

# Universidad Nacional del Litoral

Facultad de Ciencias Agrarias



Tesis para la obtención del Grado Académico de Doctor en  
Ciencias Agrarias

*“Estudio de la participación de ARNs pequeños durante la  
maduración de frutos y en el desarrollo de atributos con  
importancia comercial”*

Lic. Agustín Andrés Sabbione

Directora de Tesis: Dra. Marcela C. Dotto

Co-director de Tesis: Dr. Lucas D. Daurelio

Laboratorio de Biología Evolutiva y Molecular de Plantas

Instituto de Ciencias Agropecuarias del Litoral

-2021-

## Agradecimientos

Agradezco a la Facultad de Ciencias Agrarias de la Universidad Nacional del Litoral por permitirme realizar el “Doctorado en Ciencias Agrarias” y brindarme las instalaciones necesarias para poder llevar adelante esta tesis.

Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), por otorgarme la beca que me permitió dedicarme exclusivamente al desarrollo de este trabajo.

A mis directores, Dra. Marcela Dotto y Dr. Lucas Daurelio, por aceptar el difícil desafío de dirigirme durante esta etapa.

A Marcela, por darme la posibilidad sin conocerme de hacer una pasantía con ella y luego sumarme como becario doctoral a su grupo y transmitirme sus conocimientos sobre biología molecular, bioinformática y todo lo relacionado al trabajo en el laboratorio.

A mis compañeros de laboratorio que siempre tuvieron buena onda y buena predisposición para el trabajo en grupo. Mención especial para Seba, Nora y Germán, que siempre compartieron su experiencia conmigo y contribuyeron a que el clima de trabajo sea excelente.

A mis amigos, los de la vida y los que me llevo del doctorado. Para mí, todos ustedes son más importantes que cualquier título.

A mis padres y a mi hermana, por su amor y apoyo incondicional en cada etapa y en cada decisión de mi vida. Dicen que “la familia no se elige”, pero si se pudiera, ustedes serían siempre mi primera y única elección, porque no podría ser más afortunado de lo que soy por tenerlos en mi vida.

A Jimena, por acompañarme desde el momento en que decidí presentarme al doctorado, y estar siempre a mi lado durante el transcurso del mismo. Sin tu ayuda y cariño, hubiera sido imposible terminar esta tesis, considerate “Co-autora emocional”.

De corazón, gracias a todos.

## Publicaciones y Comunicaciones a congresos

De esta tesis doctoral se desprenden las siguientes Publicaciones científicas y Comunicaciones a congresos:

### Publicaciones:

- Sabbione, A., Daurelio, L., Vegetti, A., Talón, M., Tadeo, F., & Dotto, M. (2019). Genome-wide analysis of AGO, DCL and RDR gene families reveals RNA-directed DNA methylation is involved in fruit abscission in *Citrus sinensis*. *BMC Plant Biology*, 19(1), 1-13.
- Stalder, S., Sabbione A., Ruiz, V., Daurelio, L., Bouzo, C., Dotto, M. Genome-wide analysis of phased small interfering RNA related to tomato fruit ripening and quality. *Manuscrito en preparación.*

---

### Comunicaciones a Congresos:

- 55th Annual Meeting Argentine Society for Biochemistry and Molecular Biology. Lugar: Salta, Salta; Año: 2019.  
"Identification of small RNAs with roles in the development of organoleptic properties in strawberry fruit". Sabbione, Agustín; Daurelio, Lucas; Vegetti, Abelardo; Dotto, Marcela
- 54th Annual Meeting Argentine Society for Biochemistry and Molecular Biology. Lugar: Paraná, Entre Ríos; Año: 2018.  
"Identification and characterization of ARGONAUTE, DICER-LIKE and RNA-DEPENDENT RNA POLYMERASE protein families in sweet orange". Sabbione, Agustín; Vegetti, Abelardo; Daurelio, Lucas; Dotto, Marcela
- Workshop de Biología Celular y Molecular del ARN. Lugar: Buenos Aires; Año: 2018.  
"ARNs pequeños en el desarrollo de las propiedades organolépticas durante la maduración de frutos de tomate". Sabbione, Agustín; Daurelio, Lucas; Dotto Marcela
- Simposio de Genómica Funcional de Plantas. Lugar: Rosario; Año: 2017.  
"Análisis genómico de phasiARNs en plantas de maíz". Sabbione, Agustín; Hammell, Molly; Timmermans, Marja; Dotto, Marcela.

## *Miembros del jurado*

Guillermo Raúl Pratta

Pablo Andrés Manavella

Hernán Guillermo Rosli

# ÍNDICE

<i>Resumen</i>	11
<i>Abstract</i>	12
<i>Palabras clave</i>	13
<i>Capítulo 1: Introducción general</i>	14
1.1 - Clasificación y anatomía de frutos	15
1.2 - Importancia agronómica y comercial de los frutos	17
1.3 - Desarrollo y maduración de frutos	19
1.3.1 - Sabor y aroma	20
1.3.2 - Firmeza: pared celular	20
1.4 - Aspectos nutricionales	22
1.5 - Regulación génica durante el desarrollo y la maduración de frutos	23
1.6 - ARNs pequeños	23
1.6.1 - ARNs pequeños en fase	25
1.6.2 - Importancia de los ARN pequeños	28
1.7 - Secuenciación del ADN y tecnología Next Generation Sequencing (NGS)	29
1.7.1 - Next Generation Sequencing	30
1.7.2 - Tipos de experimentos de NGS	31
1.7.3 - Análisis Paralelo de Extremos de ARN o Degradoma	31
1.7.4 - Herramientas bioinformáticas	34
1.8 - Consideraciones generales sobre los análisis realizados	34
1.9 - Hipótesis de trabajo	36
1.10 - Objetivo general	36
1.10.1 - Objetivos específicos	36
<i>Capítulo 2: ARNs pequeños en tomate</i>	37
2.1 - Introducción	38
2.1.1 - Familia Solanaceae	38
2.1.2 - Características generales del tomate	38
2.1.3 - Maduración de frutos de tomate y aspectos nutricionales	39
2.1.3.1 Estadios de maduración de los frutos de tomate	39
2.1.3.2 - Aspectos nutricionales de los frutos de tomate	40
2.1.3.3 - Aspectos organolépticos de los frutos de tomate	41
2.1.3.4 - Vías metabólicas implicadas en la adquisición de propiedades organolépticas	43
<b>a) Sabor y aroma</b>	43

<b>b) Color</b>	44
<b>c) Textura</b>	45
2.1.4 - ARNs pequeños en frutos de tomate	46
2.1.5 - Estudios genómicos de ARNs pequeños	47
2.2 - Materiales y métodos	48
2.2.1 - Datos de secuenciación masiva	48
2.2.2 - Pre-procesamiento de datos de NGS	49
2.2.2.1 - Filtrado de secuencias contaminantes	49
2.2.3 - Análisis de miARNs de tomate	50
2.2.3.1 - Descarga y actualización de la información correspondiente a miARNs de tomate	50
2.2.3.2 - Mapeo al genoma de referencia, anotación y cuantificación de miARNs	50
2.2.4 - Análisis de expresión diferencial de miARNs	51
2.2.5 - Detección de loci en fase	51
2.2.6 - Mapeo de datos de RNA-seq al genoma de referencia	51
2.2.7 - Visualización de los resultados de alineamientos al genoma de referencia	52
2.2.8 - Análisis de expresión de genes blanco de miARNs	52
2.2.9 - Validación de genes blanco con bibliotecas de degradoma	52
2.3 - Resultados	53
2.3.1 - Análisis de miARNs	53
2.3.1.1 - Pre-procesado de los datos de secuenciación de small RNA-seq	54
2.3.1.2 - Actualización de la información de miARNs de tomate	55
2.3.1.3 - Expresión de miARNs durante la maduración de frutos de tomate	56
2.3.1.4 - Identificación y validación de genes regulados por miARNs durante la maduración	58
2.3.2 - Caracterización de loci generadores de ARNs pequeños en fase	61
2.3.2.1 - Predicción de loci PHAS y TAS	61
2.3.2.2 - Detección y validación de genes blanco	63
2.3.3 - sly-miR6026 produce phasiARNs a partir de transcritos de la familia sly-DCL2 de tomate	66
2.3.3.1 - Locus 7	67
2.3.3.2 - Locus 1	71
2.3.4 - Otros loci PHAS relacionados con el desarrollo de propiedades organolépticas	74
2.3.4.1 - Locus 6	74
2.3.4.2 - Locus 10	75
2.3.4.3 - Locus 15	77
2.4 - Discusión	78

2.4.1 - En frutos de tomate se expresan 48 familias de miARNs	78
2.4.2 - miARNs reguladores del desarrollo de propiedades organolépticas de frutos de tomate	79
2.4.3 - Se identificaron 19 loci productores de phasiARNs en frutos de tomate	81
2.4.4 - Se validaron 82 genes blanco para los phasiARNs detectados	82
2.4.5 - El microARN sly-miR6026 también genera phasiARNs en frutos	82
2.4.6 - La ubicación del locus 1 coincide con la del gen Sly-DCL2b	84
2.4.7 - Se encontraron phasiARNs reguladores de genes vinculados a las propiedades organolépticas en otros 3 loci	86
2.5 - Conclusiones	87
<i>Capítulo 3: ARNs pequeños en frutilla</i>	88
3.1 - Introducción	89
3.1.1 - Breve descripción del género <i>Fragaria</i> y origen de <i>Fragaria x ananassa</i>	89
3.1.2 - Morfología del fruto	89
3.1.3 - Aspectos nutricionales	90
3.1.4 - Maduración no climatérica	91
3.1.5 - Poscosecha	93
3.1.6 - Importancia comercial	93
3.1.7 - Genoma de frutilla y avances en bioinformática	94
3.1.8 - ARNs pequeños en plantas de frutilla	94
3.2 - Materiales y métodos	95
3.2.1 - Datos de secuenciación masiva	95
3.2.2 - Pre-procesamiento de datos de NGS	95
3.2.2.1 - Filtrado de secuencias contaminantes	95
3.2.3 - Análisis de miARNs de frutilla	96
3.2.3.1 - Descarga de la información correspondiente a miARNs de frutilla	96
3.2.3.2 - Mapeo al genoma de referencia, anotación y cuantificación de miARNs	96
3.2.4 - Análisis de expresión diferencial de miARNs	97
3.2.5 - Detección de loci en fase	97
3.2.6 - Mapeo de datos de RNA-seq al genoma de referencia	97
3.2.7 - Caracterización funcional mediante análisis de ontología génica	97
3.2.8 - Validación de genes blanco con bibliotecas de degradoma	98
3.3 - Resultados	99
3.3.1 - Análisis de microARNs	99
3.3.1.1 - Pre-procesado de los datos de secuenciación de small RNA-seq	100
3.3.1.2 - Expresión de miARNs durante la maduración de frutos de frutilla	101

3.3.1.3 - Identificación y validación de genes regulados por miARNs durante la maduración	105
3.3.2 - Análisis de descubrimiento y caracterización de loci generadores de phasiARNs	108
3.3.3 - Caracterización de transcritos regulados por ARNs pequeños	110
3.3.3.1 - Caracterización de transcritos target en base a etiquetas de ontología génica	111
3.3.3.2 - Validación de genes blanco de interés	112
3.4 - Discusión	115
3.4.1 - En frutos de frutilla se expresan 28 familias de miARNs	115
3.4.2 - miARNs reguladores del desarrollo de propiedades organolépticas de frutos de frutilla	116
3.4.3 - Se identificaron 42 loci productores de phasiARNs en frutos de frutilla	119
3.4.4 - Se validaron genes regulados por phasiARNs con potencial importancia comercial	120
3.5 - Conclusiones	121
<i>Capítulo 4: ARNs pequeños en naranja</i>	123
4.1 - Introducción	124
4.1.1 - Género Citrus	124
4.1.2 - Características generales	124
4.1.2.1 - Aspectos nutricionales, maduración y propiedades organolépticas	124
4.1.2.2 - Producción e importancia económica	126
4.1.3 - Genomas de referencia y ARNs pequeños de Citrus sinensis	127
4.2 - Materiales y métodos	128
4.2.1 - Datos de secuenciación masiva	128
4.2.2 - Pre-procesamiento de datos de NGS	128
4.2.2.1 - Filtrado de secuencias contaminantes	129
4.2.3 - Análisis de miARNs de naranjo	129
4.2.3.1 - Descarga de la información correspondiente a miARNs de naranjo	129
4.2.3.2 - Mapeo al genoma de referencia, anotación y cuantificación de miARNs	129
4.2.4 - Detección de loci en fase	130
4.2.5 - Validación de genes blanco con bibliotecas de degradoma	130
4.3 - Resultados	131
4.3.1 - Análisis de miARNs	131
4.3.1.1 - Pre-procesado de los datos de secuenciación de small RNA-seq	132
4.3.1.2 - Expresión de miARNs en frutos de naranjo	133
4.3.1.3 - Identificación y validación de genes regulados por miARNs en frutos	134
4.3.2 - Descubrimiento y caracterización de loci generadores de phasiARNs	135
4.3.2.1 - Detección y validación de genes blanco	138

4.3.2.2 - Genes relacionados a la adquisición de propiedades organolépticas regulados por phasiARNs	139
4.4 - Discusión	141
4.4.1 - En frutos de naranjo se expresan 29 familias de miARNs	141
4.4.2 - Se identificaron 63 loci productores de phasiARNs en frutos de naranjo	142
4.5 - Conclusiones	143
<i>Capítulo 5: Biosíntesis de ARNs pequeños y el proceso de abscisión en naranjo</i>	144
5.1 - Introducción	145
5.1.1 - Componentes de las rutas de biogénesis y de acción de ARNs pequeños en naranjo	145
5.1.2 - El proceso de abscisión	146
5.1.2.1 - Regulación a nivel molecular del proceso de abscisión	146
5.2 - Materiales y métodos	147
5.2.1 - Identificación de los genes AGO, DCL y RDR en naranjo	147
5.2.2 - Análisis filogenético y localización cromosomal	148
5.2.3 - Análisis de expresión de genes DCL, AGO y RDR en tejidos de plantas de naranjo	148
5.2.4 - Análisis de datos de expresión de genes DCL, AGO y RDR de naranja durante el proceso de abscisión	149
5.2.5 - Análisis cuantitativo de RT-PCR	150
5.3 - Resultados	151
5.3.1 - Identificación y análisis in silico de los genes AGO, RDR y DCL en naranjo	151
5.3.2 - Análisis filogenético para identificar los ortólogos putativos de DCL, RDR y AGO en naranja	156
5.3.3 - Ubicación cromosómica de los genes DCL, RDR y AGO de naranja	158
5.3.4 - Patrones de expresión comunes y específicos de genes AGO, DCL y RDR en diferentes tejidos	159
5.3.5 - La expresión de genes involucrados en RdDM y vías de tasiARN está regulada negativamente durante la abscisión de los frutos de naranjo.	161
5.4 - Discusión	164
5.4.1 - La familia RDR en <i>Citrus sinensis</i> presenta características distintivas y patrones de expresión específicos para cada uno	165
5.4.2 - Los miembros de la familia DCL en <i>Citrus sinensis</i> presentan diferentes patrones de expresión	166
5.4.3 - La familia AGO en <i>Citrus sinensis</i> sufrió eventos de duplicación en tándem durante la evolución	167
5.4.4 - La vía de RdDM está regulada negativamente durante la abscisión de frutos	169
5.5 - Conclusiones	171
<i>Capítulo 6: Discusión y conclusiones generales</i>	172

6.1 - ARNs pequeños como reguladores de aspectos de interés comercial en frutos	173
6.2 - Regulación mediada por ARNs pequeños de otras cualidades de importancia comerciales	175
6.3 - Perspectivas a futuro	176
<i>Bibliografía</i>	178
<i>Anexos</i>	203
Anexo I - Comandos bioinformáticos	204
Anexo II - Listado de miARNs utilizados en el análisis del capítulo 2	208

## Resumen

Este trabajo de tesis involucró el estudio de frutos de tomate, frutilla y naranja, los cuales se encuentran entre los cultivos frutihortícolas más importantes para el consumo en fresco de nuestro país, por lo que es importante que no sufran deterioro en los procesos de producción y comercialización, y que desarrollen propiedades organolépticas óptimas para el agrado del consumidor. La calidad y atractivo sensorial de los frutos están determinados en gran medida por el contenido de metabolitos secundarios, por lo que la obtención de cultivares con contenidos optimizados desde el punto de vista cualitativo y nutricional de estos metabolitos es cada vez más explorada, en particular a través del estudio de las redes regulatorias que modulan el metabolismo de estos compuestos.

Los ARNs pequeños regulan un gran número de genes involucrados en diversos procesos durante el desarrollo y crecimiento de las plantas, pero los reportes sobre su participación en redes regulatorias relacionadas con aspectos cualitativos de frutos son escasos. Existen varios tipos de ARNs pequeños originados por distintos mecanismos de biosíntesis. En este trabajo se realizaron análisis bioinformáticos para identificar módulos regulatorios que involucran la acción de miARNs y phasiARNs de manera comparativa en frutos de tres especies con características organolépticas muy distintas, y con diferentes patrones de preferencia por parte de los consumidores.

En las 3 especies analizadas, se identificaron ARNs pequeños reguladores de genes relacionados con la adquisición de características de importancia comercial, tales como el desarrollo de brillo, color, firmeza y contenido nutricional. También se realizó por primera vez la caracterización a nivel genómico de las principales familias proteicas implicadas en la biosíntesis de ARNs pequeños en naranjo, lo cual permitió relacionar a estas moléculas regulatorias con el proceso de abscisión. La información producida en esta tesis abre la puerta a futuros desarrollos tecnológicos que contribuyan a mejorar diversos aspectos cualitativos de estos frutos.

## Abstract

This thesis work involved the study of tomato, strawberry and orange fruits, which are among the most important crops commercialized for fresh consumption in our country, so it is important they do not get damaged during production and commercialization and that they develop optimal organoleptic properties for consumer satisfaction. Fruit quality and sensory attractiveness are set mostly by their secondary metabolite content; therefore, the development of cultivars with optimized secondary metabolites content from a qualitative and nutritional point of view is being actively explored, particularly through the study of regulatory networks that modulate the metabolism of these compounds.

Small RNAs regulate a large number of genes involved in several processes during plant development and growth, but there are few reports on their participation in regulatory networks related to fruit qualitative traits. There are several types of small RNAs produced by different biosynthesis mechanisms. In this work, bioinformatic analyses were performed in order to identify regulatory modules that involve the action of miRNAs and phasiRNAs, in a comparative manner, in fruits of three species with very different organoleptic characteristics and different consumer preference patterns.

In the three analyzed species, small regulatory RNAs of genes related to the acquisition of commercially important traits were identified, such as the development of brightness, color, firmness and nutritional content. Besides, the main protein families implied in small RNAs biosynthesis in orange were characterized at the genomic level for the first time, which allowed us to relate these regulatory molecules to the abscission process. The information produced in this thesis work opens the door to future technological developments that will contribute to improving several qualitative traits of these fruits.

## Palabras clave

- Frutos
- Maduración
- Abscisión
- ARNs pequeños
- Propiedades organolépticas

# Capítulo 1: Introducción general

## 1.1 - Clasificación y anatomía de frutos

Las angiospermas, vulgarmente conocidas como “plantas con flores”, son el grupo más grande de plantas terrestres, con alrededor de 300.000 especies, tanto monocotiledóneas como dicotiledóneas (Christenhusz & Byng, 2016). El origen de este grupo se ubica en el cretácico temprano, hace aproximadamente 130 millones de años, respaldado con un abundante y diverso registro fósil (Crane et al., 1995). El término “Angiosperma” proviene de las palabras griegas *angeion* (cubierta) y *sperma* (semilla), haciendo clara referencia a los frutos producidos por este grupo de plantas (Meyer, 2005). Los órganos florales de las angiospermas son su principal diferencia con las demás especies vegetales, y les han proporcionado especificidad reproductiva a lo largo de la evolución. Los óvulos se encuentran encerrados en uno o más carpelos (ovario), y luego de ser fecundadas sus ovocélulas y sus núcleos polares, darán lugar a la formación de una o más semillas. (Specht & Bartlett, 2009).

En botánica, se considera frutos a los ovarios cuyos óvulos han experimentado la fecundación arriba mencionada y se irán transformando en semillas. En algunas especies otras porciones de las estructuras florales pueden formar parte del fruto además del ovario. Su principal función es la de la propagación de la especie, dado que su desarrollo se ve acompañado por la maduración de una o más semillas en su interior

(White, 2002). Las semillas pueden dispersarse naturalmente por varios mecanismos, como la acción del viento, la adherencia a los pelos de ciertos animales, o por la digestión de frutos que presentan colores vistosos y aromas y sabores agradables para atraer a sus consumidores, quienes luego dispersan las semillas por su posterior excreción (Willson, 1993).

Estructuralmente, la pared del ovario que rodea a las semillas y conforma la parte comestible del fruto, da lugar al pericarpio. Éste, a su vez, se divide en tres porciones llamadas, desde la parte exterior hacia la interior, epicarpio, mesocarpio y endocarpio (Fig. 1.1C). En el caso de los cítricos, al mesocarpio suele llamárselo “albedo” mientras que al epicarpio se lo denomina “flavedo” (Fig. 1.1A). Los frutos se clasifican en tres grandes grupos (G. Singh, 2019):

- frutos simples, se desarrollan a partir de un ovario de la flor. Pueden ser secos, como las avellanas, o carnosos como tomate o naranja.
- frutos agregados, se forman a partir de varios carpelos dentro de la flor. Dos ejemplos son mora y frutilla.
- frutos compuestos o múltiples, se originan a partir de una inflorescencia, donde cada flor da lugar a un fruto pero se desarrollan en conjunto. Ejemplos de este grupo son el ananá y el higo.

Cada uno de estos grupos presenta subdivisiones pero sólo mencionaremos las que engloban a las especies que serán estudiadas en este trabajo: tomate (*Solanum lycopersicum*), naranja (*Citrus sinensis*) y frutilla (*Fragaria x ananassa*).

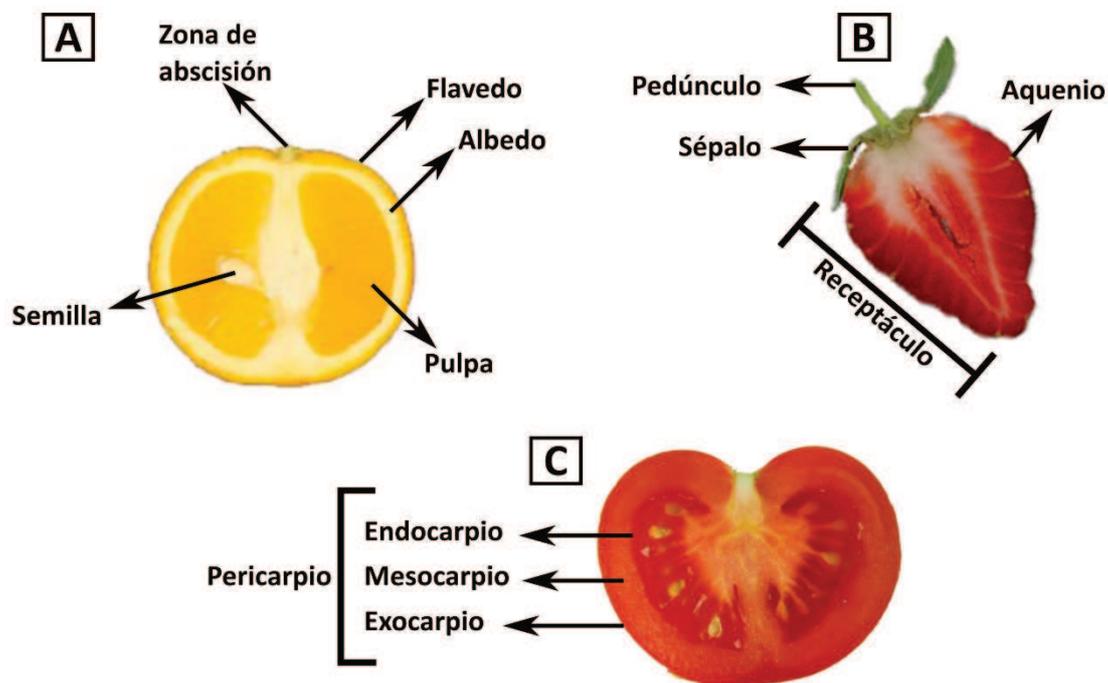


Fig. 1.1: Cortes transversales e identificación de las partes de los 3 frutos estudiados en esta tesis (Iglesias et al., 2007; Pesaresi et al., 2014; Vallarino et al., 2020).

El tomate es un fruto simple, carnosos y es considerado una baya, ya que su epicarpio es muy delgado, y su mesocarpio y endocarpio son carnosos y jugosos. Además, presenta una forma redondeada, común a la mayoría de las bayas. Los hesperidios, frutos típicos de los cítricos, son un tipo de baya modificada, conservan su forma pero la parte comestible del fruto está constituida por los carpelos hinchados recubiertos por el pericarpio. Los cítricos también presentan frutos simples carnosos.

Por otro lado, en los frutos agregados, también llamados eterios o “falsos frutos”, la porción comestible es el receptáculo, que alberga al verdadero fruto contenedor de las semillas. En el caso de las frutillas, estas estructuras son los aquenios (Fig. 1.1B; Font Quer, 1953).

## 1.2 - Importancia agronómica y comercial de los frutos

La ingestión de frutas y verduras está recomendada como parte de una dieta saludable, dado que son fuente de vitaminas C y A, de minerales e incluso de compuestos antioxidantes y proporcionan fibra dietaria (Slavin & Lloyd, 2012). De hecho, una ingesta deficiente aumenta el riesgo de padecer numerosas enfermedades, como hipertensión arterial, diabetes tipo 2, enfermedades cardíacas y ciertos tipos de cáncer (Hall et al., 2009). En líneas generales, los consumidores buscan frutos maduros, firmes y dulces, aunque cada especie tiene atributos comerciales que predominan sobre otros a la hora de la compra por parte de los consumidores. Por ejemplo, en el caso de los tomates se prioriza el color, el sabor y la firmeza (Causse et al., 2003; Oltman et al., 2016). En frutillas se busca que el ablandamiento poscosecha de las mismas sea lo más lento posible, además de preferirse colores rojo intensos y aromas agradables (Bhat et al., 2015; Jiménez-Bermúdez et al., 2002). En ambos casos, la firmeza de estos frutos, además de estar vinculada a la preferencia sensorial, está relacionada con otro importante atributo comercial por la relación que existe entre el ablandamiento y una mayor susceptibilidad al ataque de patógenos (Cantu et al., 2008; Forlani et al., 2019). Por su parte, los consumidores de cítricos priorizan el contenido de jugo, la acidez, la firmeza y el aroma (Simons et al., 2018; Soule & Grierson, 1986). En el caso particular de la comercialización de cítricos, es importante mencionar que el proceso de abscisión de frutos representa un grave problema, ya que impacta directamente en el rendimiento del cultivo y provoca pérdidas millonarias, debido a que un fruto golpeado no puede ser comercializado (Merelo et al., 2017).

