



FACULTAD DE CIENCIAS BIOQUÍMICAS Y FARMACÉUTICAS

UNIVERSIDAD NACIONAL DE ROSARIO

Tesis de Doctorado

“Bases moleculares de la biogénesis de microARNs en plantas”

Presentada por

Lic. Belén Moro

Director: Dr. Javier F. Palatnik

Co-director: Dr. Rodolfo M. Rasia

Rosario, Santa Fe, Argentina

2017

“Bases moleculares de la biogénesis de microARNs en plantas”

Belén Moro

Licenciada en Biotecnología

Facultad de Ciencias Bioquímicas y Farmacéuticas

Universidad Nacional de Rosario

Esta Tesis es presentada como parte de los requisitos para optar al grado académico de Doctor en Ciencias Biológicas, de la Universidad Nacional de Rosario y no ha sido presentada previamente para la obtención de otro título en esta u otra Universidad. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el Instituto de Biología Molecular y Celular de Rosario (IBR-CONICET), dependiente de la Facultad de Cs. Bioquímicas y Farmacéuticas, durante el período comprendido entre Abril de 2012 y Febrero de 2017, bajo la dirección del Dr. Javier F. Palatnik y la co-dirección del Dr. Rodolfo M. Rasia.

Parte del trabajo realizado en esta tesis ha sido publicado en los siguientes artículos y congresos:

Publicaciones científicas.

Schapiro AL, Bologna NG, Moro B, Zhai J, Meyers BC, Palatnik JF (2014). Construction of Specific Parallel Amplification of RNA Ends (SPARE) libraries for the systematic identification of plant microRNA processing intermediates. **Methods**.67: 36-44

doi: 10.1016/j.ymeth.2014.04.001

Presentaciones a congresos.

Nicolás G. Bologna, Belén Moro, Javier F. Palatnik. "Characterization of the loop-to-base processing mechanism responsible for the biogenesis of plant miRNAs". **XLVI Reunión anual de la SAIB, Puerto Madryn, Noviembre, 2010**

Belén Moro, Nicolas G. Bologna, Arnaldo Schapiro, Siwaret Arikít, Jixian Zhai, Blake Meyers and Javier Palatnik. "RNA recognition during microRNA processing in plants" **RNA silencing in plants (G1) Keystone symposia, Keystone, Colorado, USA, February, 2015.**

Belén Moro, Nicolas Bologna, Arnaldo Schapiro, Uciel Chorostecki, Siwaret Arikít, Jixian Zhai, Blake Meyers and Javier Palatnik. "Analysis of microRNA-processing mechanisms in plants using a genome-wide approach". **11th International Plant Molecular Biology Congress, Iguazú Falls, Brazil, October, 2015**

Rojas A, Bresso E, Schapiro A, Moro B, Mateos J, Palatnik J. Instituto de Biología Molecular y Celular de Rosario (IBR). "Importance of the precursor primary and secondary structure during microRNA processing in plants". **LII Reunión anual de la SAIB, Cordoba, Noviembre, 2016.**

AGRADECIMIENTOS

A CONICET por el apoyo económico durante estos cinco años de doctorado.

A Javier, que me ha dado más oportunidades de las que puedo contar, por su inagotable capacidad de laburo y su proporcional paciencia. No es fácil trabajar conmigo.

A Fito, por su generosidad y por la paradoja de Levinthal. I mean it!

A Flor, que hace siete años nos dimos la mano, y nunca nos la soltamos.

A Uci, cuando sea grande quiero ser como vos.

A Ara, la mejor “wingwoman” que alguien pueda querer, para hacer experimentos o ir de copas.

A Camala, por haberse convertido en parte invariable de nuestro grupo, y en mi reidora personal. Es la que compró el megáfono.

A Cabito por su energía y entusiasmo de todos los días.

A Paula por su inocencia y su bondad.

A Juli por ser la botánica entre los biólogos moleculares, la Plantipedia.

A Mati y a Rodri, que hicieron un doctorado en calma zen.

A Juan y Edgard, que me criaron en segundas nupcias. Sin ustedes nada de esto hubiera sido posible.

A Nico, que me metió en este baile, y nunca estuvo a más de un email de distancia. Fue un honor pipetear con los tres.

A Carla, no se me ocurre alguien más con quien compartir la mesada. Vamos a Toulouse?.

A Rama que es una fuente inagotable de recursos e historias del Lab.

A Arni que iluminó en el camino de las colonys, y juntos descubrimos como no estudiar el procesamiento del 172.

A Santi que se contagió de la magia de lo que hacemos, y de por qué vale la pena cambiar una AxC.

A mis laboratorios adoptivos, el 1 y el 5: Andrés, Anto, Bruno, Florcita, Guille, Irina, Lisandro, Marco, Marcos, Maru, Nahuel, Paula y Robi. Mis amigos, que me han recibido a cualquier hora, me han alimentado y me han acompañado a tomar cerveza. No sólo de ciencia vive el hombre.

Sí los puse en orden alfabético porque no quiero quejas. Excepto Irina, que es mi favorita! Por su alegría y su entrenamiento liberal en el arte de los geles de secuenciación.

A las chicas de lavado, Clau, Vivi, Sil, Lilo y a Diego. Su trabajo y buena predisposición hizo que el mío fuera posible. Gracias!

I would like to thank Dr. Blake Meyer for welcoming me into his Lab. And to the friends and colleagues I made there, Atul, Kun, Tzuu-fen and specially Tow who made laugh every day and taught how to make decent RNA libraries. Plus, they took me shopping.

A Lud y Yor, su amistad le da sentido a mi vida. Las amo.

A mi familia de los Jueves, Marcel, Mati, Yor y Anyu. La terapia después de la terapia, no imagino mi vida sin ustedes. Especialmente Anyu, the Little bit, que más quiero en este mundo.

A mis amigos del Poli, de la segunda y de la quinta, algunas cosas nunca cambian. Agus, María, Lucio, Nahuel, Sabi y Jus.

A los alquímicos, que enriquecieron mis años del Doc, por todas las horas que pasamos en comisión, ordenando el Lab del subsuelo, alguien dijo integrador?

A la banda del 555 por todas las cervezas, las patas y las carcajadas que compartimos juntos, especialmente a los infiltrados del IBR, que me tuvieron el placer de mi compañía por partida doble y por decirlo de algún modo.

A mi Familia genética y fenotípica por escuchar tantas veces (y preguntarme!) que es un precursor de miARN. Por entusiasmarse con la ciencia, o al menos con lo que yo hago.

ÍNDICE GENERAL.

ÍNDICE GENERAL.....	1
1. RESUMEN.....	5
2. ABREVIATURAS.....	6
INTRODUCCIÓN.....	9
3. Introducción.....	10
3.1. Silenciamiento génico y ARN pequeños en plantas.....	10
3.1.1. Clases de ARN pequeños en plantas.....	11
3.2. MicroARNs en plantas.....	13
3.3. Genes de miARNs y conservación evolutiva.....	14
3.4. Biogénesis de miARN.....	16
3.4.1. Procesamiento de miARNs en animales.....	16
3.4.2. Procesamiento de miARNs en plantas.....	18
3.5. Estructuras secundarias y mecanismos de procesamiento en animales y plantas.....	20
3.6. Proteínas involucradas en la biogénesis de ARNs pequeños.....	23
3.6.1. ARNasas tipo III.....	23
3.6.2. HYL1.....	26
3.6.3. SE.....	27
3.6.4. Interacción de otras proteínas con DCL1, HYL1 y SE.....	27
3.6.5. HEN1.....	28
3.6.7. ARGONAUTA.....	28
3.7. Regulación de la expresión génica por miARNs.....	31
3.7.1. Regulación por corte del ARN blanco.....	31
3.7.2. Regulación de la traducción por miARNs.....	32
3.7.3. Desestabilización del ARNm.....	33
3.8. Funciones de miARNs en plantas.....	33
4. OBJETIVOS.....	36
4.1. Objetivo general.....	36
4.2. Objetivos particulares.....	36
MATERIALES Y METODOS.....	37
5. Materiales y Métodos.....	38
5.1. Metodología para el clonado de fragmentos de ADN.....	38
5.1.1. Cepas bacterianas empleadas.....	38
5.1.2. Vectores plasmídicos utilizados.....	38

5.1.3. Minipreparaciones de ADN plasmídico.	38
5.1.4. Digestión de ADN con endonucleasas de restricción.	39
5.1.5. Electroforesis de ADN en geles de agarosa.	39
5.1.6. Purificación de fragmentos de ADN a partir de geles de agarosa.	39
5.1.7. Ligación de fragmentos de ADN.	39
5.1.8. Transformación bacteriana.	40
5.1.9. Selección y análisis de transformantes.	41
5.2. Mutagénesis sitio dirigida.	41
5.2.1. Reacción en Cadena de la Polimerasa (PCR)	41
5.2.2. Clonado de los fragmentos mutantes	42
5.3. Obtención de <i>A. thaliana</i> transformadas.	42
5.3.1. Ecotipo utilizado y condiciones de crecimiento.	42
5.3.2. Transformación de <i>A. thaliana</i>	42
5.3.3. Recolección de semillas.	43
5.3.4. Esterilización de semillas.	43
5.3.5. Selección de transformantes.	43
5.4. Blot de ARN pequeños.	43
5.4.1. Extracción de ARN.	43
5.4.2. Cuantificación del ARN purificado espectrofotométricamente.	44
5.4.3. Chequeo de la integridad del ARN purificado.	44
5.4.4. Armado del gel y corrida electroforética.	44
5.4.5. Marcado radioactivo de la sonda	45
5.4.6. Hibridación.	46
5.4.7. Detección radioactiva	46
5.5. Determinación de intermediarios.	46
5.5.1. Construcción biblioteca 5' RACE PCR.	46
5.5.2. Amplificación de intermediarios.	48
5.6. Optimización del método SPARE para análisis genómico de intermediarios de procesamiento de microARNs	48
5.6.1. Material vegetal.	48
5.6.2. Protocolo SPARE multiplex.	48
5.6.3. Puesta a punto de las reacciones de PCR.	52
5.6.4. Construcción de marcadores de peso molecular específicos.	62
5.6.5. Construcción de bibliotecas con las condiciones finales de trabajo.	63
5.7. Procesamiento de los datos crudos.	67
5.7.1. Archivo final con los resultados de las bibliotecas.	67

5.7.2. Representación gráfica de los cortes en cada precursor.	68
5.8. Transcripción <i>in vitro</i>	70
5.8.1. Preparación de muestras de ARN.....	70
5.9. Mapeos estructurales sobre el ARN	72
5.9.1. Marcado de muestras de ARN con ³² P en el extremo 5'	72
5.9.2. Mapeo por digestión enzimática	72
5.9.3. Geles de Secuenciación	73
RESULTADOS I.....	75
6. Resultados I.	76
6.1. INTRODUCCION I	76
6.2. Modificaciones realizadas para la construcción de bibliotecas multiplex en plantas silvestres.....	78
6.3. Análisis global de los resultados obtenidos en plántulas de plantas silvestres y mutantes <i>fiery1</i>	79
6.4. Identificación de la dirección de procesamiento de los precursores de miARNs.	81
6.4.1. Caracterización de precursores procesados desde la base con dos cortes de DCL1	81
6.4.2. Caracterización de precursores procesados desde la base con tres cortes de DCL1.....	83
6.4.3. Caracterización de precursores procesados desde el loop hacia la base con dos cortes de DCL1	85
6.5. Detección de cortes abortivos en precursores de miARN <i>in vivo</i>	86
6.6. Mecanismos de procesamiento caracterizados por SPARE.	89
6.7. DISCUSION I	91
RESULTADOS II.....	94
7. Resultados II	95
7.1. INTRODUCCIÓN II	95
7.2. Los precursores que se procesan desde el loop terminal pueden clasificarse en distintos grupos.....	96
7.3. Caracterización molecular de los precursores procesados en dirección loop a base con región terminal homogénea.	97
7.3.1. Identificación de los determinantes estructurales del procesamiento de miR171b.	98
7.4. Caracterización de los precursores procesados en dirección loop a base que presentan un loop terminal ramificado.	99
7.4.1. Mecanismos de procesamiento de la familia de miR157.....	99
7.4.2. Mapeo estructural de la región terminal del precursor de miR157c por digestión enzimática.	102
7.4.3. Caracterización de los determinantes estructurales de la biogénesis de miR157c.	104
7.5. Análisis global de los precursores procesados con dos cortes de DCL1.....	107
7.6. DISCUSION II	111
7.6.1. Requisitos estructurales para el procesamiento desde el loop a la base de precursores con región terminal homogénea.	111
7.6.2. Comparación entre los mecanismos de procesamiento de loop a base y base a loop de dos cortes.	112

7.6.3. Biogénesis a partir de un loop terminal ramificado	114
RESULTADOS III	116
8. Resultados III	117
8.1. INTRODUCCIÓN III	117
8.2. Conservación de miR319 y el origen de sus ARNs pequeños.....	118
8.3. Inserciones en el tallo superior de miR319a alteran el patrón de cortes de DCL1 en el precursor... 121	
8.4. El loop interno presente en la región central de miR319a previene la acumulación de otros ARN pequeños.....	125
8.5. Mejorar el apareamiento del dúplex de miR319a.2 genera DCL1 se detenga.....	127
8.6. El grado de apareamiento de los dúplex de ARN puede cuantificarse mediante su energía libre de interacción y difiere a lo largo del precursor de miR319.	130
8.7. Un modelo global de los requisitos para la acción secuencial de DCL1 en Arabidopsis	131
8.6. Mutaciones puntuales en miR319.2 permiten su acumulación sin alterar la cantidad de miR319... 132	
8.8. DISCUSIÓN III	135
8.8.1. En plantas estructura secundaria de los dúplex de ARN afectan tanto la acumulación del ARN pequeño como la procesividad de DCL1.	135
8.8.2. Relación entre la evolución de los miARNs y el procesamiento secuencial de los precursores.	136
9. CONCLUSIONES.....	140
10. ANEXO.....	142
11. BIBLIOGRAFÍA	189

1. RESUMEN.

Los microARNs (miARN) son ARN pequeños de ~20-22nt, de codificación endógena, que regulan múltiples rutas biológicas a través de la modulación de la expresión génica. En plantas controlan el desarrollo, la señalización hormonal y la respuesta al estrés. En humanos, se considera que el 60% de los genes están regulados por miARNs. Los miARNs reconocen a sus genes blancos por complementariedad de bases, y los guían a su degradación o inactivación traduccional. La especificidad de la interacción miARN-ARN mensajero (ARNm) ha permitido el diseño de miARNs artificiales para silenciar genes de interés.

Los miARNs son procesados a partir de largos precursores con extensa estructura secundaria. Estos precursores son cortados por ribonucleasas del tipo III, como DICER LIKE1 (DCL1) en plantas. DCL1 realiza al menos dos cortes sobre el ARN doble hebra (ARNdh) y libera al miARN maduro, de ~21nt, junto a la hebra complementaria del mismo, denominada miARN*. DCL1 es asistida por otras proteínas, como la proteína de unión a ARN doble hebra, HYPONATIC LEAVES1 (HYL1) y la proteína de dedos de zinc SERRATE (SE).

Tanto en plantas como en animales el precursor del miARN contiene pistas espaciales que determinan la posición del miARN a lo largo de su secuencia. Mientras que los precursores de animales presentan un tamaño uniforme, los precursores de plantas representan una colección de tamaños y formas variables que pueden ser procesados mediante cuatro mecanismos diferentes. La biogénesis de los miARNs es un proceso clave porque determina la secuencia exacta de nucleótidos del ARN pequeño funcional y por ende los genes regulados por él. Para responder a este interrogante desde un punto de vista genómico hemos desarrollado una técnica denominada SPARE (del inglés, "Specific Parallel Amplification of RNA Ends) que permite analizar los intermediarios de procesamiento de todos los precursores de *Arabidopsis thaliana*.

Los resultados obtenidos mediante la técnica de SPARE en conjunto con mutagénesis dirigida de los precursores de miR171, miR156/miR157 y miR319 nos ha permitido caracterizar mecanismos de procesamiento exclusivos de plantas que implican el reconocimiento de la porción apical del precursor y la actividad de corte sucesiva de DCL1. A su vez el análisis global del procesamiento de miARN, nos llevó a proponer que la biogénesis de miARNs en plantas depende de la dirección de procesamiento del precursor, el número de cortes necesarios para liberar el miARN maduro y el determinante estructural que es reconocido para ubicar la posición del primer corte.

2. ABREVIATURAS.

A	Adenosina.
ADN	Ácido desoxirribonucleico.
ADNasa	ADN nucleasa.
ADNc	ADN complementario.
AGO	proteína de la familia ARGONAUTA.
ARN	Ácido ribonucleico.
ARNasa	ARN nucleasa.
ARNdh	ARN doble hebra.
ARNm	ARN mensajero.
ARNsh	ARN simple hebra.
ATP	Adenosín trifosfato.
C	Citosina.
CBC	Complejo de unión a Cap.
CBP	Proteína de unión a Cap.
Col	ecotipo Col-0 de <i>Arabidopsis thaliana</i> .
DCL	Ribonucleasa DICER-like.
DDL	Dawdle.
DGCR8	Di George syndrome critical región gene 8.
diARNs	ARNs pequeños inducidos por ruptura doble hebra de ADN.
dsRBD	Dominios de unión a ARNdh.
EXP5	Exportin-5.
G	Guanina.
GTP	Guanosín trifofato.
HEN1	Hua Enhancer 1.
HESO1	Supresor de Hua Enhancer 1.
HYL1	Hyponastic leaves 1.
IR-siARNs	ARNs pequeños derivados de repeticiones invertidas.
LB	Medio de cultivo para bacterias Luria-Bertani.

LNA	“Locked nucleic acid ”.
lmiARN	MicroARN largo.
miARN	MicroARN.
min.	Minuto/s.
MOS2	Modificer of snc1 2.
MS	Medio de cultivo para plantas Murashige-Skoog.
nat-miARN	miARNs naturales.
nt	Nucleótidos.
pb	Pares de base.
PCR	Reacción en cadena de la polimerasa.
Pol	ARN polimerasa.
PTGS	silenciamiento génico transcripcional.
rasiARN	ARN pequeños asociados a regiones repetidas.
RDR	ARN polimerasa dependiente de ARN.
RdDM	Metilación del ADN dependiente de ARN.
RISC	Complejo de silenciamiento inducido por ARN.
RT	Retrotranscripción.
SE	Serrate.
seg	segundo/s.
siARN	ARN pequeño interferente.
SIC	Sickle1.
SMC	Sitio de Múltiple Clonado.
STA1	Stabilized 1.
T	Timina.
T1	Líneas transgénicas seleccionada independiente.
tasiARN	ARN pequeño de acción en trans.
T-ADN	Fragmento de ADN plasmídico transferido al genoma nuclear en las transformaciones con <i>Agrobacterium tumefaciens</i> .
TGH	Tough.

TGS	Silenciamiento génico transcripcional.
Tm	temperatura de fusión.
U	Uracilo.
wt	Silvestre.

Palabras tomadas del inglés

Loop

Bulge

mismatch

3. Introducción

3.1. Silenciamiento génico y ARN pequeños en plantas.

El control de la expresión génica es un proceso vital para los organismos multicelulares. Este proceso es utilizado por las células para ajustar el conjunto de genes expresados durante el desarrollo o para responder frente a cambios en su ambiente. En plantas y animales, la generación de los diferentes tipos celulares que forman en conjunto un organismo multicelular depende de que los genes se activen en las células adecuadas y en los momentos precisos del desarrollo.

Muchos de los mecanismos regulatorios operan mediante el control de la expresión a nivel transcripcional y post-transcripcional. En los últimos años moléculas de ARN de 20 a 25 nucleótidos (nt) de longitud han cobrado vital importancia como reguladores de la expresión génica en eucariotas. En plantas existen distintos tipos de ARN pequeños que se diferencian en su biogénesis y modos de acción, lo cual ha permitido clasificarlos.

A pesar de sus diferencias, los mecanismos de silenciamiento génico mediado por ARNs pequeños comparten cuatro pasos comunes y necesarios (i) el procesamiento de un ARNdh largo para liberar un dúplex de ARN, de un tamaño de 18-25 pares de bases (pb) de longitud, (ii) la metilación de ambos extremos 3' del dúplex, (iii) la selección de una de las hebras del dúplex de ARN y su incorporación en el complejo efector que contiene una proteína de la familia ARGONAUTA (AGO) y, por último, (iv) el reconocimiento del ácido nucleico blanco por complementariedad de bases y su silenciamiento (Bologna and Voinnet, 2014; Chapman and Carrington, 2007).

El procesamiento del ARNdh requiere la participación de ribonucleasas clase III denominadas DICER (Bologna and Voinnet, 2014). Las proteínas DICER se encuentran conservadas en eucariotas y están presentes en organismos unicelulares, vertebrados y plantas superiores. En plantas los homólogos de DICER se denominan DCL (del inglés, "*Dicer-Like*"). Luego de que DCL produce la liberación del dúplex de ARN, la ARN metiltransferasa HUA ENHANCER 1 (HEN1) metila el grupo 2'OH de la ribosa en cada extremo 3' del mismo, lo cual lo protege de la degradación por otras ARNasas (Boutet et al., 2003).

A continuación, una de las dos hebras del dúplex es incorporada en AGO y actúa como guía para el complejo RISC (por sus siglas en inglés *RNA induced silencing complex*) (Llave et al., 2002). Este complejo, lleva a cabo la búsqueda del ácido nucleico complementarios al ARN pequeño simple hebra (ARNsh). El silenciamiento génico post transcripcional (PTGS) puede darse por corte endonucleolítico sobre ARNm blanco, desestabilización del ARNm o supresión traduccional. Mientras que el silenciamiento génico transcripcional (TGS) requiere de la metilación de citosinas (C) en el ADN o de las histonas (Bologna and Voinnet, 2014).

3.1.1. Clases de ARN pequeños en plantas.

En *A. thaliana* existen cuatro proteínas DCLs, DCL1 a 4, cada una de las cuales participa en la biogénesis de distintos ARNs pequeños. DCL1 produce ARN pequeños de ~21nt, mientras que DCL2, DCL3 y DCL4 produce ARN pequeños de un tamaño de 22nt, 24nt y 21nt de longitud, respectivamente (Qi et al., 2005). Los ARNs pequeños endógenos que son responsables del PTGS se dividen en cuatro grupos según a sus orígenes, proteínas efectoras asociadas y funciones biológicas. Estos grupos son: (i) los miARNs, (ii) los ARN pequeños de interferencia derivados de repeticiones invertidas (IR-siARNs), (iii) los ARNs pequeños de interferencia de acción en trans (tasiARNs) y (iv) ARN pequeños de interferencia derivados de transcritos antisentido naturales (nat-siARNs) (Bologna and Voinnet, 2014).

Los IR-siARNs provienen de loci discretos dispersados a lo largo del genoma de las plantas. Estas repeticiones invertidas al transcribirse dan lugar moléculas de ARNsh que se pliegan sobre sí mismos dando un ARNdh de elevada complementariedad y longitud variable (Dunoyer et al., 2007; Dunoyer et al., 2010; Kasschau et al., 2007). Estos ARNdh son procesados por DCL4/DCL2, dando siARNs de 21-22nt y DCL3 24nt siARNs. Incluso DCL1 participa de esta vía, al facilitar la separación del ARNdh de la región de ARNsh del transcripto primario de los IR (Dunoyer et al., 2007; Dunoyer et al., 2010). Las funciones biológicas de estos ARNs pequeños no están del todo clara, y se ha propuesto que podrían responder a stress o expresarse en tejidos específicos (Dunoyer et al., 2010).

Los miARNs se procesan a partir ARNdh imperfectos más largos denominados precursores de miARN o pri-miARNs. Los pri-miARN se forman por hibridación intramolecular de un ARNsh, dando como resultado un ARNdh imperfecto. Esta estructura secundaria con forma de horquilla posee con un tallo doble hebra imperfecto y un loop terminal. El precursor es procesado por un complejo que contiene a la proteína DCL1 (Bologna and Voinnet, 2014). Tanto la biogénesis como la estructura secundaria de este grupo de ARNs pequeños serán ampliadas en las Secciones 3.4 y 3.5.

Los tasiARNs son codificados por genes que al transcribirse dan lugar a un ARN no codificante denominados TAS. Este transcripto es reconocido y cortado por acción de un miARN que posee una secuencia complementaria al mismo (Allen et al., 2005). Los fragmentos de corte de del ARN TAS son estabilizados por la proteína SGS3 (del inglés, "supresor of gene silencing 3") y son convertidos en ARNdh por la acción de polimerasa dependiente de ARN, en particular RDR6. Este ARNdh de complementariedad perfecta, es procesado por DCL4, generando ARNs pequeños de 21nt, cuya secuencia está especificada por la fase generada a partir del corte del miARN inicial. Algunos de estos tasiARNs se incorporan a un complejo RISC y guían el corte de los transcritos de otros genes de manera similar a la acción de un miARN (Allen et al., 2005; Gascioli et al., 2005; Xie et al., 2005; Yoshikawa et al., 2005). En Arabidopsis se han descrito ocho loci que codifican para ARNs TAS. TAS1a, TAS1b, TAS1c y TAS2 son cortados por miR173 cargado en AGO1. TAS3a, TAS3b y TAS3c son cortados por miR390 cargado en AGO7 y TAS4 es cortado por miR828

(Allen and Howell, 2010; Cuperus et al., 2010a; Montgomery et al., 2008a; Montgomery et al., 2008b; Peragine et al., 2004)

Los nat-siARNs se diferencian de los anteriores por el origen de su precursor. Estos se originan a partir de un ARNdh formado por la transcripción bidireccional convergente de dos genes parcialmente solapados (Borsani et al., 2005; Katiyar-Agarwal et al., 2006). En los casos reportados hasta el momento uno de los genes solapados presenta una expresión constitutiva, por lo que la inducción de la expresión del gen complementario activa la formación del ARNdh, el cual es procesado por DCL1 o DCL2 liberando nat-siARN primarios que inducen el corte sobre el ARNm expresado constitutivamente (Borsani et al., 2005; Katiyar-Agarwal et al., 2006). Ejemplos de esto incluyen a nat-siRNASRO5 inducido por stress por sal y a nat-siRNAATGB2 el cual sólo se acumula en respuesta a infecciones por bacterias patógenas. (Zhang et al., 2013).

Existe un segundo grupo de ARNs pequeños endógenos los cuales participan del TGS (i) siARNs asociados a regiones repetidas (rasiARNs) (ii) microARNs largos (lmiARNs) (iii) ARNs pequeños inducidos a partir de ruptura doble hebra del ADN (diARNs). La metilación del ADN dependiente de ARN (RdDM) requiere de ARNs pequeños de 24nt generados por DCL3, los rasiARNs. Estos previenen la transcripción de transposones y regiones repetidas a través de la metilación de novo de C mediante las proteínas DOMAINS REARRANGED METHYLTRANSFERASE 1 (DRM1) y DRM2. La biogénesis de estos ARNs pequeños depende dos ARN polimerasas exclusivas de plantas (Pol), Pol IV y Pol V. Estas transcribirían regiones de heterocromatina produciendo largos transcriptos de ARNsh, los cuales serían convertidos a ARNdh mediante la acción de RDR2. Estos ARNdh son procesados por DCL3, metilados por HEN1 y cargados en AGO4, los rasiARNs desencadenarían la RdDM que en última instancia resultaría en TGS en el locus de origen (Bologna and Voinnet, 2014; Matzke et al., 2009).

Trabajos anteriores en arroz y Arabidopsis mostraron que varios pri-miARNs pueden ser procesados por DCL3 para liberar lmiARN de 24nt. Estos son metilados por HEN1 y cargados en AGO4. Luego el complejo AGO4-lmiARN dirige la metilación de C tanto el locus del *MIARN* (cis) como en el locus del gen blanco del miARN (trans) (Chellappan et al., 2010; Wu et al.). Recientemente se ha demostrado que los mecanismos de reparación de rupturas doble hebra del ADN (ADNdh) dan lugar a la producción de ARNs pequeños, lo diARNs. En Arabidopsis lo diARNs de 21 y 24nt provienen tanto de la hebra sentido como antisentido en regiones cercanas a las rupturas de ADNdh (Wei et al., 2012b). Un modelo que continúa siendo especulativo, propone que Pol V generaría ARNsh a partir de regiones proximales de las rupturas de ADNdh luego una RDR convertiría el ARNsh en ARNdh el cual sería procesado por DCL2, DCL3 y DCL4 para dar diARNs. Los complejos AGO2-diARN serían reclutados al sitio de la ruptura de ADNdh donde reclutarían a la maquinaria de reparación de ADN (Wei et al., 2012b).

3.2. MicroARNs en plantas

Los miARNs son ARN pequeños de ~20-22nt, de codificación endógena que regulan otros ARNs, a los que reconocen por complementariedad de bases (Bologna and Voinnet, 2014). En plantas controlan procesos vitales como el desarrollo, señalización hormonal y respuestas al estrés (Ambros et al., 2003; Axtell, 2008; Carthew and Sontheimer, 2009; Cuperus et al., 2011; Song et al., 2010; Voinnet, 2009). En principio, los miARNs pueden controlar la expresión génica de maneras diferentes, ajustando cuantitativamente los niveles finales de las proteínas o bien eliminando los transcritos de ARN en la célula reasegurando la inactivación de ciertos genes.

Tanto en plantas como en animales, los miARNs son generados a partir de loci endógenos (Figura I.1A). Mientras que muchos miARNs de animales provienen de intrones o regiones no traducidas de ARNm codificantes (Carthew and Sontheimer, 2009; Kim et al., 2009), en plantas los genes que codifican para miARNs son unidades transcripcionales independientes (Allen et al., 2004). Al igual que los genes codificantes para proteínas, los genes que codifican miARNs (*MIARNs*) son transcritos por la Pol II en conjunto con co-activadores como el MEDIATOR y NOT2b (Kim et al., 2011; Wang et al., 2013). Luego sufren modificaciones similares a los ARNm codificantes de proteínas: poliadenilación, agregado caperuza y corte y empalme; este último proceso suele ocurrir principalmente hacia la región 3' de los pri-miARN (Schwab et al., 2013; Xie et al., 2005; Zhao and Srivastava, 2007).

Varios factores de splicing contribuyen al procesamiento del pri-miARN, incluyendo a la proteína de dedos de zinc del tipo C2H2 SE y las subunidades CAP-BINDING PROTEIN 80 (CBP80) y CBP20, del complejo de unión a la caperuza (CBC) (Laubinger et al., 2008), todos ellos se asocian al co-activador transcripcional NOT2b (Wang et al., 2013). Se ha propuesto que otras dos proteínas STABILIZED 1 (STA1) y SICKLE (SIC) participan de este proceso ya que *sta1* y *sic* muestran una reducción en la acumulación de miARNs maduros, y acumulan grandes niveles de pri-miARN con retención de intrones (Ben Chaabane et al., 2013; Zhan et al., 2012). Parte de los defectos de estas mutantes podrían deberse en a que el ARNm de DCL1 también requiere de splicing para dar lugar a la proteína funcional (Ben Chaabane et al., 2013)(Figura I.4, Sección 3.4.2).

Los pri-miARN se pliegan y forman una estructura de apareamiento imperfecto de tallo y loop (Figura I.1B). Esta estructura dentro del transcrito primario se conoce como precursor de miARN. Posteriormente el precursor de miARN es sustrato de la proteína DCL1, asociada con algunas proteínas accesorias. Este complejo produce al menos dos cortes sobre la estructura de tallo para liberar al miARN maduro junto a la hebra complementaria del mismo, denominada miARN*.

El dúplex miARN/miARN* es metilado y luego el miARN es incorporado a una proteína AGO1 mientras que el miARN* es degradado (Figura I.1C). Ya en el complejo RISC el miARN guía al complejo hacia una molécula de ARNm complementario. En plantas, usualmente cada ARNm blanco contiene un único sitio

localizado en la región codificante que es complementario al miARN (Figura I.1D)(Llave et al., 2002; Rhoades et al., 2002). Una vez que se reconoce el ARNm blanco la proteína AGO1 del complejo RISC puede introducir un corte en un enlace fosfodíester del ARNm en el sitio blanco del miARN, o inhibir la traducción del mismo (Brodersen and Voinnet, 2009; Chen, 2004).

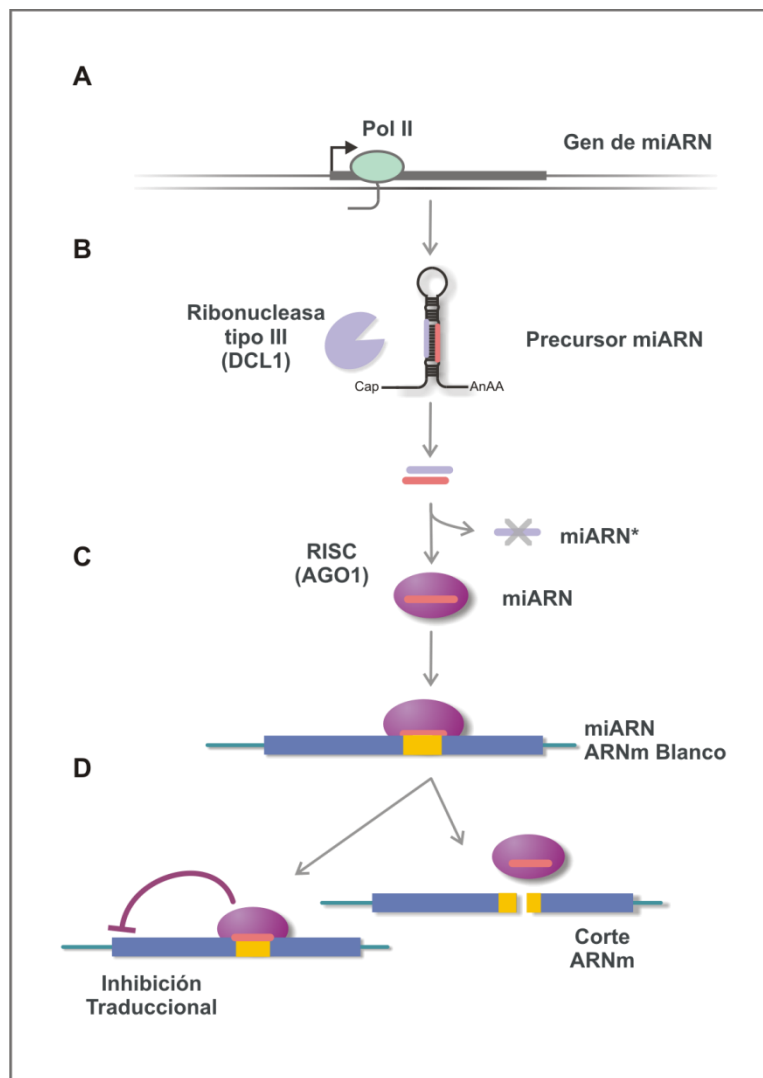


Figura I.1. Biogénesis y acción de los miARNs en plantas. Esquema que muestra el proceso de síntesis, procesamiento y mecanismo de acción de los miARNs en plantas.

A) Los pri-miARNs son transcritos por la Pol II y se pliegan sobre sí mismos en una estructura de tallo y loop imperfectos.

B) Estas estructuras son procesadas en el núcleo por DCL1 junto a otras proteínas accesorias, liberando el dúplex del miARN/miARN*.

C) El miARN es incorporado en la proteína AGO1 para formar el complejo RISC. Allí el miARN guía al ARNm blanco que es reconocido por complementariedad de bases.

D) La regulación de genes blancos ocurre principalmente por el corte endonucleolítico del ARNm blanco, aunque algunos miARNs también regulan negativamente la expresión de los ARNm al inhibir la traducción.

3.3. Genes de miARNs y conservación evolutiva

Los miARNs se registran en la base de datos de miARNs a través de números correlativos (miRBASE <http://www.mirbase.org/>). Actualmente existen más de ~8500 *MIARNs* del reino Plantae anotados y en

Arabidopsis se han identificado más de 300 *MIARNs* (miRBASE-Release 21: Febrero 2017)(Kozomara and Griffiths-Jones, 2014). Los genes de miARN se agrupan en familias, las cuales incluyen los diferentes loci que codifican para miARNs con secuencia madura similar o idéntica, y que se diferencian en el resto de las secuencias del precursor (Meyers et al., 2008). Muchos de los miARNs conservados sufrieron de duplicaciones a partir del locus original, por lo que las familias de miARNs conservados poseen múltiples *loci*. Incluso existen familias muy numerosas compuestas por 13 *loci* como es el caso de la familia miR169, también hay familias compuestas por un único gen.

Hay 21 familias de miARNs conservados en plantas dicotiledóneas (Figura I.2)(Cuperus et al., 2011; Chavez Montes et al., 2014). Sin embargo el número de familias puede variar dependiendo si miR156/157, miR165/166, miR170/171 o miR159/miR319 son considerados dos familias separadas o un a única súper familia (Cuperus et al., 2011; Meyers et al., 2008). El resto de las familias tienen una distribución taxonómica más restrictiva a determinados géneros. A su vez muchos pequeños ARNs han aparecido recientemente en la evolución y por lo tanto aparecen en un número pequeño de especies, estos son denominados miARN jóvenes (Axtell, 2008; Axtell and Bowman, 2008). En general los miARN jóvenes se encuentran representados por genes de copia única en el genoma y sus genes blanco incluyen una amplia gama de proteínas que abarcan múltiples aspectos de la biología de plantas.

En animales se han encontrado miARNs conservados en especies muy distantes como gusanos hasta humanos (Pasquinelli et al., 2000). Sin embargo no se han encontrado miARNs conservados entre animales y en plantas, lo que sugiere que han aparecido en forma independiente durante la evolución (Cuperus et al., 2011). A pesar de que los miARNs en plantas y en animales compartan componentes básicos, su biogénesis y su modo de acción difieren entre ellos.

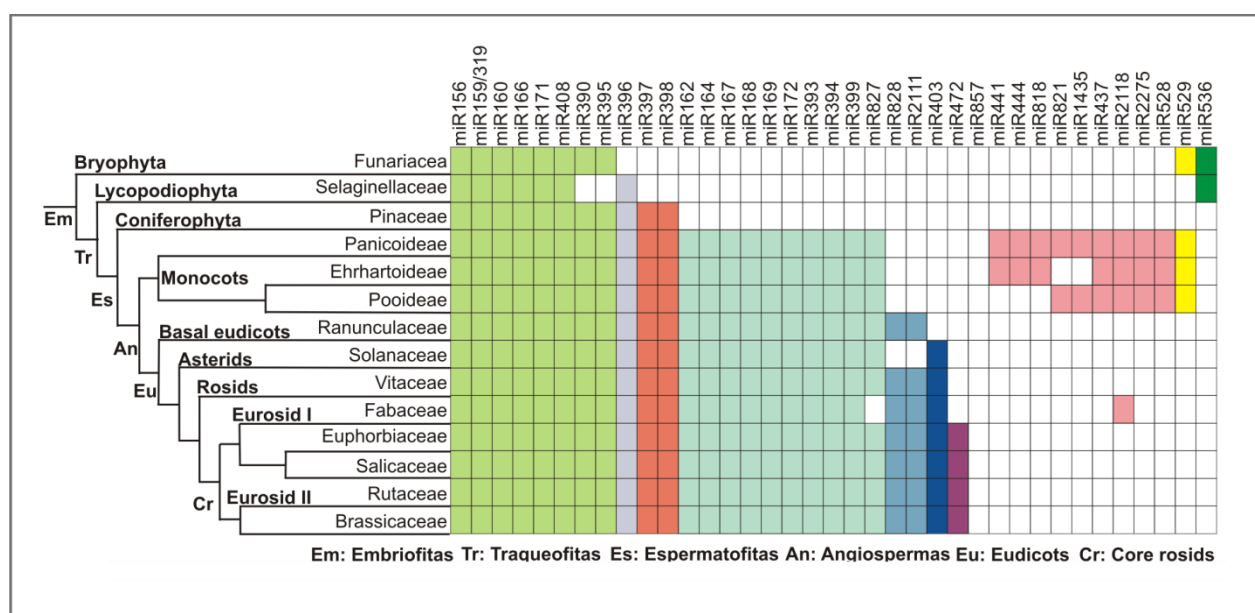


Figura 1.2. Familias de miARNs conservados en plantas. Familias de miARNs (columnas) que se encuentran conservados entre diferentes especies (filas). Las familias de miARNs específicas de una sola especie fueron omitidas. Los grupos de familias de miARNs son marcados con diferentes colores de acuerdo al rango taxonómico inferido. Tomado de (Cuperus et al., 2011).

3.4. Biogénesis de miARN

3.4.1. Procesamiento de miARNs en animales

Al igual que en plantas, los genes de miARNs de animales son transcritos por la Pol II. Los pri-miARN pueden poseer un tamaño de varias kilobases de longitud y presentar varias estructuras locales de tallo y loop que darán lugar a más de un miARN (Kim et al., 2009). Un pri-miARN típico consiste en un tallo de 33-35pb, un loop terminal pequeño y un segmento de ARNsh hacia los extremos 3' y 5'. La transición de ARN ARNdh entre el tallo y los extremos de ARNsh (ARNsh-dh), se denomina unión basal, mientras que la transición de ARNsh-dh entre el tallo y el loop terminal, se denomina unión apical (Figura 1.3A). El procesamiento de los precursores de animales, consta de dos pasos de los cuales el primero es realizado en el núcleo y el segundo en el citoplasma (Ha and Kim, 2014).

El primer corte en el núcleo es realizado por un complejo denominado microprocesador (Figura 1.3B). El mismo está constituido por una molécula de la ribonucleasa de tipo III, DROSHA y por dos moléculas de la proteína Di George síndrome critical gene 8 (DGCR8), esta proteína accesoria posee dominios de unión a ARNdh (dsRBD). El complejo microprocesador produce el primer corte sobre el pri-miARN, el cual se ubica 11pb hacia arriba de la unión basal, y 22pb hacia debajo de la unión apical (Figura 1.3B). Esto produce la liberación del pre-miARN, una estructura de tallo y loop de ~65nt que luego es exportada al citoplasma (Figura 1.3D)(Denli et al., 2004; Gregory et al., 2004; Han et al., 2004a; Kwon et al., 2016; Zeng and Cullen, 2005).

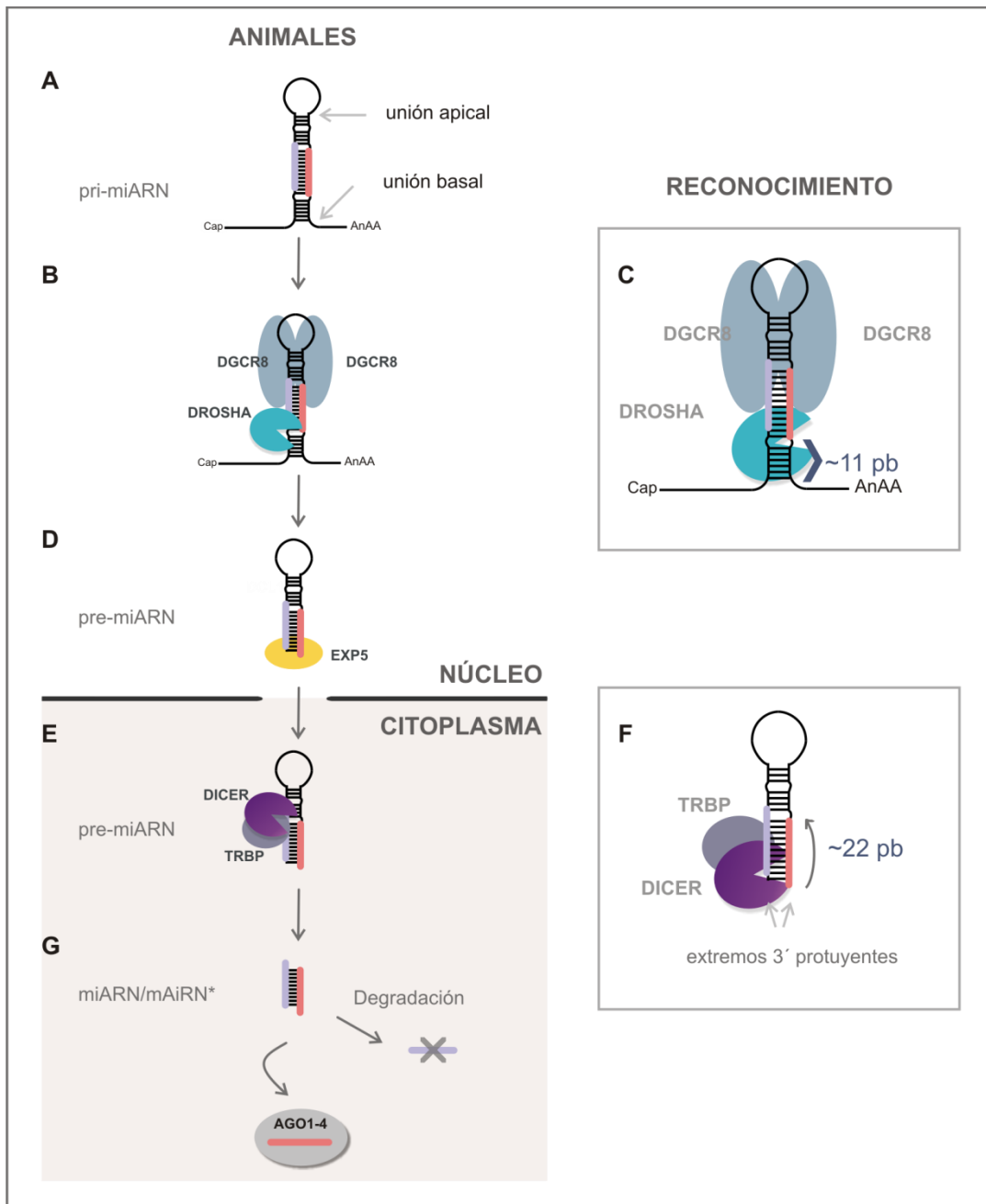
Un posible modelo para explicar el mecanismo de reconocimiento y corte del microprocesador propone que DROSHA se une a la base del pri-miARN, mientras que las dos moléculas de DGCR8 reconocen la región apical del mismo (Figura 1.3C). De esto modo las DGCR8 interaccionarían entre si frontalmente y con DROSHA hacia abajo. DROSHA tendría dos sitios de unión para DGCR8, uno en cada dominio ARNasa tipo III, mediando de este modo, el ensamblado del complejo microprocesador alrededor del pri-miARN. A continuación DROSHA reconocería la transición de ARNsh-dh presente en la unión basal del pri-miARN y realizaría un corte a ~11pb de la misma. En este sentido DROSHA no sólo aportaría la actividad catalítica para realizar el corte sino que además determinaría la posición de corte al contar 11pb de ARNdh desde la unión basal, funcionando como una regla molecular (Figura 1.3C) (Kwon et al., 2016; Nguyen et al., 2015).

La proteína de transporte nuclear EXPORTIN-5 (EXP5) reconoce el tallo inferior de ARNdh del pre-miARN con dos nucleótidos 3'protruyentes (Figura 1.3E). Luego forma un complejo con la proteína nuclear de unión a GTP, RAN-GTP. Este complejo es exportado a través de los poros nucleares al citoplasma. Una

vez allí, la hidrólisis de GTP promueve la disociación de este complejo y la liberación del pre-miARN en el citosol (Ha and Kim, 2014; Lund et al., 2004; Yi et al., 2003). Algunos pre-miARNs procesados por vías no-canónicas son transportados mediante la proteína EXPORTIN-1, sin embargo las estructuras reconocidas en el pre-miARN por la misma no están del todo claro (Ha and Kim, 2014).

En el citoplasma una segunda ribonucleasa de tipo III, DICER conjuntamente con TRPB, otra proteína de unión a ARNdh, produce un segundo corte sobre el precursor para liberar el dúplex miARN/miARN* (Figura I.3E). Este segundo corte ocurre a ~22nt de distancia del sitio donde se había producido el primer corte por DROSHA. DICER se une a los nucleótidos 3' protruyentes dejados por DROSHA y desde allí realiza un corte a una distancia fija de ~22nt (Figura I.3F). Cabe destacar que la distancia fija de corte depende de la especie y el tipo de DICER, en el caso de la DICER de *Giardia intestinalis* el corte ocurre a ~25nt (Ha and Kim, 2014; Kwon et al., 2016). Finalmente, el miARN maduro es incorporado al complejo RISC (Figura I.3G)(Ha and Kim, 2014).

Además de esta ruta canónica de procesamiento, se han descrito vías alternativas. Para algunos precursores, el requerimiento de DROSHA puede ser suplantado por la maquinaria de corte y empalme (splicing)(Okamura et al., 2007; Ruby et al., 2007). En el caso de otros, es necesaria la interacción de proteínas específicas con el loop terminal (Guil and Caceres, 2007; Trabucchi et al., 2009). Por último, existen casos en donde, el procesamiento de ciertos precursores puede estar específicamente inhibido en ciertos tejidos o determinadas condiciones (Heo et al., 2008; Obernosterer et al., 2006; Viswanathan et al., 2008).



1.3. Biogénesis de miARN en animales.

A) Dentro del pri-miARN existe una transición de ARNsh-dh en la base llamada unión basal y una transición de ARNsh-dh en el loop, llamada unión apical.

B) El complejo microprocesador conformado por DROSHA y DGCR8 realiza el primer corte en el citoplasma.

C) DROSHA reconoce la unión basal y realiza el corte a ~11pb de la mismo.

D) El pre-miARN es exportado al citoplasma.

E) El segundo corte es mediado por un complejo formado por DICER y TRBP.

F) DICER reconoce los extremos 3' protuyentes dejados por DROSHA en el pre-miARN y corta a ~22pb, liberando el dúplex miARN/miARN*

G) El miARN es incorporado en el RISC mientras que el miARN* es degradado.

3.4.2. Procesamiento de miARNs en plantas

Los miARNs se diferencian de otros ARNs pequeños por su biogénesis particular que implica su escisión de un precursor con extensa estructura secundaria en forma de tallo y loop (Jones-Rhoades et al., 2006). El

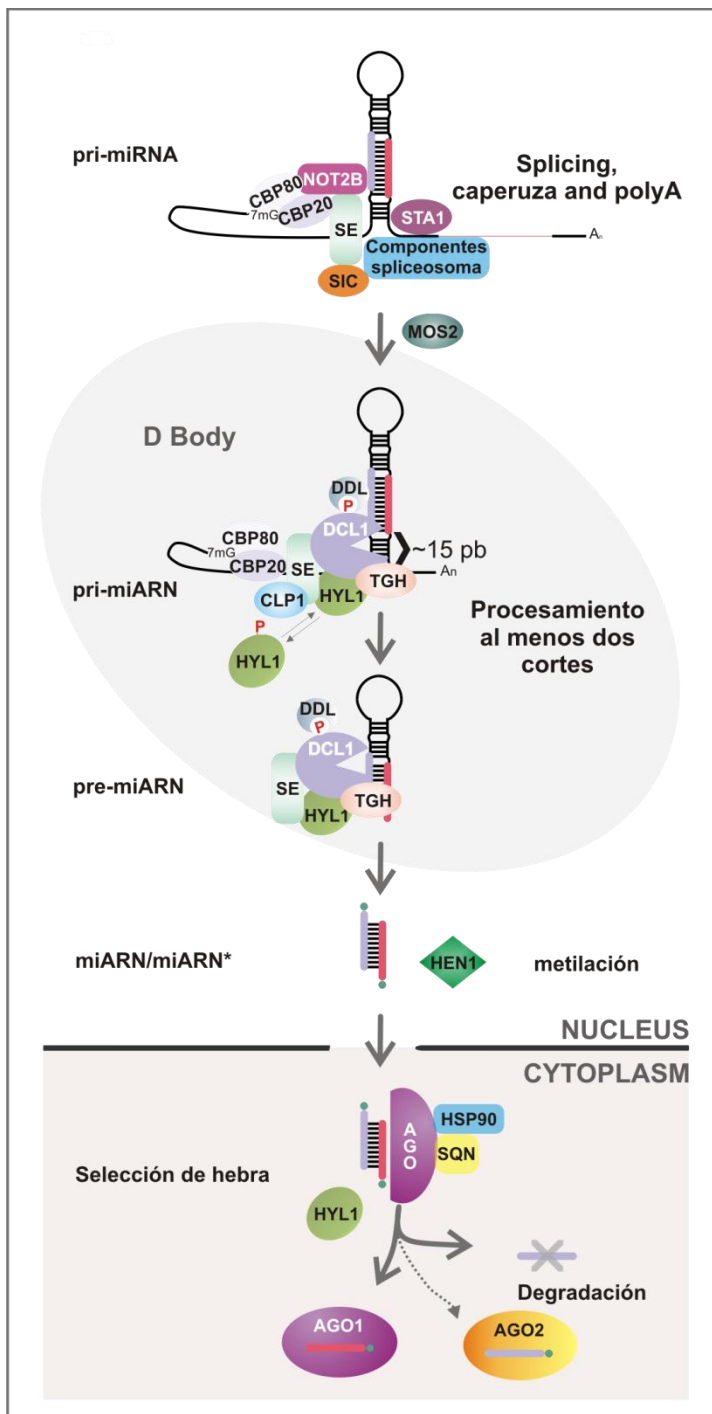
miARN maduro se ubica en uno de los brazos del precursor, apareado con una secuencia parcialmente complementaria denominada miARN* (Figura I.4). A diferencia de lo que ocurre en animales, los precursores de miARNs de plantas son procesados completamente en el núcleo mediante DCL1 (Park et al., 2002; Reinhart et al., 2002).

DCL1 es asistida por la proteína de unión a ARNdh HYL1 y SE (Figura I.4) (Laubinger et al., 2008; Vazquez et al., 2004a). Estas tres proteínas interactúan físicamente y co-localizan con el pri-miARN en los cuerpos D en núcleo (del inglés *Dicing-Bodies*) (Fang and Spector, 2007; Fujioka et al., 2007; Song et al., 2007). HYL1 y SE mejoran la eficiencia y precisión del corte de DCL1, en concordancia con esto alelos fuertes de *se* son letales mientras que mutantes en *hyl1* causan severos defectos en el desarrollo (Dunoyer et al., 2007; Han et al., 2004b; Lobbes et al., 2006; Yang et al., 2006). La inactivación de las proteínas de unión a ARN TOUGH (TGH) (Ren et al., 2012b) y MODIFIER OF SNC1 2 (MOS2) (Wu et al., 2013) aumentan los niveles de los pri-miARNs lo que sugiere que ambas proteínas participan del procesamiento de miARNs. A su vez se ha propuesto que la proteína de unión a ARN DAWDLE (DDL) reconoce el pri-miARN y mediante interacciones proteína-proteína facilita la interacción del mismo con DCL1 (Achkar et al., 2016; Bologna and Voinnet, 2014; Yu et al., 2008). La fosfatasa CPL1 participa en esta vía de una manera indirecta activando a HYL1 por desfosforilación de residuos específicos (Manavella et al., 2012a).

El dúplex miARN/miARN* liberado contiene dos nucleótidos simple hebra protruyentes en el extremo 3' los cuales son metilados por HEN1, la cual a su vez metila los extremos 3' de todos los ARNs pequeños (Figura I.4) (Boutet et al., 2003). El grupo metilo previene la uridilación y posterior degradación del dúplex del miARN (Li et al., 2005; Yu et al., 2005). El paso final de la biogénesis es la incorporación asimétrica del miARN, a partir del dúplex miARN-miARN* al complejo RISC y el miARN* es degradado. HYL1 y CPL1 facilitan este proceso, lo que sugiere que la carga en AGO1 podría estar acoplada a la biogénesis del miARN, como se ha visto en Metazoos (Eamens et al., 2009; Manavella et al., 2012a). La remoción del miARN* de AGO1 no depende de la actividad de corte de la misma, si no que requiere de su asociación con las proteínas HSP90 y SQN (Iki et al., 2012; Iki et al., 2010) las cuales probablemente causen un cambio conformacional en AGO1, como se ha visto en Metazoos (Gu et al., 2012; Kwak and Tomari, 2012). En algunos casos el miARN* puede estabilizarse y ser funcional en condiciones fisiológicas o bajo condiciones de stress como sucede con miR393* frente a la infección por *Pseudomonas* (Devers et al., 2011; Zhang et al., 2011).

El componente central de todos los complejos RISC es un miembro de la familia de proteínas AGO. En *Arabidopsis* existen distintas proteínas AGO que participan en diferentes procesos biológicos (Sección 3.6.7). La incorporación de los ARN pequeños en los diferentes complejos depende de la identidad del nucleótido del extremo 5' y de la vía de biogénesis (Mi et al., 2008; Montgomery et al., 2008a). En la mayoría de los miARNs presentan una U en el extremo 5', lo que ayudaría a la selección del ARN pequeño

por AGO1, la principal efectora de la actividad de los miARNs (Mi et al., 2008; Vaucheret et al., 2004; Voinnet, 2009).



I.4 Biogénesis de miARNs en plantas. Tomado de (Bologna and Voinnet, 2014).

3.5. Estructuras secundarias y mecanismos de procesamiento en animales y plantas.

La biogénesis de miARNs involucra el adecuado procesamiento del precursor para liberar el miARN maduro. Un error durante la biogénesis del miARN cambiaría la secuencia del miARN maduro y por lo tanto los

genes blanco regulados por él. Además de las diferencias que existen durante la biogénesis y el mecanismo de acción entre los miARNs en plantas y en animales existen diferencias en las estructuras secundarias de los precursores de ambos reinos.

En animales, existe una marcada homogeneidad en el tamaño de los precursores de miARNs, en cambio los precursores de miARNs de plantas son muy variables en tamaño y forma. Trabajos previos mostraron que un pri-miARN de animales consiste en un tallo imperfecto posee de ~65nt (Han et al., 2006) (Figura 1.5A). El mismo puede ser dividido en: un loop terminal y tallo superior pequeños, un dúplex miARN/miARN*, un tallo inferior y las secuencias flanqueantes simple hebra. El dúplex miARN/miARN* posee una longitud de ~22pb y el tallo inferior ~11pb (Han et al., 2006). Luego del primer corte llevado a cabo por DROSHA, el pre-miARN liberado está conformado sólo por el dúplex miARN/miARN* y el loop terminal.

En cambio, los precursores de miARNs de plantas presentan tamaños que varían desde los 50 a 900nt, lo que se traduce en un amplio abanico de estructuras diferentes donde el único tamaño conservado entre los precursores, es el dúplex miARN/miARN* de ~21pb (Bologna et al., 2009) (Figura 1.5B). Al parecer es la estructura secundaria por sobre la secuencia primaria del precursor la determinante del correcto procesamiento del mismo (Bologna et al., 2013a). En plantas existen cuatro mecanismos descritos por los cuales la maquinaria de procesamiento puede reconocer a precursor de miARN y cortarlo, lo cual podría explicar la heterogeneidad de estructuras de los precursores (Bologna et al., 2013b).

Un grupo de precursores de miARN presenta un tallo inferior de ARNd_h ~15/17pb seguido de un loop interno por debajo del dúplex del miARN/miARN*, el cual es reconocido por el complejo de DCL1 y especifica la posición del primer corte (Figura 1.5C) (Bologna et al., 2013b; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Zhu et al., 2013). Luego DCL1 realiza un segundo corte ~21nt del primero, y libera el dúplex del miARN/miARN*. Este mecanismo es denominado base a loop y tiene semejanza con el mecanismo de procesamiento descrito para animales (Sección 3.4.1) Este determinante está presente también en un segundo grupo de precursores los cuales presentan un tallo inferior más largo (Figura 1.5D). Estos precursores son procesados desde la base al loop, pero requieren de tres cortes consecutivos de DCL1 para liberar el miARN maduro, esto se ha denominado procesamiento secuencial (Bologna et al., 2013b; Kurihara and Watanabe, 2004).

Sin embargo otro grupo de precursores de plantas es procesado con un primer corte por debajo del loop terminal y luego DCL1 continúa en dirección a la base del precursor (Addo-Quaye et al., 2009; Bologna et al., 2009; Bologna et al., 2013b). Estos precursores procesados desde el loop hacia la base presentan una región estructurada de ARNd_h por encima del dúplex de miARN, se ha propuesto que este tallo superior es reconocido por la maquinaria de procesamiento. Como en el caso anterior, la biogénesis del ARN pequeño puede depender de dos o más cortes de DCL1, es decir que el procesamiento secuencial de un precursor

puede comenzar desde la base o loop terminal (Figura I.5E)(Addo-Quaye et al., 2009; Bologna et al., 2009; Bologna et al., 2013b).

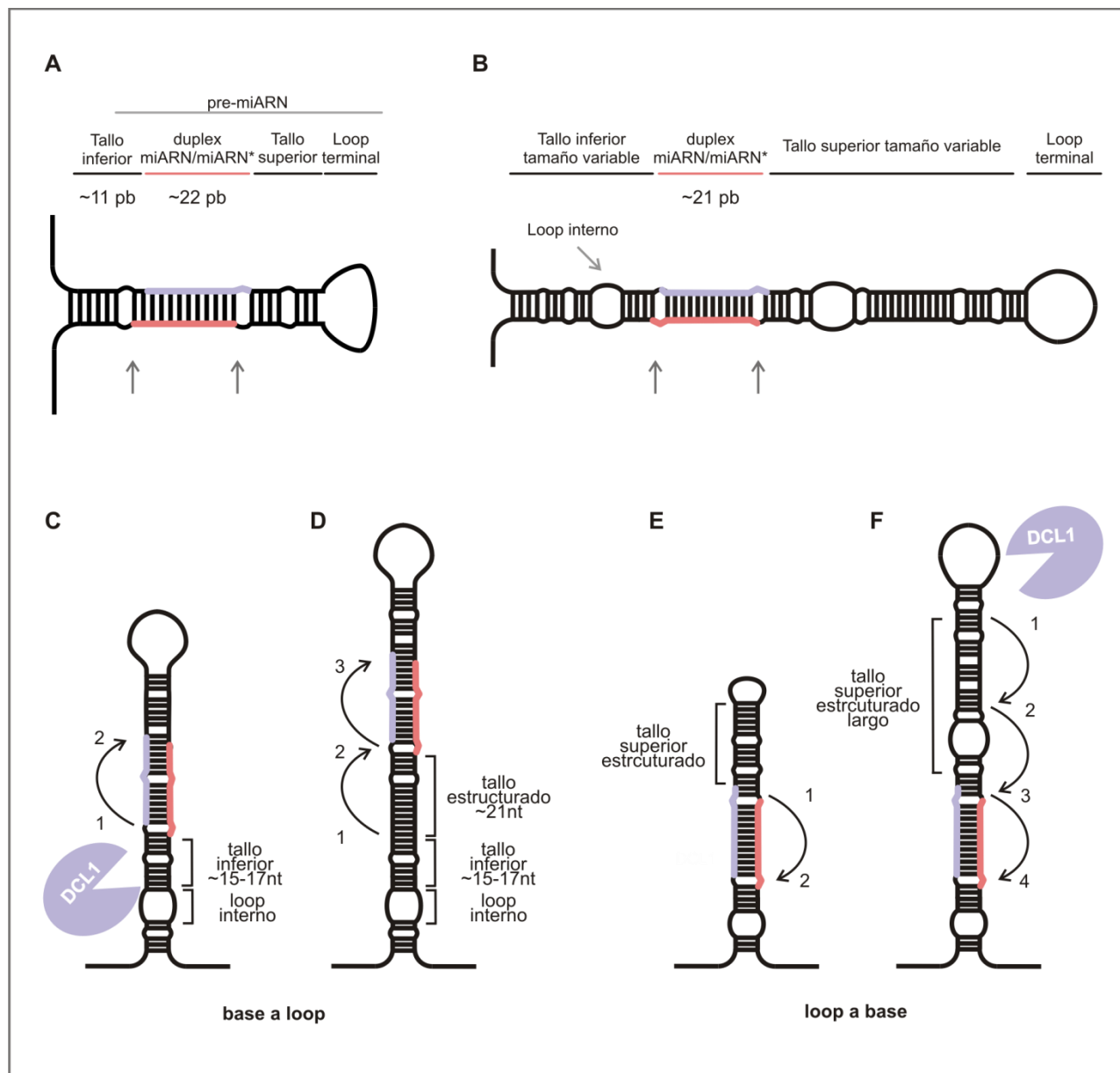


Figura I.5. Estructuras secundarias y mecanismos de procesamiento de precursores de miARN en plantas.

A) Esquema de un pri-miARN de animales. El mismo consiste en un tallo imperfecto posee de ~65nt, donde el dúplex miARN/miARN* posee una longitud de ~22pb y el tallo inferior ~11pb. El tallo superior y el loop terminal son pequeños.

B) Esquema de un pri-miARN de plantas. El mismo consiste en tallo inferior de longitud variable que puede contener loops internos, un dúplex de miARN/miARN* de ~21pb, y un tallo superior y loop terminal de longitud variables.

C) Esquema de un precursor procesado desde la base al loop con dos cortes de DCL1.

D) Esquema de un precursor procesado desde la base al loop con tres cortes de DCL1.

E) Esquema de un precursor procesado desde el loop hacia la base con dos cortes de DCL1.

F) Esquema de un precursor procesado desde el loop hacia la base con cuatro cortes de DCL1.

En todos los casos las flechas muestran los cortes sobre el precursor. Y el miARN se muestra en rojo. Adaptado de (Bologna and Voinnet, 2014).

3.6. Proteínas involucradas en la biogénesis de ARNs pequeños.

Las proteínas que participan de biogénesis de miARNs son proteínas modulares que contienen diferentes dominios con funciones específicas, y recientemente se han descrito nuevos componentes en la biogénesis de miARNs en plantas (Figure I.4). A continuación, se describen en más detalle las características de algunas de ellas que tienen un rol descrito central en la biogénesis y acción de los miARNs de plantas, incluyendo a DICER, HYL1, SE, HEN1, y AGO.

3.6.1. ARNasas tipo III.

Las enzimas ARNasa III son específicas de ARNdH. Los productos de la digestión por ARNasa III son fragmentos de ARNdH con un grupo fosfato 5' y un hidroxilo terminal 3' en cada hebra. El mecanismo propuesto de acción de estas enzimas consiste en la dimerización inter o intramolecular de los dominios ARNasa III para formar un único centro catalítico (Zhang et al., 2004). En este centro catalítico único cada dominios ARNasa corta una de las hebras del ARNdH, y la distancia entre dos dominios involucrados en el sitio catalítico de los es la que genera extremos 3' con dos nucleótidos protruyentes.

Las distintas enzimas ARNasa III presentan dominios adicionales que funcionan concertadamente con los dominios ARNasa III (Nicholson, 2014). Según el arreglo de dominios, este tipo de enzimas puede dividirse en tres clases (MacRae and Doudna, 2007). Las de clase I se encuentran en bacterias, bacteriófagos y hongos, y contienen un solo dominio tipo ARNasa III unido a un dominio dsRBD, estas enzimas no se desarrollaran en la introducción. Mientras que las clases II y III son representadas por DROSHA, DICER y las DCLs, respectivamente.

3.6.1.1. Clase II: DROSHA.

DROSHA es la ARNasa III a cargo de realizar el primer paso en el procesamiento de precursores de miARNs en animales. Se ha demostrado que la región N-terminal de la proteína no es necesaria para el procesamiento de pri-miARNs *in vitro*, sin embargo, la misma es requerida para localización nuclear (SLN) de la misma *in vivo*. Mientras que en el extremo C-terminal presenta dos dominios ARNasa III en tándem y un dsRBD (Figura I6). Los dos dominios ARNasa III dimerizan intramolecularmente para formar un único centro catalítico. El primer dominio ARNasa III (ARNasa IIIa) corta la hebra 3' mientras que el segundo dominio ARNasa III (ARNasa IIIb) corta la hebra 5' del pri-miARN (Ha and Kim, 2014; Han et al., 2004a; Han et al., 2006). El dominio dsRBD de DROSHA es necesario, pero no suficiente para asegurar la interacción con el sustrato. La actividad adicional de unión a ARN es proporcionada por DGCR8, la cual es reclutada a través de la región media de DROSHA (Ha and Kim, 2014).

3.6.1.2. Clase III: DICER.

La ARNasa de clase III son estructuralmente más complejas que las ARNasas III de tipo II. En general poseen un dominio o helicasa (DEAD/Helicasa), un pequeño dominio de interacción proteína-proteína (DUF283),

un dominio PAZ (del inglés, “Piwi Argonaute Zwiller”), dos dominios ARNasa III en tándem, y al menos un dsRBD. La clase III de las ARNasa III agrupa a las enzimas DICER. Estas enzimas se encuentran presentes en la mayoría de los eucariotas, con un número de copias y estructura de dominios que varía entre los distintos organismos. Existen dos en *Drosophila melanogaster* y una en mamíferos, *Caenorhabditis elegans*, y *Schizosaccharomyces pombe* (Park et al., 2005).

DICER presenta en el C-terminal los dominios ARNasa III en tándem los cuales forman un dímero intramolecular para crear un centro catalítico, de manera similar a lo que sucede con DROSHA (Figura 1.6). La helicasa presente en el dominio N-terminal DICER facilita el reconocimiento del pre-miARN por interacción con el loop terminal del mismo. A su vez, experimentos realizados *in vitro* con la proteína DICER de humanos indicarían que el dominio helicasa tendría funciones de interacción proteína-proteína y funcionaría como un interruptor intramolecular para mantener a DICER en un estado de baja actividad hasta que la proteína se ensamble con sus proteínas accesorias (Ma et al., 2012).

Se ha propuesto que DICER ancla el extremo 3' protruyente del pre-miARN mediante el dominio PAZ. Este dominio está conectado al dominio catalítico de tal forma que el sitio activo se queda localizado a una distancia fija, actuando como una regla molecular y determinando la longitud del producto de la reacción de corte (Ha and Kim, 2014; Lau et al., 2012; MacRae et al., 2006; Tian et al., 2014). Esto se conoce como la regla del corte 3' (MacRae and Doudna, 2007; Vermeulen et al., 2005; Zhang et al., 2002; Zhang et al., 2004). En mamíferos y moscas existe un mecanismo adicional para determinar el corte en el pre-miARN: DICER se une al extremo 5' fosfato del pre-miARN y corta a 22nt de dicho extremo, lo cual se conoce como la regla del 5' (Tian et al., 2014). La unión del extremo 5' ocurre cuando el mismo es termodinámicamente inestable, y no sucede cuando el mismo presenta un apareamiento fuerte, como un G:C. La estructura del dominio PAZ de la DICER humana tiene dos bolsillos con características básicas para unir los extremos 5' y 3' del pre-miARN (Park et al., 2011; Tian et al., 2014). Los bolsillos están suficientemente separados para que ambos puedan ocuparse simultáneamente con los extremos 5' y 3' del pre-miARN, cuando el mismo presenta dos nucleótidos 3' protruyentes. Esta disposición espacial podría explicar en parte la preferencia de la DICER de humanos de sustratos con extremos 3' protruyentes (Ha and Kim, 2014).

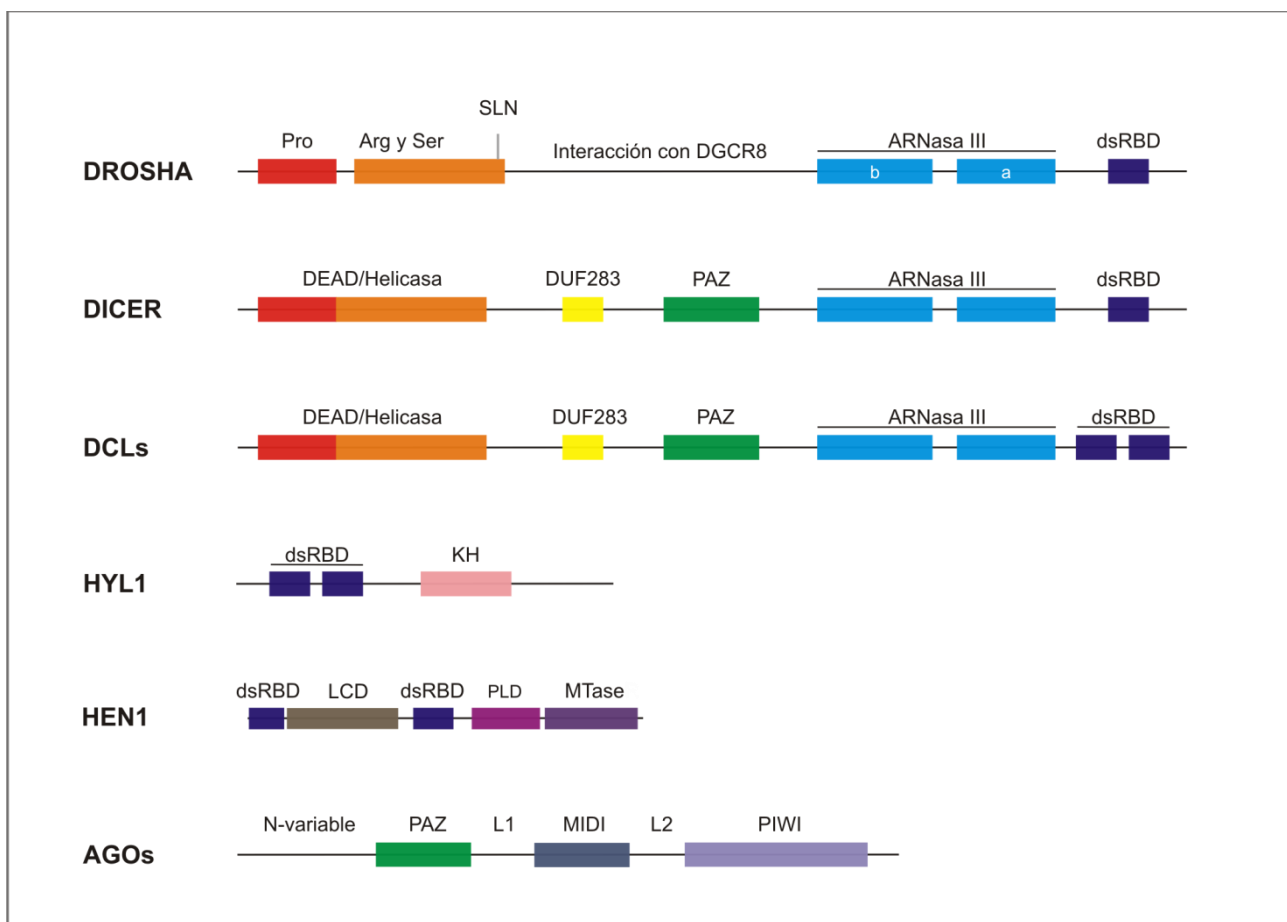
3.6.1.3. Clase III: ENZIMAS DICER-LIKE (DCLs).

