploide es menos prominente y depende nutricionalmente del gametofito, por otra parte, en las plantas vasculares el esporofito es el estadio dominante y a través de la evolución existe evidencia de la disminución progresiva del desarrollo del gametofito (Kenrick & Crane, 1997; Graham & Wilcox, 2000). Existen diferentes teorías acerca del surgimiento de la alternancia de generaciones en las plantas (Graham & Wilcox, 2000; Kenrick & Crane, 1997), sin embargo, uno de los aspectos más intrigantes de la alternancia de generaciones es que dos planes corporales morfológicamente distintos son generados a partir de un mismo genoma, fenómeno que se ha relacionado directamente con la expresión diferencial de genes durante el desarrollo (Ambrose & Ferrándiz, 2012). Como actualmente se sabe, los procesos epigenéticos contribuyen de manera importante en la regulación de la expresión génica sin modificar su secuencia de DNA y por ende, resulta una explicación atractiva para la incógnita de la evolución de la alternancia de generaciones en las plantas terrestres involucrándose de manera directa en esta expresión diferencial de genes en los diferentes estadios durante la historia de vida. Algunos estudios han demostrado que existen diferencias en los transcriptomas de los gametofitos femenino (Yu *et al.*, 2005) y masculino (Pina *et al.*, 2005) con respecto a los transcriptomas del esporofito en *Arabidopsis*. Asimismo, actualmente se ha observado que existe una acumulación diferencial de RNAs pequeños -tanto microRNAs como siRNAs- entre las células vegetativas y las células espermáticas del polen (He *et al.*, 2015), así como acumulación diferencial de siRNAs entre las células que componen al gametofito femenino (Wuest *et al.*, 2010) y después de la fertilización, una acumulación diferencial de siRNAs dependientes de Pol IV (p4-siRNAs)

entre el embrión y el endospermo (Mosher *et al.*, 2009). Por otra parte, el estudio de Olmedo-Monfil y colaboradores (2010) muestran la importancia de la participación de AGO9 y la vía *RdDM* durante la especificación del linaje germinal en el gametofito femenino en *Arabidopsis*. Todo lo anterior sugiere que procesos epigenéticos, particularmente procesos dependientes de la participación de RNAs pequeños como la vía *RdDM*, podrían estar participando de manera importante en el desarrollo del gametofito y por tanto, en la transición de la fase gametofítica a la fase esporofítica. En los resultados obtenidos se mostró que *MpAGO9* muestra un perfil de expresión amplio durante la etapa gametofítica (predominante en su ciclo de vida), similar a lo observado en el gametofito femenino de *Arabidopsis*, por tanto, la función durante la etapa gametofítica de esta proteína podría estar conservada a pesar de la distancia evolutiva entre estas especies. *Item*, el hecho de hallar que AGO9, un componente central de la *RdDM* durante este proceso de determinación celular, se encuentra conservado en el clado basal del linaje de las embriofitas sugiere que éste pudo estar involucrado en la evolución de la alternancia de generaciones.

El análisis del dominio PIWI de *MpAGO9* también arrojó resultados interesantes sobre la conservación de residuos de aminoácidos necesarios para el establecimiento de interacciones con un grupo peculiar de proteínas: las proteínas con motivos GW/WG. Se ha probado por diferentes técnicas que este grupo de proteínas son capaces de establecer interacciones con las proteínas AGO y se cree que podrían participar como co-factores durante las diferentes vías de silenciamiento mediadas por RNAs pequeños (El-Shami *et al.*, 2007; He *et al.*, 2009; Pontier *et al.*, 2012). Los resultados mostraron que *MpAGO9* es el único miembro del clado en *M. polymorpha* que conserva todos los residuos necesarios para el establecimiento de la interacción. Asimismo, *MpAGO9* conserva un residuo de W que es único para las proteínas AGO de plantas, este residuo se ha relacionado con un posible mecanismo de autoinhibición en donde las proteínas AGO pueden modular la afinidad de interacción con proteínas GW para evitar así que las vías de silenciamiento mediadas por RNAs pequeños sean comprometidas por proteínas virales que presentan esta clase de reiteraciones de GW (Poulsen *et al.*, 2013).

Las predicciones bioinformáticas sugieren fuertemente que *MpAGO9* es una proteína capaz de establecer interacciones con proteínas GW y para probar estos resultados se realizó el ensayo de interacción por complementación de la fluorescencia con una proteína GW conocida, la subunidad grande de la RNA polimerasa V de *A. thaliana*. Se lograron clonar exitosamente los genes completos de *MpAGO9*, *AtAGO4* y *AtNRPE1* en los vectores de destino y al secuenciar las construcciones se corroboró que éstas se encontraban en fase con la fusión traduccional. Sin embargo, al observar los protoplastos de las plantas agroinfiltradas no se observó restauración de la fluorescencia de la YFP. En algunos de los protoplastos observados se detectó fluorescencia, sin embargo, los resultados de los análisis de RT-PCR y Western blot

mostraron que la fluorescencia no era debida a una interacción. Posiblemente la fluorescencia observada se deba a daño celular en los protoplastos inducido por la manipulación o incluso por la exposición a la luz durante la observación al microscopio que finalmente induce la generación de especies reactivas de oxígeno que alteran la fisiología celular (Dixit & Cyr, 2003).

Como anteriormente se mencionó, los resultados del análisis por RT-PCR y Western blot de las plantas agroinfiltradas mostraron que las plantas se transformaron exitosamente, sin embargo, no se estaban produciendo y/o acumulando proteínas. Debido a que se observa transcrito en todos los tratamientos se concluye que el fallo en el sistema se dio a nivel post-transcripcional, posiblemente afectando la estabilidad de las proteínas. En el caso de las proteínas AGO, un factor que podría explicar los resultados obtenidos recae en la naturaleza misma de estas proteínas: la estabilidad de las proteínas AGO depende en gran medida de su unión con RNAs pequeños (Smibert *et al.*, 2013). Con base en lo mencionado se abren diferentes escenarios posibles: 1) Las construcciones están diseñadas para establecer la fusión traduccional con el extremo N-terminal de la YFP en el extremo C-terminal de la AGO. Al plegarse la proteína, esta fusión pudo modificar la estructura tridimensional de la proteína AGO, afectando uno o varios de sus dominios y comprometiendo su capacidad de unión a siRNAs y por ende su estabilidad, lo que pudo culminar en la degradación de la proteína; 2) Durante la agroinfiltración se co-infiltró con una cepa de *Agrobacterium* que produce p19, una proteína viral que suprime el silenciamiento a nivel post-transcripcional regulando negativamente al efector del silenciamiento AGO1, lo que permite potenciar la expresión transitoria de transgenes en los tejidos agroinfiltrados (Voinnet *et al.*, 2003; Varállyay *et al.*, 2014). Se ha reportado que p19 además de regular negativamente a AGO1, interfiere de manera directa con la metilación de miRNAs, lo que produce su uridilación y posterior degradación. Específicamente, p19 interfiere con la actividad de la metiltransferasa HEN1 (Yu *et al.*, 2006), esta enzima también se encarga de metilar siRNAs de 24 nt durante la *RdDM* (Yang *et al.*, 2006). Por tanto, la co-infiltración con p19 podría estar afectando la disponibilidad de siRNAs lo que afecta de manera directa la acumulación de las proteínas AGO, como se ha observado con los niveles de acumulación de AGO4 cuando se compromete la producción de siRNAs (Li *et al.*, 2006) y 3) Las proteínas AGO sufren modificaciones post-traduccionales que tienen efecto en su estabilidad y actividad (*i.e.* hidroxilación, fosforilación, ubiquitinación, etc.) (Meister, 2013). Existe la posibilidad de que la fusión traduccional haya modificado la accesibilidad a los sitios en donde se llevan a cabo alguna de estas modificaciones provocando alteraciones en la proteína.

Por otra parte, en el caso de *AtNRPE1*, no se ha reportado ningún estudio en donde se sobreexpresara este gen o se hayan realizado fusiones traduccionales con proteínas fluorescentes (C. Pikaard, Indiana University, comunicación personal). Los reportes en donde se evalúa su localización celular se han

realizado utilizando anticuerpos específicos para esta proteína (Li *et al.*, 2006; Pontes *et al.*, 2006; Duan *et al.*, 2015) o utilizando epítomos (Pontes *et al.*, 2006; Haag *et al.*, 2009). Actualmente se desconoce si existen modificaciones post-traduccionales de esta proteína que module su actividad y/o función (Ream *et al.*, 2014). La RNA polimerasa V es un complejo proteico de 12 subunidades (Huang *et al.*, 2015) y se ha observado que los niveles de acumulación de NRPE1 se ven afectados cuando se interfiere con otras subunidades como NRPE5a (Lahmy *et al.*, 2009). Lo anterior sugiere que es posible que exista un mecanismo que regule la acumulación de estas proteínas manteniendo un ‘equilibrio’ en la producción de las diferentes subunidades, evitando así la sobreproducción de subunidades individuales. Por otra parte, es plausible un escenario similar al de la fusión traduccional con las proteínas AGO, en donde se afecta la estructura terciaria de la proteína y posiblemente su interacción con las otras subunidades lo que afecta su acumulación (Pikaard, comunicación personal).

Finalmente, no se descarta el hecho de que este fenómeno pueda ser producto de errores durante la amplificación de estos genes. Es probable que durante la amplificación se haya introducido alguna mutación en el gen que modificara la estructura de la proteína provocando inestabilidad y su posterior degradación, este escenario resulta probable ya que estos genes son de gran tamaño, por tanto, la posibilidad de la introducción de mutaciones durante la amplificación aumenta aún utilizando una DNA polimerasa de alta fidelidad. Se ha calculado que la tasa de error de la DNA polimerasa *Phusion* es de una en doscientas clonas después de 25 ciclos de PCR amplificando un fragmento de más de 1000 pb (www.neb.com). Asimismo, el hecho de no haber detectado proteína durante el Western blot podría deberse a bajos niveles de acumulación, más que ausencia de proteína, que no pudieron ser detectados con los anticuerpos primarios.