Desde el punto de vista de los productores, es muy importante la selección del cultivar a la hora de comenzar la producción, ya que cada variedad presenta características diferentes. Por lo general, se busca que el cultivo tenga atributos deseados por el mercado, sea adecuado para el suelo y el clima del lugar elegido, que

presente buen rendimiento y posea resistencia a ciertas plagas o enfermedades, que pueden provocar pérdidas parciales o totales de producción (McAvoy, 2008). Muchas veces se obtiene alguno de estos beneficios a costa de otros, por lo que es importante lograr un balance. Por ejemplo, una variedad puede ser muy resistente frente a ciertos patógenos, pero no tener el mismo sabor o rendimiento que otras. Esto también se tiene en cuenta según el mercado al que está dirigido el fruto. Por ejemplo, en el caso de las frutillas, si se las comercializará para consumo fresco, será muy importante el sabor y el aroma, pero si en cambio su función es decorativa (por ejemplo en repostería), se priorizará la firmeza y el color del fruto.

Es fundamental, además, el correcto manejo poscosecha de acuerdo a las características particulares de cada fruto, principalmente para garantizar que el consumidor reciba un producto de buena calidad luego de los procesos de transporte y almacenamiento. Para las frutillas es muy importante el momento en que son recolectadas, ya que son frutos no climatéricos y su maduración no progresa fuera de la planta. Son frutos que se descomponen rápidamente, por lo que deben ser correctamente resguardados. Para transportarlas por largas distancias, se dispone de cámaras refrigeradas y enriquecidas en dióxido de carbono. Para los frutos de tomate se siguen prácticas similares a las mencionadas, pero en algunos casos se añade además el agregado de 1-metilciclopropeno para su traslado, ya que es un fruto climatérico y el etileno es la principal hormona regulando su maduración, incluso luego de la cosecha. Este compuesto actúa como inhibidor del etileno, ya que se une a los receptores de esta hormona con mayor afinidad, impidiendo su actividad (Arah et al., 2016). En cambio, para las naranjas, es fundamental evitar el daño de los frutos durante la recolección. Estos daños aparecen como golpes o magulladuras, producto de la caída de los frutos al suelo, ya sea por una mala manipulación o por abscisión natural de los mismos; y pueden también ocurrir durante su transporte, donde la carga y descarga deben realizarse con cuidado. Como su traslado generalmente tarda varios días, si los frutos están golpeados se observan cambios en sus propiedades físico-mecánicas, además de que se altera la percepción positiva de los consumidores (K. K. Singh & Reddy, 2006).

Todos estos fenómenos están relacionados con diferentes aspectos fisiológicos de los frutos. Algunos de ellos han sido caracterizados en detalle, mientras que otros

aún no han sido completamente estudiados. Conocer y comprender los procesos biológicos que determinan estos aspectos, permitirían el desarrollo de diversas tecnologías orientadas a mejorar las diferentes etapas de producción y comercialización de frutos.

### 1.3 - Desarrollo y maduración de frutos

El fenómeno de desarrollo de los frutos implica una serie de cambios fisiológicos y bioquímicos finamente regulados. En frutos carnosos consta de tres etapas principales: establecimiento del fruto, desarrollo propiamente dicho y maduración (Bouzayen et al., 2010). La maduración es posiblemente la etapa más estudiada del proceso: permite clasificar a los frutos en climatéricos y no climatéricos. Los frutos climatéricos presentan un pico respiratorio (CO<sub>2</sub>) junto con un aumento en la producción de la hormona gaseosa etileno, algo no observado en frutos no climatéricos. Además, los frutos climatéricos pueden continuar madurando fuera de la planta en presencia de etileno exógeno (Paul et al., 2012).

La maduración es el momento en el que se desarrollan las propiedades organolépticas que proporcionan el atractivo comercial del producto, que debe contar con sabor y aroma agradables, como también buena textura y firmeza. Muchas veces la diversidad de características sensitivas es muy amplia dentro de una misma especie, permitiendo abarcar un rango mayor de preferencias de los consumidores y también diferenciar cultivares.

La calidad de los frutos se puede evaluar en gran medida a través de su firmeza, textura, color, aroma y sabor, es decir, en base a sus propiedades organolépticas. Éstas son fundamentales para la aceptación de los consumidores, principalmente en las frutas que se consumen frescas. El contenido de metabolitos secundarios, como productos de la vía de los fenilpropanoides y de los terpenoides, rige considerablemente su aspecto nutricional y su atractivo sensorial, dado que influyen sobre el pH, el contenido de sólidos solubles totales, de carotenoides, de nutrientes y la actividad antioxidante (Bramley, 2002; Pott et al., 2019).

### 1.3.1 - Sabor y aroma

El sabor y el aroma de los frutos están íntimamente relacionados, y variaciones en uno provocan cambios en el otro. Los principales compuestos percibidos por el sentido del gusto son azúcares y ácidos orgánicos que resaltan la dulzura o acidez, mientras que otras sustancias minoritarias contribuyen a características específicas, como astringencia o amargura. En cuanto a compuestos aromáticos volátiles, los producidos principalmente en frutos son terpenos, alcoholes, aldehídos, furanonas, ésteres y cetonas (Pérez & Sanz, 2008).

Los terpenoides (o isoprenoides) participan en muchos aspectos del desarrollo vegetal, ya que son componentes de vías metabólicas que incluyen la biosíntesis de hormonas, pigmentos y metabolitos secundarios volátiles aromáticos. Existen 2 vías de biosíntesis: la citosólica (vía del ácido mevalónico, o MVA) y la plastídica (vía del 2-C-metileritritol 4-fosfato, o MEP), (Toll, 2014). En la primera, los sustratos iniciales son dos moléculas de acetil-CoA, mientras que en la segunda, son el piruvato y el gliceraldehído-3-fosfato. En los dos casos, los productos obtenidos son isopentenil-pirofosfato y dimetilalil pirofosfato.

### 1.3.2 - Firmeza: pared celular

Todas las células vegetales se encuentran recubiertas por una estructura compuesta mayoritariamente por polisacáridos llamada pared celular. Existen dos tipos: pared celular primaria y secundaria. La primera se encuentra principalmente en células en crecimiento, mientras que la segunda se forma cuando las células alcanzaron su tamaño y forma final (Brown et al., 2005). Su función no sólo es proporcionar resistencia a la planta, sino que interviene en diversos procesos como crecimiento, diferenciación celular, comunicación intercelular y actúa como barrera defensiva contra patógenos (Cosgrove, 2005).

La pared celular primaria está compuesta principalmente por polisacáridos y en menor proporción por glicoproteínas. Dentro de los polisacáridos se encuentran las pectinas, las hemicelulosas y la celulosa. Las dos primeras dan lugar a la matriz de la pared celular, donde se van a disponer los demás componentes de la misma (Fry, 2004). Las pectinas constituyen alrededor de un tercio de las macromoléculas de la pared

celular. Son polisacáridos complejos que contienen residuos de Ácido  $\alpha$ -D-galacturónico unidos a Ramnogalacturonanos RG-I y RG-II, enriquecidos en ramnosa, arabinosa y galactosa (Fry, 2004). La alta concentración de pectinas proporciona rigidez a la célula durante su crecimiento, lo que favorece su resistencia a la compresión (P. Darley et al., 2001). Tienen la capacidad de gelificar, modificar el pH y la fuerza iónica de la matriz y participan activamente en varios procesos fisiológicos, como la separación celular durante la abscisión de frutos y hojas o el ablandamiento en la maduración de frutos (Willats et al., 2001).

Las hemicelulosas son polisacáridos muy abundantes en la pared celular vegetal (constituyen alrededor de un tercio de la misma) y están formadas por una estructura unida mediante enlaces  $\beta$ -(1 $\rightarrow$ 4) glucosídicos. Engloban a xiloglucanos, heteroxilanos, heteromananos y otros glucanos con enlaces mixtos (Pauly et al., 2013). La celulosa está constituida por moléculas de glucosa unidas mediante enlaces glucosídicos tipo  $\beta$ -(1 $\rightarrow$ 4) y tiene la capacidad de formar microfibrillas de una longitud suficiente para rodear la circunferencia celular, que se intercalan en la matriz de polisacáridos, formando una red muy resistente (Cosgrove, 2005).

Las expansinas comprenden un grupo de proteínas no enzimáticas encargadas principalmente de la relajación de la pared celular (McQueen-Mason & Cosgrove, 1995). Intervienen en diversos puntos del desarrollo vegetal, como la regulación de la firmeza de los frutos (Brummell et al., 1999; Minoia et al., 2015), desarrollo radical (D.-K. Lee et al., 2003; Z. Yu et al., 2011) y crecimiento en general (Kwon et al., 2008). Por su parte, las  $\beta$ -galactosidasas son una familia de hidrolasas de la pared celular que intervienen activamente en el ablandamiento de frutos. Su función es romper los enlaces de moléculas de galactosa en cadenas laterales de xiloglucanos, RG-I y hemicelulosas, facilitando el acceso a la matriz de pectina a las demás enzimas degradativas, lo que finalmente se traduce en ablandamiento del fruto (Hangkong Liu et al., 2018). Otras enzimas que contribuyen a este proceso son las pectin metilesterasas (Xue et al., 2020), celulasas (Sexton et al., 1997) y xiloglucano endotransglucosidasa (Z. Zhang et al., 2017).

La integridad de la pared celular está íntimamente relacionada con las propiedades organolépticas de los frutos, más precisamente, con la textura y firmeza de los mismos, dos de las cualidades más importantes para los consumidores. Con la

progresión de la maduración, la pared celular tiende a desensamblarse y como consecuencia se produce el ablandamiento. Este fenómeno también influye en la vida poscosecha de los cultivos, motivo por el que esta área está siendo permanentemente estudiada (Waldron et al., 2003).

## 1.4 - Aspectos nutricionales

Las frutas y las verduras proporcionan una enorme variedad de nutrientes como vitaminas y minerales, compuestos bioactivos como flavonoides y carotenoides, y fibras dietarias (R. H. Liu, 2013). El componente mayoritario suele ser agua, variando su cantidad según el tipo de especie, seguido por hidratos de carbono, ácidos orgánicos, vitaminas, minerales y en menor medida, proteínas (Vincente et al., 2014). La composición y los valores nutricionales, como también la ausencia de sustancias tóxicas o antinutricionales, son parámetros cruciales para determinar la calidad del producto (Liener, 1994), que contribuye directamente al precio final y la apreciación de los consumidores.

El contenido de minerales influye en las propiedades organolépticas dado que el nitrógeno y el manganeso están relacionados con la coloración verde, tanto en frutos como en hojas (Deckers et al., 1996; Marsh et al., 1996), aunque también se ha reportado que un exceso de fertilizantes nitrogenados producen pérdida de coloración en uvas (Kliwer, 1977) y ablandamiento de frutos de tomate (Prange & DeEll, 2019). Se vio también que el nitrógeno favorece el aumento de sólidos solubles, alterando el sabor (Barringer et al., 1999), y además se reportó en tomate que las frutas con deficiencia de potasio presentaban un menor contenido de carotenoides (Trudel & Ozgun, 1970).

Tanto frutas como verduras contienen una gran variedad de sustancias antioxidantes, como ácido ascórbico, carotenoides, vitamina E y compuestos fenólicos, que son las encargadas de prevenir la oxidación celular descontrolada al eliminar las especies reactivas de oxígeno (ROS). Estas últimas son formas parcialmente reducidas del oxígeno, como peróxidos o superóxidos, con la capacidad de reaccionar con los diversos componentes celulares produciendo daños severos. La incorporación de vegetales a la dieta ayuda a neutralizar estas alteraciones (Vincente et al., 2014).

Otro componente nutricional importante son las fibras dietarias. Se considera como tales a las porciones comestibles de los vegetales que no pueden ser digeridas ni absorbidas en el intestino delgado (Committee, 2001). Los principales componentes de las fibras son celulosa, lignina, pectinas y polisacáridos no digeribles. Sus beneficios para la salud son muy variados. Uno de los más importantes y conocidos es la regulación de la función intestinal, además de favorecer la rápida sensación de saciedad (Vincente et al., 2014).

## 1.5 - Regulación génica durante el desarrollo y la maduración de frutos

Durante su desarrollo y posterior maduración, los frutos sufren una serie de alteraciones morfológicas y fisiológicas: incremento de tamaño, cambios de color, variaciones en su sabor y aroma. Estas modificaciones son posibles gracias a un gran número de precisos eventos moleculares que tienen lugar a lo largo de estos procesos, entre ellas, la regulación de la expresión génica. Por ejemplo, los factores de transcripción MYB están vinculados a la producción de antocianinas, responsables tanto de la coloración de frutos como de la respuesta frente a diversos tipos de estrés (Espley et al., 2007). También se sabe que el aumento en la producción de enzimas relacionadas al desensamblaje de la pared celular fomenta el ablandamiento de los frutos (Muñoz-Bertomeu et al., 2013), y que existe una regulación muy controlada y tejido-específica en proteínas directamente implicadas en la señalización de hormonas, como el caso de los factores de respuesta a etileno (ERF) en frutos de tomate (M. K. Sharma et al., 2010). Si bien la regulación se produce en múltiples niveles, a nivel postranscripcional generalmente se observa participación de algún tipo de ARN pequeño (Kamthan et al., 2015).

## 1.6 - ARNs pequeños

Una amplia variedad de procesos biológicos en plantas están regulados por ARNs pequeños, los cuales son moléculas de ARN reguladoras que poseen típicamente entre 21 y 24 nucleótidos de largo. En general, hay dos clases principales de ARNs pequeños, conocidas como microARNs (miARNs) y *short interfering RNAs* (siARNs). Sin embargo, el vasto número de fuentes de siARNs descrito hasta el momento ha dado lugar a

múltiples subgrupos dentro de la clasificación de siARN, incluyendo *trans acting* (tasiRNAs), *phased* (phasiARNs), *heterocromatic* (hc-siARNs) y *natural antisense siRNAs* (nat-siARNs), entre otros (Bologna & Voinnet, 2014; Borges & Martienssen, 2015).

Los genes *MIR*, que codifican precursores de microARNs, son transcritos por la ARN polimerasa II (Y. Lee et al., 2004), dando lugar a una estructura de tallo y burbuja (pri-microARN, por *primary microRNA*) que es procesada inicialmente en el núcleo por las proteínas *Dicer-like 1* (DCL1, ribonucleasa de tipo III), *HYL1* (*hyponastic leaves 1*, proteína de unión a ARN doble hebra) y *SE* (*Serrate*, una proteína con motivos dedos de zinc). La primera se encarga de cortar los extremos de la estructura y las demás actúan estabilizándola (Fig. 1.2; Dong et al., 2008). En primera instancia se reduce el tamaño de esta molécula generando un pre-microARN (*precursor microRNA*) al que luego el mismo complejo proteico le escinde su bucle para dar lugar a una molécula que tiene un tamaño final de 21-nt, generalmente, cuyos extremos 3' serán metilados por la proteína HEN1, para evitar su degradación, y exportados al citoplasma con ayuda del transportador HST1 (X. Yang & Li, 2012). Una vez fuera del núcleo, una de las dos hebras va a asociarse a una proteína AGO (*Argonauta*) y algunas otras proteínas estabilizadoras para formar el Complejo Silenciador Inducido por ARN (*RISC*, por sus siglas en inglés) que será el responsable de la actividad reguladora (Kidner & Martienssen, 2005).

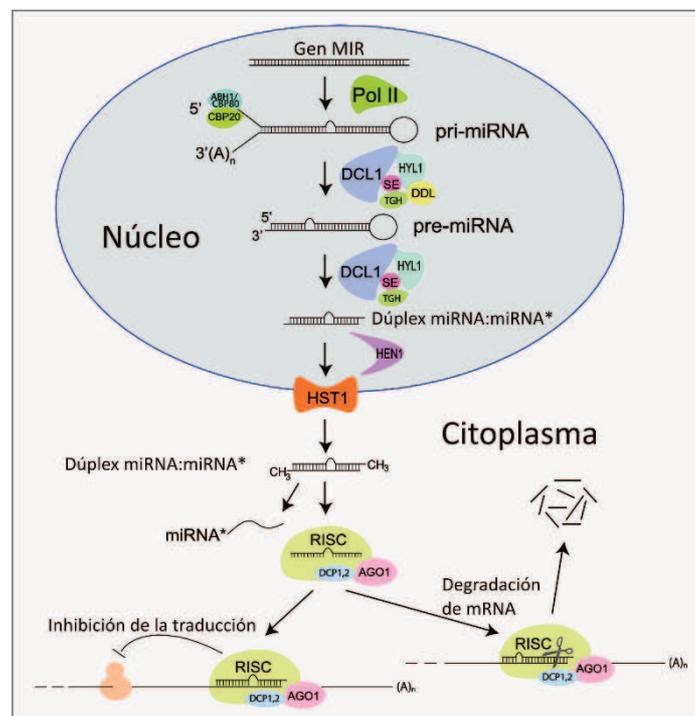


Fig. 1.2: Modelo simplificado de la biosíntesis de miRNAs en plantas. Adaptado de Yang & Li (2012).

### 1.6.1 - ARNs pequeños en fase

Los miARNs pueden desencadenar la producción de otros ARN pequeños en fase, entre los que se encuentran los tasiARNs. Estos ARNs pequeños se originan a partir de genes *TAS* que son transcritos por la ARN polimerasa II y poseen uno o dos sitios de unión para un miARN (Fig. 1.3A). Luego de efectuado el corte en el ARNm por uno (modelo “*One hit*”) o por dos miARNs (modelo “*Two hits*”) se sintetiza la doble hebra de ARN correspondiente con la ayuda de las proteínas *SGS3 (Suppressor Of Gene Silencing 3)* y *RDR6 (RNA-Dependent RNA Polymerase 6)*. Posteriormente la enzima *DCL4* será la encargada de cortar cada 21 nucleótidos, dando lugar a pequeñas moléculas de ARN que se encuentran en fase respecto al sitio de corte del miARN (entre las bases 10 y 11 del mismo) y característicamente, las hebras de cada par están desfasadas por 2 nucleótidos (Fig. 1.3A). Finalmente la estructura es separada y una simple hebra del ARN pequeño formará parte del complejo RISC (Qili Fei et al., 2013). Dentro de la clasificación de ARNs en fase existe una división entre phasiARN y tasiARN. La principal diferencia entre ellos radica en que los tasiARN surgen de un ARNm no codificante cuyo único fin es dar lugar a estas pequeñas moléculas (genes *TAS*, Fig. 1.3A) mientras que los phasiARN son originados desde un transcripto que codifica para proteínas (Fig. 1.3B; Qili Fei et al., 2015).

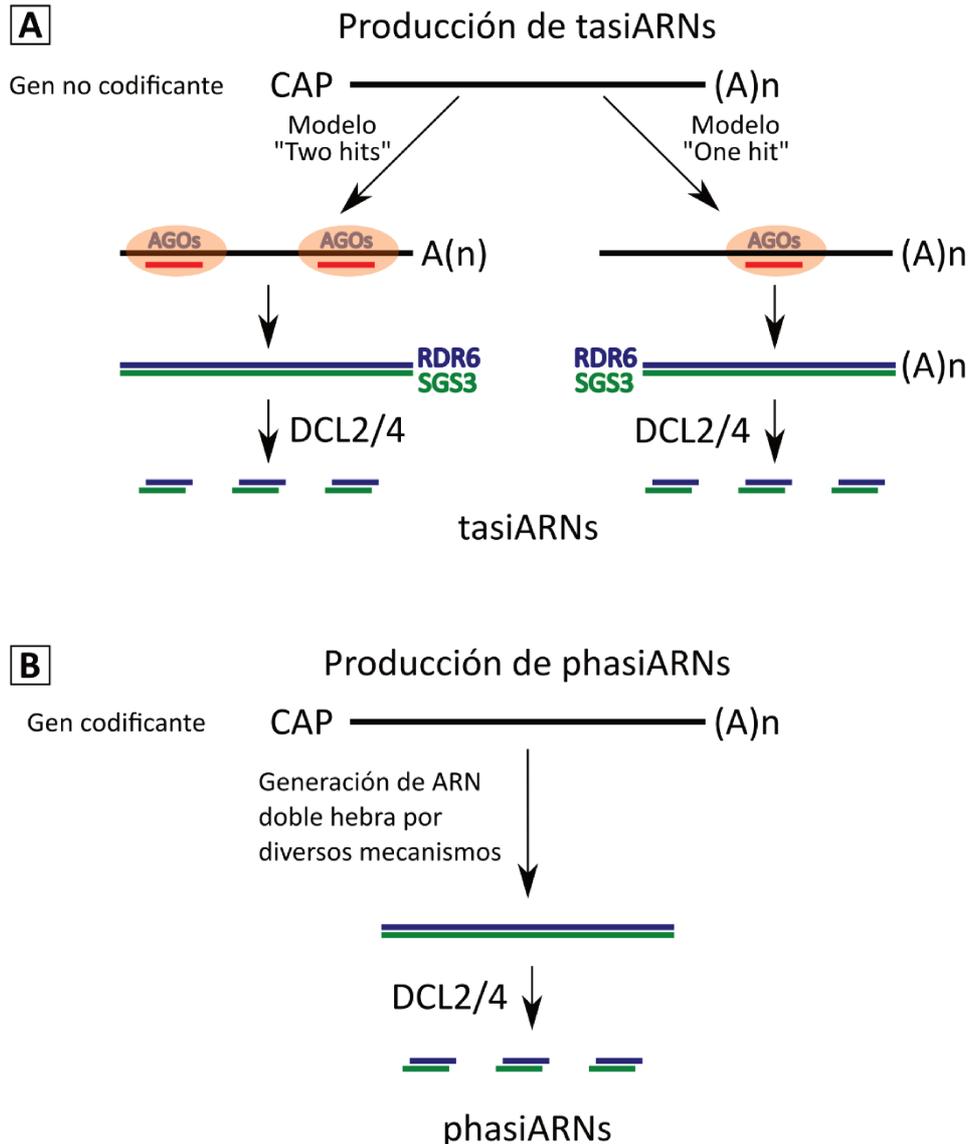


Fig. 1.3: A) Producción de tasiARNs a partir de los modelos "Two hits" y "One hit". B) Producción de phasiARNs a partir de un gen codificante. Adaptado de Yu et al. (2017).

Si bien también se ha postulado que el corte por miARNs da origen a phasiARNs, este mecanismo ha sido caracterizado principalmente para los phasiARNs producidos en órganos reproductivos de gramíneas (Das et al., 2020). Se ha verificado la existencia de phasiARNs también en otros tejidos vegetales, y se cree que la existencia de miARNs iniciadores no es el único mecanismo capaz de producir phasiARNs (Feng et al., 2019). Existen diversos mecanismos moleculares por los que puede producirse ARN doble hebra, sustrato de las enzimas de la familia DCL productoras de ARNs pequeños en fase. Por ejemplo, pueden generarse a partir de ARNs largos no codificantes (lncARNs), que se sintetizan en sentido contrario al del transcripto codificante a partir de su propio

promotor (Qili Fei et al., 2013). Los ARNs pequeños derivados de transcritos naturales antisentido (NAT-siARNs) se sintetizan a partir de la superposición de transcritos independientes, resultando también en regiones de ARN doble hebra susceptibles a la acción de enzimas DCL (Vazquez et al., 2010). Además, a través de estudios bioinformáticos se han detectado *loci* productores de phasiARNs sin un miARN disparador, por los mecanismos de origen son aún desconocidos en muchos casos (Feng et al., 2019).

En cuanto a los mecanismos regulatorios de las distintas clases de ARNs pequeños, se ha demostrado su participación en varios aspectos del crecimiento y desarrollo de las plantas, como respuesta a estrés, polaridad de las hojas, tiempo de floración (D. Jin et al., 2013; Mallory & Vaucheret, 2006; Nag & Jack, 2010; Noman et al., 2017) y resistencia contra enfermedades (Zhai et al., 2011), a través de mecanismos regulatorios que en general tienen lugar a nivel post-transcripcional. Además de actuar a este nivel, algunos ARNs pequeños pueden ejercer su función a través del mecanismo de metilación del ADN dirigida por ARN (RdDM, por sus siglas en inglés), que mediante la metilación de residuos de citosina provoca el silenciamiento génico de un determinado *locus*, actuando en estos casos a nivel transcripcional. Es un proceso regulatorio muy importante para el silenciamiento de elementos transponibles y para la estabilidad del genoma. (J. a Law & Jacobsen, 2010).

Para la biogénesis y el correcto funcionamiento de cada ARN pequeño se necesita una serie de proteínas particulares. Existen 3 familias génicas caracterizadas por su participación en estos procesos: AGO, DCL y RDR. Para la síntesis y actividad de miARNs son necesarias las proteínas AGO1 y DCL1 (Bologna & Voinnet, 2014), mientras que en la producción de phasi y tasiARNs intervienen AGO1, AGO7, DCL4 y RDR6 (Dotto et al., 2014). Por otro lado, durante el proceso de RdDM participan proteínas de las familias DCL, AGO y RDR, particularmente DCL3, AGO4 y RDR2, como así también factores de copia simple, como SGS3 y NRPD2.

Cabe destacar que también existen otros tipos de ARNs pequeños, como por ejemplo, los ARNs pequeños interferentes heterocromáticos (het-siARNs), de 24 nucleótidos de longitud, generados por una ruta diferente de biosíntesis e implicando otras proteínas de estas familias, como por ejemplo, AGO4, junto con otras vías de

silenciamiento, como RdDM a nivel ADN (Sunkar & Zhu, 2007). Por su parte, se ha reportado que los NAT-siARNs mencionados previamente, pueden presentar tamaños entre 21 y 41 nucleótidos y que en su producción participan diversas proteínas, como DCL1, DCL2, DCL4, RDR6 y componentes de la ARN polimerasa IV dependiente de ADN (Vazquez et al., 2010).

### 1.6.2 - Importancia de los ARN pequeños

Los ARNs pequeños participan de importantes procesos vinculados a diversos aspectos del desarrollo de las plantas. Por ejemplo, los tasiARNs producidos a partir del *locus TAS3* y miR390 tienen como genes blanco a diversos factores de respuesta a auxinas (ARF), por lo que son fundamentales en la regulación del crecimiento vegetal, desarrollo radicular y foliar (Xia et al., 2017). Hay evidencia de phasiARNs producidos a partir de genes codificantes para proteínas PPR, NB-LRR y MYB, las cuales abarcan funciones muy variadas, desde defensa frente a patógenos hasta respuesta a la luz UV (Ambawat et al., 2013; DeYoung & Innes, 2006; Q. Fei et al., 2013).

Sin embargo, el conocimiento sobre la participación de ARNs pequeños en la adquisición de propiedades organolépticas es muy escaso. Por ejemplo, se ha vinculado a miR156 a la coloración del fruto de tomate mediante la regulación del gen *colorless non-ripening* (CNR; Silva et al., 2014) varios miARNs han sido relacionados a la regulación de genes codificantes para proteínas encargadas del metabolismo del glutamato y biosíntesis de terpenos, moléculas asociadas a la adquisición de sabor (Karlova et al., 2013; Mohorianu et al., 2011). Por su parte, los tasiARNs generados a partir del *locus TAS4* mostraron como genes blanco de regulación a factores de transcripción MYB, implicados en la regulación de la síntesis de antocianinas en Arabidopsis (Q.-J. Luo et al., 2012), en manzana (Xia et al., 2012) y en hojas de tomate (A. Singh et al., 2016). La poca información disponible sobre este tipo de regulaciones motivó nuestro estudio sobre esta ruta y otras desconocidas que pudieran estar participando en el desarrollo y adquisición de propiedades organolépticas de frutos y de otros atributos de interés comercial.