En *Arabidopsis*, DCL1-4 cuentan con un dominio DEAD/helicasa, DUF283, PAZ, dos dominios ARNasas III y al menos un dsRBD (Figura 1.6). Por tratarse de proteínas de gran tamaño, la información estructural disponible sobre estas enzimas es limitada. Las funciones de los distintos dominios han sido determinadas combinando evidencias de actividad *in vivo* y ensayos bioquímicos.

Mutaciones en el dominio helicasa, ARNasa III o en los dsRBDs de DCL1 reducen la acumulación de miARN maduros *in vivo* (Kurihara and Watanabe, 2004; Willmann and Poethig, 2011). El dominio helicasa es necesario para el correcto procesamiento *in vivo* de al menos algunos pri-miARNs y la actividad de la

enzima dependiente de ATP *in vitro* (Liu et al., 2012; Zhu et al., 2013). Una mutante de DCL1 denominada *dcl1-7*, que presenta un dominio helicasa inactivo, no es capaz de acumular miARNs maduros pero presenta una alta eficiencia y precisión en la reacción del primer corte (Kurihara et al., 2006) lo que sugiere que probablemente que este dominio no esté involucrado en la realización del primer corte sobre el precursor (Bologna and Voinnet, 2014).

El dominio DUF283 presenta una estructura similar a un dsRBD, pero un estudio estructural y bioquímico sobre este dominio de DCL4 mostró que tenía un plegamiento no canónico y no se detectó afinidad por el ARNdh. En el mismo trabajo se determinó la capacidad de este dominio de interactuar con la proteína accesoria de unión a ARN DRB4, mientras que el dominio DUF283 de DCL1 interactuaría específicamente con HYL1 (Qin et al., 2010). Una particularidad que diferencia a DCL1, DCL3 y DCL4 de su contraparte en animales es la presencia de un segundo dsRBD en el extremo C-terminal. A pesar de ser un dominio de unión a ARNdh, el dsRBD C-terminal de DCL1 presenta un plegamiento no canónico y forma complejos heterogéneos con el ARNdh. Se ha propuesto que este dominio promueve la localización de nuclear de estas proteínas (Burdisso et al., 2012; Carbonell et al., 2012).



I.6. Representación esquemática de los dominios presentes en las proteínas DROSHA, DICER, DCLs, HYL1, HEN1 y AGOs. El esquema es ilustrativo y no representa los tamaños relativos de las proteínas.

DCL1 participa en la biogénesis de miARN y nat-siARN (Cuperus et al., 2011; Rogers and Chen, 2013). DCL2 en la formación nat-siARN y de siARN exógenos en respuesta a un transgén o defensa antiviral y (Jauvion et al., 2012). DCL3 está involucrada en el TGS mediante modificación de la cromatina, además de generar siARN en ciertas ocasiones (Pontes et al., 2006). DCL4 es la responsable de la biogénesis de tasiARNs y también participa en la defensa antiviral (Bologna and Voinnet, 2014; Henderson et al., 2006) Más allá de la sectorización en la funciones de las DCLs, estudios con doble, triple y cuádruples mutantes de DCLs han mostrado que las proteínas DCLs pueden actuar de manera redundante y compensar la ausencia de alguna de ellas en la producción de ciertos ARN pequeños, en especial en la respuesta a la infección por virus o actuar en paralelo sobre ciertos sustratos (Blevins et al., 2006; Gasciolli et al., 2005; Henderson et al., 2006; Moissiard and Voinnet, 2006). Un ejemplo de esto, es la generación de lmiARNs de 24nt a partir del reconocimiento y corte de un precursor de miARNs mediante DCL3. Se observó que la acumulación de especies de 24nt era mayor en inflorescencias, donde los niveles de DCL3 son mayores (Vazquez et al., 2008).

Un estudio filogenético acerca de la similitud entre las secuencias de las cuatro enzimas entre distintas especies de plantas propuso que DCL2-4 son evolutivamente más cercanas que DCL1 (Liu et al., 2009). Esto se consistente con la función especializada de DCL1 en la vía de biogénesis de los miARN. La pérdida de función de DCL1 resulta letal a nivel del embrión en Arabidopsis lo que sugiere que no existe otra DCL que sea capaz de procesar eficientemente a los precursores de miARNs con estructuras de horquilla o de localizarse en el tejido adecuado al momento de cumplir su función esencial (Nodine and Bartel, 2010; Schauer et al., 2002).

3.6.2. HYL1.

Durante la biogénesis de miARNs en animales las proteínas de unión a ARNdh DGCR8 y TRBP interaccionan con DROSHA y DICER respectivamente y son esenciales para producir el corte preciso de los precursores (Kim et al., 2009). En plantas la proteína de unión a HYL1 juega un similar asistiendo a DCL1 durante el procesamiento. En concordancia con esto, ensayos *in vitro* mostraron que HYL1 se une a las regiones de ARNdh de los pri-miARNs (Bologna and Voinnet, 2014). Esta proteína, también conocida como DRB1, fue descubierta originalmente a partir del fenotipo de plantas mutantes que tienen alterada la morfogénesis de las hojas, de ahí el nombre HYPONASTIC LEAVES y la respuesta a hormonas (Lu and Fedoroff, 2000). Posteriormente se demostró la participación de la misma en la biogénesis de miARNs (Vazquez et al., 2004a). En Arabidopsis, existen 4 proteínas homólogas a HYL1 denominados DRB2-DRB5, particularmente DRB4 interaccionaría con DCL4 (Hiraguri et al., 2005).

La proteína contiene dos dsRBD sobre el extremo N-terminal y un dominio de unión a ácidos nucleicos KH ubicado sobre el centro de la proteína (Figura I.6). El extremo C-terminal sería dispensable, mientras que la región N-terminal, conteniendo solamente los dos dsRBD es suficiente como para rescatar

el fenotipo de *hyl1* (Wu et al., 2007). Además de la regulación de HYL1 por fosforilación y desfosforilación, HYL1 también es regulada post-traduccionalmente. Durante el día la enzima E3 ubiquitin ligasa CONSTITUTIVEPHOTO-MORPHOGENIC1 (COP1) es traslocada desde el núcleo al citoplasma para suprimir a una proteasa aún no identificada y prevenir la degradación de HYL1 (Cho et al., 2014).

3.6.3. SE.

Posee una estructura de dedos de zinc del tipo C2H2 y fue descrita primeramente como una proteína involucrada en el desarrollo del tallo (Prigge and Wagner, 2001). Sin embargo, las plantas mutantes se presentan los fenotipos pleiotrópicos característicos de la deficiencia de miARNs debido al rol que juega SE en la biogénesis de los mismos (Lobbes et al., 2006; Yang et al., 2006). Mediante experimentos *in vitro* se demostró que SE reconoce las transiciones ARNsh-dh e interacciona con HYL1. Tanto la región N-terminal como el dominio de dedos de Zn, son necesarios para la interacción con DCL1 (Dong et al., 2008; Laubinger et al., 2008).

Las plantas mutantes se además de poseer problemas en la biogénesis de miARNs poseen afectado el mecanismo de corte y empalme de ARNm (Laubinger et al., 2008; Lobbes et al., 2006; Yang et al., 2006). Se postuló que posiblemente SE actúe como un mediador común entre CBC (el cual se encuentra unido al cap de los ARNm y a los pri-miARN) y los complejos de procesamiento del ARN, y a su vez con la maquinaria de corte y empalme (Laubinger et al., 2008).

3.6.4. Interacción de otras proteínas con DCL1, HYL1 y SE.

DDL contiene un dominio denominado forkhead-associated (FHA) el cual media las interacciones proteína-proteína con motivos que poseen fosfo-treoninas (Morris et al., 2006). Se ha demostrado que DDL interacciona con el dominio helicasa de DCL1, particularmente con una región del mismo que posee un potencial motivo fosfo-treonina, lo cual sugiere que DCL1 debería estar fosforilada para interactuar con DDL y procesar eficientemente a los pri-miARNs (Bologna and Voinnet, 2014; Yu et al., 2008). En ausencia de DDL la porción de pri-miARNs que no sean dirigidos a DCL1 probablemente sean degradados, lo que explicaría los niveles reducidos de pri-miARNs en las mutantes de *ddl* (Bologna and Voinnet, 2014; Yu et al., 2008).

TGH interacciona con los pri-miARN, DCL1, HYL1 y SE en los cuerpos D (Calderon-Villalobos et al., 2005; Ren et al., 2012b). Se ha propuesto que TGH sería capaz de modular tanto la interacción del pri-miARN con DCL1 como la actividad de DCL1 *per se* (Ren et al., 2012b), mientras que MOS2 parece ser un cofactor externo que facilita el reclutamiento de los pri-miARNs por el complejo mencionado anteriormente (Ren et al., 2012b; Wu et al., 2013). CPL1 es un interacciona directamente con SE. A diferencia de lo que ocurre en *hyl1*, *se*, *tgh*, y *mos2*, en *cpl1*, no se acumulan los pri-miARNs sin procesar sino que varios miARNs son procesados de manera incorrecta. Esto lo que llevó a proponer un modelo

donde SE reclutaría CLP1 al complejo con DCL1 donde podría activar a HYL1 al desfosforilarla (Manavella et al., 2012a).

3.6.5. HEN1.

HEN1 metila los OH 2' de cada hebra de todos los productos de las DCLs protegiéndolos de la degradación (Li et al., 2005; Yu et al., 2005). Se ha demostrado que HEN1 y SE interaccionan con la misma región de DCL1, lo cual sugiere que SE tendría que disociarse del complejo para que HEN1 pudiera unirse, lo que aseguraría que la metilación ocurra luego del procesamiento (Baranauskė et al., 2015). HEN1 de *Arabidopsis* consiste en cinco dominios, cuatro de los cuales interaccionan directamente con los ARNs pequeños: un dominio metiltransferasa (Mtase) altamente conservado pero no relacionado a ninguna otra ARN 2'-O-metiltransferasa, dos dsRBDs (dsRBD1 y dsRBD2) y un dominio que contiene un La-motive (LCD) el cual reconoce específicamente el grupo 3'OH. El dominio restante es un PPlase-like-domain (PLD) que no interacciona con los sustratos de ARN pequeño (Huang et al., 2009) (Figura I.6).

Ensayos de metilación *in vitro* y análisis estructurales de HEN1 de *Arabidopsis* en complejo con el ARNdh mostraron que los dos extremos 3' del dúplex son metilados en eventos separados (Huang et al., 2009; Li et al., 2005; Yu et al., 2005). Es probable que un monómero de HEN1 se una al dúplex de ARN pequeño a partir del extremo N-terminal de dsRBD1, luego LCD y dsRBD2 se unirían sinérgicamente del extremo del ARN pequeño por, lo que ayudaría a posicionar los extremos 3' protruyentes cerca al dominio metiltransferasa. Este podría ser uno de los motivos que explique porque HEN1 cataliza metilación sólo a partir de los dúplex de ARNdh. (Huang et al., 2009).

Los miARNs de plantas requieren de la metilación para ser estables, por lo que sólo alelos hipomórficos de *hen1* son viables en *Arabidopsis*. Ensayos de búsquedas de supresores que rescataran el fenotipo *hen1*, revelaron la existencia de una preteína denominada HEN1 SUPPRESSOR 1 (HESO1). La misma es una nucleotidil transferasa la cual puede agregar U sin templado a los extremos 3' de dúplex de ARN no metilados marcándolos para la degradación. El otro proceso responsable de la degradación de ARNs pequeños sin metilar se ha denominado truncado y requiere de la actividad de una exonucleasa que los degrada en dirección 3'-5'. Ambos procesos son impedidos por la metilación en dichos extremos por HEN1 (Ren et al., 2012a; Yu et al., 2005; Zhao et al., 2012).

3.6.7. ARGO.

Las proteínas AGO se encuentran profundamente conservadas, y se expresan en un amplio rango de tejidos en la mayoría de los organismos. Sin embargo el número de AGO codificadas en cada uno de los distintos organismos varía notablemente (Vaucheret, 2008). Por ejemplo, mientras que en animales, *C.elegans* presenta 27 AGO diferentes, en *Drosophila melanogaster* tiene dos y en humanos solo están presentes cuatro AGOs (Huntzinger and Izaurralde, 2011; Ipsaro and Joshua-Tor, 2015). Mientras que en arroz hay 18

genes *AGO* y en *Arabidopsis* diez (Mallory and Vaucheret, 2010). Estas diez AGO se encuentran clasificadas en tres grupos principales (i) AGO1, AGO5 y AGO10, (ii) AGO2, AGO3 y AGO7, y (iii) AGO4, AGO6, AGO8 y AGO 9. Sin embargo esta clasificación está basada exclusivamente en la similitud de secuencia y no necesariamente en su función (Mallory and Vaucheret, 2010; Vaucheret, 2008).

Las proteínas AGOs están constituidas por un dominio N-terminal variable y por tres dominios conservados PAZ, MID y PIWI. Juntos estos tres dominios posicionan correctamente el ARN pequeños en relación a la secuencia blanco. Los dominios PAZ, MID y PIWI están conectados por las regiones L1 y L2. (Figura I.6). El dominio MID reconoce el extremo 5' del ARN pequeño, mientras que el dominio PAZ se ancla al extremo 3' (Frank et al., 2012). La actividad catalítica para cortar al ARNm reside en el dominio PIWI, el cual tiene un plegamiento similar a las ribonucleasas tipo H (Mallory et al., 2009; Wei et al., 2012a). Sin embargo, no todas las proteínas AGO tienen los residuos catalíticos necesarios, y por ende solo algunas de ellas pueden cortar el ARN blanco (Mallory and Vaucheret, 2010). El corte del ARNm blanco ha sido demostrado experimentalmente para AGO1, AGO2, AGO7 y AGO10 las cuales participan del PTGS, y AGO4 la cual realiza TGS (Carbonell et al., 2012; Manavella et al., 2012b; Zhu et al., 2011). Finalmente el dominio N-terminal está compuesto por una región variable, la cual facilita la separación del dúplex ARN pequeño/ARN blanco luego que se produce el corte en el ARN blanco.

Estudios realizados a partir de inmuno-precipitaciones de AGO mostraron que las mismas tienen preferencia tanto por el tamaño como por identidad del nucleótido presente en el extremo 5' del ARN pequeño (Mi et al., 2008; Montgomery et al., 2008a; Zhu et al., 2011). AGO4, -6 y -9 se asocian principalmente con siARNs de 24nt, mientras que AGO1,-2,-5,-7 y 10 unen moléculas de 21-22nt. Por otro lado AGO7 y AGO10 unen casi exclusivamente a miR390 y miR165/166 respectivamente. En lo que respecta a identidad del nucleótido 5' AGO1, -2 y -5 se unen preferentemente a ARNs pequeños que presenten un 5'U, A o C respectivamente (Mi et al., 2008). Por otro lado AGO4,- 6 y -9 se asocian principalmente con ARNs pequeños que comiencen con 5'A. Existen requerimientos especiales que pueden actuar independientemente de la identidad del nucleótido 5' para determinar la carga en una AGO particular. Estos incluyen mismatches en posiciones claves o interacciones con proteínas, de esto modo los miembros de la familia *MIR165/166* se asocian con AGO10 a pesar de presentar un extremo 5'U (Zhu et al., 2011). Del mismo modo miR390 se asocia específicamente con AGO7 en lugar de AGO2, como se esperaría teniendo en cuenta que comienza con 5'A (Montgomery et al., 2008a).

El complejo RISC con el ARN pequeño incorporado puede producir una regulación a nivel transcripcional mediante la metilación de citosinas del ADN o de proteínas histonas o una regulación a nivel post-transcripcional produciendo el corte endonucleolítico sobre el ARN mensajero blanco o la represión de su traducción (Sección 3.7). Siendo la identidad de la AGO que forma parte del complejo RISC quién fijará la función final del mismo (Bologna and Voinnet, 2014).

3.6.7.1. El clado AGO1, -5, y -10.

AGO1 es un componente central de las vías de silenciamiento por miARN y tasiARN, entre otros. Se ha validado la participación de AGO1 en el corte de ARNm como también en la represión traduccional. Una fracción de las proteínas AGO1 se asocia a endomembranas, y la presencia de AGO1 en el retículo endoplasmático rugoso ha llevado a proponer que este es el sitio donde AGO1 mediaría la inhibición de la traducción (Bologna and Voinnet, 2014; Li et al., 2013). A su vez, se ha propuesto que regulación por inhibición de la traducción probablemente ocurra cuando AGO1 se encuentra asociada a proteínas que inhiben su actividad de corte (Brodersen et al., 2008; Mallory et al., 2009).

AGO5 se expresaría principalmente en las megaesporas durante la transición hacia la megagametogénesis. En concordancia con esto plantas que presentan el alelo semidominante *ago5-4* muestran defectos en la megagametogénesis. En el gametofito masculino, AGO5 se localiza preferentemente en citoplasma de las células que darán lugar al polen maduro. Es probable que allí dirija el silenciamiento mediado por miARNs o siARNs de genes necesarios para el desarrollo del gametofito masculino (Tucker et al., 2012).

AGO10 es el homólogo más cercano de AGO1 en Arabidopsis. La misma regula el desarrollo del meristema apical al unirse específicamente de los miembros de la familia *MIR165/166* (Zhu et al., 2011). Estos miARN también se asocian con AGO1 para reprimir a los factores de transcripción class-III homeodomain-leucinezipper (HD-ZIP-III), los cuales son necesarios para el establecimiento del meristema apical (Mallory et al., 2009). A diferencia de AGO1, AGO10 regula positivamente los HD-ZIP III, dispensado de la actividad de corte de dicha proteína. Mientras que AGO1 se expresa de manera ubicua AGO10 se expresa modernamente en el meristema apical del tallo, en la porción adaxial de los primordios foliares y en los tejidos que darán lugar a la vasculatura. Finalmente AGO10 participa también de la regulación traduccional de genes blanco de miARNs, los cuales incluyen a AGO1 (Mallory et al., 2009) (Zhang and Zhang, 2012)(Desarrollado en la Sección 3.8).

3.6.7.2. El clado AGO2, -3, y -7.

A pesar de pertenecer a otro clado AGO2 cumple funciones solapadas con AGO1. Por ejemplo, AGO1 y AGO2 regulan redundantemente el ARNm de la plastocianina al cargar a miR408. A su vez ambas proteínas son requeridas el silenciamiento de regiones intergénicas, pseudogenes y transposones jóvenes que se transcriben en Arabidopsis. (Maunoury and Vaucheret, 2011). Trabajos recientes mostraron que el stress genotóxico inducido por rupturas al ADNdh, da lugar a diARNs pequeños los cuales son cargados en AGO2, permitiéndole llevar adelante una función en la reparación del ADN (Wei et al., 2012b). Los niveles de AGO2 son regulados por el miR403, en un proceso que depende de AGO1 (Allen et al., 2005). A pesar de su cercanía con AGO2, no se han encontrado roles biológicos asociados a AGO3 (Bologna and Voinnet, 2014).

Como se mencionó anteriormente, AGO7 se asocia casi exclusivamente con miR390 en un mecanismo que depende del extremo 5' y la región central del dúplex miR390/miR390*. Incluso el corte de la hebra de miR390* es un paso necesario para la maduración de AGO7-miR390 en el RISC (Montgomery et al., 2008a). MiR390 cargado en AGO7 desencadena la producción de tasiARNs a partir de *TAS3*, los cuales en última instancia regulan a AUXIN-RESPONSE FACTOR 3 (ARF3) y ARF4, para asegurar la transición de fase de juvenil a adulto y el correcto establecimiento del patrón adaxial-abaxial (Axtell et al., 2006; Montgomery et al., 2008a).

3.6.7.3. El clado AGO4, -6, -8, y -9.

Los complejos de TGS involucran AGO4, AGO6 o AGO9 y están asociados a siARNs de 24nt que comienzan con 5' (Havecker et al., 2010; Mi et al., 2008; Zilberman et al., 2003). AGO4 es el principal efector de RdDM y el TGS de transposones y repeticiones. La actividad de corte AGO4 sería necesaria para la metilación de algunos loci y dispensable para otros. A su vez AGO4 carga los lmiARNs de 24nt provenientes del procesamiento por DCL3 y en última instancia causa la metilación del ADN en el loci del miARN de origen (Qi et al., 2006). Por otra parte AGO6 y AGO9 parecen dirigir el RdDM en loci y tejidos específicos (Havecker et al., 2010). La expresión de AGO6 parece estar confinada al tallo y raíz, mientras que AGO9 se expresa en los óvulos, anteras y cubierta de las semillas, participando del proceso de formación de las gametas femeninas. Se ha demostrado que AGO9 participa en el silenciamiento de los transposones en las gametas femeninas y células accesorias (Olmedo-Monfil et al., 2010). Por otro lado AGO8 tiene muy baja expresión en *Arabidopsis* y es considerada un pseudo gen (Bologna and Voinnet, 2014).

3.7. Regulación de la expresión génica por miARNs.

En humanos se considera que más del 60% de los genes están regulados por miARNs (Ha and Kim, 2014). Incluso un ARNm puede estar regulado por varios miARNs, en cambio los miARNs de plantas regulan un número limitado de genes blanco (Voinnet, 2009). A su vez, existen otras diferencias entre estos reinos, mientras que los miARN de plantas inducen principalmente el corte endonucleolítico en sus genes blanco, los miARNs de animales raramente cortan el ARNm (Carrington and Ambros, 2003; Llave et al., 2002), si no que en general reprimen la traducción o desestabilizan el ARNm (Ameres and Zamore, 2013; Bartel).

3.7.1. Regulación por corte del ARN blanco.

Esta parece ser la vía predominante de regulación en plantas. El ARN pequeño guía al complejo RISC hacia una molécula de ARNm el cual suele presentar una elevada complementariedad con el miARN. Luego del reconocimiento de ARNm blanco, la proteína AGO1 introduce un corte en un enlace fosfodiéster del ARNm. Este corte ocurre entre las posiciones 10 y 11 desde el extremo 5' del miARN, independientemente de la

longitud del miARN (Floyd and Bowman, 2004; Llave et al., 2002; Mallory et al., 2004; Palatnik et al., 2003; Xie et al., 2003).

Luego del corte mediado por el miARN, los fragmentos 3' son degradados mediante la actividad de la enzima citoplasmática 5'-3' EXORIBONUCLEASA 4 (XRN4), mientras que los fragmentos 5' son degradados por el complejo denominado Exosoma (Chekanova et al., 2007; Souret et al., 2004). A su vez la degradación de los fragmentos 5' puede ser acelerado por la uridilación en los extremos 3' mediante HESO1 (Ren et al., 2014).

3.7.2. Regulación de la traducción por miARNs.

Los miARNs en animales reconocen a sus genes blanco mediante complementariedad de bases parcial, la región semilla que incluye los nucleótidos 2-7/8 del miARN es particularmente importante para el reconocimiento (Ameres and Zamore, 2013). Incluso un mismo miARN puede unirse a uno o más sitios presentes en la región 3' no traducida de uno o varios ARNm blancos (Carrington and Ambros, 2003). Los miARNs de animales, inhiben la traducción del ARNm blanco o disminuyen la vida media de los transcriptos a los que se unen (Sección 3.7.3) (Guil and Caceres, 2007; Ha and Kim, 2014).

En plantas también se ha demostrado la existencia de un mecanismo de represión traduccional, además del corte del transcripto (Aukerman and Sakai, 2003; Brodersen et al., 2008; Chen, 2004; Dugas and Bartel, 2008; Schwab et al., 2005). Se han propuesto varios modelos acerca de cómo lo miARNs de plantas inhiben la traducción, probablemente en al impedir la iniciación de la misma. Uno de ellos plantea que el complejo RISC-miARN-ARNm podría inhibir la traducción al secuestrar al ARNm en compartimentos subcelulares donde no haya ribosomas (Xu et al., 2006). Por otro lado existe una elevada cantidad de miARN asociados a polisomas lo cual sugiere que también podría haber inhibición de la traducción en pasos posteriores a la iniciación (Lanet et al., 2009). Un estudio *in vitro* mostró que AGO1 es capaz de inhibir la traducción sin inducir la deadenilación ni desestabilizar el ARNm. A su vez ensayos con mutantes catalíticas de AGO1 mostraron que cuando el RISC se une a la región 5' sin traducir o al marco abierto de lectura, el mismo es capaz de bloquear el reclutamiento de los ribosomas o el movimiento de los mismos sobre el ARNm. Por último el complejo RISC requeriría una mayor complementariedad de bases entre el ARNm y el miARN para inhibir la traducción que su contrapartida de animales, lo cual sugiere que los mecanismos por los cuales se inhibe la traducción en ambos reinos son diferentes (Tang et al., 2003). Trabajos recientes han asociado a la represión traduccional por miARNs en plantas con a complejos asociados al retículo endoplásmico. La asociación de ARNm que no codifican para proteínas de membrana y complejos AGO-miARN con el retículo endoplasmático abre nuevas perspectivas para entender los mecanismos de regulación traduccional en plantas (Li et al., 2016; Li et al., 2013).

3.7.3. Desestabilización del ARNm

Los miARNs de animales pueden promover la desestabilización de sus ARNm blanco al reclutar deadenilasas, las cuales acortan la cola de poliA del mismo o aumentando la accesibilidad de la cola de poliA a las deadenilasas. Luego de la deadenilación el ARNm blanco es degradado en sentido 5' a 3' (Braun et al.; Fukaya and Tomari, 2012; Moretti et al., 2012). El complejo RISC también puede promover la remoción de la caperuza al reclutar factores de decapping hacia al ARNm blanco (Makino et al., 2015; Nishihara et al., 2013). En contraste con los miARNs de animales, no se han encontrado casos de miARNs de plantas que promuevan la deadenilación de sus ARNm blanco (Ameres and Zamore, 2013; Iwakawa and Tomari).

3.8. Funciones de miARNs en plantas

Muchas familias de miARNs conservados (Sección 3.3) cumplen funciones esenciales para la biología de las plantas (Jones-Rhoades and Bartel, 2004). A pesar de que muchos genes blancos de miARNs de plantas son factores de transcripción, los miARNs también participan de la regulación de genes involucrados en la transducción de señales hormonales, la respuesta a estrés y a señales del ambiente (Rubio-Somoza et al., 2009; Rubio-Somoza and Weigel, 2011; Voinnet, 2009).

De las 21 familias de miARNs conservados en plantas, 11 regulan factores de transcripción y la mayoría de ellos están involucrados en procesos de desarrollo o diferenciación celular (Jones-Rhoades and Bartel, 2004). Por ejemplo, los genes blanco del miR172 pertenecen a la familia de factores de transcripción APETALA2 (AP2) que están involucrados en el tiempo de floración y en el desarrollo de la hoja (Aukerman and Sakai, 2003; Chen, 2004). Este miARN trabaja de manera coordinada con la familia miR156/157, estos son a su vez reguladores de los factores de transcripción SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) promoviendo la floración asociada a la edad de la planta (Wang, 2014). La familia del miR319 regula factores de transcripción TEOSINTE BRANCHED1, CYCLOIDEA, PCF1/2 (TCP), la familia del miR165/166 regula los factores de transcripción HD-ZIPIII y la familia del miR396 regulan factores de transcripción GROWTH-REGULATING FACTOR (GRF). Todas estas familias intervienen en el desarrollo de la hoja (Bowman, 2004; Palatnik et al., 2003; Rodriguez et al., 2010). El miR164, participa en el establecimiento del meristema apical del tallo, en el desarrollo de la raíz y definición de los bordes de los órganos mediante la regulación de miembros de la familia de factores de transcripción NAM, ATAF1/2 and CUC2 (NAC)(Laufs et al., 2004). La sobre-expresión de miR170/171 da como resultado plantas con menos hojas y más verdes que las plantas silvestres, e incluso con menor número de ramificaciones en el tallo. Lo cual sugiere que sus genes blanco los factores de transcripción SCAREDCROW-LIKE, GRAS participan en la regulación concertada de estos procesos con los miARNs mencionados anteriormente. (Rodriguez et al., 2016; Song et al., 2010).

Los miARN miR162 y miR168 tienen como genes blanco a DCL1 y AGO1, respectivamente, en un mecanismo de regulación que requiere de la participación de otros miARN y proteínas efectoras. DCL1 estaría bajo la regulación de miR162 (Xie et al., 2003) y miR414 (Liu et al., 2009). De manera que cambios temporales o espaciales en la expresión de estos miARNs, afectarían los niveles de DCL1 y por lo tanto los niveles globales de los miARNs. Los mismos principios se pueden aplicar en la regulación de AGO1 (Rhoades et al., 2002; Vaucheret et al., 2004). En particular la homeostasis de AGO1 es alcanzada por varios loops regulatorios que permiten que las rutas de los miARN y siARN funcionen correctamente. Estos loops incluyen el corte de ARNm de AGO1 mediado por miR168 (Vaucheret et al., 2004), el aumento en la acumulación de los niveles de miR168 cuando se incrementan los niveles de AGO1 (Vaucheret, 2008) y la represión traduccional de AGO1 mediada por AGO10 (Mallory and Vaucheret, 2009).

Por otra parte, entre los genes blanco que no corresponden a factores de transcripción ni proteínas de procesamiento, hay genes involucrados en diversos aspectos de la biología vegetal. Algunos codifican proteínas pertenecientes a la familia F-Box o enzimas E2 ubiquitin ligasas (miR393, miR394 y miR399, Tabla I1), las cuales están implicadas en el marcado de proteínas para la degradación por el proteosoma (Rubio-Somoza et al., 2009). Otros genes blanco con función conocida codifican para proteínas involucradas en el metabolismo energético, la respuesta a estrés o déficit de nutrientes (miR395, miR397, miR398 y miR408, Tabla 1)(Sunkar et al., 2012).

En esta Tesis nos enfocaremos principalmente en las familias de miARNs miR171, miR156, miR157 y miR319 que regulan factores de transcripción y procesos de desarrollo esenciales para plantas, como se mencionara anteriormente. Además, estos sistemas de miARNs y sus genes blanco están conservadas ampliamente entre distintas las distintas especies de plantas.

Tablal.1. miARNs conservados y sus genes blancos en plantas

miARNs	Genes blanco
miR156	Factores de transcripción SPL
miR159	Factores de transcripción MYB, NOZZLE
miR160	Factores de transcripción ARF
miR162	DCL1
miR164	Factores de transcripción NAC
miR166	Factores de transcripción HDZip
miR167	Factores de transcripción ARF
miR168	AGO1
miR169	Factores de transcripción CCAAT-HAP2
miR171	Factores de transcripción GRAS
miR172	Factores de transcripción AP2
miR319	Factores de transcripción TCP
miR390	TAS RNA
miR393	TIR1 proteins, F-BOX proteins
miR394	Proteínas F-BOX
miR395	ATP-sulfurilasas, transportadores de sulfato
miR396	Factores de transcripción GRF
miR397	Laccases
miR398	Cu/Zn SODs, CytC oxidase protein subunit, Chaperona de cobre
miR399	Enzima E2 de conjugación de ubiquitina
miR408	Blue copper proteins, Laccases, P-TYPE ATPase, PAC1 (componente del proteosoma)

4. OBJETIVOS

4.1. Objetivo general

Se pretende estudiar varios aspectos mecanísticos de la biogénesis de los miARNs en plantas. En general, se plantea: dilucidar cuales son los determinantes estructurales presentes en el precursor del miARN que guían la biogénesis del mismo. El enfoque de esta Tesis es caracterizar los mecanismos de procesamiento exclusivos de plantas.

Los resultados obtenidos no sólo serían importantes desde el punto de vista básico sino que también tendrían aplicación directa en el diseño de miARNs artificiales, utilizados en la terapéutica y la biotecnología.

4.2. Objetivos particulares

1. Optimizar el protocolo para construir bibliotecas de intermediarios de procesamiento en plantas silvestres para describir la biogénesis de los precursores de miARNs de plantas en forma sistemática.
2. Identificar secuencias y elementos estructurales dentro de los precursores de miARNs que sean claves para dirigir el procesamiento de los mismos.
3. Determinar el mecanismo de procesamiento de precursores de miARNs que aún no se hayan descrito.
4. Caracterizar el procesamiento de precursores de aquellos precursores que presenten características inusuales.
5. Desarrollar modelos que expliquen de manera global el procesamiento de los precursores de miARN de Arabidopsis.

MATERIALES Y METODOS

5. Materiales y Métodos.

5.1. Metodología para el clonado de fragmentos de ADN.

5.1.1. Cepas bacterianas empleadas.

Escherichia coli DH5 α (Hanahan, 1983): esta cepa fue empleada para el clonado de plásmidos, la amplificación de plásmidos y para guardar los mismos en un cepario.

Agrobacterium tumefaciens ASE: Esta cepa es resistente a los antibióticos cloranfenicol y kanamicina. Se la utilizó para transformar *Arabidopsis thaliana*.

5.1.2. Vectores plasmídicos utilizados.

Plásmido pBluescript II SK(-) (Stratagene): es el vector empleado para el clonado de fragmentos de PCR. Posee el sitio de múltiple clonado en el gen *lacZ* que codifica para el fragmento N-terminal de la β -galactosidasa. Este fragmento permite la selección de colonias bacterianas blancas o azules por α complementación, dependiendo de la presencia o ausencia del fragmento a clonar, respectivamente. Además posee el gen *BLA* (ApR), que codifica para la β -lactamasa, que le confiere resistencia al antibiótico ampicilina.

Plásmido pCHF3 (Jarvis et al., 1998). Vector binario que permite la transformación de *A.thaliana* mediada por *A. tumefaciens*. Posee orígenes de replicación para *E. coli* y *A. tumefaciens*. También posee el gen de resistencia a espectinomicina para la selección bacteriana. El T-ADN está definido por los bordes derecho (RB, del inglés “right border”) e izquierdo (LB, del inglés “left border”), y es la porción de ADN transferida al genoma de la planta durante las transformación. Esta región contiene el gen *NPTII*, que confiere resistencia a kanamicina en plantas, bajo el promotor 35S del virus del mosaico del coliflor. Además presenta un sitio de múltiple clonado (SMC) flanqueado por el promotor 35S del virus del mosaico del coliflor y el terminador de la transcripción de la subunidad pequeña de Rubisco de arveja.

Secuencias transcriptas del vector:

Hebra 5': AACACGGGGGACGAGCTCGGTACCCGGGGATCC.

Hebra 3': CTGCAGAGCTTTCGTTTCGTATCATCGGTTTCGACAACGTTTCGTC AAGTTCAATGCATCAGTT TCATTGC.

5.1.3. Minipreparaciones de ADN plasmídico.

Los plásmidos se obtuvieron utilizando el “kit” comercial “Wizard® Plus SV Minipreps” (Promega), según las indicaciones del fabricante.

5.1.4. Digestión de ADN con endonucleasas de restricción.

Las digestiones de ADN con endonucleasas de restricción se realizaron utilizando enzimas y buffers comerciales. Las condiciones de reacción fueron las especificadas por el fabricante.

5.1.5. Electroforesis de ADN en geles de agarosa.

La separación de moléculas de ADN por electroforesis en geles de agarosa se realizó mediante el sistema tipo submarino en cubas de acrílico (Sambrook et al., 1989). Se utilizaron concentraciones de agarosa de 1 a 3% (p/v) según el tamaño de las moléculas de ADN a resolver. Los geles se prepararon con solución buffer 1X TAE y se agregó bromuro de etidio a una concentración final de 0,3 µg/ml. Previo a la siembra, las muestras de ADN se mezclaron con 1/5 volúmenes de solución de siembra. Las corridas electroforéticas se llevaron a cabo en solución buffer 1X TAE, empleando voltaje constante. Se utilizaron marcadores de tamaño molecular adecuados.

5.1.6. Purificación de fragmentos de ADN a partir de geles de agarosa.

Luego de la separación electroforética, el fragmento de ADN deseado se identificó mediante la visualización del gel con un transiluminador UV y se escindió la banda correspondiente empleando un bisturí. La purificación de los fragmentos de ADN del gel se llevó a cabo utilizando el “kit” de purificación de ADN Wizard® SV Gel and PCR Clean-Up System (Promega).

5.1.7. Ligación de fragmentos de ADN.

Para la ligación de fragmentos de ADN en los vectores plasmídicos deseados se siguieron dos estrategias según el origen del fragmento.

a) Clonado de productos de PCR por el método de “restricción-ligación”.

Teniendo en cuenta que la amplificación de fragmentos de ADN mediante PCR se llevó a cabo utilizando ADN polimerasas que dejan extremos romos, el clonado de dichos fragmentos se realizó según el siguiente protocolo. Primero se digirió el vector de clonado pBluescript II SK(-) con la enzima de restricción SmaI, la cual deja extremos romos. Luego se procedió a la ligación del fragmento de ADN en la siguiente mezcla de reacción: en un volumen final de 21 µl se adicionó 20 ng de pBluescript II SK(-) cortado previamente con SmaI, 2 µl de Buffer Tango (Fermentas), 2 U de T4 ADN ligasa, 3 U de SmaI (Fermentas), ATP a una concentración final de 0,09mM y aproximadamente 20 ng de inserto. La mezcla de ligación se incubó toda la noche a 16°C, y se utilizaron 5µl de la misma para transformar *E. coli* DH5α.

b) Ligación de fragmentos de ADN generados por corte con enzimas de restricción (empleado para subclonados)

En los vectores plasmídicos correspondientes ya cortados se llevó a cabo con la enzima T4 ADN ligasa (Fermentas). Las condiciones de la reacción fueron las especificadas por el fabricante. En todos los casos el volumen de reacción de ligación fue de 10ul y se utilizaron 5µl de la misma para transformar *E. coli* DH5α.

5.1.8. Transformación bacteriana.

a) Transformación de *E. coli*.

A partir de un cultivo saturado de *E. coli* DH5α, crecido toda una noche a 37°C en medio Luria-Bertani (LB; 10 g/l peptona de caseína, 5 g/l extracto de levadura y 5 g/l NaCl), se realizó una dilución 1:100 en el mismo medio y se incubó 2 horas a 37°C. Por cada transformación se tomó una alícuota de 1,5 ml de cultivo en tubos de microcentrífuga y se centrifugó durante 2 min. a 2600g a temperatura ambiente. El sedimento de células fue resuspendido suavemente en 1,5ml de 0,1M CaCl₂ frío. Luego se centrifugó nuevamente 2 min. a 2600g a temperatura ambiente. El sedimento de células fue resuspendido suavemente en 100 µl 0,1M CaCl₂ frío y sobre esto se adicionó el volumen correspondiente de solución plasmídica. Esta mezcla se incubó en hielo durante una hora, luego 90 segundos a 42°C y finalmente 2 min. en hielo. A continuación se adicionó 900 µl de medio LB y se incubó durante una hora a 37°C, de modo que el marcador de selección presente en el plásmido se exprese. Transcurrida la incubación se realizó una centrifugación de 2 min. a 2600g a temperatura ambiente, el sedimento de células fue resuspendido en 100 µl de medio LB y finalmente se sembró en placas de medio LB que contenían el antibiótico correspondiente. Las placas se incubaron toda la noche a 37°C.

b) Transformación de *A. tumefaciens*.

Un cultivo de *A. tumefaciens* (cepa ASE), crecido toda la noche a 28-30 °C en medio LB, fue diluido 1:100 en el mismo medio y cultivado hasta alcanzar una DO₆₀₀ entre 0,5 y 1. La suspensión de células fue fraccionada en alícuotas de 1,5 ml, cada una de las cuales sirve para realizar una transformación. Se centrifugó a 10.000 g por 1 min. a 4 °C, y se resuspendió el sedimento de células en 0,5 ml de agua destilada estéril a 4 °C. Se repitieron estos lavados 3 veces. Luego se resuspendieron las células en 0,5 ml de 10 % (v/v) glicerol estéril a 4 °C. Se centrifugó nuevamente a 10.000 g por 1 min. a 4 °C y se resuspendieron las células en 20 µl de 10 % (v/v) glicerol estéril. Se transfirió a una cubeta de electroporador (Biorad) y se aplicó un pulso de 2,5 kV con capacitancia de 25 µF y una resistencia de 400 Ω. A continuación se adicionó 1 ml de medio LB y se incubó durante dos horas a 28°C. Transcurrida la incubación se realizó una centrifugación de 2 min. a 2600 g a temperatura ambiente, el sedimento de células fue resuspendido en 100 µl de medio LB y finalmente se sembró en placas de medio LB que contenían el antibiótico correspondiente. Las placas se incubaron toda la noche a 28°C.

5.1.9. Selección y análisis de transformantes.

La selección de las colonias *E. coli* DH5 α transformadas se realizó suplementando las placas con medio LB con el antibiótico adecuado según el vector usado. Cuando el vector fue pBluescript II SK(-) las placas se suplementaron 100 $\mu\text{g/ml}$ de ampicilina. Además, se adicionó a las placas 100 μM IPTG y 20 $\mu\text{g/ml}$ X-Gal. Esto permitió identificar potenciales colonias que transformadas con el vector que contenía el inserto clonado por su ausencia de coloración azul.

Para bacterias (*E.coli* DH5 α) transformadas por el vector pCHF3, la selección se llevó adelante con 100 $\mu\text{g/ml}$ de espectinomicina. Para la selección de las colonias ASE transformadas con el vector binario pCHF3 las placas con medio LB se suplementaron con 50 $\mu\text{g/ml}$ de espectinomicina, 50 μg de kanamicina y 25 μg de cloranfenicol.

5.2. Mutagénesis sitio dirigida.

Para realizar las mutagénesis dirigidas se utilizó la técnica “Extensión por Superposición” (del inglés, Overlap Extension; (Pogulis et al., 1996). Este método permite introducir inserciones o deleciones sobre la región de interés mediante Reacción en Cadena de la Polimerasa (PCR) utilizando cuatro o más cebadores que se solapan entre sí.

5.2.1. Reacción en Cadena de la Polimerasa (PCR)

Las reacciones de amplificación que involucraron mutagénesis, se llevaron a cabo bajo las siguientes condiciones: en un volumen final de reacción de 25 μl se adicionó 2,5 μl de buffer *Pfu* ADN polimerasa con MgSO_4 10x (Promega), 7,5 pmol de cada oligonucleótido correspondiente, 0,1 μg de ADN plasmídico o producto de PCR y 2 U de *Pfu* ADN polimerasa. La concentración en la mezcla de reacción de cada uno de los dNTPs (dCTP, dATP, dGTP, dTTP) fue de 0,2 mM. Se completó el volumen restante con agua Milli-Q. El termociclado de las mezclas de reacción se llevó a cabo en un equipo PCR SPRINT Thermal Cycler® (Termo Electron Corporation) según el siguiente programa:

Etapa	Ciclos	Temperatura	Tiempo
1	1	95°C	30 seg
2	30	95°C	30 seg
		55°C	30 seg
		68°C	1 min
3	1	68°C	10 min

5.2.2. Clonado de los fragmentos mutantes

Para confirmar la obtención de los productos del tamaño esperado se corrió una alícuota de la reacción finalizada en un gel de agarosa teñido con Bromuro de Etidio con un marcador de peso molecular adecuado. Luego de la purificación de dichos fragmentos del gel, se procedió al clonado de los mismos en el vector pBluescript II SK(-) como se detalla en la Sección 5.1.7.a

Las mezclas de ligación se transformaron en cepas de *E. coli* DH5 α y se procedió como se detalla en la Sección 5.1.9. A partir de 3 colonias blancas seleccionadas en este medio para cada inserto se realizaron inóculos en medio LB líquido con ampicilina. La presencia de los insertos en el vector se verificó en estos clones mediante el patrón de restricción generado luego del corte de ADN plasmídico y se secuenciaron a fin de verificar que la secuencia amplificada carecía de errores.

5.3. Obtención de *A. thaliana* transformadas.

5.3.1. Ecotipo utilizado y condiciones de crecimiento.

Las plantas de *A. thaliana* utilizadas en esta tesis corresponden al ecotipo Columbia-0 (Col-0). A menos que se indique lo contrario, las mismas fueron cultivadas en una cámara de crecimiento con un régimen de 16 hs. De luz (100 μ E.m⁻².s⁻¹) y 8 hs. de oscuridad. La temperatura de crecimiento fue de 23 °C durante el ciclo luz/oscuridad, mientras que la humedad fue mantenida en 65 % de humedad relativa. Las plantas fueron regadas dos veces por semana con agua.

5.3.2. Transformación de *A. thaliana*.

La transformación de plantas de *A. thaliana* fue realizada mediante la técnica de “floral-dip” (Clough and Bent, 1998). Se crecieron las cepas de *A. tumefaciens* ASE transformadas con los distintos plásmidos binarios en 5 ml de LB conteniendo 50 μ g/ml de espectinomicina, 50 μ g de kanamicina y 25 μ g de cloranfenicol a 28 °C hasta llegar a una DO600 cercana a 2. Se recuperaron las células por centrifugación a 5.000 *g* durante 10 min. a temperatura ambiente y se resuspendieron en un medio con 5 % (p/v) sacarosa y 200 μ l/l de Silwet L-77.

Las inflorescencias de plantas de *A. thaliana* regadas el día anterior, fueron sumergidas en la suspensión de células durante 30 seg.. Luego de escurrir el exceso de medio de cultivo, se colocaron las plantas en posición horizontal dentro de bolsas plásticas para conservar la humedad. Luego de 24 hs. se sacaron de las bolsas y se prosiguió con el cultivo en la cámara de crecimiento hasta la maduración de las semillas.

5.3.3. Recolección de semillas.

La recolección de semillas se realizó a partir de plantas que finalizaron su ciclo de vida y que presentaban semillas maduras. Se retiró las plantas del contenedor de tierra y frotando las mismas con las manos se dejó caer las semillas sobre una hoja de papel. Se eliminaron los restos de silicua y hojas por tamizado. Las semillas se almacenaron en bolsas de papel en un recipiente seco hasta su posterior utilización.

5.3.4. Esterilización de semillas.

Para la esterilización de las semillas, las mismas fueron incubadas con una solución de 70% (v/v) etanol, 0,1% (v/v) tritón-X-100, durante al menos 5 y a los sumo 15 min.. Luego se realizó un enjuague con 96% (v/v) etanol y finalmente se las dejó secar en el flujo laminar. La manipulación del material vegetal en condiciones de esterilidad se realizó en una cabina de flujo laminar.

5.3.5. Selección de transformantes.

Para la selección de plantas transformadas, semillas esterilizadas fueron sembradas en placas de Petri con medio MS-agar (Murashige and Skoog, 1962) suplementado con 25 µg/ml de kanamicina. Las placas fueron mantenidas en oscuridad a 4°C por 48 horas. Transcurrido este periodo las placas fueron transferidas a la cámara de cultivo y en estas condiciones las plantas se dejaron crecer 10 días.

Luego las placas fueron analizadas y aquellas plantas que presentaban un fenotipo resistente al agente de selección, evidenciado por el desarrollo de raíz y la aparición del primer par de hojas verdaderas, fueron transferidas a contenedores con tierra. Los mismos, fueron ubicados en bandejas con agua las cuales se cubrieron con film de nylon transparente, de modo de evitar la pérdida de humedad. En estas condiciones se cultivaron en la cámara durante 2 días, momento en el que se retiró el nylon y se continuó con su crecimiento.

5.4. Blot de ARN pequeños.

5.4.1. Extracción de ARN.

Para realizar la extracción primeramente se realizó la recolección de muestras vegetales. Se recolectaron inflorescencias de al menos 15 líneas T1 transgénicas independientes que sean representativas de la población de plantas para cada muestra, por duplicado.

La extracción de ARN de tejido vegetal de *A. thaliana* se realizó utilizando el reactivo TRIzol (Invitrogen). La recolección de las muestras de tejido se realizó en crioviales que inmediatamente fueron sumergidos en N₂ líquido. El material fue reducido a un polvo fino utilizando un pilón. A continuación, sobre este polvo se adicionó TRIzol (1 ml de reactivo para un máximo de 100 mg de tejido) y se agitó de

modo de resuspender el tejido mortereado. El homogenado se centrifugó a 12000 g a 4 °C durante 10 min. y finalmente se transfirió el sobrenadante a un tubo de microcentrífuga nuevo.

A continuación se agregaron 0,2 volúmenes de cloroformo por cada volumen de TRIzol original. Esta mezcla se invirtió vigorosamente por 15 segundos y se centrifugó a 12000 g a 4 °C durante 15 min.. Se transfirió la fase superior a un nuevo tubo de microcentrífuga y se precipitó el ARN mediante el agregado de 0,5 volúmenes de isopropanol por cada volumen de TRIzol inicial. Esta mezcla se incubó por 10 min. a temperatura ambiente. Luego se centrifugó a 12000 g a 4 °C durante 10 min..

Finalizada la centrifugación se descartó el sobrenadante. Se lavó el precipitado de ARN mediante el agregado de 1 ml de 70 % (v/v) etanol frío y agitación con vortex. Luego se centrifugó a 7500 g a 4 °C durante 5 min., descartando el sobrenadante una vez finalizada la centrifugación. Este paso de lavado se repitió una vez más. El precipitado de ARN obtenido se secó en estufa 37 °C por 10 min. y luego fue resuspendido en 50 µl de agua Milli-Q para ARN.

5.4.2. Cuantificación del ARN purificado espectrofotométricamente.

La cuantificación se determinó espectrofotométricamente midiendo la absorbancia a 230, 260 y 280 nm. Se estimó la pureza de la preparación a partir de la relación de las medidas de absorbancia a Abs_{260}/Abs_{230} y Abs_{260}/Abs_{280} .

5.4.3. Chequeo de la integridad del ARN purificado.

La integridad del ARN purificado se determinó mediante la electroforesis en geles de 1,5% (p/v) agarosa de 5 µl del ARN preparado. Las bandas de ARN ribosomal se visualizaron por tinción de los geles con bromuro de etidio.

5.4.4. Armado del gel y corrida electroforética.

Se procedió a armar los geles de poliacrilamida al 17% en condiciones desnaturalizantes según las siguientes instrucciones:

37% Acryamida/ 1% Bis-acrylamida (Sigma: A6050)	10,6 ml
Urea	10,5g
H ₂ O (autoclavada dos veces)	1,75 ml
TBE 5X	12,5ml
Temed	12,5ml
APS 10% (peso/peso)	50ml

La polimerización se llevó a cabo por aproximadamente 1 hora después de la cual se procedió a pre-correr el gel por 1 hora a 180 V (E844 – CONSORT) en TBE 1X. Posteriormente, se prepararon las muestras a sembrar. Para ello, se mezcló 1 volumen de colorante de siembra y 2 volúmenes de formamida. 1,5 volúmenes de esta preparación se mezclaron con 1 volumen de la muestra de ARN (2 – 6 ug). Se incubó durante 4 min. A 95 °C y luego se enfrió en hielo.

Se sembraron las muestras en el gel y se realizó la corrida electroforética del mismo a 180 V y durante dos horas. Una vez finalizada la corrida, se tiñó el gel con 1X TBE y Bromuro de Etidio por 5 min. Se tomó una foto y el gel se destiñó con 1X TBE. Se transfirió el gel a una membrana “Nytran Super Charge” (NYLON, Schleicher & Schuell BioScience). Para ello, se empapó la misma junto con dos papeles “Whatman” (Axygen®) con 1X TBE.

Luego, se realizó la transferencia en 1X TBE a 300 mA durante 1 hora (Trans-Blot® SD Semi Dry Transfer Cell – Biorad). Terminado esto, se realizó el crosslinking de la membrana durante 5 min. Con UV (UV Transilluminator® - UVP).

5.4.5. Marcado radioactivo de la sonda

El LNA (del inglés, locked nucleic acid) es un nucleótido de ARN modificado en cual se genera un enlace que une al oxígeno del C2 con el C4. Esta modificación en la estructura, produce un aumento en las propiedades de hibridación del oligo LNA. Entre ellas, un incremento en la estabilidad térmica del dúplex y capacidad de discernir entre secuencias similares. Mediante síntesis química se pueden generar oligos que poseen bases intercaladas de LNA y ADN o ARN.

Sonda LNA 319: 5' AGGGAGCTCCCTTCAGTCCAA 3'

Sonda LNA miR319a-3p.2: 5' TTGTCTACCGCATCATTCAT 3'

Las bases LNA estas subrayadas.

Otras sondas empleadas fueron de oligos sin modificar de secuencia complementaria del ARN pequeño que se quisiera detectar.

Sonda171: 5'CGTGATATTGGCRCGGCTCAATCA3'

La R representa las bases A o G. Por ende esta sonda permite la detección simultánea de miR170/171.

Sonda 157: 5'GTGCTCTCTATCTTCTGTCAA3'

SONDA miR319a-3p.1: 5'TCAAAGAGAATCAATGATCCAA3'

Para llevar adelante el marcado radioactivo de la sonda, se preparó la siguiente mezcla de reacción:

Oligonucleótido DNA/LNA (10µM)	2 µl
--------------------------------	------

Buffer polinucleótido quinasa 10X	1 μ l
H ₂ O	3 μ l
[³² P] γ -ATP (6000ci/mmol; 10 mCi/ml)	3 μ l
Polinucleótido quinasa T4	1 μ l

Se incubó a 37°C durante 1 h, luego se agregó 40 μ l dH₂O. El oligo marcado se purificó por cromatografía de exclusión molecular (BioRad columns, MICROBIO-SPIN 6, Cat. No. 7326200).

5.4.6. Hibridación.

Se pre-hibridó la membrana a 38 °C durante 1 hora utilizando PerfectHyb™ (Sigma). Se agregó la sonda (5 pmol) a 100 μ l de PerfectHyb™ y se calentó la mezcla a 65 °C durante 10 min. tras lo cual se enfrió en hielo y centrifugó. Se agregaron 10 ml de PerfectHyb™ y se dejó hibridando la membrana con esta mezcla a 38 °C durante toda la noche.

Terminada la hibridación, se procedió a enjuagar la membrana con 4X SSC. Luego se realizaron dos lavados durante 20 min. con 4X SSC seguidos de un tercer lavado con 2X SSC y 0,1 % SDS durante 15 min.

5.4.7. Detección radioactiva

Se secó por unos minutos la membrana al aire libre. Luego se la cubrió con papel de film y fue puesta en exposición en una placa autoradiográfica.

5.5. Determinación de intermediarios.

5.5.1. Construcción biblioteca 5' RACE PCR.

Se realizó una extracción de ARN total (según punto 5.4.1) de plantas sobre-expresantes del precursor mutante al cual se desean identificar los intermediarios. Luego a partir del ARN total se llevó a cabo una purificación de ARNm mediante kit comercial PolyAtract® mRNA Isolation System III (Promega) según las indicaciones del fabricante.

A continuación se precipitó al ARNm mediante el agregado de 2 μ l de glucógeno (10mg/ μ l), 10 μ l de Acetato Sodio 3M pH 5.2 y 220 μ l de etanol absoluto. Se incubó toda la noche a -20°C. Se centrifugó 20 min a 4 °C a máxima velocidad y se lavó con etanol 80%. Luego se centrifugó 5 minutos a 4 °C a máxima velocidad. Se removió el excedente y se secó 10 minutos a 37°C. Posteriormente se resuspendió en 10 μ l de agua Milli-Q para ARN.

Luego se realizó una ligación del Oligo Adaptador de ARN mediante el siguiente protocolo: se mezcló 250ng de ARNm con 0,25 μ l de oligo ARN (100 μ M). Se incubó 5 min a 65°C. Se dejó en hielo 2 min. Se agregó a la mezcla de reacción:

10X Buffer Ligasa	1 μ l
10mM ATP	1 μ l
BSA	1 μ l
RNaseOut (40U/ μ l)	1 μ l
T4 RNA ligasa (5U/ μ l)	1 μ l
H ₂ O csp	10 μ l

Se incubó a 37°C durante 1 hora. Se centrifugó, y precipitó el ARNm mediante el protocolo descrito previamente, y se resuspendió en H₂O Milli-Q para ARN. Luego, se trató el ARN con ADNasa, con el objetivo de eliminar potenciales contaminaciones con ADN genómico. Para ello se preparó la siguiente mezcla de reacción: en un volumen final de reacción de 20 μ l se adicionó: 2 a 4 μ g de ARNm total, 2 μ l de buffer "RQ1 RNase- Free DNase" (Promega), 1 U "RQ1 RNase-Free DNase" (Promega) y agua Milli-Q para completar el volumen de reacción. La mezcla de reacción se incubó 30 min. a 37 °C. Luego se inactivó la ADNasa mediante la adición al tubo de reacción de 1 μ l de "DNase Stop Solution" (Promega) e incubación por 10 min. a 65°C. Se usaron 12 μ l para sintetizar el ADN complementario (ADNc), de los 21 μ l de reacción.

Posteriormente se realiza una retrotranscripción (RT) utilizando un oligo específico que hibrida en una región transcripta 3' de pCHF3 a modo de determinar solo los intermediarios de precursores de miARN transgénicos expresados bajo el vector CHF3 y descartando los intermediarios endógenos silvestres. La síntesis del ADNc se llevó cabo según el siguiente protocolo. En un volumen final de 25 μ l se adicionó 0,25 μ g de oligo RT CHF3-1, 12 μ l de ARN tratado con ADNasa, dNTPs a una concentración final de 0,4 Mm cada uno. El volumen de la mezcla se completó con agua Milli-Q. La mezcla de reacción se incubó durante 5 min a 65°C y a continuación en hielo durante al menos un minuto. Luego se realizó una centrifugación rápida y se adicionó a cada tubo de reacción 4 μ l de "5X First Strand Buffer", 1 μ l de 0,1 M DTT, 1 μ l "Rnase OUT Recombinant Ribonuclease Inhibitor" (Invitrogen) y 100 U de "SuperScript III Reverse Transcriptase". Se mezcló con micropipeta y se incubó 60 min. a 50 °C. Luego se realizó una centrifugación rápida y se inactivó la reacción mediante la incubación a 70 °C durante 15 min.. Finalmente se realiza la eliminación del exceso de oligo ARN GeneRacer incubando 30 min a 37°C con 0,04 μ g/ μ l de RNasa A.

Secuencia de oligo adaptador ARN:

5' CAGCUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA 3'

Oligo RT CHF3-1: 5' GTGCGCAATGAAACTGATGC 3'

5.5.2. Amplificación de intermediarios.

Se realiza una PCR con un oligo que hibrida sobre la secuencia del Oligo Adaptador ARN y un oligo que hibrida específicamente sobre la secuencias del CHF3 transcriptas 3' del precursor. Posteriormente con el objetivo de aumentar la especificidad y la eficiencia de la detección de intermediarios de procesamiento se realiza una PCR anidada.

Oligo general Adaptador: 5'- CGACTGGAGCACGAGGACACTGA-3'

Oligo general Adaptador anidado: 5'-GGACACTGACATGGACTGAAGGAGTA-3'

Oligo específico CHF3: 5'- GTGCGCAATGAAACTGATGC-3'

Oligo específico CHF3 anidado: 5'-CGAAACCGATGATACGAACG-3'

5.6. Optimización del método SPARE para análisis genómico de intermediarios de procesamiento de microARNs

5.6.1. Material vegetal.

Como material vegetal de partida se emplearon plántulas de *Arabidopsis thaliana* crecidas durante 10 días con luz continua. De este modo no hay variaciones asociadas al ritmo circadiano durante la colecta de muestra. En paralelo se crecieron las siguientes plantas mutantes en *HYL1*, *SE*, y *FIERY1*.

5.6.2. Protocolo SPARE multiplex.

5.6.2.1. Extracción de ARN

Se siguieron los pasos indicados en las Secciones 5.4.1 a 5.4.3.

5.6.2.2. Tratamiento con ADNasa

En este punto se llevan todos los ARN a la misma concentración final.

1. Preparar la siguiente reacción:

ARN (40ug de RNA en 89 ul finales)	
Buffer DNasa	10µl
DNasa (Promega)	1 µl

2. Incubar 30 minutos a 37 °C

3. Agregar 1 µl de ADNas Stop Solution

4. Incubar 10 minutos a 65 °C

5.6.2.3. Precipitación de ARN.

1. Agregar:

Glucógeno (20mg/ul)	1 ul
Acetato Sodio 3M pH 5.2	10 ul
Etanol absoluto	220 ul

2. Incubar durante la noche a -20°C.

3. Centrifugar 30 minutos a 4 °C a máxima velocidad.

4. Descartar sobrenadante.

5. Lavar con etanol 80%.

6. Centrifugar 5 minutos a 4 °C a máxima velocidad.

7. Descartar SN. Repetir el paso 5 una vez.

8. Spin y remover excedente.

9. Secar 5-10 minutos a 37°C.

10. Resuspender en 9ul agua.

5.6.2.4. Ligación con Oligo ARN.

1. Mezclar 9ul de RNA con 1ul oligo de ARN (200 uM).

2. Incubar 5 minutos a 65°C

3. Poner en hielo aproximadamente 2 minutos

4. Agregar a la mezcla de reacción (volumen final: 20 ul)

10X Buffer Ligasa	2 ul
10mM ATP	2 ul
BSA	2 ul
ARNsaOut (40U/ul)	2 ul
T4 ARN ligasa (5U/ul)	2 ul

5. Incubar a 37°C durante 1 hora

6. Centrifugar y poner en hielo.

7. Inactivación de la ligasa: 72°C durante 15´.

5.6.2.5. Purificación de ARNm con dynabeads

1. Seguir las instrucciones del kit

2. Resuspender en 12 ul agua.

5.6.2.6. Transcripción Reversa.

1. Preparar la siguiente reacción:

ARNm-ligado purificado: 12ul (dividir en 8 tubos de PCR con 4ul en cada uno).

dNTP mix 10 mM:1 µl.

MixOligoespecifico: 0,5ul (en la Sección 4.6.9 se detalla la preparación de la mix de oligos).

En total hay 8 mix de oligos diferentes con entre 15 y 16 oligos cada una (ANEXO Tabla A1).

2. Calentar a 65 °C durante 5 minutos

3. Incubar en hielo al menos 1 minuto

4. Spin down

5. Agregar:

5X First Strand Buffer	4 µl
------------------------	------

DTT 0,1 M	1 µl
-----------	------

Inhibidor de RNasa (Invitrogen)	1 µl
---------------------------------	------

SuperScript III	0,5 µl
-----------------	--------

Volumen total de: 20ul

6. Incubar 60 minutos a 50 °C

7. Spin down

8. Inactivar la reacción 15 minutos a 70 °C

9. Preparación de la mix de oligos:

Los stocks de oligos son 100uM, y en cada reacción de síntesis de ADNc se retrotranscriben 15 precursores. Para preparar cada una de las 8 mix de los oligos se coloca en un eppendorf 1ul de cada oligo y se completa con cantidad suficiente de agua mili-Q para 20ul. De esta dilución 1:20 se toman 10ul y se lleva a un volumen final de 100ul, para obtener una dilución 1:200. Este procedimiento se repite para obtener una dilución 1:2000 de la que se toman 0,5ul de para realizar la síntesis de ADNc.

5.6.2.7. PCRs

1. PCR1: el objetivo es enriquecer los intermediarios de procesamiento, sin la interferencia de los oligos indexados de gran tamaño que generan una gran cantidad de dímeros de primer. Se emplean los oligos FW1 y RV1 cuyas secuencias se detallan abajo.

2. PCR2: La función de la PCR2 es incorporar los oligo indexado al producto de la PCR1 (Tabla M1). En la PCR2 se usan los oligos FW2 y RV2 cuyas secuencias se detallan abajo y que son anidados a los oligos FW1 y RV1.

Como el protocolo completo involucra dos rondas de PCR, ambas se optimizaron (Sección 5.7 Puesta a punto de las reacciones de PCR.). El objetivo era que la técnica fuera semi-cuantitativa de modo que las lecturas obtenidas sean representativas de la abundancia real de los intermediarios de procesamiento. Luego de cada PCR se corre un gel de agarosa al 2% para evidenciar los productos de la reacción.

Oligo FW1: 5'-GTTTCAGAGTTCTACAGTCCGA-3' (contenida en el oligo FW2).

Oligo RV1: 5'-TGGAATTCTCGGGTGCCAAGG-3' (contenido en el oligo indexado, RV2)

Oligo FW2: 5'-AATGATACGGCGACCACCGACAGGTTTCAGAGTTCTACAGTCCGAC-3'. La región subrayada coincide en secuencia con el oligo FW1.

Oligo indexado, RV2:

5'-TGGAATTCTCGGGTGCCAAGGAACTCCAGTCACNNNNNNATCTCGTATGCCGCTTCTGCTTG-3'

La región **NNNNNN** es la porción variable de los oligos indexados que sirve para identificar de que biblioteca provino el fragmento. La región subrayada coincide en secuencia con el oligo RV1.

Todas estas secuencias deben ser removidas luego de la secuenciación, ya que son comunes y no aportan información acerca del procesamiento de los microARNs.

Tabla M1. Secuencias correspondientes a los 48 índex disponibles.

Índex	Secuencia	Índex	Secuencia	Índex	Secuencia
RPI1	ATCACG	RPI17	GTAGAG	RPI33	TCATTC
RPI2	CGATGT	RPI18	GTCCGC	RPI34	TCCCGA
RPI3	TTAGGC	RPI19	GTGAAA	RPI35	TCGAAG
RPI4	TGACCA	RPI20	GTGGCC	RPI36	TCGGCA
RPI5	ACAGTG	RPI21	GTTTCG	RPI37	GCCGCG
RPI6	GCCAAT	RPI22	CGTACG	RPI38	GCCTTA
RPI7	CAGATC	RPI23	GAGTGG	RPI39	GCTCCA
RPI8	ACTTGA	RPI24	GGTAGC	RPI40	GGCACA
RPI9	GATCAG	RPI25	CTAGCT	RPI41	GGCCTG
RPI10	TAGCTT	RPI26	CTAGCT	RPI42	TCTACC
RPI11	GGCTAC	RPI27	CTATAC	RPI43	TGAATG
RPI12	CTTGTA	RPI28	CTCAGA	RPI44	TGCCAT
RPI13	AGTCAA	RPI29	CTGCTG	RPI45	TGCTGG

RPI14	AGTTCC	RPI30	TAATCG	RPI46	TGGCGC
RPI15	ATGTCA	RPI31	TACAGC	RPI47	TTCGAA
RPI16	CCGTCC	RPI32	TATAAT	RPI48	TTCTCC

5.6.2.8. Purificación de la banda con el kit Wizard SV gel and PCR Clean-up system.

Siguiendo las instrucciones del fabricante se purifican las bibliotecas y se eluyen en un volumen de 20 ul de agua ultra pura. La efectividad de la purificación puede corroborarse corriendo una alícuota de 10ul en un gel de agarosa 1,5%, mientras que el remanente se envía a secuenciar.

5.6.3. Puesta a punto de las reacciones de PCR.

A la hora de iniciar la puesta a punto de este protocolo había ciertos requisitos que debían cumplirse para que la técnica fuera de utilidad:

- Reproducibilidad entre los duplicados.

Se decidió emplear una elevada cantidad de molde (ADNc) para llevar adelante las reacciones de PCR. De este modo se disminuye el error en el pipeteo y se espera asegurar la presencia de los fragmentos menos abundantes. A su vez se decidió hacer única reacción de PCR mezclando todos los ADNc correspondientes a la misma biblioteca con el objetivo de reducir la variabilidad durante las reacciones de PCR1.

La puesta a punto se inició con 5uL de ADNc por cada mix de síntesis de ADNc, para un total de 40uL. Se seleccionó el ADNc proveniente de una muestra de Col-0 ya que queríamos poner a punto el protocolo para las plantas control. Por último se quería usar el menor número de ciclos durante la PCR1 que diera amplificación en la PCR2.

- Eficiente en términos de la secuenciación.

Generalmente en técnicas de secuenciación masiva, se consume un gran número de lecturas en dímeros de primers y oligos que quedan sin utilizar durante pasos de construcción de la biblioteca. El objetivo era disminuir estas lecturas al mínimo, para ello se decidió realizar una purificación a partir de gel entre la primera y segunda PCR. Y una segunda purificación luego de la segunda PCR.

De este modo se esperaba eliminar los oligos necesarios para la síntesis del ADNc, los oligos remanentes y dímeros de primer de la primer y segunda PCR. Para delimitar las regiones a purificar se calcularon los tamaños mínimos esperados en ambas PCRs. Esto se hizo a partir de las distancias conocidas de los oligos y los mecanismos de procesamiento para cada precursor de Arabidopsis (Bologna et al., 2013b). Para la PCR1 se cortará el gel en las regiones comprendidas entre 100 y 700pb, mientras que para la PCR2 se cortará el gel en las regiones comprendidas entre 150 y 800pb. Estas regiones fueron delimitadas mediante el uso de marcadores de peso molecular específicos.

5.6.3.1. PCR1.1

Lo primero que se evaluó fue la concentración de oligos y el número de ciclos a emplear durante la PCR1. Se trabajó con la ADN polimerasa TaKaRa Ex Taq (A la cual nos referiremos de ahora en más como Taq).

MM

Oligo FW1: 0,5 uL (1uM)

Oligo RV1: 0,5 uL (1uM)

dNTPs: 8uL

buffer: 10uL

TaKaRa Ex Taq: 0,04uL

cDNA: 40uL

Agua: 41,6uL

Volumen final: 100uL

Los primeros 5 ciclos se hicieron 10% de oligo (1ul de oligo mix 1uM) y Taq polimerasa. Luego se agregó el 90% restante de oligo mix y Taq polimerasa (4,5uL de cada oligo y 0,36uL de enzima) para los siguientes 25 ciclos. En los ciclos 10, 15, 20 y 30 se detuvo la reacción y se tomaron 25ul del volumen de reacción y se corrió un gel de agarosa al 2%, concentración empleada de rutina a menos que se indique lo contrario. En la Figura M1 se muestran los resultados de la PCR1.1.

Programa de PCR:

3' 94° x 1 ciclo

30'' 94°

30'' 55° X 30 ciclos

50'' 72°

5' 72° x 1 ciclo

A menos que se indique lo contrario este el programa de PCR empleado para la PCR1.

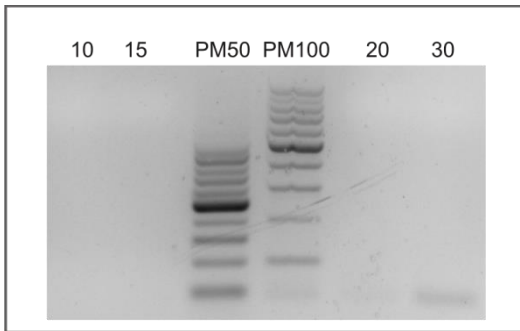


Figura M1. Gel de agarosa al 2% con los productos de la PCR1.1

El gel se corrió durante 30 minutos con dos marcadores de peso molecular de 50 y 100pb. Los nombres de las calles indican el número de ciclos de la PCR1.1 en el cual se tomó la alícuota.

A pesar de no observar amplificación en ninguna calle se cortaron las regiones comprendidas entre 100 y 700pb y se purificaron como se indicó en la Sección 5.6.2.5. con la excepción de que el volumen de elusión será de 30uL. Este procedimiento se repetirá luego de la PCR1 a menos que se indique lo contrario.

A partir de 2uL de PCR1 purificada, se realizó la PCR2 a 10 y 30 ciclos, para cada una de las condiciones (es decir, ciclos) de la PCR1. Durante la puesta a punto de la PCR2 se usarán los oligos FW1 y RV1, en lugar de los indexados, a menos que se indique lo contrario.

MM (PCR2)

Oligo FW1: 1,25 uL (1uM)

Oligo RV1: 1,25 uL (1uM)

dNTPs: 4uL

buffer: 5uL

ExTaq: 0,25 uL

moldePCR1: 2uL

Agua: 36,25uL

Volumen final: 50 uL

Programa de PCR

3' 94° x 1 ciclo

30'' 94°

30'' 55° X 10 o 30 ciclos

50'' 72°

5' 72° 1 ciclo

A menos que se indique lo contrario este es el protocolo y programa empleado en la PCR2. En la Figura M2 se muestran los resultados de la PCR2.