Los resultados del ensayo *BiFC* no permiten discernir si MpAGO9 y AtNRPE1 interactúan *in vivo*, por tanto, no se descarta el hecho de que MpAGO9 pueda establecer interacciones con esta clase de proteínas. Con base en lo anterior, se realizó una búsqueda bioinformática en el genoma de *M. polymorpha* para averiguar si existen potenciales interactores que participen en la *RdDM*. Los resultados mostraron que en el genoma de *M. polymorpha* se encuentran dos genes que codifican a la subunidad grande de la RNA polimerasa V, sin embargo, en este estudio únicamente se consideró uno de ellos (*MpNRPE1a*) pues uno carecía de repetidos GW (*MpNRPE1b*), por lo que no se consideró como un posible interacto con proteínas AGO. Los resultados, sin embargo, concuerdan con lo observado en *Physcomitrella patens* por Coruh y colaboradores (2015), en donde reportan la existencia de dos homólogos de NRPE1 y uno de ellos muestra un CTD trunco en donde no hay repetidos de GW.

Las proteínas GW son un grupo muy variado de proteínas y se ha reportado que los dominios WG/GW muestran un bajo nivel o carecen completamente de conservación en la secuencia más allá de los residuos de G y W, incluso entre especies cercanas. Asimismo, se ha observado que la carencia de similitud de secuencia entre proteínas con estos dominios no tiene influencia en su función (Karlowski, 2011). El alto nivel de divergencia y la conservación funcional en esta clase de proteínas abre todo un nuevo e interesante campo de estudio pues, como se ha observado, existe la posibilidad de que estas proteínas participen durante las vías de silenciamiento mediadas por RNAs pequeños regulando de manera directa la actividad de las proteínas AGO (Karlowski *et al.*, 2010; Karlowski, 2011; Pontier *et al.*, 2012; Poulsen *et al.*, 2013). En el caso de MpNRPE1a, la conservación de secuencia es evidente en el extremo N-terminal, no así al observarse el extremo C-terminal. Otra característica que resulta evidente, es que MpNRPE1a presenta un número menor de reiteraciones GW. Se ha reportado que el número mínimo para que una proteína GW pueda establecer interacción con la proteína AGO4 de *A. thaliana* es cuatro (He *et al.*, 2009), por tanto, el hecho de MpNRPE1 muestre un número menor de repetidos (18) en comparación con AtNRPE1 (20) podría no ser limitante en el establecimiento de interacciones con proteínas AGO. Los resultados anteriores sugieren fuertemente que MpNRPE1a podría ser una proteína con la capacidad de interactuar con proteínas AGO durante la *RdDM* en *M. polymorpha*.

Finalmente, los resultados obtenidos respaldan la hipótesis del surgimiento temprano de componentes centrales de la *RdDM* como son las proteínas AGO y la RNA polimerasa V (Huang *et al.*, 2015; Coruh *et al.*, 2015). Actualmente se cuenta con evidencia de la funcionalidad de la *RdDM* en un clado basal (Coruh *et al.*, 2015), empero, aún resta determinar si esta vía es funcional en el grupo de más temprana divergencia en las embriofitas e identificar si existen diferencias en el mecanismo entre este clado y el resto de las plantas. En el Laboratorio de Epigenética y Biología del Desarrollo (INBIOTECA-UV) se continúan realizando esfuerzos por tratar de responder estas preguntas mediante aproximaciones de genómica funcional.

## CONCLUSIONES Y PERSPECTIVAS



La familia AGO es una familia sumamente diversificada y es posible rastrear sus orígenes hasta un miembro del grupo de más temprana diversificación en las plantas, *M. polymorpha*. El clado III o 4/6/8/9 está involucrado con el silenciamiento a nivel transcripcional mediante la vía de metilación del DNA mediada por RNA, en el genoma de *M. polymorpha* es posible hallar tres genes que pertenecen a este clado. En este estudio se caracterizó a uno de estos miembros del clado, *MpAGO9* y se analizaron minuciosamente las características de la proteína que codifica: *MpAGO9* conserva los dominios característicos de las proteínas AGO y en el dominio PIWI conserva los residuos de la tétada catalítica, así como residuos únicos de las proteínas AGO de plantas y residuos importantes para el establecimiento de interacciones con proteínas GW. Asimismo, *MpAGO9* presenta dominios diversificados lo que da evidencia de su posible función durante el desarrollo gametofítico de la planta.

En el genoma de *M. polymorpha* se encuentran claros ortólogos de componentes centrales de la vía *RdDM*, lo que prueba que esta vía de silenciamiento surgió tempranamente en la evolución de las plantas terrestres. En este estudio se trató de probar la interacción *in vivo* de *MpAGO9* con una proteína GW, debido a las características estructurales y bioquímicas de las proteínas de estudio, no fue posible determinar si existe interacción por este sistema, por tanto, deberá probarse modificando el ensayo o mediante otras aproximaciones. Las predicciones bioinformáticas, sin embargo, indican que la interacción entre estas dos proteínas es muy probable.

Para tratar de probar la interacción mediante el sistema BiFC deberán tomarse algunas consideraciones. Lo primero que resulta evidente es el posible efecto de p19 sobre la estabilidad de los siRNAs y en consecuencia, sobre las proteínas AGO, por tanto, será recomendable no co-infiltrar con p19. Lo siguiente es lidiar con el mantenimiento de la estructura funcional de las proteínas. Aparentemente, con las proteínas AGO la fusión al extremo C-terminal podría estar modificando la estructura, por tanto el ensayo deberá probarse realizando fusiones al extremo N-terminal. Los vectores que se encuentran



disponibles actualmente, son los pSPYCE(M) y pSPYNE(R)173 (Waadt *et al.*, 2008). Actualmente existe un sólo reporte en donde, utilizando el par de vectores antes mencionados, muestran la interacción de la proteína AGO4 con el supresor viral 2b mediante *BiFC* (Hamera *et al.*, 2012). Por otra parte, debido a la aparente dificultad de sobreexpresar *AtNRPE1* y del posible efecto que esté surtiendo la fusión traduccional, un escenario plausible que permitiría determinar si existe interacción por este método sería fraccionar la proteína y realizar el ensayo únicamente con el dominio CTD, sitio en donde se establece la interacción.

Por otra parte, podría hacerse uso de otras técnicas para poder probar la interacción. En la literatura se han reportado ensayos de unión utilizando fusiones de proteínas con GST, mediante esta técnica se probó la interacción entre *AtAGO4* y *AtNRPE1* (El-Shami *et al.*, 2007) y entre *AtAGO4* y *AtKTF1* (He *et al.*, 2009). Considerando que actualmente contamos con anticuerpos para las proteínas AGO de *M. polymorpha*, un posible ensayo sería probar la interacción de *MpNRE1* fusionada a GST (ya sea la proteína completa o únicamente el CTD) con *MpAGO9*, utilizando como control positivo las proteínas de *A. thaliana*. Asimismo, podrían realizarse inmunolocalizaciones para observar *in vivo* la interacción.

## LITERATURA CITADA



- AALTO, A.P. & PASQUINELLI, A.E. (2012) Small non-coding RNAs mount a silent revolution in gene expression. *Current Opinion in Cell Biology*, 24:333-340
- AMBROSE, B.S. & FERRÁNDIZ, C. (2012) Development and the Evolution of Plant Form, Chapter 9. Eds.: Ambrose, B.A. & Purugganan, M.D., *Annual Plant Reviews: The evolution of plant form*. John Wiley and Sons, 408 pp.
- ALLEMAN, M., SIDORENKO, L., MCGINNIS, K., SESHADRI, V., DORWEILER, J.E., WHITE, J., SIKKINK, K., CHANDLER, V. (2006) An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature*, 442:295–298
- ARIF, M.A., FRANK, W., KHRAIWESH, B. (2013) Role of RNA Interference (RNAi) in Moss *Physcomitrella patens*. *Int J Mol Sci*, 14(1):1516-1540
- ARNOLD, K., BORDOLI, L., KOPP, J., SCHWEDE, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2):195-201
- ARTEAGA-VÁZQUEZ, M.A. & CHANDLER, V.L. (2010) Paramutation in maize: RNA mediated trans-generational gene silencing. *Current Opinion in Genetics and Development*. 20:156-163
- AXTELL, M. (2013) Classification and Comparison of Small RNAs from Plants. *Annual Review of Plant Biology*, 64:137-159
- BAUMBERGER, N. & BAULCOMBE, C. (2005) *Arabidopsis* ARGONAUTE-1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *PNAS*, 102(33): 11928-11933
- BERNSTEIN, E., CAUDY, A.A., HAMMOND, S.M., HANNON, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, 409:363-366
- BLEVINS, T., PODICHETI, R., MISHRA, V., MARASCO, M., WANG, J., RUSCH, D., TANG, H., PIKAARD, C. (2015) Identification of Pol IV and RDR2 dependent precursors of 24 nt siRNAs guiding *de novo* DNA methylation in

*Arabidopsis*. eLIFE, 10.7554/eLife.09591

BOHMERT, K., CAMUS, I., BELLINI, C., BOUCHEZ, D., CABOCHE, M., BENNING, C. (1998) AGO1 defines a novel locus of *Arabidopsis* controlling leaf development. 17(1):170-180

BONNET, E., VAN DER PEER, Y., ROUZÉ, P. (2006) The small RNA world of plants. *New Phytologist*, 451-463

BOPP, M. & VIKTOR, R. (1988) Protoplasts of *Marchantia polymorpha* and its Development. *Plant Cell Physiol.*, 29(3): 497-501

BORGES, F. & MARTIENSSEN R.A. (2015) The expanding world of small RNAs in plants. *Nature Reviews Molecular Cell Biology*, AOP, doi:10.1038/nrm4085

BORGES, F., PEREIRA, P.A., SLOTKIN, K., MARTIENSSEN, R.A., BECKER, J.D. (2010) MicroRNA activity in the *Arabidopsis* male germline. *Journal of Experimental Botany*, 62(5):1611-1620

BOWMAN, J.L., FLOYD, S.K., SAKAKIBARA, K. (2007) Green Genes- Comparative Genomics of the Green Branch of Life. *Cell*, 129:229-234

BRENNECKE J, ARAVIN, A.A., STARK, A., DUS, M., SACHIDANANDAM, R., HANNON, G.J. (2007) Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in *Drosophila*. *Cell*, 128:1089–1103.