Para contribuir al entendimiento de los procesos moleculares relacionados con los aspectos mencionados de la comercialización de frutos, en esta tesis doctoral se

llevaron adelante análisis exploratorios para identificar la existencia de rutas de ARNs pequeños que regulen los aspectos metabólicos y fisiológicos relacionados con las cualidades comerciales de los frutos de tomate, frutilla y naranja, a nivel genómico. Se utilizaron bibliotecas de secuenciación masiva de frutos (ver más adelante) de diferentes estadios de maduración para cada especie analizada. La intención de este trabajo es recopilar información que permitirá identificar ARNs pequeños que puedan usarse como herramientas agro-biotecnológicas para el mejoramiento de aspectos vinculados a la calidad de los frutos y otros aspectos importantes para la comercialización. Las posibles aplicaciones de esta información son muy variadas y abarcan distintos enfoques experimentales: modificación por transgénesis de *phas*ARNs o targets específicos (A. Singh et al., 2015, 2019), modificación por transgénesis a través de la obtención de *loci* PHAS artificiales que regulen varios targets a la vez (Cisneros & Carbonell, 2020), creación de líneas transgénicas denominadas “*mimicry*” para capturar de ARNs pequeños específicos para disminuir su actividad (Reichel et al., 2015), modificación por mutaciones de uno o varios genes simultáneamente utilizando tecnología CRISPR (R. Li et al., 2018; Xindi Li et al., 2018; T. Wang et al., 2019) y posiblemente, aplicaciones exógenas de ARNs pequeños que regulen uno o varios transcriptos importantes para mejorar la calidad sensorial de los frutos (Dalakouras et al., 2020; Dubrovina & Kiselev, 2019), entre otras.

## 1.7 - Secuenciación del ADN y tecnología Next Generation Sequencing (NGS)

La secuenciación del ADN revolucionó el campo de la biología molecular. Desde su descubrimiento en los años 70 hasta la actualidad ha evolucionado constantemente, hasta convertirse en, prácticamente, un ensayo de rutina para un gran número de experimentos. Particularmente, estos ensayos se masificaron en los últimos 15 años gracias a la aparición de nuevas tecnologías que permitieron la secuenciación de bibliotecas a precios cada vez más accesibles, sumado esto a la obtención de genomas de referencia de mejor calidad en variadas especies (J.-E. Kim et al., 2014; Phan et al., 2019).

Las dos primeras técnicas empleadas para este fin surgieron en la década del 70': el método de secuenciación de Maxam–Gilbert, desarrollado por Allan Maxam y Walter

Gilbert, y el método de secuenciación de Sanger, propuesto por Frederick Sanger y sus colaboradores. El primero implica marcaje radiactivo de las bases y modificación química específica para cada una de ellas, seguida por una ruptura de la cadena en la base alterada, mientras que el segundo conlleva el agregado de dideoxinucleótidos para detener la polimerización del ADN. En ambos casos se necesita luego una corrida electroforética para poder inferir la secuencia de nucleótidos (Maxam & Gilbert, 1992; F Sanger et al., 1977). Estos desarrollos impulsaron los primeros estudios genómicos más complejos, de hecho, Sanger y su equipo lograron secuenciar completamente el genoma del bacteriófago  $\phi$ X174 en 1977 (Frederick Sanger et al., 1977), utilizando su propia metodología, que le valió un premio Nobel en 1980.

El primer secuenciador automático fue construido en 1986 por la empresa Applied Biosystems en colaboración con el laboratorio de Leroy Hood del Instituto Tecnológico de California (Caltech) y fue llamado AB370A. Tenía la capacidad de secuenciar 500 kilopares de bases por día y fragmentos de hasta 600 pares de bases (Smith et al., 1986), basado en la estrategia de Sanger pero agregando marcadores fluorescentes a cada base nitrogenada. El desarrollo de equipos más potentes llevó a que en 1990 comience a secuenciarse el genoma humano, tarea que tardó 13 años con la tecnología disponible al momento (Collins et al., 2003).

### 1.7.1 - Next Generation Sequencing

A partir del año 2000 surge un nuevo enfoque, la pirosecuenciación, que a diferencia de métodos anteriores utilizaba una reacción enzimática que permitía medir y cuantificar la cantidad de pirofosfato generado durante la síntesis de nucleótidos. En el año 2005 la compañía 454 Life Sciences, tras una serie de adaptaciones, lanzó una versión comercial de secuenciadores basados en este principio, dando lugar a la “segunda generación” o “Next generation” de secuenciadores de ADN (Margulies et al., 2005). Varias empresas sumaron sus equipos en esta época, destacándose los equipos de Illumina y de Ion Torrent, cada una con nuevas metodologías que con el paso de los años permitieron disminuir drásticamente los costos de análisis y aumentar la fidelidad de los mismos. Se estima que la eficiencia de secuenciación aumentó aproximadamente 100000 veces con estas nuevas tecnologías (Treangen & Salzberg, 2012).

Los secuenciadores actuales integran la llamada “tercera generación”. El primero de estos equipos fue desarrollado en 2008 por la empresa Oxford Nanopore Techniques, mientras que al año siguiente Pacific Biosystems estrenó su nuevo dispositivo. Si bien cada uno tiene su mecanismo particular, se diferencian de las tecnologías previas dado que no se necesita un paso de amplificación durante la preparación de las muestras, permitiendo así la secuenciación de una molécula individual, y alcanzando una longitud de lecturas superior a las previamente reportadas (Bleidorn, 2016).

### 1.7.2 - Tipos de experimentos de NGS

Los avances tecnológicos permitieron el desarrollo de una gran variedad de experimentos utilizando técnicas de secuenciación masiva. Posiblemente el más empleado sea el *RNA-seq*, que consiste en secuenciar masivamente el ARN de una muestra en condiciones determinadas (Zhong Wang et al., 2009). Esto permite comparar niveles de expresión génica en distintas situaciones o a lo largo del tiempo, evaluar transcritos producto de splicing alternativo, entre otros análisis. Una variante de esta técnica es el *small RNA-seq*, que requiere un filtrado previo de especies de ARN pequeños durante la preparación de las bibliotecas que luego son secuenciadas. Esto permite identificar y comparar la expresión de especies no codificantes, como miARNs, ARNs pequeños o *long non-coding RNAs* (Pantaleão & Ozanne, 2018). A su vez, puede analizarse el patrón de metilación del ADN de una muestra mediante secuenciación masiva, previo tratamiento con bisulfito (*Bisulfite sequencing*; Cokus et al., 2008), y también puede acoplarse un experimento de NGS a una inmunoprecipitación de cromatina para detectar sitios de unión de proteínas específicas al ADN y modificaciones de histonas (*ChIP sequencing*; Park, 2009).

### 1.7.3 - Análisis Paralelo de Extremos de ARN o Degradoma

Una técnica muy utilizada para trabajar con ARNs pequeños es la validación de genes blanco mediante bibliotecas de Degradoma o Análisis Paralelo de Extremos de ARN (PARE, por sus siglas en inglés; German et al., 2009). Consiste en añadir un cebador específico en su extremo 5' monofosforilado a todas las moléculas de ARN poliadeniladas presentes en una muestra, de manera de capturar aquellos fragmentos de ARN sin CAP de 7-metil guanosina en su extremo 5', que corresponden a transcritos que han sido cortados por acción regulatoria de un ARN pequeño o a aquellos que están

en proceso de degradación. Luego se genera la primera hebra de *cDNA* mediante transcripción reversa usando un oligo (dT) con un adaptador en su extremo 3'. El producto se digiere con la enzima *MmeI*, cuyo sitio de reconocimiento se encuentra contenido en el adaptador ligado al extremo 5', para obtener fragmentos de igual tamaño, que son recuperados por electroforesis en gel de poliacrilamida (PAGE). A estos fragmentos se les liga un adaptador 3' de ADN doble hebra con nucleótidos degenerados. Los pasos finales implican amplificación por PCR, purificación con geles y por último secuenciación de alto rendimiento (Fig. 1.4)

Los datos obtenidos tras la secuenciación de estas bibliotecas pueden analizarse con herramientas bioinformáticas, para lo cual se seleccionó el programa CleaveLand para utilizar en este trabajo de tesis (Addo-Quaye et al., 2009). Se necesita un genoma de referencia, la secuencia de los posibles transcritos regulados por los ARNs pequeños de interés y una base de datos con las secuencias de los ARNs pequeños en estudio. El degradoma es alineado contra los transcritos de referencia, y luego se utiliza el programa GSTAr para identificar posibles sitios de unión entre los ARNs pequeños y el conjunto de posibles transcritos target. Se ha determinado que en el mecanismo de regulación por corte de los ARNs pequeños que forman parte de complejos RISC, la actividad AGO guiada por un ARN pequeño produce un corte entre los nucleótidos 10 y 11 del ARN pequeño, por lo que se puede inferir con exactitud el sitio de corte en las uniones predichas entre ARNs pequeños y transcritos. Gracias a esto, cuando la abundancia de las lecturas de degradoma que coinciden con el sitio predicho de corte para un par "ARN pequeño-transcripto target" está por encima de la abundancia de lecturas de base, se considera evidencia de que el ARN pequeño es un elemento regulador de un determinado transcripto.

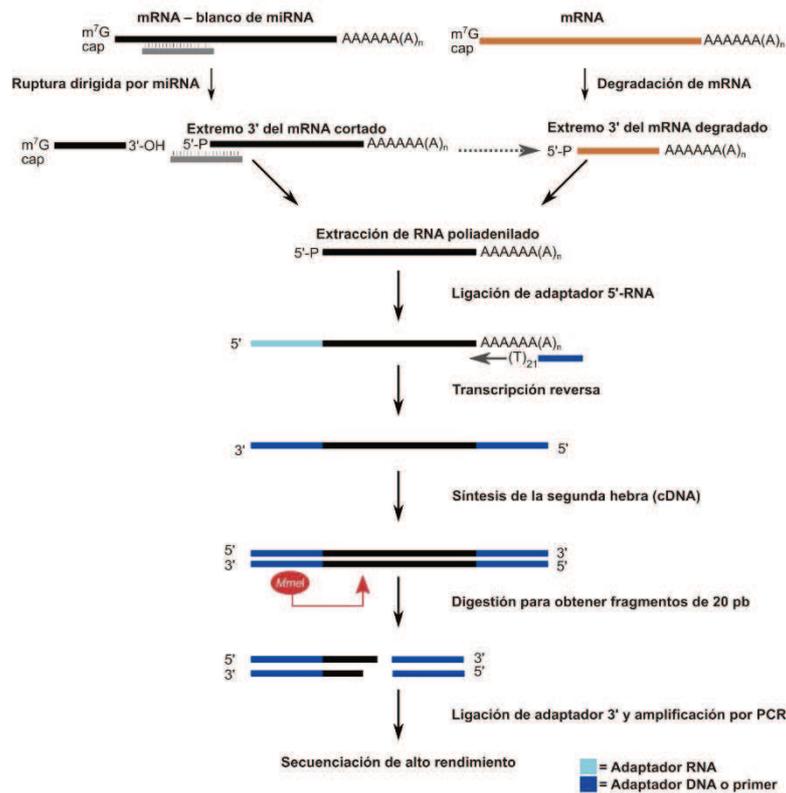


Fig. 1.4: Representación esquemática del procedimiento de construcción de biblioteca de degradomas. Adaptado de German et al. (2009).

Para representar la abundancia de *reads* detectadas en el degradoma se usan gráficos denominados T-plot, que relacionan las abundancias de todas las *reads* presentes en el degradoma que son alineadas a un determinado transcripto con respecto a la abundancia de *reads* que coincide exactamente la posición predicha de corte por un ARN pequeños. De esta manera se puede distinguir fácilmente entre cortes realizados por una especie de ARN pequeño reguladora (porque presentan mayor abundancia de *reads* coincidentes con el sitio predicho de corte), de la degradación natural de los ARNs mensajeros. Además, el programa clasifica los gráficos T-plot en 5 categorías para cada transcripto blanco de acuerdo a la “altura” del pico que coincide con el corte del transcripto mediado por el ARN pequeño (X. Xu et al., 2013):

- Categoría 0: varias *reads* en el sitio predicho de corte y su abundancia es la máxima en toda la longitud del transcripto (situación óptima).
- Categoría 1: varias *reads* en el sitio predicho de corte y su abundancia la máxima del transcripto, pero hay más de un máximo.

- Categoría 2: varias *reads* en el sitio predicho de corte y su abundancia es mayor a la mediana del transcripto, pero menor a su máxima.
- Categoría 3: varias *reads* en el sitio predicho de corte pero su abundancia no supera a la mediana del transcripto.
- Categoría 4: sólo una *read* en el sitio predicho de corte.

#### 1.7.4 - Herramientas bioinformáticas

Las tecnologías de NGS lograron que disminuyeran los costos de secuenciación y favorecieron la producción de un gran volumen de información, generando un desafío en cuanto al análisis de la misma. La bioinformática, área que emplea métodos computacionales para solucionar problemas biológicos (Hidalgo Suarez et al., 2018), evolucionó a la par de estas nuevas tecnologías, convirtiéndose en esencial para trabajar con ellas. Constantemente surgen nuevos programas para manipular la gran cantidad de datos producidos, aunque lamentablemente no suelen ser herramientas amigables con el usuario informático promedio, por lo que se requieren analistas especializados para su utilización.

El software bioinformático existente es muy variado, lo que permite la elección de aquellas herramientas que se adecúen a la necesidad de cada experimento en particular. En líneas generales, estos programas posibilitan el procesamiento de las bibliotecas de NGS (eliminación de adaptadores, control de calidad), su alineamiento contra referencias, ensamblado y anotación de genomas, visualización genómica, análisis de dominios y plegamientos proteicos, test filogenéticos, análisis estadísticos, entre muchas otras tareas (J. Zhang et al., 2011).

### 1.8 - Consideraciones generales sobre los análisis realizados

Teniendo en cuenta lo expuesto en esta Introducción general, como parte de esta tesis doctoral se llevó adelante un conjunto de análisis bioinformáticos sobre datos obtenidos por secuenciación NGS. En todos los casos se trabajó con conjuntos de datos de secuenciación de *small RNA-seq*, *RNA-seq* y de degradoma, provenientes de muestras de frutos de un mismo cultivar, para cada una de las tres especies en estudio (tomate, frutilla y naranja). Esto es importante debido a que algunas de las propiedades organolépticas de los frutos varían considerablemente entre distintos cultivares de una

misma especie. Además, se seleccionaron tres especies de frutos con características organolépticas muy diferentes entre sí, con el fin de ampliar el horizonte de búsqueda de rutas regulatorias mediadas por ARNs pequeños relacionados con un amplio espectro de cualidades importantes para la comercialización de frutos. De esta manera, si bien inicialmente se orientaron los análisis a la identificación de rutas regulatorias de las propiedades organolépticas de estas especies, en el caso de *C. sinensis* también se realizó la caracterización a nivel genómico de las familias génicas relacionadas con los procesos de ARNi en esta especie y se estudió la relación entre las rutas de ARNs pequeños y el proceso de abscisión.

## 1.9 - Hipótesis de trabajo

Existen módulos regulatorios en los que participan ARNs pequeños que contribuyen a la adquisición y el desarrollo de atributos de importancia comercial en frutos de tomate, frutilla y naranja.

### 1.10 - Objetivo general

El objetivo general del proyecto es identificar y estudiar la función de ARNs pequeños que participen en el establecimiento de atributos de importancia comercial, de manera comparativa en tomate (*Solanum lycopersicum*), frutilla (*Fragaria vesca* y *Fragaria x ananassa*) y en naranjo (*Citrus sinensis*).

#### 1.10.1 - Objetivos específicos

- Analizar datos de secuenciación de última generación (*Next Generation Sequencing*, NGS) de bibliotecas de ARNs pequeños de frutos de tomate, frutilla y naranjo, y predecir genes target a lo largo de la maduración.
- Identificar *loci* productores de ARNs pequeños en fase en frutos de tomate, frutilla y naranja.
- Predecir y validar genes blanco regulados por los ARNs pequeños identificados, a través del uso de datos de secuenciación de bibliotecas de degradoma.
- Estudiar a nivel genómico y filogenético a las familias de proteínas implicadas en la biosíntesis y actividad de ARNs pequeños en *Citrus sinensis*.
- Identificar ARNs regulatorios relacionados con atributos de importancia comercial para los frutos de tomate, frutilla y naranja.

## Capítulo 2: ARNs pequeños en tomate

## 2.1 - Introducción

### 2.1.1 - Familia Solanaceae

La familia Solanaceae está constituida por aproximadamente 90 géneros y entre 3000 y 4000 especies (Knapp et al., 2004), dentro de las cuales se encuentran tomate (*Solanum lycopersicum*), papa (*S. tuberosum*) y berenjena (*S. melongena*) como miembros destacados por su importancia agronómica y económica, dado que su consumo está ampliamente distribuido a nivel mundial. Por otro lado, algunas de estas especies son empleadas como modelo de estudio: el tomate es modelo de estudio de desarrollo de frutos carnosos y maduración climatérica (Alexander & Grierson, 2002; Kimura & Sinha, 2008), diversas variedades de Nicotiana se emplean para estudiar interacciones con patógenos (Goodin et al., 2008; L. Wang & Wu, 2013), la papa es el modelo de estudio de tubérculos por excelencia (Prat et al., 1990) y la planta ornamental petunia (*Petunia hybrida*) se emplea para estudiar la pigmentación de las flores (Spelt et al., 2002), además de haber sido esta la especie donde se descubrieron los ARNs de interferencia (Napoli et al., 1990; Van Der Krol et al., 1990).

### 2.1.2 - Características generales del tomate

El tomate (*Solanum lycopersicum*) es uno de los principales cultivos producidos y consumidos a nivel mundial en numerosas presentaciones, desde su fruto crudo hasta en puré o salsas envasadas (Alam & Goyal, 2007; Arah et al., 2016). Su origen se sitúa en la zona andina de Sudamérica, siendo posiblemente la variedad “cherry” (*Solanum lycopersicum* var. *cerasiforme*) su ancestro salvaje (Rick & Holle, 1990). Fue introducido a Europa por los conquistadores españoles en el siglo XVI y a partir de allí su cultivo se extendió a Asia, África y medio oriente (Naika et al., 2005). Es una planta perenne que crece principalmente en climas templados y puede llegar a medir hasta 3 metros de altura (Perveen et al., 2015). La forma y tamaño de sus frutos es variable: desde 2 cm de diámetro en los tomates cherry hasta alrededor de 10 cm en variedades que producen frutos más grandes. La fruta hace una transición del color verde al rojo durante la maduración, presentando este último diversas intensidades según el cultivar (Banerjee et al., 2018).

China, India, Estados Unidos y Turquía son los principales productores de esta fruta a nivel mundial (Informe de la Food And Agriculture Organization, 2014). En 2018, sólo China fue responsable del 35,5% de la producción, mientras que estos 4 países generaron el 60,9% del total de tomate consumido globalmente. Argentina ocupa la posición 30 a escala internacional y el 3er lugar en Latinoamérica, detrás de Brasil y Chile, con una producción cercana a las 653.000 toneladas anuales (un 0,37% del total general). Los principales sectores productivos de nuestro país son la región cuyana, el noroeste y la provincia de Río Negro (Lucero Reinoso & Quiroga, 2019), y aproximadamente el 70% de la producción local tiene destino industrial, en forma de conservas, puré o salsas. Respecto a este último uso, en 2017 el valor de exportación fue de 0.96 U\$D por Kg, lo que evidencia la importancia económica de esta actividad para nuestra nación (Informe del Ministerio de Agroindustria, 2018).

### 2.1.3 - Maduración de frutos de tomate y aspectos nutricionales

En el caso del tomate, la maduración es climatérica, presentando un pico de respiración acompañado por una elevada producción de etileno, en contraste con la maduración no climatérica donde los niveles respiración no presentan grandes cambios y la cantidad de etileno se mantiene baja a lo largo de todo el proceso. Además, los frutos climatéricos pueden ser arrancados de la planta y continuar madurando con la ayuda de etileno exógeno (Alexander & Grierson, 2002).

#### 2.1.3.1 Estadios de maduración de los frutos de tomate

Los estadios de maduración de frutos de tomate se caracterizan de acuerdo a su coloración, independizándose así de la cantidad de días transcurridos desde la antesis, dado que diferentes variedades presentan pequeñas diferencias de tiempo de maduración. Según el Departamento de Agricultura de los Estados Unidos (USDA), el estadio “Verde inmaduro” tiene 0% de coloración roja, “Rompierte” o “Pintón” presenta entre 1 y 10% de su superficie rojiza, “Rosa” posee entre 30 y 60% de pigmentación y “Rojo maduro” está 100% coloreado (Fig. 2.1; Quiroga & Frascina, 1997; USDA, 1991).

Además de adquirir coloración con el progreso de la maduración, los estadios del fruto de tomate varían en su composición de nutrientes, azúcares y acidez. El contenido

de licopeno se incrementa notablemente durante el fenómeno madurativo, mientras que la cantidad de clorofila se reduce. A su vez, la acidez titulable también disminuye, produciendo una pequeña suba del pH, mientras que el contenido de azúcar aumenta (Gautier et al., 2008).

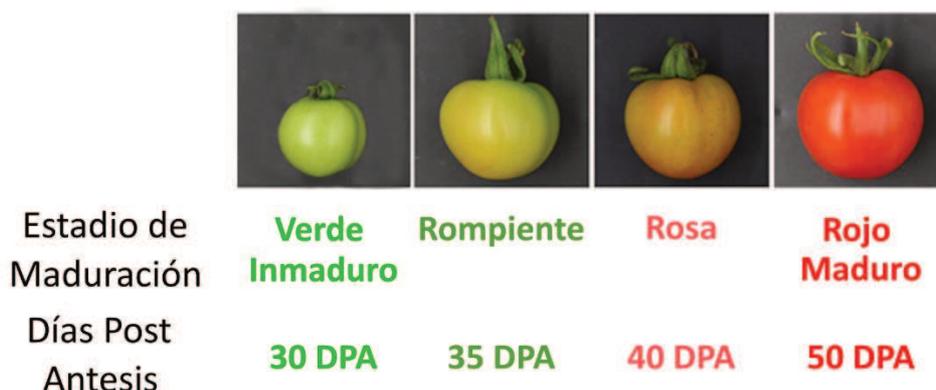


Fig. 2.1: Estadios de maduración de tomate de acuerdo a los días transcurridos luego de la antesis. Adaptado de Li et al. (2018).

### 2.1.3.2 - Aspectos nutricionales de los frutos de tomate

Los tomates contienen una gran diversidad de vitaminas y minerales. La composición de los frutos puede variar levemente entre los diversos cultivares, así como también se ve influida por factores ambientales y por las condiciones en que son almacenados (Nour et al., 2013). Son una fuente importante de oligoelementos, como zinc, que juegan un papel fundamental como cofactor de un gran número de enzimas, y contribuye al correcto funcionamiento del organismo (Informe del National Research Council, 1989). El licopeno es el principal carotenoide presente en esta fruta, y es responsable de su coloración roja en estadios maduros. Una dieta rica en este compuesto tendría efectos defensivos frente a especies reactivas de oxígeno, responsables de daño celular (Viuda-Martos et al., 2014). Además son ricos en  $\alpha$ -tomatina, que debido a su poder antioxidante, junto al licopeno, previenen la aparición de células cancerígenas de varios tipos y mejoran los marcadores coronarios (Friedman, 2013).

### 2.1.3.3 - Aspectos organolépticos de los frutos de tomate

Desde el punto de vista agronómico, existen atributos importantes tales como el tamaño, el rendimiento, la resistencia a diferentes tipos de estrés y la vida poscosecha de los frutos. La selección de cultivares con características adecuadas para la producción agronómica, en algunos casos ha dejado de lado la selección de características organolépticas óptimas desde el punto de vista del consumidor. Por lo tanto, es necesario mejorar estos aspectos para tener en cuenta las preferencias de los consumidores, quienes requieren que el producto tenga cualidades sensoriales agradables, como buen gusto y aroma, color rojo intenso y textura firme. En la definición del sabor participa un gran número de compuestos, pero fundamentalmente está regido por una combinación de azúcares (principalmente fructosa y glucosa), ácidos orgánicos (como cítrico y málico), aminoácidos (como glutamato) y numerosos compuestos volátiles, gran parte de ellos, derivados de fenilpropanoides (Peled-Zehavi et al., 2015; Tieman et al., 2012). En líneas generales, se establece que un tomate apetecible debe tener cierto grado de dulzura y componentes productores de aroma, pero una acidez intermedia (Baldwin et al., 1998; Causse et al., 2010).

Al ser un fruto de maduración climatérica, el etileno es una hormona vegetal muy importante en el establecimiento de las propiedades organolépticas del tomate. Se ha demostrado que la síntesis de etileno promueve el desarrollo de aspectos relacionados con las propiedades organolépticas de los frutos, influyendo sobre la solubilización de pectinas de la pared celular, provocando cambios en su textura (Fig. 2.2; Oeller et al., 1991). La producción de esta hormona deriva también en variaciones en los niveles de compuestos volátiles y azúcares solubles, promoviendo el desarrollo de compuestos aromáticos agradables y de sabor, entre otros cambios fisiológicos que se dan durante la maduración (Fig. 2.2; Alexander & Grierson, 2002).

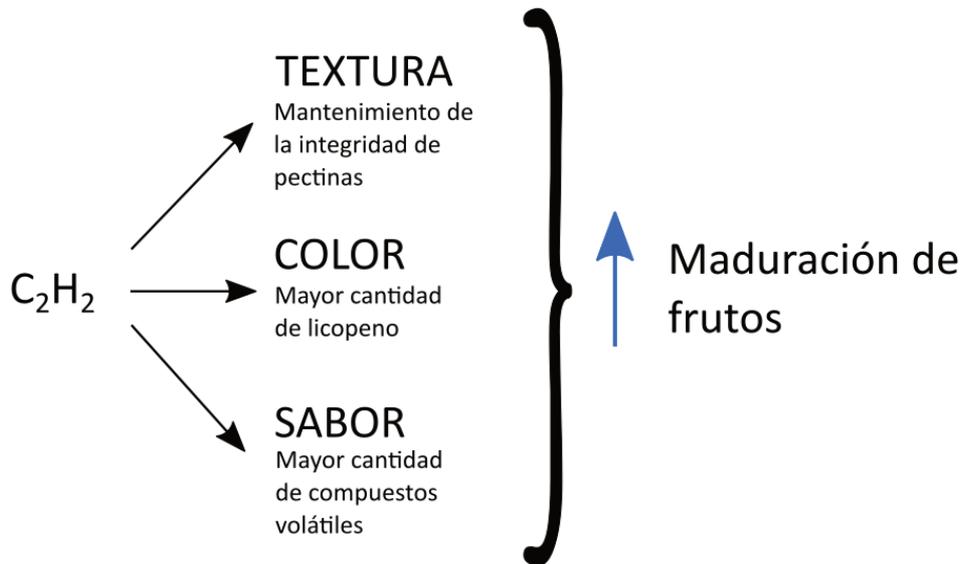


Fig. 2.2: Modelo simplificado de la influencia de etileno sobre aspectos organolépticos de frutos de tomate. Adaptado de Prasanna et al. (2007).

La percepción de esta hormona se produce a través de los factores de transcripción *Ethylene Response Factors* (Factores de Respuesta a Etileno, o ERF por sus siglas en inglés), que son específicos de plantas y actúan en el último punto de regulación de la vía de señalización de etileno. No sólo se inducen por acción de esta hormona, sino que además responden a otras hormonas, como ácido jasmónico, ácido abscísico y ácido salicílico (Pirrello et al., 2006).

Además del etileno, se ha demostrado que las auxinas tienen un papel fundamental en la maduración de los frutos de tomate. Las principales proteínas implicadas en la señalización de esta hormona son los factores de respuesta a auxinas (*AUXIN RESPONSE FACTORS*, o ARFs) y las proteínas inducidas por auxinas (*AUXIN-INDUCED proteins*, o AUX/IAAs), (Breitel et al., 2016). La inhibición de los genes *SIARF2A* y *SIARF2B*, por ejemplo, manifestó defectos madurativos y disminución en la cantidad de etileno producido (Hao et al., 2015). Además, estas dos hormonas interactúan y se regulan mutuamente durante varios puntos del desarrollo vegetal, pero aún se desconoce gran parte de esta interacción (M. Liu et al., 2018).