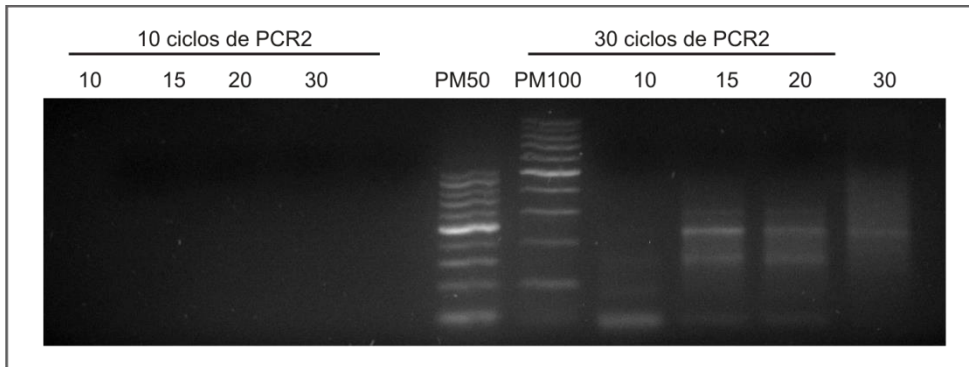


Figura M2. Gel de agarosa al 2% con los productos de la PCR2.

El gel se corrió durante 30 minutos con dos marcadores de peso molecular de 50 y 100pb. Los nombres de las calles indicar el número de ciclos de la PCR1.1 en el cual se tomó la alícuota

Sólo se observó amplificación para 30 ciclos de la PCR2, a partir de los 15 ciclos de la PCR1.1 en adelante. Sin embargo este número de ciclos es inapropiado para trabajar con oligos indexados por lo que se continuó ensayando otras condiciones para la PCR1.

5.6.3.2. PCR1.2

A continuación se decidió trabajar con una concentración final de oligo superior. Se probó entonces con una concentración 10uM en lugar de 1uM.

MM

Oligo FW1: 1 uL (10uM)

Oligo RV1: 1 uL (10uM)

dNTPs: 8uL

buffer: 10uL

ExTaq: 0,04uL

cDNA: 40uL

Agua: 31,6uL

Volumen final: 100uL

Los primeros 5 ciclos se hicieron 10% de oligo y Taq. Luego se agregó el 90% restante de oligo y Taq (4uL de cada oligo y 0,36uL de enzima). A los ciclos 10, 15, 20 y 30 se detuvo la reacción y se tomaron 25ul, para correr luego en un gel de agarosa. La Figura M3 muestra los resultados de la PCR1.2.

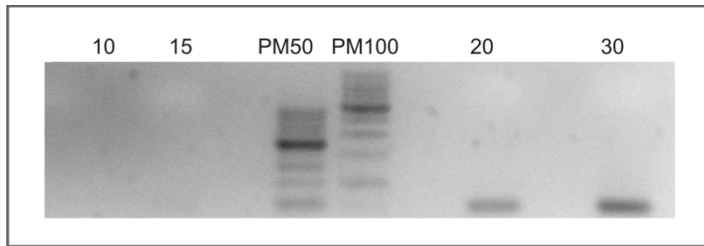


Figura M3. Gel de agarosa al 2% con los productos de la PCR1.2

El gel se corrió durante 30 minutos con dos marcadores de peso molecular de 50 y 100pb. Los nombres de las calles indican el número de ciclos de la PCR1.2 en el cual se tomó la alícuota.

A pesar de no observar amplificación en ninguna calle, se purificaron desde el gel las regiones comprendidas entre 100 y 700pb. A partir de 2ul de PCR1 purificada, se hizo la PCR2 a 10 y 30 ciclos, con el protocolo que se describió en la Sección 5.7.1., para cada una de las condiciones (es decir, ciclos) de la PCR1.2. La Figura M4 muestra los resultados de la PCR2.

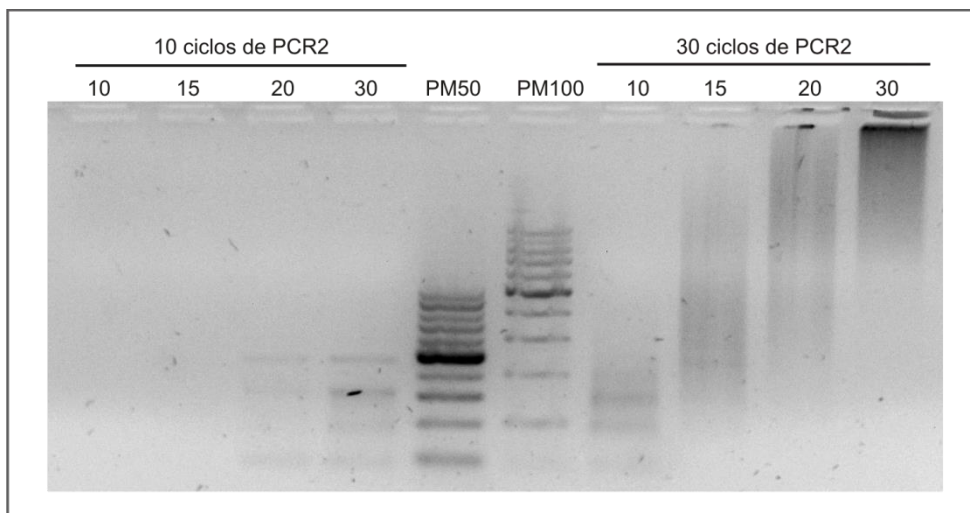


Figura M4. Gel de agarosa al 2% con los productos de la PCR2.

El gel se corrió durante 30 minutos con dos marcadores de peso molecular de 50 y 100pb. Los nombres de las calles indican el número de ciclos de la PCR1.2 en el cual se tomó la alícuota.

En las calles 3 y 4 se observa un patrón de amplificación aceptable. Sin embargo el “manchado” observado en las calles 8-10 (30 ciclos de PCR2) podría ser indicativo de problemas de amplificación del ADN simple cadena. Debido a esto se decidió hacer la PCR1 con temperaturas de hibridación más bajas de 55°C.

5.6.3.3. PCR1.3

Se repitieron las condiciones de la PCR1.2. La única modificación fue en el programa de PCR, donde se emplearon dos temperaturas de hibridación: 45°C y 50°C. Para ambas temperaturas se tomaron 25ul a los 15, 20 y 30 ciclos. La Figura M5 muestra los resultados de la PCR1.3

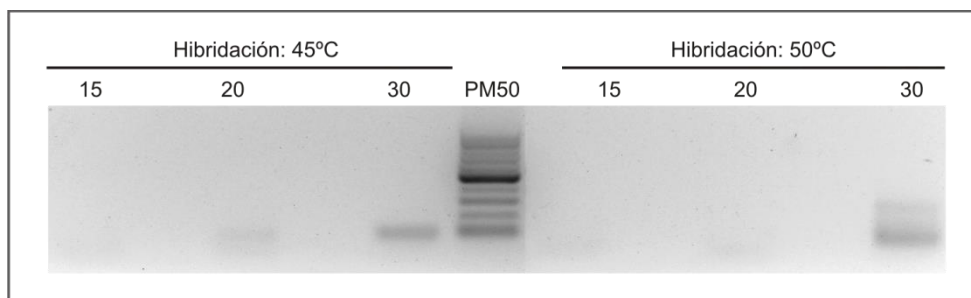


Figura M5. Gel de agarosa al 2% con los productos de la PCR1.3

El gel se corrió durante 30 minutos con el marcador de peso molecular de 50pb. Los nombres de las calles indican el número de ciclos de la PCR1 en el cual se tomó la alícuota.

A partir de 2ul de PCR1.3 purificada, se realizó la PCR2 a 10 y 30 ciclos, en un volumen final de 25 ul en lugar de 50uL. A su vez se realizó la PCR2 de la PCR1.2, con 10ul (A) y 20 ul (B), del ciclo 15 de la PCR1.2 purificada. Esto último se hizo con el objetivo de evaluar el efecto del aumento de molde en el resultado de la PCR2 y seguir trabajando con una temperatura de hibridación de 55°C. La Figura M6 muestra los resultados de las condiciones ensayadas en la PCR2.

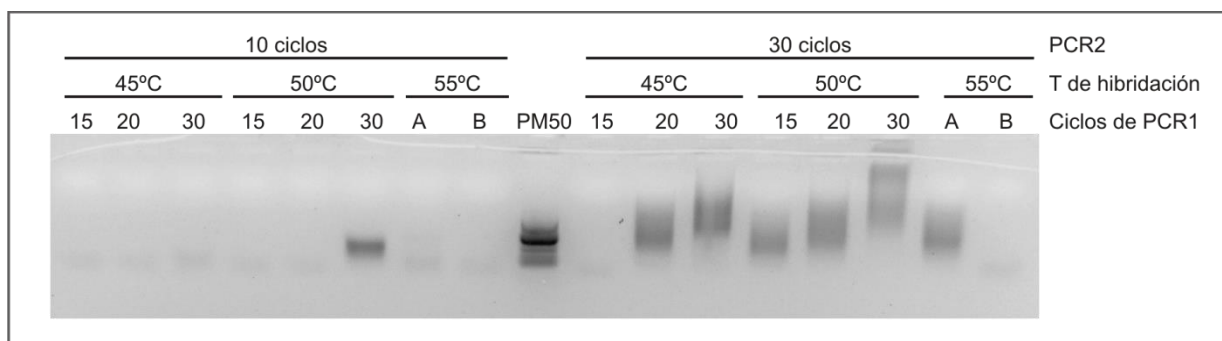


Figura M6. Gel de agarosa al 2% con los productos de la PCR2.

El gel se corrió durante 15 minutos con el marcador de peso molecular de 50pb. Los nombres de las calles indican el número de ciclos de la PCR1.3 en el cual se tomó la alícuota (15, 20 y 30) o la cantidad de producto purificado de la PCR1.2 (A y B).

En todos los casos el manchado a los 30 ciclos de la PCR2, persistió. Debido a ello decidimos hacer una búsqueda bibliográfica para encontrar posibles soluciones a este problema. Las recomendaciones más frecuentes fueron:

- Reducir el número de ciclos de PCR.
- Aumentar la temperatura de hibridación.
- Emplear un perfil térmico de “touchdown” PCR.
- Trabajar en condiciones de “hot start” PCR.
- Para ADNc complejos se recomienda optimizar el buffer de la PCR, utilizar DMSO (desde un 5-10%) y utilizar una polimerasa muy procesiva.

En función de dichas recomendaciones se decidió cambiar la polimerasa (TaKara ExTaq) por la enzima Phusion. Esta polimerasa es procesiva y extremadamente fiel, lo suficiente como para emplearse en la construcción de bibliotecas que requieren de secuenciación masiva (Zhai et al., 2014). A su vez, emplea elevadas temperaturas de hibridación (~63°C) debido a la elevada concentración de sales presentes en los buffers que vienen con la enzima, HF (de alta fidelidad) y GC. Este último es recomendado para templados difíciles, con posibilidad de formar estructura secundaria y es compatible con el uso de DMSO.

5.6.3.4. PCR1.4

Se cambió la polimerasa TaKaRa Ex Taq por la Phusion. En esta primera prueba se siguieron las recomendaciones del manual de la enzima.

MM

Oligo FW1: 2,5uL (10uM)

Oligo RV1: 2,5uL (10uM)

dNTPs: 4uL

buffer: 10uL

DMSO: 1,5uL

Phusion: 0,5 uL

cDNA: 20uL o 8uL

Agua: c.s.p: 50uL

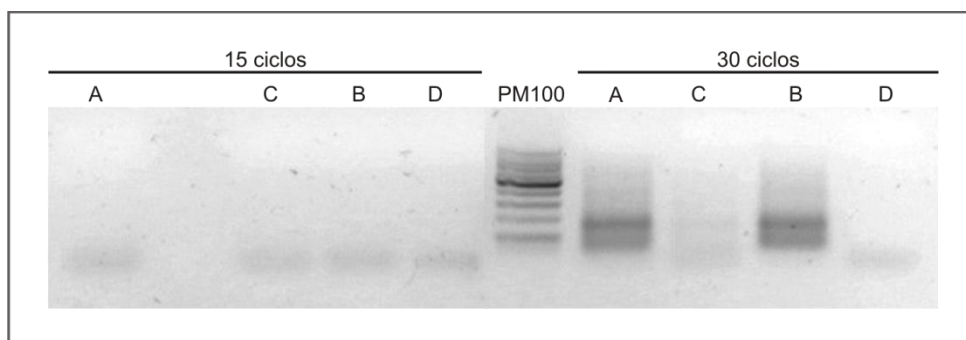
Programa de PCR:

60'' 98°	x1 ciclo
30'' 98°	
30'' 63°	X 15 o 30 ciclos.
50'' 72°	

Las variables a probar fueron la cantidad de inicial de ADNc (20uL o 8uL totales) en 50 ul de reacción y los dos tipos de buffers, high fidelity (HF) y GC (con DMSO). Estas se resumen en la Tabla M.2. Además de esto, la PCR1.4 se hizo a 15 y 30 ciclos. Los resultados de la PCR1.4 se presentan en la Figura M7.

Tabla M2. Condiciones evaluadas con polimerasa Phusion.

	A	B	C	D
cDNA	20uL	20uL	8uL	8uL
Buffer	GC	HF	GC	HF

**Figura M7. Gel de agarosa al 2% con los productos de la PCR1.4.**

El gel se corrió durante 15 minutos con el marcador de peso molecular de 100pb. Los nombres de las calles hacen referencia a las condiciones enumeradas en la Tabla M2.

Se cortaron las regiones comprendidas entre los 100 y 700pb para las reacciones de PCR de 15 ciclos A-D (izquierda del gel). A partir de 2uL de PCR1.4 purificada, se realizó la PCR2. Teniendo en cuenta que se habían obtenido mejores resultados al trabajar con buffer GC y DMSO (Figura M7, calles A y C, derecha del gel, 30 ciclos de PCR1.4) estos se emplearon en la PCR2.

MM

Oligo FW1: 1uL (10uM)

Oligo RV1: 1uL (10uM)

dNTPs: 2uL

buffer GC: 4uL

DMSO: 0,6uL

Phusion: 0,2 uL

molde: 2uL

Agua: c.s.p: 20uL

Programa de PCR:

60'' 98° x1 ciclo

30'' 98°

30'' 63°X 10 ciclos.

50'' 72°

La Figura M8 muestra los resultados de la PCR2.

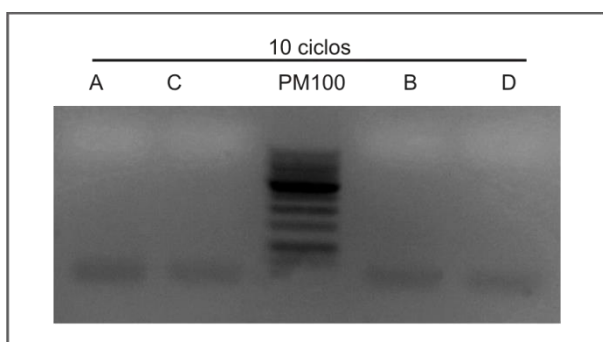


Figura M8. Gel de agarosa al 2% con los productos de la PCR1.

El gel se corrió durante 15 minutos con el marcador de peso molecular de 100pb. Los nombres de las calles hacen referencia a las PCR1 de 15 ciclos purificadas, según las condiciones enumeradas en la Tabla M2

Como no se hubo amplificación en ninguna calle, se decidió repetir la PCR2 cambiando las cantidades iniciales de la PCR1.4 de 15 ciclos purificada (Figura M7).

5.6.3.5. PCR2.1

Se hizo la PCR2 a partir de la PCR1.4 de 15 ciclos, con cantidades variables de molde: 25ul de la reacción A, 8ul y 16ul de la reacción B. A su vez se usaron 2ul de la reacción A a 30 ciclos de la PCR1.4, como control de amplificación. El volumen final de la reacción fue 50ul y se tomaron 25ul a los 10 y 30 ciclos de la PCR2 para constatar amplificación y ausencia de la mancha. En la Figura M9 se muestran los resultados de la PCR2.1.

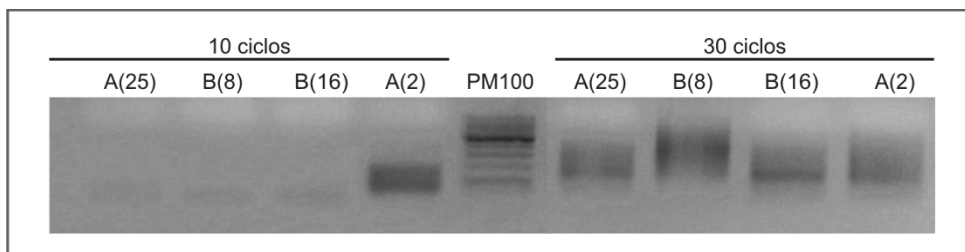


Figura M9. Gel de agarosa al 2% con los productos de la PCR2.1

El gel se corrió durante 15 minutos con el marcador de peso molecular de 100pb. Los nombres de las calles hacen referencia a las PCR1.4 de 15 ciclos purificada, según las condiciones enumeradas en la Tabla R1.2. Los números entre paréntesis son las cantidades de partida de la PCR1.4 purificada (25, 8 y 16) mientras que A (2) corresponde a la PCR1.4 purificada de 30 ciclos.

A los 30 ciclos de la PCR2.1 se detecta amplificación en todas las calles. A la derecha del gel, en las calles A (25) y B (16) se ve una colección de bandas, definidas y acotadas, siendo este el resultado esperado. En función de esto, se seleccionaron las siguientes condiciones para la PCR1: 20uL de ADNc y el buffer GC, condición A.

Sin embargo 30 ciclos de PCR2 son inadecuados para trabajar con oligos indexados, por lo que se continuó poniendo a punto las condiciones de la PCR2, a partir del molde A de la PCR1.4 purificada.

5.6.3.6. PCR2.2

En función de los resultados anteriores, se decidió hacer la PCR2 con cantidad variable de molde de la PCR1.4 purificada (A). Para esto se repitió la PCR1.4, a 15, 20 y 25 ciclos. Se corrió un gel de agarosa al 2% (no se observó amplificación, gel no mostrado) y se cortó la región comprendida entre 100 y 700pb.

Luego, se usaron 2 y 10ul para cada una de las condiciones (es decir, ciclos) de la PCR1.4 como molde de la PCR2.2 El objetivo era trabajar con un número bajo de ciclos en la PCR1, durante la etapa exponencial de amplificación. La PCR2.2 se hizo a 10 ciclos. La Figura M10 muestra los resultados obtenidos.

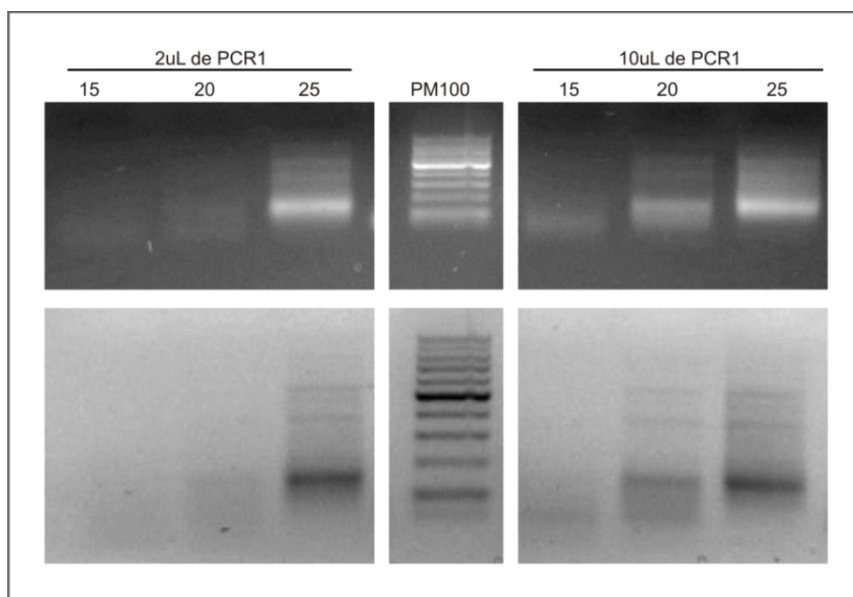


Figura M10. Gel de agarosa al 2% con los productos de la PCR2.2

El gel que se muestra en la parte superior, se corrió durante 15 minutos, y el que se muestra en la porción inferior, durante media hora. Se usó un marcador de peso molecular de 100pb. Los nombres de las calles hacen referencia al número de ciclos de la PCR1.4 de 15 ciclos purificada.

La calle que mejor amplificó fue la 6, donde se usaron 20 ciclos de PCR1, 10 ciclos de PCR2 y 10uL de PCR1 purificada como molde. Estas condiciones se seleccionaron para la construcción de las bibliotecas.

5.6.4. Construcción de marcadores de peso molecular específicos.

Estos marcadores de peso molecular se hicieron con el objetivo de delimitar la zona de corte en los geles, especialmente para la PCR1 donde se trabaja a 20 ciclos y no se detectan las bandas en el gel (Figura M7). Para su construcción se emplearon plásmidos vacíos y oligos genéricos que hibridaran en por fuera del sitio de múltiple clonados de los mismos, disponibles en el laboratorio.

Para la PCR1:

- Límite inferior: 90pb. Se construyó un marcador de 92pb, Figura M11, calle 1.
- Banda mayoritaria esperada: 120pb. Se construyó un marcador de 117pb, Figura M11, calle 2.

Para la PCR2:

- Límite inferior: 150pb. Se construyó un marcador de 144pb, Figura M11, calle 4.
- Banda mayoritaria esperada: 220pb. Se construyó un marcador de 224pb, Figura M11, calle 5.

Los marcados mencionados anteriormente fueron preparados mediante PCR siguiendo el siguiente protocolo:

MM

5X buffer GoTaq (transparente)	10 μ L
dNTPs	2 μ L
oligo FW	2 μ L
oligo RV	2 μ L
GoTaq	0.2 μ L
Molde (plásmido vacío 1/10)	1 μ L
H2O	31,9 μ L

Volumen final: 50 μ L

Programa:

95 °C 60 seg

95°C 30 seg

55°C 30 seg x 35 ciclos

72°C 40 seg

Los resultados se muestran en la Figura M11.

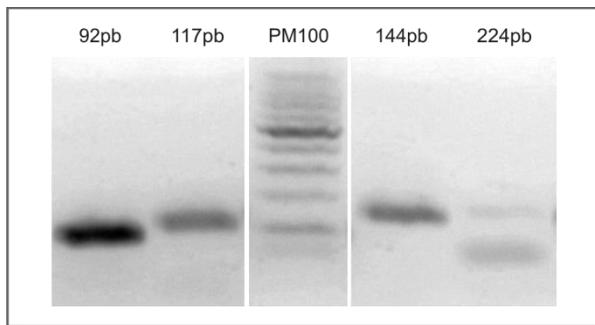


Figura M11. Gel de agarosa al 1,5% con los productos de la PCR.

El gel se corrió durante 15 minutos, con un marcador de peso molecular de 100pb. Los nombres de las calles hacen referencia al tamaño esperado del producto de PCR.

En todas las calles se obtuvo una banda del tamaño esperado a excepción de la calle 5. Por lo tanto se decidió reemplazar el marcador de 224pb por un marcador de 240pb que estaba disponible previamente en el laboratorio.

5.6.5. Construcción de bibliotecas con las condiciones finales de trabajo.

Se tomaron 5ul de cada mix individual de los ADNc, para dar 40ul totales, de los cuales se usaron 20ul como molde de la PCR1 y el remanente se reservó en caso de que hubiera que repetir la primer PCR. Este protocolo se aplicó a las muestras provenientes de Col-0 (cuadruplicados biológicos), *se* (duplicados), *hyl1* (duplicados) y *fiery1* (duplicados), se construyeron un total de 10 bibliotecas.

5.6.5.1. PCR1.

MM:

5x buffer GC: 10uL

dNTPs: 4uL

Oligo FW1: 1,5uL

Oligo RV1: 1,5 uL

Molde: 20ul de ADNc

DMSO: 1,5 ul

Phusion: 0,5 ul

H₂O: 11ul

Volumen final: 50ul

Programa de PCR:

98C 1 min x 1 ciclo

98C 30s

63C 30s x 20 ciclos

72C 50s

72C 10 min

Se corrió un gel de agarosa al 1,5% (no mostrado). Se cortaron las regiones comprendidas entre 100 y 700pb y se purificaron en un volumen final de 35uL.

5.6.5.2. PCR2.

En cada caso se usaron 10uL de la PCR1 purificada.

MM

5X GC buffer: 10ul

dNTPs: 4ul

Oligo FW2: 0.75uL

Oligo Indexado (RV2): 0.75uL

Molde: 10ul PCR1 purificada.

DMSO: 1,5ul

Phusion: 0,5ul

H₂O: 22,5ul

Volumen final: 50ul

Los oligos indexados empleados en la PCR2, se resumen en la Tabla M3

Programa de PCR

98C 1 min x 1 ciclo

98C 30s

63C 30s x 10 ciclos

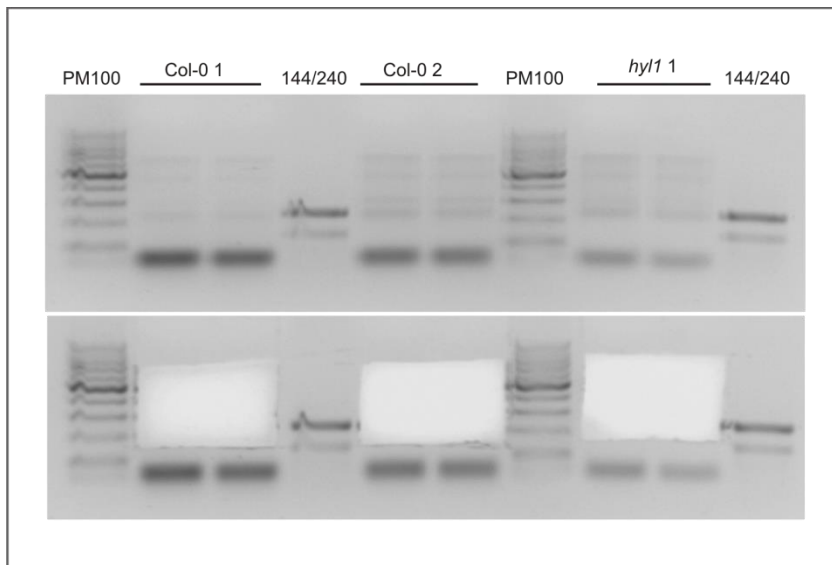
72C 50s

72C 10 min

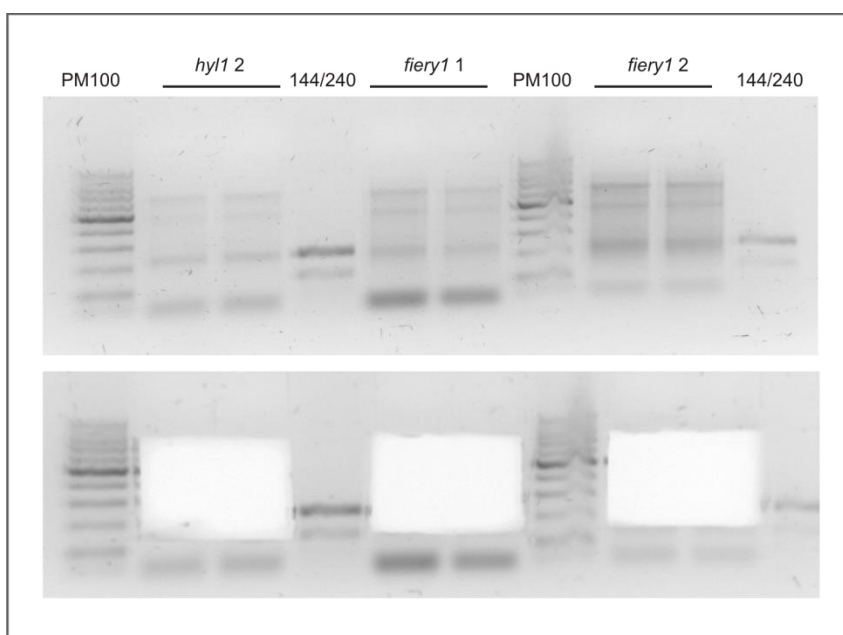
Tabla M3. Número de índex empleado en cada biblioteca generada con el nuevo protocolo de SPARE.

Col-0 1	Col-0 2	Col-0 3	Col-0 4	<i>fiery1 1</i>	<i>fiery1 2</i>	<i>hyl1 1</i>	<i>hyl1 2</i>	<i>se 1</i>	<i>se 2</i>
5	6	11	12	9	10	7	8	3	4

Los productos de la PCR2 se corrieron en geles de agarosa al 2%, dividiéndose el volumen total de la reacción en dos pocillos. Los resultados se resumen en las Figuras M12-14.

**Figura M12. Gel de agarosa al 2% con los productos de la PCR2.**

El gel se corrió durante 30 minutos, con un marcador de peso molecular de 100pb y los marcadores específicos de 144 y 240pb. Los nombres de las calles hacen referencia a las bibliotecas amplificadas durante la PCR2. En el gel inferior se muestra la región cortada entre ~150pb y 800pb.

**Figura M13. Gel de agarosa al 2% con los productos de la PCR.**

El gel se corrió durante 30 minutos, con un marcador de peso molecular de 100pb y los marcadores específicos de 144 y 240pb. Los nombres de las calles hacen referencia a las bibliotecas amplificadas durante la PCR2. En el gel inferior se muestra la región cortada entre ~150pb y 800pb.

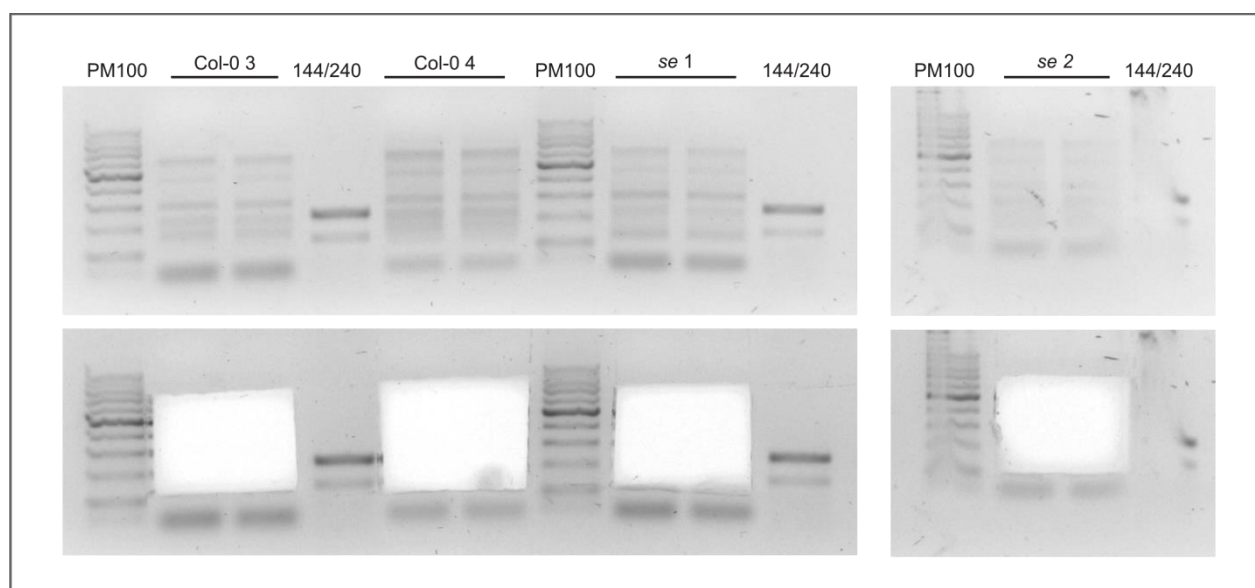


Figura M14. Gel de agarosa al 2% con los productos de la PCR.

El gel se corrió durante 30 minutos, con un marcador de peso molecular de 100pb y los marcadores específicos de 144 y 240pb. Los nombres de las calles hacen referencia a las bibliotecas amplificadas durante la PCR2. En el gel inferior se muestra la región cortada entre ~150pb y 800pb.

En todos los casos se obtuvieron productos de amplificación, siendo las mismas bandas discretas de diferentes tamaños, lo cual coincide con lo esperado. La región cortada desde 150-800pb se purificó como se indica en la Sección 5.6.2.8. y se eluyeron en un volumen de 22uL.

5.6.5.3. Cuantificación de las bibliotecas generadas

Empleando el equipo NanoDrop se cuantificaron 2uL de las bibliotecas, midiendo absorbancia a 260nm. Los resultados se resumen en la Tabla M4.

Tabla M4. Concentraciones de las bibliotecas generadas.

Col-0 1	Col-0 2	Col-0 3	Col-0 4	<i>fiery 1 1</i>	<i>fiery1 2</i>	<i>hyl1 1</i>	<i>hyl1 2</i>	<i>se 1</i>	<i>se 2</i>
5,9 ng/uL	5,4 ng/uL	7,9 ng/uL	12,9 ng/uL	7,7 ng/uL	11ng/uL	5,5ng/uL	5,8 ng/uL	8,6ng/uL	5,4ng/uL

Finalmente se corrió un gel con la PCR2 purificada para corroborar la presencia de ADN post-purificación. Se corrieron aproximadamente 10uL de cada biblioteca purificada, quedando 10uL para enviar a secuenciar. Los resultados se muestran en las Figuras M15-16.

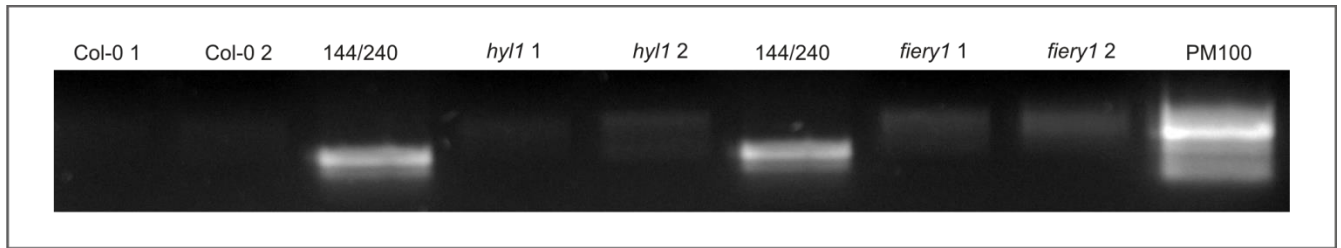


Figura M15. Gel de agarosa al 1,5 % con los productos purificados de la PCR2.

El gel se corrió durante 15 minutos, con un marcador de peso molecular de 100pb y los marcadores específicos de 144 y 240pb. Los nombres de las calles hacen referencia a las bibliotecas purificadas.

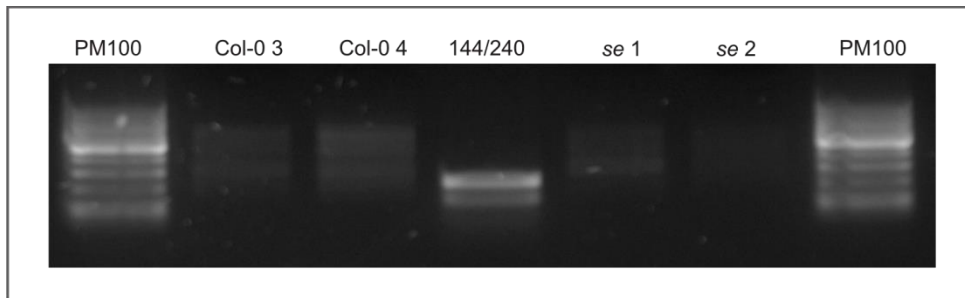


Figura M16. Gel de agarosa al 1,5 % con los productos purificados de la PCR2.

El gel se corrió durante 15 minutos, con un marcador de peso molecular de 100pb y los marcadores específicos de 144 y 240pb. Los nombres de las calles hacen referencia a las bibliotecas purificadas.

5.6.5.4. Precipitación de las bibliotecas generadas.

Luego de constatar la presencia de ADN luego de la purificación, las bibliotecas se precipitaron. Para esto, 10ul de cada biblioteca se llevaron a 100ul con agua, y a continuación se agregaron 220uL de etanol absoluto, 10uL de NaAc 3M y 2uL de Glicoblue.

Las muestras se dejaron precipitando durante la noche a -20°C en tubos eppendorf de 2mL de panza redondeada. Luego se centrifugaron a máxima velocidad durante media hora en la centrifuga refrigerada, se descartó el sobrenadante y se agregó 1mL de etanol absoluto. Se envolvieron y se enviaron a secuenciar.

5.7. Procesamiento de los datos crudos.

5.7.1. Archivo final con los resultados de las bibliotecas.

Una vez obtenidos los datos crudos de la secuenciación, se realizó el procesamiento de los datos (en el contexto de la tesis doctoral del Dr. Uciel Chorostecki). Para la obtención de un archivo final de Excel se siguieron los siguientes pasos:

1. Remover secuencias adaptadoras (trimming)
2. Agrupar secuencias iguales y determinar abundancia (conversión de fast-q file a tag-count file).
3. Mapear las secuencias únicas contra las secuencias de precursores (sub-genoma definido previamente).

4. Combinar los resultados del mapeo con la abundancia de cada secuencia.
5. Filtrado de secuencias menores a 30pb. Este paso es importante ya que los ARNs pequeños maduros son muy contaminantes. Esto se visto para bibliotecas del tipo degradoma o PARE del inglés, Paralell Anlisys of RNA Ends(Addo-Quaye et al., 2009).
6. Normalización de los datos. El número de lecturas totales varía de biblioteca a biblioteca, es por ello que es necesario normalizar los resultados obtenidos para poder comparar entre las diferentes muestras. La Tabla M5 resumen las lecturas obtenidas en cada biblioteca antes de normalizar.

Tabla M5. Lecturas obtenidas en cada una de las bibliotecas generadas.

Col-0 1	Col-0 2	Col-0 3	Col-0 4	<i>fiery 1 1</i>	<i>fiery1 2</i>	<i>hyl1 1</i>	<i>hyl1 2</i>	<i>se 1</i>	<i>se 2</i>
56.280	121.253	39.262	81.407	45.2434	79.4067	45.103	25.677	38.554	74.277

Teniendo en cuenta que la biblioteca Col-0 2 fue la que tuvo el mayor número de lecturas se empleó 120K para la normalización. Para normalizar, cada lectura obtenida en cada biblioteca se realizó la siguiente conversión:

Normalización: N° de lecturas* 120.000/ N° lecturas totales.

5.7.2. Representación gráfica de los cortes en cada precursor.

En todos los casos, la secuencia de interés que corresponde al precursor de microARN es la representada como: 5'-N1NNNNNNNNNNNNNNNNNNNNNNNN-3' (Figura M17A). El corte en el precursor se encontrará entre los nucleótidos N1 y "N0", estando este último ausente en la secuencia obtenida y debe recuperarse ubicando esta secuencia en el contexto de la estructura secundaria del precursor (Figura M17B). Cada posición de corte se indicará con una flecha de color verde, cuando el corte provenga de una biblioteca de Col-0 y roja cuando el corte provenga de una biblioteca de *fiery1*.

Cada flecha, estará acompañada de la abundancia normalizada de dicha posición de corte. En el caso de las muestras provenientes de plantas salvajes, para graficar el corte en una determinada posición, el mismo debe estar presente en al menos dos de las cuatro bibliotecas. Mientras que en el en caso de *fiery1* los cortes deben estar presentes en todas las bibliotecas (dos) para considerarlo como un corte positivo. La sumatoria de todos los cortes provenientes de cada biblioteca, permiten reconstruir el patrón de procesamiento de cada precursor y a partir de este, inferir el mecanismo de procesamiento del mismo (Figura M17CD y Tabla M6).

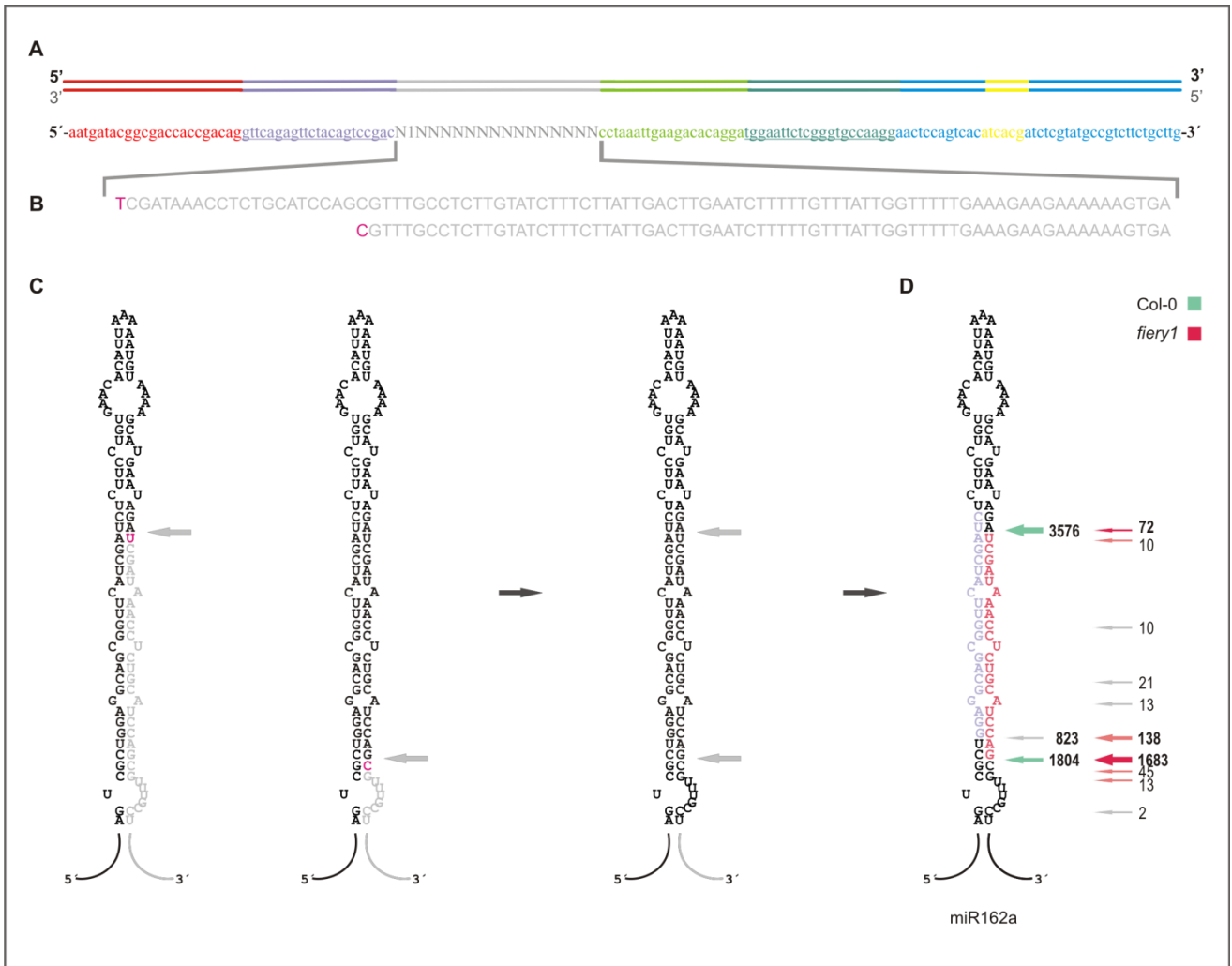


Figura M17. Reconstrucción del patrón de procesamiento a partir de las lecturas de fragmentos individuales.

A) Esquema y detalle de los fragmentos finales obtenidos en bibliotecas de SPARE multiplex.

B) Detalle de dos secuencias correspondientes a dos intermediarios de procesamiento del precursor de miR162a.

C) Ubicación de los cortes asociados a las secuencias en el contexto de la estructura secundaria del precursor de miR162a.

D) Patrón final de procesamiento del precursor de miR162a reconstruido para las bibliotecas de Col-0 y *fiery1*

Debido a la posición relativa de los oligos en el precursor, el método permite detectar un número variable de cortes dependiendo del mecanismo de procesamiento responsable de la biogénesis de cada microARN. La información resumida en la Tabla M6 permite asociar el patrón de cortes de un precursor con uno de los cuatro mecanismos de procesamiento descritos previamente (Bologna et al., 2013b). De este modo se determinación del mecanismo de procesamiento para cada precursor.

Tabla M6. Número de intermediarios de SPARE esperados para cada de mecanismo de procesamiento.

Mecanismo de procesamiento	Número de Intermediarios	Ejemplo
Tallo a burbuja corto	1	miR171a
Tallo a burbuja secuencial	1	miR169b
Burbuja a tallo corto	2	miR156a
Burbuja a tallo secuencial	4	miR159b

5.8. Transcripción *in vitro*

5.8.1. Preparación de muestras de ARN

5.8.1.1. Moldes para transcripción *in vitro*.

Para generar el molde para la transcripción *in vitro* del precursor de miR157c se partió de un plásmido que contenía un clon genómico de dicho precursor. Con el objetivo de incorporar la secuencia del promotor T7 se realizó una primer PCR donde uno de los oligos presentaba la secuencia del promotor T7 sobre el extremo 5'. El programa de PCR empleado se detalla en la Sección 5.2.2. El producto esperado era de 195pb.

Para evidenciar los productos de la PCR se corrió un gel de agarosa al 2%. La banda deseada se cortó y se purificó como se detalla en la Sección 5.1.6. A continuación se realizó una nueva PCR sobre este inserto purificado con el objetivo de aumentar la masa final de molde para la transcripción. Esta PCR se hizo en las mismas condiciones que la primer PCR. Luego se corrió un gel de agarosa al 2% donde se sembraron 5µL del volumen de reacción para constatar amplificación. El remanente de del volumen de PCR se purificó empleando una extracción con una mezcla de fenol-cloroformo-isoamílico que se detalla a continuación.

1. Llevar el volumen de reacción a 100µL finales
2. Agregar 100µL de la meza de 25 partes de fenol, 24 partes de cloroformo y 1 parte de alcohol isoamílico
3. Mezclar usando un vortex
4. Centrifugar 2min. a máxima velocidad
5. Descartar la fase orgánica
6. Agregar 100µL de cloroformo
7. Mezclar usando un vortex
8. Centrifugar 2min. a máxima velocidad
9. Descartar la fase orgánica
10. Agregar 10µL de NaCl 3M y 300µL de etanol absoluto
11. Dejar precipitando durante 20 min. a -80°C
12. Centrifugar 30min. a máxima velocidad
13. Hacer un lavado con etanol al 70% v/v
14. Centrifugar 5min. a máxima velocidad
15. Secar en una estufa a 37°C durante 7min.
16. Resuspender en 20µL de agua
17. Cuantificar el molde

5.8.1.2. Transcripción *in vitro*.

Para realizar la transcripción *in vitro* se realiza una mezcla inicial que contiene: 10µL del ADN molde, 4µL del buffer comercial de la ARN polimerasa y 6µL de agua. Esta mezcla se calienta a 95°C durante 2min. y luego se deja 15min. a temperatura ambiente.

La mezcla de reacción para transcripción lleva 16µL del buffer comercial, 20µL de la mezcla de NTPs 10mM cada uno, 2,5µL de ARNasa OUT, 3µL de la T7 ARN polimerasa (Thermo) y 38,5 ul de agua. Una vez transcurridos los 15 minutos, se agrega la mezcla de reacción a la mezcla inicial. La transcripción se realiza a 37°C durante 5hs. A continuación los productos de la reacción se mezclaron con 1/5 volúmenes de solución de siembra y se corrieron en un gel de poliacrilamida al 6% en condiciones desnaturizantes (Se usaron las condiciones detalladas en la Sección 5.4.4 modificando la cantidad de 19% Acrylamida/1% Bis-acrylamida para llegar a una concentración del 6%). La corrida electroforética del mismo se realizó a 180 V y durante una hora. Una vez finalizada la corrida, se tiñó el gel con 1X TBE y Bromuro de Etidio por 5 min, y se escindió la banda del tamaño esperado de 175nt.

5.8.1.3. Purificación de ARN a partir de geles de poliacrilamida

Para realizar la purificación del ARN a partir de la banda de poliacrilamida, se colocó la misma en un tubo eppendorf y se centrifugó el mismo a 13000rpm durante 2min. Luego de agregar 300µL de buffer TEN (cuya composición se detalla más abajo) y se mezcla empleando un vortex. Se incuban los fragmentos de acrilamida con el buffer de extracción a 4°C durante toda la noche.

Al día siguiente se centrifuga a 4°C a 13000rpm durante 2min y se toman ~300µL de la fase acuosa y se trasvasa a otro tubo eppendorf el cual se guarda a 4°C. Mientras tanto se agregan 300µL más en el tubo eppendorf que contiene los restos de acrilamida, se mezcla empleando un vortex y se incuban 3hs a temperatura ambiente. Una vez transcurridas las 3hs, se centrifuga a 4°C a 13000rpm durante 2min y se toman ~300µL de la fase acuosa y se trasvasa al tubo eppendorf que se había guardado previamente. El remanente de poliacrilamida se descarta.

A los 600µL de buffer TEN de extracción se agregan 900µL de etanol absoluto y se deja precipitando 1h. a -80°C. Luego se centrifuga a 4°C a 13000rpm durante 1 hora. Se descarta el sobrenadante y el sedimento se lava con 1mL de etanol al 75%v/v. Se centrifuga a 4°C a 13000 rpm durante 5min. Se descarta el sobrenadante y el sedimento se seca a 37°C durante 7 minutos. Finalmente se resuspende en 20µL de agua y se cuantifica.

Para la preparación del buffer de extracción TEN se mezclan 5mL de Tris 1M pH=8, 1mL de EDTA 0,5M pH=8, 50mL de NaCl 3M y se lleva a volumen final de 500mL.

5.9. Mapeos estructurales sobre el ARN

5.9.1. Marcado de muestras de ARN con ^{32}P en el extremo 5'

Para realizar los ensayos de mapeo estructural es necesario marcar las muestras de ARN en el extremo 5' con fosforo radioactivo. Para esto tomamos 300pmoles de ARN y realizamos una desfosforilación empleando la enzima Fosfatasa Antartika (NEB) según las instrucciones del fabricante. Luego realizamos una extracción fenol/cloroformo/alcohol isoamílico como se describe en la Sección 5.8.1.1 y resuspendimos en 15 μl de agua.

Tomamos 5 μl de esta muestra y procedimos a realizar la fosforilación del extremo 5' empleando la enzima Polinucleótido Quinasa (Thermo Scientific) de acuerdo a las instrucciones del fabricante y utilizando como sustrato ^{32}P - γ -ATP. Finalmente purificamos estas muestras por gel de poliacrilamida desnaturalizante al 8% como se describe en la Sección 5.4.4 y la extracción del ARN marcado se realizó según la Sección 5.8.1.3. Resuspendimos las muestras en agua estimando una concentración final de 1 μM de acuerdo al porcentaje de precursor marcado de largo completo visualizado en el gel.

5.9.2. Mapeo por digestión enzimática

Antes de realizar el tratamiento con las enzimas específicas se calienta el precursor marcado a 95°C durante 5 min. y se enfría inmediatamente en un baño de hielo/agua durante 5min. para favorecer el plegamiento por hibridación intramolecular. Esto es lo que se conoce como ARN activado y se usa en una concentración de 1pmol/ μL .

Luego se preparan 6 reacciones. Las primeras dos reacciones se emplean como referencias. Para la reacción de hidrolisis alcalina (OH^-) se mezclan 1,5 μL de ARN activado y 3 μL de NaHCO_3 200mM pH=7. Se calienta a 95°C durante 3min. y se le agrega 5 μL del buffer de siembra. Para la reacción con la enzima T1 en condiciones desnaturalizantes se mezclan 1,5 μL de ARN activado y 7,5 μL de buffer CEU (cuya composición se detalla más adelante). Se calienta a 50°C durante 5min. y se le agrega 1 μL de T1 en una concentración de 1U/ μL . Se deja cursar la reacción durante 15min. a 50°C y luego se realiza una precipitación durante de la noche como se detalla en la Sección 5.9.2.1. Para la reacción con T1 (Fermentas) en condiciones nativas se mezcla 1,5 μL de ARN activado, 1 μL de buffer de plegado (cuya composición se detalla más abajo), 0,3 μL de ARN de transferencia de *E.coli* 0,1mg/mL (Sigma) y 6,2 μL de agua. Esto se incuba a temperatura ambiente durante 15min. y luego se agrega 1 μL de T1 a una concentración de 0,01U/ μL . La reacción se deja cursar durante 15min. a temperatura ambiente y luego se precipita durante la noche.

Para la reacción con la enzima S1 (Fermentas) se mezclan 1,5 μL de ARN activado, 1 μL de buffer de plegado, 0,3 μL de ARN de transferencia de *E.coli* 0,1mg/mL y 5,2 μL de agua. La mezcla se incuba 15 min. a temperatura ambiente. Luego se agrega 1 μL de S1 a una concentración de 1U/ μL y 1 μL de ZnCl_2 y la

reacción ocurre durante 15 minutos a temperatura ambiente y se deja precipitando durante la noche. Para la reacción con la enzima V1 se mezclan 1,5µL de ARN activado, 1µL de buffer de plegado, 0,3µL de ARN de transferencia de *E.coli* 0,1mg/mL y 6,2 µL de agua. Se incuba 15 minutos a temperatura ambiente, y luego se agrega 1µL de la enzima V1 (Ambion) a una concentración de 0,001 U/µL. La reacción ocurre durante 15 minutos y se deja precipitando durante la noche. La última reacción sirve como control negativo y consiste en mezclar 1,5µL de ARN activado y 8,5µL de agua e incubar durante media hora a temperatura ambiente antes de poner a precipitar durante la noche.

Composición del buffer CEU: citrato 20 mM pH 5, EDTA 1 mM y urea 7 M. Composición del buffer de plegado: Tris 20 mM pH 7,5, KCl 25 mM y MgCl₂ 10 mM.

5.9.2.1. Precipitación de las muestras

Las reacciones se llevan a un volumen de 100µL con agua y a cada tubo se le agrega 10µL de NaCl 3M, 300µL de etanol absoluto y 1µL de Glycoblué, lo que permite seguir ocularmente el sedimento. Se deja precipitando toda la noche y al día siguiente se centrifuga a máxima velocidad durante 45min a 4°C. Se descarta el sobrenadante y se lavan los sedimentos con 1mL de etanol al 80% v/v. A continuación se centrifuga a máxima velocidad durante 7min. a 4°C, se descarta el sobrenadante y se repite el paso anterior. Luego se seca el remanente en un termobloque a 37°C hasta que no se observe presencia de etanol.

El sedimento se resuspende en 5µL de solución de siembra la cual contiene formamida 80%v/v, Xilene cyanol 0,025%p/v azul de bromo fenol 0,025%p/v, solución reguladora TBE (Tris 100 mM, borato de sodio 100 mM, EDTA 2 mM) y EDTA 50 mM pH 8. Finalmente se sembraron las muestra en geles de secuenciación como se describe en la Sección 5.9.3.

5.9.3. Geles de Secuenciación

Para la preparación del gel, se utilizaron geles de 6 y 8 % acrilamida conteniendo urea 7 M. La polimerización de los geles se realizó en vidrios de 40 cm de largo, con separadores de 0,4 mm y peines de 32 calles. Luego de verter la mezcla de polimerización dentro de los vidrios se dejó incubando durante 30 min. Una vez polimerizados, se colocaron los geles en las cubas de electroforesis vertical junto con la solución reguladora de corrida (TBE 1X). Se hizo una pre-corrida aplicando un potencial de 3000 V durante 30 min previo a sembrar las muestras.

Se prepararon las muestras para sembrar en geles de secuenciación mezclando con solución de siembra desnaturalizante. A continuación se calientan durante 1 min a 95°C, se centrifuga y se siembra 2,5 µl inmediatamente en cada pocillo del gel. Luego de sembrar las muestras los geles se corrieron aplicando el mismo potencial durante 3h. para los geles de 8% de poli(acrilamida) y 4h para los geles de 6 %. Finalizada la electroforesis transferimos los geles a papel de filtro y cubrimos con film plástico para secar en secador

de geles a 80°C durante 40 min. Finalizado el secado colocamos el gel en contacto con placas sensibles a la radioactividad durante toda la noche. Finalmente escaneamos las placas con un escáner óptico.

RESULTADOS I.

Muchacho, que no sabes que es secarse en una timba y armarse, para volverse a meter.

“Muchacho” – Adriana Varela

6. Resultados I.

“Diseño de una herramienta genómica para estudiar el procesamiento de los miARNs de Arabidopsis”

6.1. INTRODUCCION I

Los miARNs se definen por su mecanismo de biogénesis único, el cual involucra el corte preciso a partir de su precursor plegado (Bologna and Voinnet, 2014). En animales, un complejo formado por DROSHA y DGCR8 reconoce el pri-miARN en el núcleo, y produce el primer corte. El pre-miARN resultante es trasladado al citoplasma por las proteínas EXP5 y RAN-GTP. Una vez allí, un complejo formado por DICER y TRBP, realizan el segundo corte, que libera al miARN maduro (Ha and Kim, 2014). Debido a que los intermediarios de procesamiento deben ser exportados al citoplasma, los mismos son relativamente estables y poseen una vida media considerable, lo que en la práctica significa que son fácilmente detectados en blots para ARN pequeños.

En *Arabidopsis thaliana*, el componente principal de la maquinaria de procesamiento, es la ribonucleasa de tipo III DCL1. La misma interacciona con la proteína de unión a ARNdH HYL1 (Han et al., 2004b; Vazquez et al., 2004a) y con una proteína de dedos de Zn del tipo C2H2, SE (Lobbes et al., 2006; Yang et al., 2006). A diferencia de lo que ocurre en animales, DCL1 posee la actividad catalítica para realizar los todos los cortes necesarios para liberar el miARN maduro (Axtell et al., 2011; Bologna and Voinnet, 2014; Rogers and Chen, 2013). Por ende, en plantas todos los cortes se producen en el núcleo y los intermediarios de procesamiento tienen una vida media corta, comparada con lo que sucede en animales.

Dada esta diferencia entre plantas y animales, muchos de los enfoques técnicos usados para estudiar el procesamiento de miARNs en animales no resultan apropiados para plantas. Es por esto que la mayoría de los mecanismos de procesamiento en plantas, han sido descritos empleando mutagénesis dirigida y caracterizados mediante el uso de una versión modificada de la técnica 5'RACE (Bologna et al., 2009; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010). Si bien estos métodos son informativos, consumen una gran cantidad de tiempo, y sólo brindan información acerca de precursores particulares. Debido a esto, en nuestro laboratorio se desarrolló un método basado en secuenciación masiva de alta sensibilidad el cual permite estudiar el procesamiento de todos los precursores de Arabidopsis en simultáneo (Bologna et al., 2013b). Esta técnica se denominó SPARE (del inglés Specific Parallel Analisis of 5'RNA Ends). El objetivo de la misma es dilucidar los mecanismos de procesamiento de los precursores de miARNs de Arabidopsis. El número y posición de cortes detectados para la biogénesis de cada miARN son

utilizados para deducir el mecanismo de procesamiento de cada precursor (Figura R1.1B) (Bologna et al., 2013a).

Brevemente, la técnica consiste en emplear los extremos 5' fosfato, que deja DCL1 cortar los precursores, para realizar una reacción de ligación ARN-ARN entre los intermediarios de procesamiento y un oligo de ARN de secuencia conocida. A continuación se emplean oligos específicos para retro-transcribir los intermediarios de cada precursor, con una secuencia común a todos. En este punto todos los fragmentos generados durante el procesamiento presentan los mismos extremos: en el 5' está presente la secuencia del oligo de ARN y en el 3' la secuencia común incorporada durante la retro-transcripción. Con esta información se diseñan los oligos FW1 y RV1, que se emplean a continuación en una reacción de PCR. Finalmente los fragmentos son secuenciados (Figura R1.1A) (Bologna et al., 2013b; Schapire et al., 2014). Este último paso en la construcción de la biblioteca fue realizado en colaboración con el Laboratorio del Dr. Blake Meyers, el cual se especializa en estudios de secuenciación de alto rendimiento.

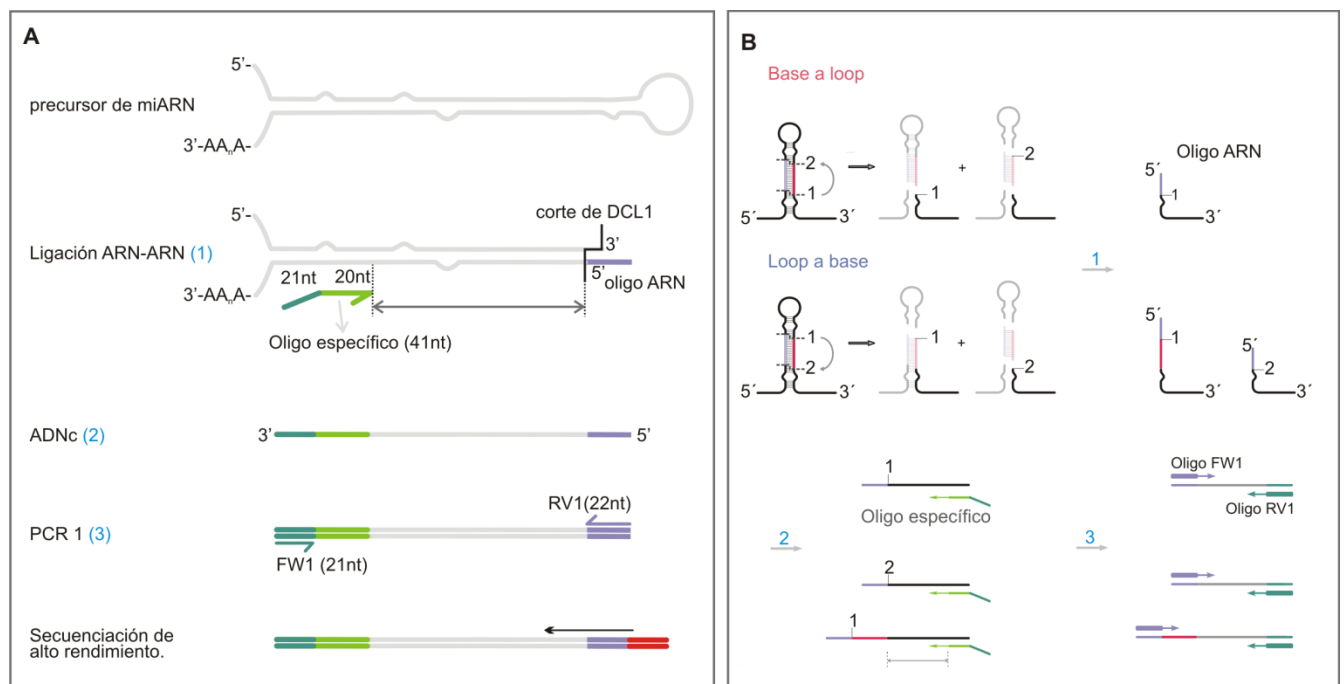


Figura R1.1 Esquema de la construcción de bibliotecas de SPARE.

A) Reacciones empleadas para la construcción de bibliotecas SPARE.

B) Detección diferencial de intermediarios de procesamiento para precursores procesados desde la base al loop o del loop a la base. Para los precursores procesados desde la base se obtiene un único fragmento, a diferencia de los precursores procesados desde el loop donde se obtienen tantos fragmentos como cortes hay en el precursor. Adaptado de (Schapire et al., 2014).

Si bien la técnica de SPARE desarrollada previamente en el laboratorio permitió identificar el mecanismo de procesamiento de varios precursores, consideramos que la misma necesitaba optimizarse. Por cuestiones metodológicas la mayoría de los resultados obtenidos previamente provinieron de plantas mutantes en *FIERY1* en lugar de las plantas silvestres, Col-0. Las mutantes *fiery* tienen reducida la actividad de las ARNasas *XRN*, lo que permite que los intermediarios del metabolismo de ARN se acumulen,

facilitando su detección (Gy et al., 2007). Debido a esto, nos planteamos los siguientes objetivos específicos:

- 1- Optimizar el método de SPARE para poder dilucidar los mecanismos de procesamiento de precursores de miARNs en plantas silvestres.
- 2- Determinar el mecanismo de procesamiento de precursores de miARNs que aún no se hayan descrito.
- 3- Utilizar el método de SPARE para evaluar si los precursores de miARNs sufren cortes no-productivos *in vivo*, además de aquellos producidos por DCL1 para liberar el miARN maduro.
- 4- Evaluar los posibles mecanismos que generen estos cortes adicionales.

6.2. Modificaciones realizadas para la construcción de bibliotecas multiplex en plantas silvestres.

El principal interés en modificar la técnica de SPARE radicaba en aumentar la sensibilidad, poder trabajar con plantas Col-0 y detectar cortes adicionales. Para ello se hicieron los siguientes cambios:

- Optimización general del protocolo:

Se aumentó la cantidad de ARN total de partida de 10µg a 40µg y luego de varias pruebas se decidió cambiar la ADN polimerasa utilizada en las PCRs, eligiéndose la enzima Phusion® High-Fidelity DNA Polymerase (NEB), entre otros cambios realizados. En la Sección 5.6-5.7 de Materiales y Métodos, puede verse una discusión general sobre los esfuerzos para optimizar la técnica, así como el protocolo final obtenido.

- Diseño de los oligos para la síntesis de ADNc.

En la construcción de la biblioteca original se emplearon oligos que hibridaban en el brazo 3' del precursor, 30nt corriente abajo del extremo inferior del miARN o miARN*. Con el objetivo de detectar cortes que se produjeran aún más corriente abajo, se decidió aumentar esta distancia. Se diseñaron un total de 111 oligos a ~100nt, para retro-transcribir todos los precursores conservados y el conjunto de los jóvenes que hubieran sido validados hasta el momento (Bologna et al., 2013b)(Anexo Tabla A1).

- Construcción de bibliotecas multiplex:

El paso final de la técnica consiste en secuenciar las bibliotecas generadas (Figura R1.1A). Cuando se desea comparar entre replicas biológicas o diferentes condiciones, es necesario construir más de una biblioteca. Para poder secuenciar todas las bibliotecas en una misma calle, los fragmentos provenientes de cada biblioteca se marcan con secuencias específicas de 6nt denominadas índice (Sección 5.6.2. Materiales y métodos, Tabla M1). Metodológicamente, esto se hace agregando una segunda reacción de PCR con los

las cuatro bibliotecas para ser tomado en cuenta. Mientras que en el caso de *fiery1* los cortes deben estar presentes en las dos bibliotecas para considerarlo como un corte positivo. La sumatoria de todos los cortes provenientes de cada biblioteca, permiten reconstruir el patrón de procesamiento de cada precursor e inferir el mecanismo de procesamiento del mismo.

Se analizaron muestras de plántulas de *Arabidopsis* de diez días. En este estadio particular y teniendo en cuenta los criterios antes mencionados, en las bibliotecas de Col-O se detectaron en total 71 precursores, de los cuales se pudo establecer la dirección de procesamiento de 45. A su vez hubo 40 precursores de los cuales no se recuperaron intermediarios de procesamiento, por lo que se los calificó como no detectados (Figura R1.3, panel izquierdo). En el caso de las bibliotecas provenientes de *fiery1* se detectaron 86 precursores, de los cuales se pudo establecer la dirección de procesamiento de 71 y 25 no fueron detectados (Figura R1.3, panel derecho).

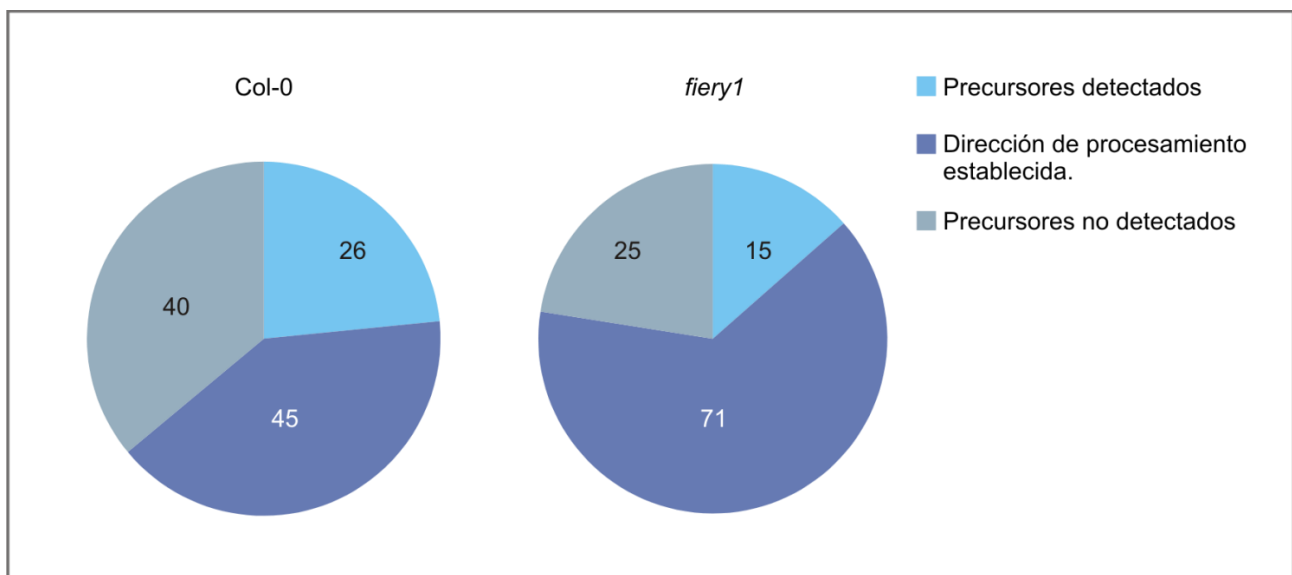


Figura R1.3. Análisis global del procesamiento de precursores de miARN obtenidos de plántulas de *Arabidopsis*.

Los precursores detectados pero sin poder confirmar la dirección de procesamiento en forma experimental se muestran en celeste, mientras que los precursores detectados y con dirección de procesamiento establecida se indican en azul. Finalmente, los precursores sin detectar en plántulas de *Arabidopsis* se muestran en gris para las bibliotecas de Col-O y *fiery1*.

La diferencia entre el número los precursores detectados en Col-O y *fiery1*, puede deberse a que estos precursores se expresen en bajos niveles o sus intermediarios de procesamiento sean particularmente susceptibles a la degradación por ARNasas, las cuales tienen su actividad reducida en plantas mutantes en *FIERY1* (Gy et al., 2007). Es posible que aquellos precursores que no fueron detectados en ningún caso, se expresen bajo condiciones de crecimiento particulares o en otro tipo de tejidos. Este es el caso del precursor de miR172a, el cual se expresa en plantas adultas y tiene su máximo de expresión durante la floración (Wang, 2014). En concordancia con su función biológica y patrón de expresión, es razonable no haber detectado intermediarios de procesamiento provenientes de este precursor.

6.4. Identificación de la dirección de procesamiento de los precursores de miARNs.

Con los datos obtenidos de las bibliotecas de Col-0 y *fiery1* se caracterizó el mecanismo de procesamiento de 45 y 75 precursores respectivamente (ver ANEXO A1-A43 con información detallada de estos resultados). Entre estos precursores encontramos 16 precursores que no habían sido identificados previamente y decidimos analizarlos en más detalle.

6.4.1. Caracterización de precursores procesados desde la base con dos cortes de DCL1

Los precursores procesados desde la base tienen un único corte detectado en las bibliotecas de SPARE, el cual coincide con el extremo inferior del miARN o miARN* (Figura R1.4A). En la Figura R1.4B se muestran los resultados obtenidos para los precursores *demir173*, *miR395f*, *miR397b* y *miR399a*. Cada corte está representado por una flecha y el número de lecturas que se recuperaron para el mismo, los cuales se relacionan con la abundancia del intermediario detectado. El corte responsable de la biogénesis del miARN se indica en color verde oscuro cuando proviene de las bibliotecas de Col-0 y en rojo oscuro cuando proviene de las bibliotecas de *fiery1*.

Los precursores de *miR167c*, *miR396a* y *miR399d* también forman parte del grupo de nuevos precursores procesados en dirección de base a loop (ANEXO A13, A30 y A34). Todos estos precursores tienen un segmento de ARNdH de 15-17pb por debajo del miARN y justo por encima de un loop interno, indicado como un recuadro verde en la Figura R1.4AB. Este segmento sería reconocido por la maquinaria de procesamiento de miARNs para posicionar el sitio del primer corte (Bologna et al., 2013b; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Zhu et al., 2013).

MIR395f, *miR397b* y *miR399a* son miembros de familias más grandes, compuestas por otros genes que codifican para el mismo miARN maduro. El resto de los miembros descritos para cada una de las familias, comparten la misma dirección de procesamiento. El *MIR173* es un *MIARN* joven presente en *Arabidopsis thaliana* y *Arabidopsis lyrata* (Taylor et al., 2014). Este desencadena la producción de tasiARNs a partir del corte de sus genes blanco, *TAS1* y *TAS2* (Vazquez et al., 2004b). Mientras que *MIR395*, *MIR397* y *MIR399* están presentes en las angiospermas en general (Cuperus et al., 2011).