BRINK, R.A. (1956) A Genetic Change Associated with the *R* Locus in Maize Which is Directed and Potentially Reversible. *Genetics*, 41:872-889

BROSSEAU, C. & MOFFETT, P. (2015) Functional and Genetic Analysis Identify a Role for *Arabidopsis* ARGONAUTE5 in Antiviral RNA Silencing. *The Plant Cell*, 27(5)

CASTEL, S.E. & MARTIENSSEN, R.A. (2013) RNA interference (RNAi) in the Nucleus: roles for small RNA in transcription, epigenetics and beyond. *Nat Rev Genet*, 12(2):100-112

CHANDLER, V.L. (2007) Paramutation: From Maize to Mice. *Cell*, 128:641-645

CRAMER, P., BUSHNELL, D.A., KORNBERG, R.D. (2001) Structural Basis of Transcription: RNA Polymerase II at 2.8 Ångstrom Resolution. *Science*, 8:1863-1875

COE, E.H. (1959) A Regular Continuing Conversion-Type Phenomenon at the *B* Locus in Maize. *PNAS*, 45:828-832

COGONI, C. & MACINO, G. (1997) Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in *Neurospora crassa*. *PNAS*, 94:10233-10238

- CORUH, C., HYUN CHO, S., SHAHID, S., LIU, Q., WIERZBICKI, A., AXTELL, M. (2015) Comprehensive Annotation of *Physcomitrella patens* Small RNA Loci Reveals That the Heterochromatic Short Interfering RNA Pathway Is Largely Conserved in Land Plants. *The Plant Cell*, doi: 10.1115/tcp.15.00228
- DIXIT, R. & CYR, R. (2003) Cell damage and reactive oxygen species production induced by fluorescence microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy. *Plant J.*, 36(2):280-90
- DOYLE, M., JASKIEWICZ, L., FILIPOWICZ, W. (2012) Dicer Proteins and Their Role in Gene Silencing Pathways. *The Enzymes*, 32:1-35
- DOWEN, R.H., PELIZZOLA, M., SCHMITZ, R.J., LISTER, R., DOWN, J.M., NERY, J.R., DIXON, J.E., ECKER, J.R. (2012) Widespread dynamic DNA methylation in response to biotic stress. *PNAS*, E2183-E2191
- DURÁN-FIGUEROA N. & VIELLE-CALZADA, J-P. (2010) ARGONAUTE9 dependent silencing of transposable elements in pericentromeric regions of *Arabidopsis*. *Plant Signaling and Behavior*, 5(11): 1476-1479
- DUAN, C-H., ZHANG, H., TANG, K., ZHU, X., QIAN, W., HOU, Y-H., WANG, B., LANG, Z., ZHAO, Y., WANG, X., WANG, P., ZHOU, J., LLANG, G., LIU, N., WANG, C., ZHU, J-K. (2015) Specific but interdependent functions for *Arabidopsis* AGO4 and AGO6 in RNA-directed DNA methylation. *The EMBO Journal*, 34:581-592
- DURAND, E.J. (1908) The Development of the Sexual Organs and Sporogonium of *Marchantia polymorpha*. *Bulletin of the Torrey Botanical Club*. 35(7):321-335
- EDGAR, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Res*; 32(65):1792-7
- EL-SHAMI, M., PONTIER, D., LAHMY, S., BRAUN, L., PICART, C., VEGA, D., HAKIMI, M.A., JACOBSEN, S.E., COOKE, R., LAGRANGE, T. (2007) Reiterated WG/GW motif form functionally and evolutionary conserved ARGONAUTE-binding platforms in RNAi related components. *Genes Dev*, 21(20):2539-44
- EUN, C., LORKOVIC, Z.J., NAUMANN, U., LONG, Q., HAVECKER, E.R., SIMON, S.A., MEYERS, B.C., MATZKE, A.J.M., MATZKE, M.(2011) AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in *Arabidopsis*. *PLoS One*, 6(10) e25730
- ERHARD, K.F. & HOLLICK, J.B. (2011) Paramutation: a process for acquiring trans-generational regulatory states. *Current Opinion in Plant Biology*, 14:1-7
- ERHARD, K.F., STONAKER, J.L., PARKINSON, S.E., LIM, J.P., HALE, C.J., HOLLICK, J.B. (2009) RNA polymerase IV functions in paramutation in *Zea mays*. *Science*, 323:1201-1205
- FAEHNLE, C.R., ELKAYAM, E., HAASE, A.D., HANNON, G.J., JOSHUA-TOR, L. (2013) The making of a Slicer: Activation

of Human Argonaute-1. *Cell Rep*, 3(6):1901-1909

FEDOROFF, N.V. (2012) Transposable Elements, Epigenetics and Genome Evolution. *Science*, 338:758-767

FEI, Q., XIA, R., MEYERS, B. (2013) Phased, Secondary, Small Interfering RNAs in Posttranscriptional Regulatory Networks. *The Plant Cell*, 25:2400-2415

FIRE, A., XU, SQ, MONTGOMERY, M.K, KOSTAS, S.A., DRIVER, S.E., MELLO, C.C.(1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 391:806-810

FINN, R.D., CLEMENS, J., CLEMENTS, J., ARNDT, W., MILLER, B.J., WHEELER, T.J., SCHREIBER, F., BATEMAN, A., EDDY, S.R. (2015) HMMER web server: 2015 update. *Nucle. Acids Res.*, 43 (W1): W30-W38.

FRANK, F., HAUVER, J., SONENBERG, N., NAGAR, B. (2012) *Arabidopsis* Argonaute MID domains use their nucleotide specificity to sort small RNAs. *The EMBO Journal*, 31:3588-3595

FLORES-MARTÍNEZ, D.O. (2015) Identificación y caracterización del patrón de expresión del gen *ARGONAUTE4* en *Marchantia polymorpha*. Tesis de Licenciatura, Universidad Veracruzana.

FLORES-SANDOVAL, E., DIERSCHKE, T., FISHER, T.J., BOWMAN, J.L. (2015) Efficient and inducible use of artificial microRNAs in *Marchantia polymorpha*. *Plant and Cell Physiology*, doi:10.1093

GUO, S. & KEMPHUES, K.J. (1995) *par-1*, a Gene Required for Establishing Polarity in *C. elegans* Embryos, Encodes a Putative Ser/Thr Kinase That Is Asymmetrically Distributed. *Cell*, 81:611-620

GRAHAM, L.K & WILCOX, L.W. (2000) The origin of alternation of generations in land plants: a focus on matrotrophy and hexose transport. *Phil. Trans. R Soc. Lond.*, 355:757-767

GROßHANS, H. & FILIPOWICZ, W. (2008) The expanding world of small RNAs. *Nature*, 451:414-416

HAAG, J.R., PONTES, O., PIKAARD, C.S. (2009) Metal A and metal B sites of nuclear RNA polymerases: Pol IV and Pol V are required for siRNA-dependent DNA methylation and gene silencing. *PLoS One*, 4(1):e4110.doi10.1371/journal.pone.0004110

HAAG, J.R. & PIKAARD, C.S. (2011) Multisubunit RNA polymerase IV and V: purveyors of non-coding RNA for plant gene silencing. *Nature Reviews Molecular Cell Biology*, 12:483-492

HALE, C.J., STONAKER, J.L., GROSS, S.M., HOLLICK, J.B. (2007) A Novel Snf2 Protein Maintains *trans* Generational Regulatory States Established by Paramutation in Maize. *PLOS Biology*, doi:10.1371/journal.ppbio.0050275

- HAMERA, S., SONG, X., SU, L., CHEN, X., FANG, R. (2012) Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. *The Plant Journal*, 69:104-115
- HAMILTON, A.J. & BAULCOMBE D.C. (1999) A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. *Science*, 286:950-951
- HAMMOND, S.M., BOETTCHER, S., CAUDY, A.A., KOBAYASHI, R., HANNON, G.J. (2001) Argonaute2, a Link Between Genetic and Biochemical Analyses of RNAi. *Science*, 293: 1146-1150
- HAVECKER, E.R., WALLBRIDGE, L.M., HARDCASTLE, T.J., BUSH, M.S., KELLY, K.A., DUNN, R.M., CSHWACH, F., DOONAN, J.H., BAULCOMBE, D.C. (2010) The *Arabidopsis* RNA-Directed DNA Methylation Argonautes Functionally Diverge Based on their Expression and Interaction with Target Loci. *The Plant Cell*, 22:321-334
- HE, X-J., HSU, Y-F., ZHU, S., WIERZBICKI, A.T., PONTES, O., PIKAARD, C.S., LIU, H-L., WANG, C-S., JIN, H., ZHU, J-K (2009) An Effector of RNA-Directed DNA Methylation in *Arabidopsis* Is An ARGONAUTE-4 and RNA-Binding Protein. *Cell*, 137(3):498-508
- HE, H., YANG, T., WU, W., ZHENG, B. (2015) Small RNAs in pollen. *Sci. China Life Sci.*, 58(3):246-252
- HERR, A.J., JENSEN, M.B., DALMAY, T., BAULCOMBE, D.C. (2005) RNA Polymerase IV Directs Silencing of Endogenous DNA. *Science*, 308: 118-120
- HÖCK, J & MEISTER, G. (2008) The Argonaute protein family. *Genome Biology*, 9:210
- HOLLICK, J.B. (2010) Paramutation and development. *Annu. Rev. Cell Dev. Biol.*, 26:557-579
- HIRSCH, S., BAUMBERGER, R., GROSSNIKLAUS, U. (2012) Epigenetic Variation, Inheritance and Selection in Plant Populations. CSH Laboratory Press, 77: 97-104
- HUANG, Y., KENDALL, T., FORSYTHE, E.S., DORANTES-ACOSTA, A., LI, S., CABALLERO-PÉREZ, J., CHEN, X., ARTEAGA-VÁZQUEZ, M., BEILSTEIN, M.A., MOSHER, R. (2015) Ancient Origin and Recent Innovations of RNA Polymerase IV and V. *Mol. Bio. Evol.*, 32(7):1788-1799
- HUNTER, C., SUN, H., POETHIG, R.S. (2003) The *Arabidopsis* Heterochronic Gene ZIPPY is an ARGONAUTE Family Member. *Current Biology*, 13:1734-1739
- HUR, J.K., ZINCHENKO, M.K., DJURANOVIC, S., GREEN, R. (2013) Regulation of Argonaute Slicer Activity by Guide RNA 3' End Interactions with the N-terminal Lobe. *Journal of Biological Chemistry*, 288(11):7829-7840
- ISHIZAKI, K., CHIYODA, S., YAMATO, K.T., KOHCHI, T. (2008) *Agrobacterium*-mediated transformation of the haploid

liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant Cell Physiology*, 49(7):1084-91