#### 2.1.3.4 - Vías metabólicas implicadas en la adquisición de propiedades organolépticas

Los principales atributos a considerar cuando se habla de propiedades organolépticas de frutos son sabor, aroma, color y textura. Se conocen las vías metabólicas implicadas en cada uno, y a continuación se describen brevemente las más importantes.

##### a) Sabor y aroma

Estas dos cualidades están muy relacionadas y dependen de la interacción y el contenido de metabolitos primarios y secundarios que involucran azúcares, aminoácidos, ácidos orgánicos y compuestos aromáticos volátiles. Estos últimos derivan del metabolismo de la fenilalanina, los ácidos grasos y los carotenoides. En cada una de estas vías se conocen enzimas regulatorias clave, como la descarboxilasa de aminoácidos aromáticos (AADC) y fenilacetaldehído reductasa (PAR) en la vía de la fenilalanina, dioxigenasas de ruptura de carotenoides (CCD) en la vía de estos pigmentos, 13-lipoxigenasa (13-LOX) y 13-hidroperóxido liasa (13-HPL) en la vía de ácidos grasos (Fig. 2.3; Klee, 2010) Además se destaca la enzima alcohol deshidrogenasa (ADH), que contribuye a la producción de compuestos volátiles al interconvertir aldehídos y alcoholes provenientes de lípidos y aminoácidos (Fig. 2.3; Díaz de León-Sánchez et al., 2009). Dentro de los principales compuestos volátiles que contribuyen al aroma y sabor de los frutos se destacan hexanol y *cis*-3-hexanal, ambos provenientes del metabolismo de ácidos grasos, 2-feniletanol generado a partir del metabolismo de fenilalanina, y  $\beta$ -damascenona, proveniente de la vía de los carotenos, entre otras moléculas representadas de color rojo en la Figura 2.3.

Debido a la gran cantidad de compuestos implicados en la adquisición de sabor, la modificación de uno sólo o unos pocos no produce cambios sustanciales en la percepción de esta propiedad, por lo que la utilización de técnicas de ingeniería genética para mejoramiento, donde se modifica la expresión de un único gen no ha sido muy efectiva.

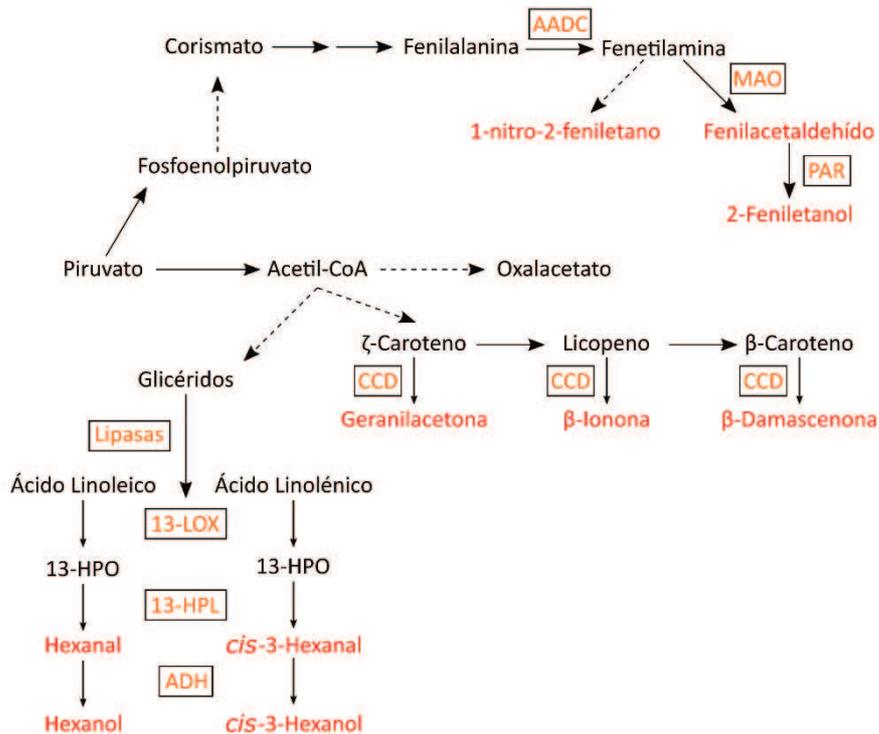


Fig. 2.3: Resumen de las principales vías metabólicas de compuestos volátiles involucrados en la adquisición de sabor y aroma. Los compuestos aromáticos volátiles se muestran en rojo. Las líneas punteadas indican múltiples pasos de reacción. Las enzimas implicadas en cada reacción se indican con un recuadro negro. AADC, decarboxilasa de aminoácidos aromáticos; MAO, monoaminoxidasa; PAR, fenilacetaldehído reductasa; CCD, dioxigenasa de ruptura de carotenoides; 13-LOX, 13-lipoxigenasa; 13-HPL, 13-hidroxiperoxido liasa; ADH, alcohol deshidrogenasa. Adaptado de Klee (2010).

## b) Color

Este atributo ayuda a los consumidores a seleccionar visualmente los frutos, por lo que es de vital importancia para los productores la obtención de tomates con su color rojo característico. Los pigmentos, principalmente los carotenoides, además de cumplir importantes funciones vinculadas a la fotosíntesis y presentar beneficios nutricionales, son los responsables de la coloración de ciertos órganos de las plantas, volviéndolos más atractivos para la polinización o la dispersión de semillas (Ronen et al., 2000). Se sintetizan a partir de la vía de isoprenoides, en los cromoplastos de flores y frutos, proceso que se representa en la Figura 2.4, donde el color del texto indica el color de cada pigmento y las enzimas implicadas se muestran encuadradas a la izquierda de cada paso. El geranyl-geranyl difosfato (GGDP) es el precursor de los principales pigmentos, como carotenos, violaxantina y licopeno, y a su vez este último es el sustrato de una serie de reacciones para producir luteína (Fig. 2.4).

En estadios tempranos, las frutas verdes tienen en su mayoría  $\beta$ -caroteno, luteína, y violaxantina, para luego comenzar a acumular licopeno y adquirir coloración rojiza (Fraser et al., 1994). Durante este proceso hay una fina regulación transcripcional de los genes que codifican para las enzimas fitoeno sintasa (Psy) y fitoeno desaturasa (Pds): los niveles de expresión de Psy y Pds aumentan a medida que progresa la maduración (Bramley, 1997). Los transcriptos codificantes de licopeno  $\beta$ -ciclaza y  $\epsilon$ -ciclaza (CrtL) presentan expresión muy baja al inicio del proceso, pero se incrementa considerablemente en estadios maduros (Pandurangaiah et al., 2016).

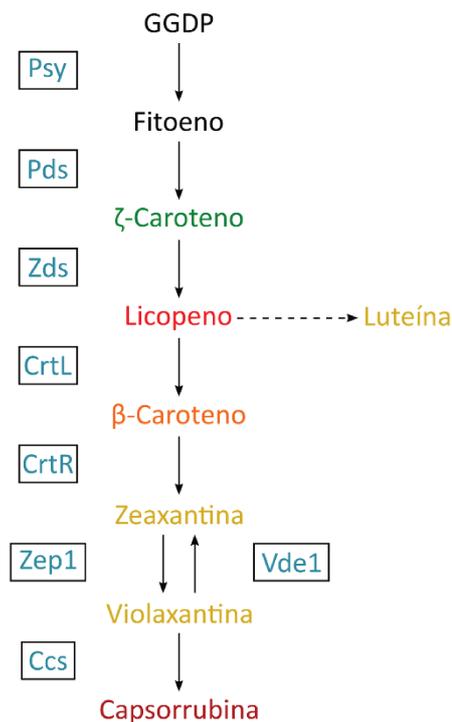


Fig. 2.4: Resumen de la principal vía metabólica de pigmentos involucrados en la adquisición de color. Los compuestos coloreados se representan con su color característico. Las líneas punteadas indican múltiples pasos de reacción. Las enzimas implicadas en cada reacción se indican con un recuadro negro. GGDP, geranil-geranil difosfato; Psy, fitoeno sintasa; Pds, fitoeno desaturasa; Zds,  $\zeta$ -caroteno desaturasa; CrtL, licopeno ciclaza; CrtR,  $\beta$ -caroteno hidroxilasa; Zep1, zeaxantina epoxidasa; Vde1, violaxantina depoxidasa; Ccs, capsantina/capsorrubina sintasa. Adaptado de Ronen et al. (2000).

### c) Textura

La textura no sólo influye en la aceptación de los consumidores, sino que también es importante en la vida poscosecha, la transportabilidad y la resistencia de los frutos. Los cambios en la firmeza del fruto están asociados con una disminución de la rigidez de la pared celular, siendo las proteínas del tipo poligalacturonasas las principales responsables de este fenómeno (Crookes & Grierson, 1983), aunque

participan también otros tipos de enzimas, por ejemplo, las que modifican los distintos componentes de la pared celular primaria, como las hemicelulosas y pectinas (Muñoz-Bertomeu et al., 2013).

La cutícula, de composición lipídica, otorga brillo a los frutos además de resistencia frente a patógenos y a la pérdida de agua (Kosma et al., 2010; Riederer & Schreiber, 2001). Es precisamente a través de la regulación de la hidratación del fruto que actúa como otro controlador de la firmeza de los mismos, reforzando la pared celular y proporcionando extensibilidad al órgano (Domínguez et al., 2011; Lara et al., 2019). Su síntesis se puede dividir en dos grandes procesos: producción de cutina y de ceras cuticulares. En ambos casos participan varios tipos de enzimas, que involucran diversas vías metabólicas (Joubès & Domergue, 2018).

#### 2.1.4 - ARNs pequeños en frutos de tomate

Los miARNs son la subclase de ARNs pequeños más estudiada. Se sabe que muchos de ellos tienen influencia directa en el desarrollo de frutos al regular la expresión génica a nivel postranscripcional, como se ha demostrado para miR159 a través de la regulación de un factor de transcripción GAMYB; para miR156 mediante la regulación de factores de transcripción del tipo *Squamosa Promoter binding protein-like* (SPL) o para miR172 por medio de la regulación del gen *Apetala 2* (AP2; E. M. da Silva et al., 2017; Karlova et al., 2014). Además, también se ha asociado a los miARNs con la maduración de frutos (Moxon et al., 2008), mediante la modificación de su firmeza, como es el caso de sly-miR157 que controla específicamente al gen SPL-CNR (W. Chen et al., 2015), la regulación en su tamaño a través de sly-miR396 y la expresión de factores de transcripción del tipo *Growth-Regulating Factors* (GRFs; Cao et al., 2016), y el fino control de genes como los pertenecientes a la familia *AUXIN RESPONSE FACTORS* (ARFs) por parte de miR160, fundamentales para el correcto desarrollo del proceso madurativo (Cedillo-Jimenez et al., 2020).

Otras clases de ARNs pequeños que han sido vinculados a la resistencia frente a patógenos son algunos siRNAs que aumentan su acumulación ante *Phytophthora sojae* o *Botrytis cinerea* (F. Wu et al., 2017; L. Yang & Huang, 2014). También se han identificado phasiARNs que están involucrados en la respuesta a estrés por sequía, a

través de la intervención de miR1514a para modular la expresión de un factor de transcripción de la familia NAC (Sosa-Valencia et al., 2017) o la promoción de la floración y reproducción por parte de otros *loci* productores de phasiARNs (Y. Liu et al., 2020). Sin embargo, los estudios de estas últimas moléculas regulatorias en tomate suelen realizarse en hojas (W. Luan et al., 2020; F. Wu et al., 2017), y los estudios que analicen su existencia y función en frutos aún son escasos.

### 2.1.5 - Estudios genómicos de ARNs pequeños

El genoma de tomate es diploide y posee 12 pares de cromosomas. La variedad *Heinz 1706* fue la escogida para organizar un genoma de referencia, que recientemente fue actualizado a su versión SL4.0 empleando técnicas de secuenciación y ensamblaje modernas (Hosmani et al., 2019). Cuenta con alrededor de 34.000 genes, de los cuáles cerca de 31.000 están validados por *RNA-seq* y un tamaño aproximado de 900 megapares de bases (Mpb), además de tener muy buenas anotaciones genómicas que facilitan en gran medida trabajar con esta especie (Causse et al., 2013).

En los últimos 15 años, con el desarrollo de nuevas técnicas de secuenciación masiva y la disminución en los costos de las mismas, se incrementó en gran medida la cantidad de estudios sobre ARNs pequeños en numerosas especies (Cui et al., 2017; W. Wang & Luan, 2015; J. H. Zuo et al., 2011). La principal ventaja de esta metodología es que permite analizar en simultáneo un gran número de ARNs pequeños en un contexto genómico y particularmente, en frutos de tomate se ha estudiado la actividad de miARNs en relación a su función como reguladores durante el desarrollo y la maduración de este órgano (Lopez-Gomollon et al., 2012; Moxon et al., 2008), analizando en algunos casos los genes blanco de cada uno de ellos (Karlova et al., 2013). Existen también trabajos que aprovecharon la tecnología de secuenciación masiva para comparar los perfiles de miARNs en frutos de tomate provenientes de plantas salvajes y mutantes para genes específicos ligados a la maduración (C. Gao et al., 2015), y otros que se enfocan mayormente en el etileno como principal responsable de la maduración, regulado por ARNs pequeños y modificaciones epigenéticas (Zhong et al., 2013).

Los demás tipos de ARNs pequeños en tomate no han sido estudiados en tanta profundidad aún, y las investigaciones existentes apuntan fundamentalmente a su relación con la respuesta de la planta frente a diferentes tipos de estrés, utilizando a las

hojas como órgano de análisis (W. Luan et al., 2020; Prigigallo et al., 2019; F. Wu et al., 2017).

En esta tesis doctoral se llevaron a cabo análisis genómicos de ARNs pequeños durante la maduración de frutos de tomate, orientando los esfuerzos a identificar ARNs regulatorios que participen en la adquisición de propiedades organolépticas de frutos y otros atributos de interés comercial, aspectos que aún no han sido caracterizados en tomate y tampoco en la mayoría de los frutos. Este trabajo exploratorio sienta las bases para futuros proyectos que utilicen estos resultados para el diseño de herramientas biotecnológicas que pueden utilizarse para modular la expresión de genes de interés para el mejoramiento de las cualidades organolépticas de los frutos.

## 2.2 - Materiales y métodos

Los comandos utilizados para los procedimientos descritos en esta sección se detallan en el Anexo I.

### 2.2.1 - Datos de secuenciación masiva

Se realizaron búsquedas bibliográficas en Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>) de publicaciones científicas que incluyeran el uso y producción de datos de secuenciación masiva (NGS) correspondientes a ARNs pequeños (*small RNA-seq*) de frutos de tomate en distintos estadios de madurez y que cuenten con datos provenientes de réplicas biológicas. También se hicieron búsquedas de datos de secuenciación masiva en la base de datos Sequence Read Archive (SRA, <https://www.ncbi.nlm.nih.gov/sra>), ya que es posible encontrar datos que no estén asociados a una publicación y por lo tanto no habrían sido identificados a través de búsquedas en Pubmed. Se utilizó la herramienta fastq-dump del paquete SRA-toolkit (<http://ncbi.github.io/sra-tools/>) para convertir el formato SRA original a formato fastq y de esta manera se descargaron los conjuntos de datos de secuenciación de bibliotecas de *small RNA-seq*, de *RNA-seq* o de degradomas de frutos de tomate *S. lycopersicum* cv. Ailsa Craig (Tabla 2.1). Cuando fue posible, se analizaron datos correspondientes al mismo cultivar ya que las propiedades organolépticas de los frutos pueden variar considerablemente entre cultivares de una misma especie.

Tabla 2.1: Números de acceso de los datos de NGS utilizados

Estadio	Small RNA-seq	RNA-seq	Degradoma
<b>Verde inmaduro</b>	SRR1344486 <sup>[1]</sup>	SRR1759281 <sup>[2]</sup>	SRR627531 <sup>[3]</sup>
	SRR1344487 <sup>[1]</sup>		SRR627532 <sup>[3]</sup>
<b>Rompiente</b>	SRR1344488 <sup>[1]</sup>	SRR1759282 <sup>[2]</sup>	
	SRR1344489 <sup>[1]</sup>		
<b>Rosa</b>	SRR1344490 <sup>[1]</sup>	SRR1759283 <sup>[2]</sup>	
	SRR1344491 <sup>[1]</sup>		
<b>Rojo maduro</b>	SRR1344492 <sup>[1]</sup>	SRR1759284 <sup>[2]</sup>	SRR627534 <sup>[3]</sup>
	SRR1344493 <sup>[1]</sup>		SRR627535 <sup>[3]</sup>

Referencias: [1] Gao et al., 2014, [2] Ye et al., 2015, [3] Karlova et al., 2013.

## 2.2.2 - Pre-procesamiento de datos de NGS

Se evaluó la calidad de secuenciación de los archivos fastq descargados utilizando la herramienta FastQC v 0.11.5 (Andrews, 2010). En los casos en que fue necesario filtrar secuencias de acuerdo a su calidad, se utilizó el programa FASTQ Quality Filter v. 0.0.14 del paquete FASTX\_toolkit (Gordon, 2011) conservando las secuencias con un *score* de calidad mayor o igual a 28 (-q 28) en el 85% de su longitud (-p 85). Para cortar la secuencia correspondiente a los adaptadores de secuenciación se utilizó el programa fastx\_clipper v. 0.0.14 (Gordon, 2011), modificando los parámetros por *default* para obtener *reads* de al menos 18 nucleótidos de longitud (opción “-l 18”), mantener solo aquellas *reads* en las que estaba presente el adaptador (parámetro -c) y cortar la secuencia correspondiente al adaptador solo cuando estaban presentes al menos 5 nucleótidos del mismo en una determinada secuencia (opción -M 5).

### 2.2.2.1 - Filtrado de secuencias contaminantes

Con el fin de eliminar las secuencias de ARN ribosomales (ARNr), ARN de transferencia (ARNt), ARN nucleares pequeños (snARN) y ARN nucleolares pequeños (snoARN), que no son de interés para el análisis, se descargó la base de datos de Rfam

14.0 (Griffiths-Jones et al., 2003) de la que sólo se conservaron las secuencias pertenecientes a *Solanum lycopersicum* y se reemplazaron las letras U por T, correspondientes a uracilo y timina, respectivamente, mediante reemplazo masivo de caracteres usando el comando “sed”. Para realizar el filtrado de las secuencias contaminantes se utilizaron los archivos fastq sometidos al pre-procesado detallado en el punto anterior y se utilizó el programa Bowtie v 1.1.1 (Langmead, 2010) para alinear dichas secuencias sobre el conjunto de secuencias contaminantes de tomate obtenidas de la base de datos Rfam. El conjunto de secuencias no deseadas en formato fasta, fue indexado usando la herramienta bowtie-build v 1.1.1 y para el alineamiento con bowtie se utilizó el modo “end to end” (-v) permitiendo 2 mismatches, y con la opción “--un” se guardaron las reads limpias que no alinearon a la referencia de secuencias contaminantes.

## 2.2.3 - Análisis de miARNs de tomate

### 2.2.3.1 - Descarga y actualización de la información correspondiente a miARNs de tomate

Se descargó la información correspondiente a miARNs de tomate de la base de datos de miRbase (versión 21; Kozomara & Griffiths-Jones, 2014). Con el fin de actualizar esta lista, se realizaron búsquedas bibliográficas en la literatura más reciente de esta temática y se incorporaron al análisis los miARNs ausentes en miRBase. La información se almacenó en formato fasta para poder indexarla posteriormente usando bowtie-build para emplear el programa Bowtie, y utilizarla como referencia para los alineamientos.

### 2.2.3.2 - Mapeo al genoma de referencia, anotación y cuantificación de miARNs

Luego del pre-procesado, se alinearon los archivos en formato fastq al genoma de referencia (SL 3.00, cv. Heinz 1706) usando Bowtie v 1.1.1 en el modo “end to end” permitiendo 2 mismatches (-v 2) y eliminando *reads* que alinean a la referencia más de 50 veces (-m 50). Posteriormente, con la herramienta fastx\_collapser v. 0.0.14 (Gordon, 2011), utilizando las opciones por *default*, se colapsaron y cuantificaron las *reads* repetidas de los archivos resultantes. Se alineó cada uno de estos ficheros colapsados a la lista de miARNs actualizada, empleando Bowtie v 1.1.1, en el modo “end to end”, sin permitir mismatches (-v 0) ni *reads* que alineen más de una vez (-m 1). Para conservar sólo aquellos miARNs que tuvieron alineamientos perfectos, se utilizó la herramienta

Samtools v 1.5 (Heng Li et al., 2009), mediante la cual se descartaron las *reads* no alineadas (valor de flag: 4), las que alinearon a la hebra reversa (valor de flag: 16) y aquellas cuya posición de inicio de alineamiento no haya sido 1 (columna 4). La información obtenida se organizó como tabla indicando los niveles de expresión normalizados en *reads* por millón (RPM) de cada miARN para cada réplica de cada estadio de madurez.

#### 2.2.4 - Análisis de expresión diferencial de miARNs

Para evaluar la expresión diferencial de cada miARN a lo largo de la maduración de los frutos se utilizó el paquete EdgeR 3.6 (Robinson et al., 2009) dentro del software Rstudio 1.1.456 (Allaire, 2012), utilizando la versión 3.6.0 de R. Se realizaron comparaciones entre los distintos estadios de madurez y se consideró como diferencia estadísticamente significativa a aquellas comparaciones con un valor de q-value  $\leq 0,05$  y una tasa de cambio mayor a 2.

#### 2.2.5 - Detección de loci en fase

Los archivos fastq pre-procesados correspondientes a réplicas biológicas de un mismo estadio fueron combinados en un solo archivo ya que una mayor cantidad de lecturas puede beneficiar la identificación de clusters de ARNs pequeños en fase. Luego se utilizó el programa Shortstack 3.8.3 (Axtell, 2013), permitiendo 2 mismatches, en el modo “--nohp” para desactivar la búsqueda de nuevos miARNs y desactivando también la opción que marca como “no alineadas” aquellas *reads* que presentan más de *n* alineamientos posibles y no pueden ser asignados a un lugar en particular (“--ranmax none”). Sólo se guardaron para posteriores análisis aquellos *loci* con al menos 5 *reads*, tamaño entre 100 y 1000 pares de bases y *p-score* mayor a 30. Se conservaron sólo los ARNs pequeños de 21 nucleótidos, y se los separó de acuerdo a la hebra a la que alinearon. Se eliminaron entradas duplicadas en base a la posición de inicio de mapeo, priorizando siempre los alineamientos que no tuvieron mismatches.

#### 2.2.6 - Mapeo de datos de RNA-seq al genoma de referencia

Los datos de secuenciación de *RNA-seq* obtenidos de bibliotecas correspondientes a frutos de *S. lycopersicum* cv. Ailsa Craig en estadios Verde Inmaduro, Rompiente y Rojo Maduro fueron descargados de la base de datos SRA (Tabla 2.1) y las

secuencias se filtraron por calidad como se indica en el punto 2.2. Para alinearlas al genoma de referencia se utilizó el programa HISAT2 2.1.0 (D. Kim et al., 2015) con la opción “--no-softclip” para evitar que se oculten las porciones de las *reads* que no alineen de extremo a extremo a la referencia.

### 2.2.7 - Visualización de los resultados de alineamientos al genoma de referencia

Para visualizar gráficamente los niveles de expresión y la ubicación en el genoma de los datos de *RNA-seq* y de *small RNA-seq* se empleó el programa *Integrative Genomics Viewer* (IGV; Thorvaldsdóttir et al., 2013). Para esto fue necesario ordenar los archivos en formato bam de acuerdo a su posición de mapeo usando la herramienta samtools sort v. 1.5 y luego indexarlos empleando samtools index v. 1.5. Los datos se visualizaron sobre la versión 3.00 del genoma de tomate utilizando el archivo de anotaciones ITAG 3.2 (ambos disponibles en la base de datos SolGenomics: <https://solgenomics.net/>), previamente cargado en IGV viewer.

### 2.2.8 - Análisis de expresión de genes blanco de miARNs

Para evaluar la expresión durante la maduración de los genes blanco de miARNs se utilizaron datos de *RNA-seq* de frutos de tomate ya publicados de la variedad Ailsa Craig correspondientes a los 4 estadios de maduración analizados en las bibliotecas de small RNA-seq (Ye et al., 2015). Se analizó la expresión diferencial de los genes de interés utilizando una dispersión artificial de 0.05, siguiendo las recomendaciones del manual de usuario de EdgeR (Y. Chen et al., 2020).

### 2.2.9 - Validación de genes blanco con bibliotecas de degradoma

Los genes blanco de los miARNs analizados y de los phasiARNs detectados fueron validados bioinformáticamente con bibliotecas de degradoma de fruto (estadio inmaduro y maduro, dos réplicas biológicas cada uno, Tabla 2.1) y el software CleaveLand v4.4 (Addo-Quaye et al., 2009), conservando sólo los resultados con valor  $p \leq 0,1$  y tipo de degradoma 0 o 1. Los datos de secuenciación correspondientes a bibliotecas de degradoma fueron pre-procesados como se explica en el punto 2.2 (Pre-procesamiento de datos de NGS). Luego se convirtió el formato fastq a fasta con la

herramienta FASTQ-to-FASTA v. 0.0.14 (Gordon, 2011) y las réplicas fueron concatenadas.

Se consideró como posibles blancos de miARNs y phasiARNs al conjunto de transcritos codificados en el genoma de tomate y para esto se descargó el archivo de anotaciones ITAG 3.2 de la base de datos SolGenomics (<https://solgenomics.net/>). Además fue necesario reemplazar espacios y otros caracteres en la línea correspondiente al identificador de secuencia en este archivo para lograr que el programa CleaveLand reporte el gráfico *T-plot* (ver más adelante), para lo cual se realizó una serie de reemplazos masivos por el carácter “\_” (guión bajo). Los genes blanco de phasiARNs fueron clasificados en base a su función molecular de acuerdo a la categorización por ontología génica (GO) provista por SolGenomics. Se utilizó el sitio web <http://pantherdb.org/> para analizar enriquecimiento de términos.

## 2.3 - Resultados

### 2.3.1 - Análisis de miARNs

Al comenzar el presente trabajo de tesis existían varias publicaciones relacionadas con la identificación de miARNs en frutos de tomate y algunas que realizaban análisis de expresión durante la maduración de tomate. Sin embargo, aquellas que identificaban miARNs en tomate utilizaban frutos correspondientes a un único estadio de madurez (Khalidun et al., 2016; Y. Wang et al., 2017) y los análisis de expresión a través de distintos estadios de maduración habían sido realizadas para algunos miARN seleccionados y usando expresión por hibridación tipo Northern blot (Moxon et al., 2008; X. Zhang et al., 2011). Por otro lado, no existía información acerca de phasiARNs presentes durante la maduración de frutos de tomate. Por lo tanto, se realizó una serie de análisis a nivel genómico destinados a estudiar la presencia y los cambios en la acumulación de ARNs pequeños en distintos estadios de la maduración de tomate usando datos de secuenciación masiva (Fig. 2.5). Por un lado se llevó a cabo un análisis orientado a identificar aquellos miARNs presentes durante la maduración y por otro se identificaron *loci* generadores de phasiARNs. En ambos casos se continuó con un análisis de identificación y validación de los transcritos regulados por estos ARNs pequeños (Fig. 2.5).