Además de los cortes de biogénesis discutidos hasta ahora, se encontraron otros dos tipos de cortes. El primero, fue denominado corte impreciso, y es representado con una flecha verde claro en Col-0 y una flecha rosa en *fiery1*. El corte impreciso es cercano al corte de biogénesis y libera un ARN pequeño que difiere en el miARN maduro en uno o dos nucleótidos. Esto tiene una verificación experimental, ya que las bibliotecas de ARN pequeños de plantas silvestres también registran una heterogeneidad similar en las secuencias de estos pequeños ARN. Es decir, que consideramos a cualquier corte que caiga en un rango +/- dos nucleótidos del corte de biogénesis como un corte impreciso (Liu et al., 2012). Lo más probable es que

estos cortes se generen como resultado de cierta heterogeneidad durante el reconocimiento y corte de DCL1 en el procesamiento (Jeong and Green, 2012; Lu et al., 2010; Nakano et al., 2006).

También encontramos un tercer tipo de corte que denominamos espurio. Los cortes espurios son eventos de corte cuyo origen es, en principio, desconocido, y que no se corresponde con la liberación del miARN con actividad biológica (indicados en color gris en la figura R1 4B). Más aun, en general estos cortes no se correlacionan con la generación de un ARN pequeño, lo cual no sería sorprendente porque los precursores de miARNs también podrían ser sujetos a la acción de otras rutas involucradas en el metabolismo del ARN.

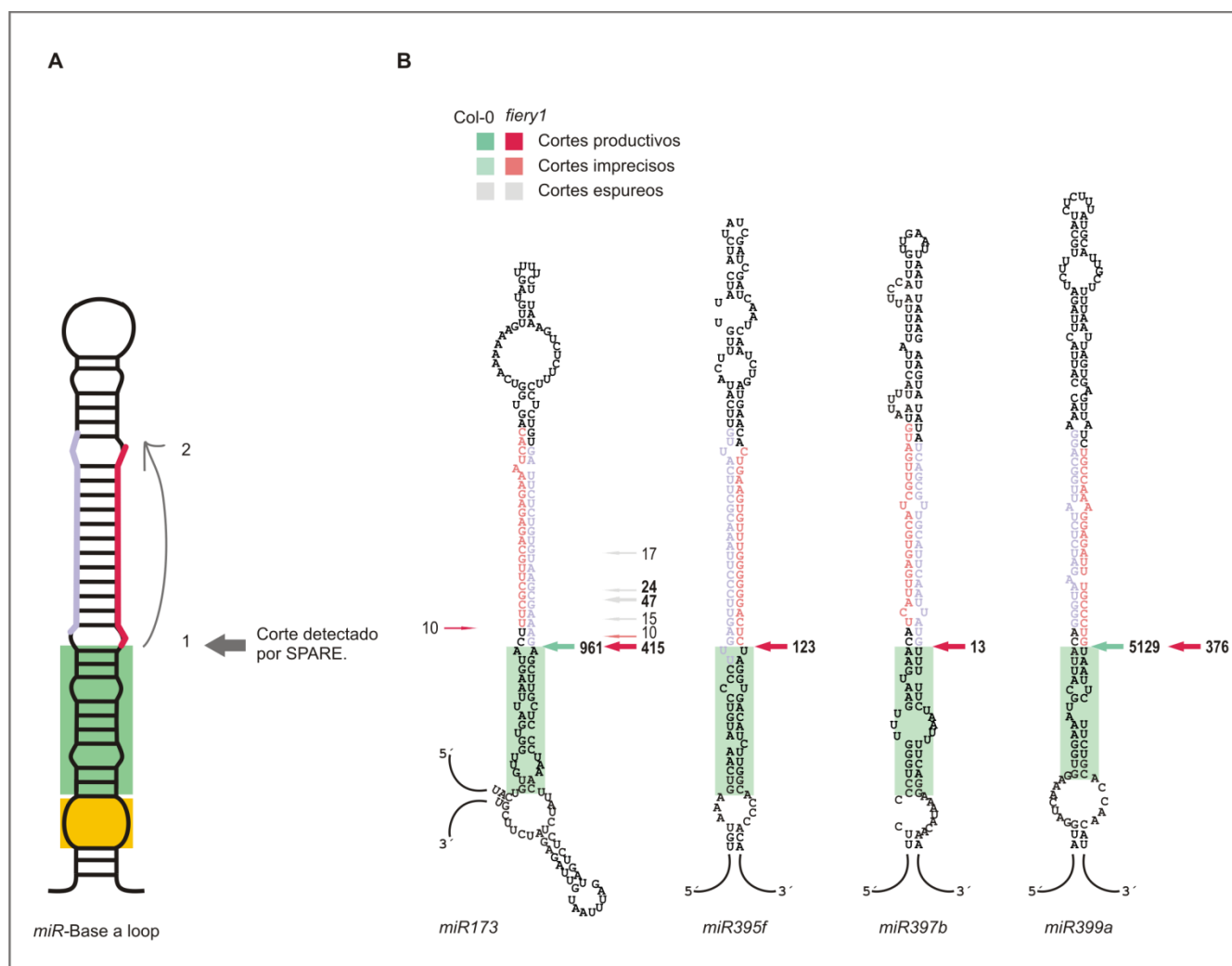


Figura R1.4. Identificación de sitios de corte en precursores de miARNs procesados en dirección de base a loop con dos cortes.

A) Representación esquemática de un precursor procesado desde la base al loop terminal con dos cortes de DCL1. Se muestra la posición del primer y segundo corte, así como el único corte detectado por SPARE. El tallo inferior reconocido por la maquinaria de procesamiento se marca con un recuadro verde.

B) Estructuras secundarias predichas para los precursores de miR173, miR395f, miR397b y miR399a. El miARN está en color rojo y el miARN* en celeste. Las flechas indican las posiciones de los diferentes cortes: de biogénesis verde oscuro y rojo oscuro, impreciso, verde claro y rosa, y espurio, en color gris, según se hayan obtenido en Col-0 o *fiery1* respectivamente. Cada flecha está acompañada de un número que representa la abundancia de dicho corte. En todos los casos los 3 cortes mayoritarios están indicados en negrita.

6.4.2. Caracterización de precursores procesados desde la base con tres cortes de DCL1.

Los precursores procesados desde la base de manera secuencial, presentan un único corte en las bibliotecas de SPARE. Sin embargo este corte se encuentra ~21 nucleótidos corriente abajo del dúplex miARN/miARN* (Figura R1.5A). Se encontraron tres nuevos precursores de miARNs con este patrón de cortes. En la Figura R1.5B se muestran los resultados para los precursores de miR169i y miR402 (el precursor de miR169k se muestra en el ANEXO A18).

En estos precursores el loop interno se encuentra a ~35pb por debajo del dúplex de miARN/miARN*. El tallo inferior está compuesto por un primer segmento de ARNdH de ~15pb por encima del loop interno, el cual es reconocido por la maquinaria de procesamiento para posicionar el primer corte (Figura R1.5B, recuadro verde). A continuación existe un segundo segmento de ARNdH de ~21pb, que coincide con la distancia esperada entre dos cortes de DCL1. El mismo dúplex de ARNdH de ~35pb se ha encontrado en otros precursores que se procesan desde la base con tres cortes de DCL1 (Bologna et al., 2013b).

Es común encontrar ARNs pequeños derivados de los precursores que se procesan de modo secuencial (Addo-Quaye et al., 2009; Bologna et al., 2009; Bologna et al., 2013b; Zhang et al., 2010). Se hizo una búsqueda en bibliotecas de ARNs pequeños disponibles públicamente (<https://mpss.danforthcenter.org/>) y se encontraron ARNs pequeños cuyos extremos provenían del primer y segundo corte de DCL1 (Figura R1.5 B, líneas grises). Por lo tanto, el patrón de cortes obtenido por SPARE es concordante con la identificación de ARN pequeños a partir de las células vegetales, lo que muestra que a partir de aproximaciones experimentales diferentes se obtienen resultados coincidentes.

Tanto *MIR169i*, como *MIR169k* pertenecen a una familia numerosa de *MIARNs*, ampliamente distribuida en las angiospermas en general (Cuperus et al., 2011). El resto de los miembros descritos de esta familia, comparten la misma dirección de procesamiento con tres cortes por DCL1, a excepción del precursor *demir169a* el cual se procesa desde la base pero con dos cortes (Bologna et al., 2013b). Por otro lado, el *MIR402* es un miARN de evolución reciente presente sólo en *Arabidopsis thaliana* y *Arabidopsis lyrata* (Taylor et al., 2014). El mismo afecta la germinación de las semillas bajo condiciones de stress, al regular a el ARNm de la proteína DEMETER-LIKE Protein3 (Kim et al., 2010).

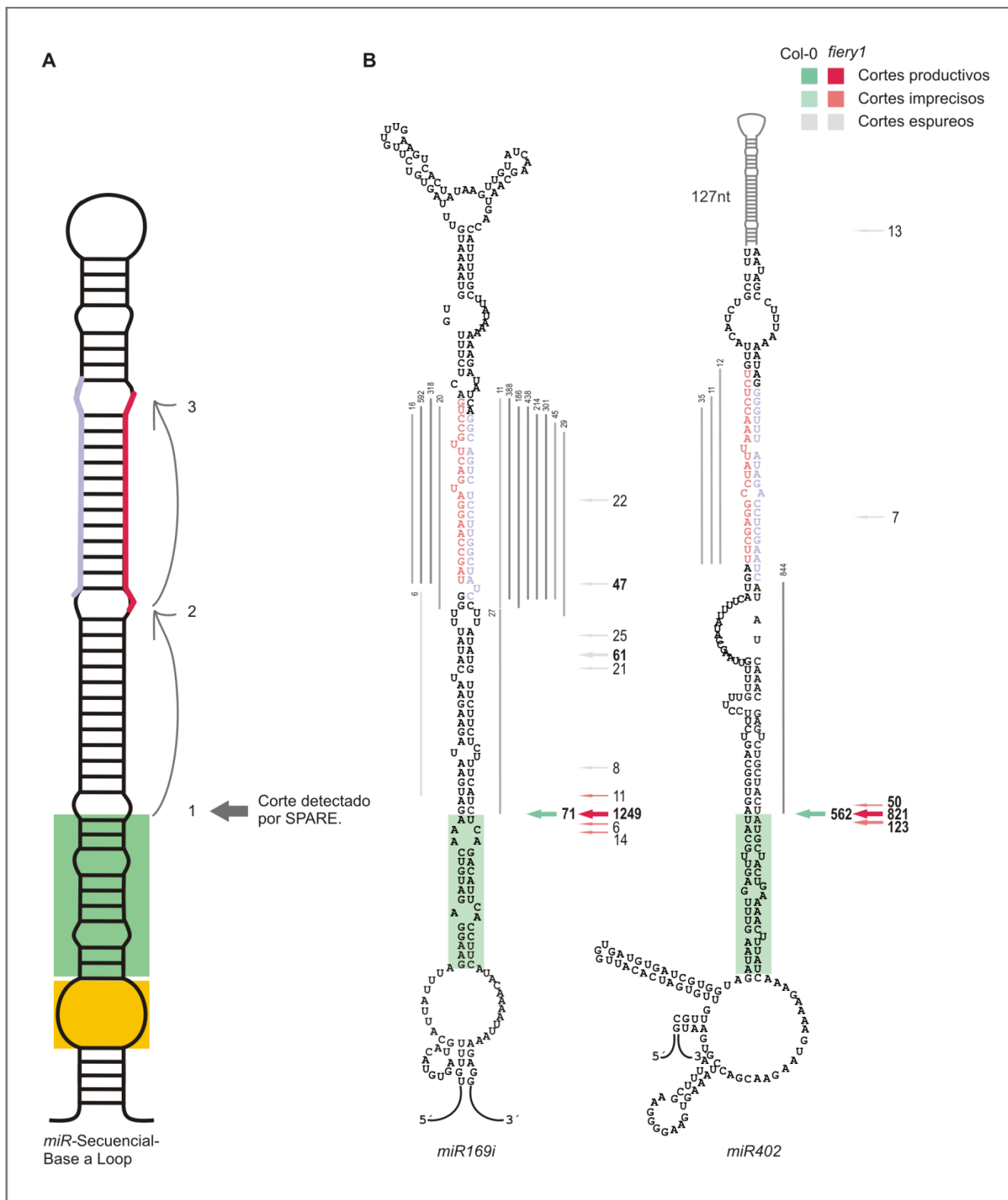


Figura R1.5. Identificación de sitios de corte en precursores de miARNs procesados secuencialmente desde la base al loop.

A) Representación esquemática de un precursor procesado desde la base al loop con tres cortes de DCL1. Se muestra la posición del primer, segundo y tercer corte, así como el único corte detectado por SPARE. El tallo inferior reconocido por la maquinaria de procesamiento se marca con un recuadro verde.

B) Estructuras secundarias predichas para los precursores de miR169i y miR402. El miARN está en color rojo y el miARN* en celeste. Las flechas indican las posiciones de los diferentes cortes: de biogénesis verde oscuro y rojo oscuro, impreciso, verde claro y rosa, y espurio, en color gris, según sean en Col-0 o *fiery1* respectivamente. Cada flecha está acompañada de un número que representa la abundancia de dicho corte. En todos los casos los 3 cortes

mayoritarios están indicados en negrita. ARNs pequeños derivados de estos precursores se indican con una línea gris, y la abundancia de los mismos se indica con un número.

6.4.3. Caracterización de precursores procesados desde el loop hacia la base con dos cortes de DCL1

Los precursores procesados desde el loop hacia la base tienen dos cortes en las bibliotecas de SPARE. Estos se encuentran en los extremos del miARN/miARN* (Figura R1.6A). Se hallaron cinco nuevos precursores con este tipo de cortes. En la Figura R1.6B se muestran los resultados obtenidos para los precursores *demir162a* y *miR408*. Los precursores de *miR157a*, *miR157b* y *miR157c*, que comparten este mecanismo de procesamiento, se desarrollaran en el capítulo de RESULTADOS II (ANEXO A3).

Se ha visto que muchos precursores procesados desde el loop tienen una región superior estructurada de ~42nt, que incluye un segmento de ARNd_h y un loop terminal pequeño, por encima del dúplex del miARN (Bologna et al., 2013b). Si bien tanto el precursor de *miR162a* como el precursor de *miR408* presentan estas características (Figura R1.6B, recuadro verde), los determinantes estructurales de este tipo de mecanismo de procesamiento no han sido evaluados experimentalmente por el momento.

El precursor de *miR408* había sido descrito previamente como un precursor procesado desde la base al loop (Bologna et al., 2013b) sin embargo, en estas bibliotecas resulta claro que hay dos posiciones de corte para el mismo. Una re-evaluación de los datos originales, mostró que el precursor de *miR408* presentaba pocas lecturas, en contraste con las 20.190 lecturas para las bibliotecas de Col-0. Las modificaciones técnicas permitieron obtener un mayor número de lecturas para este precursor, lo cual permitió corregir la dirección de procesamiento adjudicada.

Tanto *MIR162a* como *MIR408* están conservados en angiospermas en general, y cumplen diversas funciones (Cuperus et al., 2011). *miR408* participa de la homeostasis del Cu al regular los niveles de plastocianina (Maunoury and Vaucheret, 2011). Por otro lado *miR162* es un miARN de vital importancia, ya que es el regulador negativo de DCL1, la enzima responsable de la biogénesis de miARNs (Xie et al., 2003). Es el único miembro de la familia, cuya dirección de procesamiento se conoce y pudo establecerse tanto en plantas silvestres como en *fiery1*. También se detectaron los cuatro cortes que requiere el precursor de *miR159a* para liberar el miARN maduro. Este mecanismo de procesamiento secuencial es compartido por otros miembros de la familia (Bologna et al., 2009). Este miARN regula a factores de transcripción MYB (Palatnik et al., 2007)(ANEXO A5).

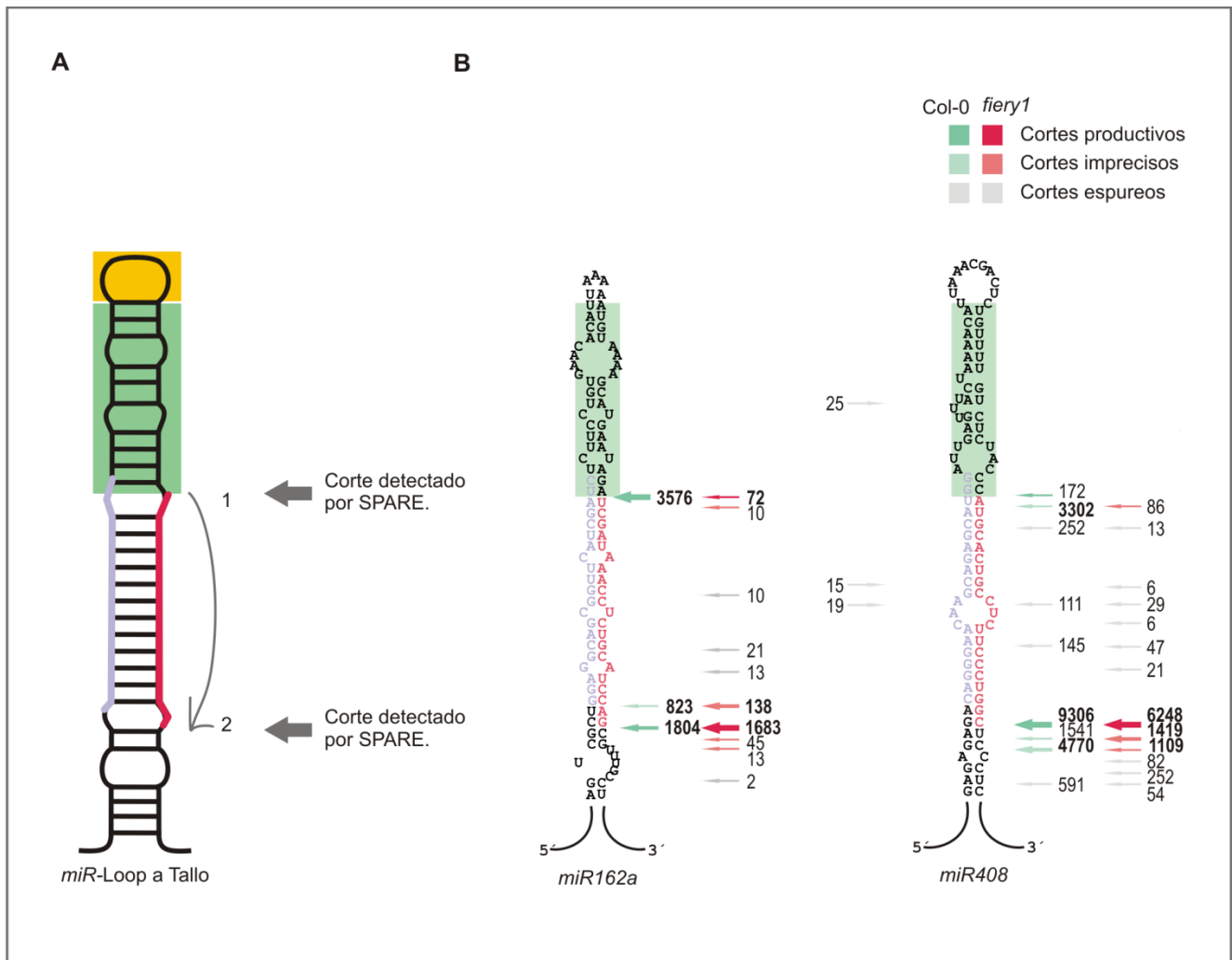


Figura R1.6. Identificación de sitios de corte en precursores de miARNs procesados en dirección loop a base.

A) Representación esquemática de un precursor procesado desde el loop a la base con dos cortes de DCL1. Se muestra la posición del primer y segundo corte, ambos detectados por SPARE. La región superior, compuesta por un segmento de ARNdH, en color verde y el loop terminal, en color amarillo, serían reconocidos por la maquinaria de procesamiento.

B) Estructuras secundarias predichas para los precursores de miR162a y miR408. El miARN se muestra en color rojo y el miARN* en celeste. Las flechas indican las posiciones de los diferentes cortes: de biogénesis verde oscuro y rojo oscuro, impreciso, verde claro y rosa, y espurio, en color gris, según sean en Col-0 o *fiery1* respectivamente. Cada flecha está acompañada de un número que representa la abundancia de dicho corte. En todos los casos los 3 cortes mayoritarios están indicados en negrita.

6.5. Detección de cortes abortivos en precursores de miARN *in vivo*.

En las bibliotecas de SPARE, se obtuvieron cortes que flanqueaban al miARN/miARN*. Estos cortes son los responsables de la biogénesis del miARN maduro y dada la estrategia diseñada también permiten dilucidar la dirección de procesamiento. Sin embargo, para ciertos precursores de miARNs, además se recuperaron cortes que caían dentro del dúplex miARN/miARN* (Figura R1.4B precursor de miR173, Figura R1.5B precursor de miR169i y miR402 y Figura R1.6B precursor de miR162a y miR408). Este tipo de cortes espurios son abortivos ya que anulan la posibilidad de generar el miARN a partir del precursor.

A continuación, nos concentramos en aquellos precursores que presentaban cortes abortivos con una abundancia del al menos el 5% respecto a los cortes productivos responsables de la biogénesis.

Encontramos cortes en ocho precursores que satisfacían este criterio. Los precursores de miR164b, miR395c y miR158a, son procesados desde la base al loop con dos cortes, a partir de reconocimiento de un tallo inferior de ARNdh ~ 15 pb (Figura R1.7A, línea verde). En estos precursores encontramos cortes improductivos dentro de la región del miARN/miARN*. A continuación procedimos a evaluar si estos cortes podían deberse a la actividad de la maquinaria de procesamiento de miARNs mediante el reconocimiento de determinantes estructurales secundarios presentes en el precursor.

En el caso del precursor de miR164b, vimos que los cortes improductivos se encontraban a una distancia de ~ 15 -18pb de un loop interno contenido dentro del tallo inferior de ARNdh (Figura R1.7A, recuadro amarillo). Un patrón similar puede observarse en el precursor de miR158a. A partir de aquí, se puede proponer que la maquinaria de procesamiento estaría reconociendo equivocadamente un determinante estructural secundario, similar al que produce el miARN endógeno.

El precursor de miR400 es procesado desde el loop, probablemente debido al reconocimiento de la región terminal de ~ 38 nt (Figura R1.7B, línea gris). En este caso los cortes abortivos coinciden con el reconocimiento de un segmento de ARNdh presente en el tallo inferior, ~ 15 pb por debajo del dúplex del miARN y por encima de un loop interno (Figura R1.7B, recuadro amarillo). A diferencia de lo mencionado en el párrafo anterior, aquí la dirección de procesamiento y la dirección de los cortes abortivos serían opuestas.

Los precursores de miR393a, miR393b, miR396 y miR864 se procesan desde la base, a través del reconocimiento de un tallo inferior de ARNdh de ~ 15 pb, indicado con una línea verde (Figura R1.7C). Sin embargo ninguno presenta un loop inferior que pudiera ser responsable de los cortes abortivos. Por el contrario, todos presentan un loop terminal ramificado de gran tamaño, ubicado a ~ 15 -19pb de los cortes improductivos. Recientemente, se ha mostrado que ciertos loops terminales pueden ser reconocidos por la maquinaria de procesamiento (Zhu et al., 2013), lo que daría una explicación al mecanismo de origen de estos cortes.

Trabajos anteriores han mostrado que otros precursores de miARNs son susceptibles a cortes abortivos, entre ellos se destacan el precursor de miR472, el cual no fue detectado en nuestras bibliotecas y el precursor de miR825. Para precursor se detectó un corte abortivo localizado a seis nucleótidos del extremo inferior del miARN, que es idéntico al que se encuentra publicado, lo cual refleja la reproducibilidad los datos obtenidos a partir de bibliotecas de SPARE (Bologna et al., 2013b)(ANEXO 41). Es interesante destacar que los cortes improductivos se encontraron tanto en precursores conservados como jóvenes, que se procesan en dirección de base a loop y de loop a la base. A su vez no parece haber una distinción entre en miARNs que responden a stress o regulen factores de transcripción asociados al desarrollo. Es tentador pensar que los cortes improductivos representen un nivel adicional de regulación de la expresión génica, y cuya función biológica sea controlar las cantidades finales de miARN maduro.

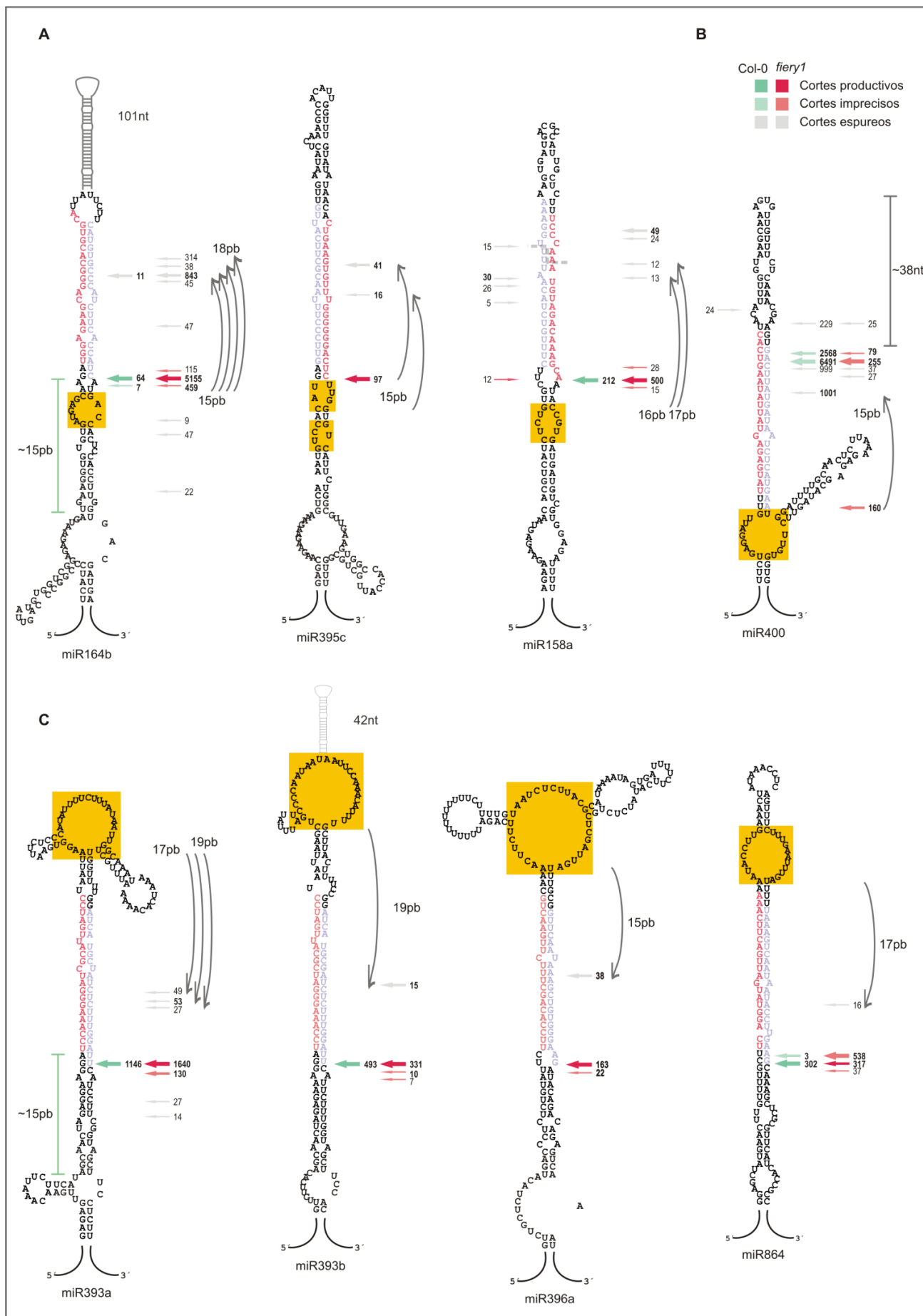


Figura R1.7. Detección de cortes improductivos en precursores de miARNs *in vivo*.

A) Estructuras secundarias predichas para los precursores de miR164b, miR395c, miR158a y miR400. El miARN se muestra en color rojo y el miARN* en celeste. Estos precursores son procesados desde la base al loop, a partir de reconocimiento de un tallo inferior de ARNdh de ~15pb, indicado con una línea verde. Las flechas indican las posiciones de los diferentes cortes: de biogénesis verde oscuro y rojo oscuro, impreciso, verde claro y rosa, y espurio, en color gris, según sean en Col-0 o *fiery1* respectivamente. Los cortes abortivos coinciden con el reconocimiento de un loop interno (recuadro amarillo). En todos los casos los 3 cortes mayoritarios están indicados en negrita.

B) Estructura secundaria predicha para el precursor de miR400. El miARN se muestra en color rojo y el miARN* en celeste. El mismo es procesado desde el loop a la base, probablemente a partir de reconocimiento de la región terminal ~38nt, línea gris. Las flechas indican las posiciones de los diferentes cortes: de biogénesis verde oscuro y rojo oscuro, impreciso, verde claro y rosa, y espurio, en color gris, según sean en Col-0 o *fiery1* respectivamente. Los cortes abortivos se coinciden con el reconocimiento de una burbuja interna (recuadro amarillo). En todos los casos los 3 cortes mayoritarios están indicados en negrita.

C) Estructuras secundarias predichas para los precursores de miR393a, miR3393b, miR396b y miR864. El miARN se muestra en color rojo y el miARN* en celeste. Todos estos precursores son procesados desde la base al loop, a partir de reconocimiento de un tallo inferior de ARNdh de ~15pb, línea verde. Las flechas indican las posiciones de los diferentes cortes: de biogénesis verde oscuro y rojo oscuro, impreciso, verde claro y rosa, y espurio, en color gris, según sean en Col-0 o *fiery1* respectivamente. Los cortes abortivos coinciden con el reconocimiento de una burbuja terminal ramificada (recuadro amarillo). En todos los casos los 3 cortes mayoritarios están indicados en negrita.

6.6. Mecanismos de procesamiento caracterizados por SPARE.

A continuación se resumen los precursores cuyo mecanismo de procesamiento se ha caracterizado mediante la técnica de SPARE. Teniendo en cuenta los resultados provenientes de las bibliotecas de Col-0 y *fiery1* y, se han dilucidado los mecanismos de procesamiento de 75 precursores de miARNs conservados y 16 miARNs jóvenes. De este total, 16 se obtuvieron como resultado de la técnica de SPARE desarrollada en este capítulo (Figura R1.4, Figura R1.5, Figura R1.6) (Bologna et al., 2013b).

En la Figura R1.8 cada columna representa una familia de *MIARNs* conservados en Arabidopsis, mientras que las filas representan a los miembros de una familia. Las flechas violetas indican aquellos precursores que son procesados desde el loop hacia la base, ya sea con dos cortes (un solo bloque) o cuatro cortes de DCL1 (tres bloques). Por otro lado, las flechas rosas muestran los precursores que son procesados desde la base al loop, ya sea con dos cortes (un bloque) o tres cortes de DCL1 (dos bloques). En todo los casos el miARN se representa como un bloque relleno.

A su vez se marcan con flechas vacías los precursores de miR398a y miR166c cuyo mecanismo de procesamiento se ha establecido mediante el uso de otras técnicas (Mateos et al., 2010; Zhu et al., 2013). El precursor de miR159c fue marcado con una X ya que se ha demostrado que el mismo se trata una versión no funcional o pseudo gen (Palatnik et al., 2007). Los precursores cuyo mecanismo de procesamiento se desconoce se marcaron con un signo de interrogación (?).

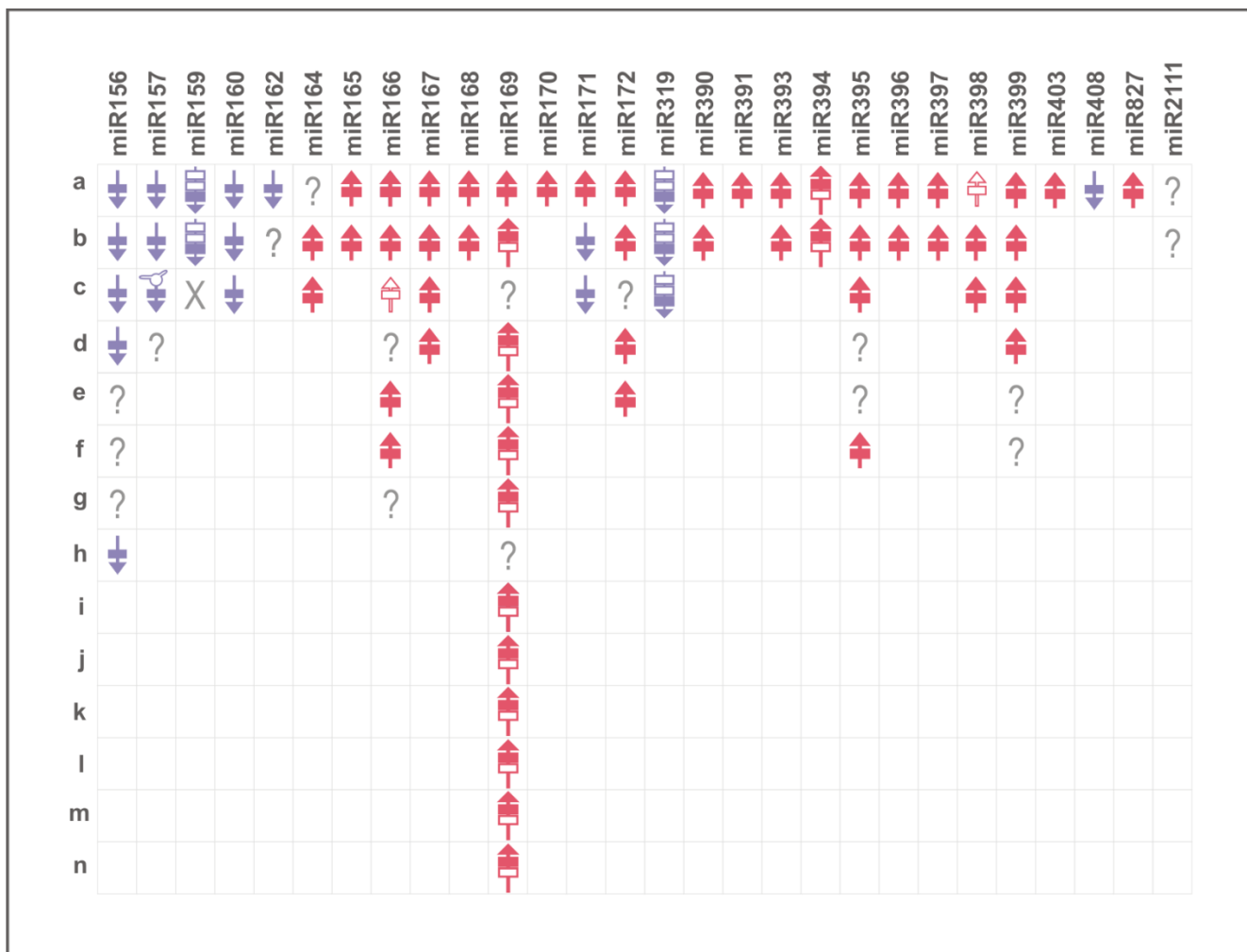


Figura R1.8. Mecanismos de procesamiento de los *MIARNs* de Arabidopsis.

Cada columna representa una familia de *MIARNs* conservados de Arabidopsis, mientras que las filas representan a los miembros de cada familia. Las flechas indican la dirección de procesamiento de cada precursor, mientras que el número de bloques refleja si se trata de un mecanismo de procesamiento de dos cortes (un bloque) o secuencial (dos o más bloques).

6.7. DISCUSION I

A partir de los datos genómicos obtenidos de las bibliotecas de SPARE se caracterizaron los mecanismos de procesamiento responsables de la biogénesis de 91 miARNs en *Arabidopsis*. Los mismos pueden ser agrupados en cuatro rutas dependiendo de la dirección y el número de cortes requeridos para liberar al miARN maduro (Figura D1, panel izquierdo). La clasificación de los precursores según su mecanismo de procesamiento, puso de manifiesto los determinantes estructurales que guían los cortes de DCL1 (Bologna et al., 2013b).

El mecanismo de procesamiento de la base al loop con dos cortes depende del reconocimiento de un tallo inferior de ARNdh de ~15pb, justo por encima de un loop interno (Bologna et al., 2013b; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Zhu et al., 2013). Este mecanismo de procesamiento es el más frecuente en *Arabidopsis* y el que más se asemeja al procesamiento de animales (Ha and Kim, 2014; Han et al., 2006). Aquellos precursores que poseen un tallo inferior de ARNdh de ~35pb, requieren de tres cortes en total para liberar el miARN maduro. Es por ello que además del reconocimiento inicial, dependen de la actividad secuencial de DCL1 (Bologna et al., 2013b).

Por otro lado los precursores de miARNs de plantas pueden procesarse desde el loop hacia la base con un número variable de cortes. La mayoría de ellos requiere sólo de dos cortes en los extremos del miARN y presentan una región terminal homogénea de ~42nt que incluye un segmento de ARNdh y un loop pequeño (Bologna et al., 2013b). Por último, la súper familia de miR319/159 requiere cuatro cortes para completar la biogénesis del miARN. Estos precursores presentan un tallo superior largo y estructurado, el cual se ha demostrado que es esencial para su biogénesis (Bologna et al., 2009; Bologna et al., 2013b).

Estos mecanismos de procesamiento son productivos, y dan lugar a un miARN cuya secuencia es definida y capaz de regular genes blancos a los cuales reconoce por complementariedad de bases (Bologna and Voinnet, 2014). Sin embargo la mera presencia de un determinante estructural, no garantiza que el corte generado a partir del mismo, de lugar a un miARN. Se ha propuesto que el complejo que contiene a DCL1 es capaz de reconocer transiciones de ARNsh-dh y realizar un corte a ~15-17pb (Zhu et al., 2013). A su vez experimentos de procesamiento *in vitro* han demostrado la posibilidad de cortes no productivos en el precursores de miR166c (Zhu et al., 2013). Aquí, utilizando la técnica de SPARE, hemos visto que estos cortes no productivos son relativamente frecuentes *in vivo* en muchos precursores de miARNs. En principio, el procesamiento abortivo puede iniciarse desde loop internos inferiores o loops terminales ramificados, ambos ubicados a esta distancia del dúplex de miARN/miARN* (Figura D1, panel derecho). Además, la dirección del procesamiento abortivo puede ser igual u opuesta, a la del procesamiento productivo.

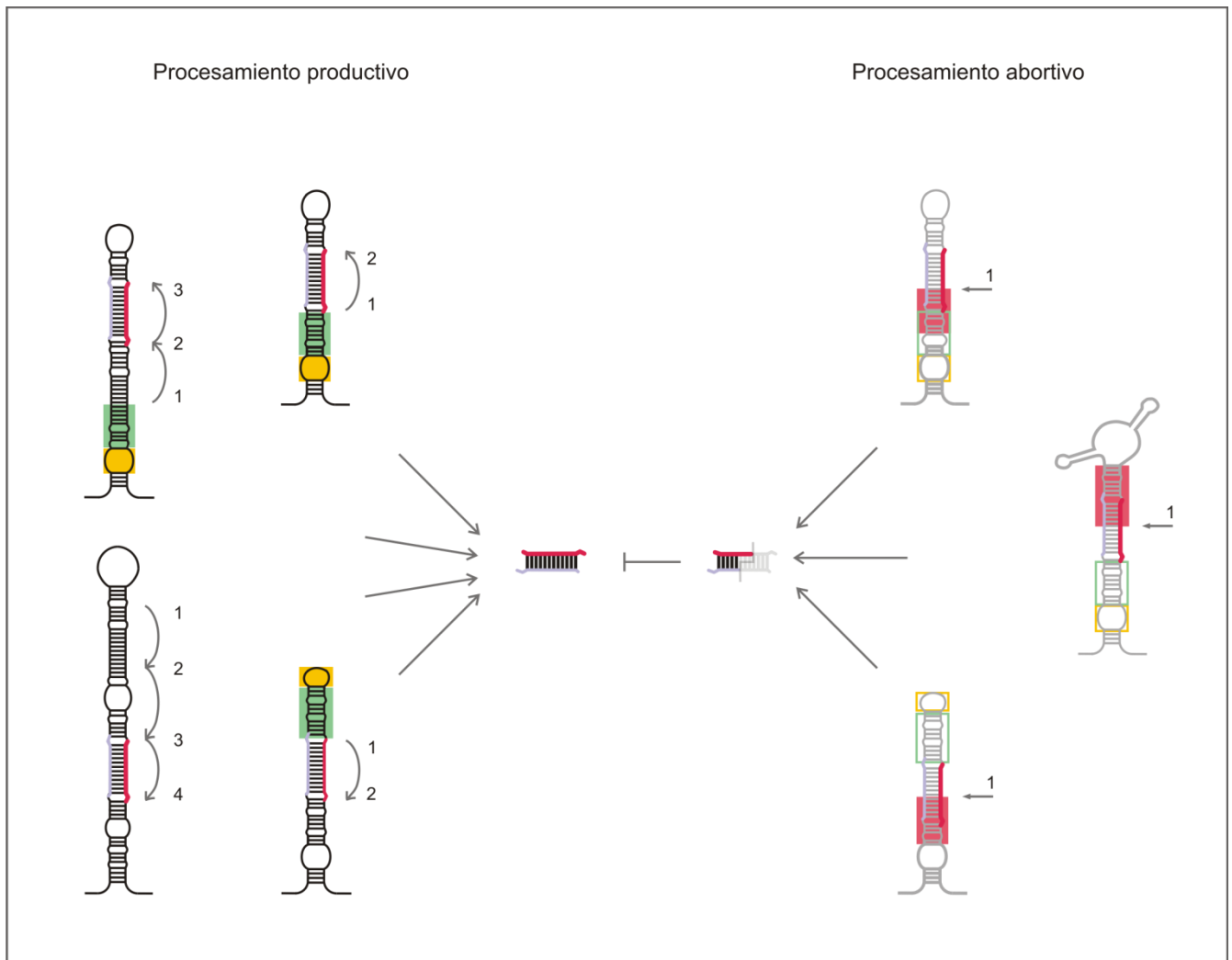


Figura D1. Determinantes estructurales reconocidos por la maquinaria de biogénesis de miARNs responsables del procesamiento productivo y abortivo.

En el panel izquierdo se muestran los cuatro mecanismos de procesamiento presentes en plantas que dan lugar a un miARN maduro. En el panel derecho se muestran las estructuras secundarias alternativas reconocidas por la maquinaria de procesamiento responsable de los cortes abortivos. En todos los casos el miARN se muestra en rojo.

El análisis efectuado aquí, muestra que buena parte de los cortes espurios coinciden con un reconocimiento de determinantes estructurales alternativos localizados en los precursores. Estos cortes improductivos *in vivo* demuestran que la maquinaria de procesamiento es capaz de reconocer múltiples determinantes estructurales independientemente de que estos estén asociados a la biogénesis de miARNs o no.

La plasticidad de la maquinaria de procesamiento podría explicar la diversidad en tamaños y formas en los precursores de miARN en plantas. Además, los determinantes estructurales presentes en los precursores pueden variar mientras las proteínas de procesamiento puedan continuar reconociéndolos como tal. El fort-da adaptativo entre múltiples determinantes estructurales y la maquinaria de procesamiento puede ayudar a constituir la base evolutiva de que exista más de un tipo de mecanismo de procesamiento en plantas.

MIR161 es un microARN joven el cual presenta dos miARNs parcialmente solapados. Sorprendentemente ambos ARNs pequeños tienen función biológica al regular a genes de la familia PPR (Allen et al., 2004). La biogénesis de miR161.1 y miR161.2 está dada por el reconocimiento dual de dos determinantes estructurales presentes en el tallo inferior del precursor de miR161 (ANEXO A37)(Bologna et al., 2013b). Es probable que este mecanismo de biogénesis particular, represente un estadio intermedio en la evolución del mecanismo de procesamiento “final” de miR161.

Del mismo modo un precursor de miARN podría ser reconocido y procesado mediante dos mecanismos opuestos, hasta que uno de los determinantes estructurales se pierda, definiéndose un único mecanismo de procesamiento para ese miembro de la familia. Esto puede haber sucedido para los miembros de la súper familia de miR170/miR171, donde miR170 y miR171a son procesados desde la base y miR171b y miR171c son procesados desde el loop. Esto implicaría que en determinado momento existió un precursor de esta familia con dos determinantes estructurales diferentes (por encima o por debajo del miARN/miARN*) capaces de generar un miR170/171 funcional.

La flexibilidad en el reconocimiento de la maquinaria de procesamiento de determinantes estructurales podría sugerir que la misma sea capaz de realizar cortes directamente sobre otros ARNs como el ARNm. Un sólo corte de DCL1 en un ARNm sería suficiente para promover la degradación de los fragmentos cortados del mensajero o generar una proteína alternativa en un potencial caso que llegue a traducirse. Cualquiera fuera el mecanismo involucrado, la presencia de estructuras secundarias en ARNm que mimetizaran determinantes estructurales presentes de los precursores, sumarían un nivel adicional en la regulación de la expresión génica.

Este tipo de actividad sobre ARNm ha sido demostrada en animales por el complejo que contiene a DROSHA (Macias et al., 2012). Es interesante que al menos uno de estos cortes ocurra sobre el ARNm de DGCR8, que es la proteína accesoria a DROSHA en el complejo microprocesador. Lo que indica que la actividad de DROSHA sobre este ARNm tiene un rol biológico (Ha and Kim, 2014). Sería interesante ver si DCL1 también tiene actividad sobre ARNm en Arabidopsis.

RESULTADOS II.

“Cuando me convenza que la suerte me corre a la par de la pasión”

“Cuando”-Juan Carlos Baglietto.

7. Resultados II

“Caracterización del mecanismo de biogénesis de microARNs que ocurre desde el loop terminal hacia la base del precursor”

7.1. INTRODUCCIÓN II

Los miARNs derivan de largos precursores que se pliegan sobre si mismos en estructuras de tipo tallo y loop. Tanto en plantas como en animales el precursor contiene determinantes estructurales que delimitan la ubicación del miARN a lo largo de su secuencia (Bologna and Voinnet, 2014; Ha and Kim, 2014). Mientras que los precursores de animales presentan tamaños y formas homogéneas, los precursores de plantas presentan estructuras secundarias muy variables tanto en formas como tamaño (Bologna et al., 2009). Esto representa un desafío para la maquinaria de procesamiento responsable de la biogénesis de todos los microARNs.

Dependiendo del precursor, el procesamiento puede empezar en la base del mismo y luego dirigirse hacia el loop terminal, o puede comenzar en el loop terminal y dirigirse a la base. La mayoría de los miARN son liberados mediante un segundo corte de DCL1 que ocurre a ~21nt del primero. Sin embargo algunos precursores requieren de la actividad secuencial de DCL1 y el miARN es liberado luego de tres o cuatro cortes consecutivos (Addo-Quaye et al., 2009; Bologna et al., 2009; Bologna et al., 2013b; Kurihara and Watanabe, 2004). A su vez se ha demostrado que DCL1 puede reconocer un loop terminal ramificado de gran tamaño y realizar cortes abortivos a ~15-17pb del mismo. Este procesamiento de loop a base improductivo compromete cuantitativamente la biogénesis de algunos miARNs, disminuyendo las cantidades de miARN maduro (Zhu et al., 2013).

En el capítulo anterior se describió el mecanismo de procesamiento de más de 90 precursores de miARNs en Arabidopsis. Mientras que los requisitos para el procesamiento de base a loop han sido estudiados en detalle (Cuperus et al., 2010b; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Zhu et al., 2013) se sabe poco acerca de los precursores cuyo procesamiento comienza en el loop terminal y sigue hacia la base. Sin embargo este mecanismo de procesamiento se ha encontrado en ~25% de los precursores de miARNs en Arabidopsis. En por ello que en este capítulo nos centraremos en los requerimientos estructurales que guían el reconocimiento de precursores procesado en dirección loop a base.

7.2. Los precursores que se procesan desde el loop terminal pueden clasificarse en distintos grupos.

Los precursores procesados desde el loop terminal a la base con dos cortes de DCL1 incluyen a miembros de las familias de miR156/157, miR160, miR171, miR162, y miR408 (Figura R2.1AB). Todos estos miARNs conservados desempeñan funciones importantes en plantas través de la regulación de sus genes blancos: los factores de transcripción de las familias “SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL, miR156/157), “AUXIN RESPONSE FACTOR” (ARF, miR160), y SCARECROW-LIKE (SCL, miR171), el metabolismo de ARNs pequeños a través de la regulación de DCL1 (miR162), o el control de niveles de cupro-proteínas (miR408) (Axtell and Bowman, 2008; Cuperus et al., 2011; Jones-Rhoades et al., 2006).

A diferencia de lo que ocurre con los precursores procesados desde la base, los precursores que se procesan desde el loop terminal carecen de una región inferior estructurada (Bologna et al., 2013b; Mateos et al., 2010; Song et al., 2010). Por el contrario, la porción estructurada del precursor se ubica en la región terminal del mismo. La región terminal es definida por el primer corte de DCL1 y consiste en las secuencias de ARN liberadas luego de dicho corte. La región terminal puede dividirse en de un tallo de ARNd_h y un loop terminal definido de tamaño variable (Figura R2.1D) (Bologna et al., 2013b).

Para comenzar el análisis de estos precursores decidimos estudiar la distribución de nucleótidos presentes en la región terminal de aquellos precursores cuyo mecanismo de procesamiento estuviera validado experimentalmente como de loop a base (ANEXO A1-A3, A6-A7, A20 y A36). Observamos, que 14 precursores de miARNs presentan una región terminal homogénea de ~42nt.

Podría definirse un segundo grupo de miARNs procesados del loop terminal, compuesto por los miembros de la súper familia de miR319/miR159, cuyos precursores presentan un tallo terminal de gran tamaño ~120nt. Dichos *MIARNs* requieren un total de cuatro cortes para liberar el miARN maduro lo cual requiere de la actividad secuencial de DCL1, que será desarrollada en el capítulo siguiente (Bologna et al., 2009; Bologna et al., 2013b)(ANEXO A5 y A23).

Finalmente, observamos que el miARN miR157c se destaca por presentar un loop terminal particularmente grande de 85nt (Figura R2.1). Esto sugiere que los precursores procesados desde el loop terminal hacia la base con dos cortes, podrían presentar diferentes determinantes estructurales necesarios para el reconocimiento por la maquinaria de procesamiento.

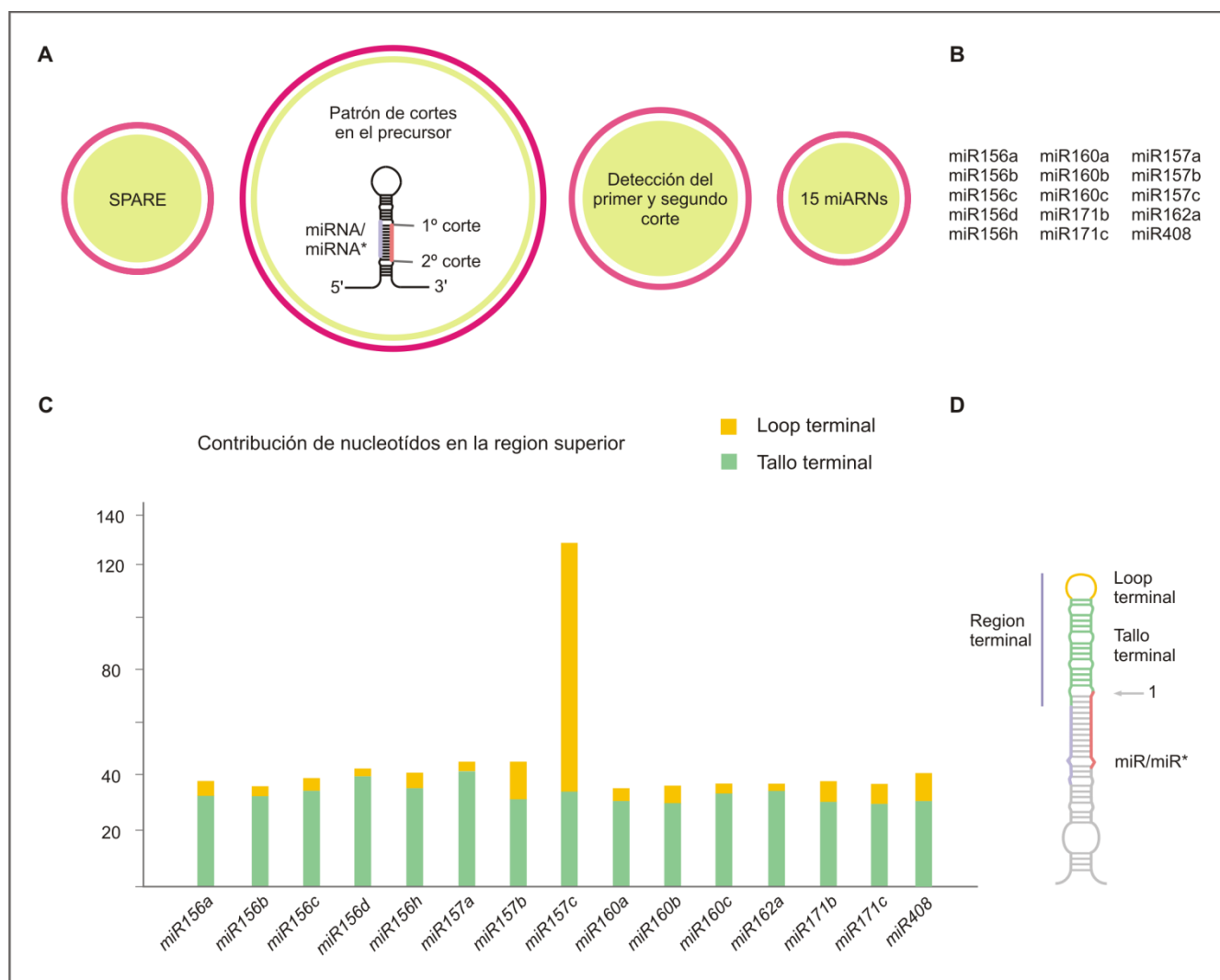


Figura R2.1. La distribución de nucleótidos en la región terminal permite diferenciar a los precursores procesados en dirección loop a base.

- A)** Esquema del patrón de cortes de SPARE para los precursores procesados desde el loop terminal.
- B)** MiARNs conservados cuya biogénesis depende de dos cortes de DCL1.
- C)** Distribución de nucleótidos presentes en la región terminal de los precursores procesados en dirección loop a base.
- D)** Representación esquemática de un precursor donde se muestra el primer corte de DCL1 que define la región terminal, constituida por el tallo (verde) y el loop terminal (amarillo). El miARN se muestra en rojo.

7.3. Caracterización molecular de los precursores procesados en dirección loop a base con región terminal homogénea.

El análisis de la Figura R2.1C reveló que en 14 de los 15 precursores la región terminal presentaba un tamaño homogéneo. Una predicción a partir de esta observación es que tanto el tamaño del tallo como el loop terminal son importantes para el procesamiento de este tipo de precursores. Para evaluar esta hipótesis empleamos como modelo el precursor de miR171b (Figura R2.2A), cuyo mecanismo de procesamiento ha sido descrito por SPARE y validado por mutagénesis dirigida (Bologna et al., 2013b). MiR171b pertenece a la súper familia de miR170/171, compuesta por cuatro miembros en Arabidopsis, de

los cuales dos son procesados desde la base, miR170 y miR171 y dos son procesados desde el loop terminal, miR171b y miR171c (Bologna et al., 2013b; Song et al., 2010).

7.3.1. Identificación de los determinantes estructurales del procesamiento de miR171b.

A partir del precursor silvestre de miR171b se generó una versión mutante denominada *miR171b-super-tallo* en la cual se extendió el tallo superior en cinco pares de bases (Figura R2.2 A, bases en color verde). Para ello en cada hebra del precursor se insertaron cinco nucleótidos a modo de generar un ARNdH. La secuencia insertada se eligió de modo que presente un mismatch en el centro, ya los precursores de miARN no suelen tener largos segmentos de ARNdH perfectos.

Ambos vectores se expresaron en plantas de Arabidopsis bajo el control del promotor vírico 35S. A partir de muestras de inflorescencias de T1 independientes se realizó una extracción de ARN y blots para ARNs pequeños. Estos mostraron que *miR171b-super-tallo* no acumulaba miARN maduro (Figura R2.2 B). A continuación se generó una segunda versión mutante de miR171b llamada *miR171b-super-loop*, en la cual se insertaron ocho bases en el loop terminal (Figura R2.2 A, bases en color amarillo). Esta mutante también se sobre-expresó en plantas de Arabidopsis y se hicieron blots para ARNs pequeños (Figura R2.2 B). Como sucedió en el caso anterior, las plantas sobre-expresantes fueron incapaces de acumular el miARN maduro como consecuencia de la inserción en el loop terminal.

Estos resultados mostraron que modificar el tamaño de la región terminal del precursor de miR171b impide la acumulación del miARN maduro. En este aspecto, los cambios generados en el tallo y el loop terminal son equivalentes ya que ambos impiden el procesamiento adecuado del precursor de miR171b. A partir de esto, proponemos que la posición del primer corte se establecería midiendo la distancia desde el loop terminal en pares de bases. Lo cual explicaría por qué no se toleran mutaciones que modifiquen dicho determinante estructural.

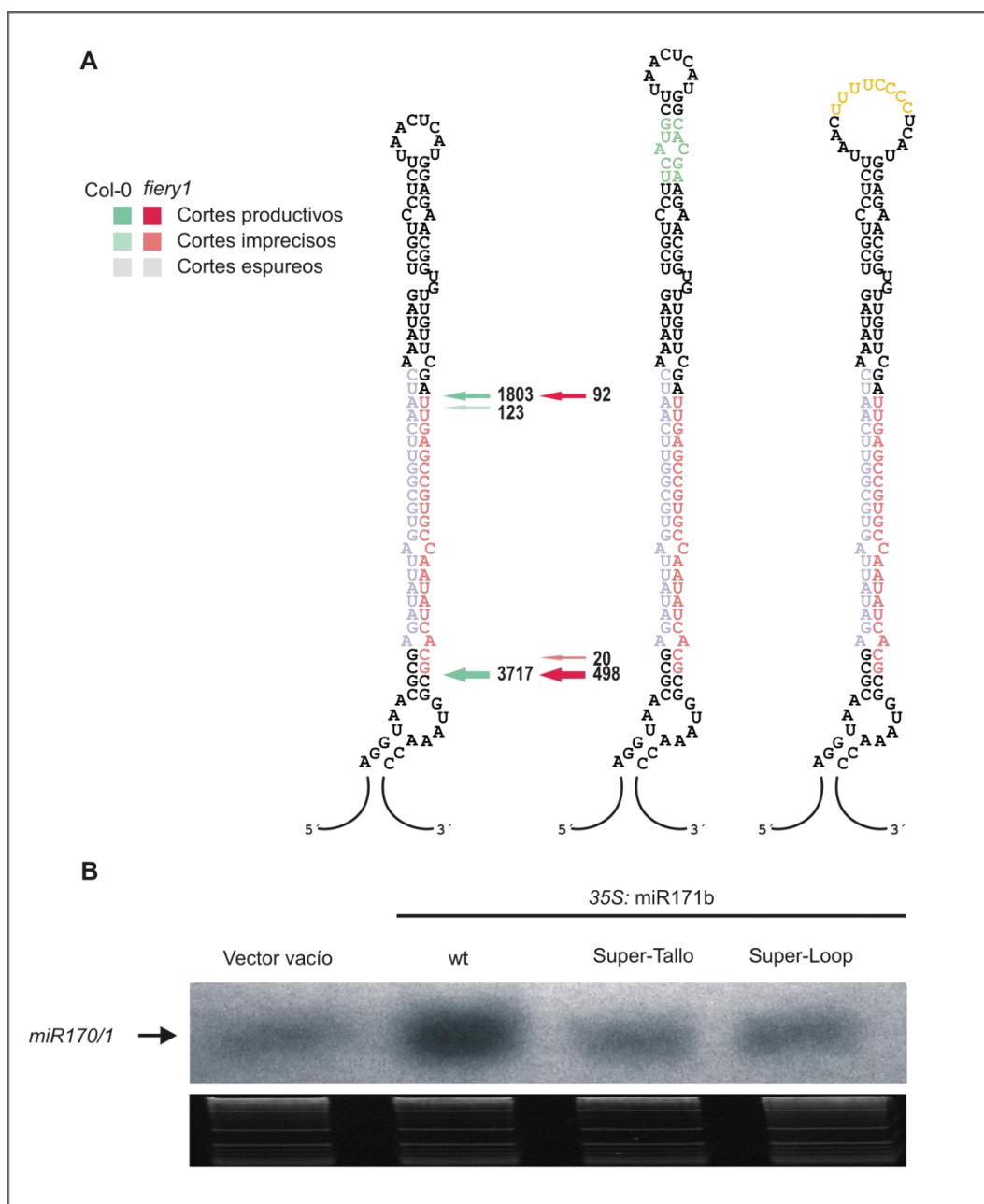


Figura R2.2. Modificaciones en la región terminal de miR171b impiden la acumulación de miARN maduro.

A) Estructura secundaria predicha para el precursor de miR171b, *miR171b-super-tallo* y *miR171b-super-burbuja*. Se muestran además los cortes obtenidos por SPARE en plantas silvestres y deficientes en *FIERY1* para miR171b.

B) Blots para ARNs pequeños de líneas transgénicas sobre-expresantes de los diferentes precursores. Al menos 20 líneas independientes se colectaron por muestra.

7.4. Caracterización de los precursores procesados en dirección loop a base que presentan un loop terminal ramificado.

7.4.1. Mecanismos de procesamiento de la familia de miR157.

De los cuatro miembros de la familia del miARN miR157 se determinó el mecanismo de procesamiento de tres de ellos, miR157a, miR157b y miR157c (Figura R2.3A). El cuarto miembro, miR157d no fue detectado en ninguna biblioteca.

Mientras que los precursores de miR157a y miR157b son estructuralmente similares al precursor de miR171b, el precursor de miR157c presenta una región terminal de 125nt en vez de los usuales ~42nts. A pesar de esta diferencia significativa en los tamaños de las regiones terminales, estos precursores son procesados en dirección loop a base con dos cortes de DCL1 (Figura R2.1C y Figura R2.3A). Es por ello que decidimos estudiar esta familia en más detalle.

Las versiones silvestres de los *MIARNS* miR157a, miR157b y miR157c se clonaron desde ADN genómico y se expresaron en plantas de Arabidopsis bajo el promotor constitutivo 35S. Se analizaron más de 50 líneas T1 independientes y se realizaron blots para ARNs pequeños. Los distintos niveles de miR157 acumulados por cada uno de los precursores, mostraron que existe un procesamiento diferencial entre estos tres precursores, siendo el precursor de miR157a procesado más eficientemente y el precursor de miR157c el más pobre (Figura R2.3B).

En la Figura R2.3C se muestra un resumen de los cortes encontrados para los precursores de miR157a-c. Notamos que el 96% de los cortes del miR157a estaba flanqueando exactamente el miARN/miARN*, mientras que sólo el 85% de los cortes de miR157c tenían dicha precisión. Esta falta de precisión durante la biogénesis de miR157c podría estar relacionada con su baja eficiencia de procesamiento, evidenciada mediante por sobreexpresión de los precursores en Arabidopsis. Esto nos llevó a preguntarnos si las diferencias en la región terminal de dichos precursores, eran los responsables de las diferencias en la eficiencia de procesamiento y el patrón de cortes entre estos precursores.

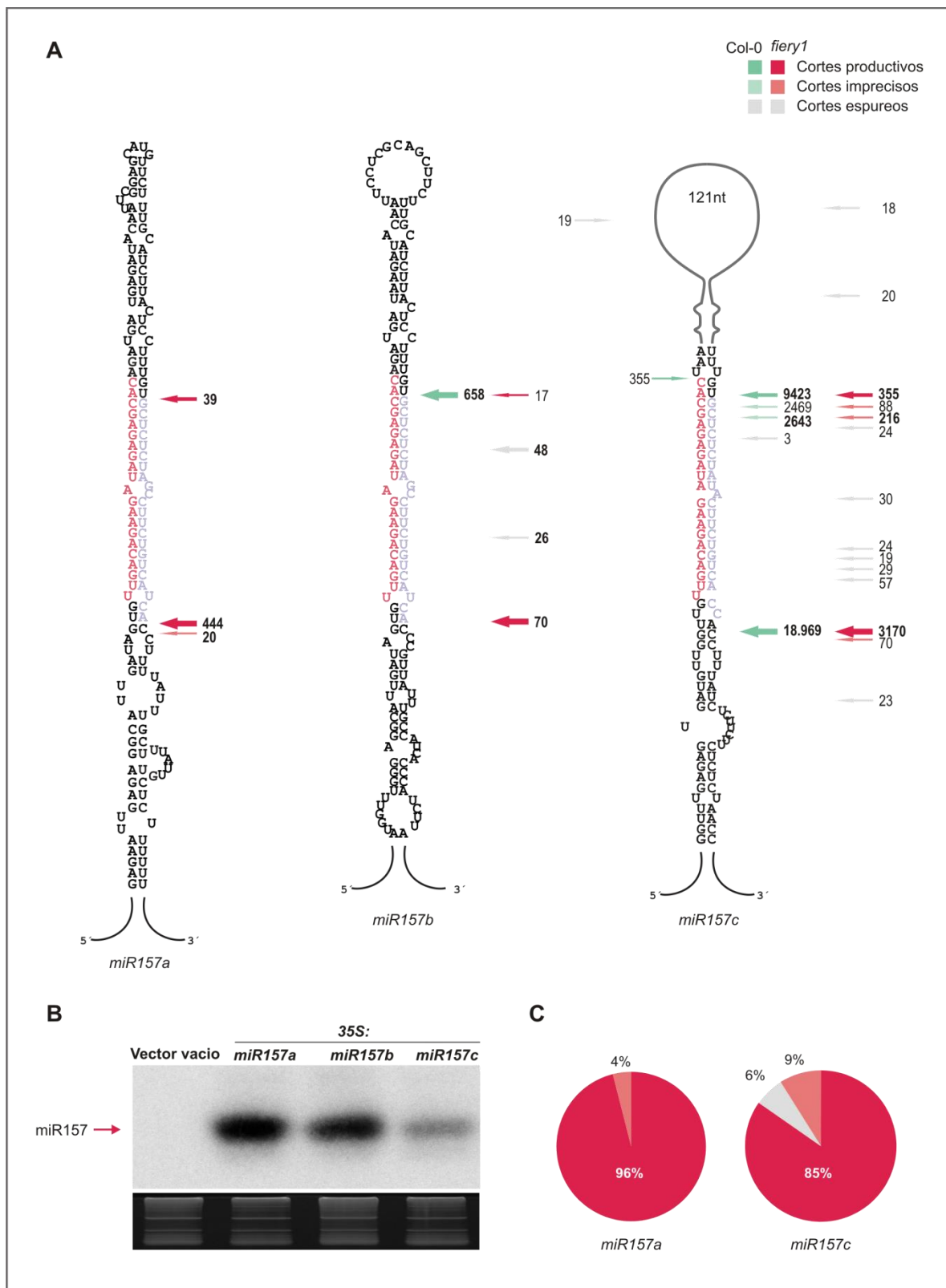


Figura R2.3. Los precursores de miR157a, miR157b y miR157c son procesados con diferentes eficiencias *in vivo*.

A) Estructura secundaria predicha para los precursores de miR157a, miR157b y miR157c. Se muestra el patrón de cortes de SPARE en plantas silvestres y *fiery1*. El miARN se muestra en rojo y el miARN* en celeste.

B) Blots para ARNs pequeños a partir de líneas transgénicas sobre-expresantes de los diferentes precursores. Al menos 25 inflorescencias independientes se colectaron por muestra

C) Gráficos que muestran el porcentaje de los cortes productivos, imprecisos y espurio en *fiery1* para los precursores de miR157a y miR157c.

7.4.2. Mapeo estructural de la región terminal del precursor de miR157c por digestión enzimática.

La predicción informática de los tallos presentes en los precursores de miARNs de plantas generalmente da como resultado una única estructura estable, como es de esperar de una región de ARNdH. Sin embargo es común que la predicción de estructura secundaria de los loops terminales de gran tamaño dé como resultado un conjunto de estructuras posibles con valores de energía libre de Gibbs similares. Considerando además, la importancia de la región terminal en el procesamiento desde el loop hacia la base decidimos determinar experimentalmente la estructura secundaria de la región terminal del precursor de miR157c.

Una de las técnicas empleadas para determinar la estructura secundaria de moléculas de ARN está basada en la existencia de ARNasas que digieren selectivamente el ARN en función de su estructura secundaria. La Nucleasa V1 corta el ARN en regiones donde la hebra se encuentra en conformación de hélice, la cual puede deberse al apareamiento de dos bases complementarias o al apilamiento de las bases nitrogenadas consecutivas en una región de ARNsh (Chaulk et al., 2014; Lowman and Draper, 1986). La enzima requiere de entre cuatro y seis nucleótidos consecutivos en conformación de hélice para poder cortar.

Por otro lado, la Nucleasa S1 digiere selectivamente regiones de ARNsh independientemente de conformación de la hebra (Ehresmann et al., 1987). A su vez, la ARNasa S1 posee actividad exo y endonucleolítica. En combinación estas dos enzimas permiten obtener un mapa de la estructura secundaria cada región de los precursores de miARN, como se resume en la Tabla R2.1.

Tabla R2.1. Análisis de los resultados de digestión enzimática de ARN

Digestión por V1	Digestión por S1	Estructura secundaria	Código de color en las Figuras
+	-	ARN doble hebra	Verde
-	+	ARN simple hebra	Amarillo
+	+	Bases apiladas	Gris

El esquema general de trabajo consiste en realizar una transcripción *in vitro* del precursor de miR157c, y marcar el extremo 5' con ^{32}P . A continuación se desnaturaliza el ARN y se permite que el mismo se pliegue adoptando una estructura secundaria estable por hibridación intramolecular. Luego se incuba con las ARNasas, que cortan el ARN por hidrólisis a la altura del grupo fosfato del enlace fosfodíester. Se obtienen dos fragmentos, uno con un extremo 3'OH y el otro con un 5' H_2PO_4 . Los productos de las digestiones se resuelven por electroforesis en geles desnaturizantes y las bandas se revelan por auto radiografía (Figura R2.4A, Sección 5.9 de materiales y métodos). Las condiciones de digestión se optimizaron a modo de obtener un único corte por molécula de ARN.

En la Figura R2.4BC se muestran los dos geles de poliacrilamida obtenidos para el precursor de miR157c tratado con las enzimas V1 y S1. Como referencia se usó una escalera de hidrólisis alcalina, donde

se escinden los enlaces fosfodíester entre todas las bases y se obtienen fragmentos que se diferencian entre ellos por una base. Una segunda referencia es la digestión del ARN en condiciones desnaturalizantes con la enzima T1, que corta específicamente luego de las Guaninas (G) desapareadas, T1(d). De este modo se obtiene una banda por cada nucleótido de G presente en el precursor y el patrón de bandas permite identificar la región en secuencia a la que pertenecen. Por último se usó la T1 en condiciones nativa, T1(n), para ubicar aquellas G que estuvieran desapareadas en la estructura plegada del precursor de miR157c (Figura R2.4DE).

La predicción de la estructura secundaria del precursor de miR157c se hizo con el programa Mfold (Zuker, 2003). El mismo arrojó tres estructuras posibles que se diferenciaban por la disposición en el espacio de las bases justo por encima del dúplex miARN/miARN* (posición 23, Figura R2.4D-F). Las estructuras secundarias fueron denominadas: *Viserion* (S1, $\Delta G = -77,58$ kcal/mol), *Drogon* (S2, $\Delta G = -75,66$ kcal/mol) y *Rhaegal* (S3, $\Delta G = -74,85$ kcal/mol) en honor a los dragones de la serie "Game of Thrones", por analogía precursor de miR157c sería el dragón de tres cabezas.

Para dilucidar la estructura secundaria de la región terminal del precursor de miR157c se empleó principalmente la información proveniente de la base 104 y las bases 108-128 (Figura R2.4B, gel al 8% y Figura R2.4C, detalle del gel al 6%). Para cada base en estas posiciones en las tres estructuras secundarias predichas, se contrastó la conformación determinada experimentalmente, simple hebra o doble hebra y la predicción informática. Cuando se ambas conformaciones coincidían, la base se marcó con una línea verde, y cuando los resultados eran opuestos, con una línea roja (Figura R2.4DEF). Para la estructura *Viserion* se encontraron ocho bases que coincidían con conformación experimental y siete bases que diferían de la misma (Figura R2.4D). Para la estructura *Drogon* siete bases se coincidían con la conformación determinada experimentalmente mientras que ocho bases diferían de la misma (Figura R2.4E). Por último para la estructura *Rhaegal* la conformación predicha de las 15 bases analizadas coincidió con la determinación experimental. Esto nos permitió asignar la estructura *Rhaegal* como la estructura secundaria *in vitro* de la región terminal de miR157c (Figura R2.4E).