ISHIZAKI, K., NISHIHAMA, R., YAMATO, K.T., KOHCHI, T. (2015) Molecular Genetic Tools and Techniques for *Marchantia polymorpha* Research. *Plant and Cell Physiology*, doi: 10.1093

ITO, H., GAUBERT, H., BUCHER, E., MIROUZE, M., VAILLANT, I., PASZKOWSKI, J. (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*, 472:115-120

JACOBSEN, S.E., RUNNING, M.P. & MEYEROWITZ, E. (1999) Disruption of an RNA helicase/RNase III gene in *Arabidopsis* causes unregulated cell division in floral meristems. *Development*, 126:5231-5243

JAUBERT, M.J., BHATTACHARJEE, S., MELLO, A.F.S., PERRY, K.L., MOFFETT, P. (2011) AGO2 mediates RNA silencing anti-viral defenses against Potato Virus X in *Arabidopsis*. *Plant Physiology*, 156: 1556-1564

KARLOWSKI, W.M., ZIELEZINSKI, A., CARRÉRE, J., PONTIER, D., LAGRANGE, T., COOKE (2010) Genome-wide computational identification of WG/GW Argonaute-binding proteins in *Arabidopsis*. *Nucleic Acids Research*, 38(13): 4231-4245

KARLOWSKI, W.M. (2011) Beyond sequence similarity- The curious case of GW/WG protein domain. *Journal of Biotechnology, Computational Biology and Bionanotechnology*. 92:303-305

KAPOOR, M., ARORA, R., LAMA, T., NOHAWAN, A., KHURANA, J.P., TYAGI, A.K., KAPOOR, S. (2008) Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics*, 9:451

KENRICK, P. & CRANE, P.R. (1997) The origin and early evolution of plants on land. *Nature Reviews*, 389:33-39

KERPPOLA, T.K. (2009) Visualization of molecular interactions using bimolecular fluorescence complementation analysis: Characteristics of protein fragment complementation. *Chemical Society Reviews*.38:2876-2886

KIBBE, W.A. (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.*, 35(Web Server issue):W43-6

KIM, V.N., HAN, J., SIOMI, M.C. (2009) Biogenesis of small RNAs in animals. *Nature Reviews Molecular Cell Biology*, 10:126-139

KODAMA, Y. & HU, D.H. (2012) Bimolecular fluorescence complementation (BiFC): A 5-year update and future perspectives. *BioTechniques*, 53:285-298

- KUBOTA, A., ISHIZAKI, K., HOSAKA, M., KOHCHI, T. (2013) Efficient *Agrobacterium*-mediated Transformation of the Liverwort *Marchantia polymorpha* using regenerating thalli. *Biosci. Biotechnol. Biochem.*, 77(1):167-172
- KUHN, C.D. & JOSHUA-TOR, L. (2013) Eukaryotic Argonautes come into focus. *Cell*, 38(5):263-271
- KURIHARA, Y. & WATANABE, Y. (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *PNAS*, 101(34):12753-12758
- KWAK, P.B. & TOMARI, Y. (2012) The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nature Structural and Molecular Biology*, 19: (2):145-153
- LAHMY, S., PONTIER, D., CAVEL, E., VEGA, D., EL-SHAMI, M., KANNO, T., LAGRANGE, T. (2009) PolV (PolIVb) function in RNA-directed DNA methylation requires the conserved active site and an additional plant-specific subunit. *PNAS*, 106(3):941-946
- LARA-MONDRAGÓN, C. M. & ARTEAGA-VÁZQUEZ, M. A. (2015) A day in the life of a *Marchantia* Lab. The Node, The Company of Biologists LTD en <http://thenode.biologists.com/a-day-in-the-life-of-a-marchantia-lab/lablife>
- LAW, J.A. & JACOBSEN, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Review Genetics*, 11:204-220
- LAW, J.A., VASHISHT, A.A., WOHLSCHEGEL, J.A., JACOBSEN, S. (2011) SHH1, a Homeodomain Protein Required for DNA Methylation, as well as RDR2, RDM4 and Chromatin Remodeling Factors Associate with RNA Polymerase IV. *PLoS Genetics*, 7(7): e1002195
- LEE, T., GUARAZADA, S.G.R., ZHAI, J., LI, S., SIMON, S.A., MATZKE, M.A., CHEN, X., MEYERS, B. (2012) RNA polymerase V-dependent small RNAs in *Arabidopsis* originate from small, intergenic loci including most SINE repeats. *Epigenetics*, 7(7):781-795
- LI, C.F., PONTES, O., EL-SHAMI, M., HENDERSON, I.R., BERNATAVICHUTE, Y.V., CHAN, S.W.L., LAGRANGE, T., PIKAARD, C.S., JACOBSEN, S. (2006) An ARGONAUTE4-Containing Nuclear Processing Center Colocalized with Cajal Bodies in *Arabidopsis thaliana*. *Cell*, 126:93-106
- LINGEL, A., SIMON, B., IZAURRALDE, E., SATTLER, M. (2004) Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol*, 11:576-577
- LUO, J. & HALL, B.D. (2007) A multistep process gave rise to RNA polymerase IV of land plants. *J. Mol. Evol.*, 64(1):101-112
- MA, J.B., YE, K., PATEL, D.J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the



PAZ domain. *Nature*, 429(6989):318-22

MALLORY, A.C. & VAUCHERET, H. (2006) Functions of micro RNAs and related small RNAs in plants. *Nature Genetics Supplement*, 38:S31-S36

MALLORY, A. & VAUCHERET, H. (2010) Form, Function, and Regulation of ARGONAUTE Proteins. *The Plant Cell*, 22:3879-3889

MARCHLER-BAUER, A., LU, S., ANDERSON, J.B., CHITZAS, F., DERBYSHIRE, M.K., DEWEESE-SCOTT, C., FONG, J.H., GEER, L.Y., GEER, R.C., GONZALES, N.R., GWADZ, M., HURWITZ, D.I., JACKSON, J.D., LANCZYCKI, C.J., MARCHLER, G.H., MULLOKANDOV, M., OMELCHENKO, M.V, ROBERTSON, C.L., SONG, J.S., THANKI, N., SCHWEDE, T., KOPP, J., GUEX, N., PEITSCH, M.C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acid Res.* 31(13):3381–3385

MARTINEZ, J., PATKANIOWSKA, A., URLAUB, H., LÜHRMANN, R., TUSCHL, T. (2002) Single-stranded Antisense siRNAs Guide Target RNA Cleavage in RNAi. *Cell*, 110:563-574

MATZKE, M.A & MOSHER, R.A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Gen*, 15:394-408

McCLINTOCK, B. (1950) The origin and behaviour of mutable loci in maize. *PNAS*, 36(6): 344-355

MCCUE, A.D., PANDA, K., NUTHIKATTU, S., CHOUDURY, S.G., THOMAS, E.N., SLOTKIN, R.K. (2015) ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *The EMBO Journal*, 34:20-35

MEISTER, G. (2013) Argonaute proteins: functional insights and emerging roles. *Nature Reviews*, 14:447-460

METTE, M.F., AUFSATZ, W., VAN DER WINDEN, J., MATZKE, M.A., MATZKE, A.J.M. (2000) Transcriptional silencing and promoter methylation triggered by double stranded RNA. *EMBO Journal*, 19:5194-5201

MI, S., CAI, T., HU, Y., CHEN, Y., HODGES, E., NI, F., WU, F., WU, L., LI, S, ZHOU, H., LONG, C., CHEN, S., HANNON, G.J., QI, Y. (2008) Sorting of Small RNAs into *Arabidopsis* Argonaute Complexes is Directed by the 5' Terminal Nucleotide. *Cell*, 133:116-127

MIROUZE, A., & PASZKOWSKI, J. (2011) Epigenetic contribution to stress adaptation in plants. *Current Opinion in Plant Biology*, 14(3): 267-274

MISHLER, B.D. (2001) The biology of Bryophytes. *America Journal of Botany*, 88(11):2129-2131