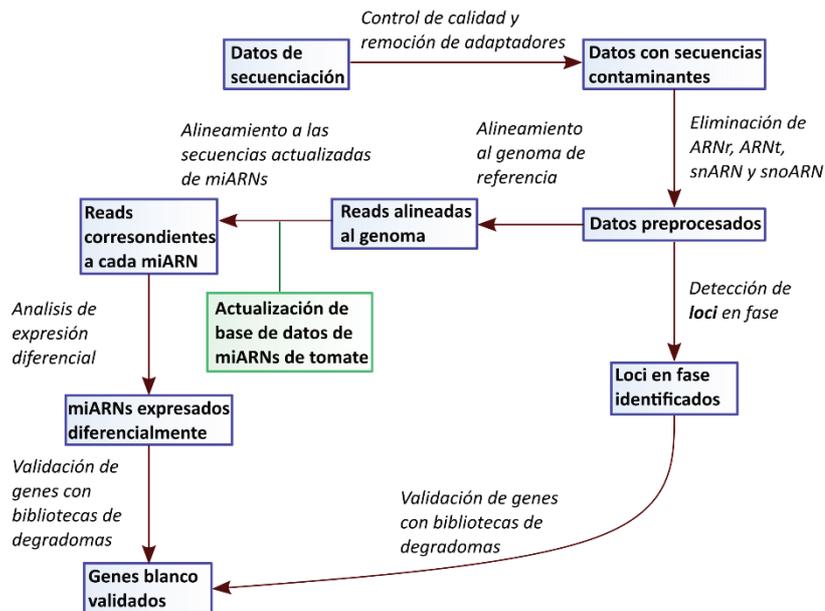


Fig 2.5. Procedimiento seguido para el procesamiento de los datos de secuenciación de ARNs pequeños de tomate.

### 2.3.1.1 - Pre-procesado de los datos de secuenciación de *small RNA-seq*

El análisis de datos de secuenciación de ARNs pequeños provenientes de bibliotecas de *small RNA-seq* de frutos de tomate (*Solanum lycopersicum* cv. Ailsa Craig) se realizó con datos correspondientes a cuatro estadios de maduración: verde inmaduro, rompiente, rosa y rojo maduro (Tabla 2.1). El control de calidad de los datos crudos (punto 2.2 de Materiales y métodos) realizado con el programa FastQC v 0.11.5 permitió determinar que la calidad de secuenciación por base se ubicaba en la región de máxima calidad para todas las posiciones secuenciadas y la calidad por secuencia de cada biblioteca era superior a 30, y además, el porcentaje de bases no asignadas (N) fue mínimo en todos los casos, por lo que se decidió continuar con el análisis sin realizar filtrados por calidad de los datos de secuenciación (Fig. 2.5; Tabla 2.2). Se identificó a la secuencia “TGAATTC” como el inicio de la porción correspondiente a los adaptadores usados en la preparación de las bibliotecas de *small RNA-seq* y se eliminaron las porciones de adaptador presentes en las secuencias (punto 2.2 de Materiales y métodos). Luego se realizó el filtrado de las secuencias correspondientes a ARNr, ARNt, snARN y snoARN (punto 2.2 de Materiales y métodos; Fig. 2.5). En la Tabla 2.2 se puede observar que luego de eliminar adaptadores se retuvo más del 99% de las secuencias originales en todas las bibliotecas analizadas y luego de filtrar las secuencias contaminantes, el porcentaje de *reads* limpias en todas las bibliotecas se mantuvo entre

82% y 91% de las secuencias iniciales (Tabla 2.2). Estos valores obtenidos luego de seguir procedimientos estándar, reflejan que la calidad de secuenciación de las bibliotecas y la cantidad de *reads* obtenidas para los análisis posteriores son adecuadas para el objetivo que se pretendía llevar a cabo (Fig. 2.5).

Tabla 2.2: Pre-procesado y mapeo de bibliotecas de *small RNA-seq* de tomate

	<i>Reads</i> crudas	<i>Reads</i> tras eliminar adaptadores (%)*	<i>Reads</i> limpias (%)*	<i>Reads</i> alineadas al genoma (%)**
<b>Verde Inmaduro - R1</b>	22.303.888	22.242.801 (99,7)	20.191.831 (90,5)	15.540.616 (77,0)
<b>Verde Inmaduro - R2</b>	21.404.460	21.354.949 (99,8)	19.477.156 (91,0)	14.949.863 (76,8)
<b>Rompiente - R1</b>	20.152.218	20.106.725 (99,8)	18.159.840 (90,1)	14.252.176 (78,5)
<b>Rompiente - R2</b>	22.577.145	22.529.993 (99,8)	20.395.012 (90,3)	15.997.189 (78,4)
<b>Rosa - R1</b>	18.758.904	18.717.205 (99,8)	16.871.972 (89,9)	12.945.904 (76,7)
<b>Rosa - R2</b>	21.395.963	21.302.979 (99,6)	19.303.074 (90,2)	14.759.033 (76,5)
<b>Rojo Maduro - R1</b>	19.762.026	19.653.759 (99,5)	16.213.900 (82,0)	12.382.941 (76,4)
<b>Rojo Maduro - R2</b>	15.967.713	15.894.361 (99,5)	13.098.077 (82,0)	10.000.020 (76,3)

\* Porcentaje respecto del número de *reads* crudas

\*\* Porcentaje respecto del número de *reads* limpias

### 2.3.1.2 - Actualización de la información de miARNs de tomate

Una inspección detallada de la información descargada desde miRbase (versión 21) nos permitió determinar que los 113 miARN contenidos en esta base de datos, no reflejaban el total de los miARN de tomate descritos hasta el momento. Además, esta base de datos no contaba con un archivo de anotaciones (.gff3 o similar) para la versión 3.0 del genoma de referencia de tomate. Por lo tanto, se realizó una exhaustiva búsqueda bibliográfica que permitió completar la lista con un total de 158 miARNs, los cuales pertenecen a 65 familias distintas y se presentan en el Anexo II. Como se detalla en este anexo, las familias de miARNs ausentes en la lista original corresponden a miR827, miR1863, miR1919, miR0211, miR399, miR4414, miR477, miR530, miR157, miR170, miR2111, miR2118, miR398, miR408 y miR894, los cuales se agregaron según la información publicada en la bibliografía consultada (Karlova et al., 2013; Y. Luan et al.,

2014; Moxon et al., 2008; Sato et al., 2012; Suresh et al., 2014; F. Wu et al., 2017; J. Zuo et al., 2012). Luego, estos datos fueron utilizados como se describe más adelante para realizar la anotación y cuantificación de miARNs durante la maduración de tomate.

### 2.3.1.3 - Expresión de miARNs durante la maduración de frutos de tomate

Para analizar los niveles de expresión de los miARNs durante la maduración de frutos de tomate, se alinearon las *reads* pre-procesadas y libres de secuencias contaminantes, usando como referencia la versión 3.0 del genoma de tomate (Fig. 2.5). El porcentaje de *reads* alineadas para los datos provenientes de cada una de las bibliotecas fue variable entre 76,3% y 77% (Tabla 2.2), lo cual corresponde a porcentajes típicos de mapeo para este tipo de análisis (Hongjun Liu et al., 2014; Y. Luan et al., 2015). Como se mencionó anteriormente, la base de datos miRbase no contaba con un archivo en formato .gff o similar para anotar y cuantificar los miARNs contenidos en el total de *reads* mapeadas, por lo tanto se utilizó la lista de miARNs de tomate actualizada (Anexo II) para anotar los miARNs conocidos y cuantificar el número de *reads* correspondientes a cada uno de ellos, dentro de las *reads* mapeadas al genoma de referencia (Fig. 2.5). Este análisis permitió establecer que hay miembros pertenecientes a 48 de las 65 familias de miARNs analizadas que están presentes durante la maduración de tomate (Fig. 2.6A), mientras que no se detectó la presencia de ningún miembro de las 17 familias restantes (Fig. 2.6A). Luego se realizó el análisis de identificación y validación de genes target (Fig. 2.6B-C, ver más adelante).

Además se realizó un análisis de expresión diferencial entre los distintos estadios de madurez analizados (punto 2.4 de Materiales y métodos), a través del cual se determinó que algunos miembros de 7 familias de miARNs expresados en frutos mostraron expresión diferencial durante la maduración. Estas familias de miARNs corresponden a las familias sly-miR319, sly-miR396, sly-miR156, sly-miR159, sly-miR166, sly-miR171 y sly-miR482 (Fig. 2.7: Tabla 2.3) las cuales han sido descritas y caracterizadas en diversas especies, indicando que son familias de miARNs conservados en la evolución y que la regulación génica mediada por ellos ha sido seleccionada para participar en diversos aspectos del desarrollo y el crecimiento de distintas especies (Sunkar & Jagadeeswaran, 2008; Willmann & Poethig, 2007).

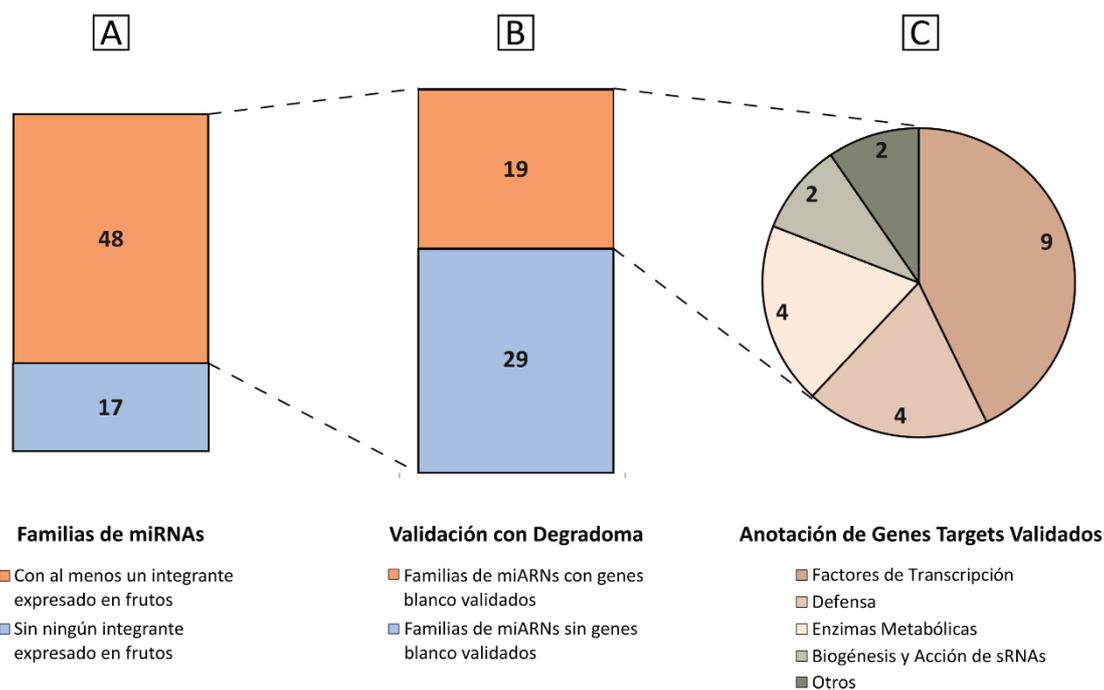


Fig. 2.6. Análisis de miARNs. (A) Total de familias con alguno de sus miembros expresado en al menos uno de los estadios de maduración analizado. (B) Total de familias con genes blanco validados. (C) Distribución de los procesos biológicos en los que participan los genes blanco validados.

Los perfiles de expresión de estos miARNs durante la maduración son diversos y diferentes para cada uno de ellos (Fig. 2.7). Sin embargo, es posible distinguir tres patrones de expresión generales entre el inicio (estadio verde inmaduro) y el final de la maduración (estadio rojo maduro). Los miARNs sly-miR319c, sly-miR396a y sly-miR396b llegan al final de la maduración con una acumulación menor a la inicial en el estadio verde inmaduro, mientras que sly-miR156f tiene el perfil opuesto, ya que se acumula a un mayor nivel en el estadio rojo maduro en comparación con el estadio verde inmaduro inicial (Fig. 2.7). Por su parte, sly-miR159, sly-miR166c, sly-miR171e y sly-miR482c presentan niveles comparables al inicio y al final de la maduración, pero evidencian niveles máximos de expresión en alguno de los estadios de maduración intermedios (Fig. 2.7).

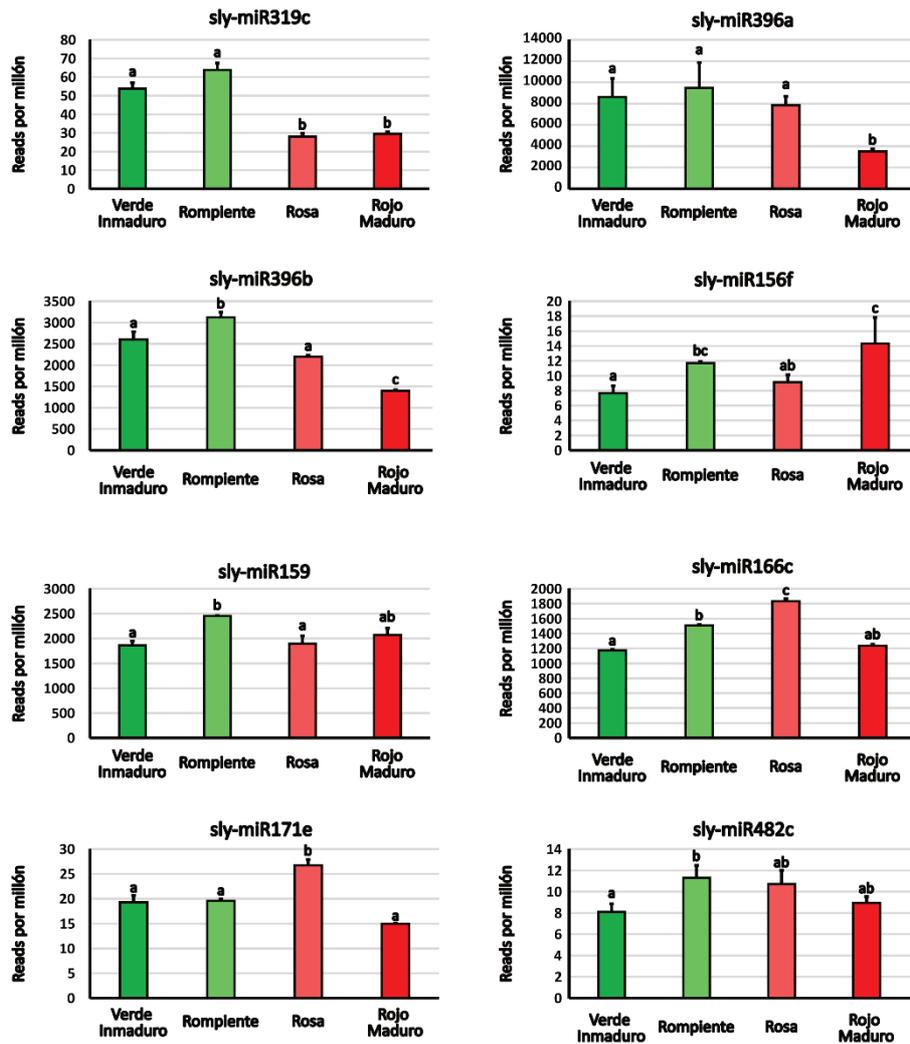


Fig. 2.7. Niveles de expresión de miembros de familias de miARNs conservadas entre especies con expresión diferencial entre las bibliotecas analizadas. Los valores representados corresponden al promedio de las *reads* normalizadas por millón de *reads* y la desviación estándar para dos réplicas biológicas en cada estado. Las letras sobre las barras indican diferencia significativa ( $p$ -value ajustado < 0,05 y tasa de cambio > 2).

#### 2.3.1.4 - Identificación y validación de genes regulados por miARNs durante la maduración

Para estudiar los procesos biológicos regulados por miARNs durante la maduración, se hizo un análisis de identificación y validación de los genes *target* regulados por aquellos miARNs presentes en frutos (Fig. 2.6) usando el programa Cleaveland v4.4 (Addo-Quaye et al., 2009). Este programa permite realizar predicciones de posibles genes regulados por un conjunto de ARNs pequeños sobre el total de los transcritos codificados en un genoma y una posterior validación de los mismos usando

datos de secuenciación provenientes de bibliotecas de degradoma. Se utilizaron bibliotecas de degradoma correspondientes a los estadios verde inmaduro y rojo maduro (Tabla 2.1) de frutos de tomate de la variedad Ailsa Craig. Se seleccionaron para análisis posteriores aquellos resultados cuyo *p-value* fue menor a 0,1 con gráficos *T-plot* de categoría 0 y 1, y además se realizó una inspección visual de todos los archivos *T-plots* generados por el programa. Esto permitió conocer los procesos biológicos regulados por los miARNs presentes en frutos de tomate, validándose los genes blanco de 19 familias de miARNs (Fig. 2.6B; Tabla 2.3), mientras que no se detectaron genes blanco para 29 familias de miARNs presentes en frutos (Fig. 2.6B).

De acuerdo a su anotación, se clasificaron los genes blanco Según el proceso biológico del que participan (Fig. 2.6C; Tabla 2.3). La mayoría de los genes blanco regulados por miARNs en frutos corresponden a factores de transcripción y también se encontraron genes que codifican para proteínas relacionadas con la defensa de la planta, para enzimas metabólicas y para enzimas implicadas en la biogénesis y acción de ARNs pequeños (Fig. 2.6C; Tabla 2.3). De acuerdo a los genes regulados por estos miARNs y de acuerdo a la bibliografía existente, tres de estas familias de miARNs están relacionadas con distintos aspectos del desarrollo y la maduración de los frutos: sly-miR156, sly-miR159, sly-miR396 (Tabla 2.3), por lo que resultan interesantes por su participación en la regulación de algunos aspectos clave de estos procesos.

Tabla 2.3: Validación de genes blanco de miARNs expresados en frutos

	miARNs presentes en frutos	Expresión diferencial	Gen blanco
Factores de transcripción	sly-miR156	Sí	Squamosa promoter binding protein
	sly-miR159	Sí	GAMYB-like protein
	sly-miR166	Sí	Homeobox leucine-zipper protein
	sly-miR171	Sí	Transcription factor GRAS
	sly-miR319	Sí	TCP transcription factor
	sly-miR396	Sí	Growth-regulating factor
	sly-miR172	No	APETALA 2b
	sly-miR169	No	Nuclear transcription factor Y
Defensa	sly-miR482	Sí	Disease resistance protein (CC-NBS-LRR class)
	sly-miR6024	No	Disease resistance protein
	sly-miR6022	Sí	Leucine-rich repeat receptor-like protein kinase family protein
	sly-miR6023	Sí	Leucine-rich repeat receptor-like protein kinase family protein
Enzimas metabólicas	sly-miR160	No	Auxin Response Factor
	sly-miR1919	No	Purple acid phosphatase
	sly-miR394	No	F-box family protein
	sly-miR396	Sí	Cytosine-5 DNA methyltransferase
Biogénesis y acción de ARNs pequeños	sly-miR403	No	Argonaute 2a
	sly-miR6026	Sí	Dicer-like 2d
Otros	sly-miR9476	Sí	Rer1 family protein
	sly-miR393	No	Transport inhibitor response 1

El análisis de expresión durante la maduración de los genes blanco se realizó utilizando datos de *RNA-seq* de frutos de tomate ya publicados (punto 2.8 de Materiales y métodos). En la Figura 2.8 se grafican los niveles de expresión durante la maduración de los 3 genes blanco analizados: Solyc01g009070 (*GAMYB-like 1*, blanco de sly-miR159, Figura 2.8A), Solyc08g075950 (*Growth-regulating factor 3*, blanco de sly-miR396a,

Figura 2.8B) y Solyc05g012040 (*Squamosa promoter binding protein 6b*, blanco de sly-miR156. Figura 2.8C). El nivel de expresión de *GAMYB-like 1* es mayor al inicio de la maduración y se mantiene en niveles bajos en los demás estadios. Se vio que la expresión de *GRF3* fue baja y similar en los 4 estadios estudiados, mientras que *SPB6b* se expresó de manera relativamente constante durante la maduración, con un leve aumento en el estadio maduro (49 días post antesis, Fig. 2.8C).

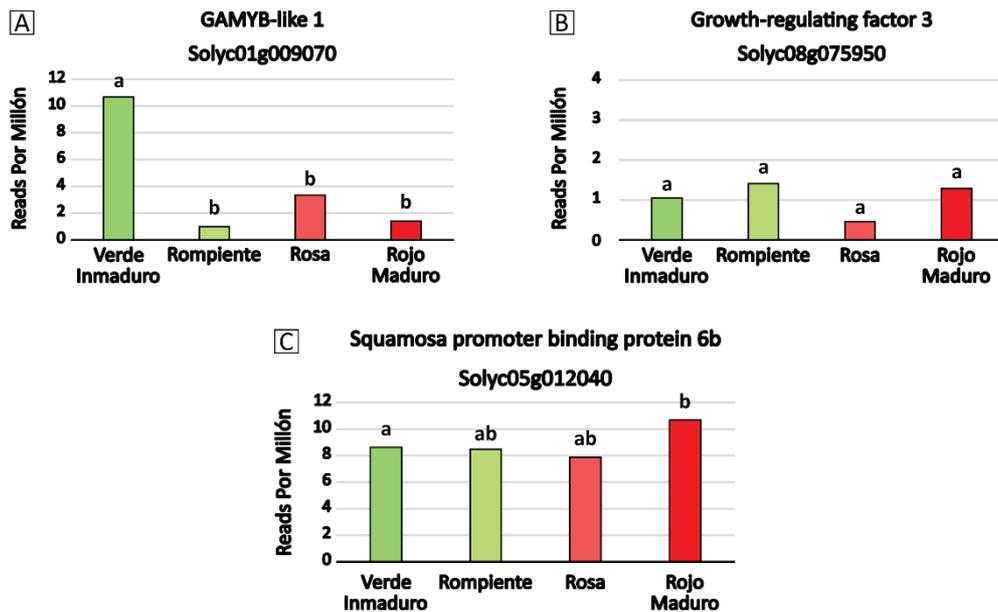


Fig. 2.8. Niveles de expresión de transcritos regulados por miARNs seleccionados. (A) Expresión de *GAMYB-like 1*, regulado por sly-miR159. (B) Expresión de *GRF3*, regulado por sly-miR196. (C) Expresión de *SPL6b*, regulado por sly-miR156. Los valores representados corresponden a las *reads* normalizadas por millón de reads. Las letras sobre las barras indican diferencia significativa ( $p$ -value ajustado < 0,05).

## 2.3.2 - Caracterización de loci generadores de ARNs pequeños en fase

### 2.3.2.1 - Predicción de loci PHAS y TAS

Como parte de la caracterización de los ARNs pequeños presentes durante la maduración de frutos de tomate, realizamos un análisis para identificar *loci* generadores de ARNs pequeños en fase. Utilizando la herramienta Shortstack 3.8.3 se detectaron 19 *loci* presentes en frutos de tomate que no habían sido descritos hasta el momento (Tabla 2.4). Dentro de los criterios elegidos para seleccionar aquellos *loci* fidedignos incluidos en la extensa lista resultante del análisis, establecimos que cada *locus* tenga un *Phasing-score* (P-score) menor o igual a 30 y una longitud entre 100 y 1000 pb, de

acuerdo a procedimientos similares reportados previamente (Das et al., 2020; Feng et al., 2019; Nakamura et al., 2019).

Como se mencionó en la introducción de este capítulo, los phasiARNs están típicamente asociados a regiones genómicas que coinciden con genes que codifican para proteínas, mientras que los tasiARNs provienen de transcritos no codificantes (Qili Fei et al., 2015). Teniendo en cuenta esta distinción, nuestro análisis detectó la presencia de 15 *loci PHAS*, que coinciden en su ubicación genómica con genes codificantes conocidos y anotados en la versión ITAG 3.2 (correspondiente a la versión 3.0 del genoma), lo que permitiría clasificar a estos siARNs como phasiARNs. Sorprendentemente, 5 de estos *loci PHAS* se superponen con el gen que codifica para sly-DCL2b (*loci 1, 9, 11, 18 y 19*; Tabla 2.4) y el *locus 7* coincide con la ubicación genómica del gen sly-DCL2d (Tabla 2.4). Por otro lado, cuatro de los *loci* detectados no coinciden con ningún gen anotado hasta el momento, por lo que podrían clasificarse como *loci TAS* (*loci 3, 4, 6 y 8*, destacados en rojo en la Tabla 2.4), mientras que los 9 *loci PHAS* restantes se solapan con genes que codifican para proteínas diversas, tales como factores de transcripción y proteínas de defensa (Tabla 2.4).

**Tabla 2.4: Descripción de los *loci* en fase detectados en frutos de tomate**

<b>Locus</b>	<b>Ubicación genómica</b>	<b>Modelo génico</b>	<b>Gen asociado</b>
<b>Locus 1</b>	SL3.0ch11:2716549..2718986	Solyc11g008540.2.1	DCL2b
<b>Locus 2</b>	SL3.0ch05:6248547..6249413	Solyc05g013160.3.1	Ribulose-1 methyltransferase
<b>Locus 3</b>	SL3.0ch01:4336179..4337135	-	-
<b>Locus 4</b>	SL3.0ch08:63488258..63488701	-	-
<b>Locus 5</b>	SL3.0ch00:14433784..14434274	Solyc00g090430.2.1	LOW QUALITY: DUF789 family protein
<b>Locus 6</b>	SL3.0ch08:59849551..59849996	-	-
<b>Locus 7</b>	SL3.0ch11:2700083..2702516	Solyc11g008520.2.1	DCL2d
<b>Locus 8</b>	SL3.0ch06:689091..689662	-	-
<b>Locus 9</b>	SL3.0ch11:2716829..2717511	Solyc11g008540.2.1	DCL2b
<b>Locus 10</b>	SL3.0ch08:14341621..14342313	Solyc08g048290.3.1	Inositol transporter
<b>Locus 11</b>	SL3.0ch11:2715289..2715821	Solyc11g008540.2.1	DCL2b
<b>Locus 12</b>	SL3.0ch09:70938942..70939873	Solyc09g091100.3.1	Coenzyme F420 hydrogenase/dehydrogenase beta subunit-like
<b>Locus 13</b>	SL3.0ch09:70627278..70627708	Solyc09g090710.3.1	pfkB-like carbohydrate kinase family protein
<b>Locus 14</b>	SL3.0ch00:18367288..18367989	Solyc00g218560.1.1	LOW QUALITY: T-box transcription factor
<b>Locus 15</b>	SL3.0ch06:422035..424175	Solyc06g005410.3.1	Disease resistance protein (TIR-NBS-LRR class)
<b>Locus 16</b>	SL3.0ch11:50593599..50593871	Solyc11g064980.1.1	LOW QUALITY: Carboxy-terminal domain RNA polymerase II polypeptide A - small phosphatase 1
<b>Locus 17</b>	SL3.0ch10:59280355..59281331	Solyc10g076270.2.1	Regulatory-associated protein of TOR 1
<b>Locus 18</b>	SL3.0ch11:2719457..2719960	Solyc11g008540.2.1	DCL2b
<b>Locus 19</b>	SL3.0ch11:2718743..2719201	Solyc11g008540.2.1	DCL2b

### 2.3.2.2 - Detección y validación de genes blanco

Posteriormente se identificaron los siARNs producidos a partir de estos 19 *loci*, se extrajo su secuencia a partir de los datos resultantes del alineamiento realizado por Shortstack y se les asignó un nombre de acuerdo al *locus* de origen y a la hebra a partir

de la cual cada siARN se produce: S para hebra “sentido” y AS para hebra “anti sentido”. De esta forma, los nombres tienen el formato “Locus\_hebra\_orden”, por ejemplo: L1\_S\_1.