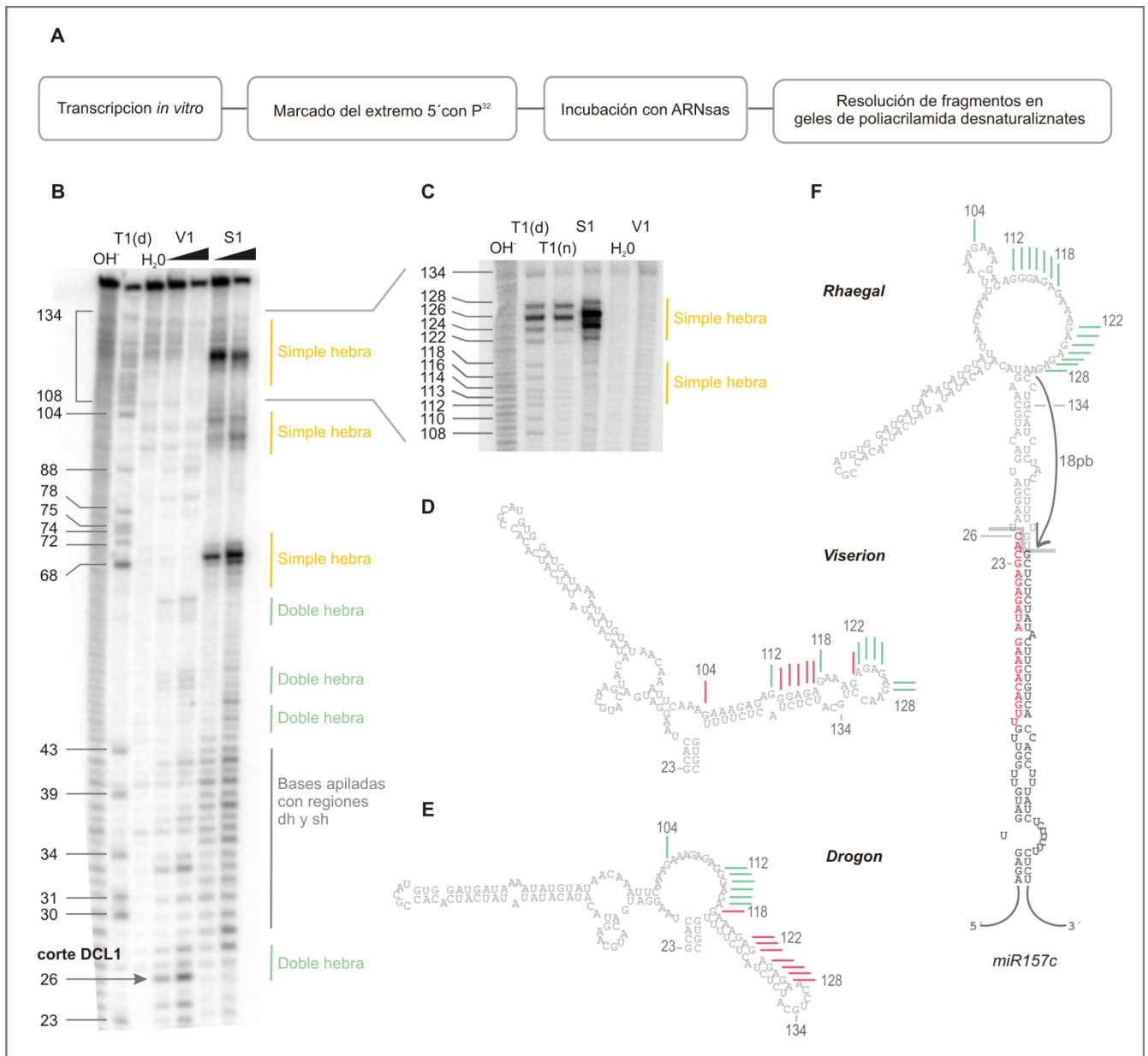


Figura R2.4. El precursor de miR157c posee un tallo de ~18pb y un loop terminal ramificado.

- A)** Esquema de trabajo para el mapeo estructural de la región terminal de *miR157c* por digestión enzimática.
- B)** Gel de poliacrilamida al 8% en condiciones desnaturalizantes de los fragmentos obtenidos por digestión.
- C)** Gel de poliacrilamida al 6% en condiciones desnaturalizantes de los fragmentos obtenidos por digestión.
- D)** Estructura secundaria alternativa brindadas por el programa M fold: *Viserion*
- E)** Estructura secundaria alternativa brindadas por el programa M fold: *Drogon*
- F)** Estructura secundaria de la región terminal de miR157c determinada experimentalmente. La predicción de esta estructura secundaria de obtuvo con el programa M fold, *Rhaegal*. El miARN se muestra en rojo.
- Las bases identificadas como ARNsh experimentalmente (A-C) están marcadas con líneas en las predicciones de estructura secundarias (D-F). Las líneas verdes indican una coincidencia de la estructura experimental y predicha, mientras que las rojas muestran una inconsistencia. La estructura mostrada en F es la que tiene una coincidencia entre la estructura predicha y la experimental.

7.4.3. Caracterización de los determinantes estructurales de la biogénesis de miR157c.

Una vez determinada la estructura secundaria del precursor de miR157c, se identificó la presencia de un loop terminal ramificado y un tallo de ARNdh de ~18pb por encima del dúplex del miARN/miARN*. Con el

objetivo de estudiar en más detalle la biogénesis de miR157c, se diseñaron una serie de mutantes del precursor (Figura R2.5AC).

En primera instancia se clonó una versión del precursor de miR157c en la que se dejaron cuatro pares de bases por debajo del dúplex miARN/miARN* para generar el vector, *miR157c-CORTO* (Figura R2.5A, flecha naranja). Esta versión se introdujo en plantas de *Arabidopsis* bajo el control del promotor 35S en conjunto con el precursor silvestre, y se colectaron muestras 20 de inflorescencias de T1 independientes. El blot para ARNs pequeños mostró que los niveles de miR157 eran iguales para los dos precursores, independientemente de la extensión del tallo inferior (Figura R2.5B). Este resultado nos permitió concluir que el tallo inferior del precursor de miR157c es prescindible para la biogénesis del mismo. Se han obtenidos resultados similares para los precursores de miR319a y miR171b con lo cual es probable de que el tallo inferior sea dispensable para muchos precursores procesados desde el loop hacia la base (Bologna et al., 2009; Bologna et al., 2013b).

Luego decidimos ensayar si el loop terminal ramificado de miR157c cumplía alguna función *in vivo*. Para esto construimos el precursor *miR157c-loop-156b*, (Figura R2.5C) donde el loop terminal fue reemplazado por un loop pequeño de cuatro nucleótidos, como el que tiene el precursor de miR156b. Este precursor fue introducido como control en plantas de *Arabidopsis*, junto con los precursores de miR157c y *miR157c-loop-156b*. El blot para ARNs pequeños mostró que el precursor de *miR157c-loop-156b* acumulaba la mayor cantidad de miARN maduro de todos los precursores sobre-expresados (Figura R2.5F). El hecho de que *miR157c-loop-156b* fuera procesado *in vivo* más eficientemente que la versión silvestre es indicativo de que el loop terminal ramificado afecta negativa y cuantitativamente la biogénesis de miR157c.

Por ultimo hemos diseñado una versión del precursor de miR157c donde se ha acortado el tallo superior en cinco pares de bases, quedando el mismo con ~13pb. El objetivo de esta mutante es comprobar *in vivo* el rol de del tallo superior de ARNdH de ~18pb la biogénesis de miR157c. Se han transformado plantas de *Arabidopsis thaliana* con este vector y al momento de disponer de la semillas T1 se harán caracterizaciones moleculares similares a las otras mutantes presentadas en este capítulo.

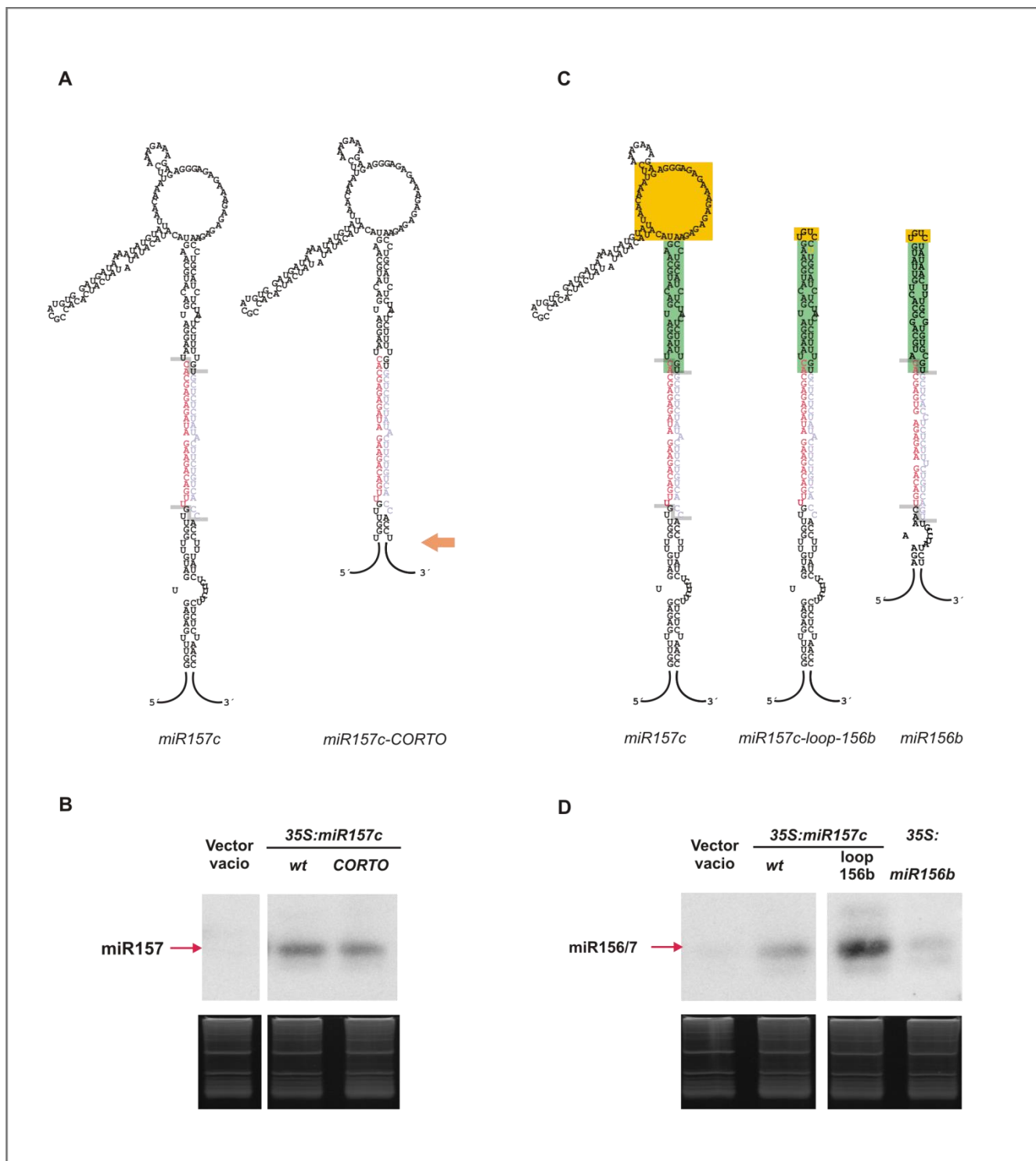


Figura R2.5. El loop terminal ramificado afecta cuantitativamente la biogénesis de miR157c.

A) Estructura secundaria predicha para el precursor de miR157c y *miR157c-CORTO*. La flecha naranja indica las regiones delecionadas. El miARN se muestra en rojo.

B) Blots para ARNs pequeños de líneas transgénicas sobre-expresantes de los diferentes precursores. Al menos 25 líneas independientes se colectaron por muestra

C) Estructura secundaria predicha para los precursores de miR157c, *miR157c-loop-miR156b* y miR156b. Las mutaciones se muestran en amarillo y el tallo terminal se muestra en verde.

D) Blots para ARNs pequeños de líneas transgénicas sobre-expresantes de los diferentes precursores. Al menos 25 líneas independientes se colectaron por muestra

7.5. Análisis global de los precursores procesados con dos cortes de DCL1.

Los resultados obtenidos para el precursor de miR171b nos permitieron proponer un reconocimiento de la región terminal del precursor por la maquinaria de procesamiento. Un análisis de esta región mostró que la misma está compuesta por un loop pequeño de ocho nucleótidos y un tallo superior de 17pb para el precursor de miR171b (Figura R2.6A). Por otro lado, la región inferior del precursor de miR172a consiste en un tallo inferior de ARNdh de ~15pb seguido de un loop interno definido de nueve nucleótidos (Figura R2.6A). Trabajos anteriores han demostrado la importancia del tallo y el loop interno en el procesamiento de miR172a (Mateos et al., 2010; Werner et al., 2010). La presencia de características estructurales similares en la región superior de miR171b y la región inferior de miR172a nos llevó a inspeccionar en detalle las estructuras secundarias presentes todos los otros precursores procesados desde el loop a la base y la base al loop, con dos cortes de DCL1 validados experimentalmente.

En primera medida cuantificamos el número de pares de bases presentes en los tallos superiores de los precursores procesados desde el loop (Figura R2.6C, izquierda). Los mismos tienen un mínimo de 14pb y un máximo de 22pb. Sin embargo la mayoría de los precursores presenta entre ~16-17pb en el tallo superior. A continuación repetimos este procedimiento con los tallos inferiores de los precursores procesados desde la base (Figura R2.6C, derecha). El número de pares de bases presentes en los tallos inferiores variaron desde 12-17pb, mientras que la mayoría de los precursores presentaron entre 14 y 15pb en el tallo inferior. Este valor concuerda con lo reportado en el bibliografía (Bologna et al., 2013b; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Zhu et al., 2013). Una excepción en el número de pares de bases fue el precursor de miR167c, ya que el mismo tenía sólo nueve pares de bases en el tallo inferior (Anexo A8). Un gráfico de box plot, reveló que en este sentido el precursor de mir167c presentaba un valor anómalo de pares de bases en el tallo inferior (Figura R26D). Sería interesante estudiar en mayor detalle la estructura secundaria del precursor de miR167c y su procesamiento para corroborar si existe una correlación entre estas características inusuales y su procesamiento *in vivo*. Cabe destacar que si medimos el tallo superior y el inferior desde el extremo 3 del miARN/miARN* la distancia del tallo es similar y equivalente a los ~15pb tanto para los precursores procesados desde el loop terminal como para los que lo hacen desde la base, lo que unificaría el mecanismo de corte de los precursores.

En segunda instancia, cuantificamos el número de nucleótidos presentes en los loops terminales de los precursores procesados desde el loop (Figura R2.6B, izquierda). El tamaño de los loops terminales definidos varía de tres a 15nt, presentando la mayoría de los precursores un tamaño entre cuatro y seis nucleótidos. Por otro lado el precursor de miR157c presentó el mayor tamaño de todos con 85nt para su loop terminal ramificado (Figura R2.1, R2.3 y R2.4). Al momento de evaluar el tamaño de los loops internos presentes en la región inferior de los precursores procesados desde la base, encontramos un primer grupo estaba presentaba un loop interno definido cuyo tamaño oscilaba entre cuatro y 18nt, mientras que la mayoría de los precursores presentaron un tamaño de nueve nucleótidos. Este loop interno

definido, es comparable en tamaño con la mayoría de los loops terminales de los precursores procesados desde arriba (Figura R2.6B, indicado con corchetes). Mientras que un segundo grupo presentaba un loop interno ramificado, similar al descrito en la Sección 6.5 (Figura R2.6B, indicado con corchetes). En este caso los tamaños del loop interno ramificado variaban de 19 a 36nt, y no se encontró ningún tamaño de loop que fuera representativo de la mayoría. Cabe destacar que si bien estos loops internos ramificados son de mucho menor tamaño que el loop ramificado de miR157c, su estructura secundaria es similar. Teniendo en cuenta los resultados obtenidos en la Sección 7.4.3. sería interesante generar una versión mutante de un precursor procesado desde de la base, donde se reemplace el loop interno ramificado por un loop definido para estudiar su efecto en la biogénesis.

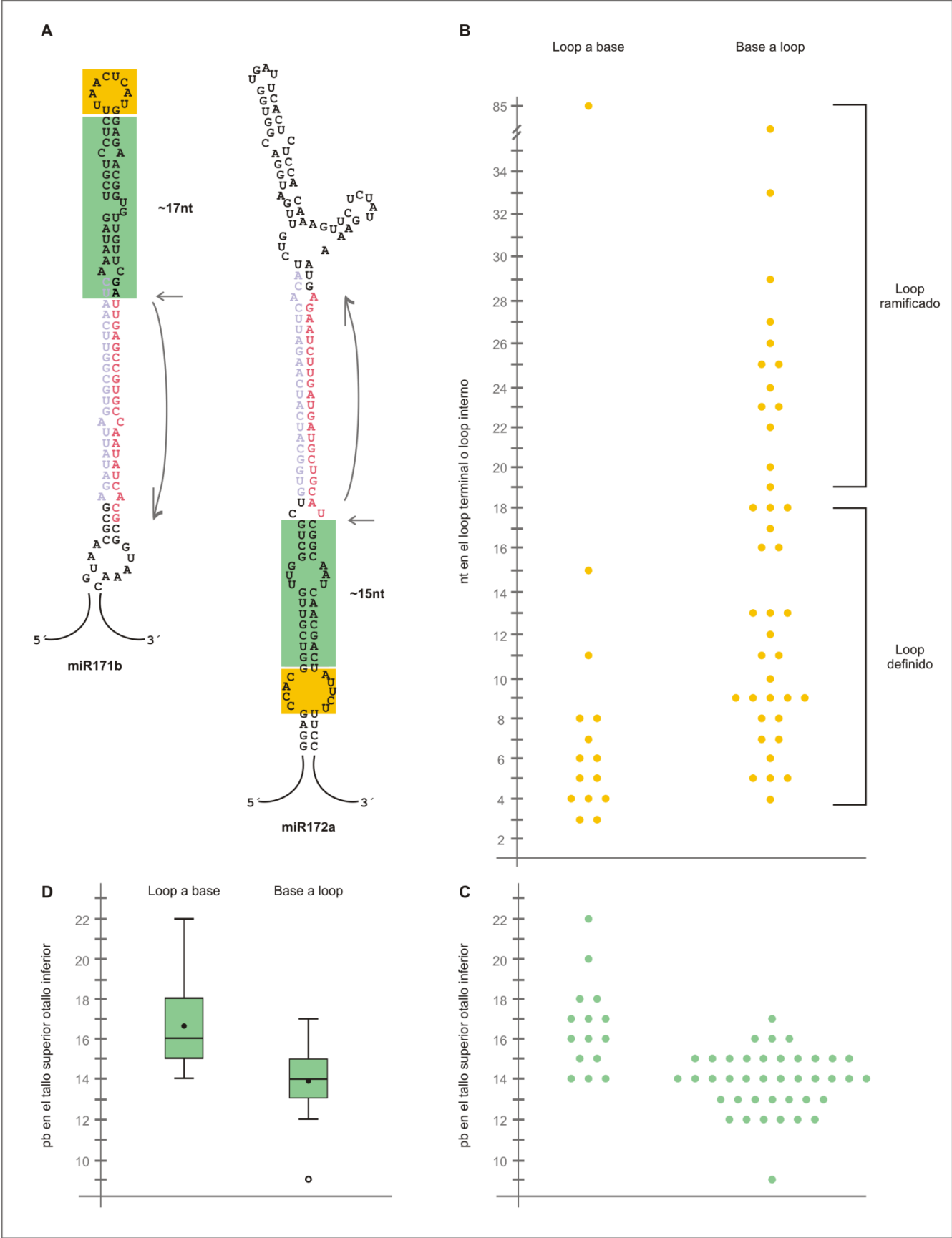


Figura R2.6. Los precursores procesados con dos cortes de DCL1 presentan características estructurales similares en la región terminal y la región inferior.

A) Estructura secundaria predicha para los precursores de miR171b y miR172a. El loop terminal y el loop interno de cada uno se marca en con un recuadro amarillo, mientras que el tallo superior e inferior están delimitados por un

recuadro verde. El primer corte en cada uno de los precursores es indicado con una flecha. El microARN se muestra en color rojo.

B) Gráfico de dispersión que muestra el número de nucleótidos presentes en el loop terminal (izquierda) o el loop interno (derecha) de los precursores procesados desde el loop o desde la base, respectivamente. Se ha hecho una subdivisión ambos loops en función de su estructura secundaria predicha en: loops definidos (abajo) y loops ramificados o no definidos (arriba).

C) Gráfico de dispersión que muestra el número de pares de bases presentes en el tallo terminal (izquierda) o el tallo inferior (derecha) de los precursores procesados desde el loop o desde la base, respectivamente.

D) Gráfico de box plot donde se grafica la distribución de pares de bases presentes en el tallo terminal (izquierda) o el tallo inferior (derecha) de los precursores procesados desde el loop o desde la base, respectivamente. Los valores anómalos se muestran como un círculo vacío.

7.6. DISCUSION II

7.6.1. Requisitos estructurales para el procesamiento desde el loop a la base de precursores con región terminal homogénea.

En la primera parte de este capítulo se caracterizó el procesamiento de precursores que se procesan por dos cortes de DCL1 en dirección loop a base. A partir del precursor de miR171b se generaron dos mutantes. En una se extendió el tamaño de la región de ARNdH por encima del dúplex del miARN, *miR171b-supertallo* y en otra se duplicó el tamaño del loop terminal, *miR171b-superloop* (Figura R2.2). En ambos casos, al modificar el tamaño de la región terminal de miR171b se anuló la acumulación de miARN maduro.

Alineamientos de los ortólogos de miR171b provenientes de otras dicotiledóneas mostraron que las regiones terminales de estos precursores se mantienen constantes en las diferentes especies vegetales (Chorostecki, Moro et al., enviado a publicación). Por el contrario las regiones terminales de precursores de miR172a que se procesa desde la base al loop no presentaron un tamaño uniforme (Chorostecki, Moro et al., enviado a publicación). Esta contraposición, sugiere que el requisito de una región terminal homogénea es exclusivo de los precursores procesados del loop a la base, de los cuales miR171b fue usado como modelo.

Un análisis global de las características estructurales de los precursores de Arabidopsis que son procesados desde el loop a la base con dos cortes, mostró que estos presentan una región terminal de ~42nt (Bologna et al., 2013b). En la mayoría de los casos dicha región terminal está compuesta un tallo de ARNdH de ~16-17pb y un loop terminal pequeño (R26.BC)(ANEXO A1-A3, A6-A7, A20 y A345). El precursor de miR156a tiene un tallo superior de ~17pb y un loop terminal de 6nt. Experimentos realizados con este precursor mostraron que mutaciones que alteren la estructura secundaria del tallo superior, comprometen la biogénesis del miARN (Kim et al., 2016).

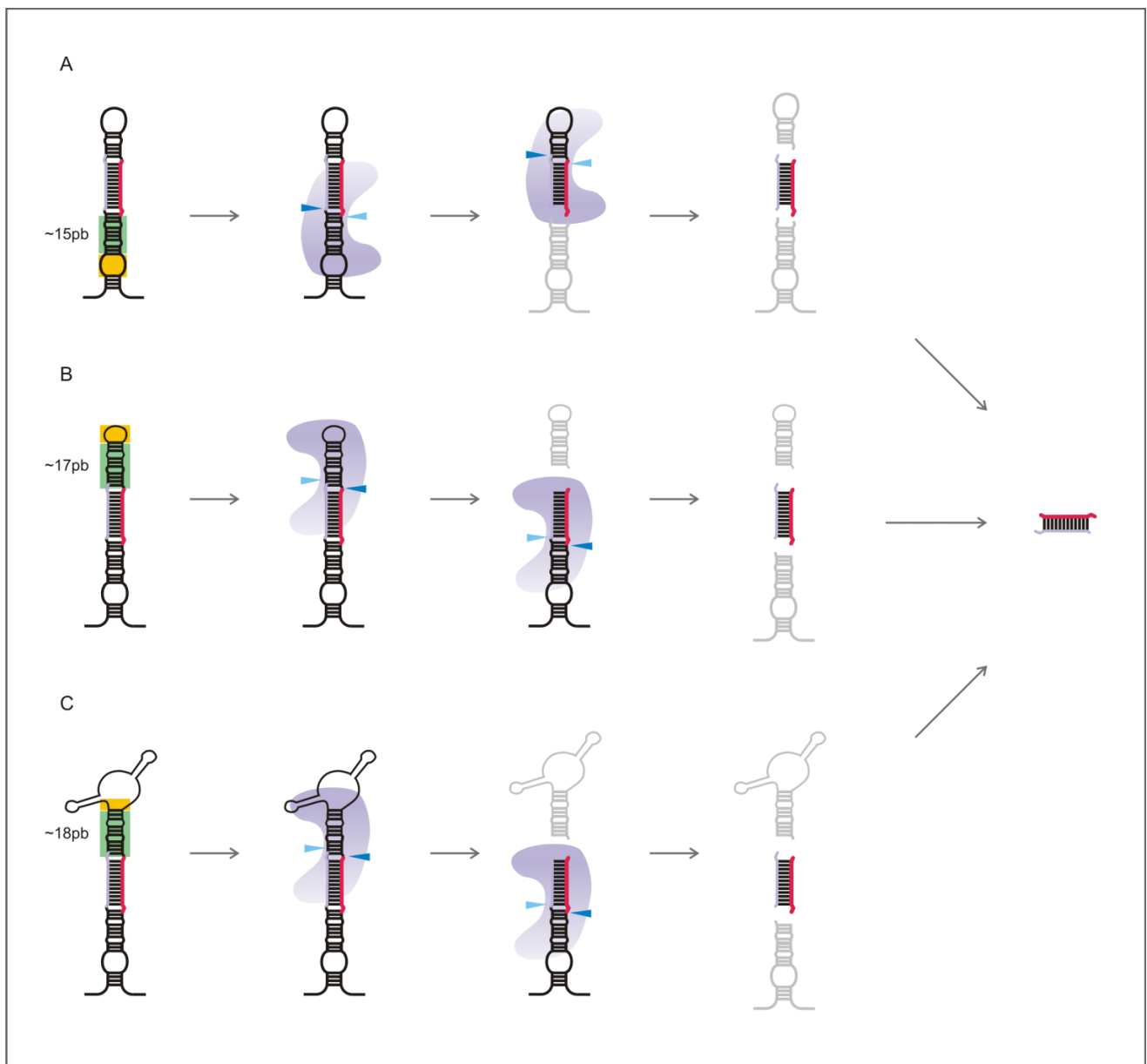
Las mutaciones incluyen cerrar loops internos desapareados, lo que equivale a aumentar el número de bases apareadas, o generar regiones simple hebra al romper apareamientos de bases. Estas modificaciones redujeron cuantitativamente la acumulación de miR156, lo cual puede deberse a que la maquinaria de procesamiento sería capaz de tolerar pequeños cambios estructurales en los dúplex de ARN (Meyers et al., 2010; Werner et al., 2010). En contraposición con esto, modificaciones como producidas en las mutantes *miR171b-supertallo* y *miR171b-superloop* de miR171b, impidieron la acumulación del miARN.

Por último cabe destacar que existe más de una combinación del número de nucleótidos en el loop y pares de bases en el tallo superior que da como resultado una región terminal homogénea de ~42nt (R26.BC)(ANEXO A1-A3, A6-A7, A20 y A345). Esto sugeriría que si bien el primer corte sería posicionado midiendo la distancia en pares de bases desde el loop terminal, la estructura secundaria de la región

terminal, podría jugar un papel en el ajuste fino de la precisión o eficiencia de DCL1 al momento de realizar el primer corte.

7.6.2. Comparación entre los mecanismos de procesamiento de loop a base y base a loop de dos cortes.

Los resultados obtenidos en la Figura R2.6BC mostraron que existe una transición de ARNsh-dh a ~15-17pb del primer corte tanto para los precursores procesados desde el loop como desde la base. Esto nos llevó a pensar de que el loop terminal en los precursores procesados desde arriba y el loop interno en los precursores procesados desde abajo podrán ser análogos para el reconocimiento por la maquinaria de procesamiento (Figura D2.1AB).



D2.1 Los precursores procesados con dos cortes de DCL1 tienen determinantes estructurales análogos.

- A)** Representación esquemática de un precursor procesado desde la base por el reconocimiento de un loop interno seguido de un tallo de ARNdh de ~15pb por debajo del dúplex del miARN.
- B)** Representación esquemática de un precursor procesado desde el loop terminal por el reconocimiento de dicho loop terminal seguido de un tallo de ARNdh de ~17pb por encima del dúplex de miARN.
- C)** Representación esquemática del precursor de miR157c, el cual es procesado desde el loop terminal ramificado, probablemente por el reconocimiento de la transición de ARNsh-dh entre la región próxima al loop y el tallo de ARN dh de ~18pb por encima del miARN.

Si bien la presencia de un segmento de ARNdh de ~15-17pb por debajo o por encima del dúplex del miARN parece ser un rasgo común para el procesamiento desde el loop o desde la base (Figura R2.6CD), el tamaño del loop terminal y el loop interno resultó ser más variable. Para los dos mecanismos de procesamiento pudimos encontrar dos tipos de loops. Los loops terminales definidos se encontraron en 14 de los 15 precursores procesados desde arriba y en 27 de los 40 precursores procesados desde abajo, es decir que son el tipo predominante. Mientras que de los precursores procesados desde el loop hacia la base, sólo miR157c posee un loop ramificado de gran tamaño y 13 precursores procesados desde la base poseen un loop interno ramificado antes del tallo inferior, o al menos no tienen un loop interno claramente definido (Figura R2.6B). Versiones mutantes de miR172a donde se ha reemplazado el loop interno definido por un loop interno ramificado son procesadas *in vivo*, sin embargo, acumulan una menor cantidad de miARN maduro (Werner et al., 2010). Este puede ser un motivo por el cual los loops internos ramificados sean menos frecuentes como iniciador de la biogénesis en los precursores procesados desde la base al loop en Arabidopsis.

A partir de lo expuesto hasta ahora, proponemos un modelo unificado para la biogénesis de los precursores procesados por DCL1 con dos cortes. En aquellos precursores procesados desde la base al loop, DCL1 reconocería la presencia del loop interno inmediatamente por debajo del segmento de ARNdh y realizaría el primer corte a ~15pb (Figura D2.1A). Trabajos anteriores mostraron que eliminar el loop interno de miR172a disminuye cuantitativamente la acumulación del miARN *in vivo*, mientras que eliminar el loop interno del precursor de miR166c anula el primer corte *in vitro* (Mateos et al., 2010; Zhu et al., 2013). Esto pone de manifiesto la importancia de los loop internos en la biogénesis de estos miARNs

En el caso de los precursores procesados desde el loop a la base, la maquinaria de procesamiento reconocería el loop terminal y a partir de allí realizaría un corte a ~17pb (Figura D2.1B). Como mencionamos anteriormente, esta distancia sería de ~15pb en el caso de contar desde el extremo 3' del dúplex de miARN. En concordancia con este modelo, modificar tanto el tallo superior como el loop terminal, alterarían el reconocimiento de DCL1 e impedirían que el primer corte se produjera en el sitio adecuado (Figura R2.2)(Kim et al., 2016). Para los precursores de miR167a y miR172a se ha visto que los determinantes estructurales necesarios para realizar el primer corte y el segundo son independientes (Mateos et al., 2010; Song et al., 2010). Por lo que una vez generado los extremos simple hebra, ya sea a través del reconocimiento del loop terminal o un loop interno, en ambos mecanismos de procesamiento, el

segundo corte se realizaría contando ~21pb para liberar el miARN maduro. Esta división de tareas entre los determinantes estructurales que dan lugar al primer y segundo corte, es comparable al procesamiento de miARNs en animales (Ha and Kim, 2014).

7.6.3. Biogénesis a partir de un loop terminal ramificado

A partir de los datos genómicos se determinó que el precursor de miR157c es procesado desde el loop hacia la base con dos cortes de DCL1 (Figura R2.3A). Sin embargo este precursor se destaca de los mencionados en la sección anterior, por su extensa región terminal, la cual se debe a un loop terminal ramificado de 85nt. En el capítulo de RESULTADOS I los loops terminales ramificados se presentaron como estructuras que desencadenaban cortes abortivos a ~17pb. Sin embargo en este caso, el loop terminal ramificado se encuentra ~18pb del dúplex de miARN y dirige la biogénesis del mismo.

Mutagénesis dirigida sobre el precursor de miR157c mostró que el tallo inferior del precursor es prescindible para la biogénesis del mismo (Figura R2.5AB). Por otro lado, remover el loop terminal ramificado, mejoró la eficiencia de procesamiento para el precursor *miR157c-loop-156b* (Figura R2.5EF). A su vez, este resultado sugiere que la posición del primer corte en el precursor de miR157c se establece por un mecanismo diferente a miR171b. En este caso la región terminal sufrió una delación de 81pb y la posición del primer corte no parece haberse modificado. Sería interesante mapear la posición del primer corte en esta mutante, mediante el método 5'RACE modificado, para corroborar si la posición del primer corte se ha modificado y si la reducción el tamaño del loop terminal fue acompañada por una disminución el número de cortes imprecisos para el primer corte de DCL1.

Se ha propuesto que en los loops terminales ramificados la región de transición de ARNsh-dh es la que atrae a la maquinaria de procesamiento (Zhu et al., 2013). Este reclutamiento es similar al que ocurre en animales cuando DROSHA reconoce la unión basa en el pri-miARN y realiza el primer corte (Ha and Kim, 2014). Por lo tanto, en el precursor de miR157c DCL1 reconocería la transición ARNsh-dh presente entre el loop terminal ramificado y el tallo superior para realizar el primer corte por encima del dúplex del miARN (Figura D2.1C). Aún no es claro como DCL1 se posiciona exactamente en la transición de ARNsh-dh, de los loops ramificados, pudiendo ser este el motivo por el cual el primer corte de miR157c presenta el mayor porcentaje de impresión de todos los precursores procesados desde el loop terminal.

Estudios *in vitro* han demostrado que el dominio helicasa de DCL1 es necesario para realizar el corte abortivos desde del loop terminal ramificado del precursor de miR166c (ANEXO A10). Sin embargo este mismo dominio resulta dispensable para realizar el primer corte en el precursor de miR166c desde el loop interno definido que se encuentra en la base del precursor (Zhu et al., 2013). En particular se ha propuesto que la función del dominio helicasa sería la desenrollar la gran cantidad de ARNsh presente el loop terminal ramificado y exponer la transición de ARNsh-dh necesaria para que DCL1 para realice el

primer corte (Liu et al., 2012). Esta podría ser una de las bases que expliquen las diferencias encontradas entre modificar las regiones terminales en los precursores de miR157c y miR171b.

Tanto el precursor de miR157c como el precursor de miR171b son procesados desde el loop terminal hacia la base con dos cortes de DCL1. Sin embargo los determinantes estructurales reconocidos en cada caso son diferentes. Esto plantea el interrogante particular de si comparten el mismo mecanismo de procesamiento, y una pregunta más general de qué es lo que define un mecanismo de procesamiento. Con la información experimental disponible hasta el momento un mecanismo de procesamiento era definido en función de dos parámetros, la dirección del mismo, de loop a base o de base a loop y el número de cortes requeridos para liberar el miARN maduro, dos o más. Sin embargo en el contexto de los resultados presentados para el precursor de miR157c proponemos ampliar esta definición teniendo en cuenta además los determinantes estructurales responsables de la dirección de procesamiento. En este sentido, el precursor de miR157c podría ser denominado con un precursor procesado desde el loop terminal hacia la base con dos cortes “no canónico”. Un caso similar se encuentra en el precursor del miR166b, el cual tiene un único corte detectado por SPARE en el extremo inferior del dúplex del miARN (Anexo A10). En principio esto define su mecanismo de biogénesis como de base a loop con dos cortes. Sin embargo una versión mutante del precursor de miR166b donde se ha eliminado el tallo inferior de ARNdh por completo, es capaz de acumular miARN en cantidades comparables al precursor silvestre. Esto quiere decir que el determinante estructural está por fuera del tallo inferior y el loop interno. Según la nueva definición propuesta, el mecanismo de procesamiento de dicho precursor sería de base a loop con dos cortes de DCL1 “no canónico” (Chorostecki, Moro et al., enviado a publicación).

Sépalo de una vez, lo engañamos, aquí todo es falso -La muerte.
Mi mentira de amor vale más que ese horror que usted llama verdad -El enamorado.
"El enamorado y la muerte"-Opereta de Alejandro Dolina.

8. Resultados III

“Relación entre la estructura del precursor de miARNs y la actividad procesiva de DICER-LIKE1”

8.1. INTRODUCCIÓN III

Como se mencionó en la introducción existe una gran diversidad de ARNs pequeños en plantas. La regulación génica a través de PTGS depende de la acción de cuatro grupos de ARNs pequeños, los miARNs, siARN, los nat-siARNs y los tasiARNs. Los siARN, los nat-siARNs y los tasiARNs requieren de la actividad de corte secuencial de DCL2, DCL3 y DCL4, para liberar un conjunto de ARNs pequeños de entre ~21-24nt. Mientras que la mayoría de los miARNs son procesados por DCL1 mediante sólo dos cortes (Bologna and Voinnet, 2014).

Sin embargo de los 75 precursores de miARN conservados presentes en Arabidopsis cuyo mecanismo de procesamiento se ha establecido, 18 de ellos se procesan de modo secuencial, y requieren al menos tres cortes de DCL1 (Bologna et al., 2013b; Kurihara and Watanabe, 2004)(Figura R1.8). Estos precursores liberan dos tipos de dúplex de ARN: el dúplex del miARN/miARN* y un dúplex compuesto por los otros ARNs pequeños derivados del precursor. Si bien estos ARNs pequeños se escinden en la misma proporción que el miARN, en general los mismos son detectados con mucha menor frecuencia *in vivo* que al miARN maduro (Bologna et al., 2009; Zhang et al., 2010).

Las bases moleculares responsables de esta asimetría en la acumulación de ARNs pequeños derivados de un mismo precursor se desconocen. A su vez esta situación no es exclusiva de los miARNs. A pesar que los TAS1a, TAS1b, TAS1c y TAS2 son cortados entre seis y ocho veces en fase por DCL4 los siARNs predominantes son siR1511 y siR255, detectándose por Northern blot para ARNs pequeños. Otros siARNs derivados de estos TAS comienzan con 5´U y tienen 21nt, al igual que siR1511 y siR255, sin embargo no se acumulan (Allen et al., 2005; Yoshikawa et al., 2005). Situaciones similares se han encontrados para los ARNs pequeños derivados de TAS3 y TAS4 (Allen et al., 2005; Rajagopalan et al., 2006).

Mientras que DCL2, DCL3 y DCL4 tienen preferencias por largos ARNdhs de complementariedad casi perfecta, DCL1 procesa sustratos de ARN plegados de manera imperfecta (Bologna and Voinnet, 2014). Esto hace que sea importante estudiar los requisitos estructurales presentes en los precursores de miARN que dan lugar al procesamiento secuencial, ya que podrían contribuir a entender las diferencias en el modo de acción entre las diferentes DCLs. A su vez, cómo se mencionó en el párrafo anterior, no está claro cuáles

son las características estructurales que determinan que un miARN se acumule por encima de otro ARN pequeño.

Para responder a estos interrogantes, empleamos como modelo a la familia de miR319. Este *MIARN* está conservado evolutivamente e incluso está presente en musgos (Axtell and Bartel, 2005; Talmor-Neiman et al., 2006). El mismo presenta un mecanismo de procesamiento secuencial desde el loop terminal hacia la base, que libera tres dúplex de ARNs pequeños. Dicho mecanismo se encuentra conservado entre *Arabidopsis* y el musgo *Physcomitrella patens* (Addo-Quaye et al., 2009; Bologna et al., 2009). Finalmente miR319, regula a un grupo de factores de transcripción denominados TCP en angiospermas, los cuales controlan el desarrollo de la planta, en particular el crecimiento y morfogénesis de la hoja. Como consecuencia de esto, la sobre-expresión de dicho miARN produce defectos fenotípicos fácilmente observables (Palatnik et al., 2003).

8.2. Conservación de miR319 y el origen de sus ARNs pequeños.

El precursor de miR319 es procesado mediante cuatro cortes secuenciales, los cuales pueden originar potencialmente seis ARNs pequeños. El miARN maduro es liberado entre el tercer y cuarto corte. (Addo-Quaye et al., 2009; Axtell and Bartel, 2005; Bologna et al., 2009)(Figura R3.1A). Siendo miR319 un miARN muy antiguo, decidimos iniciar el análisis del mismo focalizándonos en la conservación evolutiva y su relación con los ARNs pequeños derivados del mismo.

Con este fin, obtuvimos las secuencias de 41 precursores del miARN miR319 de 13 especies diferentes de dicotiledóneas, disponibles en miRBase (Version 21, <http://www.mirbase.org/>), cuyo genoma se encontraba disponible en Phytozome (<https://phytozome.jgi.doe.gov/pz/>). Considerando que las secuencias que se encuentran por debajo del dúplex miARN/miARN* son dispensables para la biogénesis de miR319, se tomaron sólo bases contenidas entre el miARN y el miARN*, incluyéndolos (Bologna et al., 2009). Estas secuencias se ingresaron en el programa M-coffee (<http://tcoffee.crg.cat/apps/tcoffee/do:mcoffee>) el cual alinea diferentes biomoléculas combinando varios algoritmos de alineamientos, obteniéndose como resultado un alineamiento “promedio”.

A partir del alineamiento se pudieron observar cuatro islas de conservación (Figura R3.1B). *Arabidopsis* tiene tres genes que codifican para miR319, *MIR319a-c*. A continuación se hizo una búsqueda de secuencias asociadas a estos locus en bibliotecas de ARNs pequeños disponibles públicamente en miRBase (Tabla R3.1). Al ubicar las secuencias en el contexto del alineamiento, se vio que los ARNs pequeños se originaban de las cuatro regiones conservadas. Sin embargo miR319 era el más abundante de todos los ARN pequeños detectados en las bibliotecas, lo cual tiene sentido considerando su función biológica como regulador de los factores de transcripción TCP (Palatnik et al., 2003). Mientras que el

segundo ARN en abundancia correspondía a miR319-3p.2, el cual es liberado entre el primer y segundo corte de DCL1.

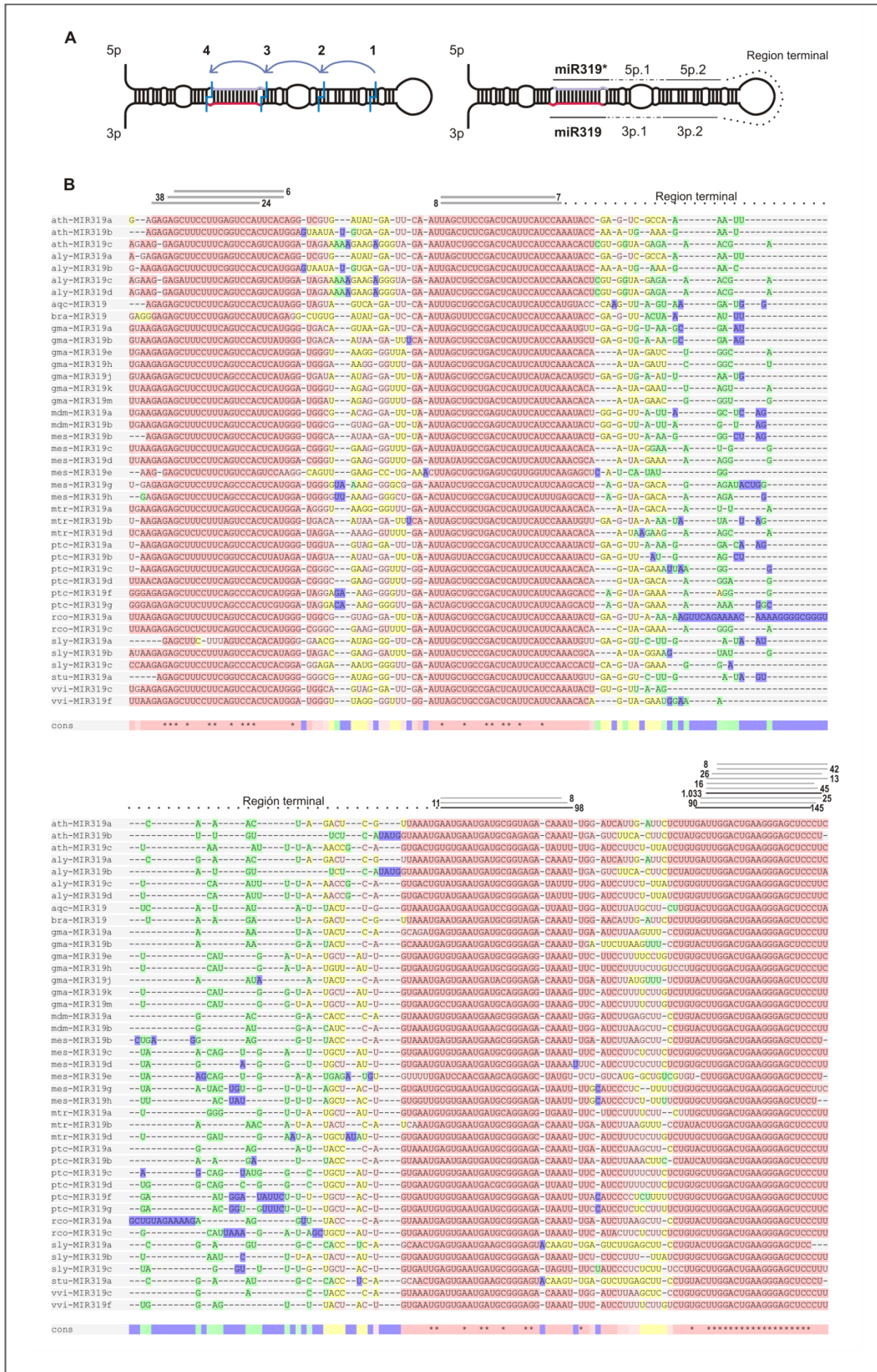


Figura R3.1. ARNs pequeños derivados del precursor de MIR319 y su conservación en secuencia.

A) Esquema representativo del procesamiento secuencial de loop a base de miR319. El precursor es procesado con cuatro cortes de DCL1 (panel izquierdo), que pueden generar 6 ARNs pequeños (panel derecho). En estos ARN pequeños se puede detectar a su vez cierta heterogeneidad en sus secuencias producto del corte inexacto por DCL1.

B) Alineamiento de los precursores de miR319 obtenido con el programa M-coffee a partir de las secuencias de precursores de miR319 de diferentes especies vegetales. Los ARN pequeños más frecuentes originados de los precursores de miR319a-c de Arabidopsis se muestran como una línea. Los números al lado de las líneas corresponden al número de lecturas de cada ARN pequeño (Tabla 3.1).

Tabla 3.1. ARNs pequeños derivados de los precursores de miR319a, miR319b and miR319c de

Arabidopsis. Las secuencias únicas se muestran en celeste.

ath-MIR319a		ath-MIR319b		ath-MIR319c	
Secuencia	lecturas presenta	Secuencia	lecturas presenta	Secuencia	lecturas presenta
GAGAGCTTCCTTGAGTCCAT	24 RNA-3p	AGAGCTTCTTCGGTCCACTC	50 miR-5p	GGAGATTCTTTCAGTCCAGTC	20 miR-5p
AGCTTCTCTTGAGTCCATTCAC	38 RNA-3p	TGACTCTCGACTCATTTCATCCA	7 RNA-5p.2	TGTGAATGATGCGGGGATAT	44 RNA-3p.2
GCTTCTCTTGAGTCCATTCAC	6 RNA-3p	GACTCTCGACTCATTTCATCC	6 RNA-5p.2	TTGGACTGAAGGGAGCTCCT	75 miR-3p
TAGCTTCCGACTCATTTCATCCA	8 RNA-3p.2	AATGAATGATGCGAGAGACAA	296 RNA-3p.2	TTGGACTGAAGGGAGCTCCT	65 miR-3p
TAGCTTCCGACTCATTTCATCC	7 RNA-5p.2	AATGAATGATGCGAGAGACAAAT	229 RNA-3p.2	TTGGACTGAAGGGAGCTCCT	16 miR-3p
AATGAATGATGCGGTAGACAAAT	98 RNA-3p.2	AATGAATGATGCGAGAGACA	47 RNA-3p.2	TTGGACTGAAGGGAGCTCCT	10 miR-3p
AATGAATGATGCGGTAGACAAA	11 RNA-3p.2	AATGAATGATGCGAGAGAC	32 RNA-3p.2		
AATGAATGATGCGGTAGACAA	8 RNA-3p.2	AATGAATGATGCGAGAGACAAA	29 RNA-3p.2		
GATTGGACTGAAGGGAGCTCC	145 miR-3p	AATGAATGATGCGAGAGACAAAT	16 RNA-3p.2		
ATTGGACTGAAGGGAGCTCCC	90 miR-3p	ATGAATGATGCGAGAGACAA	27 RNA-3p.2		
ATTGGACTGAAGGGAGCTCCT	25 miR-3p	ATGAATGATGCGAGAGACAAA	19 RNA-3p.2		
TTGGACTGAAGGGAGCTCCCT	1033 miR-3p	CTTGGACTGAAGGGAGCTCCC	23 miR-3p		
TTGGACTGAAGGGAGCTCCC	45 miR-3p	CTTGGACTGAAGGGAGCTCCT	12 miR-3p		
TTGGACTGAAGGGAGCTCCC	16 miR-3p	TTGGACTGAAGGGAGCTCCCT	1033 miR-3p		
TTGGACTGAAGGGAGCTCCT	13 miR-3p	TTGGACTGAAGGGAGCTCCC	45 miR-3p		
TGGACTGAAGGGAGCTCCCT	26 miR-3p	TTGGACTGAAGGGAGCTCCT	16 miR-3p		
GGACTGAAGGGAGCTCCCT	42 miR-3p	TGGACTGAAGGGAGCTCCCT	26 miR-3p		
GACTGAAGGGAGCTCCCT	8 miR-3p	GGACTGAAGGGAGCTCCCT	42 miR-3p		

8.3. Inserciones en el tallo superior de miR319a alteran el patrón de cortes de DCL1 en el precursor.

Si bien el precursor de miR319 requiere de cuatro cortes para liberar el miARN maduro, pueden distinguirse dos tipos de cortes diferentes. El primer corte es un endocorte, en el que se generan los extremos simple hebra que DCL1 reconoce para realizar el segundo corte a ~21nt, generando nuevos extremos simple hebra y así sucesivamente. En este sentido los cortes dos, tres y cuatro serian equivalentes ya que todos se producen a partir del mismo determinante estructural.

A pesar de que Arabidopsis presente tres genes que codifican para miR319, *MIR319a*, *MIR319b* y *MIR319c*, el precursor de miR319a es procesado con mayor eficiencia y acumula mayor cantidad de miARN maduro al expresarse desde un promotor 35S (Desarrollado en la Tesis del Dr. Nicolás Bologna). En función de esto decidimos trabajar sobre el precursor de miR319a.

Para evaluar la hipótesis mencionada en el primer párrafo diseñamos dos mutantes sobre el precursor de miR319a donde se agregaron cinco nucleótidos en cada hebra del precursor, a modo de generar un ARNdH con un mismatch en el centro. Las mutantes fueron denominadas: *miR319a-INS25* e *miR319a-INS41*, haciendo referencia a la posición del precursor donde se realizó la inserción, a 25 y 41 nucleótidos

corriente abajo del loop terminal. Estas posiciones se eligieron para que la inserción en *miR319a-INS25* estuviera antes del segundo corte y en *miR319a-INS41* antes del tercer corte. No se diseñaron mutantes entre el tercer y cuarto corte ya que esto alteraría la secuencia del miARN maduro (Figura R3.2A).

Se generaron plantas transgénicas que sobre-expresaran tanto los precursores mutantes como el precursor silvestre de miR319a. La sobre-expresión de miR319 genera cambios en la morfología de la hoja, particularmente se producen pliegues en los márgenes de la misma. Alrededor del 70-75% de las plantas transgénicas que sobre-expresan el precursor silvestre desde el promotor 35S poseen importantes alteraciones en la curvatura de la hoja (clasificado como “fuerte: ++”), mientras que el 25-30% restante, muestran defectos más suaves (fenotipo clasificado “débil: +”)(Bologna et al., 2009)(Figura R2.3D). En todos los casos, la intensidad de los defectos en el desarrollo, se correlaciona con los niveles de miR319 (Bologna et al., 2009; Palatnik et al., 2003; Schommer et al., 2008).

Se evaluaron los fenotipos de 50 T1 independientes que sobre-expresaran los constructos de *miR319a-INS25* e *miR319a-INS41*. Todas las plantas presentaron el fenotipo silvestre de las plantas transformadas con el vector vacío (Figura R3.2 C), mientras que la sobre-expresión del precursor de miR319 provoco cambios fuertes en el ~70% de las plantas, lo cual es consistente con experimentos anteriores. Esto nos llevó a pensar que los precursores de *miR319a-INS25* y *miR319a-INS41* no estarían acumulando miR319.

Se recolectaron 25 inflorescencias por duplicado, tanto de las líneas control, como mutantes. Se realizó la extracción de ARN y un blot para ARN pequeños contra miR319a. Cabe destacar que la sonda contra miR319, también detecta al miARN miR159, debido a su elevada similitud en secuencia, y a la abundancia de miR159 endógeno. A pesar de esto, ambos miARNs son fácilmente diferenciables por sus distintas movilidades electroforéticas (Palatnik et al., 2007).

Los blots para ARNs pequeños mostraron que ni la mutante *miR319a-INS25*, ni *miR319a-INS41* acumulaban miR319, en concordancia con los fenotipos obtenidos (Figura R3.2B). Para este resultado existen dos posibles explicaciones, la primera es que las inserciones impiden el reconocimiento del precursor por la maquinaria de procesamiento y el mismo no es procesado. La segunda, es que las inserciones hayan producido alteraciones en el patrón de cortes de DCL1, los cuales dan lugar a un ARN pequeño cuya secuencia es diferente de miR319.

Para responder este interrogante decidimos mapear los intermediarios de procesamiento de *miR319a-INS25*, empujando una versión modificada del protocolo de 5'RACE. Brevemente, el método consiste en una ligación entre un oligo adaptador de ARN y el extremo 5' del ARN cortado. Luego, se realiza una reacción de transcripción reversa y finalmente una amplificación por PCR usando un oligo específico para el precursor y un oligo que hibrida contra el adaptador. Este paso permite amplificar específicamente los intermediarios del precursor deseado y evitar la interferencia de los intermediarios endógenos del mismo.

Los productos obtenidos son clonados y secuenciados lo cual permite conocer la secuencia de los intermediarios en el procesamiento de cada precursor (Sección 5.5. Materiales y Métodos).

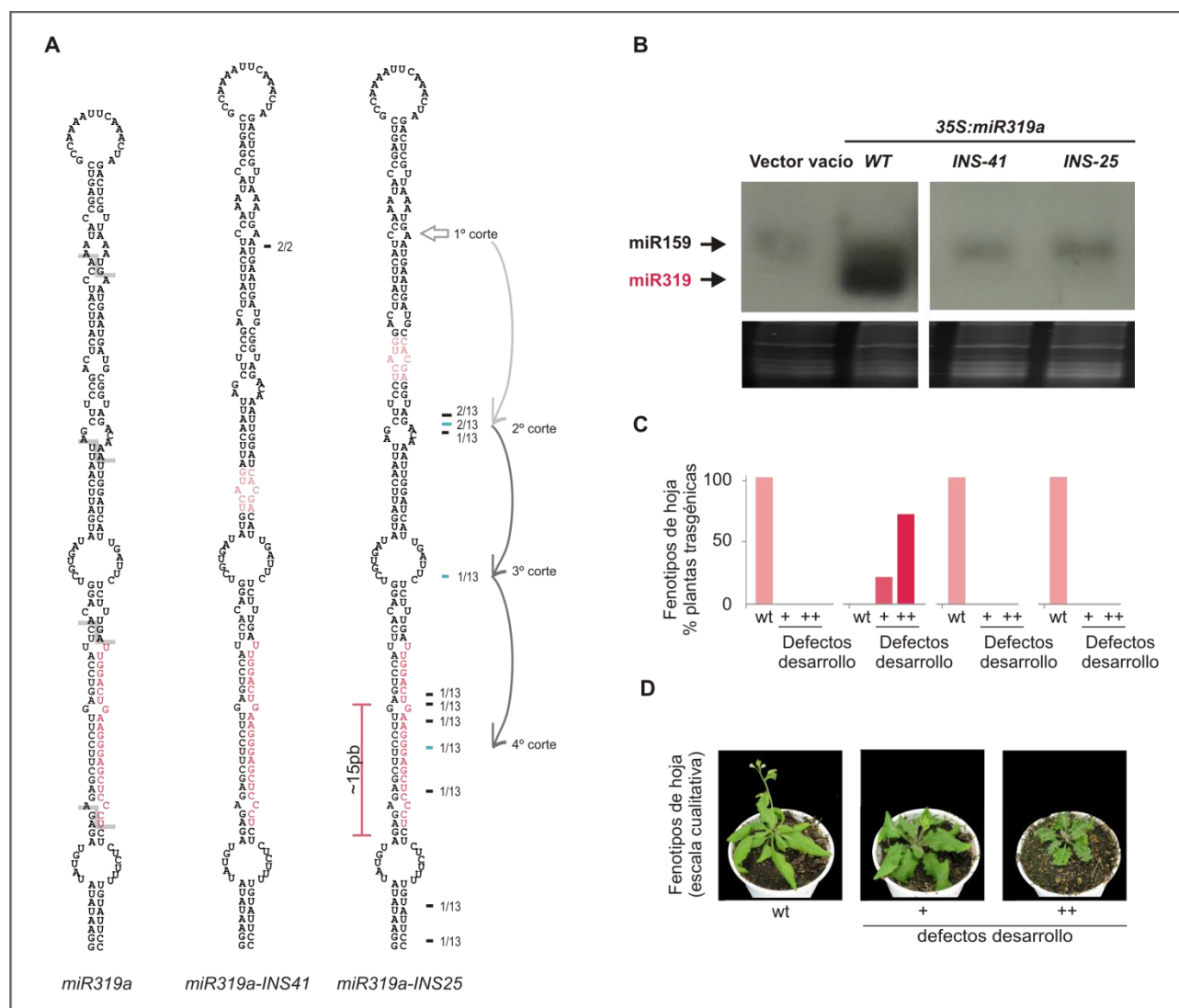


Figura R3.2. Inserciones en el tallo superior de miR319a alteran la posición de los cortes de DCL1.

A) Predicción de estructura secundaria para precursor silvestre de miR319a, miR319a-INS41 y miR319a-INS25. Se muestran a su vez, los cortes en el precursor obtenidos por 5'RACE, y el patrón de procesamiento reconstruido para miR319a-INS25 a la derecha. Las bases insertadas se muestran en rosa y el miARN en rojo.

B) Blot para ARNs pequeños contra miR319 realizado a partir de plantas sobre-expresantes de miR319a, miR319a-INS41 y miR319a-INS25. Al menos 25 inflorescencias independientes constituyen una muestra.

C) Distribución de los defectos del desarrollo en las hojas de plantas transgénicas sobre-expresantes de cada vector. Al menos 50 plantas fueron evaluadas para cada constructo.

D) Plantas representativas que presenta en fenotipo silvestre (wt), el fenotipo débil (+) y el fenotipo fuerte (++) debido a la sobre-expresión de miR319.

A partir de plantas sobre-expresantes de *miR319a-INS25* se construyeron bibliotecas de los intermediarios de procesamiento. Con la información obtenida de las secuenciaciones se ubicó en la estructura del precursor la posición de los cortes (Figura R3.2A, líneas negras). Para *miR319a-INS25*, se secuenciaron 13 clones que correspondieron a nueve intermediarios de procesamiento.

Cinco de los 13 cortes se encuentran en una región aproximadamente cinco nucleótidos corriente arriba del lugar donde ocurre segundo corte en el precursor silvestre de *mir319a* (Figura R3.2A). A partir de esta región, se pueden encontrar cortes a ~20nt y ~40nt corriente abajo. Esta distancia es consistente con los cortes secuenciales de DCL1. Como la posición del primer corte no pudo ser detectada para *mir319a-INS25*, alternativamente, mapeamos la misma *mir319a-INS41*, cuyas mutaciones son análogas a *mir319a-INS25*. Para *mir319a-INS41* se secuenciaron dos clones, que posicionaron el primer corte dos nucleótidos por debajo de la posición esperada para el primer corte en el precursor silvestre, pero dentro de la variabilidad asociada a los cortes de DCL1 (Jeong and Green, 2012; Lu et al., 2010; Nakano et al., 2006). Esto nos lleva a pensar que en el precursor de *mir319a-INS25* la posición del primer corte tampoco ha sido modificada.

El agregado de cinco nucleótidos luego del primer corte, no impide que el precursor de *miR319a* sea reconocido por la maquinaria de procesamiento. Por el contrario el complejo de DCL1 procesaría el precursor, pero se produce un corrimiento concomitante en la posición del corte de DCL1 de cinco nucleótidos, luego del sitio de la inserción. Por lo tanto estos resultados confirmarían la hipótesis de que después del primer corte del precursor, los cortes subsiguientes se realizan midiendo una distancia de ~21nt desde el extremo libre. Resultados similares se obtuvieron al extender el largo del tallo inferior de *miR172a*, lo cual modificó en la posición de corte de DCL1, resultando en un miARN no funcional (Werner et al., 2010). Esto sugiere que la maquinaria de procesamiento es capaz de contar el largo del tallo inferior o superior, en un modo que resulta análogo al complejo microprocesador de animales (Ha and Kim, 2014; Han et al., 2006).

Los cortes encontrados en el precursor permitirían explicar los resultados obtenidos en el blot de ARN pequeños y los fenotipos de las plantas (Figura R3.2ABC). El dúplex liberado por el tercer y cuarto corte, no posee la secuencia del miARN *miR319*, sino que esta secuencia está corrida al moverse el marco de los cortes por DCL1. Lo cual impide que el mismo sea capaz de regular a los genes blanco *TCP*, motivo por el cual no se obtiene el fenotipo esperado para la sobre-expresión de *mir319a-INS25* y no se detecta el *miR319* en el blot de ARNs pequeños.

Además de los cortes mencionados hasta ahora, también se encontraron cortes en la región central del miARN/miARN*. Este tipo de cortes abortivos, ya han sido descritos previamente en versiones mutantes del precursor de *miR319* donde se ha eliminado el loop o partes del tallo superior (Bologna et al., 2013b). La hipótesis es que modificaciones que afecten la estructura secundaria del tallo superior, interfieren con el procesamiento loop a base y el precursor comienza a ser reconocido desde la base. *miR319a* presenta un loop interno inferior, localizada a aproximadamente cinco nucleótidos del dúplex miARN/miARN* (Figura R3.2A, línea roja). Este loop interno sería reconocido por la maquinaria de procesamiento que produciría los cortes en el centro del dúplex de *miR319a*, en un proceso análogo al descrito en la Sección 6.5, donde

mediante el reconocimiento de determinantes estructurales alternativos, se producen cortes que disminuyen la cantidad de miARN maduro.

8.4. El loop interno presente en la región central de miR319a previene la acumulación de otros ARN pequeños.

A partir del alineamiento de miR319 y la ubicación de los ARNs pequeños detectados por secuenciación (Figura R3.1B), notamos que no había lecturas asociadas a los ARNs pequeños derivados del segundo y tercer corte, ya sea miR319-3p.1 o miR319-5p.1. A pesar de que esta región no está conservada, no existe una razón obvia por la cual no se acumulen ARNs pequeños asociados a la misma en Arabidopsis. De hecho se han encontrado ARNs pequeños provenientes de la región media del precursor miR319 del musgo *Physcomitrella patens* (Axtell et al., 2007; Talmor-Neiman et al., 2006).

A continuación nos preguntamos cuáles serían los motivos que impedirían la acumulación de miR319.3p.1 o miR319.5p.1. Un análisis de la estructura secundaria en la región del dúplex de miR319a.1 del precursor de miR319a, reveló que el mismo contiene un loop interno de ~12nt. En Arabidopsis, los dúplex de miARNs tienen como máximo cuatro bases desapareadas, por lo que pensamos que seis bases consecutivas sin aparear probablemente desestabilizarían el dúplex de miR319a.1 *in vivo*. En función de esto, generamos una versión mutante de miR319a donde el loop interno estuviera cerrado por completo. Para esto se insertaron cuatro mutaciones puntuales en la secuencia de miR319a-5p.1 (Figura R3.3A, cuadrados naranja) y la mutante se denominó *miR319a-LIC*, por loop interno cerrado.

Tanto la versión silvestre de miR319a como *miR319a-LIC* se transformaron en Arabidopsis. A partir de T1 independientes se colectaron inflorescencias, se extrajo ARN y se hicieron blots para ARNs pequeños contra miR319a-3 p.1 y miR319. En las plantas transformadas con el vector vacío y con el vector que expresa el precursor miR319a no se pudo detectar miR319a-3p.1 en los blots para ARNs pequeños en las condiciones empleadas (Figura R3.3B). Esto es consistente con los datos obtenidos a partir de las bibliotecas de ARNs pequeños, donde el mismo no se había detectado (Figura R3.1B y Tabla R3.1). Por otro lado, en las plantas que sobre expresan el constructo *miR319a-LIC* se observa una banda correspondiente a miR319a-3p.1. Esto nos hizo pensar que generar un dúplex de miR319a.1 con elevada complementariedad, permitiría que el mismo se acumulara y fuera detectable en ensayos de Northern blot.

Cuando se empleó la sonda contra miR319, se detectó que las plantas sobre-expresantes del precursor de miR319a acumulaban el microARN maduro (Figura R3.3C). Sin embargo, las plantas sobre-expresantes de *miR319a-LIC* acumulaba una menor cantidad de miR319. Se evaluaron los fenotipos de al menos 35 transgénicas independientes que sobre-expresaran los distintos constructos, teniendo en cuenta la escala desarrollada en la Sección 8.3. Las plantas que sobre-expresaban el precursor de miR319a tenían un mayor

porcentaje de hojas arrugadas a comparación con las transgénicas de *miR319a-LIC*, en concordancia con los resultados obtenidos en los blots para ARNs pequeños.

Esto sugiere que si bien mejorar el apareamiento del dúplex de miR319a.1 produjo que el mismo se acumulara, a su vez causó una menor acumulación de miR319. El mapeo de los intermediarios de procesamiento para *miR319a-LIC* mostró que no se habían producido cambios significativos en la posición del tercer corte de DCL1 (Figura R3.2B)(Sección 5.5 Materiales y métodos). Sin embargo se encontraron cortes en la región central del miARN/miARN*, como los mencionados en la Sección 8.3 (Figura R3.2A, línea roja).

El procesamiento abortivo desde la base del precursor de miR319a no sido ha detectado en las bibliotecas de SPARE (Bologna et al., 2013b)(ANEXO A23) ni mediante 5'RACE modificada a partir de plantas sobre-expresantes del mismo (Bologna et al., 2009). Lo cual sugiere que esta no es la vía primaria mediante la cual el complejo de DCL1 reconoce y corta al precursor, si no que la misma se vuelve competitiva cuando el procesamiento secuencial de loop a base es entorpecido (Figura R3.2A, R3.3A) (Bologna et al., 2013b). En conjunto estos resultados nos llevaron a hipotetizar que las modificaciones realizadas en la región central del precursor de miR319a podrían enlentecer o detener parcialmente la actividad procesiva de DCL1, lo que daría tiempo al reconocimiento y corte del precursor de miR319 desde la base (Bologna et al., 2013b).

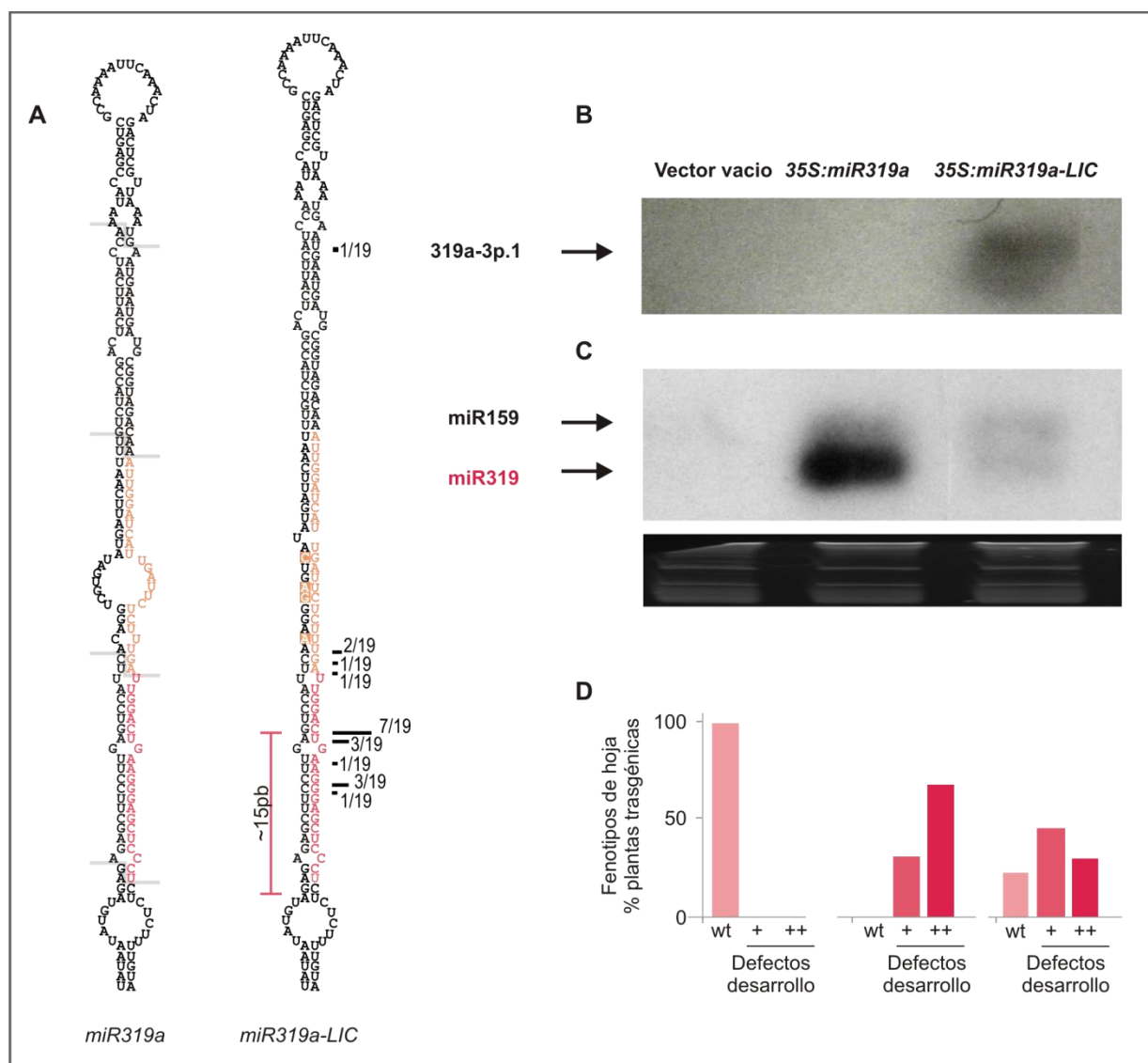


Figura R3.2. Mejorar el apareamiento del dúplex miR319a.1 resulta en su acumulación.

A) Predicción de estructura secundaria para el precursor silvestre de miR319a y miR319a-LIC. Las bases mutadas en miR319a-5p.1 se marcan con un cuadrado naranja. El miARN se muestra en rojo y el miR319a-3p.1 en naranja. Se muestra el patrón de cortes sobre el precursor de miR319a-LIC mapeado por 5'RACE.

B) Blot para ARNs pequeños contra miR319a-3p.1 realizado a partir de plantas sobre-expresantes de miR319a y miR319a-LIC. Al menos 20 inflorescencias independientes constituyen una muestra.

C) Blot para ARNs pequeños contra miR319 realizado a partir de plantas sobre-expresantes de miR319a y miR319a-LIC. Al menos 20 inflorescencias independientes constituyen una muestra.

D) Distribución de los defectos del desarrollo en las hojas de plantas transgénicas sobre-expresantes de cada vector. Al menos 40 plantas fueron evaluadas para cada constructo.

8.5. Mejorar el apareamiento del dúplex de miR319a.2 genera DCL1 se detenga.

A continuación, estudiamos en más detalle el primer segmento del precursor de miR319 que libera un dúplex de ARN pequeños miR319a.2. Este dúplex sería el resultado de los dos primeros cortes por DCL1 y se corresponden con una región de alta conservación evolutiva (Figura R3.1). Para estudiar la importancia de este segmento utilizamos una mutante *miR319a-B4* que tiene un mejor apareamiento (Bologna et al., 2009)(Figura R3.4ABC, *miR319a-B4*). Pudimos reproducir trabajos anteriores donde la mejora de la complementariedad de esta región aumenta la acumulación de ARN pequeños (Bologna et al., 2009)(Figura

R3.4ABC, *miR319a-B4*) *in vivo*, lo cual es consistente con los resultados mostrados en la sección anterior donde mejoramos la interacción en la región intermedia del precursor de miR319 (Figura R3.3).

A continuación realizamos mutagénesis dirigida sobre la región inferior del precursor de miR319a-B4, y extendimos el tallo de ARNdH a una distancia de ~15pb del loop interno (Figura R3.4A, *miR319a-reconocimiento-doble*). En principio, este precursor podría ser cortado por DCL1 comenzando desde la base o desde el loop, y en ambos casos generaría los mismos ARN pequeños.