- MLOTSHWA, S., PRUSS, G.J., PERAGINE, A., ENDRES, M.W., LI, JUNJIE, CHEN, X., POETHIG, R.S., BOWMAN, L.H., VANCE, V. (2008) DICER-LIKE2 plays a primary role in transitive silencing of transgenes in *Arabidopsis*. PLoS ONE, doi: 10.1371/journal.pone.0001755
- MOLNAR, A., MELNYK, C., BAULCOMBE, D. (2011) Silencing signals in plants: a long journey for small RNAs. *Genome Biology*, 12:215
- MONTGOMERY, T.A., HOWELL, M.D., CUPERUS, J.T., LI, D., HANSEN, J.E., ALEXANDER, A.L., CHAPMAN, E.J., FAHLGREN, N., ALLEN, E., CARRINGTON, J.C. (2008) Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. *Cell*, 133: 128-141
- MOSHER, R.A., MELNYK, C.W., KELLY, K.A., DUNN, R.M., STUDHOLME, D.J., BAULCOMBE, D.C. (2009) Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature*, 460:283-287
- MOSHER, R.A., SCHWACH, F., STUDHOLME, D., BAULCOMBE, D.C. (2008) PolVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *PNAS*, 105(8):3145-3150
- NAKANISHI, K., WEINBERG, D.E., BARTEL, D.P., PATEL, D.J. (2012) Structure of yeast Argonaute with guide RNA. *Nature*, 486(7403)
- NAPOLI, C., LEMIEUX, C., JORGENSEN, R. (1990) Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell*, 2:279-289
- OLMEDO-MONFIL, V., DURÁN-FIGUEROA, N., ARTEAGA-VÁZQUEZ, M., DEMESA-ARÉVALO, D.A., GRIMANELLI, D., SLOTKIN, R.K., MARTIENSSEN, R.A., VIELLE-CALZADA, J.P. (2010) Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* 464, 628-632
- ONODERA, Y., HAAG, J.R., REAM, T., COSTA-NUNES, P., PONTES, O., PIKAARD, C. (2005) Plant Nuclear RNA Polymerase IV Mediates siRNA and DNA Methylation-Dependent Heterochromatin Formation. *Cell*, 120: 613-622
- PARK, W., LI, J., SONG, R., MESSING, J., CHEN, X. (2002) CARPEL FACTORY, a Dicer Homolog, and HEN1, a Novel Protein, Act in microRNA Metabolism in *Arabidopsis thaliana*. *Current Biology*, 12:1484-1495
- PINA, C., PINTO, F., FEIJÓ, J.A., BECKER, J.D. (2005) Gene Family Analysis of the *Arabidopsis* Pollen Transcriptome Reveals Biological Implications for Cell Growth, Division Control, and Gene Expression Regulation. *Plant Physiology*, 138: 744-756
- PONTES, O., LI, C.F., COSTA NUNES, P., HAAG, J., REAM, T., VITINS, A., JACOBSEN, SE, PIKAARD, CS (2006) The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell*,

126(1):79-92

- PONTIER, D., PICART, C., ROUDIER, F., GARCIA, D., LAHMY, S., AZEVEDO, J., ALART, E., LAUDIÉ, M., KARLOWSKI, W.M., COOKE, R., COLOT, V., VOINNET, O., LAGRANGE, T. (2012) NERD, a pPlant-Specific GW Protein, Defines an Additional RNAi-Dependent Chromatin-Based Pathway in *Arabidopsis*. *Molecular Cell*, 48:121-132
- POPOVA, O.V., DIHN, H.Q., AUFSATZ, W., JONAK, C. (2013) The RdDM pathway is required for basal heat tolerance in *Arabidopsis*. *Molecular Plant*, 6(2):396-410
- POULSEN, C., VAUCHERET, H., BRODERSEN, P. (2013) Lessons on RNA Silencing Mechanisms in Plants from Eukaryotic Argonaute Structures. *Plant Cell*, 25: 22-37
- QI, Y., HE, X., WANG, X.-J., KOHANY, O., JURKA, J., HANNON, G.J. (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature Letters*, 443: doi:10.1038
- QIAN, Y., CHENG, Y., CHENG, X., JIANG, H., ZHU, S., CHEN, B. (2011) Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Reports*, 30(7):1347-1363
- QIU, Y.-L. & PALMER, J.D. (1999) Phylogeny of early land plants: insights from genes and genomes. *Trends in Plant Science*. 4(1):26-30
- QIU, Y.-L., LI, L., WANG, B., CHEN, Z., DOMBROVSKA, O., LEE, J., KENT, L., LI, R., JOBSON, R.W., HENDRY, T.A., TAYLOR, D.W., TESTA, C.M., AMBROS, M. (2007) A non-flowering land plant phylogeny inferred from nucleotide sequences of seven chloroplasts, mitochondria and nuclear genes. *Int. J. Plant Sci.*, 168(6):691-708
- RAVINDRAN (2012) Barbara McClintock and the discovery of jumping genes. *PNAS*, 109:20198-20199
- REAM, T.S., HAAG, J.R., PIKAARD, C.S. (2014) Plant Multisubunit RNA Polymerases IV and V. , Chapter 13 en Murakami, K.S & Trakselis, M.A. (eds.) *Nucleic Acid Polymerases, Nucleic Acids and Molecular Biology*, 30: doi 10.1007/978-3-642-39796-7\_13
- RENSING, S.A., LANG, D., ZIMMER, A.D., TERRY, A., SALAMOV, A., SHAPIRO, H., NISHIYAMA, T., PERROUD, P., LINDQUIST, E.A., KAMISUGI, Y., TANAHASHI, T., SAKAKIBARA, K., FUJITA, T., OISHI, K., SHIN-I, T., KUROKI, Y., TOYODA, A., SUZUKI, Y., HASHIMOTO, S., YAMAGUCHI, K., SUGANO, S., KOHARA, Y., FUJIYAMA, A., ANTEROLA, A., AOKI, S., ASHTON, N., BARBAZY, W.B., BARKER, E., BENNETZEN, J.L, BLANKENSHIP, R., ET AL. (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*, 319(5859):64-69.
- ROMANO, N. & MACINO, G. (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by

- transformation with homologous sequences. *Mol Microbiol*, 6(22):3343-53
- SCHAUER, S.E., JACOBSEN, S.E., MEINKE, D.W., RAY, A. (2002) DICER-LIKE 1: blind men and elephants in *Arabidopsis* development. *Trends in Plant Science*, 7(11): 487-491
- SCHIRLE, N.T. & MCRAE, I.J (2012) The Crystal Structure of Human Argonaute2. *Science*, 336:1037-1040
- SEN, G.L. & BLAU, H.M. (2006) A brief history of RNAi: the silence of the genes. *The FASEB Journal*, 1293-1299
- SHENG, G., ZHAO, H. , WANG, J., RAO, Y., TIAN, W., SWARTS, D., VAN DER OOST, J., PATEL, D.J., WANG, Y. (2013) Structure-based cleavage mechanism of *Thermus thermophilus* Argonaute DNA guide strand-mediated DNA target cleavage. *PNAS*, 111(2):652-657
- SINGH, M., GOEL, S., MEELEY, R.B., DANTEC, C., PARRINELLO, H., MICHAUD, C., LEBLNC, O., GRIMANELLI, D. (2011) Production of viable gametes without meiosis in maize deficient of an ARGONAUTE protein. *Plant Cell*, 23(2):443-58
- SINGH, R.K., GASE, K., BALDWIN, I.T., PANDEY, S.P. (2015) Molecular evolution and diversification of the Argonaute family of proteins in plants. *BMC Plant Biology*, 15(23): doi: 10.1186
- SLOTKIN, R.K. & MARTIENSSSEN, R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Reviews*, 8: 272-285
- SMIBERT, P., YANG, J-S., AZZAM, G., LIU, J-L., LAI, E.C. (2013) Homeostatic control of Argonaute stability by microRNA availability. *Nat. Struc. Mol. Biol.*, 20(7):789-795
- SONG, J.J., SMITH, S.K., HANNON, G.J., JOSHUA-TOR, L. (2004) Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science*, 305:1434-1437
- STAM, M. (2009) Paramutation: A Heritable Change in Gene Expression by Allelic Interactions *In Trans*. *Molecular Plant*, 2(4):578-588
- STEERE, W. C. (1940) Liverworts of southern Michigan. *Cranbrook Institute of Science Bulletin* 17, Bloomfield, MI: Cranbrook Press. 97 p.
- SUGANO, S., TAKAGI, S., SHIMADA, M., HARA-NISHIMURA, T., KOHCHI, T. (2014) CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. *Plant Cell Physiol.* 55(3):475-81
- SWARTS, D.C, MARAKOVA, K., WANG Y., NAKANISHI, K., KETTING, R.F., KOONIN, E.V., PATEL, D.J., VAN DER OOST, J. (2014) The evolutionary journey of Argonaute proteins. *Nat Struc & Mol Biol*, 21(9):743-753

- TABARA, H., SAKISSIAN, M., KELLY, W.G., FLEENOR, J., GRISHOK, A., TIMMONS, L., FIRE, A., MELLO, C.C. (1999) The *rde-1* Gene RNA Interference, and transposon silencing in *C. elegans*. *Cell*, 99(2):123-132
- TAKEDA, A., IWASAKI, S., WATANABE, T., UTSUMI, M., WATANABE, Y. (2008) The mechanism selecting the guide strand from small RNA duplexes is different among Argonaute proteins. *Plant Cell Physiol.*, 49: 493–500
- TSUBOYAMA, S., KODAMA, Y. (2014) AgarTrap: a simplified *Agrobacterium*-mediated transformation method for sporelings of the liverwort *Marchantia polymorpha* L. *Plant Cell Physiology*, 55(1):229-236
- TSUBOYAMA-TANAKA, S. & KODAMA, Y. (2015) AgarTrap-mediated genetic transformation using intact gemmae/gemmalings of the liverwort *Marchantia polymorpha* L.J. *Plant Res.*, 128(2):337-44
- TRICKER, P.J., GIBBINGS, J.G, RODRÍGUEZ-LÓPEZ, C.M., HADLEY, P., WILKINSON, M.J. (2012) Low relative humidity triggers RNA-directed de novo DNA methylation and suppression of genes controlling stomatal development. *Journal of Experimental Botany*, 3799-3813
- VAN DER KROL, A.R., MUR, L.A., BELD, M., MOL, J.N.M., STUITJE, A.R.(1990) Flavonoid Genes in Petunia: Addition of Limited Number of Gene Copies May Lead to a Suppression of Gene Expression. *The Plant Cell*, 2:291-299
- VAUCHERET, H. (2008) Plant ARGONAUTES. *Trends in Plant Science*, 13(7):1360-1385
- VÁRALLYAY, E., OLÁH, E., HAVELDA, Z. (2014) Independent parallel functions of p19 plant viral suppressor of RNA silencing required for effective suppressor activity. *Nucleic Acid Res.*, 42(1): 599-608
- VERDE, A., JIA, S., GERBER, S., SUGIYAMA, T., GYGI, S., GREWAL, S.I.S., MOAZED, D. (2004) RNAi-Mediated Targeting of Heterochromatin by the RITS Complex. *Science*, 302(5658):672-676
- VOINNET, O., RIVAS, S., MESTRE, P., BAULCOMBE, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal*, 33: 949-956
- WAADT, R., SCHMIDT, L.K., LOHSE, M., HASHIMOTO, K., BOCK, R., KUDLA, J. (2008) Multicolor bimolecular fluorescence complementation reveals simultaneous formation of alternative CBL/CIPK complexes in plants. *Plant Journal*, 56:505-516
- WALTER, M., CHABAN, C., SHÜTZE, K., BATISTIC, O., WECKERMANN, K., NÄKE, C., BLAZEVIC, D., GREFEN, C., SHUMACHER, K., OECKING, C., HARTER, K., KUDLA, J. (2004) Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *The Plant Journal*. 40: 428-438
- WASSENEGGER, M. (2006) The Role of the RNAi Machinery in Heterochromatin Formation. *Cell*, 122:13-16