Para estudiar la importancia biológica de los siARNs identificados, se realizó un análisis de búsqueda y validación de genes blanco, análogo al realizado y descrito previamente para miARNs. La selección de aquellos genes blanco cuyos gráficos de *T-plot* pertenecen a las categorías 0 y 1 asignadas por Cleaveland y tienen un *p-value*  $\leq 0,1$  resultó en un total de 128 genes validados para el estadio Verde Inmaduro, 118 para el estadio Rojo Maduro y 42 validados en ambos estadios. Posteriormente se realizó una selección manual inspeccionando individualmente los *T-plots* correspondientes a cada gen blanco y se retuvieron aquellos gráficos que presentaron un solo pico o un pico más alto que los demás y separados por más de 500 pb de picos contiguos. Como resultado se obtuvieron 82 genes validados como targets de los phasiARNs de frutos de tomate, que fueron clasificados de acuerdo a su función molecular en base a la categorización por ontología génica (GO, Fig. 2.9). Se determinó que los phasiARNs provenientes de los 19 loci en fase identificados en este trabajo regulan una amplia variedad de procesos biológicos y no se encontró enriquecimiento de ninguna categoría GO cuando se realizó este análisis.

### Distribución de genes blanco por función

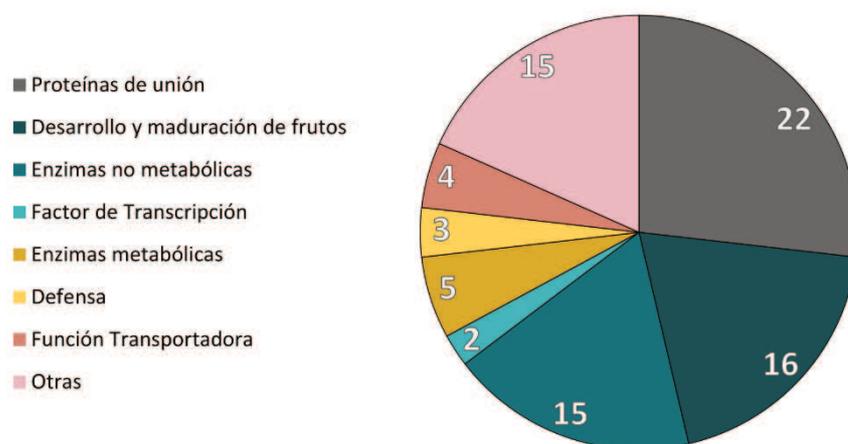


Fig. 2.9: Distribución de los procesos biológicos en los que participan los genes blanco validados para los phasiARNs detectados.

Un análisis detallado de las funciones moleculares de cada uno de los genes validados como targets de phasiARNs, permitió identificar que 16 de ellos se destacan por tener funciones relacionadas al desarrollo y maduración de frutos (Tabla 2.5). La mayoría de estos phasiARNs pertenecen al *locus 7*, y uno de ellos (L7\_S\_30) regula dos genes que codifican para enzimas del tipo lipasas GDSL. Cuatro phasiARNs correspondientes al *locus 1* regulan genes directamente relacionados con la adquisición de propiedades organolépticas de frutos. De los phasiARNs con genes blanco validados restantes, dos se producen a partir del *locus 10* y uno a partir de los *loci 6, 9 y 15*. Estos *loci* PHAS se describen a continuación con más detalle.

**Tabla 2.5: phasiARNs implicados en desarrollo, maduración y propiedades organolépticas de frutos**

phasiARN	Gene ID	Anotación
L1_AS_34	Solyc10g085830.2.1	Caffeic acid O-methyltransferase
L1_S_144	Solyc06g060180.2.1	Terpene synthase
L1_S_81	Solyc06g071690.3.1	R2R3MYB transcription factor 50
L1_S_99	Solyc12g009420.2.1	Poligalacturonase
L6_S_2	Solyc11g013480.2.1	Auxin Response Factor 17
L7_AS_128	Solyc07g021040.1.1	Alcohol dehydrogenase
L7_AS_3	Solyc04g050930.3.1	Violaxanthin de-epoxidase
L7_S_146	Solyc02g088510.3.1	Saccharopine dehydrogenase
L7_S_149	Solyc03g093610.1.1	Ethylene response factor
L7_S_180	Solyc07g066250.1.1	GRAS family transcription factor
L7_S_30	Solyc02g071640.2.1	Lipase, GDSL
L7_S_30	Solyc02g071645.1.1	Lipase, GDSL
L9_AS_21	Solyc01g087035.1.1	Asparagine synthetase 2
L10_S_48	Solyc01g008230.3.1	TCP transcription factor 15
L10_S_8	Solyc00g084760.1.1	Myb SANT-like domain-containing protein
L15_S_20	Solyc08g080400.1.1	GRAS family transcription factor

### 2.3.3 - sly-miR6026 produce phasiARNs a partir de transcritos de la familia sly-DCL2 de tomate

Ante la identificación de varios phasiARNs provenientes de los *loci 1 y 7* que resultaron interesantes por su posible papel en la regulación del desarrollo de propiedades organolépticas de frutos de tomate (Tabla 2.5), se procedió a la caracterización detallada de la estructura y función de estos *loci*. De esta manera pudimos establecer que 3 miembros de la familia DCL2 de tomate (DCL2a, 2b y 2d), implicada en procesos de síntesis de ARN pequeños y defensa antiviral, presentan sitios de unión para sly-miR6026 y que es capaz de dar lugar a siARNs en fase de 21 nucleótidos.

El análisis de expresión de este miARN indica que está presente en todos los estadios de maduración de frutos analizados, detectándose un nivel máximo de acumulación en el estadio rompiente y una disminución a medida que avanza la maduración hasta alcanzar un nivel de expresión similar al detectado al inicio de la maduración (Fig. 2.10A). Luego se realizó un análisis usando datos de degradoma y se demostró que sly-miR6026 es capaz de mediar el corte del transcripto de sly-DCL2d (Fig. 2.10B). De manera similar, se determinó que sly-miR6026 es capaz de mediar el corte de los transcritos de sly-DCL2a y sly-DCL2b.

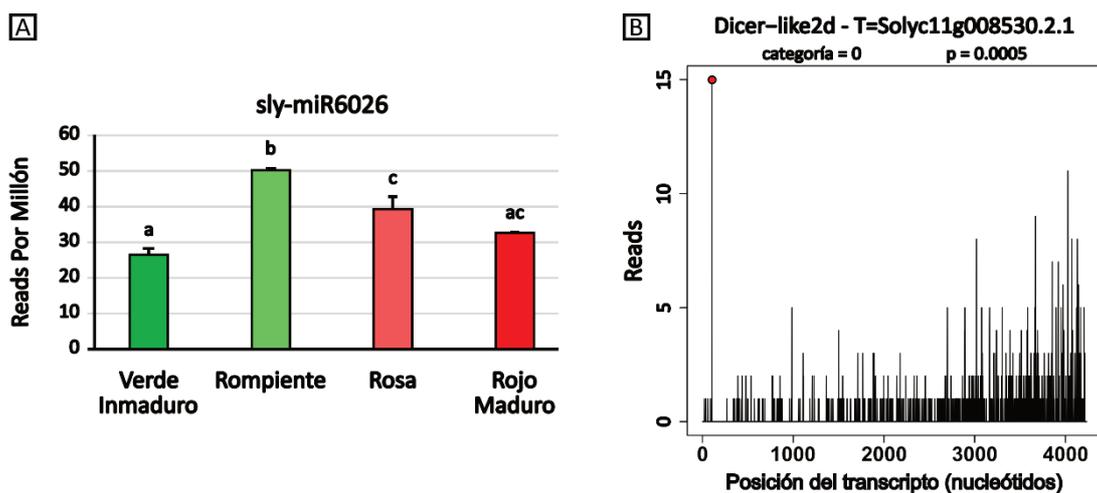


Fig. 2.10. (A) Perfil de expresión de sly-miR6026 en frutos de distintos estadios de madurez. (B) Validación del corte del transcripto de sly-DCL2d mediado por sly-miR6026. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland.

### 2.3.3.1 - Locus 7

El análisis de validación de genes blanco con el software Cleaveland y bibliotecas de degradoma permitió la detección de seis genes blanco para los phasiARNs del *locus* 7. El sitio de corte validado por Cleaveland indica que la unión de sly-miR6026 al transcripto de sly-DCL2d desencadena la producción de los phasiARNs detectados con Shortstack (Fig. 2.11). En la Figura 2.11 se representan con colores en tonos verdes aquellos phasiARNs presentes en los datos de secuenciación analizados y se indican los nombres asignados. Se puede visualizar el corrimiento de 2 nucleótidos entre los phasiARNs producidos a partir de ambas hebras, como resultado del mecanismo de acción de las enzimas DCL (Vazquez, 2006).



Fig. 2.11: Representación esquemática del corte de sly-miR6026 en el transcripto correspondiente a sly-DCL2d y los phasiARNs generados a partir de las dos hebras.

Luego se analizó el nivel de acumulación de los distintos phasiARNs producidos a partir del *Locus 7* en los distintos estadios de madurez analizados (Tabla 2.5). Se puede observar el detalle del nivel de acumulación de los phasiARNs provenientes de *locus 7*, siendo más abundantes en general aquellos producidos a partir de la hebra “sentido” (S), dentro de los cuales se destaca L7\_S\_17 como el más abundante, mientras que otros como L7\_S10; L7\_S\_12; L7\_S\_16; L7\_S\_27 y L7\_S\_30 presentan un nivel de acumulación intermedio (Tabla 2.6).

**Tabla 2.6: Información de los phasiARNs del *locus 7* durante la maduración**

Posición	Nombre	Hebra antisentido (Reversa complementaria)	Número de Reads normalizadas			
			V. inmaduro	Rompiente	Rosa	Rojo
2700337	L7_AS_11	CAAGGGCTTCCAAGTATAAC	0,81	0	0	0
2700463	L7_AS_17	AAACAGCTATGTAAGGCGACG	0,40	0	0	0
2700484	L7_AS_18	GAACAACAGTCGGAACCAAGA	1,44	2,07	0	2,21
2700778	L7_AS_32	CCTCATGATCGTCCACTAGTT	12,28	3,31	13,77	7,83

Posición	Nombre	Hebra sentido	Número de Reads normalizadas			
			V. inmaduro	Rompiente	Rosa	Rojo
2700335	L7_S_10	AAGTTATCAGTTGGAAGCCCT	867,42	1325,91	1045,92	747,23
2700356	L7_S_11	TGAGATTGCATTAAGCAGAA	1,06	1,36	0,72	4,23
2700377	L7_S_12	CACAATTGTTACTTGAGAC	464,81	722,56	687,44	556,48
2700398	L7_S_13	TGGTCTGGAAAGACCTTGAT	4,01	4,73	4,36	2,21
2700440	L7_S_15	CTATGCTTCCCTTCTCGTAA	0,36	0	0	0
2700461	L7_S_16	ACCGTCGCCTTACATAGCTGT	308,68	357,84	262,36	201,87
2700482	L7_S_17	TTTCTTGGTCCGACTGTTGT	4454,88	6686,14	5514,08	3903,40
2700692	L7_S_27	TGCTCTGATGTTGCATACCGA	112,50	89,30	88,23	73,05
2700755	L7_S_30	CTGGAATGCTGCTACATGGCA	104,60	138,23	129,42	106,47
2700776	L7_S_31	ACAACACTAGTGACGATCATGA	6,13	2,77	3,65	2,21

Como se mencionó antes, es particularmente interesante el phasiARN L7\_S\_30 producido por este *locus* (Tabla 2.5). El análisis de la acumulación de L7\_S\_30 durante la maduración indica que los niveles de este phasiARN se mantienen constantes durante la maduración (Fig. 2.12A). Este phasiARN fue validado como regulador del transcripto del gen Solyc02g071640.2.1, que codifica para una lipasa del tipo GDSL (Tabla 2.5). En la Fig. 2.12B se resumen los resultados de Cleaveland para este gen como target de L7\_S\_30 y se puede observar el sitio de corte mediado por este phasiARN en la región 3' del transcripto, validado por los datos de secuenciación de bibliotecas de degradoma. La expresión del gen GDSL en los estadios analizados fue baja y relativamente constante, excepto para el estadio rosa, donde no se registró expresión (Fig. 2.12C). Las lipasas GDSL participan en la formación de la cutícula en los frutos (Yeats et al., 2012), por lo que la regulación de la expresión de esta enzima por parte de L7\_S\_30 podría estar

relacionada con características como el brillo y la vida poscosecha de los frutos de tomate.

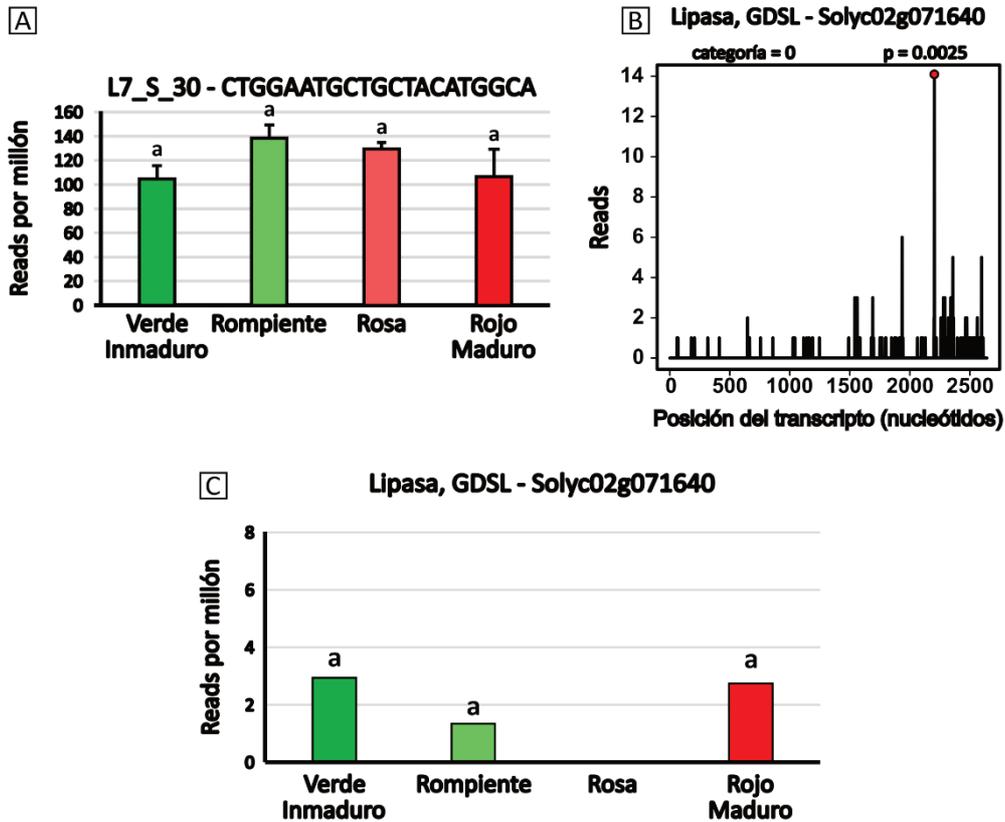


Fig. 2.12: (A) Perfil de expresión del phasiARN L7\_S\_30 en frutos de distintos estadios de madurez. (B) Validación del corte del transcrito de Solyc02g071640 mediado por L7\_S\_30. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcrito en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland. (C) Perfil de expresión del transcrito de Solyc02g071640 en frutos de distintos estadios de madurez.

En este *locus* también se detectaron otros phasiARNs que no están en fase con el corte predicho para miR6026. Se validaron 5 genes blanco además del mencionado previamente para L7\_S\_30 (Fig. 2.13). Los phasiARN y sus targets validados corresponden a L7\_AS\_3 que regula la expresión de un transcrito que codifica para una enzima Violaxantina de-epoxidasa (Solyc04g050930.3); L7\_AS\_128 regula al transcrito correspondiente a la enzima Alcohol deshidrogenasa (Solyc07g021040.1); L7\_S\_146 regula la expresión del transcrito de la enzima Sacaropina deshidrogenasa (Solyc02g088510.3); L7\_S\_149 regula al transcrito que codifica para un Factor de

Respuesta a Etileno (Solyc03g093610.1) y L7\_S\_180 al transcripto de un factor de transcripción de la familia GRAS (Solyc07g066250.1; Fig. 2.13).

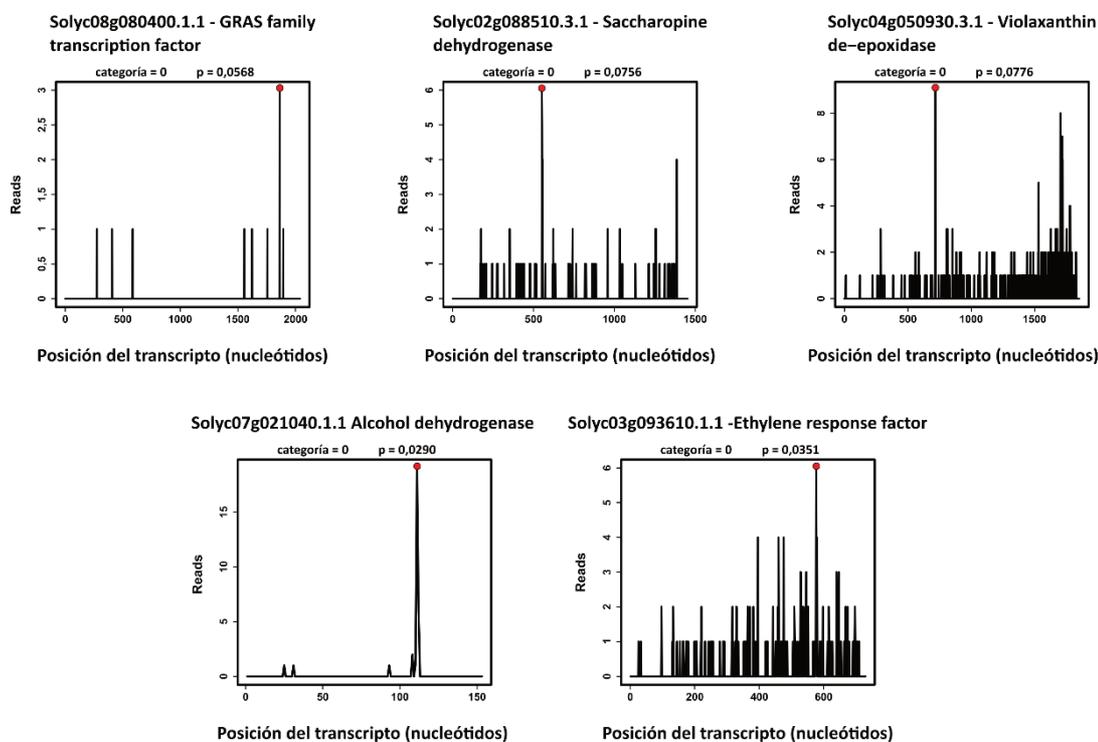


Fig. 2.13: Validación de los cortes para los transcritos de Solyc08g080400.1.1, Solyc02g088510.3.1, Solyc04g050930.3.1, Solyc07g021040.1.1 y Solyc03g093610.1.1. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte de cada transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland.

La violaxantina es un carotenoide implicado en la adquisición de la coloración de los frutos durante la maduración, y la enzima violaxantina de-epoxidasa cataliza la interconversión entre violaxantina, anteraxantina y zeaxantina (Hieber et al., 2000). La enzima alcohol deshidrogenasa participa en la producción de alcoholes y aldehídos, compuestos ligados a los aromas de los frutos (R. K. Singh et al., 2018). La sacaropina deshidrogenasa participa en el metabolismo de la lisina, hidrolizando sacaropina y produciendo glutamato (Arruda & Barreto, 2020), aminoácido relacionado con la adquisición de sabor en los frutos de tomate (Sorrequeta et al., 2010). Los factores de respuesta a etileno controlan la transcripción de genes regulados por esta hormona, y por lo tanto participan activamente en el proceso madurativo de los frutos (M. Liu et al., 2016). Por último, la familia de factores de transcripción GRAS participa en numerosos

procesos fisiológicos, como transducción de señales, señalización de la luz y respuesta a estrés biótico y abiótico, entre otros, (Niu et al., 2017) y además se demostró que miembros de esta familia en tomate intervienen en el control del tamaño de los frutos (M. Li et al., 2018).

### 2.3.3.2 - Locus 1

El análisis de validación de genes blanco permitió la detección de 12 genes blanco para los phasiARNs del *locus 1*. La ubicación de este *locus* coincide con la del intrón 5 del gen Solyc11g008540.2.1, correspondiente a sly-DCL2b (SL3.0ch11:2717620-2717767). Con el fin de visualizar la región que da origen a estos phasiARNs en el contexto del genoma, la anotación del gen de sly-DCL2b y los datos de *RNA-seq*, se utilizó el software IGV Viewer. Se detectó la presencia de *reads* correspondientes a transcritos a pesar de tratarse de una región intrónica, típicamente no detectadas en los datos de secuenciación de *RNA-seq*. En el estadio verde inmaduro se observa un alto número de *reads* que avalan la existencia de ARNs pequeños generados tanto en la hebra sentido como en la antisentido (Fig. 2.14).

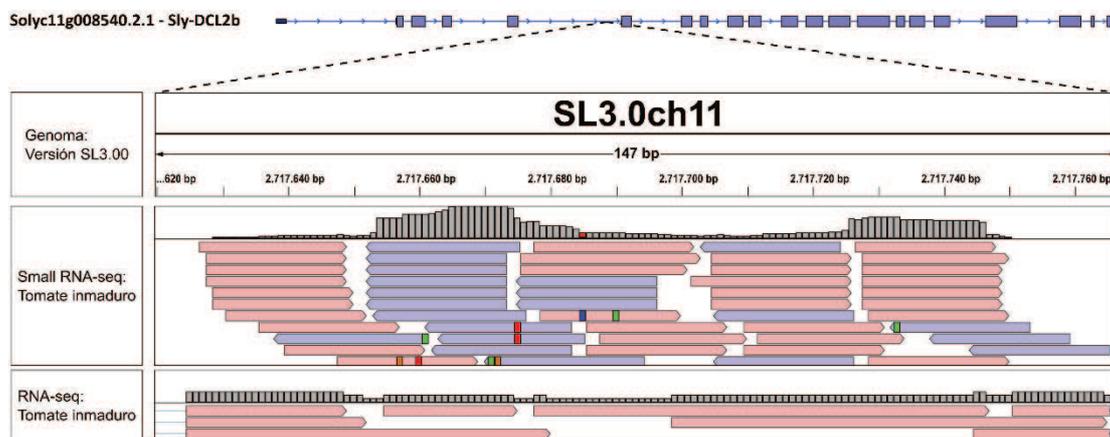


Fig. 2.14: Representación gráfica del *locus 1* con el detalle de las *reads* mapeadas correspondientes a bibliotecas de *small RNA-seq* y *RNA-seq* del estadio inmaduro. Las flechas rosadas indican *reads* alineadas en la hebra “sentido” y las flechas celestes indican *reads* alineadas a la hebra “antisentido”. Los pequeños rectángulos de colores en su interior indican si un nucleótido difiere en las *reads* alineadas más de un 20% de las veces de la secuencia de referencia. El color verde indica adenina, el azul citosina, el naranja guanina y el rojo timina.

Dos phasiARNs se destacan dentro de este *locus* de acuerdo a sus genes blanco: L1\_S\_81 y L1\_S\_99, los cuales regulan la expresión de transcritos que codifican para un factor de transcripción de la familia MYB (Solyc06g071690) y para una enzima con actividad poligalacturonasa (Solyc12g009420), respectivamente (Tabla 2.5, Fig. 2.15). Los perfiles de acumulación de L1\_S\_81 y de L1\_C\_99 son constantes durante la maduración, siendo el segundo más abundante que el primero (Fig. 2.15 A y B, izquierda) y la validación de sus genes blanco a través de Cleaveland indica que hay una alta acumulación de transcritos cortados en los sitios de corte mediados por estos phasiARNs en las bibliotecas de degradoma analizadas (Fig. 2.15 A y B, derecha). Los niveles de expresión de los genes blanco se grafican en las figuras 2.15C (Solyc06g071690) y 2.15D (Solyc12g009420). Se observa que el transcripto perteneciente a la familia MYB presenta un pico de expresión en el estadio inmaduro y luego disminuye marcadamente, mientras que la enzima poligalacturonasa sólo mostró niveles de expresión génica en los dos primeros estadios, alcanzando su máximo en el estadio Rompiente. Estos dos genes blanco son atractivos desde un punto de vista productivo dado que el primero pertenece a una familia génica que participa en la vía de los fenilpropanoides, algunos de los cuales corresponden a metabolitos secundarios implicados en la adquisición de aroma y sabor de los frutos, y compuestos con capacidad antioxidante fundamentales en la dieta humana (J. Liu et al., 2015) y el segundo regula el desensamblaje de la red de polisacáridos de la pared celular durante la maduración de los frutos, permitiendo así su ablandamiento (Fangling Jiang et al., 2019).