Generamos plantas transgénicas que sobre-expresen el precursor silvestre de miR319a, *miR319a-B4* y *miR319a-reconocimiento-doble*. A partir de 20 inflorescencias de plantas T1 independiente, se realizó la extracción de ARN, y blots para ARNs pequeños contra 319-3p.2 y miR319. Ni las plantas transformadas con el vector vacío ni con el precursor de miR319a acumularon el miR319a-3p.2. A diferencia de estas y en concordancia con los resultados publicados previamente las plantas que sobre-expresaban *miR319a-B4* acumularon el ARN pequeño 319a-3p.2 (Bologna et al., 2009). A su vez el precursor de *miR319a-reconocimiento-doble* también acumuló miR319a-3p.2. Por otro lado el blot contra miR319 mostró que *miR319a-B4* acumulaba la menor cantidad del miARN de los tres precursores. Mientras que la mutante *miR319a-reconocimiento doble*, acumuló más cantidad de miARN que su contrapartida *miR319a-B4*, pero menos cantidad que el precursor silvestre.

En conjunto estos resultados nos permitieron verificar que el procesamiento del precursor de miR319a desde la base puede iniciarse independiente del reconocimiento del loop terminal. A su vez el mismo es capaz de recuperar, al menos parcialmente, la acumulación del miARN maduro en *miR319a-reconocimiento-doble*. En principio hubiéramos esperado que las cantidades acumuladas de 319a-3p.2 entre *miR319a-B4* y *miR319a-reconocimiento-doble* fueran similares, ya que no se modificó la región superior en el precursor de *miR319a-reconocimiento-doble*. Sin embargo *miR319a-reconocimiento-doble* acumula una menor cantidad de 319a-3p.2, lo cual sugiere que este precursor no puede ser procesado simultáneamente por dos DCL1 en direcciones opuestas. En concordancia con esto, trabajos *in vitro* con el precursor de miR166c mostraron que existe una correlación inversamente proporcional entre la frecuencia de los cortes de biogénesis (desde la base al loop) y los cortes abortivos (desde el loop terminal ramificado hacia la base)(Zhu et al., 2013). Lo cual indicaría que esta restricción podría ser extensiva a otros precursores de miARNs de plantas e incluso ocurrir *in vivo*.

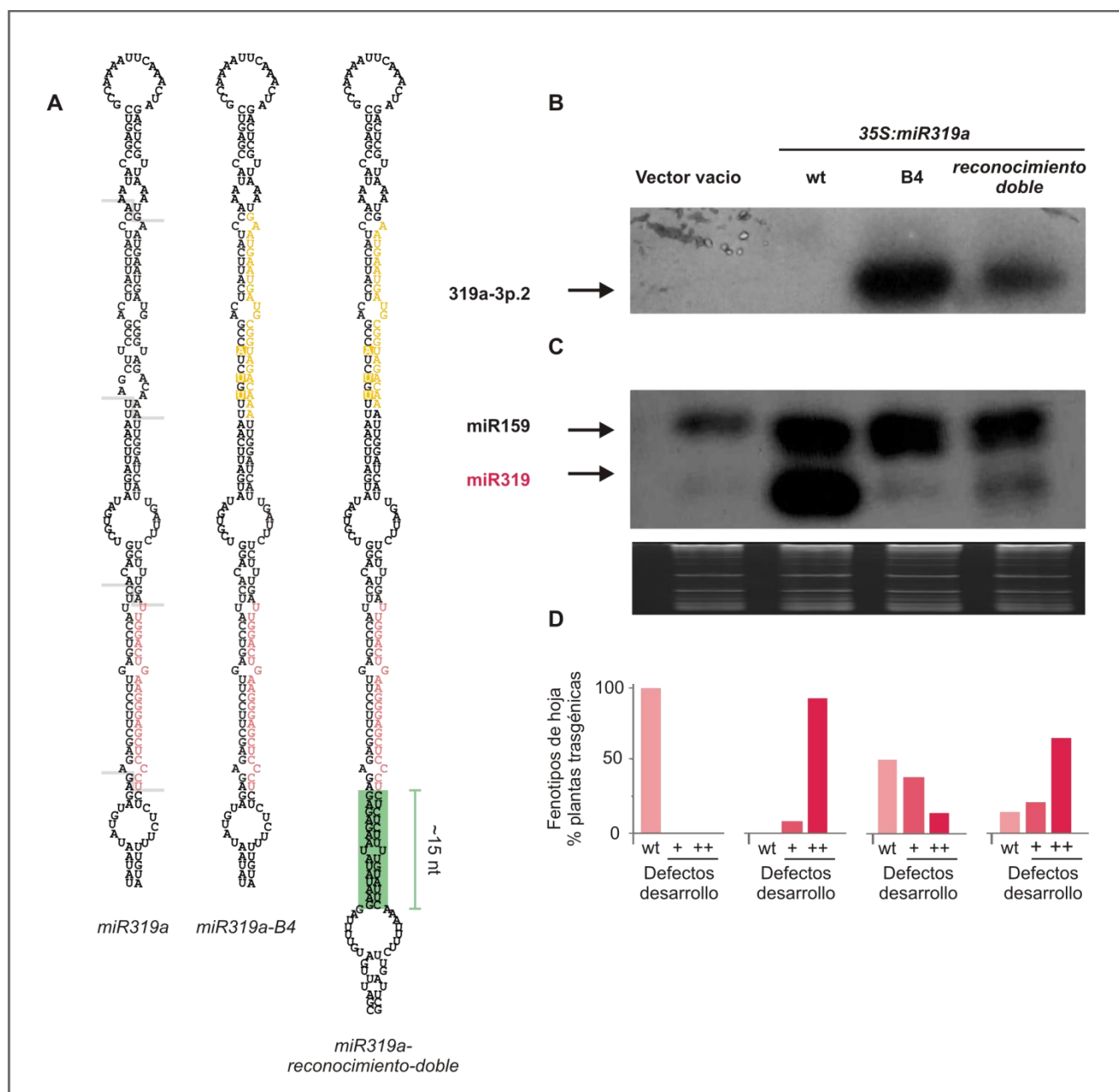


Figura R3.4. Mejorar el apareamiento del dúplex miR319a.1 resulta en su acumulación.

A) Predicción de estructura secundaria para el precursor silvestre de miR319a, miR319a-B4 y miR319a-reconocimiento doble. Las bases mutadas en miR319a-5p.2 se marcan con un cuadrado amarillo y la extensión del tallo inferior de miR319a-reconocimiento-doble se muestra con un cuadrado verde. El miARN se muestra en rojo y el miR319a-3p.2 en amarillo.

B) Blot para ARNs pequeños contra miR319a-3p.2 realizado a partir de plantas sobre-expresantes de miR319a, miR319a-B4 y miR319a-reconocimiento-doble. Al menos 20 inflorescencias independientes constituyen una muestra.

C) Blot para ARNs pequeños contra miR159 realizado a partir de plantas sobre expresantes de miR319a, miR319a-B4 y miR319a-reconocimiento doble. Al menos 20 inflorescencias independientes constituyen una muestra.

D) Distribución de los defectos del desarrollo en las hojas de plantas transgénicas sobre-expresantes de cada vector. Al menos 40 plantas fueron evaluadas para cada constructo.

8.6. El grado de apareamiento de los dúplex de ARN puede cuantificarse mediante su energía libre de interacción y difiere a lo largo del precursor de miR319.

Para el precursor de miR319a observamos que existe una relación cualitativa entre el grado de apareamiento de un dúplex de ARN y su acumulación. Parámetros empíricos como el cálculo del ΔG de dúplex de ARN se han empleado previamente para predecir de manera confiable interacciones entre miARNs y genes blanco (Allen et al., 2005; Rajagopalan et al., 2006; Rodriguez et al., 2010). Es por ello que decidimos cuantificar el grado de apareamiento de los dúplex de ARN estudiados hasta ahora mediante el cálculo de su energía libre de interacción. Para esto empleamos el programa Two State Melting que permite calcular el ΔG de dúplex de ARN de complementariedad parcial (<http://unafold.rna.albany.edu/?q=DINAMelt/Two-state-melting>).

En el precursor de miR319a, el dúplex de ARN con mayor valor de $|\Delta G|$, es decir más estable corresponde al miARN/miARN*. Además existe una considerable diferencia energética entre el dúplex de miR319a y los otros dos dúplex, siendo miR319a.1 el que presenta el menor valor de $|\Delta G|$, es decir más inestable (Figura R3.5). Los cálculos de energía para los dúplex derivados de miR319a-LIC y miR319a-B4 muestran que los dúplex mutantes de miR319a.1 y miR319a.2 son más estables que sus contrapartidas silvestres (Figura R3.5). A su vez el dúplex de miR319a.2 proveniente de miR319a-B4, es casi tan estable como el dúplex de miR319a. Lo cual podría ser uno de los motivos por los cuales se produzca el detenimiento de DCL1 luego de liberar este dúplex.

En función de lo desarrollado hasta ahora, parecería existir en primer instancia una relación cuantitativa entre el grado de apareamiento del dúplex de ARN y su acumulación. En segundo lugar el grado de apareamiento del dúplex de ARN podría afectar también la procesividad con la cual la maquinaria de procesamiento se mueve a lo largo del precursor de miR319a realizando cortes en fase. Esta última parte de la hipótesis requiere que DCL1 sea capaz de censar la estabilidad de los dúplex de ARN a medida que realiza los cortes.

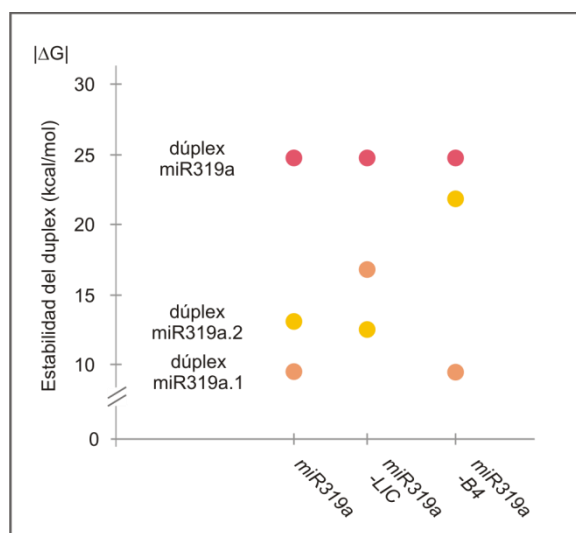


Figura R3.5. La energía de interacción de los dúplex de ARN refleja su estabilidad *in vivo*.

Gráfico de dispersión que muestra la estabilidad de cada dúplex liberado por cortes de DCL1, para miR319a, miR319a-LIC, y miR319a-B4. El dúplex miARN319a, representado en rojo está compuesto por el miR319a/miR319a*, el dúplex de miR319a.2, representado en amarillo está compuesto por los ARNs pequeños miR319a-3p.2/ miR319a-5p.2 y el dúplex de miR319a.1, representado en naranja está compuesto por los ARNs pequeños miR319a-3p.2/ miR319a-5p.2. En el eje de las ordenadas se representa el valor absoluto de la energía libre de Gibbs en kcal/mol.

8.7. Un modelo global de los requisitos para la acción secuencial de DCL1 en Arabidopsis

A partir de los cortes presentes en las bibliotecas de SPARE, obtuvimos las secuencias de los ARNs pequeños provenientes de los precursores procesados secuencialmente tanto conservados como jóvenes. Todos los precursores procesados de manera secuencial requieren al menos tres cortes de DCL1, pero difieren en la presencia de un cuarto corte (Figura R3.6B). Por ende no todos presentan un tercer dúplex correspondiente al miARN.2 como miR319a. Debido a esto, sólo graficamos los valores de $|\Delta G|$ obtenidos para el dúplex de miARN (Figura R3.6A, panel superior círculos rojos) y para el miARN.1 (Figura R3.6A, panel superior círculos naranjas).

Encontramos de manera consistente que el dúplex correspondiente al miARN (Figura R3.6A, panel superior círculos rojos) interacciona más fuertemente que los otros dúplex (Figura R3.6A, panel superior círculos naranjas). Esto lo observamos tanto para los precursores de miARNs evolutivamente antiguos, como para los jóvenes que evolucionaron recientemente. El caso particular que observamos para el precursor de miR319 sería entonces general para todos los precursores procesados secuencialmente, donde el dúplex correspondiente al miARN se distingue de otros dúplex secundarios por su mayor estabilidad.

Es interesante que los dúplex de los miARN conservados tienen valor de $|\Delta G|$ promedio superior al valor de $|\Delta G|$ promedio de los dúplex de miARNs jóvenes (Figura R3.6A). Y a su vez, los dúplex de ARNs pequeños secundarios tienen mayor estabilidad en los *MIARNs* jóvenes que en los precursores conservados. Si se calcula la diferencia de estabilidad entre el dúplex del miARN y el secundario ($|\Delta\Delta G|$, puntos azules en la figura R3.6A) se observa que la diferencia es menor para los miARNs jóvenes que los conservados evolutivamente. Estas diferencias en el patrón de energía de los distintos dúplex entre miARNs conservados y jóvenes es significativa (T- Student $P < 0,05$). Esto sugiere que durante la evolución de los precursores que se procesan secuencialmente se seleccionarían mutaciones que cambien la energía relativa de los distintos dúplex de ARN pequeños.

Cabe destacar que el dúplex miR394a presenta el valor de $|\Delta G|$ más bajo de todos los miARNs conservados. Incluso este valor es comparable al dúplex de miARN.1 de miR394b, aunque sigue valiendo la regla que el dúplex de miARN es más estable a otros dúplex secundarios. Sería interesante estudiar la biogénesis de miR394a en más detalle para esclarecer si estas características particulares son producto de la interacción con proteínas específicas durante su procesamiento o la carga del dúplex en alguna de las proteínas AGO. Para miR390 se ha demostrado que la presencia de una A en posición 5' y los mismatches en región central del dúplex son esenciales para la interacción específica con AGO7 (Endo et al., 2013).

Sin embargo parece existir un límite tanto superior como inferior para el valor de $|\Delta G|$. El mayor valor de $|\Delta G|$ para un dúplex de miARN corresponde es de 27,4 kcal/mol para el dúplex de miR169b. Mientras que el valor de $|\Delta G|$ esperado para un dúplex compuesto por un ARN exactamente complementario a la secuencia de miR169b maduro tendría un valor de 38 kcal/mol el cual es un valor significativamente mayor. Los motivos para esto no están del todo claro, pero una posible explicación podría ser que un dúplex de miARN de complementariedad perfecta podría hacer que el precursor se vuelva sustrato de otras DCLs (DCL2-4), que poseen preferencia por ARNdhs perfectos. Otro factor a tener en cuenta es que la incorporación del miARN a AGO1 requiere que el mismo sea separado del miARN*, por ende un dúplex demasiado estable, podría dificultar este paso (Bologna and Voinnet, 2014; Vaucheret, 2008).

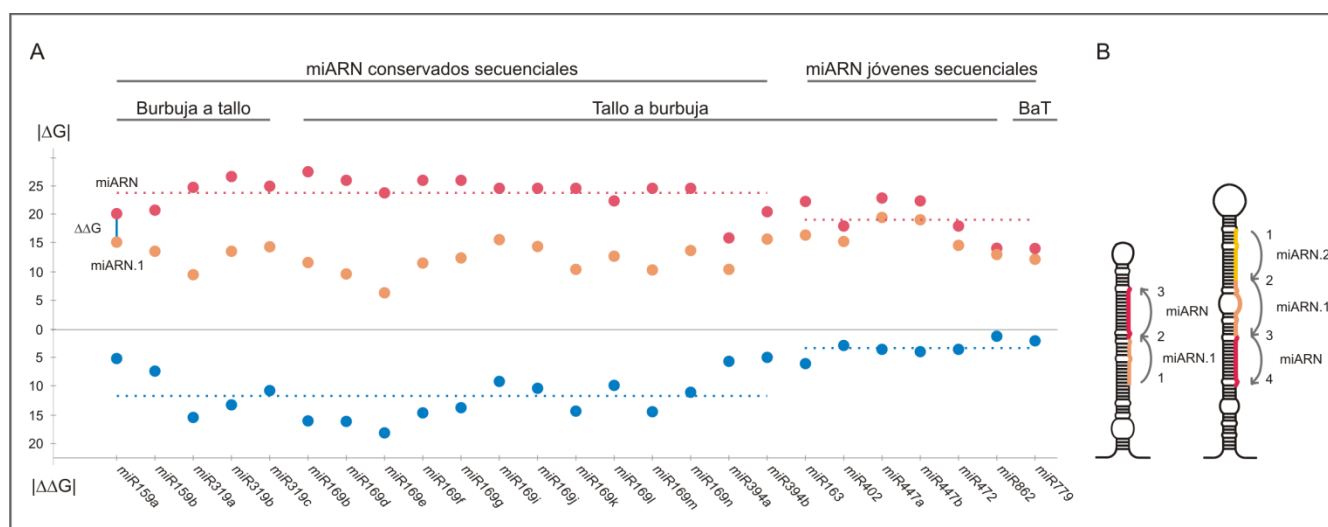


Figura R3.6. Análisis global de la energía de interacción de los dúplex derivados de *MIARNS* secuenciales en *Arabidopsis*.

A) Gráfico de dispersión que muestra la estabilidad de los dúplex miARN (puntos rojos) y miARN.1 (puntos naranjas) liberado por cortes de DCL1, para todos los precursores de miARNs conservados y jóvenes procesados secuencialmente. Se muestra a su vez, la estabilidad promedio de los dúplex de miARN conservados y jóvenes (línea punteada roja). En el eje de las ordenas se representa el valor absoluto de la energía libre de Gibbs ($|\Delta G|$) para la interacción de cada dúplex. En la parte inferior de gráfico se muestra la diferencia entre la estabilidad del dúplex de miARN y el de miARN.1 (puntos azules), representada por el valor absoluto de diferencia de las energías libres de Gibbs ($\Delta\Delta G$), junto con la estabilidad promedio de los dúplex de miARN conservados y jóvenes (línea punteada azul). Nótese que para la porción inferior del gráfico la escala de $|\Delta G|$ está invertida.

B) Representación esquemática de un precursor de miARN procesado secuencialmente desde la base con tres cortes que liberan dos dúplex de ARNs pequeños (izquierda) y un precursor procesado desde el loop terminal con cuatro cortes, que liberan tres dúplex de ARNs pequeños (derecha). En los gráficos de dispersión sólo se tuvieron en cuenta los valores de los dúplex miARN y miARN.1.

8.6. Mutaciones puntuales en miR319.2 permiten su acumulación sin alterar la cantidad de miR319.

A partir del gráfico presentado en la Figura R3.5 se pueden deducir los límites energéticos para que un dúplex de ARN empieza a acumularse significativamente, lo cual sería a partir de ~ 16 kcal/mol. Como prueba de concepto diseñamos dos mutantes sobre miR319a-5p.2 denominadas, miR319a-G:U y miR319a-

A:U que mejoran el apareamiento del dúplex de miR319a.2 (Figura R3.7A, recuadro amarillo). El cálculo de energía libre asociado al dúplex fue de 16,8 kcal/mol y 17,1 kcal/mol respectivamente (Figura R5.7D).

Se generaron plantas transgénicas sobre-expresantes del precursor de miR319a, *miR319aU:G* y *miR319aC:G*. A partir de plántulas de 10 días, se hizo una extracción de ARN y blots para ARN pequeños. El blot contra la región de miR319a-3p.2 reveló que tanto *miR319a-U:G* como *miR319a-C:G* acumulaban el ARN pequeño miR319a-3p.2 a diferencia de *miR319a* (Figura R3.5B). Esto nos llevó a pensar que el límite empírico deducido anteriormente, podía emplearse para predecir si un ARN pequeño va a acumularse *in vivo* significativamente. Incluso puede verse que miR319a-C:G acumula más cantidad de miR319a-3p.2 que miR319a-U:G. Lo cual está de acuerdo con el cambio del par de bases no canónico G:U por un par canónico, G:C tuvo un impacto *in vivo* sobre las cantidades finales de miR319a-3p.2.

El blot contra miR319 (Figura R3.5C) mostró que los niveles de miR319 eran similares para todas las plantas que sobre expresaban alguna versión de *miR319a*. Este resultado nos indica que la mutación incorporadas en *miR319a-U:G* y *miR319a-C:G* no tuvieron un efecto negativo en el procesamiento de sus precursores. Sin embargo la mutación puntual en miR319a-5p.2, sí tuvo un efecto cualitativo sobre estabilidad de miR319a-3p.2, generando un precursor de miARN capaz de acumular dos ARNs pequeños con secuencias completamente diferentes. Esto sugiere que la acumulación de un ARN pequeño por aumento de su estabilidad y el detenimiento del procesamiento secuencial de precursores, son dos procesos independientes que pueden desacoplarse. A su vez el precursor de *miR319a-C:G* podría llegar a usarse como molde para expresar dos miARN artificiales de secuencias no relacionadas para generar “knock down” de dos genes blanco diferentes en simultáneo.

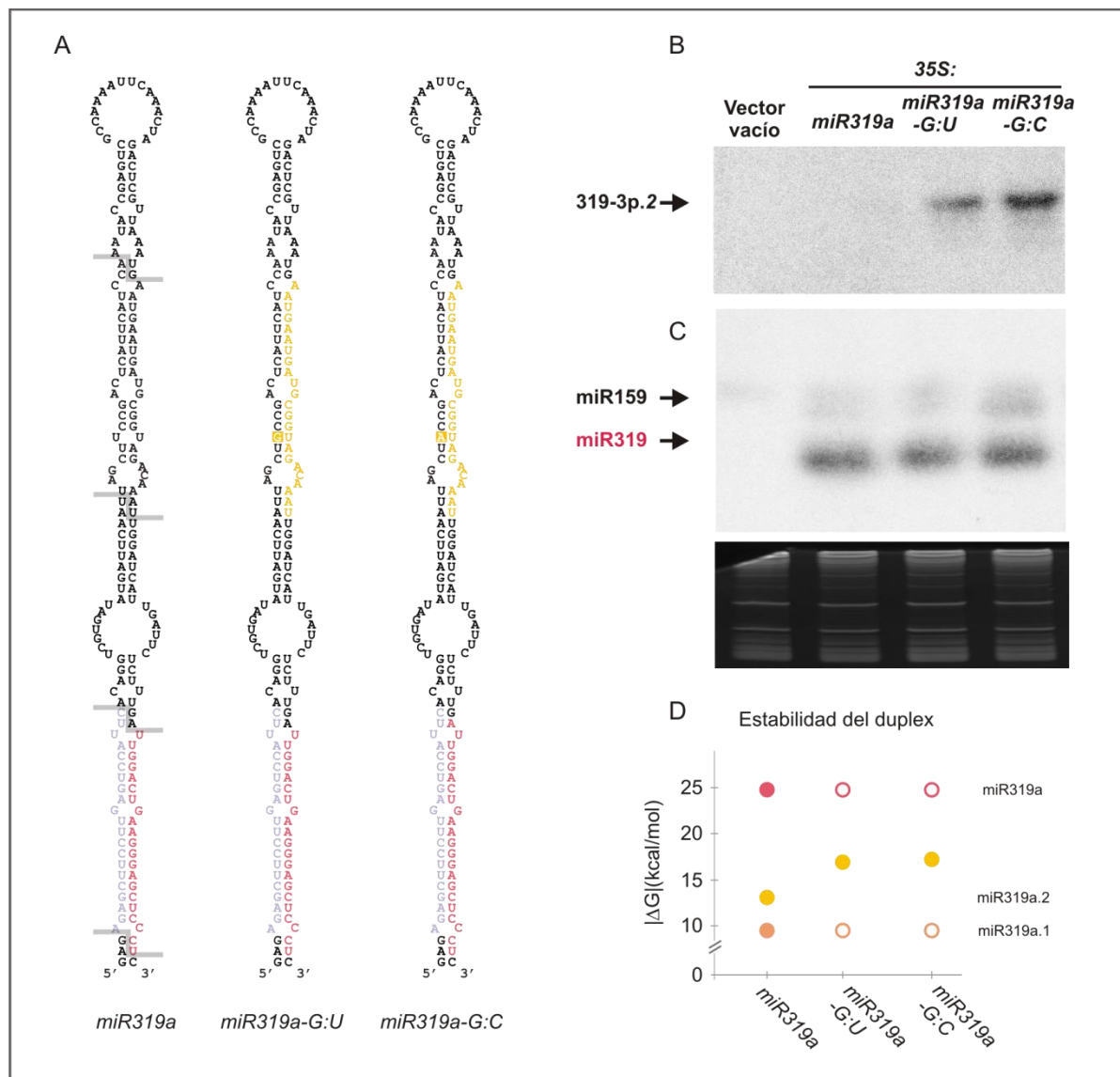


Figura R3.7. Mutaciones puntuales en el RNA1 permiten su acumulación sin afectar la biogénesis de miR319.

A) Predicción de estructura secundaria para los precursores de miR319a, miR319a-G:U y miR319a-G:C. Las bases mutadas se marcan con un cuadrado amarillo y el miARN se muestra en rojo.

B) Blot para ARNs pequeños contra miR319-3p.2 a partir de líneas transgénicas sobre-expresantes de los precursores de miR319a, miR319a-G:U y miR319a-G:C. Al menos 20 inflorescencias independientes constituyen una muestra.

C) Blot para ARNs pequeños contra miR319 a partir de líneas transgénicas sobre-expresantes de los precursores de miR319a, miR319a-G:U y miR319a-G:C. Al menos 20 inflorescencias independientes constituyen una muestra.

D) Gráfico de dispersión que muestra la estabilidad de cada dúplex liberado por cortes de DCL1, para el precursor de miR319a, miR319a-G:U y miR319a-G:C. En el eje de las ordenadas se representa el valor absoluto de la energía libre de Gibbs ($|\Delta G|$) para la interacción de cada dúplex. Los dúplex de miR319a y miR319a.1 para miR319a-G:U y miR319a-G:C se muestran como un círculo vacío ya que no se han mapeado los cortes en el precursor y los mismos se obtuvieron por predicción.

8.8. DISCUSIÓN III

8.8.1. En plantas estructura secundaria de los dúplex de ARN afectan tanto la acumulación del ARN pequeño como la procesividad de DCL1.

Los resultados presentados en este capítulo nos permitieron proponer que la estructura secundaria de los dúplex de ARN, presentes en los precursores procesados secuencialmente, afectan tanto la acumulación del ARN pequeño como la procesividad con la cual DCL1 se mueve a lo largo del precursor generando cortes sucesivos. Trabajos anteriores mostraron que la estructura secundaria del dúplex de miARN, específicamente la presencia de bulges asimétricos, son suficientes para reprogramar al complejo efector RISC, y desencadenar siARNs a partir del ARNm blanco del miARN, en un proceso análogo a la producción de tasiARNs a partir de los genes TAS (Cuperus et al., 2010a; Chen et al., 2010; Manavella et al., 2012b). A su vez la estructura secundaria del dúplex de miARN miR165/6 produce que el mismo sea secuestrado en AGO10, y en lugar de cargarse en AGO1, en un circuito regulatorio que permite el desarrollo del meristema apical (Zhu et al., 2011).

La relación entre la estructura secundaria de un dúplex de ARN y su acumulación puede explicarse si se tiene en cuenta una vez escindido del precursor, un dúplex de ARN debe metilarse e incorporarse en alguna proteína AGO para ser estable (Bologna and Voinnet, 2014). Por ende si el dúplex posee un valor de $|\Delta G|$ menor al límite empírico de ~ 16 kcal/mol, el mismo podría ser inestable y tendería a separarse en sus hebras individuales. Tanto HEN1 como AGO1 tienen como sustrato un dúplex de ARN por la separación del mismo causaría que las hebras individuales no sean reconocidas por estas proteínas y e ingresen en las vías de degradación de ARN pequeños sin metilar (Bologna and Voinnet, 2014).

La segunda parte del modelo propone que la estructura secundaria del dúplex de ARN influenciaría además la actividad procesiva de DCL1. Para el precursor de miR319a el dúplex del miR319a es el más estable, y luego de que este es liberado no se producen otros cortes corriente abajo en el precursor. A su vez esta tendencia se observó para los otros precursores procesados secuencialmente en Arabidopsis (Figura R3.6A). Lo cual nos llevó a pensar que DCL1 realiza cortes posesivamente hasta que encuentra un dúplex de ARN suficientemente estable como para ser considerado el dúplex de miARN, con límite inferior de $|\Delta G| \sim 20$ kcal/mol, o bien se encuentra con la pérdida de estructura de ARNdh presente en los precursores de miARNs.

Esta hipótesis implica que DCL1 sería capaz de censar la estabilidad de los dúplex de ARN a medida que realiza los cortes. Se ha propuesto que el dominio helicasa de DCL1 hidroliza ATP para promover el cambio conformacional necesario para realizar el segundo corte y liberar el miARN maduro en el precursor (Liu et al., 2012). El dominio helicasa funcionaría como una traslocasa dependiente de ATP que proveería la energía necesaria para que la enzima se desplace a lo largo del sustrato de ARN. Durante este proceso, el

dominio helicasa podría censar la estabilidad de los dúplex de ARN a mientras se desplaza sobre ellos, y en función de esto promover o no los cambios conformacionales necesarios para que DCL1 pueda continuar cortando a lo largo del precursor.

Como se mencionó en la Sección 8.6 la acumulación de un ARN pequeño por aumento de su estabilidad y el detenimiento de la maquinaria de procesamiento luego de liberarlo, serían son dos procesos independientes y tendrían dos límites empíricos diferentes de $|\Delta G| \sim 16 \text{kcal/mol}$ y $|\Delta G| \sim 20 \text{kcal/mol}$ respectivamente. Concordante con esto mutaciones aumenten la energía de interacción de un dúplex de ARN por encima del valor de $\sim 20 \text{kcal/mol}$ provocan que el mismo se acumule y afectan la actividad procesiva de DCL1 (miR319a-B4, Figura R3.4). Mientras que mutaciones puntuales que mejoren el apareamiento del dúplex de ARN dentro de los límites de $\sim 16 \text{kcal/mol}$ y $\sim 20 \text{kcal/mol}$, permiten la acumulación del ARN pequeño, sin alterar las cantidades finales del miARN maduro (miR319a-U:G y miR319a-C:G, Figura R3.7).

8.8.2. Relación entre la evolución de los miARNs y el procesamiento secuencial de los precursores.

Una teoría acerca del origen de los miARNs propone que los mismos se generaron como duplicaciones invertidas de genes codificantes para proteínas (Allen et al., 2004; Rajagopalan et al., 2006) (Figura D3A). En caso de transcribirse, este evento de duplicación invertida daría lugar a una estructura secundaria de tipo tallo y loop, con apareamiento perfecto o casi perfecto. Dada la estructura secundaria del proto-miARN, el mismo podría ser reconocido y procesado por las proteínas DCL3 y DCL4, dando lugar a un conjunto de siARNs, potencialmente capaces de regular al gen de origen (Figura D3BE).

Con el tiempo se acumularían mutaciones en la secuencia del proto-miARN, para dar lugar a un miARN joven. Estas mutaciones incluyen un incremento en del número de bases desapareadas en el precursor, el acortamiento del tallo o aumento del loop terminal (Vazquez et al., 2008). Estos cambios permitirían que los miARN jóvenes empiecen a ser reconocidos y procesados por el complejo de DCL1 (Figura D3CF). En concordancia con esto, se han encontrado miARNs de evolución reciente como miR822, miR839 y miR869, los cuales posiblemente se generaron a partir de duplicaciones invertidas de los genes blancos, que son procesados por DCL4 en lugar de DCL1 (Ben Amor et al., 2009; Rajagopalan et al., 2006).

A continuación sucesivas mutaciones traerían como resultados nuevos cambios en la estructura secundaria. Lo que a su vez permitirían que los miARN jóvenes comenzaran a ser procesados preferentemente por DCL1, dando eventualmente lugar a un precursor de miARN conservado. Se ha demostrado que genes de miARNs de evolución reciente dan frecuentemente lugar a lmiARN de $\sim 24 \text{nt}$, al ser procesados preferentemente por DCL3, mientras que miARNs conservados dan predominantemente ARNs pequeños canónicos de $\sim 21 \text{nt}$ luego del corte por DCL1 (Vazquez et al., 2008)(Figura D3DG).

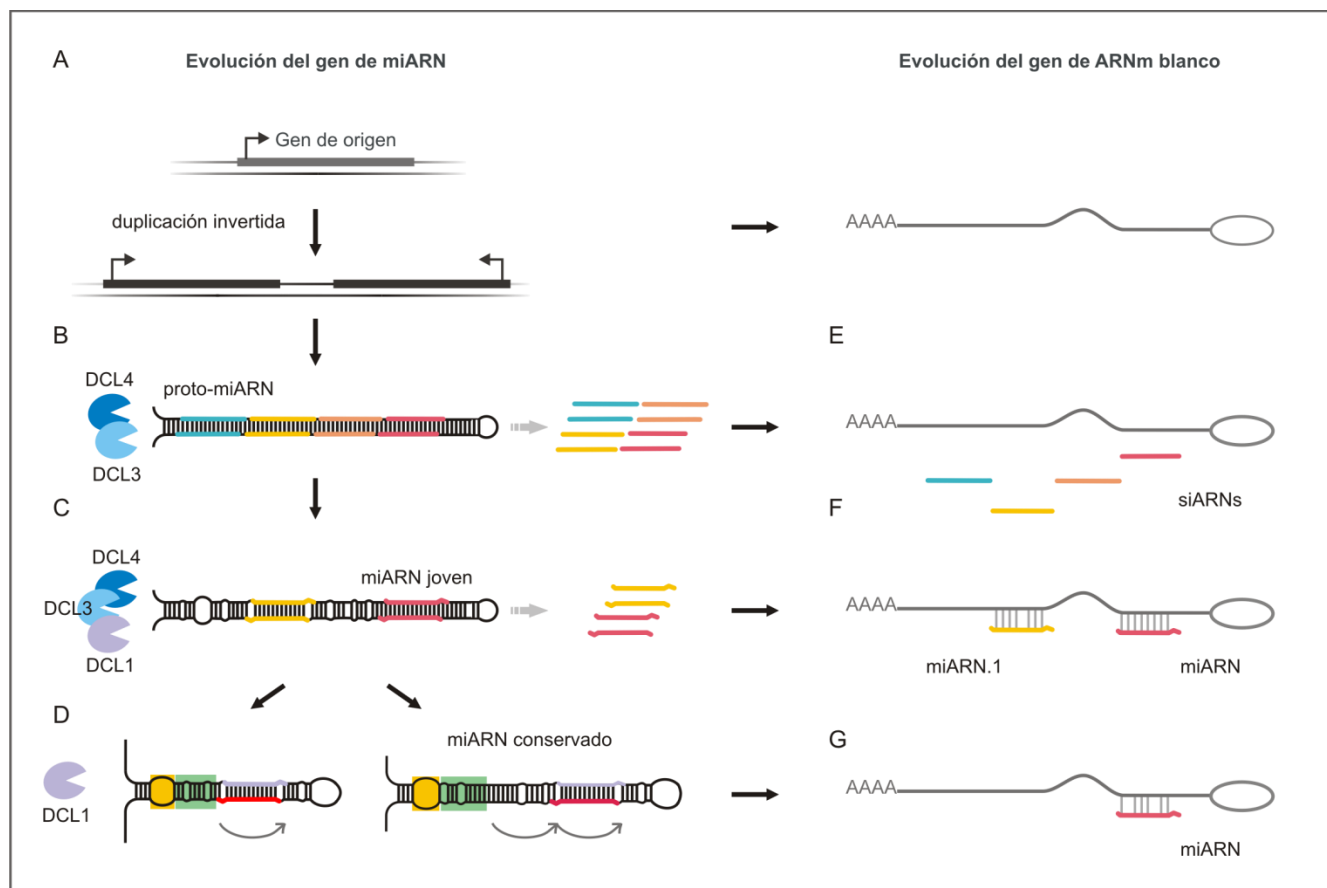


Figura D3. Relación entre la evolución de los miARNs y el procesamiento secuencial de los precursores.

- A)** Un gen codificante para proteínas sufre una duplicación invertida.
- B)** La duplicación invertida es capaz de transcribirse y dar lugar a un proto-miARN el cual puede ser procesado por DCL3 y DCL4 dando lugar a un conjunto heterogéneo de siARNs pequeños
- C)** Mediante la acumulación de mutaciones el proto-miARN da lugar a un miARN joven el cual comienza a ser procesado por DCL1, acotado el número de ARNs pequeños derivados del mismo.
- D)** Nuevas mutaciones producen cambios en la estructura secundaria del precursor de miARN que generan el cambio de enzima de procesamiento de DCL4/DCL3 \rightarrow DCL1. Un único miARN es producido desde el precursor de miARN conservado que puede requerir dos o más cortes. El miARN se muestra en rojo.
- E)** Los ARNs pequeños derivados del proto-miARN son complementarios al gen original, pudiendo controlar los niveles finales de su ARNm.
- F)** Los ARNs pequeños derivados del precursor de miARN joven retienen la complementariedad con el ARNm del gen blanco original.
- G)** El miARN maduro controla los niveles finales del ARNm mediante el corte del ARNm o la represión traduccional.

La diversidad actual de los sistemas regulatorios miARN-gen blanco, puede explicarse si a su vez se considera que el locus del miARN, puede sufrir duplicaciones dando lugar a familias multigénicas, como es común en miARNs conservados, y sólo retienen complementariedad parcial con sus genes blanco originales. Por otro lado el gen blanco también puede sufrir duplicaciones adicionales, aumentando el número de genes bajo el control del miARN, e incluso algunos de ellos puede perder el sitio blanco de unión del miARN (Allen et al., 2004).

Esta teoría sobre el origen de los miARNs predice que en los precursores de evolución reciente pueden encontrarse largos segmentos de elevada similitud con el gen blanco. Esto último se ha demostrado para *MIR161* y *MIR163*, dos genes de evolución relativamente reciente (Allen et al., 2004). Tomando a *MIR163* como modelo de un miARN joven procesado secuencialmente es posible usar la teoría del origen de los miARNs para explicar los resultados obtenidos en la Figura R3.6 (ANEXO A38). El valor de $|\Delta G|$ para miR163 es de 22,2 kcal/mol, mientras que para miR163.1 es de 16,3 kcal/mol, lo cual da como resultado un valor de $|\Delta\Delta G|$ de 5,9 kcal por mol. Este número es aproximadamente la mitad del $|\Delta\Delta G|$ promedio de los miARN conservados. Por ende la baja diferencia entre las energías de interacción entre el miR163 y el miR163.1 puede deberse a que *MIR163* no ha acumulado suficientes mutaciones en la parte del tallo que corresponde al dúplex de miR163.1. En concordancia con esto, se ha demostrado que miR163-5p.1 se acumula in vivo y es detectable mediante Northern blots (Kurihara and Watanabe, 2004). Esto es diferente de lo que ocurre en los precursores de miARN conservados, los cuales presentan un mayor número de mismatches en las diferentes regiones del tallo (Vazquez et al., 2008) y acumulan preferentemente el miARN maduro (Figura R3.3 y R3.4). En última instancia, un mayor número de mismatches se traduce en un menor valor de $|\Delta G|$ para el miARN.1 y por ende en una mayor diferencia energética con el dúplex del miARN.

Si bien la teoría mencionada previamente permite explicar la presencia de mismatches en los dúplex de miARN.1 no explica por qué los dúplex de los miARN conservados son más estables que los dúplex de los miARN jóvenes. Esto sugiere que existiría una presión de selección negativa sobre las mutaciones que empeoren el apareamiento entre el miARN y el miARN*, a medida que el precursor diverge de las secuencias del gen original durante la evolución de los locus de *MIARNs*.

Como se mencionó miR163 presenta extensa complementariedad con su gen de origen, la cual incluye la secuencia del sitio target (Allen et al., 2004). Esta característica se ha visto para otros miARN jóvenes y sus genes blancos (Rajagopalan et al., 2006) y es posible que sirva como un mecanismo compensatorio, que permita mantener la presión de selección positiva sobre un miARN joven hasta que este acumule suficientemente mutaciones para ser procesado con mayor precisión (Figura D3FG).

Non, je ne regrette rien

9. CONCLUSIONES

- Se desarrolló un protocolo optimizado de la técnica de SPARE que permite aumentar la sensibilidad significativamente y analizar plantas silvestres.
- Con los datos obtenidos de las bibliotecas de Col-0 y *fiery1* se caracterizó el mecanismo de procesamiento de 45 y 75 precursores respectivamente. De los cuales 16 precursores que no habían sido identificados previamente.
- A partir de los datos genómicos obtenidos de las bibliotecas de SPARE se caracterizaron en total los mecanismos de procesamiento responsables de la biogénesis de 91 miARNs en Arabidopsis. Los mismos pueden ser agrupados en cuatro rutas dependiendo de la dirección y el número de cortes requeridos para liberar al miARN maduro.
- Mediante SPARE también se detectaron cortes no productivos en los precursores de miARNs, y se observó que estos son relativamente frecuentes *in vivo*. En principio, el procesamiento abortivo puede iniciarse desde loop internos inferiores o loops terminales ramificados, ambos ubicados a una distancia de ~15-17pb del dúplex de miARN/miARN*. Además, la dirección del procesamiento abortivo puede ser igual u opuesta, a la del procesamiento productivo.
- Estos cortes improductivos *in vivo* demuestran que la maquinaria de procesamiento es capaz de reconocer múltiples determinantes estructurales independientemente de que estos estén asociados a la biogénesis de miARNs o no. La plasticidad de la maquinaria de procesamiento podría explicar la diversidad en tamaños y formas en los precursores de miARN en plantas.
- Los precursores procesados desde el loop terminal hacia la base pueden clasificarse en grupos. Un primer grupo está representado por el precursor de miR171b, el cual posee una región terminal homogénea. Otro grupo está representado por la familia de miR319/159, con un extenso tallo superior, como consecuencia de la actividad secuencial de DCL1. Y por último el precursor de miR157c que presenta un loop terminal de gran tamaño.
- Para los precursores procesados desde el loop con una región terminal homogénea, la misma tiene un tamaño de ~42nt. La región terminal está compuesta por un tallo de ARNdH ~16-17pb y un loop terminal pequeño. Proponemos que la posición del primer corte se establecería midiendo la distancia desde el loop terminal en pares de bases.
- El requisito de una región terminal homogénea para el precursor de miR171b es extensivo a otras especies vegetales y a su vez parecería ser un requisito exclusivo de los precursores procesados del loop a la base.
- Proponemos un modelo unificado para la biogénesis de los precursores procesados por DCL1. En aquellos precursores procesados desde la base al loop, DCL1 reconocería la presencia del loop interno inmediatamente por debajo del segmento de ARNdH y realizaría el primer corte a ~15pb. En el caso de los precursores procesados de arriba hacia abajo, la maquinaria de procesamiento reconocería el loop terminal

y a partir de allí realizaría un corte a ~17pb. Una vez generado el primero corte o endocorte, el segundo corte en ambos tipos de precursores se realizaría contando ~21nt desde el extremos libre.

- El precursor de miR157c se procesa desde el loop hacia la base con dos cortes de DCL1. Este precursor presenta un tallo superior de ~18pb seguido de un loop terminal ramificado desde donde se produciría el primer corte. Es la primera vez que se describe y caracteriza este mecanismo de procesamiento en plantas para generar un miARN funcional. Mientras que el tallo inferior que el loop terminal ramificado afecta cuantitativamente la biogénesis de miR157c.

- Es probable que la posición del primer corte en el precursor de miR157c se establezca por un mecanismo diferente a miR171b. En el precursor de miR157c DCL1 reconocería la transición de ARNsh-dh presente entre el loop terminal ramificado y el tallo superior para realizar el primer corte por encima del dúplex del miARN.

- En el precursor de miR319a el primer corte es un endocorte, en el que se generan los extremos libres para que DCL1 realice cortes sucesivos en una fase de ~21nt. La acción secuencial de DCL1 requiere un ARNdh pero con apareamiento parcial.

- Una misma molécula de precursor de miR319 y probablemente otros, no puede ser procesado simultáneamente por dos DCL1 en direcciones opuestas *in vivo*.

- El grado de apareamiento o desapareamiento de un dúplex de ARN puede cuantificarse mediante del cálculo del ΔG , y estos valores emplearse para hacer predicciones sobre la acumulación de ARNs pequeños *in vivo*.

- La estructura secundaria de los dúplex de ARN, presentes en los precursores procesados secuencialmente, afectan tanto la acumulación del ARN pequeño como la procesividad con la cual DCL1 de mueve por el precursor. Sin embargo estos son dos procesos independientes y pueden desacoplarse ya que tendrían dos límites empíricos diferentes $|\Delta G| \sim 16 \text{kcal/mol}$ para acumularse eficientemente y $|\Delta G| \sim 20 \text{kcal/mol}$ para detener la procesividad de DCL1.

- En el caso de los dúplex de miARNs, hay una tendencia al apareamiento, lo que implica un mayor valor de $|\Delta G|$. Mientras que para los dúplex adicionales de miARN.1 hay una tendencia hacia dúplex preferentemente desapareados, con un menor valor de $|\Delta G|$. A su vez esta diferencia energética es mayor en el caso de los miARN conservados respecto de los jóvenes.

- Durante la evolución de los miARNs habría selección sobre el grado de apareamiento de los distintos dúplex de ARN pequeños.

- Esta hipótesis implica que DCL1 sería capaz de censar la estabilidad de los dúplex de ARN a medida que realiza los cortes. Nuestra propuesta que el dominio helicasa de DCL1 podría censar la estabilidad de los dúplex de ARN a mientras se desplaza sobre ellos, y en función de esto promover o no los cambios conformacionales necesarios para que DCL1 pueda continuar cortando a lo largo del precursor.

10. ANEXO

Tabla A1. Secuencias de los oligos específicos y composición de las mix de oligos.

En mayúscula se muestran las secuencias específicas de cada precursor de arabidopsis.

Mix 1		
n° MIX	miARN	secuencia
1	miR156a	ccttggcaccgagaattccaTTAGATCGTATCTTCTTACC
2	miR156b	ccttggcaccgagaattccaACAGACGAGATGATAAGAAG
3	miR156c	ccttggcaccgagaattccaTGAGAGATGAAGAACACATG
4	miR156d	ccttggcaccgagaattccaTAGATGCAACATATGTATGC
5	miR156e	ccttggcaccgagaattccaCTTCGACCTACTTTGATCCG
6	miR156f	ccttggcaccgagaattccaAATATGCTGATTCATGTTTG
7	miR156g	ccttggcaccgagaattccaTCTAACCATACACAGAGACG
8	miR156h	ccttggcaccgagaattccaACGTACCTTACTTGATAGTG
9	miR157a	ccttggcaccgagaattccaTTCATAACTGTTTCAATCAC
10	miR157b	ccttggcaccgagaattccaATCTGCATTCTGATAGTTGC
11	miR157c	ccttggcaccgagaattccaCACCCATGTTAGTATTACGC
12	miR157d	ccttggcaccgagaattccaCACTTTTCTCACACCAAAC
13	miR158a	ccttggcaccgagaattccaTCTAGTTTTGAGCGAGATCC
14	miR158b	ccttggcaccgagaattccaAACAAGCAGTCAGTGAAATC
15	miR159a	ccttggcaccgagaattccaTCAAATTATAGCGAATAATC

Mix 2		
n° MIX	miARN	secuencia
1	miR156c.2	ccttggcaccgagaattccaGGAACCAAATCAATGGCAGC
2	miR159b	ccttggcaccgagaattccaCTACTCAAGATCCATCATCC
3	miR159c	ccttggcaccgagaattccaTGCAAATAAACATGACAACC
4	miR160a	ccttggcaccgagaattccaCTACACATGATGAGGCAATG
5	miR160b	ccttggcaccgagaattccaGTTATAGACAATTAGACATC
6	miR160c	ccttggcaccgagaattccaCTGTTTGCTTATTCAAATGG
7	miR161	ccttggcaccgagaattccaCTAATTAAATCAAATCGATC
8	miR162a	ccttggcaccgagaattccaAGTAATCGGACTTGACTCTG
9	miR162b	ccttggcaccgagaattccaGCTAAAAGATGAATACTTTG
10	miR163	ccttggcaccgagaattccaGGCATGAATTTAATTACATG
11	miR164a	ccttggcaccgagaattccaGTCGAACACAAATGATTTAAC
12	miR164b	ccttggcaccgagaattccaCGATCTAGGCTAGCTCGTAC
13	miR164c	ccttggcaccgagaattccaCAATGTAACTTCATGTCTC
14	miR165a	ccttggcaccgagaattccaAAGCCATGCAAGAAAGATTC
15	miR165b	ccttggcaccgagaattccaAGACGCCAATGGTAGTTACC

Mix 3		
n° MIX	miARN	secuencia
1	miR156c.3	ccttggcaccgagaattccaGATAACATTTTCATACCTTTG
2	mir162a.2	ccttggcaccgagaattccaAACAAGCACTTCATTAAGT
3	miR166a	ccttggcaccgagaattccaATGAATCTGAGAAAGTAAGG
4	miR166b	ccttggcaccgagaattccaTATATCACATGGATTCATAG
5	miR166c	ccttggcaccgagaattccaATTAATCTAATAACAAGATC
6	miR166d	ccttggcaccgagaattccaGCTCTCACTTCAGGATCTAC
7	miR166e	ccttggcaccgagaattccaGAAATTGAAGTTGCTTGAAC
8	miR166f	ccttggcaccgagaattccaTACATTGCTGCGGATTGATG
9	miR166g	ccttggcaccgagaattccaACATGGTTATACTCTAGATG
10	miR167a	ccttggcaccgagaattccaAGAAAGAGAAGTAAGCTCAC
11	miR167b	ccttggcaccgagaattccaTGGAGAGTGTGTCAAAGCAG
12	miR167c	ccttggcaccgagaattccaAATATAATTAATCTCTGCTG
13	miR167d	ccttggcaccgagaattccaCTTTCTCATGAAATGAAGTG
14	miR168a	ccttggcaccgagaattccaAAACAATTTTCAGATTCAAAG
15	miR168b	ccttggcaccgagaattccaAACCCAATACCGAATCAATC

Mix 4		
n° MIX	miARN	secuencia
1	mir167d.2	ccttggcaccgagaattccaACAAGTAAATGTATCGTTAC
2	miR169a	ccttggcaccgagaattccaCTTTCTGCATTGTTCCCTTAG
3	miR169b	ccttggcaccgagaattccaAAATACTCATAACGGTCGATG
4	miR169c	ccttggcaccgagaattccaCTCATTATATTAGACCATCC
5	miR169d	ccttggcaccgagaattccaTATTAGCATTAGCATTCACC
6	miR169e	ccttggcaccgagaattccaTATATACATTTCAACGATAC
7	miR169f	ccttggcaccgagaattccaTTGAGACAAATTAAACATCG
8	miR169g	ccttggcaccgagaattccaAAATCTGATCATTCAAATCG
9	miR169h	ccttggcaccgagaattccaCATTGACAAAGTCCACTATG
10	miR169i	ccttggcaccgagaattccaGCTCAAAGTCATCAACATTG
11	miR169j	ccttggcaccgagaattccaATGCTTTCTAAATCGAATGC
12	miR169k	ccttggcaccgagaattccaATCGTCAACATTCGCTCACC
13	mir169l	ccttggcaccgagaattccaAGTGATTTCGGAAGACAG
14	miR169m	ccttggcaccgagaattccaTCGAAATCATGAACATTATC
15	miR169n	ccttggcaccgagaattccaACCAACTGCGAAAATTTGAC

Mix 5		
n° MIX	miARN	secuencia
1	miR170	ccttggcaccgagaattccaATTGAGTGATCGATGAGTAC
2	mir171a	ccttggcaccgagaattccaCTATAGGTAAACAATATAAC
3	mir171b	ccttggcaccgagaattccaGAAATATCAAAGCCATTAATC
4	miR171c	ccttggcaccgagaattccaTTGATAATACCTCATCTCTG
5	mir172a	ccttggcaccgagaattccaGATATGTAAACATAAAGGTG
6	miR172b	ccttggcaccgagaattccaATATGTAAACATGTTCAAAC
7	miR172c	ccttggcaccgagaattccaTACCTCCGATCTGTGAATTC
8	miR172d	ccttggcaccgagaattccaAAGTTTCACCTCAAGTTATC
9	miR172e	ccttggcaccgagaattccaGTGCATGATCAAGATATTGC
10	miR173	ccttggcaccgagaattccaACCCTAATGAGATACTTTCC
11	mir319a	ccttggcaccgagaattccaCAAAATGTTAATTTTACCAG
12	miR319b	ccttggcaccgagaattccaACTTATTTATATTCATATCG
13	miR319c	ccttggcaccgagaattccaTCCAGTTTCAGTTCAATTCG
14	miR390a	ccttggcaccgagaattccaAAGATAGCTTAAATGGACAG
15	miR390b	ccttggcaccgagaattccaGATTTGAACTTCAACAATTC

Mix 6		
n° MIX	miARN	secuencia
1	miR391	ccttggcaccgagaattccaTTATGGTGTTACTATGTAAG
2	miR393a	ccttggcaccgagaattccaCTGTTGTAGGCTTGAGATAC
3	miR393b	ccttggcaccgagaattccaCTTGTTGATATGACTGGATC
4	mir394a	ccttggcaccgagaattccaATTACCCTAGATCGAGGCTC
5	miR394b	ccttggcaccgagaattccaGATAATACCTAGTTTCTTC
6	miR395a	ccttggcaccgagaattccaTTTATATCTTTAAGCCATTC
7	miR395b	ccttggcaccgagaattccaATTAGCTAGTGTCATCATTG
8	miR395c	ccttggcaccgagaattccaGTCCACACCATGAATCCATG
9	miR395d	ccttggcaccgagaattccaTCACTCATTTTTGTGGATCG
10	miR395e	ccttggcaccgagaattccaTTTTTGTGGATCGTTAATC
11	miR395f	ccttggcaccgagaattccaTCACTCATGAATGATAGATC
12	miR396a	ccttggcaccgagaattccaCTACAATATAGTTGGTAGTC
13	miR396b	ccttggcaccgagaattccaTCCTGTGTCTTCAATTTAGG
14	mir398c.3	ccttggcaccgagaattccaGTCGCATTTGAAACCATTTG
15	mir824	ccttggcaccgagaattccaCAACAAAGTCACTGCATTAC

Mix 7		
n° MIX	mirARN	secuencia
1	miR397a	ccttggcaccgagaattccaGCCCTAAATAATATCTGATG
2	miR397b	ccttggcaccgagaattccaAGAAACTAAATGTTGGAGTC
3	miR398a	ccttggcaccgagaattccaAGATACAAAATAGAGGTTCC
4	miR398b	ccttggcaccgagaattccaTACTACTGTGATTTTCATCTG
5	miR398c	ccttggcaccgagaattccaAGCCACGGGCCACGGCGTTG
6	miR399a	ccttggcaccgagaattccaAGGACTTGAACATCGTCATC
7	miR399b	ccttggcaccgagaattccaCAGTCTGTTCTATTCGGTTCG
8	miR399c	ccttggcaccgagaattccaAACCGCACTAGTTTTGTAGC
9	miR399d	ccttggcaccgagaattccaAGATTCCAAGATTGATCTAG
10	miR399e	ccttggcaccgagaattccaTTAATTTGAAGAGGCTCTAG
11	miR399f	ccttggcaccgagaattccaGTTAGAACTTAGAATCGTTCG
12	miR400	ccttggcaccgagaattccaCTCTACCTTACCATAATCAC
13	mir402	ccttggcaccgagaattccaGACTCTTTTCATGTGTATTC
14	miR403	ccttggcaccgagaattccaTAGATCTTGTTTACGATTCC
15	miR408	ccttggcaccgagaattccaTGAATGACAGAGAGGTAGAC

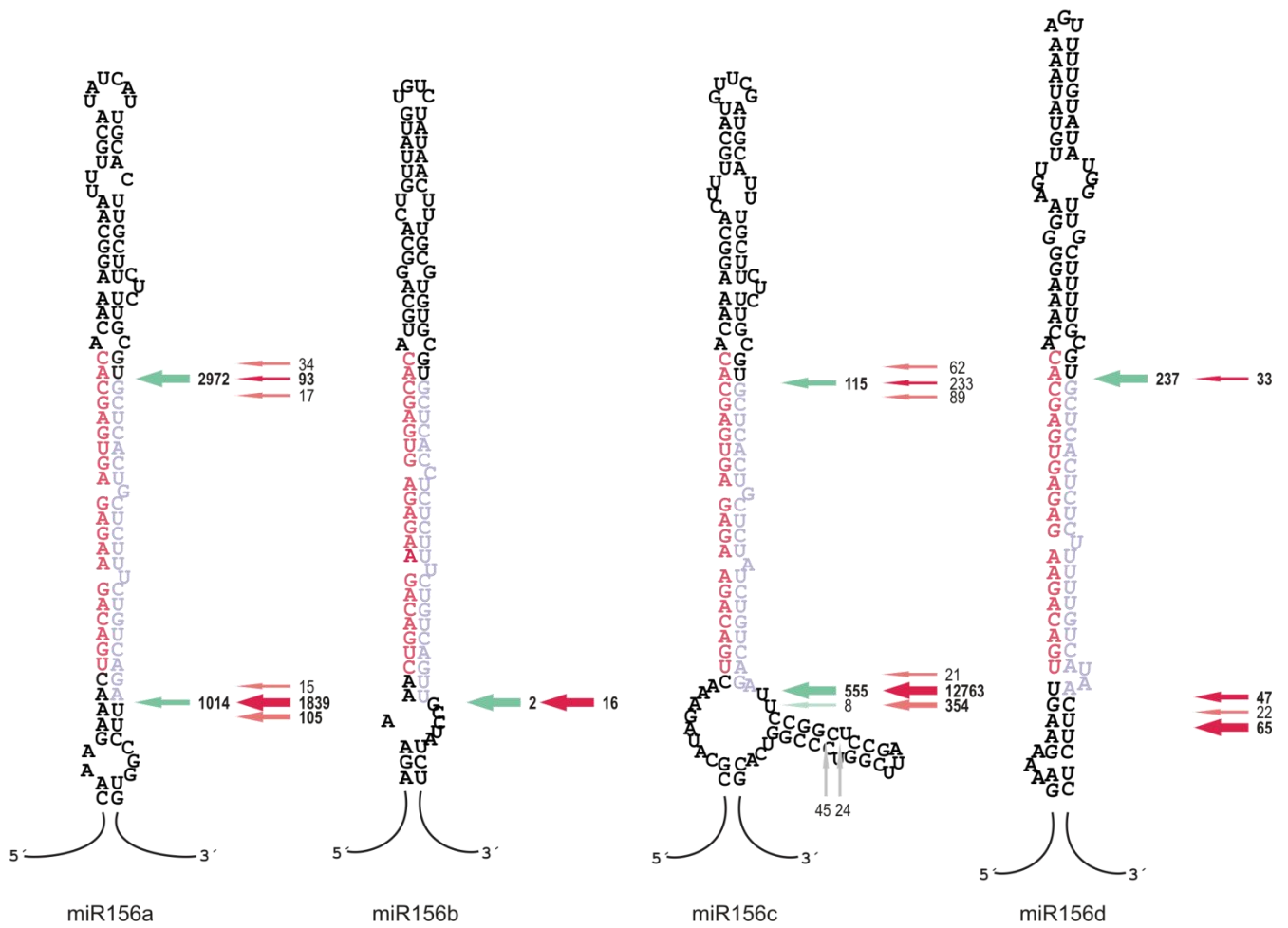
Mix 8		
n° MIX	miARN	secuencia
1	mir398c.2	ccttggcaccgagaattccaCTACGTTTGTACTACTTTGTG
2	miR447a	ccttggcaccgagaattccaATCTATGATATCGATGCAAC
3	miR447b	ccttggcaccgagaattccaTGAATCTATGATATCGATGC
4	miR472	ccttggcaccgagaattccaACTGAAAGTCTAGCGACTAG
5	miR771	ccttggcaccgagaattccaTATCTTGACCATGGAGACAG
6	miR779	ccttggcaccgagaattccaTCTCATCTCGAACGGAATGC
7	miR824	ccttggcaccgagaattccaAAAAGGCAACAAAGTCACTG
8	miR825	ccttggcaccgagaattccaAATCCATATAGTTCTCTAGC
9	miR827	ccttggcaccgagaattccaTTCGATTTGCCAGGTGATGC
10	miR862	ccttggcaccgagaattccaCTGAACCGAGTGTATATGAG
11	miR864	ccttggcaccgagaattccaCGCTGCTGACTTCAATATAC
12	miR2111a	ccttggcaccgagaattccaGATATTCAGTCTTAAATATC
13	miR2111b	ccttggcaccgagaattccaCCACTTCGAATGACTAGACC
14	miR5023	ccttggcaccgagaattccaTCCACAAGCTATGTCTATGC
15	miR5648	ccttggcaccgagaattccaGACTTTGAATGTTGCAGAAC
16	miR5657	ccttggcaccgagaattccaTCATGGTTCCCTTCTCGACC

ANEXO. Patrón de cortes de SPARE encontrado en los precursores analizados en bibliotecas de Col-0 y *fiery1*.

ANEXO A1

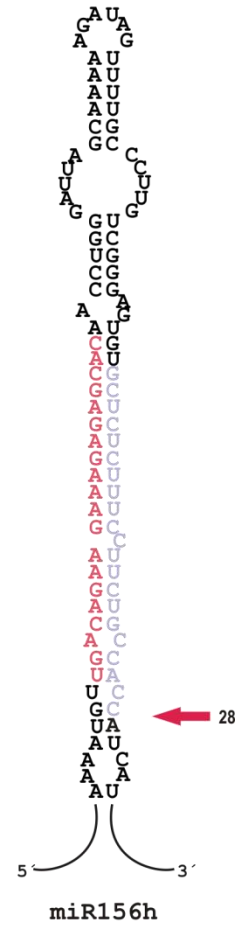
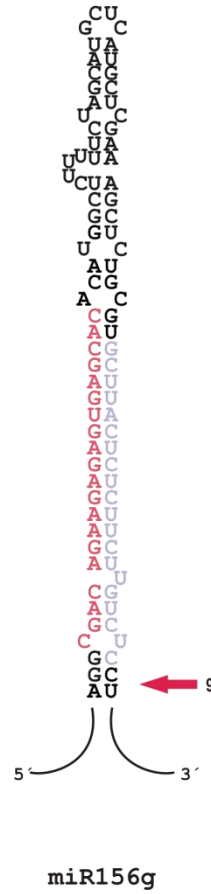
Col-0 *fiery1*

- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



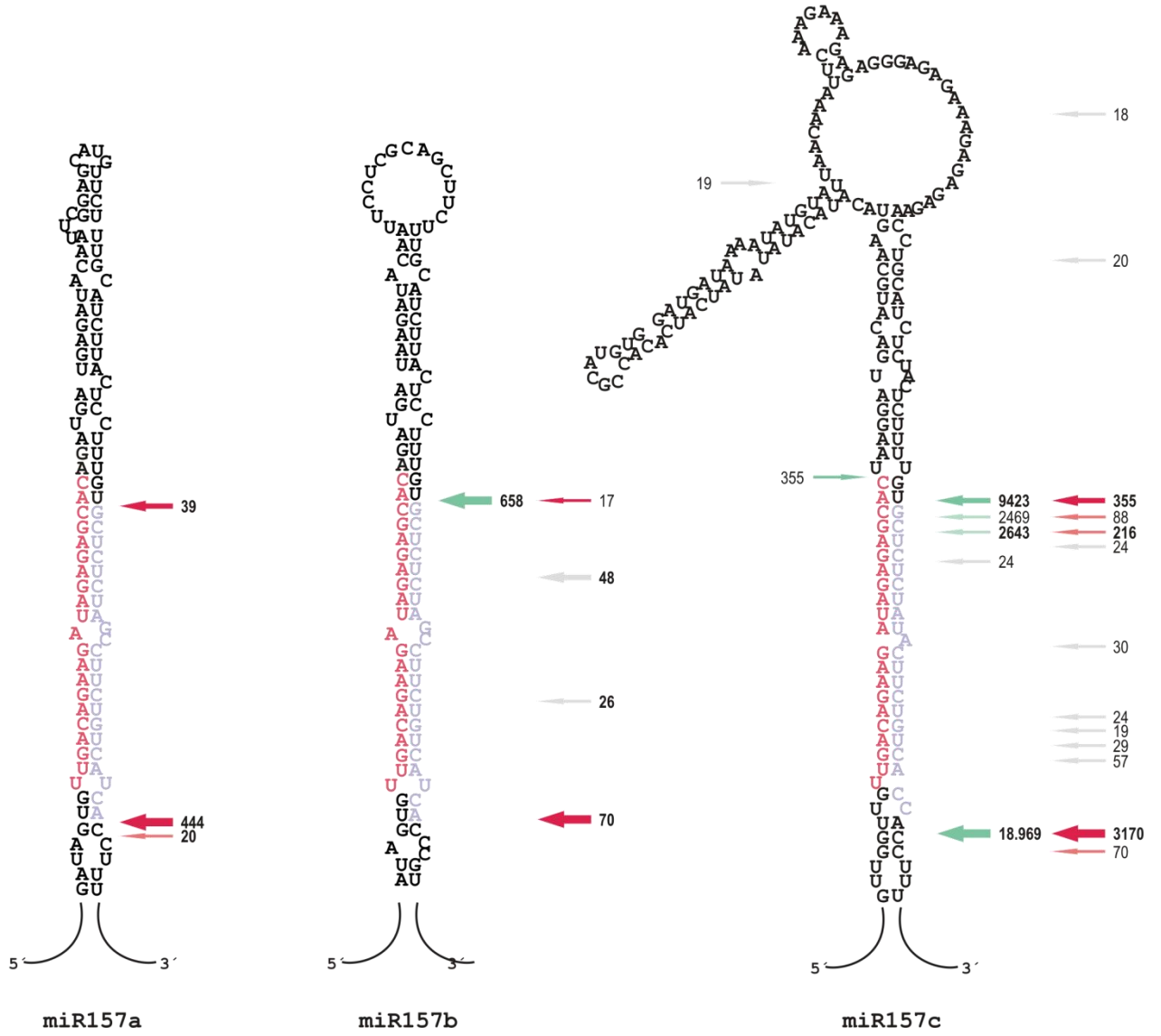
ANEXO A2

Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



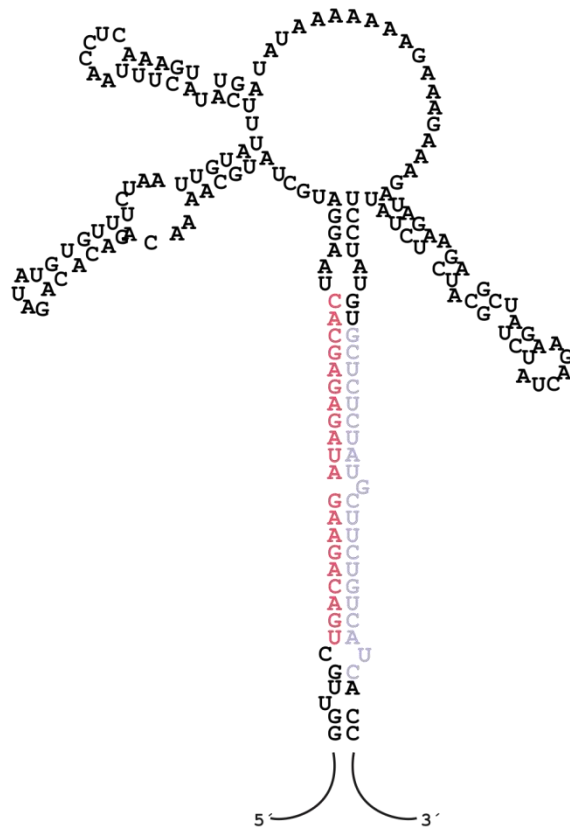
ANEXO A3

Col-0 *fiery1*
█ Cortes productivos
█ Cortes imprecisos
█ Cortes spureos



ANEXO A4

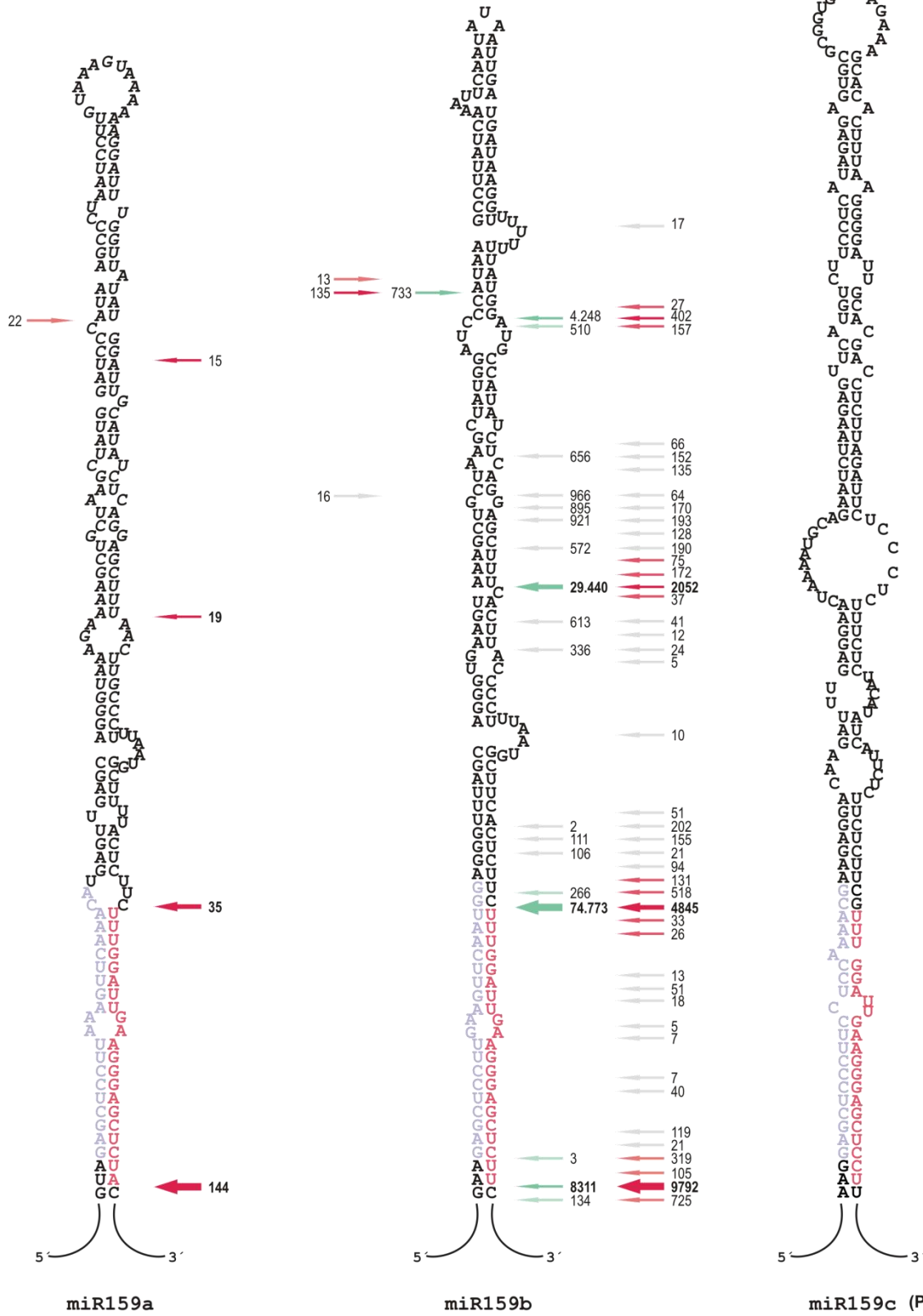
Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



miR157d

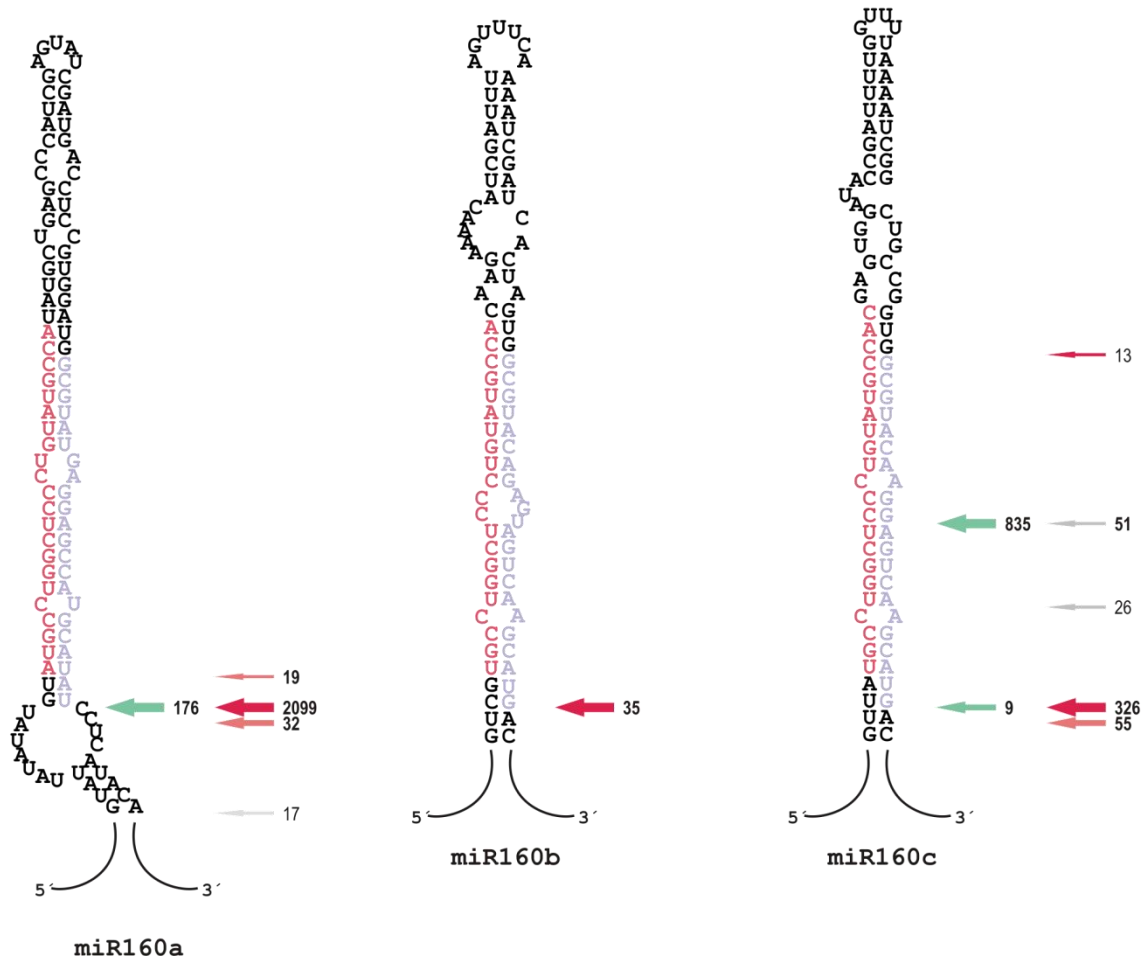
ANEXO A5

Col-0 *fiery1*
█ Cortes productivos
█ Cortes imprecisos
█ Cortes spureos



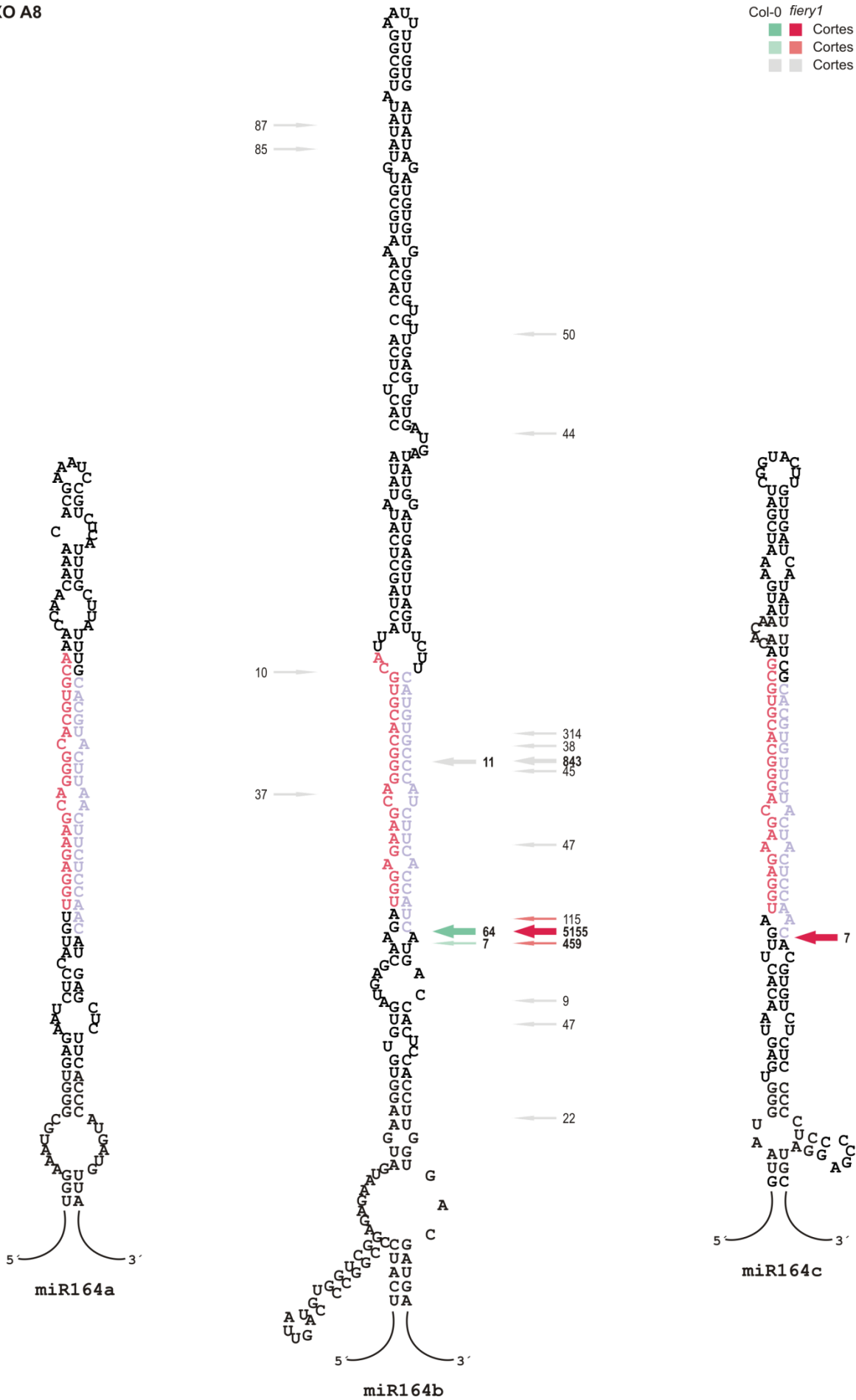
ANEXO A6

Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A8

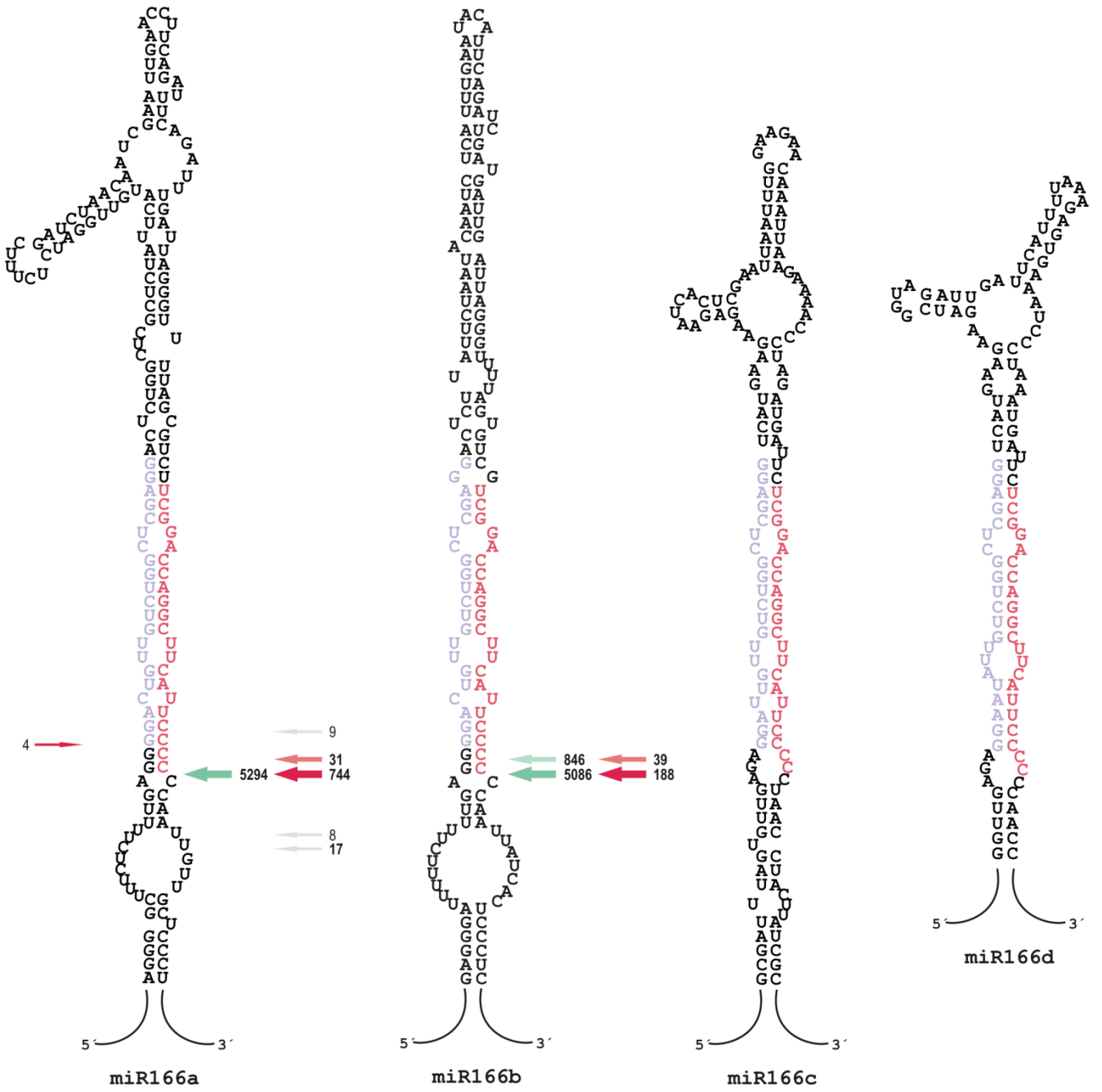
Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A10