- WASSENEGGER, M., HEIMES, S., RIEDEL, L., SÄNGER, H.L. (1994) RNA-Directed De Novo Methylation of Genomic Sequences in Plants. *Cell*, 76: 567-576
- WEI, K.F., WU, L.J., CHEN, J., CHEN, Y.F., XIE, D.X. (2012) Structural Evolution and Functional Diversification Analyses of Argonaute Protein. *Journal of Cellular Biochemistry*, 113:2576-2585
- WEI, L., GU, L., SONG, X., CUI, X., LU, Z., ZHOU, M., WANG, L., HU, F., ZHAI, MEYERS, B.C., CAO, X. (2014) Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice. *PNAS*, 111(10):3877-3882
- WICKETT, N., MIRARAB, S., NGUYEN, N., WARNOW, T., CARPENTER, E., MATASCI, N., AYYAMPALAYAM, S., BARKER, M., BURLEIGH, J., GITZENDANNER, M., RUEHFEL, B., WAFULA, E., DER, J., GRAHAM, S., MATHEWS, S., MELKONJAN, M., SLOTIS, D., SOLTIS, P., MILES, N., ROTHFELS, C., POKORNY, L., SHAW, J., DEGIRONIMO, C., CHEN, T., DEYHOLOS, M.K., BAUCOM, R., KUTCHAN, T., AUGUSTIN, M., WANG, M., WANG, J., ZHANG, Y., TIAN, Z., YAN, Z., WU, X., SUN, X., SHU WONG, G., LEEBENS-MACK, J. (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. *PNAS*, E4859-E4868
- WOODHOUSE, M.R. , FREELING, M., LISCH, D.(2006) The *mop1* (mediator of paramutation1) mutant progressively reactivates one of the two genes encoded by the MuDR transposon in maize. *Genetics*, 172:579–592
- WUEST, S.E., VIJVERBERG, K., SCHMIDT, A., WEISS, M., GHEYSELINCK, J., LOHR, M., WELLMER, F., RAHNENFÜRER, J., VON MERING, C., GROSSNIKLAUS, U. (2010) *Arabidopsis* Female Gametophyte Gene Expression Map Reveals Similarities between Plnat and Animal Gametes. *Current Biology*, 20:506-512
- XIE, Z., ALLEN, E., WILKEN, A., CARRINGTON, J.C. (2005) DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *PNAS*, 102(36):12984-9
- XU, R., WANG, Y., ZHENG, Y., LU, W., WU, C., HUANG, J., YAN, K., YANG, G., ZHENG, C. (2015) Salt-induced transcription factor MYB74 is regulated by the RNA-directed DNA methylation pathway in *Arabidopsis*. *Journal of Experimental Botany*, doi:10.1093
- YANG, Z., EBRIGHT, Y.W., YU, B., CHEN, X. (2006) HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2'OH of the 3' terminal nucleotide. *Nucleic Acid Research*, 34(2): 667-675
- YAMASHITA, R.A., ZHANG, N., ZHENG, C., BRYANT, S.H. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acid Res*; 39:225-9
- YAO, Y., BILICHAK, A., GOLUBOV, A., BLEVINS, T., KOVALCHIK, I. (2010) Differential sensitivity of *Arabidopsis* siRNA biogenesis mutants to genotoxic stress. *Plant Cell Reports*, 29:1401-1410
- YE, R., WANG, W., IKI, T., LIU, C., WU, Y., ISHIKAWA, M., ZHOU, X., QI, Y. (2012) Cytoplasm Assembly and Selective

Nuclear Import of *Arabidopsis* ARGONAUTE4/siRNA Complexes. *Molecular Cell*, 46: 859-870

YOSHIKAWA, M., PERAGINE, A., PARK, M.Y., POETHING, S. (2005) A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.*, 19(18):2164-2175

YU, H-J., HOGAN, P., SANUDARESAN, V. (2005) Analysis of the Female Gametophyte Transcriptome of *Arabidopsis* by Comparative Expression Profiling. *Plant Physiology*, 139:1853-1869

YU, B., CHAPMAN, E.J., YANG, Z., CARRINGTON, J.C., CHEN, X. (2006) Transgenically expressed viral RNA silencing suppressors interfere with microRNA methylation in *Arabidopsis*. *FEBS Letters*, 580: 3117-3120

YU, A., LEPERE, G., JAY, F., WANG, J., BAPAUME, L., WANG, Y., ABRAHAM, A-L., PENTERMAN, J., FISHER, R.L., VOINNET, O., NAVARRO, L. (2013) Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *PNAS*, 110(2):2389-2394

YUAN, Y.R., PEI, Y., MA, J.B., KURYAVYI, V., ZHADINA, M., MEISTER, G., CHEN, H.Y., DAUTER, Z., TUSCHL, T., PATEL, D.J. (2005) Crystal Structure of *A. aeolicus* Argonaute, a Site-Specific DNA-Guided Endoribonuclease, Provides Insights into RISC-Mediated mRNA Cleavage. *Molecular Cell*, 19(3):405-419

ZHANG, H., XIA, R., MEYERS, B.C., WALBOT, V. (2015) Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Current Opinion in Plant Biology*, 27:84-90

ZHAI, J., BISCHOR, S., WANG, H., FENG, S., LEE, TZUU-FEN, TENG, C., CHEN, X., PARK, SOO YOUNG, LIU, L., GALLEGO-BARTOLOME, J., LIU, W., HENDERSON, I.R., MEYERS, B.C., AUSIN, I., JACOBSEN, S.E. (2015) A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis. *Cell*. 163:445-455

ZHENG, Q., ROWLEY, M.J., BÖHMDORFER, G., SANDHU, D., GREGORY, B.D., WIERZBICKI, A.T. (2013) RNA polymerase V targets transcriptional silencing components to promoters of protein-coding genes. *The Plant Journal*, 73:179-189

ZHENG, X., ZHU, J., KAPOOR, A., ZHU, J-K. (2007) Role of *Arabidopsis* AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *The EMBO Journal*, 26:1691-1701

ZHU, H., HU, F., WANG, R., ZHOU, X., SZE, AH. , LIOU, L.W., BAREFOOT, A., DICKMAN, M., ZHANG, X. (2011) *Arabidopsis* Argonaute10 Specifically Sequesters miR166/165 to Regulate Shoot Apical Meristem Development. *Cell*, 145:242-256

ZILBERMAN, D., CAO, X., JACOBSEN, S.E. (2003) ARGONAUTE4 Control of Specific siRNA Accumulation and DNA and Histone Methylation. *Science*, 299: 716-719

ZILBERMAN, D., CAO, X., JOHANSEN, L.K., XIE, Z., CARRINGTON, J.C., JACOBSEN, S.E. (2004) Role of *Arabidopsis ARGONAUTE4* in RNA-directed DNA Methylation Triggered Inverted Repeats. *Current Biology*, 14: 1214-1220

ZIELEZINSKI, A. & KARLOWSKI, W.M. (2011) Agos-a universal web tool for GW Agonate-binding domain prediction. *Bioinformatics*. 27(9)-1318-9



# APÉNDICES

**APÉNDICE 1.** Listado de secuencias utilizadas en la reconstrucción filogenética para la identificación del homólogo de *AGO9* en *M. polymorpha*.

Nombre	Identificador TAIR
Argonauta 1 (AGO1)	AT1G48410.2
Argonauta 2 (AGO2)	AT1G31280.1
Argonauta 3 (AGO3)	AT1G31290
Argonauta 4 (AGO4)	AT2G27040
Argonauta 5 (AGO5)	AT2G27880
Argonauta 6 (AGO6)	AT2G32940
Argonauta 7 (AGO7)	AT1G69440
Argonauta 8 (AGO8)	AT5G21030
Argonauta 9 (AGO9)	AT5G21150
Argonauta 10 (AGO10)	AT5G43810
DmeAGO401	---

**APÉNDICE 2.** Listado de especies utilizadas en la reconstrucción filogenética de los parálogos de *AGO9* en las embriofitas.