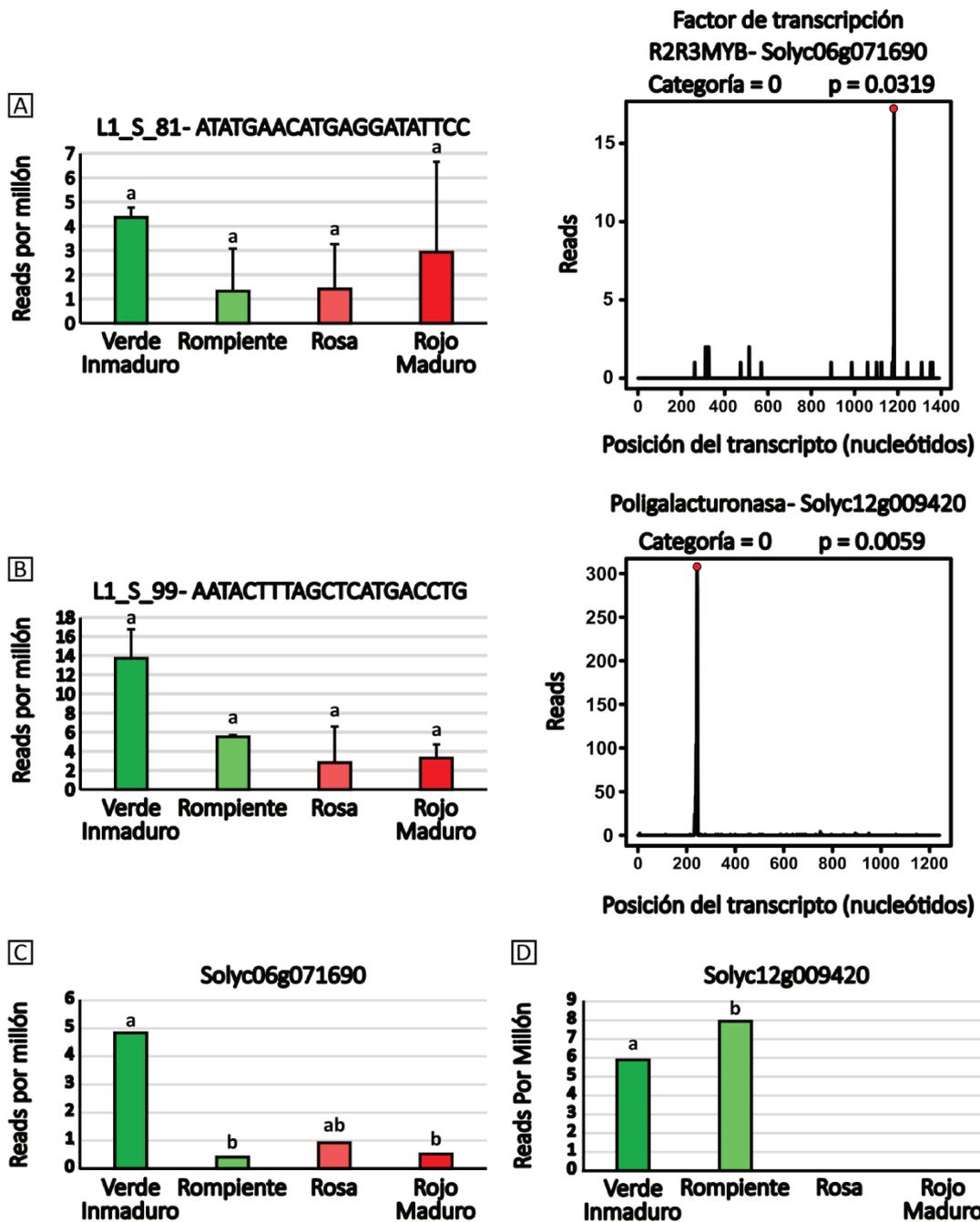


Fig. 2.15: (A) Perfil de expresión del phasiARN L1\_S\_81 en frutos de distintos estadios de madurez y la validación del corte del transcripto de Solyc06g071690. (B) Perfil de expresión del phasiARN L1\_S\_99 en frutos de distintos estadios de madurez y la validación del corte del transcripto de Solyc12g009420. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland. (C) Perfil de expresión del gen Solyc06g071690 en frutos de distintos estadios de madurez. (D) Perfil de expresión del gen Solyc12g009420 en frutos de distintos estadios de madurez. Las letras sobre las barras indican diferencia significativa (*p-value* ajustado < 0,05), mientras que su ausencia implica que no hay cambios estadísticamente significativos entre ninguno de los estadios.

## 2.3.4 - Otros loci PHAS relacionados con el desarrollo de propiedades organolépticas

### 2.3.4.1 - Locus 6

El análisis de validación de genes blanco con el software Cleaveland y bibliotecas de degradoma permitió la detección de un gen blanco para los phasiARNs del *locus* 6, cuya anotación corresponde a un factor de respuesta a auxinas (Solyc11g013480.2.1). En la Figura 2.16A se representan los niveles de expresión correspondientes a L6\_S\_2 para cada estadio de maduración, donde puede verse que no hay diferencias significativas entre los distintos estadios analizados, manteniéndose aproximadamente constante durante la maduración. El pico del gráfico *T-plot* obtenido para este gen blanco fue catalogado como “categoría 0”, lo que indica que la mayoría de las *reads* detectadas para este transcripto en las bibliotecas de degradoma analizadas coinciden con el sitio de corte mediado por L6\_S\_2. Los niveles de expresión para el gen ARF17 se grafican en la Figura 2.16B, donde se pueden observar niveles bajos en los estadios iniciales, seguidos por un incremento significativo conforme progresa la maduración.

Si bien el tomate es un fruto climatérico donde la mayoría de los aspectos relacionados con la maduración están regulados por la hormona etileno, existen publicaciones que indican que las auxinas están involucradas en el control del metabolismo de azúcares durante la maduración de tomate, por lo que la regulación de la expresión de ARF17 mediada por L6\_S\_2 resulta interesante (Sagar et al., 2013).

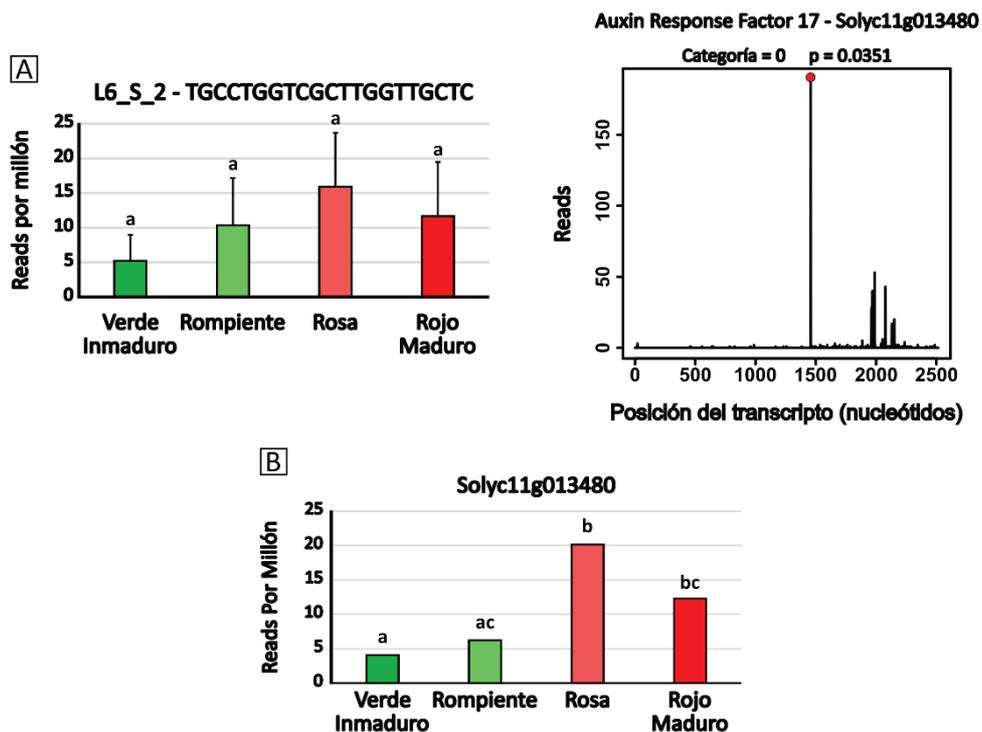


Fig. 2.16: (A) Perfil de expresión del phasiARN L6\_S\_2 en frutos de distintos estadios de madurez y validación del corte del transcripto Solyc11g013480. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland. (B) Perfil de expresión del transcripto Solyc11g013480 en frutos de distintos estadios de madurez. Las letras sobre las barras indican diferencia significativa (*p-value* ajustado < 0,05), mientras que su ausencia implica que no hay cambios estadísticamente significativos entre ninguno de los estadios.

#### 2.3.4.2 - Locus 10

Se validaron 5 genes blanco de phasiARNs pertenecientes al *locus 10*, cuyos transcriptos codifican para proteínas con funciones en diversos procesos celulares que se detallan en la Tabla 2.7. Se destaca L10\_S\_48 que regula el factor de transcripción TCP 15 (Fig. 2.17), cuya expresión en frutos ha sido demostrada aunque no se conoce aún su función (Parapunova et al., 2014), mientras que los demás genes no presentan relación con la adquisición de propiedades organolépticas o el desarrollo y maduración de los frutos.

Tabla 2.7: phasiARNs del *locus 10* con genes blanco detectados

phasiARN	Modelo génico	Gen asociado
L10_AS_30	Solyc02g081410.3.1	Beta catenin-like protein 1
L10_AS_28	Solyc06g068060.3.1	Sister chromatid cohesion protein PDS5 like
L10_AS_30	Solyc11g007700.2.1	S-adenosyl-L-methionine dependent methyltransferase superfamily protein
L10_S_51	Solyc08g067040.3.1	Transducin WD40 repeat-like superfamily protein
L10_S_48	Solyc01g008230.3.1	TCP transcription factor 15

En la Figura 2.17A se muestra el perfil de expresión sin cambios durante la maduración para L10\_S\_48, junto con el gráfico *T-plot* que resume el resultado del análisis de validación de corte del transcripto que codifica para SITCP15, mediado por L10\_S\_48. Los niveles de expresión del gen blanco de este phasiARN se representan en la Figura 2.17B, donde se observa que los estadios intermedios alcanzan los niveles máximos de expresión, mientras que el nivel de acumulación de los estadios inmaduro y maduro son significativamente más bajas y comparables entre sí.

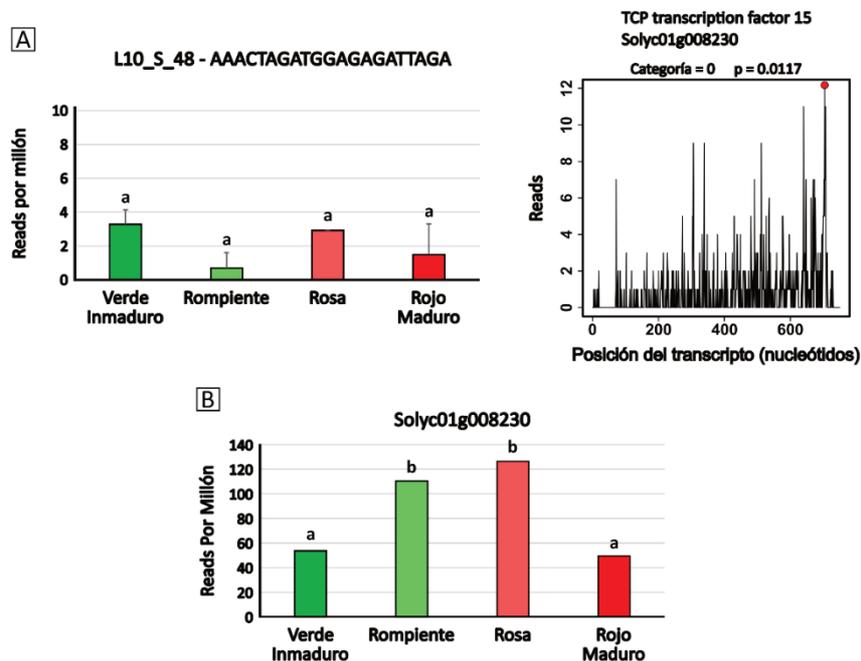


Fig. 2.17: (A) Perfil de expresión del phasiARN L10\_S\_48 en frutos de distintos estadios de madurez y la validación del corte del transcripto de Solyc01g008230.3.1. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland. (B) Perfil de expresión del gen Solyc01g008230 en frutos de distintos estadios de madurez. Las letras sobre las barras indican diferencia significativa (*p-value* ajustado < 0,05),

mientras que su ausencia implica que no hay cambios estadísticamente significativos entre ninguno de los estadios.

### 2.3.4.3 - Locus 15

Se validaron 12 genes blanco de phasiARNs pertenecientes a este *locus*, dentro de los que se destaca a L15\_S\_20 que es capaz de cortar al transcripto de un factor de transcripción de la familia GRAS (Solyc08g080400.1.1), mientras que los demás genes encontrados tienen funciones muy variadas pero no se relacionan con el desarrollo, la maduración o la adquisición de propiedades organolépticas de frutos. En la Figura 2.18 se representan los niveles de expresión de L15\_S\_20, que se mantienen constantes en los diferentes estadios de maduración analizados. El sitio de corte validado en el transcripto que codifica para el factor de transcripción GRAS se encuentra hacia el extremo 3' del transcripto y coincide con la acumulación de *reads* que indican que el corte del transcripto fue mediado por L15\_S\_20. El análisis de expresión de este gen se realizó usando datos de *RNA-seq* y se determinó que el mismo presenta niveles de expresión muy bajos durante la maduración (datos no mostrados). Los 11 genes restantes se detallan en la Tabla 2.8 junto con los phasiARN responsables del corte de sus transcriptos.

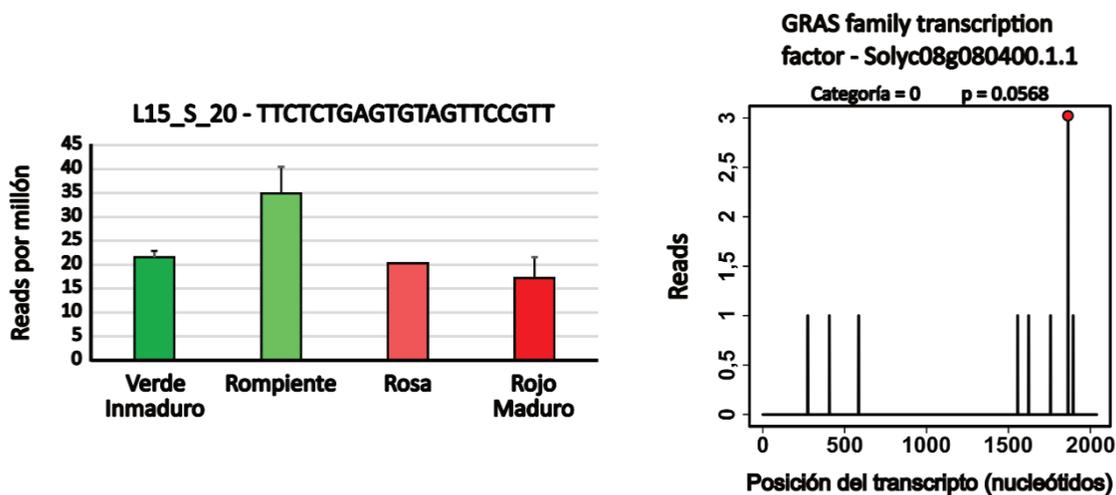


Fig. 2.18: Perfil de expresión del phasiARN L15\_S\_20 en frutos de distintos estadios de madurez y la validación del corte del transcripto de Solyc08g080400.1.1. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland.

Finalmente, es importante mencionar que ninguno de los phasiARNs asociados a los *loci* 2, 3, 4, 5, 8, 9, 11, 12, 13, 14, 16, 17, 18 y 19 tuvo como genes blanco validados

a transcritos implicados en el desarrollo y la maduración de frutos o en la adquisición de propiedades organolépticas, por lo que se decidió no profundizar su análisis.

**Tabla 2.8: phasiARNs del locus 15 con genes blanco validados**

<b>phasiARN</b>	<b>Modelo génico</b>	<b>Gen asociado</b>
<b>L15_AS_52</b>	Solyc05g018010.2.1	Disease resistance protein
<b>L15_AS_63</b>	Solyc03g043710.1.1	Clade IX lectin receptor kinase
<b>L15_S_124</b>	Solyc06g008720.3.1	CW-type zinc finger protein
<b>L15_S_154</b>	Solyc07g006670.1.1	HXXXD-type acyl-transferase family protein
<b>L15_S_154</b>	Solyc07g006680.1.1	HXXXD-type acyl-transferase family protein
<b>L15_S_159</b>	Solyc09g063150.3.1	Glutathione s-transferase
<b>L15_S_18</b>	Solyc09g090950.2.1	Far-red impaired responsive FAR1 family protein
<b>L15_S_20</b>	Solyc08g080400.1.1	GRAS family transcription factor
<b>L15_S_45</b>	Solyc02g024070.3.1	Homeobox leucine-zipper protein
<b>L15_S_6</b>	Solyc07g064740.3.1	Rab3 GTPase-activating protein catalytic subunit
<b>L15_S_9</b>	Solyc11g007300.2.1	Pleiotropic drug resistance ABC transporter
<b>L15_S_96</b>	Solyc09g090670.3.1	Oxidation resistance 1

## 2.4 - Discusión

### 2.4.1 - En frutos de tomate se expresan 48 familias de miARNs

Los datos analizados correspondientes a 4 estadios diferentes de maduración de frutos de tomate evidencian la expresión de miembros de 48 familias de miARNs de tomate. Un estudio previo había detectado un menor número de familias expresadas en frutos (J. Zuo et al., 2012), probablemente debido a que la base de datos de miARNs de tomate analizada no estaba actualizada, y a que analizaron sólo tres estadios de maduración, disminuyendo así el margen de búsqueda.

Se validaron genes blanco para miembros de 19 familias de miARNs expresadas en frutos, donde se vio que la mayoría regula factores de transcripción, proteínas de defensa y enzimas metabólicas. Los resultados obtenidos concuerdan con estudios publicados previamente en los cuales se realizaron análisis similares a los de esta tesis

doctoral, pero donde no se habían realizado análisis de expresión detallados en distintos estadios de maduración (Cedillo-Jimenez et al., 2020; Karlova et al., 2013). Sorprendentemente, nuestro análisis determinó que hay miembros de 29 familias de miARNs expresadas en frutos para los que no fue posible validar ningún transcripto como target. Es posible que esto esté relacionado con bajos niveles de expresión de los genes blanco o con expresión en zonas puntuales del fruto que resultan difíciles de ser detectados en bibliotecas construidas empleando los frutos completos. También es importante tener en cuenta que la validación de transcriptos target utilizando Cleaveland y bibliotecas de degradoma solo permite detectar aquellas regulaciones mediadas por corte de transcriptos, dejando de lado la regulación mediada por inhibición de la traducción tras la unión imperfecta de un miARN (Addo-Quaye et al., 2009; Axtell et al., 2011).

#### 2.4.2 - miARNs reguladores del desarrollo de propiedades organolépticas de frutos de tomate

Dentro de los miARNs expresados en frutos, se destacan algunos que regulan transcriptos con funciones relacionadas al desarrollo de las propiedades organolépticas de los frutos de tomate. En este contexto, podemos mencionar a miR159, para el cual existen reportes que demuestran sus múltiples funciones que abarcan desde la respuesta a estrés por sequía (López-Galiano et al., 2019) hasta la transición floral y la formación de frutos (E. M. da Silva et al., 2017), siempre a través de la regulación de genes GAMYB, pertenecientes a la familia de proteínas R2R3 MYB (Tabla 2.3). Para sly-miR159 se observó un patrón de expresión con diferencias significativas entre algunos estadios de maduración y se destacan los altos niveles de acumulación para sly-miR159 durante toda la maduración (Fig. 2.7). Los datos transcriptómicos muestran que el gen blanco tiene un nivel de expresión bajo durante la maduración, con un mayor nivel en el estadio inicial de la maduración, que luego disminuye considerablemente, para mantenerse relativamente constante en el resto de los estadios analizados (Fig. 2.8A). Es posible que los niveles de expresión elevados de este miARN durante toda la maduración estén relacionados con los bajos niveles de expresión de su gen blanco, aunque también es posible que esta regulación esté contribuyendo a la expresión en sectores especializados del gen del factor de transcripción GAMYB-like en fruto. Este

miARN ha sido sobreexpresado en tomate (E. M. da Silva et al., 2017) dando lugar a frutos partenocárpicos y a la formación temprana de los mismos. Además, un estudio reciente demostró que la sobreexpresión del gen SIMYB75, aumenta el contenido de antocianinas, sólidos solubles, flavonoides y compuestos fenólicos en comparación con los frutos control (Jian et al., 2019). Toda esta información apunta a que los factores de transcripción de la familia MYB y los miARNs de la familia miR159 son candidatos interesantes para la mejora de atributos de interés comercial de los frutos de tomate.

Tanto sly-miR396a como sly-miR396b presentan también altos niveles de expresión durante la maduración (Fig. 2.7). Los genes blanco de estos miARNs son, principalmente, los factores de transcripción de la familia GROWTH-REGULATING FACTOR (GRF, Tabla 2.3), implicados en la proliferación celular y el desarrollo de órganos (Rodríguez et al., 2016), aunque también regulan la expresión de ciertas proteínas relacionadas con la metilación del ADN (Karlova et al., 2013). Los datos transcriptómicos indican que estos genes blanco mantienen sus niveles de expresión bajos y constantes durante la maduración, como se muestra en la Fig. 2.8B para GRF3. Esto concuerda con otros reportes que establecen que la expresión de estos genes es baja en frutos de tomate (D. Cao et al., 2016). Probablemente en este caso también los bajos niveles de expresión de estos transcritos se mantengan a través de la elevada expresión, independientemente del estadio, de sly-miR396a y sly-miR396b (Fig. 2.7). Un estudio llevado a cabo en tomate mostró que al regular negativamente sly-miR396a y sly-miR396b la expresión de los genes GRF aumenta, dando lugar a frutos con un tamaño levemente mayor sin alterar el proceso madurativo, aunque no se ha realizado un estudio detallado de la influencia en las propiedades organolépticas de los frutos (D. Cao et al., 2016).

La regulación mediada por miR156 de los factores de transcripción de la familia SPL se ha caracterizado en detalle en *Arabidopsis* principalmente. La regulación de estos genes está implicada en la transición de la fase vegetativa a la floral, ya sea por corte (G. Wu & Poethig, 2006) como por inhibición de la traducción de los transcritos SPL (Gandikota et al., 2007). Los datos analizados en el presente trabajo muestran que la expresión de este miARN aumenta a medida que la maduración progresa (Fig. 2.7), al igual que su gen blanco (Fig. 2.8C). Una posible explicación para esta observación es que

esta regulación esté confinada a determinados sectores dentro del fruto que determinan que la correlación de expresión inversa entre el miARN y su transcripto blanco no se evidencie al analizar la expresión en frutos completos. Por otro lado, un estudio en tomate mostró que la sobreexpresión de miR156 altera el fenotipo de las plantas, principalmente disminuyendo su tamaño y el rendimiento de frutos (X. Zhang et al., 2011) Teniendo en cuenta estos resultados, controlar los niveles de expresión de miR156 podría ser una forma de mejorar la cantidad y el tamaño de los frutos producidos.

Los demás miARNs detectados y analizados en este estudio regulan genes que no están relacionados con la adquisición de propiedades organolépticas o no están suficientemente caracterizados en cuanto a su participación en la maduración de los frutos (Tabla 2.3), por lo que no profundizaremos en sus funciones, pero pueden resultar interesantes candidatos para explorar su influencia en los caracteres de interés agronómico de los frutos de tomate: la familia sly-miR319 regula la expresión de factores de transcripción TCP, implicados en un gran número de procesos biológicos, como producción de ramificaciones (Aguilar-Martínez et al., 2007), crecimiento de hojas (Danisman et al., 2012) y desarrollo floral (Nag et al., 2009; Schommer et al., 2014); miR171 controla en varias especies la expresión de la familia de factores de transcripción GRAS (X. Zhu et al., 2015), miR166 tiene como genes blanco a los factores de transcripción del tipo *HD-ZIP III* (Xuyan Li et al., 2017; Y. Y. Yang et al., 2020) y la familia miR482 actúa sobre proteínas del tipo *coiled-coil nucleotide-binding site leucine-rich repeat* (CC-NBS-LRR), directamente relacionadas con la respuesta frente al ataque de patógenos (De Vries et al., 2015).

#### 2.4.3 - Se identificaron 19 loci productores de phasiARNs en frutos de tomate

Hasta el momento existen escasos reportes acerca de phasiARNs de tomate, y aún no existe bibliografía que presente un análisis de phasiARNs durante la maduración de frutos. Wu y colaboradores detectaron 91 *loci PHAS* analizando muestras provenientes de hojas de tomate, y demostraron que la expresión de 55 phasiARNs generados a partir de estos *loci* aumenta frente a la presencia del hongo *Botrytis cinérea* (F. Wu et al., 2017). A su vez, Luan y colaboradores (W. Luan et al., 2020) identificaron

137 phasiARNs, también presentes en hojas de tomate, siendo la mayoría producidos a partir de genes de defensa del tipo NBS-LRR y cuya expresión se modifica en respuesta a ácido abscísico.

En esta tesis doctoral se identificaron por primera vez 19 *loci* PHAS expresados durante la maduración de frutos de tomate. La ubicación de 14 de estos *loci* coincide con genes anotados en el genoma de tomate. Es destacable que 5 de ellos superponen su ubicación genómica con el gen *sly-DCL2b*, de 11 kpb (Tabla 2.4). De estos 19 *loci* PHAS se validaron 82 genes blanco regulados por phasiARNs derivados de ellos, dentro de los cuales reportamos 16 que tendrían una relación directa con las propiedades organolépticas de frutos de tomate (Tabla 2.5) y que se discuten en las secciones siguientes.

#### 2.4.4 - Se validaron 82 genes blanco para los phasiARNs detectados

Estudios previos destacan el rol de los phasiARNs en la regulación de la expresión de proteínas de defensa frente al estrés biótico en diversas especies (Qili Fei et al., 2013; F. Wu et al., 2017) y un trabajo demuestra la relación entre estos ARNs pequeños y el estrés abiótico, puntualmente la resistencia a sequía, en *Populus trichocarpa* (Shuai et al., 2016). También se han detectado phasiARNs en tejidos florales, relacionados con la reproducción, en varias especies (Kakrana et al., 2018; Y. Liu et al., 2020; Xia et al., 2019). Sin embargo, no se encontraron reportes que analicen phasiARNs en frutos y los vinculen a la maduración de los mismos o al desarrollo de propiedades organolépticas.

En este trabajo de tesis validamos 82 genes blanco para los phasiARNs identificados en frutos de tomate y la clasificación funcional de los mismos determinó que la mayoría de estos genes blanco codifica para proteínas de unión, desarrollo y maduración de frutos y enzimas no metabólicas (Fig. 2.9), incluyendo algunos vinculados a la adquisición y regulación de propiedades organolépticas de frutos y factores de transcripción.

#### 2.4.5 - El microARN *sly-miR6026* también genera phasiARNs en frutos

Al inicio de este trabajo de tesis doctoral identificamos la producción de phasiARNs a partir de la unión de *miR6026* al transcripto de *sly-DCL2a*, *sly-DCL2b* y *sly-DCL2d* (Tabla 2.3; Fig. 2.10B), lo cual fue posteriormente reportado también para

muestras de hoja de tomate, en dos publicaciones donde se caracterizó la producción de estos phasiARNs y en la respuesta a infecciones con virus (T. Wang et al., 2018; Zhengming Wang et al., 2018). Nuestro análisis determinó que los phasiARNs iniciados por la unión de miR6026 a los transcritos de sly-DLC2 también se producen en frutos de tomate e identificamos genes blanco para varios de los phasiARNs generados a partir de estos *loci*. Si bien nuestro análisis se orientó a validar targets relacionados con el desarrollo de propiedades organolépticas, resta conocer si algunos de estos phasiARNs modifican su expresión ante el ataque de virus u otros patógenos en frutos de tomate. Además, Wang y colaboradores detectaron que la proteína DCL2b es fundamental para la síntesis de ARNs pequeños de 22 nucleótidos, incluidos algunos miARNs que podrían funcionar como generadores de phasiARNs, (T. Wang et al., 2018) lo cual no fue analizado en el marco de esta tesis doctoral, ya que centramos nuestros esfuerzos en el análisis de siARNs de 21-nt de longitud. Por otro lado, en ninguno de los dos trabajos se analizan los genes blanco regulados por los phasiARNs detectados.

Dentro de los genes blanco identificados, nos resultó de particular interés el phasiARN L7\_S\_30, regulador de una enzima de la familia de lipasas GDGL (Fig. 2.12A), relacionada directamente con la formación de la cutícula del fruto y por lo tanto, con propiedades protectoras de la integridad del mismo y la adquisición de brillo (Petit et al., 2014). Los niveles de expresión de este gen se muestran bajos y relativamente constantes, excepto en el estadio rosa, donde no se detectó expresión (Fig. 2.12C). El análisis de expresión de L7\_S\_30 indica que este phasiARN se encuentra en todos los estadios en cantidades similares (Fig. 2.12A), resultado coherente con lo observado para el transcripto blanco. Es importante señalar que las bibliotecas analizadas corresponden a frutos completos y no exclusivamente a la cutícula de los mismos, sitio donde se expresa preferentemente este gen (Girard et al., 2012). Cabe destacar que este grupo de enzimas no se ha estudiado al detalle en tomate, por lo que resulta interesante su caracterización a futuro en relación con su influencia en el aspecto del fruto y su regulación por parte de este ARN pequeño.