Col-0 *fiery1*

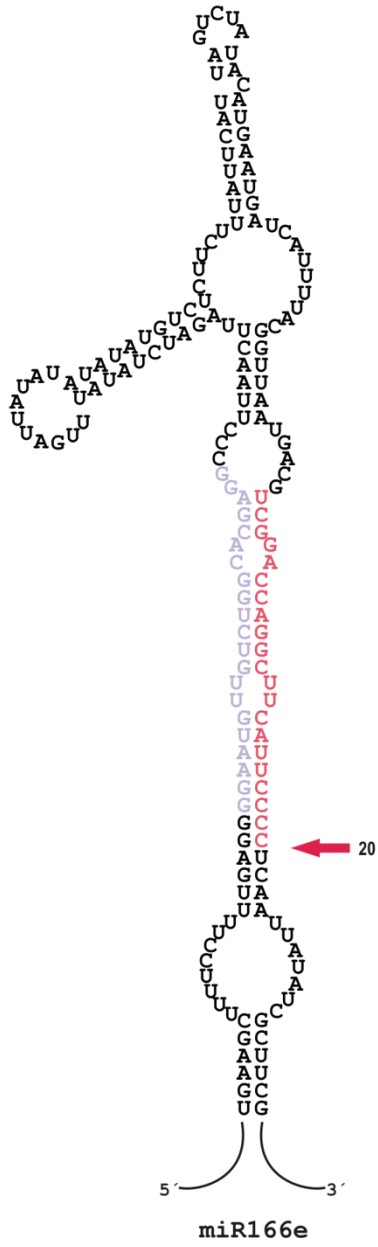
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A11

Col-0 *fiery1*

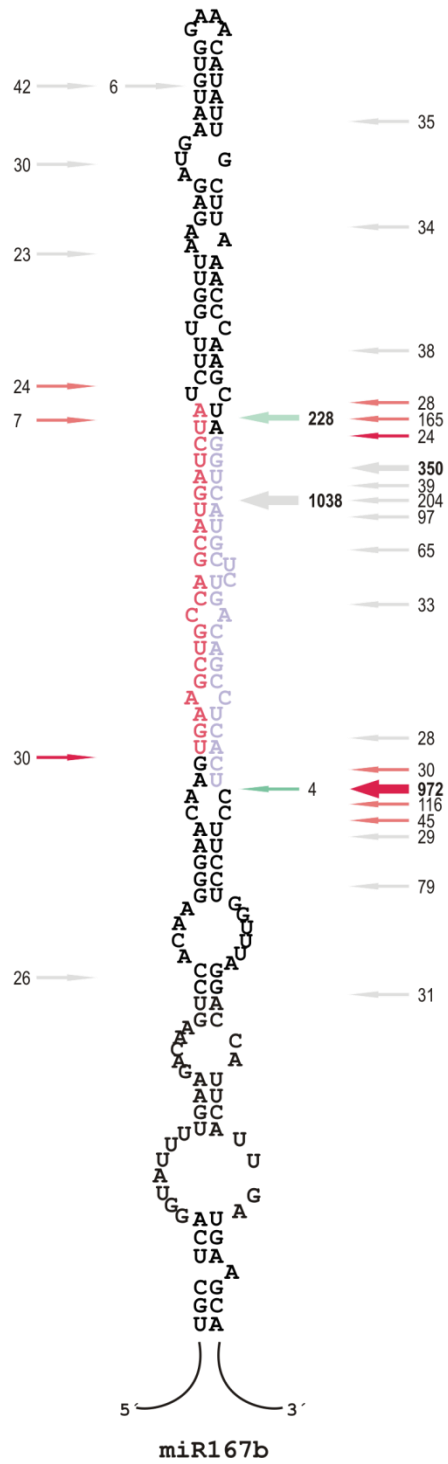
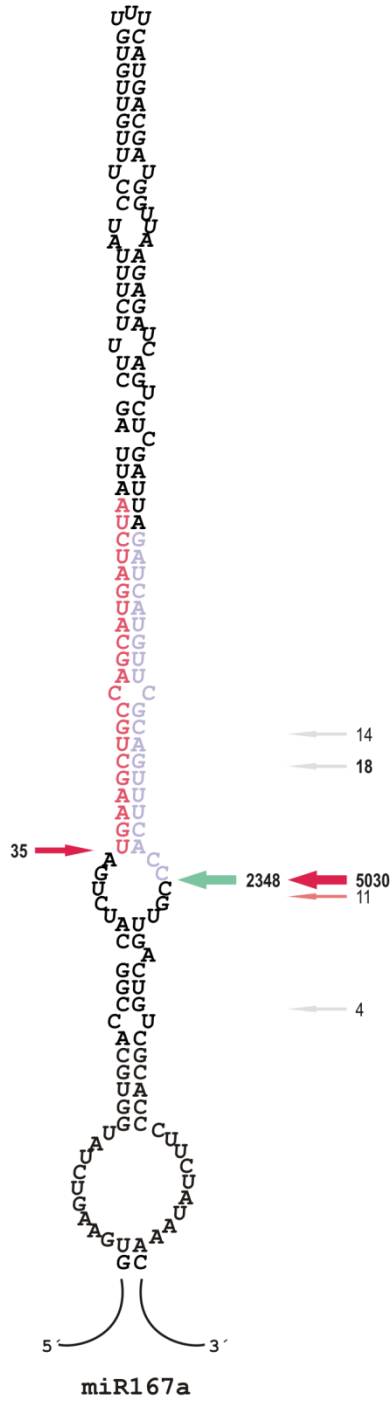
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A12

Col-0 *fiery1*

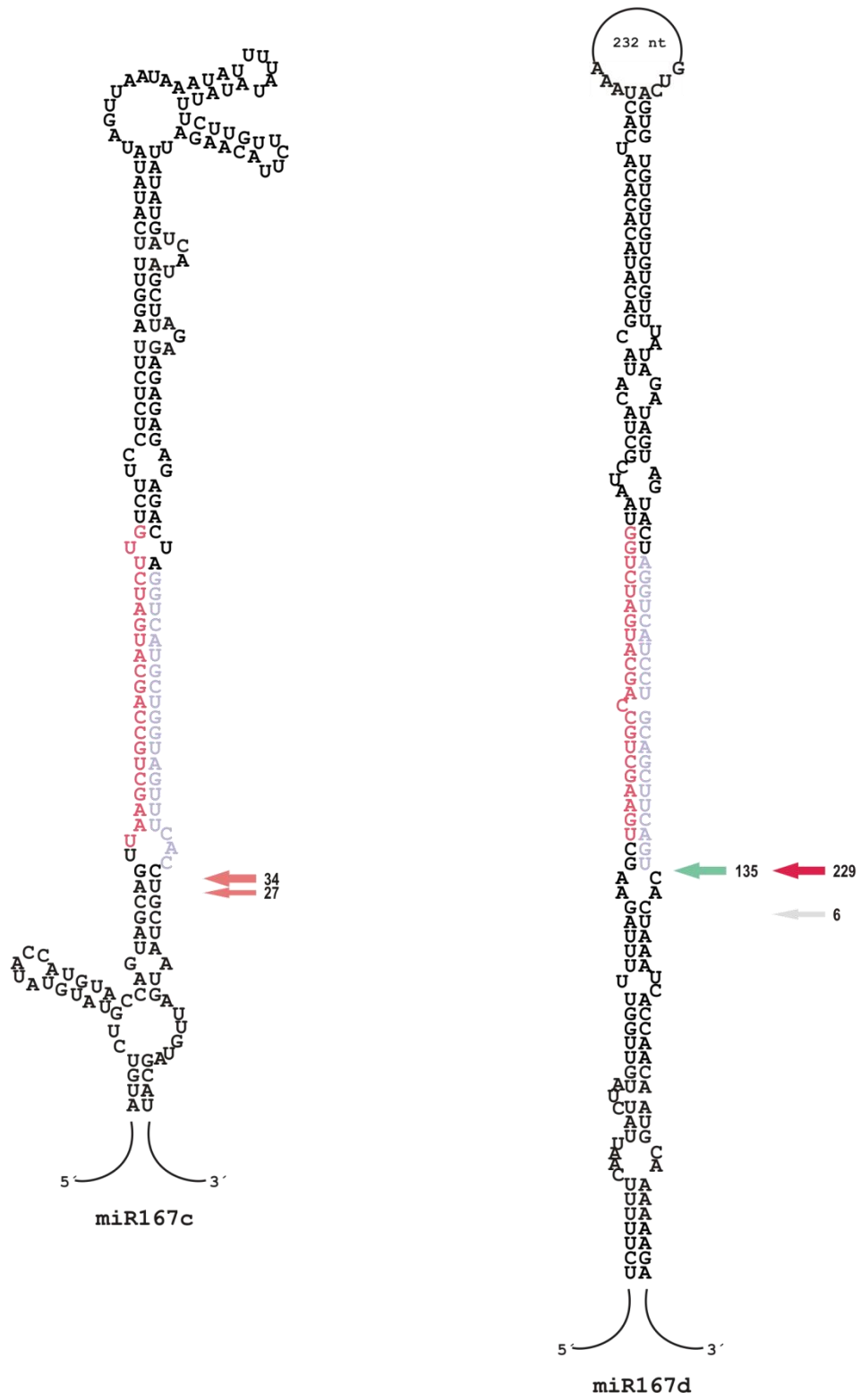
- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



ANEXO A13

Col-0 *fiery1*

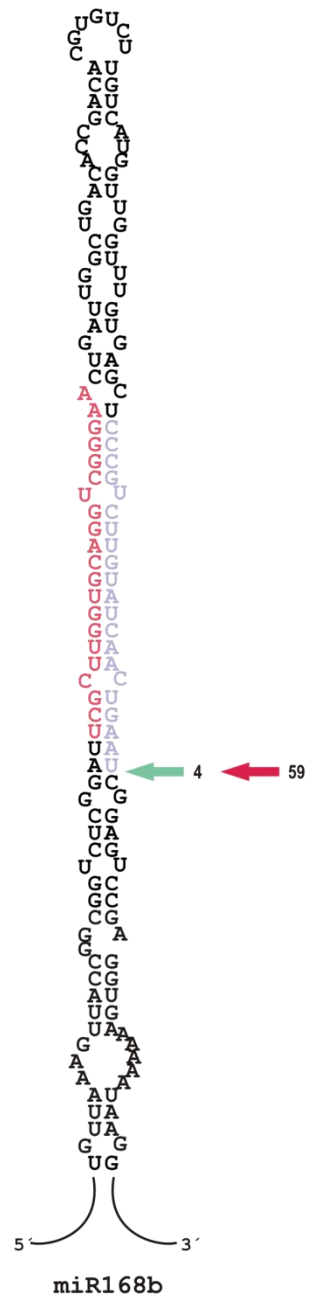
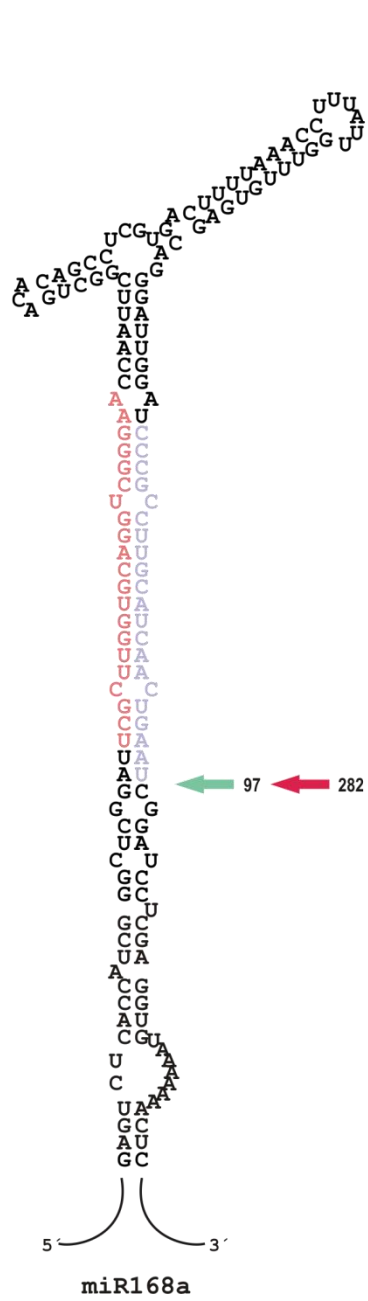
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A14

Col-0 *fiery1*

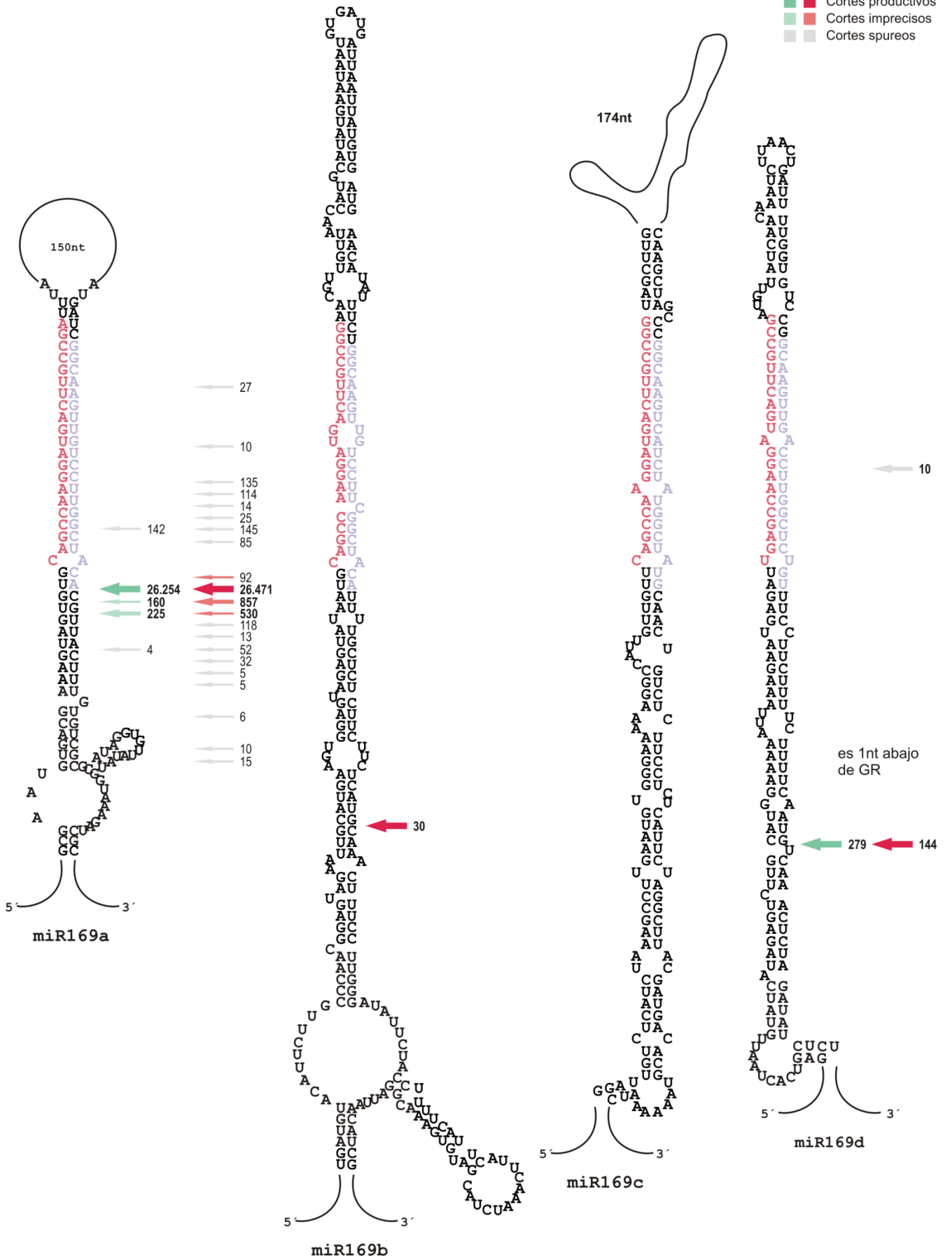
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A15

Col-0 *fiery1*

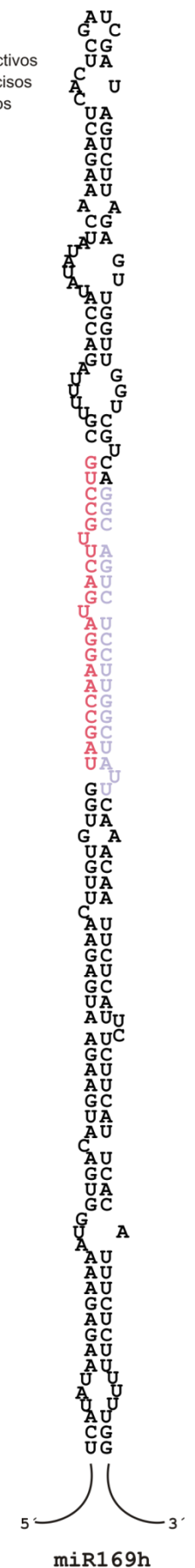
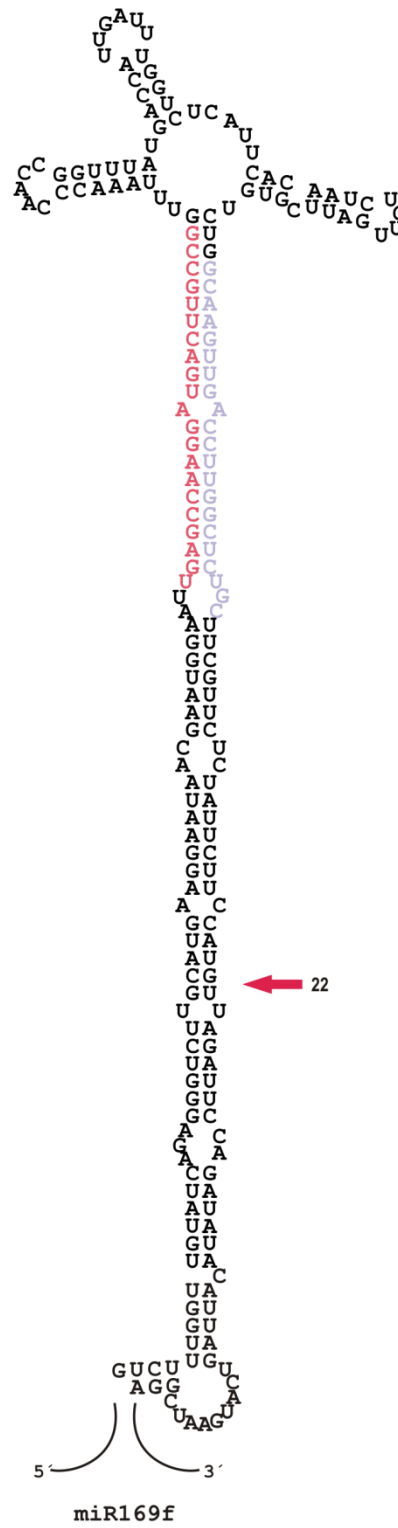
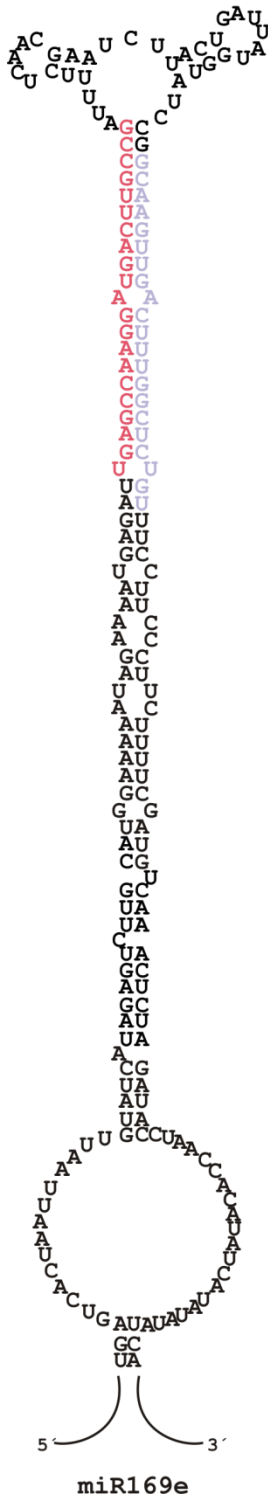
- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



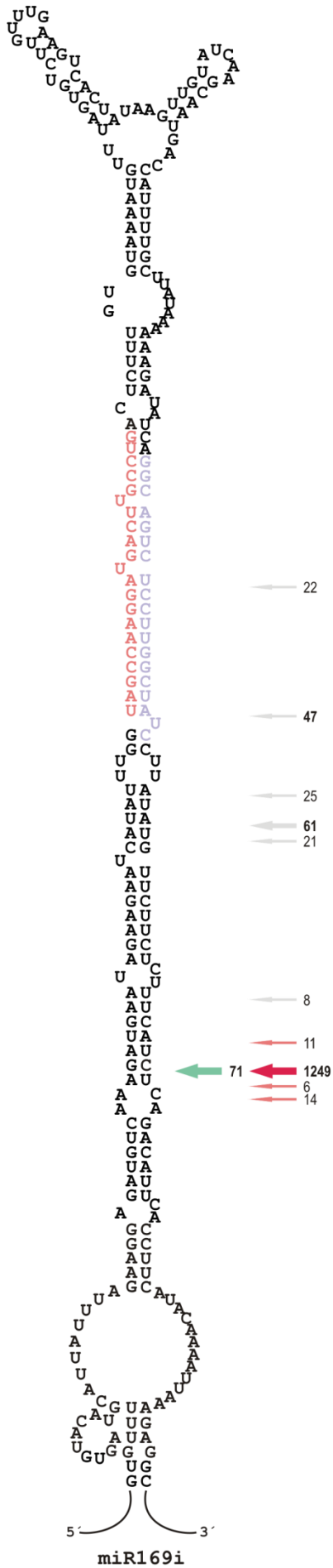
ANEXO A16

Col-0 *fiery1*

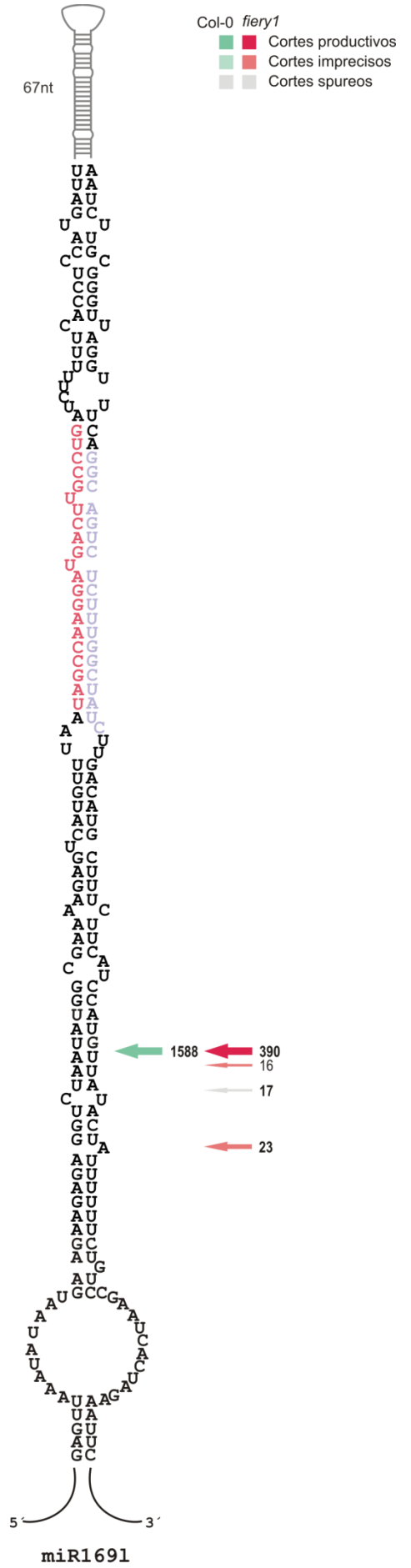
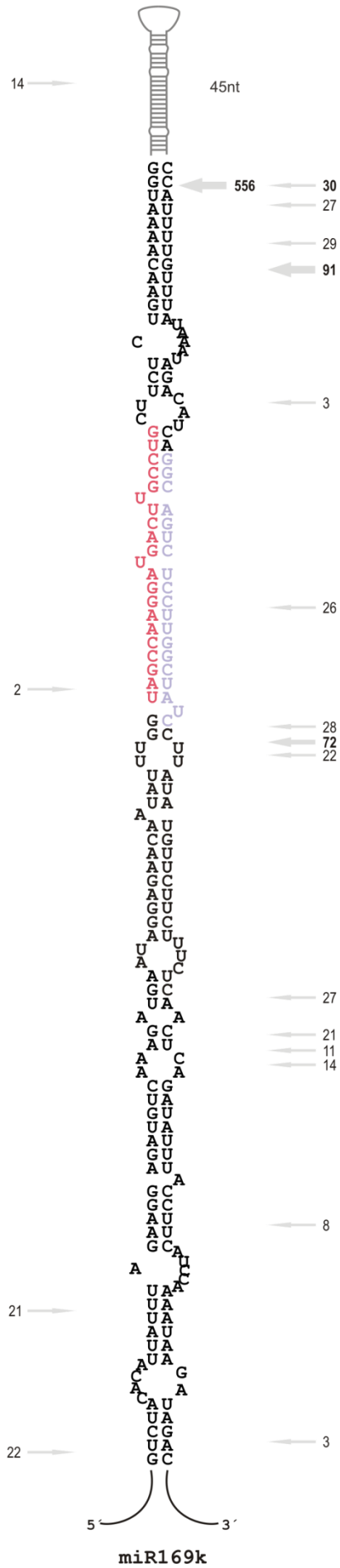
- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



ANEXO A17



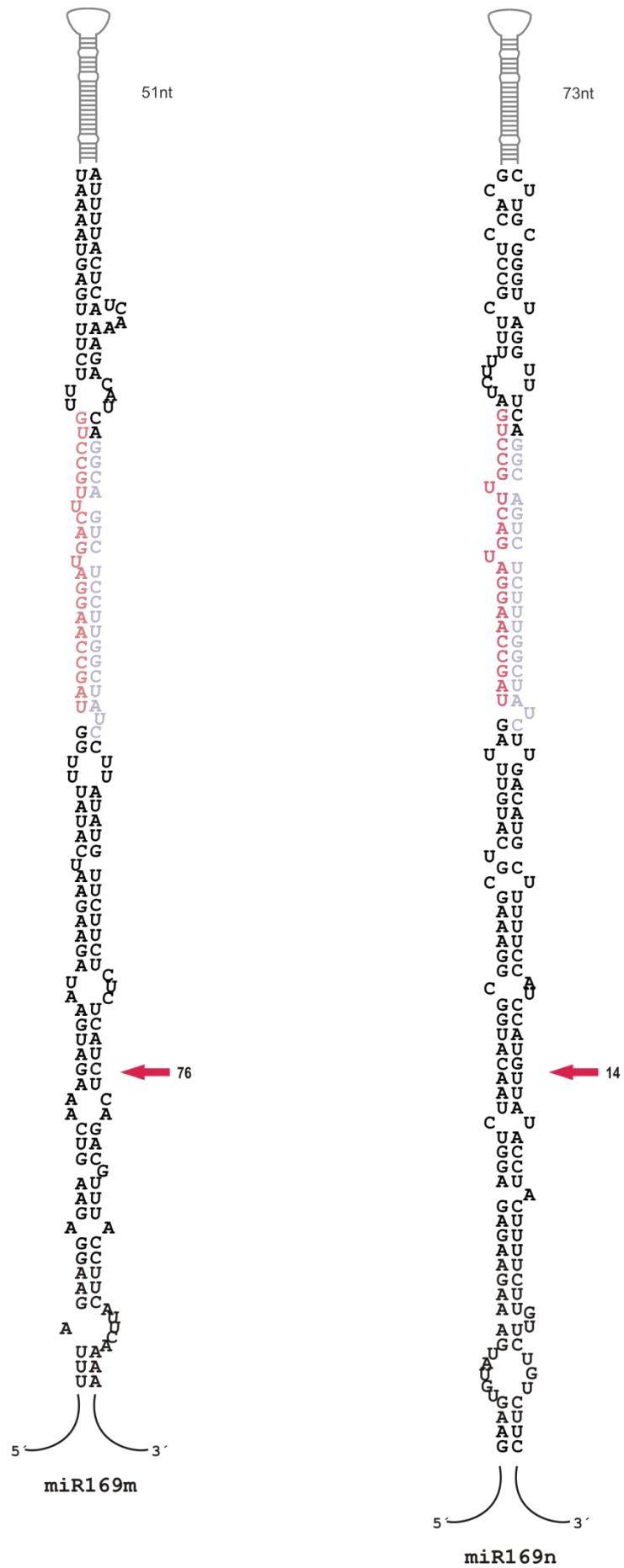
ANEXO A18



Col-0 *fiery1*

- Cortes productivos
- Cortes imprecisos
- Cortes spureos

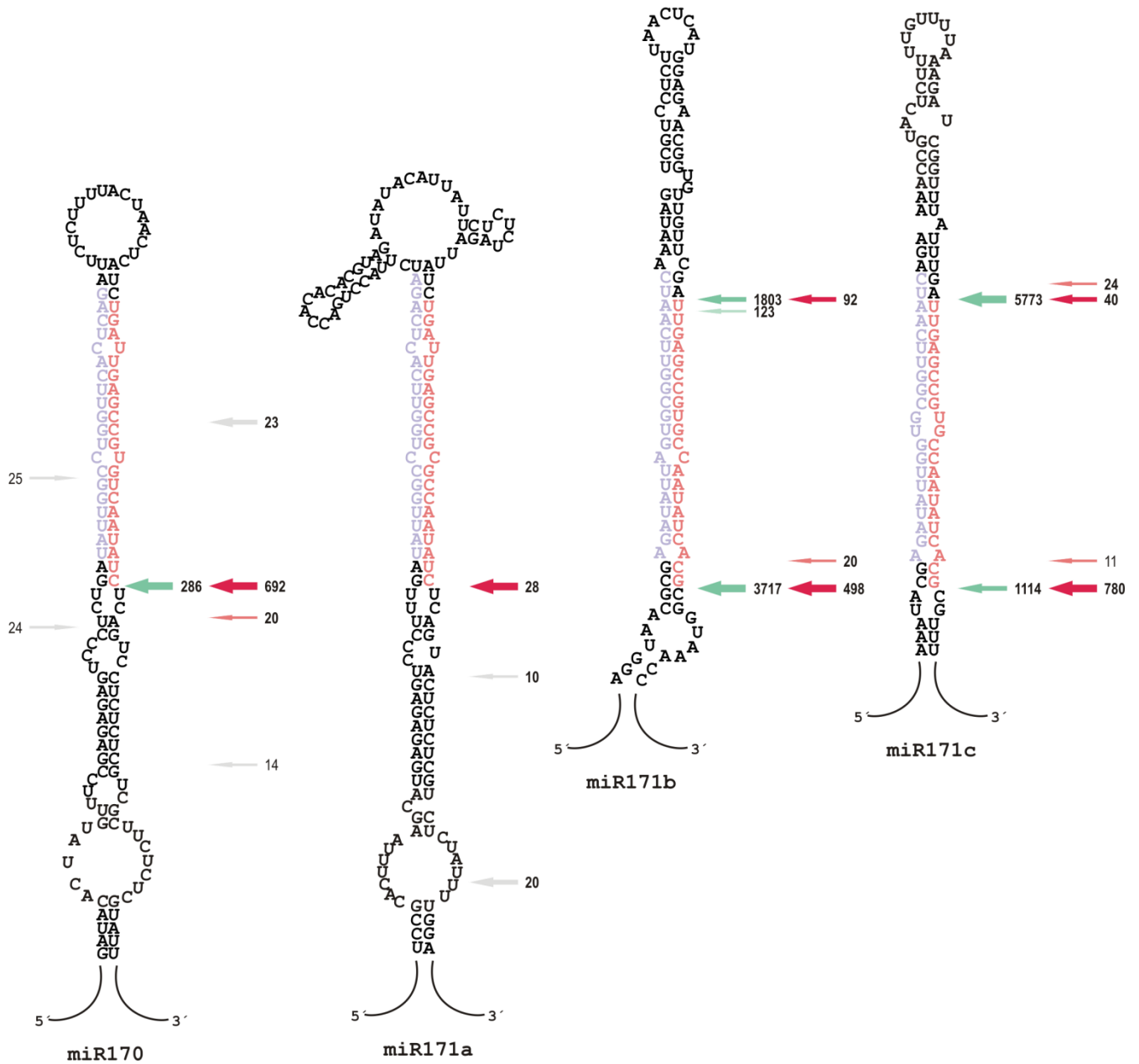
ANEXO A19



ANEXO A20

Col-0 *fiery1*

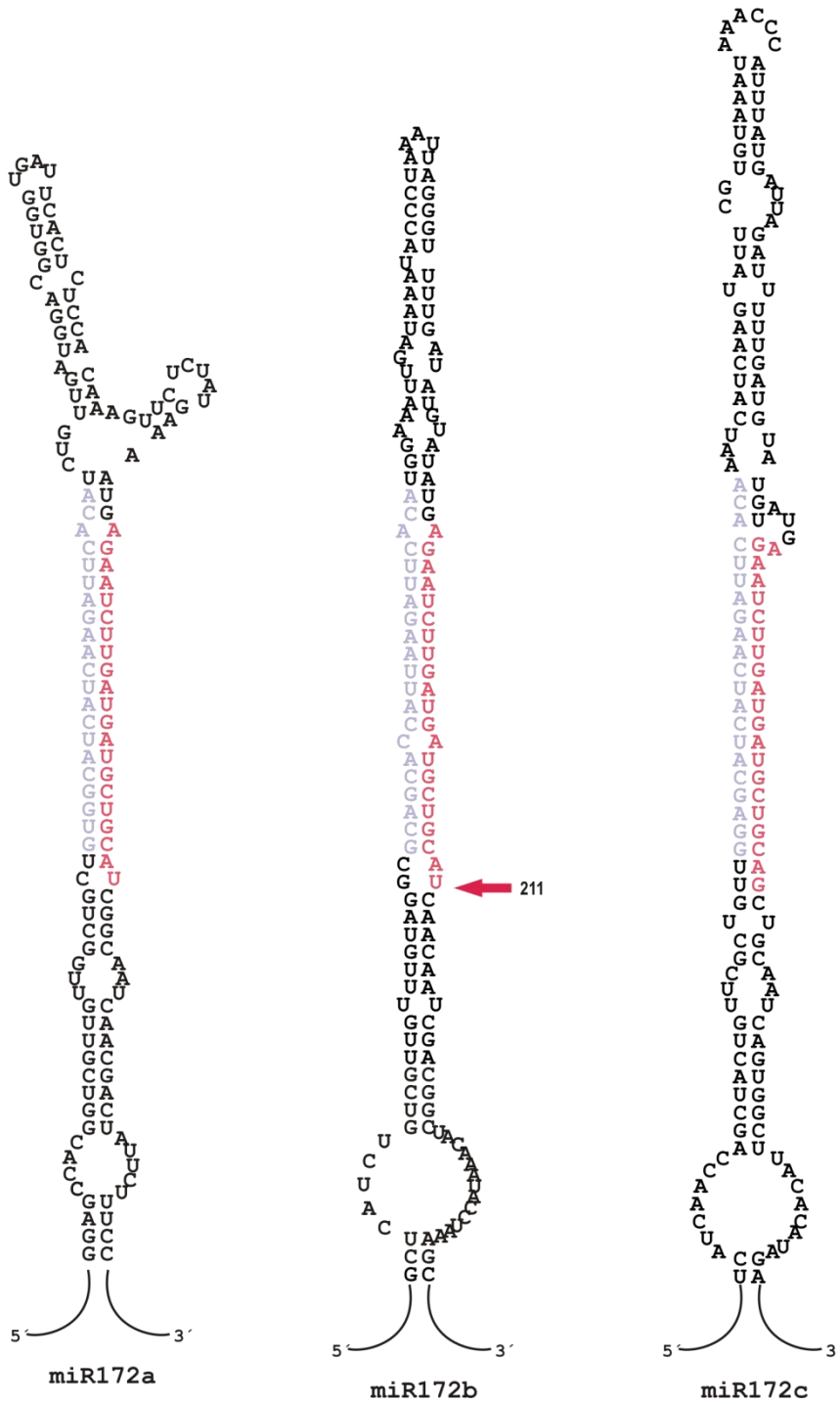
- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



ANEXO A21

Col-0 *fiery1*

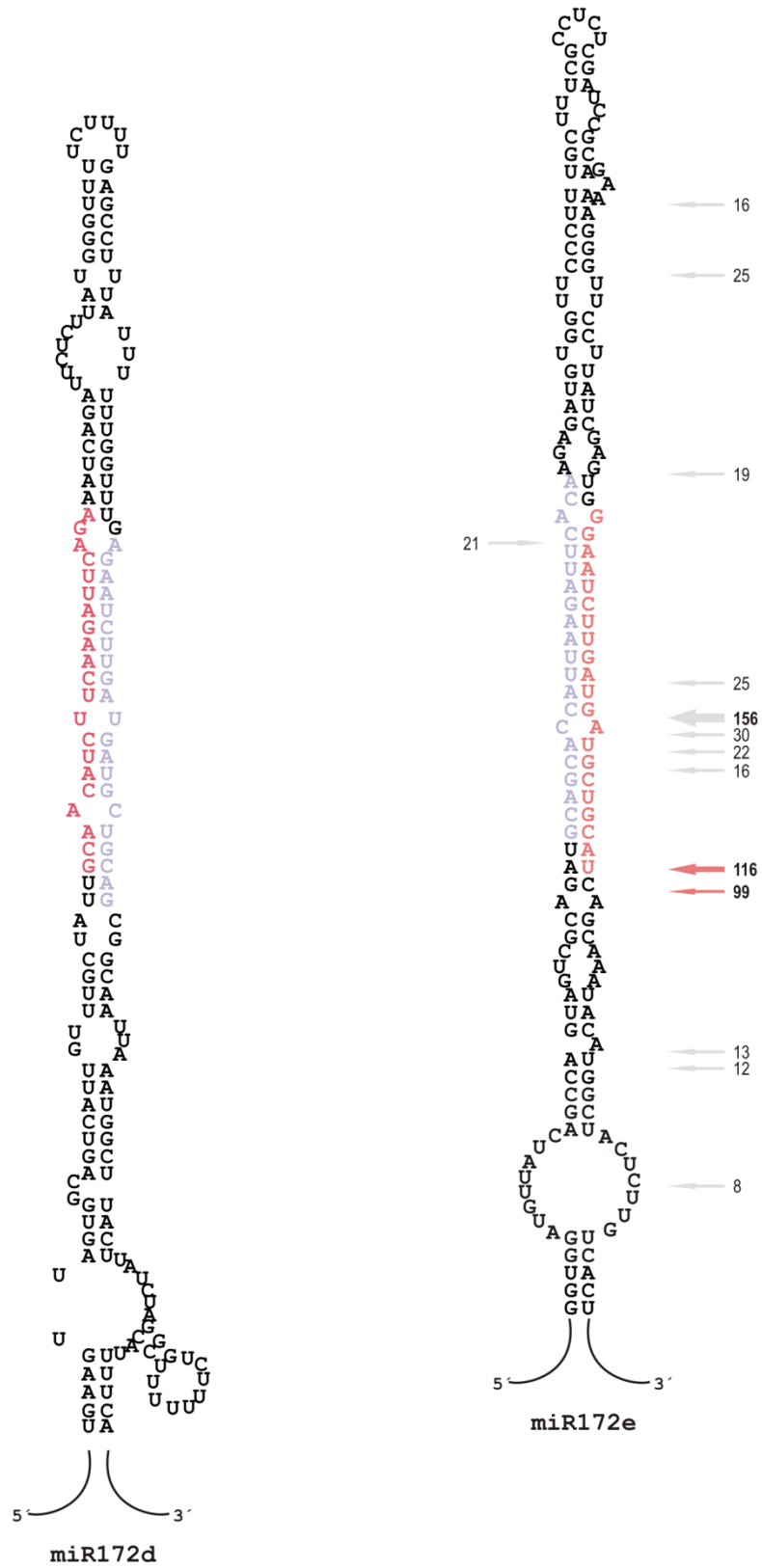
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



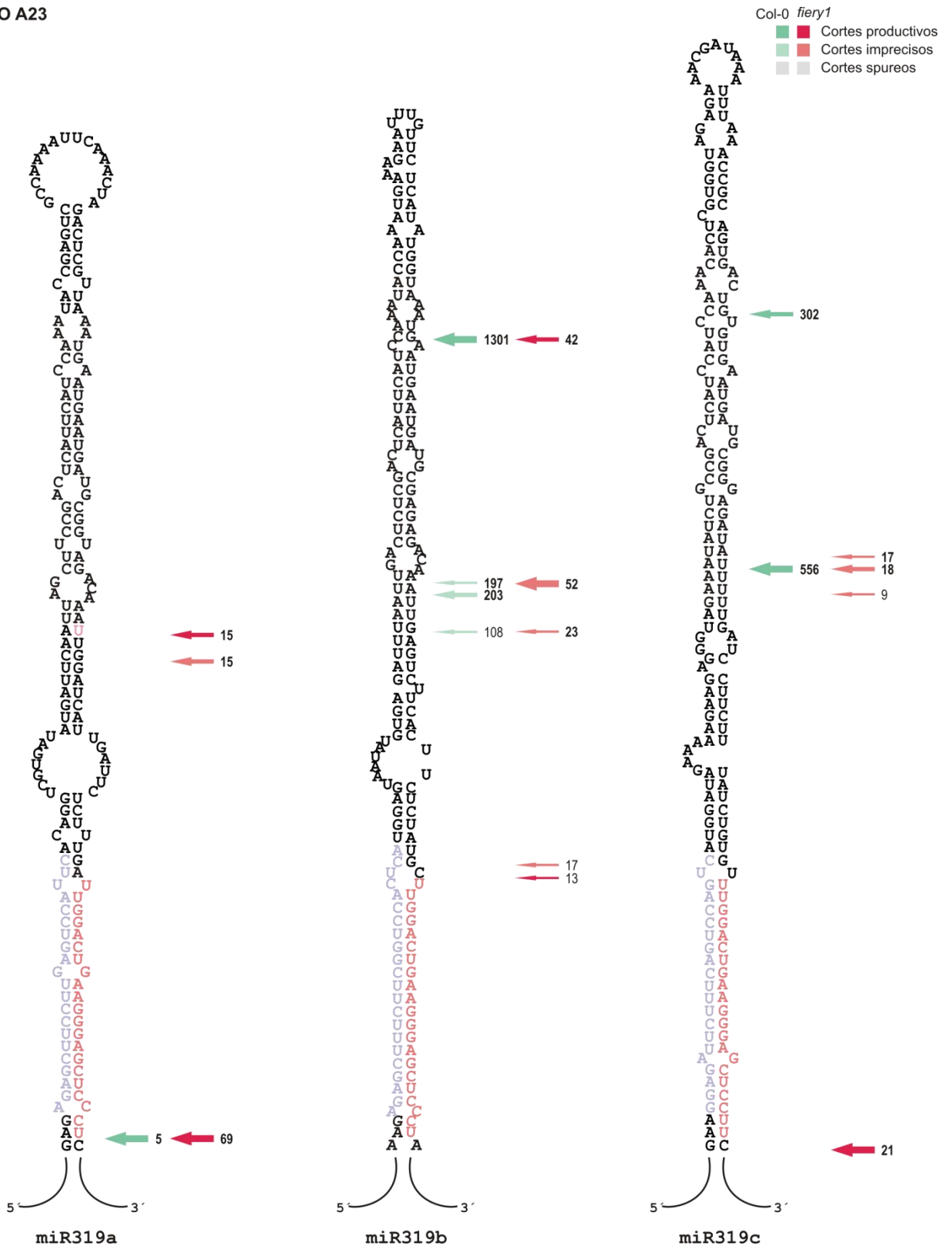
ANEXO A22

Col-0 *fiery1*

- Cortes productivos
- Cortes imprecisos
- Cortes spureos

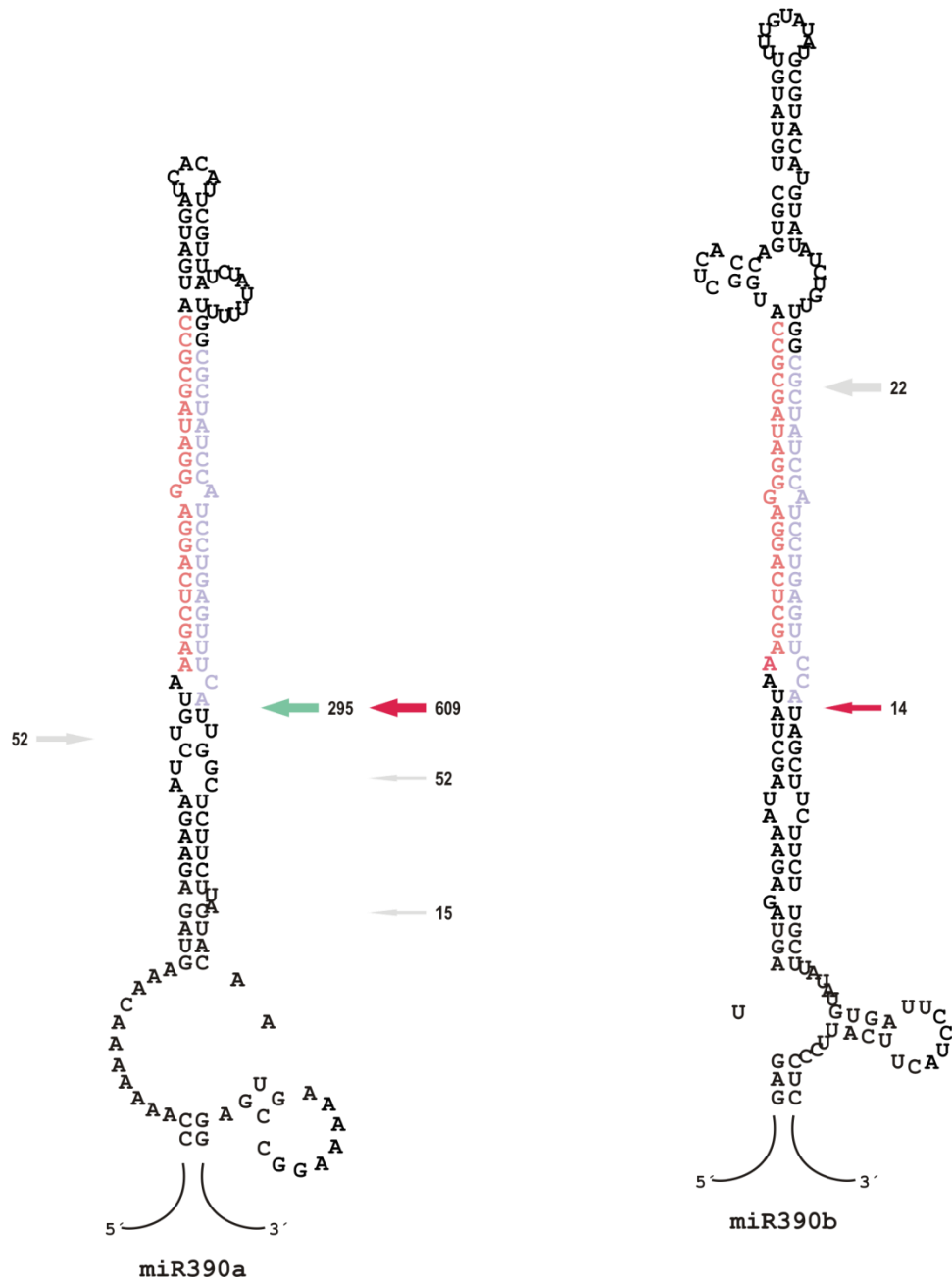


ANEXO A23



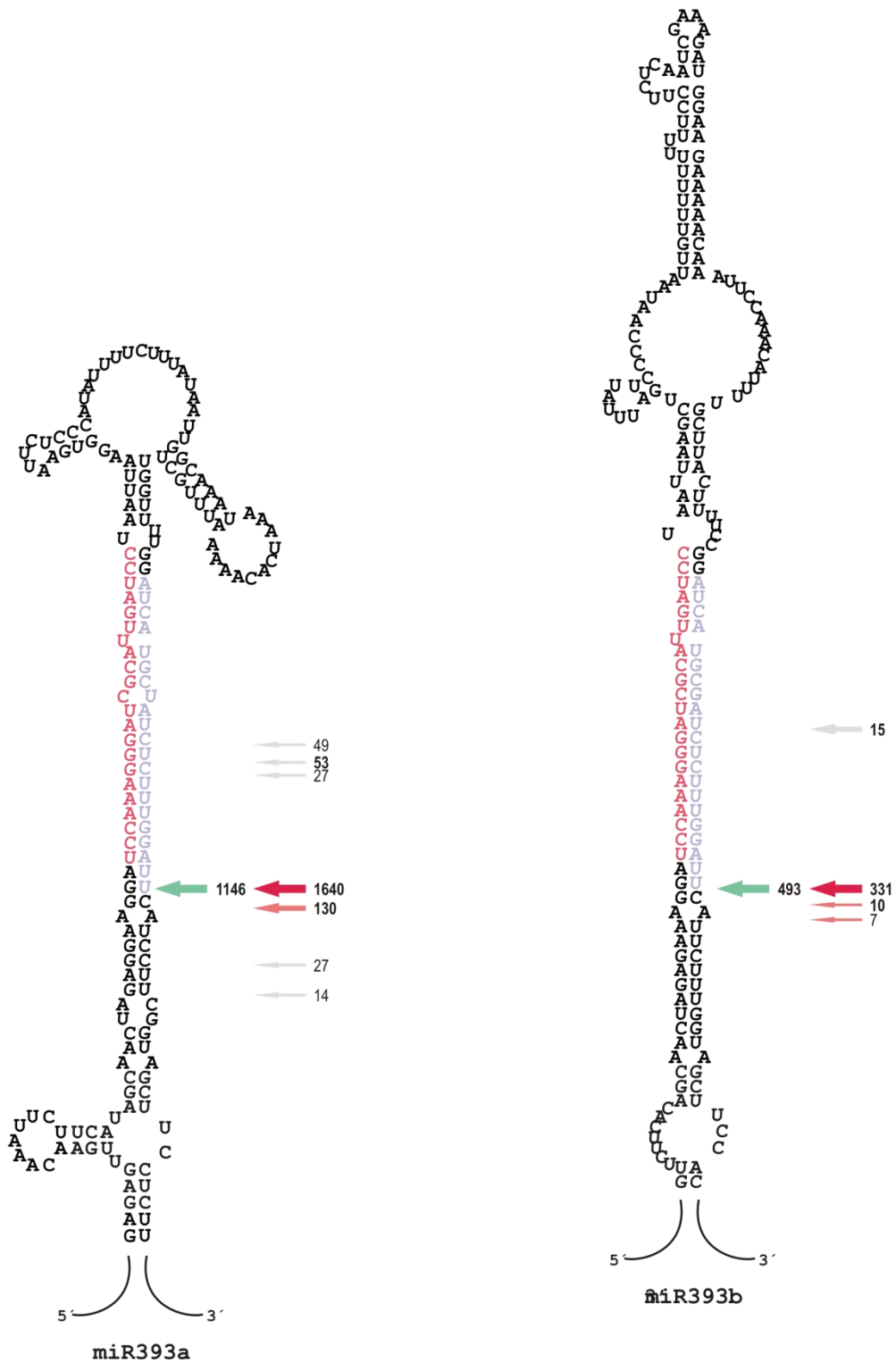
ANEXO A24

Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A26

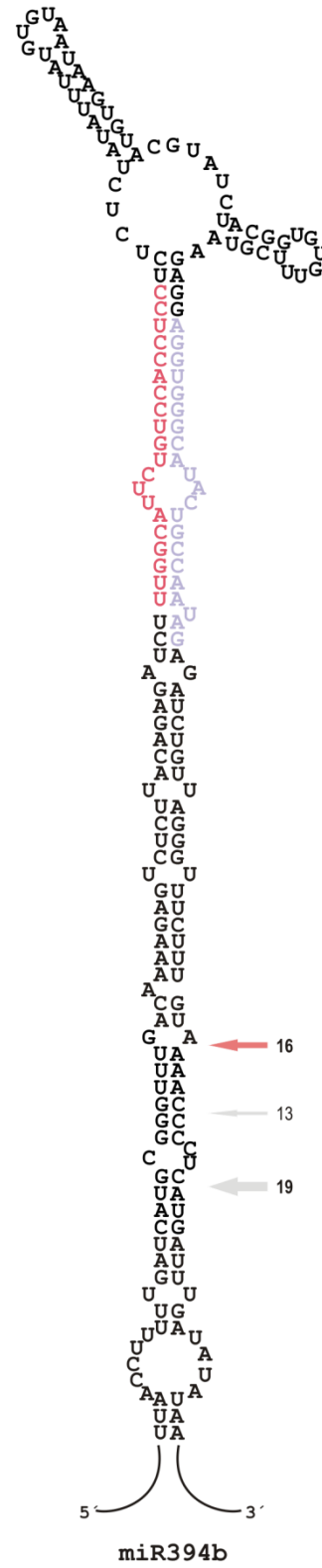
Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A27

Col-0 *fiery1*

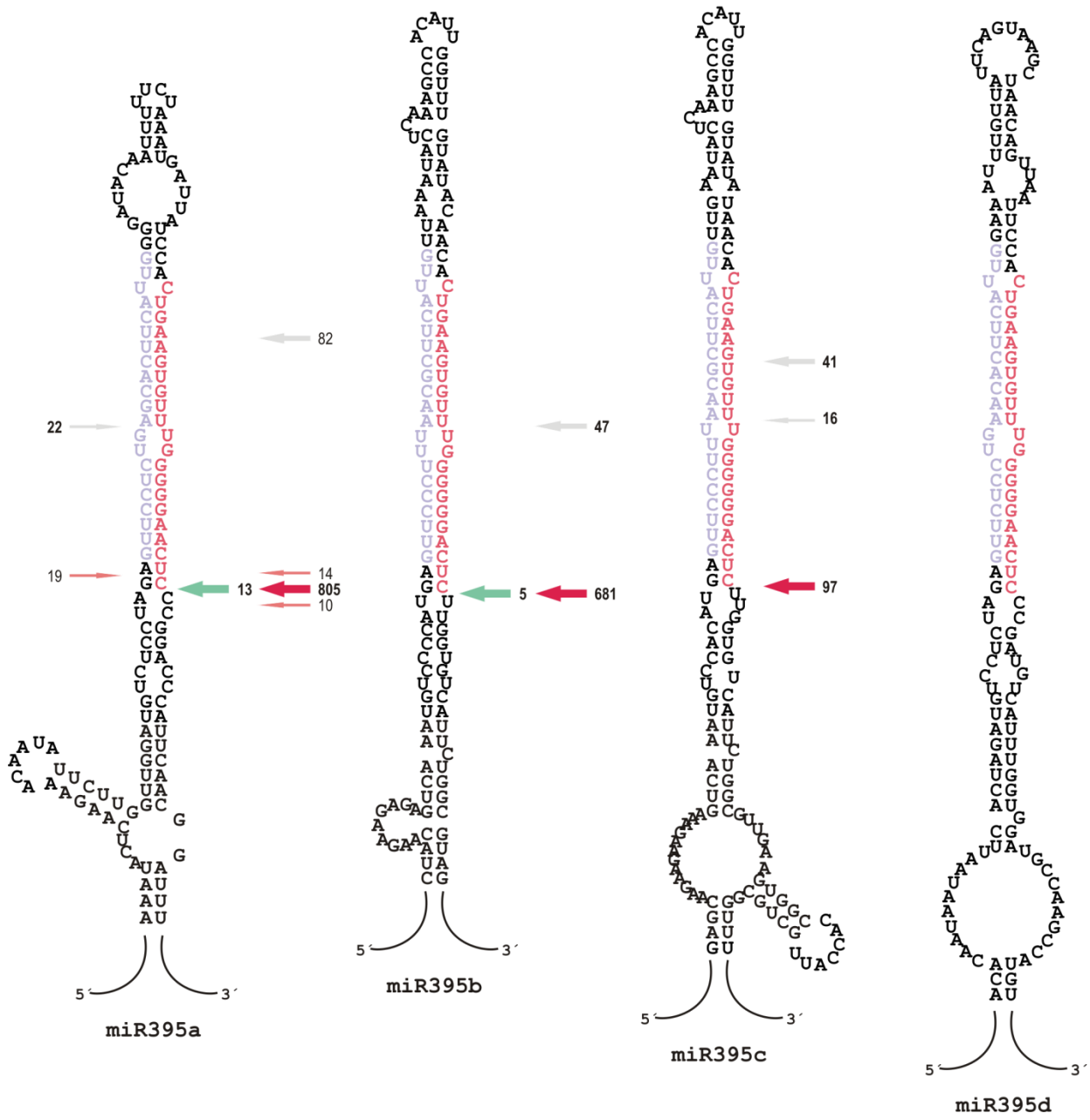
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A28

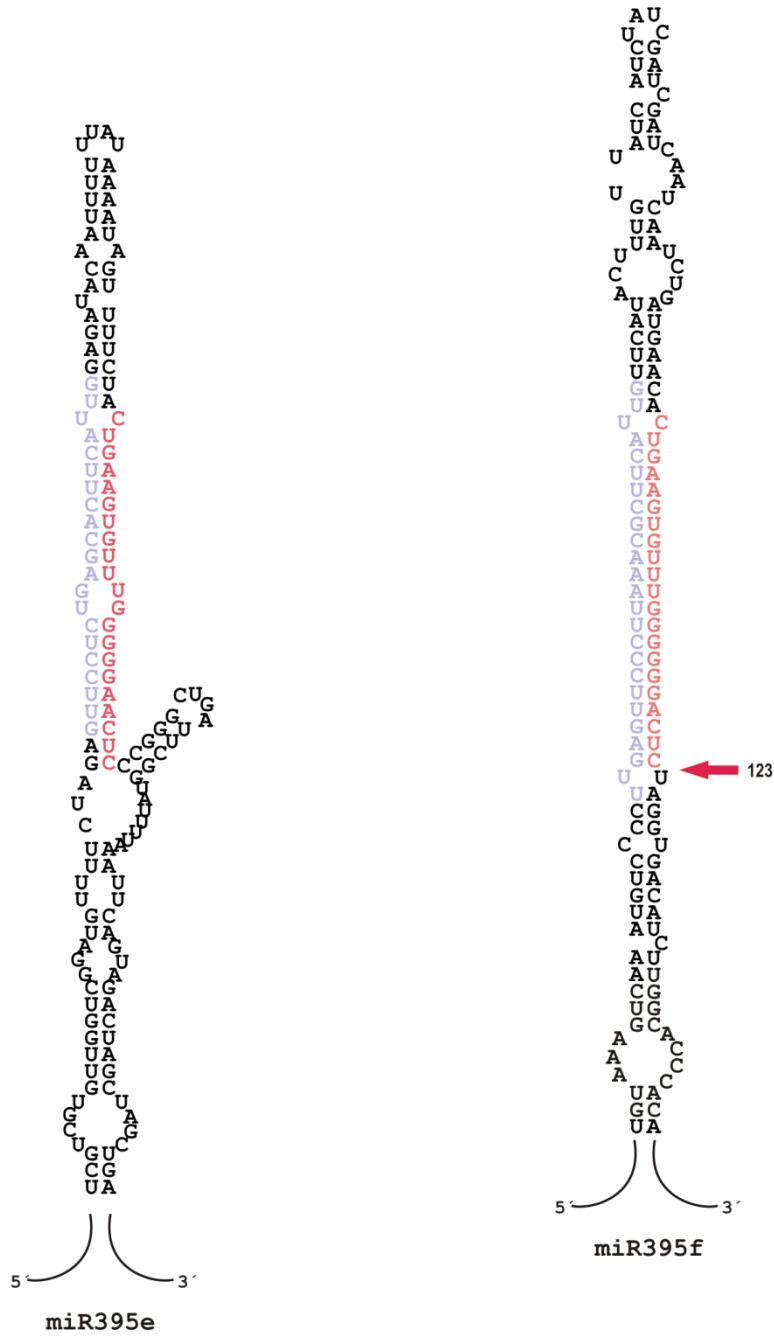
Col-0 *fiery1*

- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



ANEXO A29

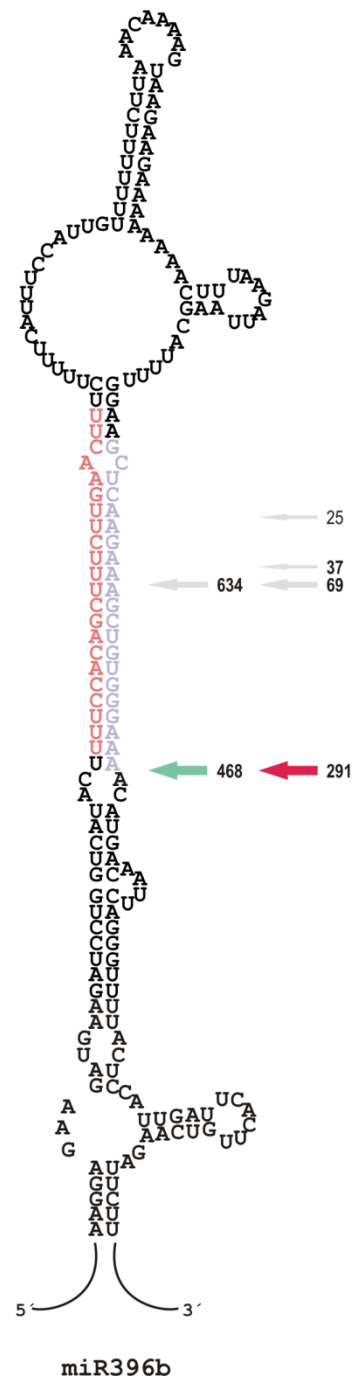
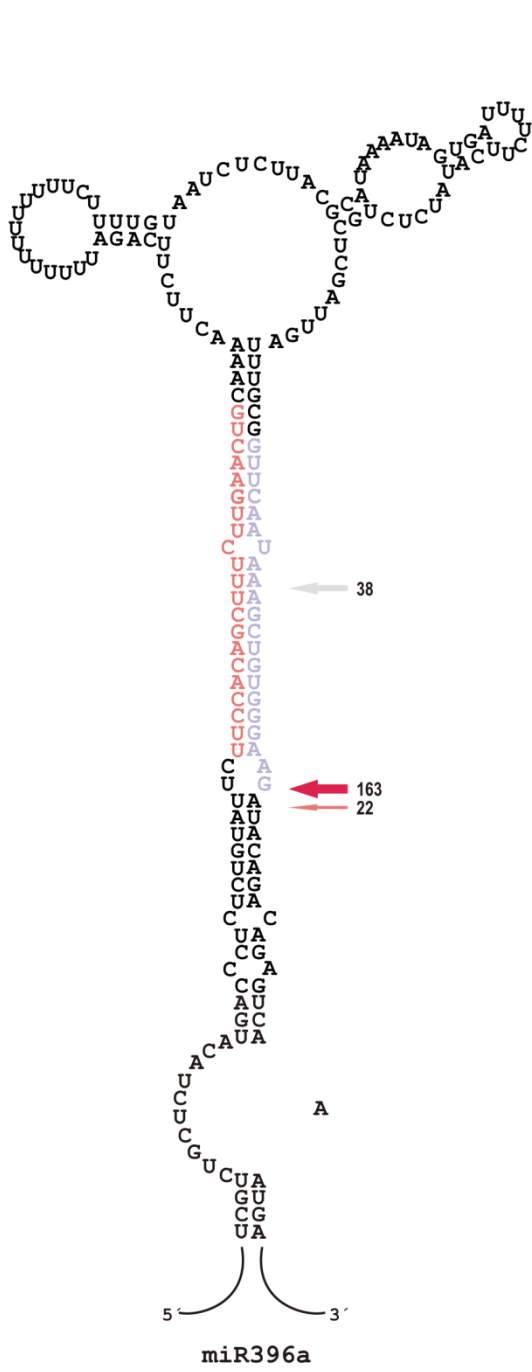
Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A30