Especie	Nombre
<i>Arabidopsis thaliana</i>	AGO9
<i>Arabidopsis lyrata</i>	AGO21205
<i>Oryza sativa</i>	AGO703
<i>Zea mays</i>	AGO105
<i>Sorghum bicolor</i>	AGO2608
<i>Medicago truncatula</i>	AGO1302
<i>Glycine max</i>	AGO1217
<i>Triticum aestivum</i>	AGO802
<i>Picea glauca</i>	AGO7501
<i>Selaginella moellendorffii</i>	AGO1604
<i>Physcomitrella patens</i>	AGO1505
<i>Marchantia polymorpha</i>	AGO10998
<i>Drosophila melanogaster</i>	AGO401

**APÉNDICE 3.** Listado de oligonucleótidos utilizados

Nombre	Secuencia 5' → 3'	Tamaño	Tm	% GC
MpAGO9_3UTR_F	GGCTTCGTGGAGAGTAATGGAG	22	64	55
MpAGO9_3UTR_R	CCGCAGATTTTACACCGAACTTGA	24	63	43
MpEF1_F	ATGCATCTCGACGGACTTGACCTC	24	59	54
MpEF1_R	TTCAAGTACGCCTGGGTGCTCGACC	23	62	60
MpAGO9_full_F	ATGATGAAACTGTACGAAATCTATG	25	52	32
MpAGO9_full_R	TCAAGCATAGAACATTTTCTTGTTTC	25	53	32
MpAGO9_COOH_pDONR221_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGATGAAACTGTAC-GAAATCTATG	56	53	38
MpAGO9_COOH_pDONR221_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTAGCATAGAA-CATTTTCTTGTTTC	52	52	42
MpNRPE1_full_F	ATGCTGGTGTGCGTGAAAAATC	23	58	43
MpNRPE1_full_R	CTAGAACGATGGCGGTTTCG	19	56	58
MpNRPE1_COOH_pDONR221_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGCTGGTGTGCGT-GAAAAATC	54	59	43
MpNRPE1_COOH_pDONR221_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTGAACGATGGCGGTTTC-GGGCTC	51	64	57
AtAGO4_full_F	ATGGATTCAACAAATGGTAACGGAGC	26	59	42
AtAGO4_full_R	TTAACAGAAGAACATGGAGTTGGC	24	56	42
AtAGO4_COOH_pDONR221_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGGATTCAACAAAT-GGTAACGGAGC	57	60	44
AtAGO4_COOH_pDONR221_R	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGGATTCAACAAAT-GGTAACGGAGC	51	58	51
AtNRPE1_full_F	ATGGAGGAAGAATCTACATCAGAGATTC	28	57	39
AtNRPE1_full_R	TTATGTCTGCGTCTGGGACGG	21	60	57
AtNRPE1_COOH_pDONR221_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGGAGGAAGAATC-TACATCAGAGATTC	59	58	42
AtNRPE1_COOH_pDONR221_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTGTCTGCGTCTGG-GACGG	48	57	58
M13_F	CGCCAGGGTTTTCCAGTCACGAC	23	53	48
M13_R	TCACACAGGAAACAGCTATGAC	22	50	45
C-YFP_R	TGCCGTTCTTCTGCTTGT	18	55	50
N-YFP_R	TCGGCGCGGGTCTTGTAGTT	20	62	60
AtNRPE1_inner_F	GAGAGAAGCCAGTGGGGAA	19	57	58

**APÉNDICE 4.** Protocolo de extracción de RNA total TRI Reagent (SIGMA-ALDRICH)

- Colocar tejido en un tubo Eppendorf de 1.5 mL y adicionar 1 mL de TRI Reagent. Homogenizar utilizando vórtex.
- Centrifugar a 13,400 rpm durante 10 minutos a 4°C.
- Transferir la fase acuosa a un tubo nuevo.
- Adicionar 200 µL de Cloroformo por mL de TRI Reagent. Mezclar con vórtex por 15 segundos y dejar reposar a temperatura ambiente por 15 minutos.
- Centrifugar a 13,400 rpm durante 15 minutos a 4°C.
- Recuperar la fase acuosa y transferir a un tubo nuevo.

**APÉNDICE 4.** Protocolo de extracción de RNA total TRI Reagent (SIGMA-ALDRICH)

- Colocar tejido en un tubo Eppendorf de 1.5 mL y adicionar 1 mL de TRI Reagent. Homogenizar utilizando vórtex.
- Centrifugar a 13,400 rpm durante 10 minutos a 4°C.
- Transferir la fase acuosa a un tubo nuevo.
- Adicionar 200 µL de Cloroformo por mL de TRI Reagent. Mezclar con vórtex por 15 segundos y dejar reposar a temperatura ambiente por 15 minutos.
- Centrifugar a 13,400 rpm durante 15 minutos a 4°C.
- Recuperar la fase acuosa y transferir a un tubo nuevo.
- Adicionar 700 µL de Isopropanol. Dejar reposar a -80°C durante 10 minutos.
- Centrifugar a 13,400 rpm durante 10 minutos a 4°C.
- Remover el sobrenadante y adicionar 1 mL de Etanol al 75%. Mezclar por inversión y centrifugar a 13, 400 rpm durante 5 minutos a 4°C.
- Remover el sobrenadante y dejar secar la pastilla al aire.
- Resuspender la pastilla en agua libre de nucleasas (10 – 30 µL).

**APÉNDICE 5.** Lavado de RNA total con Fenol ácido.

- Llevar el volumen final de la muestra a 300 µL con agua libre de nucleasas.
- Adicionar 300 µL de Fenol ácido (Ambion, cat. No. AM9720). Homogenizar con vórtex durante 30 segundos.
- Adicionar 300 µL de Cloroformo. Homogenizar con córtex durante 30 segundos.
- Incubar en hielo durante 20 minutos.
- Centrifugar a 14,000 rpm por 20 minutos a 4°C.
- Recuperar la fase superior. Calcular el volumen recuperado con la micropipeta.
- Adicionar 1/10 del volumen recuperado de Acetato de sodio 3M, pH 5.
- Adicionar 3 volúmenes de Etanol absoluto grado reactivo.
- Incubar a -80°C durante 30 minutos.
- Centrifugar a 14,000 rpm durante 10 minutos a 4°C.
- Lavar la pastilla con 500 µL de Etanol al 75%.
- Centrifugar a 14,000 rpm por 10 minutos a 4°C.
- Descartar el sobrenadante cuidadosamente con la micropipeta y dejar secar la pastilla al aire.
- Resuspender la pastilla en agua libre de nucleasas (10 – 30 µL).

**APÉNDICE 6.** Síntesis de cDNA con el kit RevertAid RT (Thermo Scientific)

- Adicionar los siguientes componentes en un tubo estéril y libre de nucleasas (mantener en hielo): RNA total (0.1 ng-5 µg), 1 µL de Oligo dT, agua libre de nucleasas para llevar el volumen a 12 µL.
- Incubar a 65°C durante 5 minutos. Enfriar con hielo.
- Adicionar los siguientes componentes a la reacción anterior: 4 µL de 5x Reaction Buffer, 1 µL de Ribolock RNase Inhibitor, 2 µL de 10 mM dNTP Mix y 1 µL de RevertAid RT.
- Mezclar y centrifugar brevemente.
- Incubar a 42°C durante 60 minutos.
- Detener la reacción calentando a 70°C durante 5 minutos.

**APÉNDICE 7.** Extracción de DNA con buffer CTAB

- Colectar entre 200 y 300 mg de tejido, congelar en nitrógeno líquido.
- Adicionar 500 µL de Buffer CTAB (1 M TRIS, 0.5 M EDTA pH 8, 5 M NaCl, PVP 40) y moler los tejidos hasta formar una pasta fina.
- Incubar la mezcla a 56°C durante 15 minutos.
- Centrifugar a 13,000 rpm durante 5 minutos. Transferir el sobrenadante a un tubo nuevo.
- Adicionar 250 µL de una solución de fenol: cloroformo: alcohol isoamílico (24:24:1).
- Centrifugar a 13,000 rpm durante 5 minutos.
- Adicionar 50 µL de acetato de sodio 7.5 M, seguido de 500 µL de etanol frío. Mezclar por inversión.
- Centrifugar a 13,000 rpm durante 3 minutos. Descartar el sobrenadante.
- Lavar la pastilla con 500 µL de etanol al 75%.
- Centrifugar a 13,000 rpm durante 3 minutos. Descartar el sobrenadante. Dejar secar la pastilla al aire durante 5 minutos.
- Resuspender el DNA en agua libre de nucleasas (30 – 40 µL).

**APÉNDICE 8.** Purificación de productos de PCR con 30% PEG/ 30 mM MgCl<sub>2</sub>

- Adicionar 3 volúmenes de buffer TE pH 8 al volumen final de producto de PCR.
- Adicionar 2 volúmenes de 30% PEG/ 30 mM MgCl<sub>2</sub>
- Homogenizar con vórtex y centrifugar inmediatamente a 13,000 rpm por 15 minutos.
- Remover el sobrenadante.
- Resuspender la pastilla en 30 µL de TE pH 8.

**APÉNDICE 9.** Extracción de plásmidos por lisis alcalina

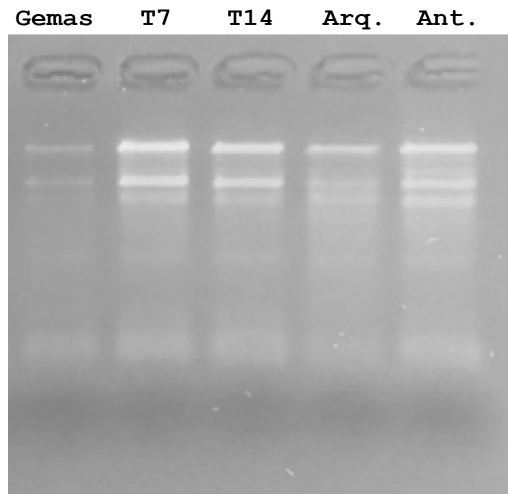
- Resuspender las bacterias en 100  $\mu$ L de Solución de Resuspensión (25 mM Tris-HCl, 10 mM EDTA, 50 mM Glucosa + RNasa) fría.
- Adicionar 200  $\mu$ L de Solución de Lisis (0.2 N NaOH, 1% SDS) fresca. Mezclar por inversión.
- Adicionar 150  $\mu$ L de Solución de Neutralización (5 M Acetato de potasio, Ácido acético glacial) fría. Mezclar por inversión. Incubar por 3 minutos a temperatura ambiente.
- Centrifugar a 12,000 rpm durante 3 minutos. Transferir el sobrenadante a un tubo nuevo.
- Adicionar 450  $\mu$ L de Isopropanol para precipitar el DNA plasmídico. Mezclar por inversión. Incubar por 1 minuto a temperatura ambiente.
- Centrifugar a 14,000 rpm por 1 minuto. Descartar el sobrenadante.
- Adicionar 700  $\mu$ L de Etanol al 70%. Mezclar por inversión.
- Centrifugar a 14,000 rpm por 1 minuto. Descartar el sobrenadante.
- Dejar secar la pastilla al aire durante 5-10 minutos.
- Resuspender la pastilla en agua libre de nucleasas (25-30  $\mu$ L).

**APÉNDICE 10.** Purificación de productos de PCR con el QIAquick Gel Extraction Kit (Qiagen).

- Cortar la banda que corresponde al producto de PCR directamente del gel con un escalpelo nuevo.
- Pesar el pedazo de gel en un tubo. Adicionar 3 volúmenes de buffer QG a 1 volumen de gel.
- Incubar a 50°C durante 10 minutos hasta que la agarosa se disuelva completamente.
- Adicionar 1 volumen de Isopropanol a la mezcla.
- Colocar la muestra en una columna y centrifugar a 13,000 rpm durante 1 minuto. Desechar el sobrenadante.
- Adicionar 500  $\mu$ L de buffer QG y centrifugar por 1 minuto.
- Centrifugar nuevamente bajo las mismas condiciones.
- Adicionar a la columna 750  $\mu$ L de buffer PE y centrifugar. Desechar el sobrenadante. Dejar reposar la columna durante 3 minutos a temperatura ambiente.
- Centrifugar nuevamente bajo las mismas condiciones.
- Colocar la columna en un tubo nuevo.
- Adicionar 20-30  $\mu$ L de agua libre de nucleasas justo en el centro de la columna y centrifugar.

# ANEXOS

**ANEXO 1.** RNA total de *Marchantia polymorpha*. Electroforesis en agarosa al 2% teñida con Bromuro de Etidio. Se muestran las muestras de RNA de diferentes tejidos, gemas, talos de 7 (T7) y de 14 días (T14), Arquegonióforos (Arq.) y Anteridióforos (Ant.).



**ANEXO 2.** Resultados obtenidos de la secuenciación de los vectores de destino AtAGO4:nYFP, MpAGO9:nYFP y AtNRPE1:cYFP.

Diagrama de los resultados obtenidos de la secuenciación de la construcción **AtAGO4:nYFP**. En azul se muestra un fragmento de los que corresponde a AtAGO4, en azul fuerte el sitio de recombinación generado por la reacción LR, en lila el epítipo Myc, en amarillo la secuencia que corresponde a la nYFP y en verde el marco abierto de lectura de la fusión. La construcción se envió a secuenciar con el oligo reverso de la nYFP.

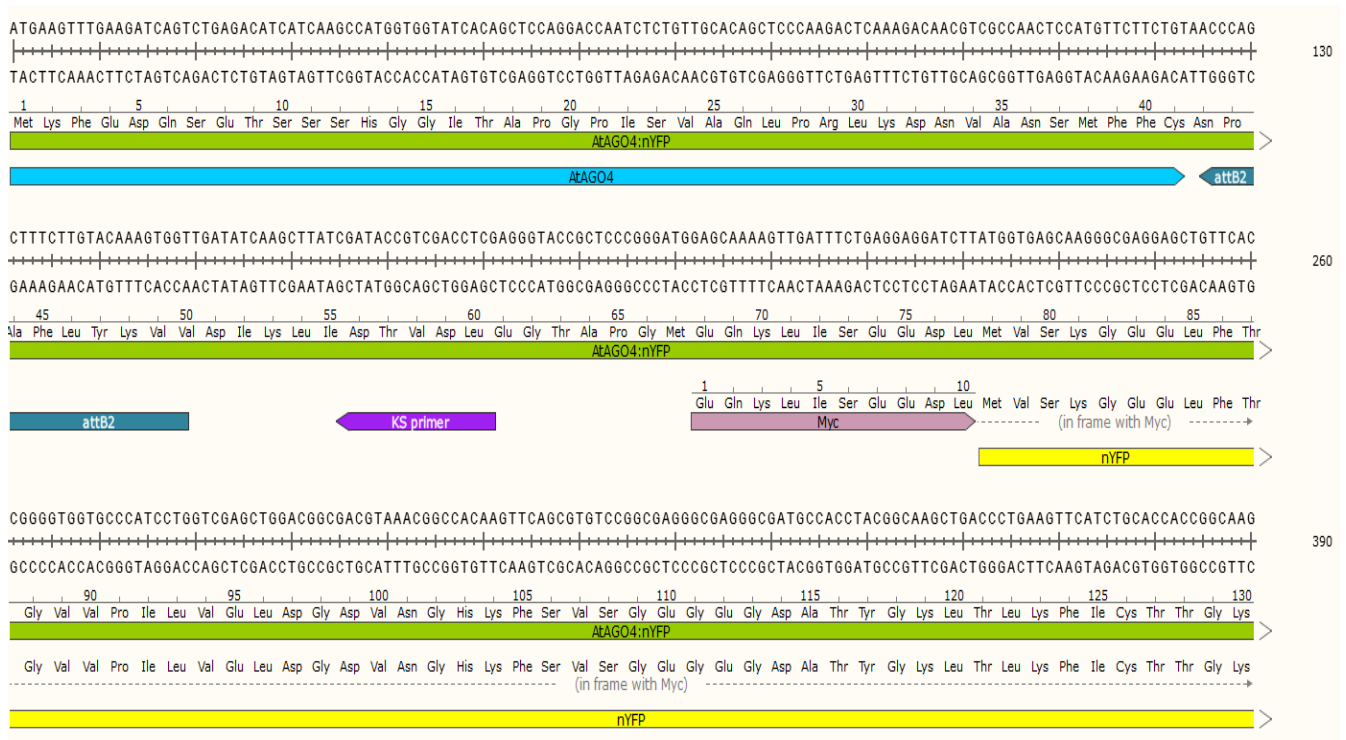


Diagrama de los resultados obtenidos de la secuenciación de la construcción **MpAGO9:nYFP**. En azul se muestra un fragmento de MpAGO9, en azul fuerte el sitio de recombinación formado por la reacción LR, en lila el epítipo de Myc, en amarillo la secuencia que corresponde a nYFP y en verde el marco abierto de lectura de la fusión.

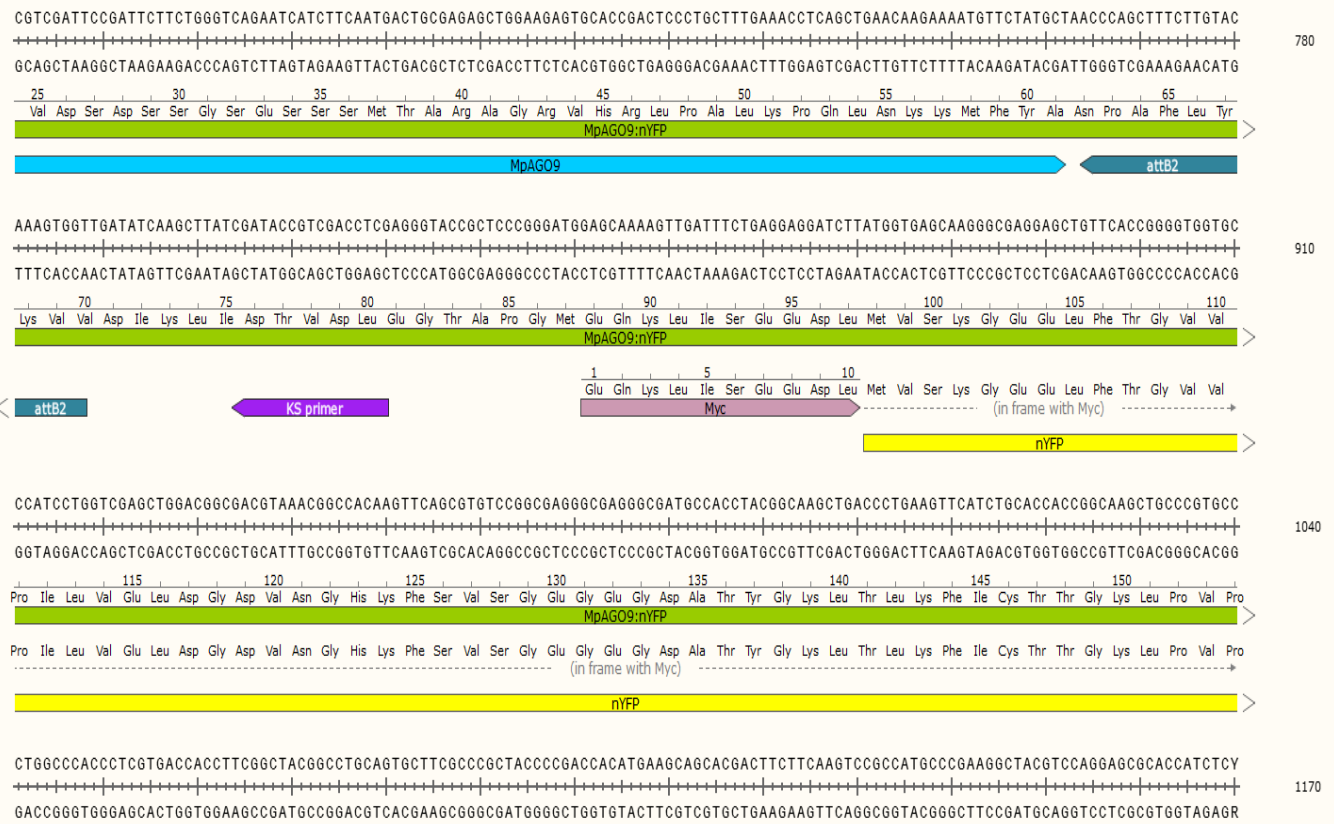


Diagrama de los resultados obtenidos de la secuenciación de la construcción **AtNRPE1:cYFP**. En morado se muestra un fragmento de AtNRPE1, en azul fuerte el sitio de recombinación formado por la reacción LR. En verde se muestra el marco abierto de lectura de la fusión. Debido al sitio de hibridación del oligonucleótido reverso cYFP en la secuenciación no se logró obtener la secuencia del epítipo ni de la cYFP (era un sitio muy cercano y hubo errores normales al inicio de la secuenciación). La secuencia obtenida, sin embargo, al ser comparada con otras construcciones secuenciadas en el vector pSPYCE es igual.