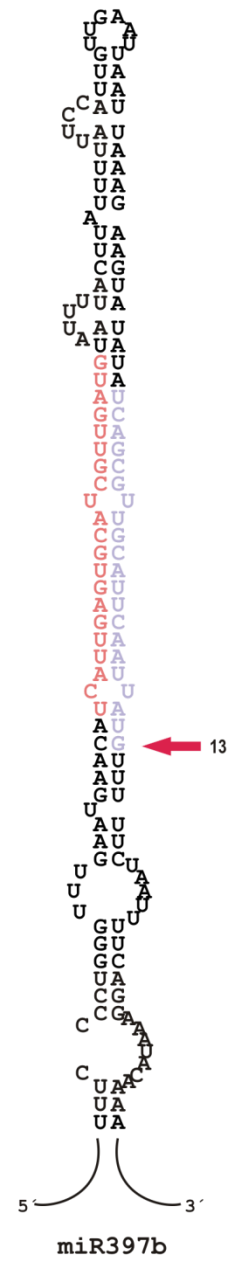
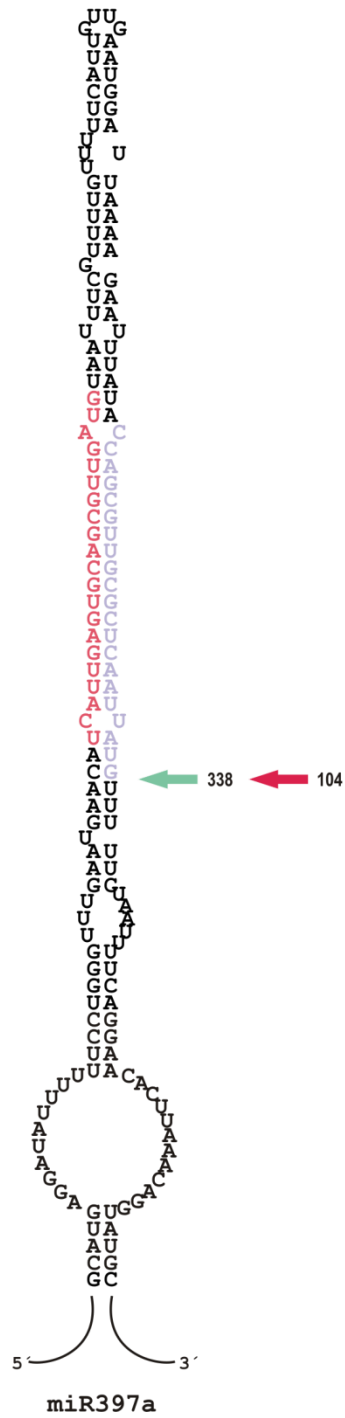
Col-0 *fiery1*

- Cortes productivos
- Cortes imprecisos
- Cortes spureos



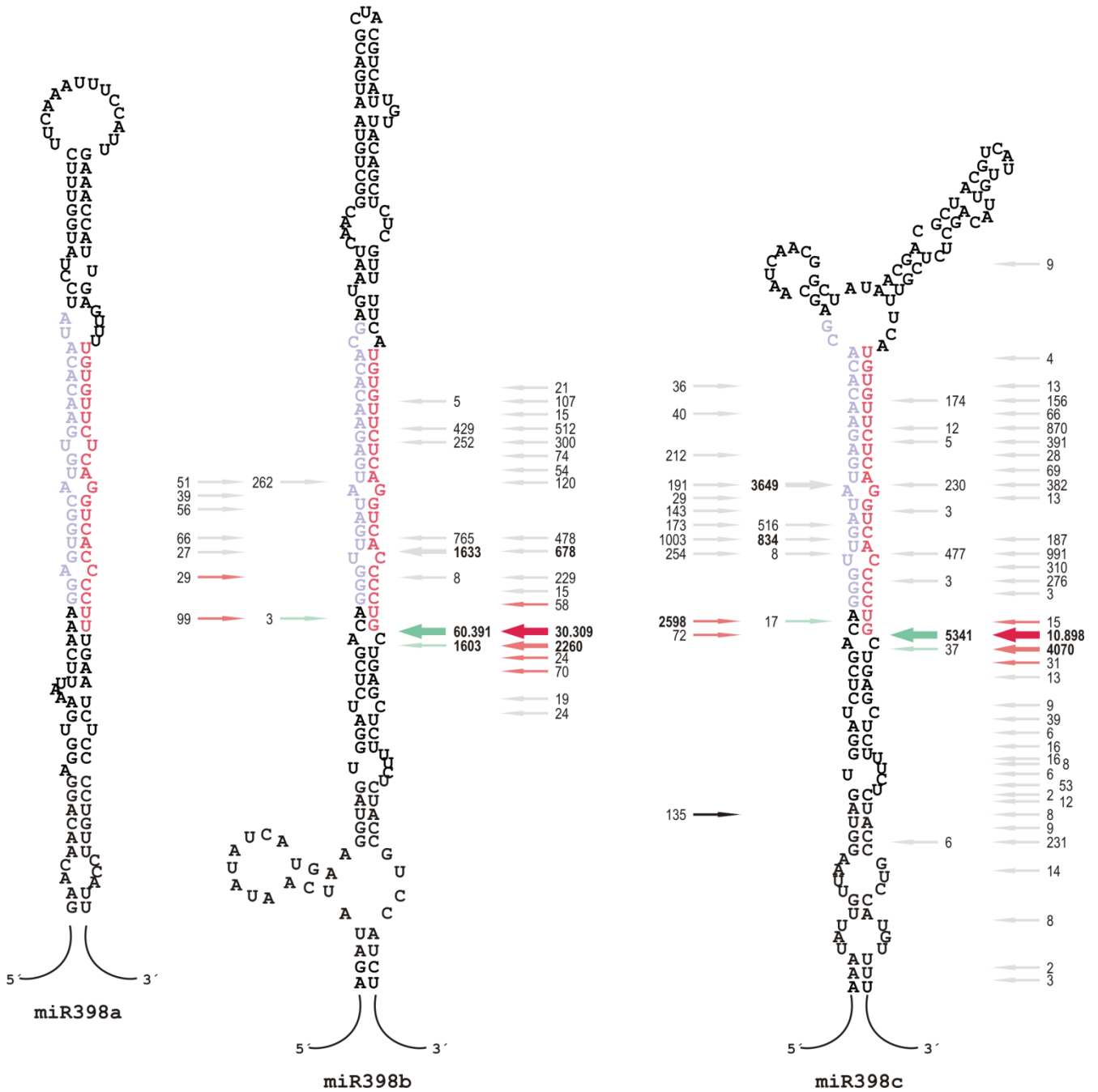
ANEXO A31

Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



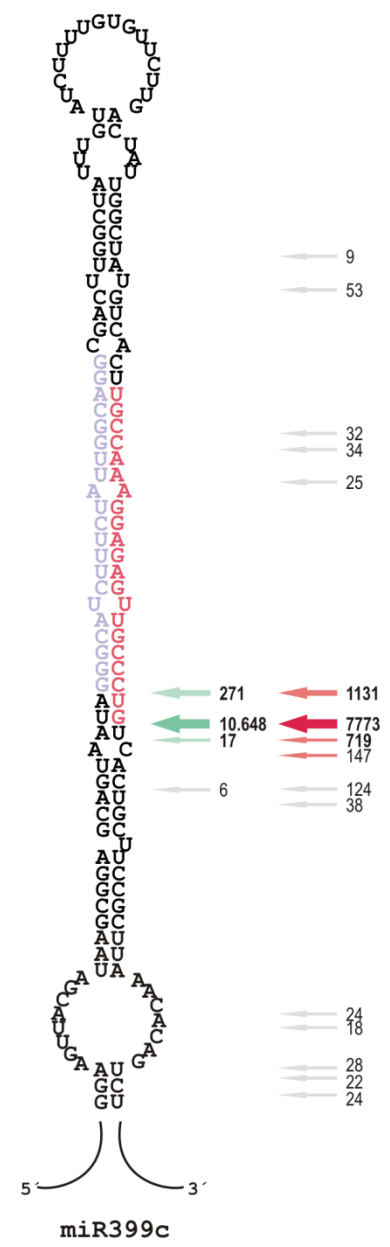
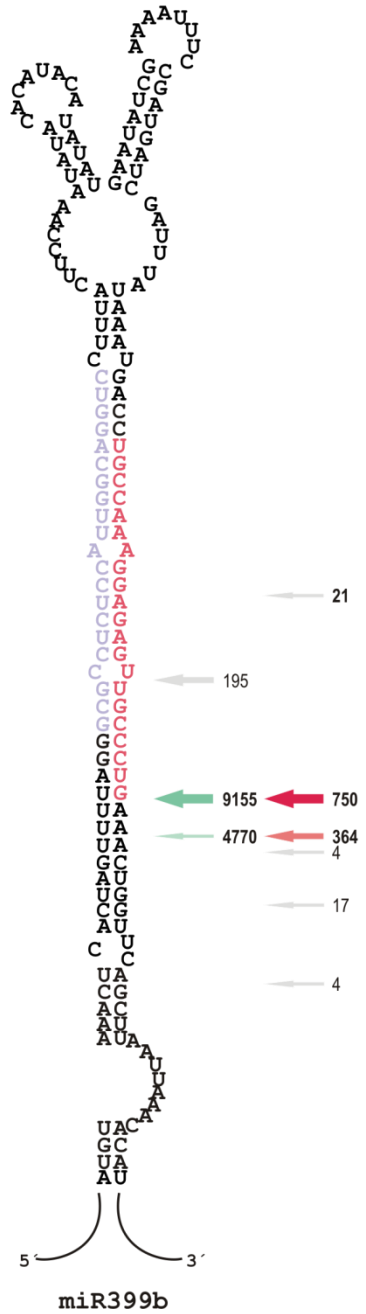
ANEXO A32

Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A33

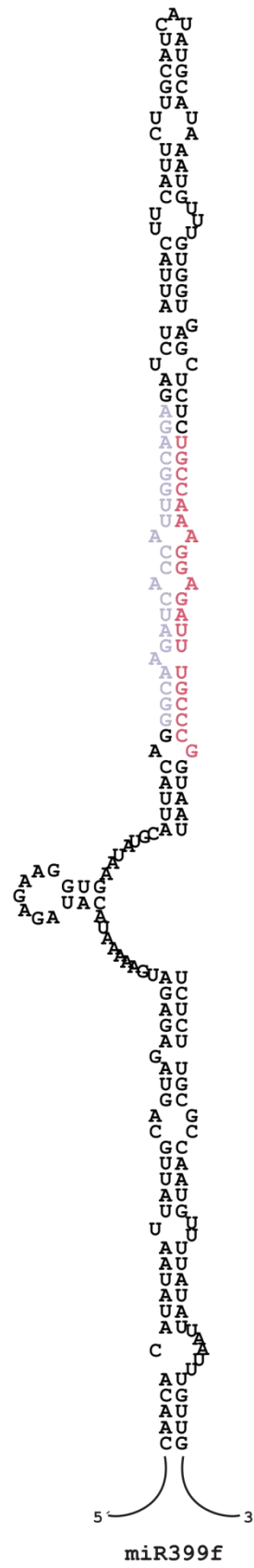
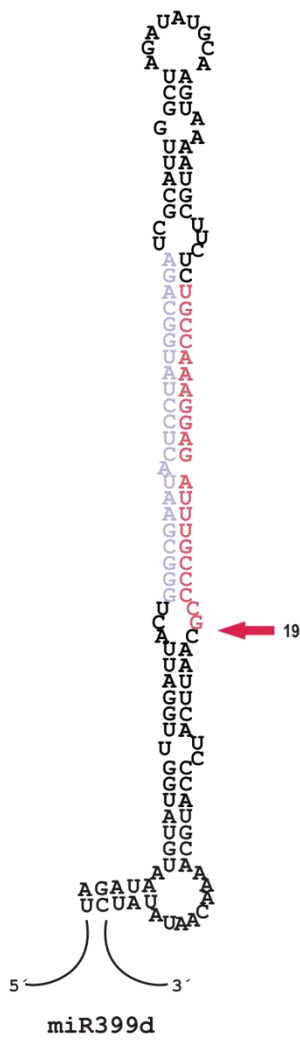
Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A34

Col-0 *fiery1*

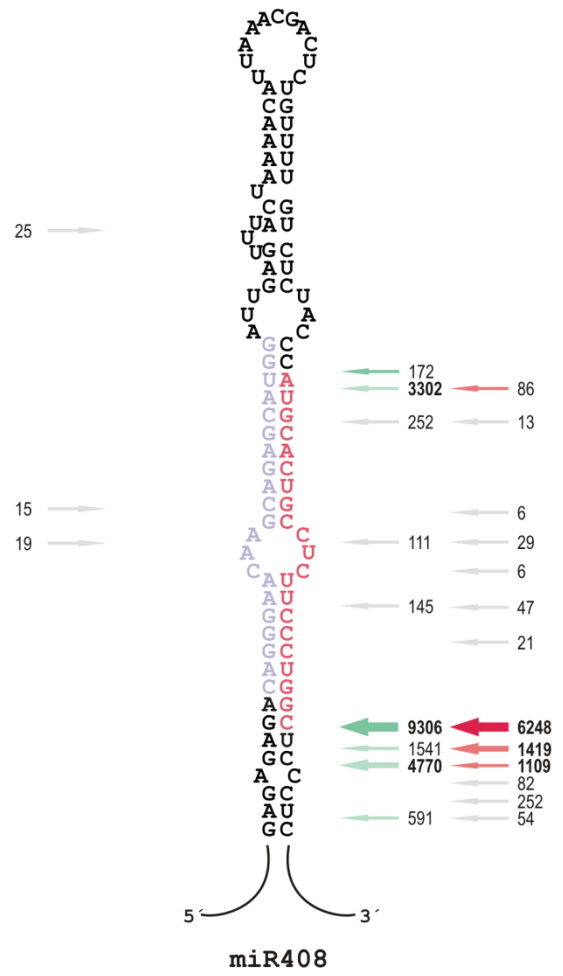
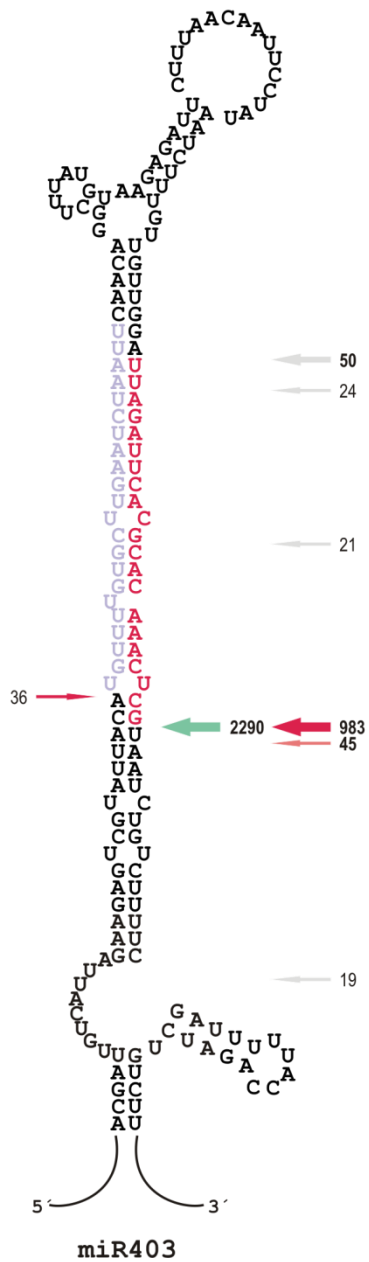
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A35

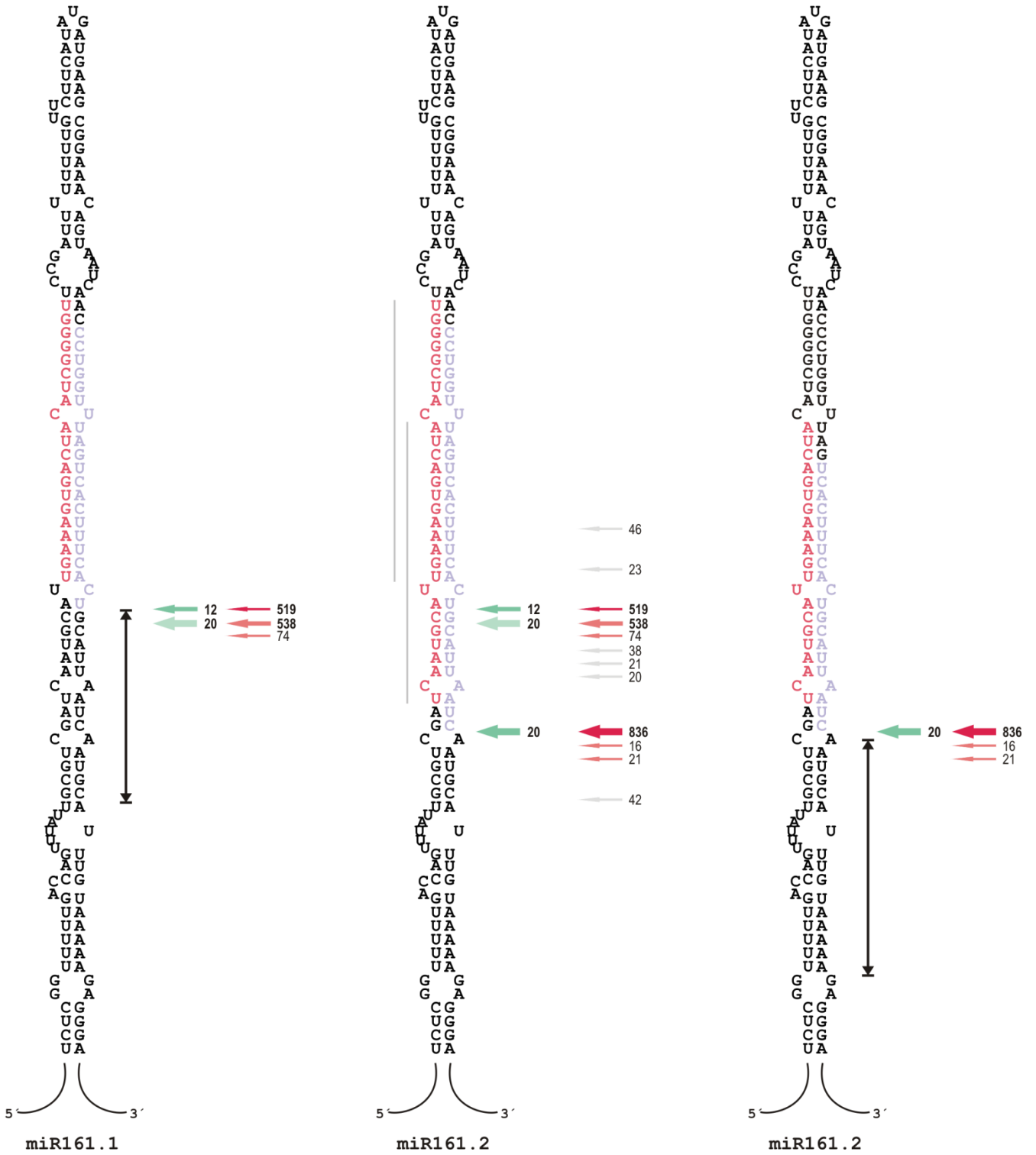
Col-0 *fiery1*

- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos

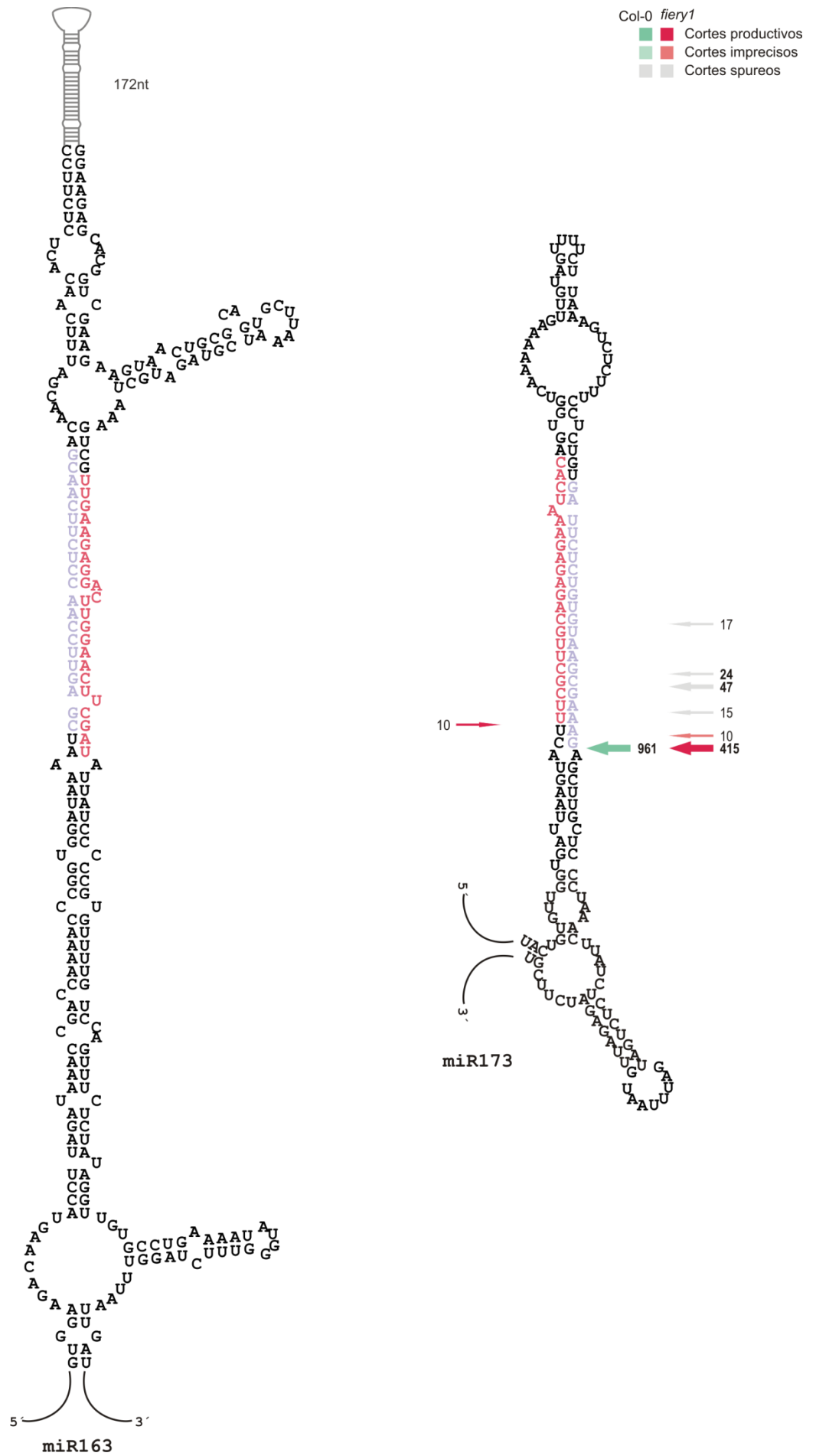


ANEXO A37

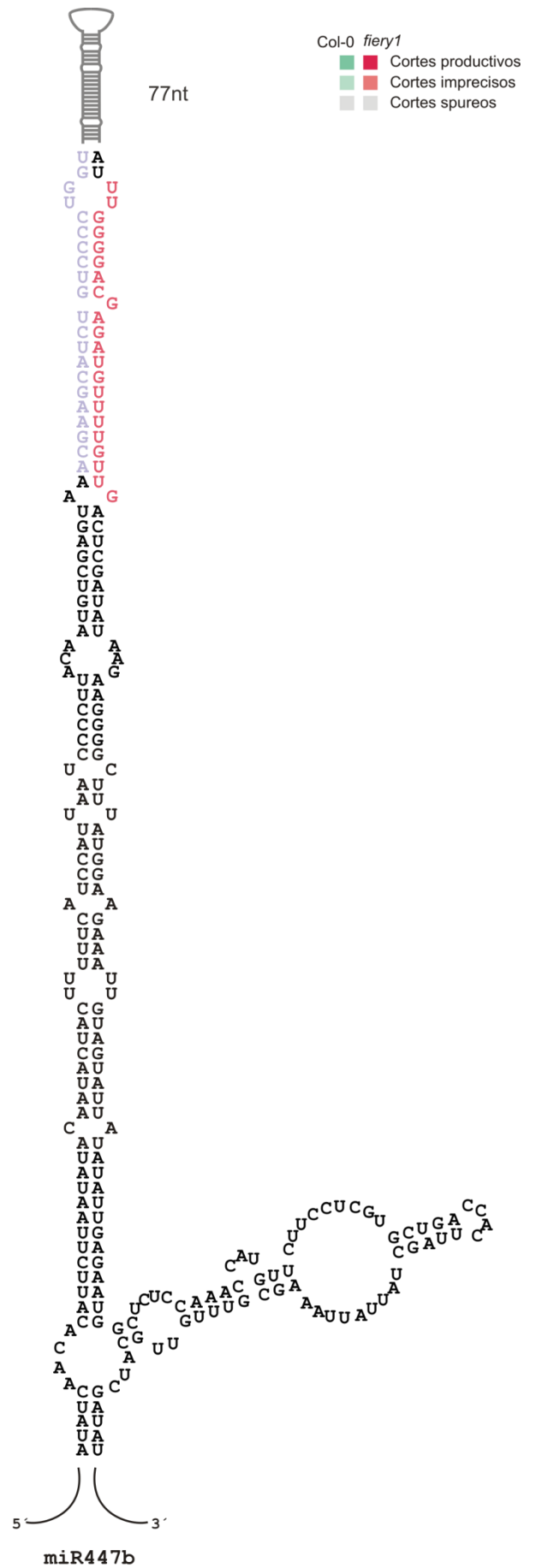
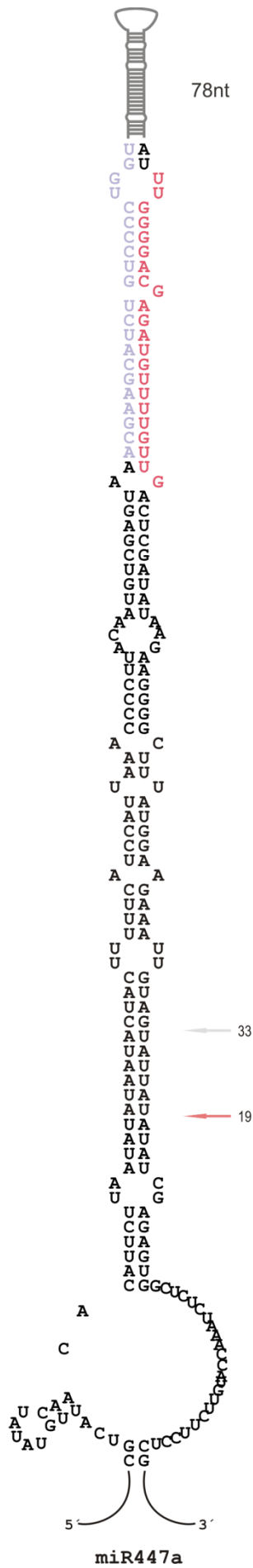
Col-0 *fiery1*
■ Cortes productivos
■ Cortes imprecisos
■ Cortes spureos



ANEXO A38



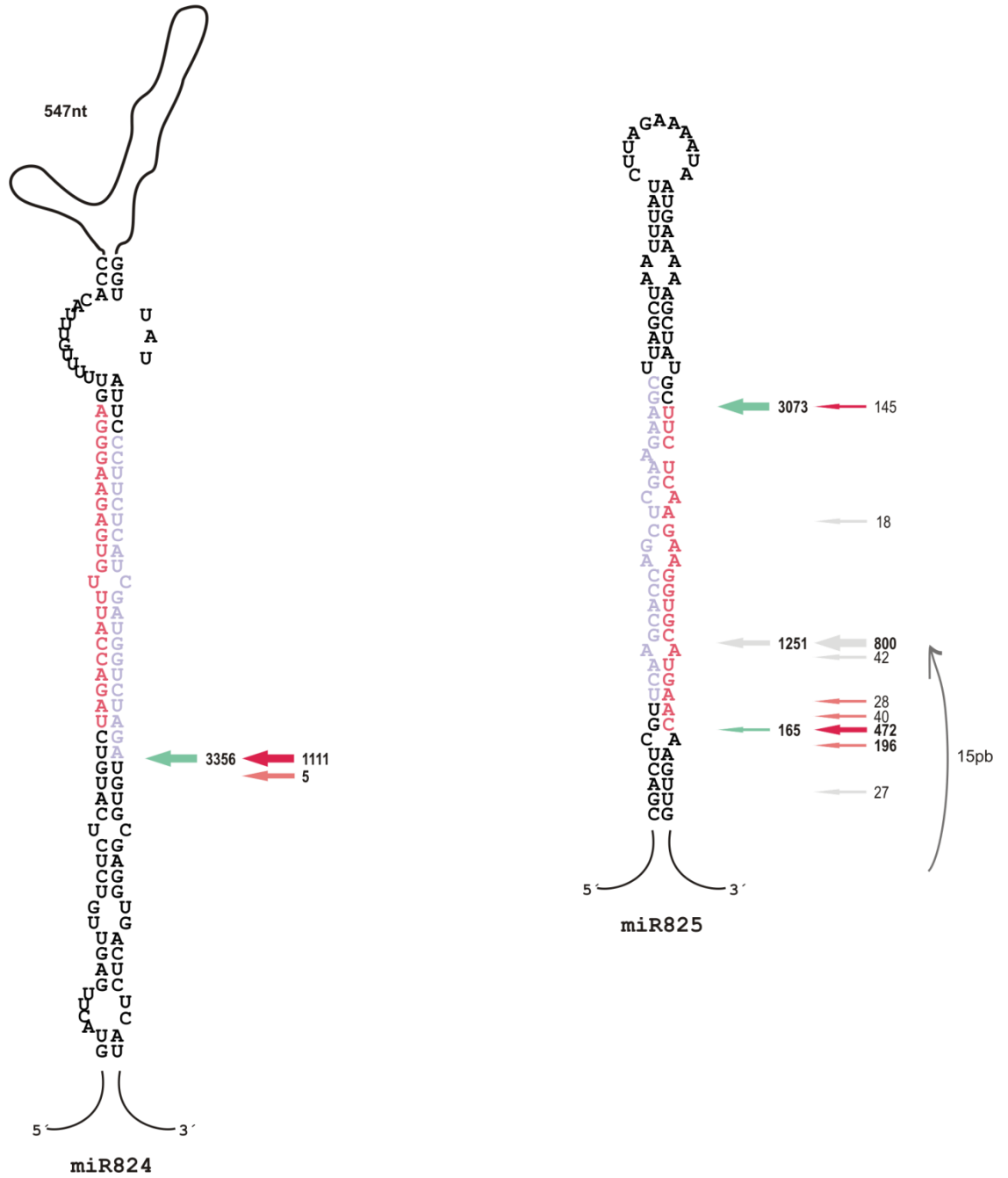
ANEXO A40



ANEXO A41

Col-0 *fiery1*

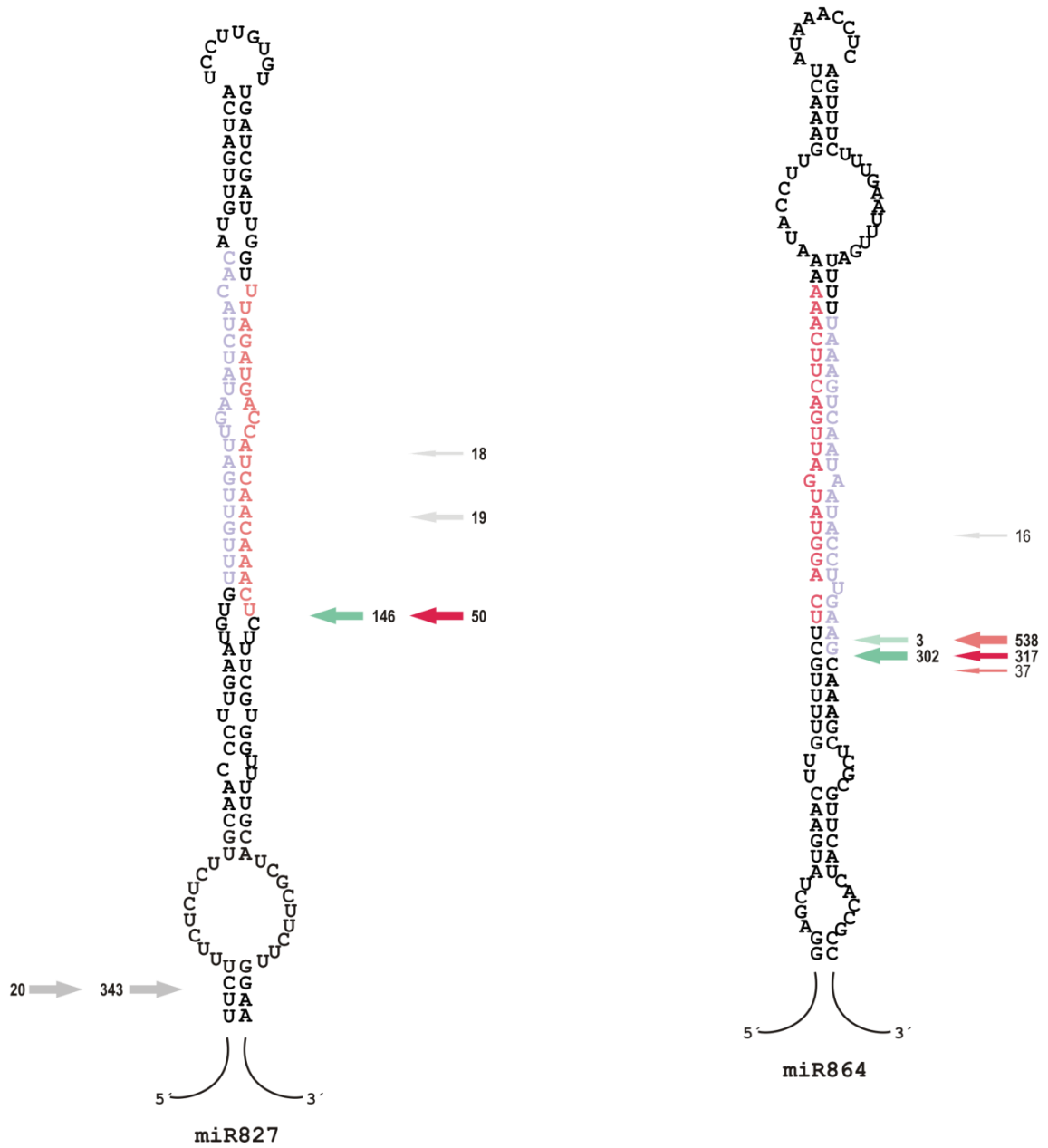
- Cortes productivos
- Cortes imprecisos
- Cortes spureos



ANEXO A42

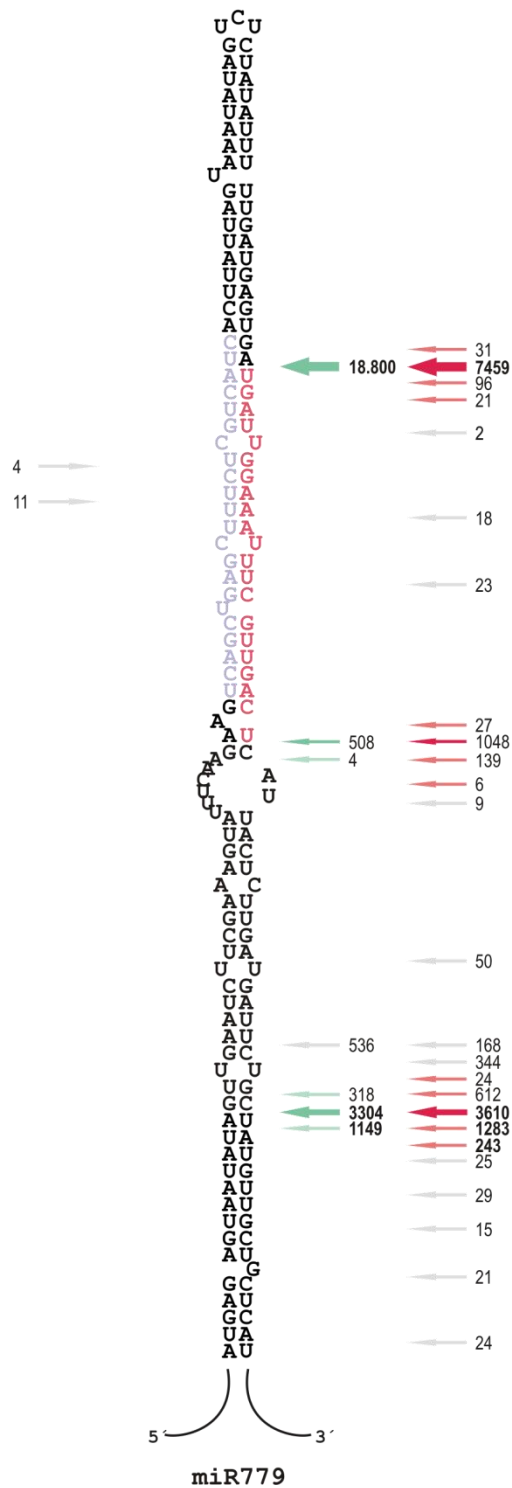
Col-0 *fiery1*

- Cortes productivos
- Cortes imprecisos
- Cortes spureos



Col-0 *fiery1*

- █ Cortes productivos
- █ Cortes imprecisos
- █ Cortes spureos



11. BIBLIOGRAFÍA

- Achkar, N.P., Cambiagno, D.A., and Manavella, P.A. (2016). miRNA Biogenesis: A Dynamic Pathway. *Trends in Plant Science* 21, 1034-1044.
- Addo-Quaye, C., Snyder, J.A., Park, Y.B., Li, Y.F., Sunkar, R., and Axtell, M.J. (2009). Sliced microRNA targets and precise loop-first processing of MIR319 hairpins revealed by analysis of the *Physcomitrella patens* degradome. *Rna* 15, 2112-2121.
- Allen, E., and Howell, M.D. (2010). miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *Seminars in Cell & Developmental Biology* 21, 798-804.
- Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121, 207-221.
- Allen, E., Xie, Z., Gustafson, A.M., Sung, G.H., Spatafora, J.W., and Carrington, J.C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat Genet* 36, 1282-1290.
- Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., *et al.* (2003). A uniform system for microRNA annotation. *Rna* 9, 277-279.
- Ameres, S.L., and Zamore, P.D. (2013). Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol* 14, 475-488.
- Aukerman, M.J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* 15, 2730-2741.
- Axtell, M.J. (2008). Evolution of microRNAs and their targets: are all microRNAs biologically relevant? *Biochim Biophys Acta* 1779, 725-734.
- Axtell, M.J., and Bartel, D.P. (2005). Antiquity of microRNAs and their targets in land plants. *Plant Cell* 17, 1658-1673.
- Axtell, M.J., and Bowman, J.L. (2008). Evolution of plant microRNAs and their targets. *Trends Plant Sci* 13, 343-349.
- Axtell, M.J., Jan, C., Rajagopalan, R., and Bartel, D.P. (2006). A two-hit trigger for siRNA biogenesis in plants. *Cell* 127, 565-577.
- Axtell, M.J., Snyder, J.A., and Bartel, D.P. (2007). Common functions for diverse small RNAs of land plants. *Plant Cell* 19, 1750-1769.
- Axtell, M.J., Westholm, J.O., and Lai, E.C. (2011). Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome Biol* 12, 221.
- Baranauskė, S., Mickutė, M., Plotnikova, A., Finke, A., Venclovas, Č., Klimašauskas, S., and Vilkaitis, G. (2015). Functional mapping of the plant small RNA methyltransferase: HEN1 physically interacts with HYL1 and DICER-LIKE 1 proteins. *Nucleic Acids Research* 43, 2802-2812.
- Bartel, D.P. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* 136, 215-233.
- Ben Amor, B., Wirth, S., Merchan, F., Laporte, P., d'Aubenton-Carafa, Y., Hirsch, J., Maizel, A., Mallory, A., Lucas, A., Deragon, J.M., *et al.* (2009). Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res* 19, 57-69.
- Ben Chaabane, S., Liu, R., Chinnusamy, V., Kwon, Y., Park, J.-h., Kim, S.Y., Zhu, J.-K., Yang, S.W., and Lee, B.-h. (2013). STA1, an *Arabidopsis* pre-mRNA processing factor 6 homolog, is a new player involved in miRNA biogenesis. *Nucleic Acids Research* 41, 1984-1997.

- Blevins, T., Rajeswaran, R., Shivaprasad, P.V., Beknazariants, D., Si-Ammour, A., Park, H.S., Vazquez, F., Robertson, D., Meins, F., Jr., Hohn, T., *et al.* (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res* **34**, 6233-6246.
- Bologna, N.G., Mateos, J.L., Bresso, E.G., and Palatnik, J.F. (2009). A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. *Embo J* **28**, 3646-3656.
- Bologna, N.G., Schapire, A.L., and Palatnik, J.F. (2013a). Processing of plant microRNA precursors. *Brief Funct Genomics* **12**, 37-45.
- Bologna, N.G., Schapire, A.L., Zhai, J., Chorosteki, U., Boisbouvier, J., Meyers, B.C., and Palatnik, J.F. (2013b). Multiple RNA recognition patterns during microRNA biogenesis in plants. *Genome Res* **23**, 1675-1689.
- Bologna, N.G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annu Rev Plant Biol* **65**, 473-503.
- Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* **123**, 1279-1291.
- Boutet, S., Vazquez, F., Liu, J., Beclin, C., Fagard, M., Gratias, A., Morel, J.B., Crete, P., Chen, X., and Vaucheret, H. (2003). Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. *Curr Biol* **13**, 843-848.
- Bowman, J.L. (2004). Class III HD-Zip gene regulation, the golden fleece of ARGONAUTE activity? *Bioessays* **26**, 938-942.
- Braun, Joerg E., Huntzinger, E., Fauser, M., and Izaurralde, E. GW182 Proteins Directly Recruit Cytoplasmic Deadenylation Complexes to miRNA Targets. *Molecular Cell* **44**, 120-133.
- Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**, 1185-1190.
- Brodersen, P., and Voinnet, O. (2009). Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* **10**, 141-148.
- Burdisso, P., Suarez, I.P., Bologna, N.G., Palatnik, J.F., Bersch, B., and Rasia, R.M. (2012). Second double-stranded RNA binding domain of dicer-like ribonuclease 1: structural and biochemical characterization. *Biochemistry* **51**, 10159-10166.
- Calderon-Villalobos, L.I.A., Kuhnle, C., Dohmann, E.M.N., Li, H., Bevan, M., and Schwechheimer, C. (2005). The Evolutionarily Conserved TOUGH Protein Is Required for Proper Development of Arabidopsis thaliana. *The Plant Cell* **17**, 2473-2485.
- Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K.B., Montgomery, T.A., Nguyen, T., Cuperus, J.T., and Carrington, J.C. (2012). Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. *Plant Cell* **24**, 3613-3629.
- Carrington, J.C., and Ambros, V. (2003). Role of microRNAs in plant and animal development. *Science* **301**, 336-338.
- Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**, 735-743.
- Cuperus, J.T., Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Burke, R.T., Takeda, A., Sullivan, C.M., Gilbert, S.D., Montgomery, T.A., and Carrington, J.C. (2010a). Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. *Nat Struct Mol Biol*.

- Cuperus, J.T., Fahlgren, N., and Carrington, J.C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell* **23**, 431-442.
- Cuperus, J.T., Montgomery, T.A., Fahlgren, N., Burke, R.T., Townsend, T., Sullivan, C.M., and Carrington, J.C. (2010b). Identification of MIR390a precursor processing-defective mutants in Arabidopsis by direct genome sequencing. *Proc Natl Acad Sci U S A* **107**, 466-471.
- Chapman, E.J., and Carrington, J.C. (2007). Specialization and evolution of endogenous small RNA pathways. *Nat Rev Genet* **8**, 884-896.
- Chaulk, S.G., Xu, Z., Glover, M.J.N., and Fahlman, R.P. (2014). MicroRNA miR-92a-1 biogenesis and mRNA targeting is modulated by a tertiary contact within the miR-17~92 microRNA cluster. *Nucleic Acids Research* **42**, 5234-5244.
- Chavez Montes, R.A., de Fatima Rosas-Cardenas, F., De Paoli, E., Accerbi, M., Rymarquis, L.A., Mahalingam, G., Marsch-Martinez, N., Meyers, B.C., Green, P.J., and de Folter, S. (2014). Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. *Nat Commun* **5**, 3722.
- Chekanova, J.A., Gregory, B.D., Reverdatto, S.V., Chen, H., Kumar, R., Hooker, T., Yazaki, J., Li, P., Skiba, N., Peng, Q., *et al.* (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell* **131**, 1340-1353.
- Chellappan, P., Xia, J., Zhou, X., Gao, S., Zhang, X., Coutino, G., Vazquez, F., Zhang, W., and Jin, H. (2010). siRNAs from miRNA sites mediate DNA methylation of target genes. *Nucleic Acids Res* **38**, 6883-6894.
- Chen, H.M., Chen, L.T., Patel, K., Li, Y.H., Baulcombe, D.C., and Wu, S.H. (2010). 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proc Natl Acad Sci U S A* **107**, 15269-15274.
- Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* **303**, 2022-2025.
- Cho, S.K., Chaabane, S.B., Shah, P., Poulsen, C.P., and Yang, S.W. (2014). COP1 E3 ligase protects HYL1 to retain microRNA biogenesis. *Nature Communications* **5**, 5867.
- Denli, A.M., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231-235.
- Devers, E.A., Branscheid, A., May, P., and Krajinski, F. (2011). Stars and Symbiosis: MicroRNA- and MicroRNA*-Mediated Transcript Cleavage Involved in Arbuscular Mycorrhizal Symbiosis. *Plant Physiology* **156**, 1990-2010.
- Dong, Z., Han, M.H., and Fedoroff, N. (2008). The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. *Proc Natl Acad Sci U S A* **105**, 9970-9975.
- Dugas, D.V., and Bartel, B. (2008). Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. *Plant Mol Biol* **67**, 403-417.
- Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A., and Voinnet, O. (2007). Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. *Nat Genet* **39**, 848-856.
- Dunoyer, P., Schott, G., Himber, C., Meyer, D., Takeda, A., Carrington, J.C., and Voinnet, O. (2010). Small RNA duplexes function as mobile silencing signals between plant cells. *Science* **328**, 912-916.
- Eamens, A.L., Smith, N.A., Curtin, S.J., Wang, M.-B., and Waterhouse, P.M. (2009). The Arabidopsis thaliana double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. *Rna* **15**, 2219-2235.
- Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P., and Ehresmann, B. (1987). Probing the structure of RNAs in solution. *Nucleic Acids Research* **15**, 9109-9128.
- Endo, Y., Iwakawa, H.O., and Tomari, Y. (2013). Arabidopsis ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly. *EMBO Rep* **14**, 652-658.

- Fang, Y., and Spector, D.L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. *Curr Biol* 17, 818-823.
- Floyd, S.K., and Bowman, J.L. (2004). Gene regulation: ancient microRNA target sequences in plants. *Nature* 428, 485-486.
- Frank, F., Hauver, J., Sonenberg, N., and Nagar, B. (2012). *Arabidopsis* Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. *The EMBO Journal* 31, 3588-3595.
- Fujioka, Y., Utsumi, M., Ohba, Y., and Watanabe, Y. (2007). Location of a possible miRNA processing site in Smd3/Smb nuclear bodies in Arabidopsis. *Plant Cell Physiol* 48, 1243-1253.
- Fukaya, T., and Tomari, Y. (2012). MicroRNAs Mediate Gene Silencing via Multiple Different Pathways in Drosophila. *Molecular Cell* 48, 825-836.
- Gascioli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr Biol* 15, 1494-1500.
- Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-240.
- Gu, S., Jin, L., Huang, Y., Zhang, F., and Kay, Mark A. (2012). Slicing-Independent RISC Activation Requires the Argonaute PAZ Domain. *Current Biology* 22, 1536-1542.
- Guil, S., and Caceres, J.F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* 14, 591-596.
- Gy, I., Gascioli, V., Laressergues, D., Morel, J.B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell* 19, 3451-3461.
- Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* 15, 509-524.
- Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004a). The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-3027.
- Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125, 887-901.
- Han, M.H., Goud, S., Song, L., and Fedoroff, N. (2004b). The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc Natl Acad Sci U S A* 101, 1093-1098.
- Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology* 166, 557-580.
- Havecker, E.R., Wallbridge, L.M., Hardcastle, T.J., Bush, M.S., Kelly, K.A., Dunn, R.M., Schwach, F., Doonan, J.H., and Baulcombe, D.C. (2010). The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* 22, 321-334.
- Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat Genet* 38, 721-725.
- Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol Cell* 32, 276-284.
- Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., Koiwa, H., Seki, M., Shinozaki, K., and Fukuhara, T. (2005). Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. *Plant Mol Biol* 57, 173-188.

- Huang, Y., Ji, L., Huang, Q., Vassilyev, D.G., Chen, X., and Ma, J.-B. (2009). Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature* 461, 823-827.
- Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* 12, 99-110.
- Iki, T., Yoshikawa, M., Meshi, T., and Ishikawa, M. (2012). Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. *The EMBO Journal* 31, 267-278.
- Iki, T., Yoshikawa, M., Nishikiori, M., Jaudal, M.C., Matsumoto-Yokoyama, E., Mitsuhashi, I., Meshi, T., and Ishikawa, M. (2010). In Vitro Assembly of Plant RNA-Induced Silencing Complexes Facilitated by Molecular Chaperone HSP90. *Molecular Cell* 39, 282-291.
- Ipsaro, J.J., and Joshua-Tor, L. (2015). From guide to target: molecular insights into eukaryotic RNA-interference machinery. *Nat Struct Mol Biol* 22, 20-28.
- Iwakawa, H.-o., and Tomari, Y. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends in Cell Biology* 25, 651-665.
- Jarvis, P., Chen, L.J., Li, H., Peto, C.A., Fankhauser, C., and Chory, J. (1998). An Arabidopsis mutant defective in the plastid general protein import apparatus. *Science* 282, 100-103.
- Jauvion, V., Rivard, M., Bouteiller, N., Elmayan, T., and Vaucheret, H. (2012). RDR2 Partially Antagonizes the Production of RDR6-Dependent siRNA in Sense Transgene-Mediated PTGS. *PLoS ONE* 7, e29785.
- Jeong, D.H., and Green, P.J. (2012). Methods for validation of miRNA sequence variants and the cleavage of their targets. *Methods*.
- Jones-Rhoades, M.W., and Bartel, D.P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell* 14, 787-799.
- Jones-Rhoades, M.W., Bartel, D.P., and Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57, 19-53.
- Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-wide profiling and analysis of Arabidopsis siRNAs. *PLoS Biol* 5, e57.
- Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A., Jr., Zhu, J.K., Staskawicz, B.J., and Jin, H. (2006). A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci U S A* 103, 18002-18007.
- Kim, J.Y., Kwak, K.J., Jung, H.J., Lee, H.J., and Kang, H. (2010). MicroRNA402 Affects Seed Germination of Arabidopsis thaliana Under Stress Conditions via Targeting DEMETER-LIKE Protein3 mRNA. *Plant and Cell Physiology* 51, 1079-1083.
- Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* 10, 126-139.
- Kim, W., Kim, H.-E., Jun, A.R., Jung, M.G., Jin, S., Lee, J.-H., and Ahn, J.H. (2016). Structural determinants of miR156a precursor processing in temperature-responsive flowering in Arabidopsis. *Journal of Experimental Botany* 67, 4659-4670.
- Kim, Y.J., Zheng, B., Yu, Y., Won, S.Y., Mo, B., and Chen, X. (2011). The role of Mediator in small and long noncoding RNA production in Arabidopsis thaliana. *The EMBO Journal* 30, 814-822.
- Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 42, D68-73.
- Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *Rna* 12, 206-212.
- Kurihara, Y., and Watanabe, Y. (2004). Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci U S A* 101, 12753-12758.

- Kwak, P.B., and Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol* *19*, 145-151.
- Kwon, S.C., Nguyen, Tuan A., Choi, Y.-G., Jo, Myung H., Hohng, S., Kim, V.N., and Woo, J.-S. (2016). Structure of Human DROSHA. *Cell* *164*, 81-90.
- Lanet, E., Delannoy, E., Sormani, R., Floris, M., Brodersen, P., Crété, P., Voinnet, O., and Robaglia, C. (2009). Biochemical Evidence for Translational Repression by Arabidopsis MicroRNAs. *The Plant Cell* *21*, 1762-1768.
- Lau, P.W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and MacRae, I.J. (2012). The molecular architecture of human Dicer. *Nat Struct Mol Biol* *19*, 436-440.
- Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Ratsch, G., and Weigel, D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* *105*, 8795-8800.
- Laufs, P., Peaucelle, A., Morin, H., and Traas, J. (2004). MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development* *131*, 4311-4322.
- Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Curr Biol* *15*, 1501-1507.
- Li, S., Le, B., Ma, X., Li, S., You, C., Yu, Y., Zhang, B., Liu, L., Gao, L., Shi, T., *et al.* (2016). Biogenesis of phased siRNAs on membrane-bound polysomes in Arabidopsis. *Elife* *5*, e22750.
- Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., Ji, L., Pan, Z., Cao, X., Mo, B., *et al.* (2013). MicroRNAs Inhibit the Translation of Target mRNAs on the Endoplasmic Reticulum in Arabidopsis. *Cell* *153*, 562-574.
- Liu, C., Axtell, M.J., and Fedoroff, N.V. (2012). The helicase and RNaseIIIa domains of Arabidopsis Dicer-Like1 modulate catalytic parameters during microRNA biogenesis. *Plant Physiol* *159*, 748-758.
- Liu, Q., Feng, Y., and Zhu, Z. (2009). Dicer-like (DCL) proteins in plants. *Functional & Integrative Genomics* *9*, 277-286.
- Lobbes, D., Rallapalli, G., Schmidt, D.D., Martin, C., and Clarke, J. (2006). SERRATE: a new player on the plant microRNA scene. *EMBO Rep* *7*, 1052-1058.
- Lowman, H.B., and Draper, D.E. (1986). On the recognition of helical RNA by cobra venom V1 nuclease. *Journal of Biological Chemistry* *261*, 5396-5403.
- Lu, C., and Fedoroff, N. (2000). A Mutation in the Arabidopsis HYL1 Gene Encoding a dsRNA Binding Protein Affects Responses to Abscisic Acid, Auxin, and Cytokinin. *The Plant Cell* *12*, 2351-2365.
- Lu, T., Lu, G., Fan, D., Zhu, C., Li, W., Zhao, Q., Feng, Q., Zhao, Y., Guo, Y., Huang, X., *et al.* (2010). Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res* *20*, 1238-1249.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. *Science* *303*, 95-98.
- Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science* *297*, 2053-2056.
- Ma, E., Zhou, K., Kidwell, M.A., and Doudna, J.A. (2012). Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing. *Journal of Molecular Biology* *422*, 466-476.
- Macias, S., Plass, M., Stajuda, A., Michlewski, G., Eyras, E., and Caceres, J.F. (2012). DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. *Nat Struct Mol Biol* *19*, 760-766.
- MacRae, I.J., and Doudna, J.A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol* *17*, 138-145.
- MacRae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural Basis for Double-Stranded RNA Processing by Dicer. *Science* *311*, 195-198.

- Makino, S., Mishima, Y., Inoue, K., and Inada, T. (2015). Roles of mRNA Fate Modulators Dhh1 and Pat1 in TNRC6-dependent Gene Silencing Recapitulated in Yeast. *Journal of Biological Chemistry* *290*, 8331-8347.
- Mallory, A., and Vaucheret, H. (2010). Form, function, and regulation of ARGONAUTE proteins. *Plant Cell* *22*, 3879-3889.
- Mallory, A.C., Hinze, A., Tucker, M.R., Bouche, N., Gascioli, V., Elmayan, T., Laressergues, D., Jauvion, V., Vaucheret, H., and Laux, T. (2009). Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing. *PLoS Genet* *5*, e1000646.
- Mallory, A.C., Reinhart, B.J., Jones-Rhoades, M.W., Tang, G., Zamore, P.D., Barton, M.K., and Bartel, D.P. (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *Embo J* *23*, 3356-3364.
- Mallory, A.C., and Vaucheret, H. (2009). ARGONAUTE 1 homeostasis invokes the coordinate action of the microRNA and siRNA pathways. *EMBO Rep* *10*, 521-526.
- Manavella, P.A., Hagmann, J., Ott, F., Laubinger, S., Franz, M., Macek, B., and Weigel, D. (2012a). Fast-Forward Genetics Identifies Plant CPL Phosphatases as Regulators of miRNA Processing Factor HYL1. *Cell* *151*, 859-870.
- Manavella, P.A., Koenig, D., and Weigel, D. (2012b). Plant secondary siRNA production determined by microRNA-duplex structure. *Proc Natl Acad Sci U S A* *109*, 2461-2466.
- Mateos, J.L., Bologna, N.G., Chorostecki, U., and Palatnik, J.F. (2010). Identification of microRNA processing determinants by random mutagenesis of Arabidopsis MIR172a precursor. *Curr Biol* *20*, 49-54.
- Matzke, M., Kanno, T., Daxinger, L., Huettel, B., and Matzke, A.J.M. (2009). RNA-mediated chromatin-based silencing in plants. *Current Opinion in Cell Biology* *21*, 367-376.
- Maunoury, N., and Vaucheret, H. (2011). AGO1 and AGO2 act redundantly in miR408-mediated Plantacyanin regulation. *PLoS ONE* *6*, e28729.
- Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., Bowman, J.L., Cao, X., Carrington, J.C., Chen, X., Green, P.J., *et al.* (2008). Criteria for Annotation of Plant MicroRNAs. *Plant Cell*.
- Meyers, B.C., Simon, S.A., and Zhai, J. (2010). MicroRNA Processing: Battle of the Bulge. *Current Biology* *20*, R68-R70.
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., *et al.* (2008). Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* *133*, 116-127.
- Moissiard, G., and Voinnet, O. (2006). RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. *Proc Natl Acad Sci U S A* *103*, 19593-19598.
- Montgomery, T.A., Howell, M.D., Cuperus, J.T., Li, D., Hansen, J.E., Alexander, A.L., Chapman, E.J., Fahlgren, N., Allen, E., and Carrington, J.C. (2008a). Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell* *133*, 128-141.
- Montgomery, T.A., Yoo, S.J., Fahlgren, N., Gilbert, S.D., Howell, M.D., Sullivan, C.M., Alexander, A., Nguyen, G., Allen, E., Ahn, J.H., *et al.* (2008b). AGO1-miR173 complex initiates phased siRNA formation in plants. *Proc Natl Acad Sci U S A* *105*, 20055-20062.
- Moretti, F., Kaiser, C., Zdanowicz-Specht, A., and Hentze, M.W. (2012). PABP and the poly(A) tail augment microRNA repression by facilitated miRISC binding. *Nat Struct Mol Biol* *19*, 603-608.
- Morris, E.R., Chevalier, D., and Walker, J.C. (2006). DAWDLE, a Forkhead-Associated Domain Gene, Regulates Multiple Aspects of Plant Development. *Plant Physiology* *141*, 932-941.
- Murashige, T., and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* *15*, 473-497.

- Nakano, M., Nobuta, K., Vemaraju, K., Tej, S.S., Skogen, J.W., and Meyers, B.C. (2006). Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA. *Nucleic Acids Res* *34*, D731-735.
- Nguyen, Tuan A., Jo, Myung H., Choi, Y.-G., Park, J., Kwon, S.C., Hohng, S., Kim, V.N., and Woo, J.-S. (2015). Functional Anatomy of the Human Microprocessor. *Cell* *161*, 1374-1387.
- Nicholson, A.W. (2014). Ribonuclease III mechanisms of double-stranded RNA cleavage. *Wiley Interdisciplinary Reviews: RNA* *5*, 31-48.
- Nishihara, T., Zekri, L., Braun, J.E., and Izaurralde, E. (2013). miRISC recruits decapping factors to miRNA targets to enhance their degradation. *Nucleic Acids Research* *41*, 8692-8705.
- Nodine, M.D., and Bartel, D.P. (2010). MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev* *24*, 2678-2692.
- Obernosterer, G., Leuschner, P.J., Alenius, M., and Martinez, J. (2006). Post-transcriptional regulation of microRNA expression. *Rna* *12*, 1161-1167.
- Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* *130*, 89-100.
- Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., Grimanelli, D., Slotkin, R.K., Martienssen, R.A., and Vielle-Calzada, J.-P. (2010). Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* *464*, 628-632.
- Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature* *425*, 257-263.
- Palatnik, J.F., Wollmann, H., Schommer, C., Schwab, R., Boisbouvier, J., Rodriguez, R., Warthmann, N., Allen, E., Dezulian, T., Huson, D., *et al.* (2007). Sequence and expression differences underlie functional specialization of *Arabidopsis* microRNAs miR159 and miR319. *Dev Cell* *13*, 115-125.
- Park, J.-E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5[prime] end of RNA for efficient and accurate processing. *Nature* *475*, 201-205.
- Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., and Poethig, R.S. (2005). Nuclear processing and export of microRNAs in *Arabidopsis*. *Proc Natl Acad Sci U S A* *102*, 3691-3696.
- Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol* *12*, 1484-1495.
- Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P., *et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* *408*, 86-89.
- Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev* *18*, 2368-2379.
- Pogulis, R.J., Vallejo, A.N., and Pease, L.R. (1996). In Vitro Recombination and Mutagenesis by Overlap Extension PCR. In *In Vitro Mutagenesis Protocols*, M.K. Trower, ed. (Totowa, NJ: Humana Press), pp. 167-176.
- Pontes, O., Li, C.F., Costa Nunes, P., Haag, J., Ream, T., Vitins, A., Jacobsen, S.E., and Pikaard, C.S. (2006). The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* *126*, 79-92.
- Prigge, M.J., and Wagner, D.R. (2001). The *Arabidopsis* SERRATE Gene Encodes a Zinc-Finger Protein Required for Normal Shoot Development. *The Plant Cell* *13*, 1263-1280.
- Qi, Y., Denli, A.M., and Hannon, G.J. (2005). Biochemical specialization within *Arabidopsis* RNA silencing pathways. *Mol Cell* *19*, 421-428.

- Qi, Y., He, X., Wang, X.J., Kohany, O., Jurka, J., and Hannon, G.J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443**, 1008-1012.
- Qin, H., Chen, F., Huan, X., Machida, S., Song, J., and Yuan, Y.A. (2010). Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *Rna* **16**, 474-481.
- Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes Dev* **20**, 3407-3425.
- Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. *Genes Dev* **16**, 1616-1626.
- Ren, G., Chen, X., and Yu, B. (2012a). Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. *Curr Biol* **22**, 695-700.
- Ren, G., Xie, M., Dou, Y., Zhang, S., Zhang, C., and Yu, B. (2012b). Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. *Proc Natl Acad Sci U S A*.
- Ren, G., Xie, M., Zhang, S., Vinovskis, C., Chen, X., and Yu, B. (2014). Methylation protects microRNAs from an AGO1-associated activity that uridylates 5' RNA fragments generated by AGO1 cleavage. *Proceedings of the National Academy of Sciences* **111**, 6365-6370.
- Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., and Bartel, D.P. (2002). Prediction of plant microRNA targets. *Cell* **110**, 513-520.
- Rodriguez, R.E., Mecchia, M.A., Debernardi, J.M., Schommer, C., Weigel, D., and Palatnik, J.F. (2010). Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. *Development* **137**, 103-112.
- Rodriguez, R.E., Schommer, C., and Palatnik, J.F. (2016). Control of cell proliferation by microRNAs in plants. *Curr Opin Plant Biol* **34**, 68-76.
- Rogers, K., and Chen, X. (2013). Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. *Plant Cell*.
- Rubio-Somoza, I., Cuperus, J.T., Weigel, D., and Carrington, J.C. (2009). Regulation and functional specialization of small RNA-target nodes during plant development. *Curr Opin Plant Biol*.
- Rubio-Somoza, I., and Weigel, D. (2011). MicroRNA networks and developmental plasticity in plants. *Trends Plant Sci* **16**, 258-264.
- Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**, 83-86.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning: a laboratory manual* (Cold spring harbor laboratory press).
- Schapiro, A.L., Bologna, N.G., Moro, B., Zhai, J., Meyers, B.C., and Palatnik, J.F. (2014). Reprint of: construction of Specific Parallel Amplification of RNA Ends (SPARE) libraries for the systematic identification of plant microRNA processing intermediates. *Methods* **67**, 36-44.
- Schauer, S.E., Jacobsen, S.E., Meinke, D.W., and Ray, A. (2002). DICER-LIKE1: blind men and elephants in Arabidopsis development. *Trends in Plant Science* **7**, 487-491.
- Schommer, C., Palatnik, J.F., Aggarwal, P., Chetelat, A., Cubas, P., Farmer, E.E., Nath, U., and Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. *PLoS Biol* **6**, e230.
- Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev Cell* **8**, 517-527.
- Schwab, R., Speth, C., Laubinger, S., and Voinnet, O. (2013). Enhanced microRNA accumulation through stemloop-adjacent introns. *EMBO reports* **14**, 615-621.
- Song, L., Axtell, M.J., and Fedoroff, N.V. (2010). RNA secondary structural determinants of miRNA precursor processing in Arabidopsis. *Curr Biol* **20**, 37-41.

- Song, L., Han, M.H., Lesicka, J., and Fedoroff, N. (2007). Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proc Natl Acad Sci U S A* *104*, 5437-5442.
- Souret, F.F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. *Mol Cell* *15*, 173-183.
- Sunkar, R., Li, Y.-F., and Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends in Plant Science* *17*, 196-203.
- Talmor-Neiman, M., Stav, R., Frank, W., Voss, B., and Arazi, T. (2006). Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss. *Plant J* *47*, 25-37.
- Tang, G., Reinhart, B.J., Bartel, D.P., and Zamore, P.D. (2003). A biochemical framework for RNA silencing in plants. *Genes Dev* *17*, 49-63.
- Taylor, R.S., Tarver, J.E., Hiscock, S.J., and Donoghue, P.C.J. (2014). Evolutionary history of plant microRNAs. *Trends in Plant Science* *19*, 175-182.
- Tian, Y., Simanshu, Dharendra K., Ma, J.-B., Park, J.-E., Heo, I., Kim, V.N., and Patel, Dinshaw J. (2014). A Phosphate-Binding Pocket within the Platform-PAZ-Connector Helix Cassette of Human Dicer. *Molecular Cell* *53*, 606-616.
- Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A.D., Filipowicz, W., Ramos, A., Gherzi, R., and Rosenfeld, M.G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* *459*, 1010-1014.
- Tucker, M.R., Okada, T., Hu, Y., Scholefield, A., Taylor, J.M., and Koltunow, A.M.G. (2012). Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in *Arabidopsis*. *Development* *139*, 1399-1404.
- Vaucheret, H. (2008). Plant ARGONAUTES. *Trends in Plant Science* *13*, 350-358.
- Vaucheret, H., Vazquez, F., Crete, P., and Bartel, D.P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* *18*, 1187-1197.
- Vazquez, F., Blevins, T., Ailhas, J., Boller, T., and Meins, F., Jr. (2008). Evolution of Arabidopsis MIR genes generates novel microRNA classes. *Nucleic Acids Res* *36*, 6429-6438.
- Vazquez, F., Gascioli, V., Crete, P., and Vaucheret, H. (2004a). The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr Biol* *14*, 346-351.
- Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gascioli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crete, P. (2004b). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. *Mol Cell* *16*, 69-79.
- Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J., and Khvorova, A. (2005). The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA* *11*, 674-682.
- Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. *Science* *320*, 97-100.
- Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* *136*, 669-687.
- Wang, J.-W. (2014). Regulation of flowering time by the miR156-mediated age pathway. *Journal of Experimental Botany* *65*, 4723-4730.
- Wang, L., Song, X., Gu, L., Li, X., Cao, S., Chu, C., Cui, X., Chen, X., and Cao, X. (2013). NOT2 proteins promote polymerase II-dependent transcription and interact with multiple MicroRNA biogenesis factors in Arabidopsis. *Plant Cell* *25*, 715-727.

- Wei, K.-F., Wu, L.-J., Chen, J., Chen, Y.-f., and Xie, D.-x. (2012a). Structural evolution and functional diversification analyses of argonaute protein. *Journal of Cellular Biochemistry* *113*, 2576-2585.
- Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.G., and Qi, Y. (2012b). A role for small RNAs in DNA double-strand break repair. *Cell* *149*, 101-112.
- Werner, S., Wollmann, H., Schneeberger, K., and Weigel, D. (2010). Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. *Curr Biol* *20*, 42-48.
- Willmann, M.R., and Poethig, R.S. (2011). The effect of the floral repressor FLC on the timing and progression of vegetative phase change in *Arabidopsis*. *Development* *138*, 677-685.
- Wu, F., Yu, L., Cao, W., Mao, Y., Liu, Z., and He, Y. (2007). The N-terminal double-stranded RNA binding domains of *Arabidopsis* HYPONASTIC LEAVES1 are sufficient for pre-microRNA processing. *Plant Cell* *19*, 914-925.
- Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., and Qi, Y. DNA Methylation Mediated by a MicroRNA Pathway. *Molecular Cell* *38*, 465-475.
- Wu, X., Shi, Y., Li, J., Xu, L., Fang, Y., Li, X., and Qi, Y. (2013). A role for the RNA-binding protein MOS2 in microRNA maturation in *Arabidopsis*. *Cell Res* *23*, 645-657.
- Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., and Carrington, J.C. (2005). Expression of *Arabidopsis* MIRNA genes. *Plant Physiol* *138*, 2145-2154.
- Xie, Z., Kasschau, K.D., and Carrington, J.C. (2003). Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. *Curr Biol* *13*, 784-789.
- Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H. (2006). *Arabidopsis* DCP2, DCP1, and VARICOSE Form a Decapping Complex Required for Postembryonic Development. *The Plant Cell* *18*, 3386-3398.
- Yang, L., Liu, Z., Lu, F., Dong, A., and Huang, H. (2006). SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. *Plant J* *47*, 841-850.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* *17*, 3011-3016.
- Yoshikawa, M., Peragine, A., Park, M.Y., and Poethig, R.S. (2005). A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev* *19*, 2164-2175.
- Yu, B., Bi, L., Zheng, B., Ji, L., Chevalier, D., Agarwal, M., Ramachandran, V., Li, W., Lagrange, T., Walker, J.C., *et al.* (2008). The FHA domain proteins DAWDLE in *Arabidopsis* and SNIP1 in humans act in small RNA biogenesis. *Proc Natl Acad Sci U S A* *105*, 10073-10078.
- Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* *307*, 932-935.
- Zeng, Y., and Cullen, B.R. (2005). Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J Biol Chem* *280*, 27595-27603.
- Zhai, J., Arikait, S., Simon, S.A., Kingham, B.F., and Meyers, B.C. (2014). Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. *Methods* *67*, 84-90.
- Zhan, X., Wang, B., Li, H., Liu, R., Kalia, R.K., Zhu, J.K., and Chinnusamy, V. (2012). *Arabidopsis* proline-rich protein important for development and abiotic stress tolerance is involved in microRNA biogenesis. *Proc Natl Acad Sci U S A* *109*, 18198-18203.
- Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *Embo J* *21*, 5875-5885.
- Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell* *118*, 57-68.

- Zhang, W., Gao, S., Zhou, X., Xia, J., Chellappan, P., Zhang, X., and Jin, H. (2010). Multiple distinct small RNAs originate from the same microRNA precursors. *Genome Biol* *11*, R81.
- Zhang, X., Lii, Y., Wu, Z., Polishko, A., Zhang, H., Chinnusamy, V., Lonardi, S., Zhu, J.-K., Liu, R., and Jin, H. (2013). Mechanisms of Small RNA Generation from Cis-NATs in Response to Environmental and Developmental Cues. *Molecular Plant* *6*, 704-715.
- Zhang, X., Zhao, H., Gao, S., Wang, W.C., Katiyar-Agarwal, S., Huang, H.D., Raikhel, N., and Jin, H. (2011). Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. *Mol Cell* *42*, 356-366.
- Zhang, Z., and Zhang, X. (2012). Argonautes compete for miR165/166 to regulate shoot apical meristem development. *Current Opinion in Plant Biology* *15*, 652-658.
- Zhao, Y., and Srivastava, D. (2007). A developmental view of microRNA function. *Trends in Biochemical Sciences* *32*, 189-197.
- Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012). The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. *Curr Biol* *22*, 689-694.
- Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X. (2011). Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* *145*, 242-256.
- Zhu, H., Zhou, Y., Castillo-Gonzalez, C., Lu, A., Ge, C., Zhao, Y.T., Duan, L., Li, Z., Axtell, M.J., Wang, X.J., *et al.* (2013). Bidirectional processing of pri-miRNAs with branched terminal loops by Arabidopsis Dicer-like1. *Nat Struct Mol Biol* *20*, 1106-1115.
- Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* *299*, 716-719.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* *31*, 3406-3415.