Se encontraron también phasiARNs en este *locus* que no comparten la fase con el corte predicho de sly-miR6026, y regulan genes relacionados a la maduración, como un Factor de Respuesta a Etileno, y a la producción de compuestos volátiles, como la

enzima Alcohol deshidrogenasa (Fig. 2.13). Existen reportes que vinculan esta proteína a la producción de aroma en varias especies (Echeverría et al., 2004; Manríquez et al., 2006) y a la respuesta ante situaciones de estrés, principalmente frente a hipoxia (R. K. Singh et al., 2018). En tomate, se ha podido relacionar a esta familia de enzimas directamente con el sabor de los frutos (Speirs et al., 1998), pero hay pocos estudios moleculares que intenten mejorar este atributo a través de la sobreexpresión o modificación de estos genes. Teniendo en cuenta esto, un phasiARN con la capacidad de regular la expresión de un transcripto que codifica para la enzima Alcohol deshidrogenasa aparece como un candidato interesante para analizar en profundidad en el futuro.

#### 2.4.6 - La ubicación del locus 1 coincide con la del gen Sly-DCL2b

Se detectó un *locus PHAS* cuya ubicación coincide con la del intrón 5 del gen Solyc11g008540.2.1, correspondiente a sly-DCL2b (tabla 2.4). Usando el software de visualización genómica *IgV Viewer*, se detectó la presencia de *reads* en ambas hebras para esta región. También se encontró evidencia de ARNm al analizar bibliotecas de *RNA-seq* en esta región, donde habitualmente no debería detectarse (Fig. 2.14). Puede tratarse de ARNs largos no codificantes, o *lncRNAs*, moléculas de ARN de 200 o más pares de bases que no codifican proteínas funcionales (J. Jin et al., 2013). Se sabe que regulan diversos procesos, como la floración, la transcripción, el *splicing* y la metilación del ADN dirigida por ARN (Chekanova, 2015). A su vez, pueden ser procesados para dar lugar a la producción de ARNs pequeños o regular el procesamiento de otros ARNs para generar estas moléculas (Wilusz et al., 2009), por lo tanto, es posible que los ARNs pequeños de este *locus* provengan de un lncRNA aún no caracterizado.

Para este *locus* se destacan los phasis L1\_S\_81 y L1\_S\_99, encargados de la regulación de un factor de transcripción de la familia MYB y de una enzima con actividad poligalacturonasa, respectivamente (Fig. 2.15). Los factores MYB constituyen una gran familia presente en todos los eucariotas. En plantas están ampliamente distribuidos y cumplen un gran número de funciones, como respuesta a ácido abscísico, iniciación de tricomas y producción de antocianinas, entre otras (Ambawat et al., 2013; Ramsay & Glover, 2005). En tomate se han encontrado 127 miembros de esta familia, aunque la función de la mayoría de ellos es aún desconocida (Z. Li et al., 2016). Diversos reportes

relacionan a estas proteínas con la coloración del fruto y las flores en varias especies, como manzana (Espley et al., 2007), uva (Kobayashi et al., 2002), y petunia (Mol et al., 1998). También se encontró que la sobreexpresión de un gen MYB de uva en tomate evidenció cambios en el desarrollo de la planta, regulación negativa del metabolismo de fenilpropanoides y aumento en la cantidad de  $\beta$ -caroteno (Mahjoub et al., 2009). Tanto L1\_S\_81 como su gen blanco Solyc06g071690 presentan niveles de expresión bajos durante la maduración y que no se correlacionan con los perfiles esperados para una regulación por corte de transcripto (Fig. 2.15A y 2.15C), por lo que es posible que también en este caso el análisis de frutos completos esté dificultando el análisis de esta regulación que puede estar restringida a determinados dominios o tejidos dentro del fruto o a los múltiples tipos de regulación de la expresión que afectan a esta familia de factores de transcripción (C. Li et al., 2015). La identificación de un phasiARN con la capacidad de regular genes de la familia MYB puede abrir las puertas a cultivos mejorados genéticamente, por lo que sería fundamental estudiar en detalle la función del gen blanco y su regulación por este phasiARN.

Las poligalacturonasas hidrolizan los enlaces  $\alpha$ -1,4 glucosídicos entre los residuos de ácido galacturónico, componentes de la red de pectina integrante de la pared celular vegetal (Fry, 2004; Illera et al., 2018). La degradación de esta matriz de polisacáridos da como resultado el ablandamiento de los frutos. Mercado y colaboradores silenciaron un gen de una poligalacturonasa de frutilla, y descubrieron que las frutas obtenidas eran significativamente más firmes que el control (Mercado et al., 2009). También se vio que en tomates con reducida actividad de estas proteínas, la vida poscosecha mejoraba, al igual que la resistencia de los frutos al transporte (Schuch et al., 1991). Además, se ha detectado una mayor expresión de estos genes en células de zonas de abscisión durante este fenómeno en frutos de naranjo (Merelo et al., 2017), mientras que en tomate se encontró que su silenciamiento contribuye a retrasar el desprendimiento de las hojas (C. Z. Jiang et al., 2008). En este estudio se encontró expresión de este gen en los estadios verde inmaduro y rompiente (Fig. 2.15D), por lo que su actividad sería menor a medida que los frutos maduran. Estos patrones no coinciden con los de L1\_S\_99 (Fig. 2.15B), que también disminuye en estadios maduros, tal vez porque la regulación mediante phasiARNs no es la principal sobre este

transcripto. El phasiARN detectado podría utilizarse para a reducir la acumulación de transcriptos que codifican para poligalacturonasas, por lo que un incremento de estos ARNs pequeños disminuiría de forma significativa el desensamblado de la pared celular, lo que podría generar frutos con mayor vida poscosecha.

#### 2.4.7 - Se encontraron phasiARNs reguladores de genes vinculados a las propiedades organolépticas en otros 3 loci

Nuestro análisis permitió determinar que los *loci* 6, 10 y 15 producen un *phasiARN* cada uno con la capacidad de regular la expresión de un gen importante desde el punto de vista de la maduración o la adquisición de propiedades organolépticas. El ARN pequeño correspondiente al *locus* 6 tiene como gen blanco a un factor de respuesta a auxinas, ARF17 (Solyc11g013480) (Fig. 2.16A). Al ser un fruto climatérico, la maduración del tomate está principalmente relacionada con la hormona etileno, pero hay reportes que vinculan a las auxinas a la regulación del establecimiento de los frutos (De Jong et al., 2009). Además, en otro estudio se vio que los genes ARFs (y en particular SIARF2) regularían positivamente la maduración de los frutos (Hao et al., 2015). Esta familia de proteínas ya ha sido caracterizada en tomate (R. Kumar et al., 2011), pero no hay estudios donde se reporte una regulación mediante phasiARNs.

Dentro del *locus* 10 se validaron genes blanco para 5 phasiARNs (Tabla 2.7), y se detectó a L10\_S\_48, cuyo gen blanco es el factor de transcripción SITCP15 (Fig. 2.17). Si bien no está demostrado por el momento que este gen esté relacionado con la maduración de los frutos, sí se sabe que se expresa preferentemente en este órgano (Parapunova et al., 2014). En este trabajo se vio que la expresión de L10\_S\_48 es baja y constante durante la maduración, lo que se correlaciona con los niveles altos de SITCP15, que son particularmente elevados en los estadios rompiente y rosa, presentando diferencias significativas con verde inmaduro y rojo maduro (Fig. 2.17B).

Finalmente, se validaron 12 genes blanco para el *locus* 15 con diversas funcionalidades (Tabla 2.8). Dentro de estos, L15\_S\_20 es considerado como candidato para análisis posteriores, dado que regula a un gen de la familia de factores de transcripción *GRAS* (Fig. 2.18), aunque la expresión de este transcripto es baja en frutos y es detectable sólo en los estadios tempranos. Huang y colaboradores sobreexpresaron

y reprimieron en tomate a un miembro de la familia *GRAS* (SIGRAS24) y también sobreexpresaron a un miembro de la familia *sly-miR171* (Huang et al., 2017). Observaron varias modificaciones fenotípicas, como tiempo de floración, tamaño de la planta y establecimiento y maduración de frutos. Otro estudio logró silenciar a esta familia de miARNs, obteniendo como resultado plantas de tomate con desarrollo anormal y unos pocos frutos pequeños y sin semillas (Kravchik et al., 2019). Resulta interesante analizar funcionalmente la regulación de este factor de transcripción de la familia *GRAS* mediada por *L15\_S\_20*, para conocer en detalle la participación en el desarrollo de frutos de tomate.

Teniendo en cuenta el conjunto de resultados expuestos aquí, nuestros análisis aportan una gran cantidad de módulos regulatorios miARN/target o phasiARN/target que podrían resultar útiles como herramientas para el mejoramiento genético de atributos de interés agronómico relacionados con los aspectos cualitativos de los frutos de tomate y con las propiedades organolépticas atractivas para los consumidores.

## 2.5 - Conclusiones

- Se actualizó la base de datos de miARNs de tomate, encontrando un total de 158 miARNs pertenecientes a 65 familias.
- Al menos un miembro de 48 familias de miARNs se expresan durante la maduración de frutos.
- Se detectaron 19 *loci* productores de ARNs pequeños en fase, de los cuales 15 corresponden a *loci* PHAS y 4 a *loci* TAS.
- Se validaron 82 genes blanco para los phasiARNs encontrados, y 16 de ellos tienen funciones relacionadas con el desarrollo y la maduración de frutos.
- Como genes blanco de regulación de ARNs pequeños se destacan varios genes relacionados con la adquisición de propiedades organolépticas de frutos de tomate, detallados en la Tabla 2.5.

## Capítulo 3: ARNs pequeños en frutilla

## 3.1 - Introducción

### 3.1.1 - Breve descripción del género *Fragaria* y origen de *Fragaria x ananassa*

El género *Fragaria*, perteneciente a la familia Rosacea, está formado por 23 especies aceptadas (*Fragaria*, en The Plant List, 2013) cuyos genomas son desde diploides ( $2n = 2x = 14$ ) hasta decaploides ( $2n = 10x = 70$ ). Los miembros salvajes de este género se distribuyen a lo largo del hemisferio norte y Sudamérica, mientras que la frutilla comercial (*Fragaria x ananassa*) se cultiva en todo el mundo. Esta última es una especie híbrida octaploide que tiene 7 juegos de cromosomas, 56 cromosomas en total. Se originó en Francia hace aproximadamente 250 años, cuando accidentalmente se produjo una cruce de dos especies americanas, un clon femenino de *F. chiloensis subsp. chiloensis* de fruta blanca proveniente de Sudamérica y *F. virginiana subsp. virginiana*, hermafrodita con fruto rojo, de América del Norte (Fig. 3.1; Hancock et al., 2010, <https://www.rosaceae.org/species/fragaria/all>). Es una especie alopoliploide con cuatro sub-genomas, uno de los cuales se cree que corresponde al genoma de la especie salvaje *Fragaria vesca* (Edger et al., 2019).

### 3.1.2 - Morfología del fruto

Desde un punto de vista botánico, la frutilla es un fruto agregado, ya que la parte pulposa se origina a partir del receptáculo que sostiene los múltiples pistilos de las flores. Los frutos verdaderos son los aquenios, distribuidos por fuera del receptáculo (Fig. 3.1). Cada uno de ellos deriva de un ovario y contienen una semilla en su interior (Ariza et al., 2011; Trejo-Téllez, Libia, Gómez-Merino, 2004). Para simplificar la terminología, de ahora en adelante la palabra “fruto” hará referencia al receptáculo.

***Fragaria × ananassa***



1cm

***Fragaria chiloensis***



1cm

***Fragaria virginiana***



1cm

Fig. 3.1: Comparación de especies del género *Fragaria*. Adaptado de Vallarino et al.(2018).

### 3.1.3 - Aspectos nutricionales

Los diversos cultivares de *Fragaria x ananassa* producen frutos grandes, muy coloridos y con un mayor rendimiento y vida poscosecha que las especies salvajes. Estos aspectos, muy apreciados por productores y consumidores, se lograron a costa de la variedad e intensidad de los aromas que presentan las frutillas silvestres (Negri et al., 2015). La calidad nutricional y el atractivo sensorial dependen en gran medida del contenido de metabolitos secundarios. Este fruto es rico en polifenoles como las antocianinas, una gran familia de compuestos que además de ser los principales responsables de la coloración (Yoshida et al., 2002), son importantes desde el punto de vista nutricional por su conocido efecto antioxidante (Meyers et al., 2003). Existen numerosos reportes en la literatura donde se ha caracterizado la participación de factores de transcripción de la familia MYB en la regulación de varias reacciones de las vías de los fenilpropanoides y de los flavonoides, que se indican con un asterisco en la Fig. 3.2 (Medina-Puche et al., 2014). Los productos de la ruta de los fenilpropanoides, flavonoides y terpenoides influyen directamente sobre los parámetros que determinan la calidad de los frutos, tal como el contenido de sólidos solubles totales, el pH, el contenido de vitaminas C y E, la actividad antioxidante y el contenido de nutrientes (Fig. 3.2, do Nascimento Nunes, 2009; Hui et al., 2010).

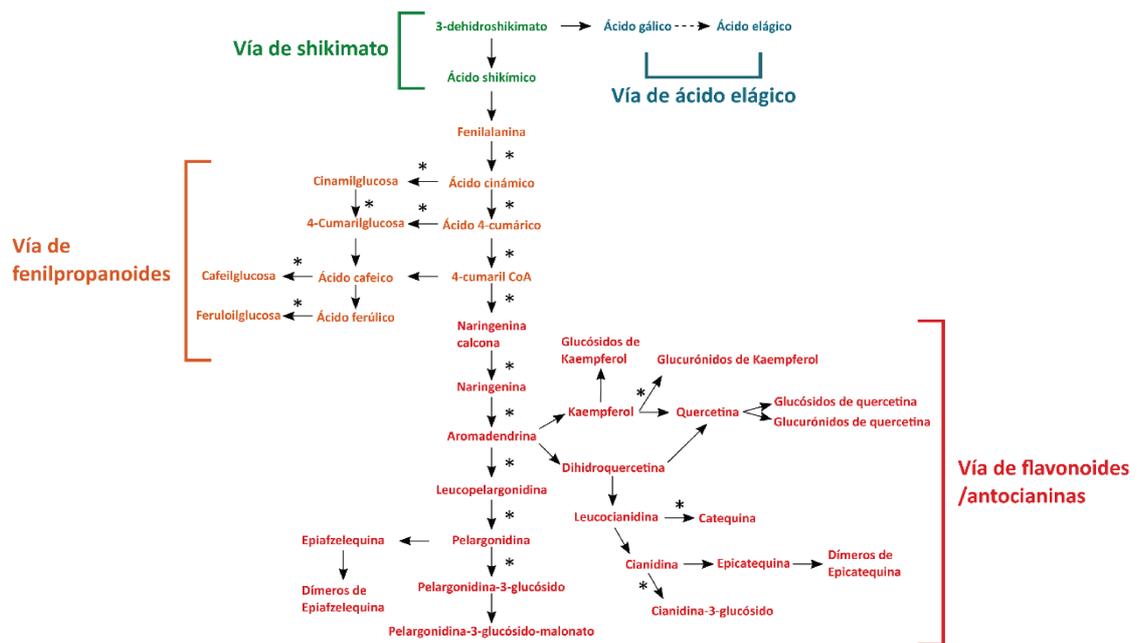


Fig. 3.2: Esquema de las vías metabólicas de shikimato, ácido elágico, fenilpropanoides y flavonoides. Los asteriscos indican regulación de factores de transcripción MYB en ese punto (Medina-Puche et al., 2014). Adaptado de Härtl et al. (2017).

### 3.1.4 - Maduración no climatérica

El avance de la maduración de frutilla se clasifica según el color y el tamaño de los frutos: verde, blanco, rosa y rojo. En el estadio verde el receptáculo se puede distinguir fácilmente de los aquenios, y el fruto es muy pequeño, por lo que se denomina a este periodo “fase de crecimiento”. En el estadio blanco comienza la “fase de maduración”, el rosa evidencia un fruto de mayor tamaño con el 50% de su coloración final, y por último, en el estadio rojo se alcanzan la tonalidad y las dimensiones definitivas (Fig. 3.3). Al igual que las naranjas, las frutillas son frutos no climatéricos. Su maduración no presenta un pico de etileno ni de CO<sub>2</sub> (aumento de respiración) y se detiene cuando el fruto se desprende de la planta. Se sabe considerablemente menos de maduración no climatérica en relación a la climatérica y por el momento no se ha detectado la participación obligatoria de un compuesto en particular que actúe de forma análoga al etileno. El crecimiento del receptáculo está regulado por los aquenios, principales productores de ácido indol acético, una de las principales auxinas (Seymour et al., 2012). Además, Given y col. (1988) demostraron que si se remueven los aquenios sólo a una mitad de una fruta inmadura, esta mitad mostraba una maduración acelerada respecto a la otra, ya que la concentración de auxinas disminuye. Se ha comprobado

también, que además de las auxinas, el ácido abscísico es un compuesto regulador fundamental en la maduración de los frutos de esta especie (H. F. Jia et al., 2011).

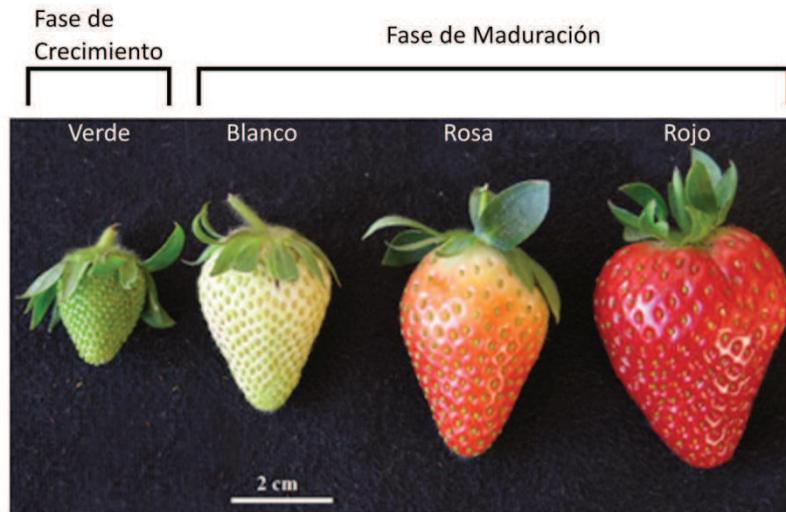


Fig. 3.3: Estadios de maduración de frutos de frutilla. Adaptado de Basson et al. (2010).

Las frutillas adquieren su característico color rojo hacia el final de la maduración, gracias a la aparición de antocianinas. Estos pigmentos son flavonoides, metabolitos secundarios derivados de la vía de los fenilpropanoides, siendo la enzima Fenilalanina amonio liasa la encargada de comenzar la síntesis a partir de fenilalanina (Y. Jiang & Joyce, 2003). A la par de la coloración se adquieren el sabor y el aroma, cualidades que varían mucho entre diferentes cultivares. Estas propiedades organolépticas son muy complejas para estudiar, dado que están determinadas por alrededor de 350 compuestos volátiles, como aldehídos, alcoholes, terpenos y ésteres (Forney et al., 2000).

Los frutos maduros pierden progresivamente su firmeza, debido al desensamblaje de la pared celular de las células del parénquima (Heng Koh & Melton, 2002). Pectina, celulosa y varios xiloglucanos son los principales blancos de enzimas líticas, que llevan al ablandamiento de los frutos (M. L. Tucker, 2014). Este fenómeno causa rechazo en los consumidores y genera un problema para el almacenamiento y posterior comercialización de los frutos una vez cosechados, lo que lo convierte en un punto interesante para obtener cultivares mejorados genéticamente. Por ejemplo, Paniagua y col. (2016) encontraron que silenciando un gen que codifica para una enzima

$\beta$ -galactosidasa se obtienen frutos más firmes, aunque levemente más pequeños (Paniagua et al., 2016). Se realizaron estudios similares y con resultados acordes con una Pectato liasa (Jiménez-Bermúdez et al., 2002) y con una Poligalacturonasa (Mercado et al., 2009).

### 3.1.5 - Poscosecha

Las frutillas tienen una vida poscosecha muy corta, principalmente porque son atacadas por patógenos, como *Botrytis cinerea* y *Rhizopus stolonifer*. Almacenadas en frío (0°C) se conservan por aproximadamente 2 semanas. Sus elevados niveles de agua y su bajo pH las vuelven particularmente susceptibles a hongos. Por ello se han desarrollado diversos mecanismos para recubrir los frutos con sustancias comestibles que sean inocuas para los consumidores y que a su vez mantengan la integridad del producto, por lo que en algunos países normalmente se usan como cobertores sustancias compuestas por distintas proteínas o polisacáridos (Gol et al., 2013; Vu et al., 2011). Fundamentalmente actúan como una barrera contra la humedad impidiendo el ataque y desarrollo de microorganismos sin alterar ni la calidad nutricional ni las propiedades organolépticas del fruto. Además, se vio que los aceites esenciales tienen propiedades antifúngicas (Lacroix, 2007), lo que los convierte en otro elemento de protección natural bien visto a los ojos de los consumidores.

### 3.1.6 - Importancia comercial

De acuerdo a datos del año 2018 (Food and Agriculture Organization of The United Nations – FAO) el principal productor de frutillas es China, seguido por Estados Unidos y México. Estos 3 países fueron responsables del 59% de la producción global, correspondiendo el 35.6% sólo a China. En Argentina se cosecharon alrededor de 15000 toneladas de este fruto (0.18% del total mundial), llegando al puesto 33 de esta lista. En nuestra nación, este cultivo representa una actividad económicamente muy importante. Las principales provincias productoras son Santa Fe, Tucumán, Buenos Aires (que concentran cerca del 70% de la producción), Jujuy y Corrientes (<https://www.argentina.gob.ar/noticias/la-argentina-es-el-tercer-pais-productor-de-frutillas-de-sudamerica>). De acuerdo al Instituto Nacional de Estadísticas y Censos de la República Argentina (INDEC), en el año 2018 la exportación alcanzó un monto de U\$S

1.708.581. Las principales variedades producidas localmente son “Camarosa”, “Sweet Charlie”, “Selva”, “Milsei Tudla”, “Chandler” y “Pájaro” (García et al., 2002).

### 3.1.7 - Genoma de frutilla y avances en bioinformática

La complejidad genética de *Fragaria x ananassa* dificultó mucho la obtención de un genoma de referencia de calidad que pueda ser utilizado para llevar adelante análisis bioinformáticos precisos en esta especie. El procedimiento más común en estudios genómicos iniciales fue emplear como referencia a la especie silvestre diploide *F. vesca*, que ha sido estudiada en profundidad y cuenta con un mayor número de recursos y bases de datos disponibles. Su genoma tiene varias versiones: la actual es la v4.0, publicada en 2018 (Edger et al., 2018), con anotaciones mejoradas respecto a su antecesora.

A principios de 2019 se publicó una versión casi completa del genoma de *Fragaria x ananassa* (Edger et al., 2019), siendo la primera en lograr distribuir la información en cromosomas para esta especie. Su tamaño aproximado es de 813.4 Mb, y si bien es muy reciente y aún resta mejorar la anotación, este nuevo genoma secuenciado de la variedad “Camarosa” permitirá realizar análisis más específicos y con mejor resolución respecto a los que se hacían previamente.

### 3.1.8 - ARNs pequeños en plantas de frutilla

Recientemente se han publicado varios trabajos sobre microARNs en *Fragaria x ananassa* (D. Li et al., 2019) y en *F. vesca* (Liang et al., 2018), pero los números disminuyen mucho cuando se trata de estudios de otro tipo de ARNs pequeños. Por ejemplo, Xia y col. (2015) analizaron por técnicas informáticas y de secuenciación masiva la diversidad de ARNs pequeños presentes en distintos tejidos de *F. vesca*. Descubrieron que un microARN que regula genes de la familia F-box es el encargado de la producción de phasiARNs en esta especie.

En este trabajo se propuso realizar un análisis genómico exploratorio con la intención de detectar *loci* PHAS generadores de especies de ARNs pequeños y los posibles genes regulados por ellos. Para eso se emplearon bibliotecas de *small RNA-seq* y *RNA-seq* de *Fragaria x ananassa* de frutos del cultivar Hongjia al inicio y al final de la maduración y una biblioteca de degradoma de *F. vesca* para validar los resultados.

Además se realizó un análisis de expresión diferencial buscando las secuencias de los miARNs conocidos de *F. vesca* en las bibliotecas de *Fragaria x ananassa* para evaluar la conservación de miARNs entre estas dos especies y estudiar las variaciones en los niveles de expresión con la progresión de la maduración de los frutos.

## 3.2 - Materiales y métodos

Los comandos utilizados para los procedimientos descritos en esta sección se detallan en el Anexo I.

### 3.2.1 - Datos de secuenciación masiva

Se realizaron búsquedas bibliográficas en Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>) y en la base de datos *Sequence Read Archive* (SRA, <https://www.ncbi.nlm.nih.gov/sra>), para identificar datos obtenidos por NGS a partir de frutos de frutilla en distintos estadios de madurez y que cuenten con réplicas biológicas. Se utilizó la herramienta *fastq-dump* del paquete *SRA-toolkit* (<http://ncbi.github.io/sra-tools/>) como se mencionó previamente y se descargaron los conjuntos de datos de secuenciación de bibliotecas de *small RNA-seq*, de *RNA-seq* de frutos de frutilla *Fragaria x ananassa* cv. Hongia, y datos de degradomas de *Fragaria vesca* cv. *Yellow wonder 5AF7* (Tabla 3.1).

### 3.2.2 - Pre-procesamiento de datos de NGS

Se evaluó la calidad de secuenciación de los archivos *fastq* descargados utilizando la herramienta *FastQC* v 0.11.5 (Andrews, 2010). Para cortar la secuencia correspondiente a los adaptadores de secuenciación se utilizó el programa *Fastx clipper* v. 0.0.14 (Gordon, 2011), modificando los parámetros por *default* como se indica en el punto 2.2 de materiales y métodos del Capítulo 2.

#### 3.2.2.1 - Filtrado de secuencias contaminantes

Con el fin de eliminar las secuencias de ARN ribosomales (ARNr), ARN de transferencia (ARNt), ARN nucleares pequeños (snARN) y ARN nucleolares pequeños (snoARN), que no son de interés para el análisis, se las descargó de la base de datos *Genebank* (<https://www.ncbi.nlm.nih.gov/genbank/>). Se reemplazaron las letras U por T, correspondientes a uracilo y timina, respectivamente, mediante reemplazo masivo de caracteres usando el comando “sed”. Para realizar el filtrado de las secuencias

contaminantes se utilizaron los archivos fastq sometidos al pre-procesado detallado en el punto anterior y se utilizó el programa Bowtie v 1.1.1 (Langmead, 2010) para retener aquellas *reads* que no alinean sobre el conjunto de secuencias contaminantes de frutilla obtenidas de *Genebank*. El conjunto de secuencias contaminantes fue indexado usando la herramienta bowtie-build v 1.1.1 y para el alineamiento con bowtie se utilizó el modo “*end to end*” (-v) permitiendo 2 mismatches, y con la opción “--un” se guardaron las *reads* limpias que no alinearon a la referencia de secuencias contaminantes.

### 3.2.3 - Análisis de miARNs de frutilla

#### 3.2.3.1 - Descarga de la información correspondiente a miARNs de frutilla

Se descargó la información correspondiente a miARNs de frutilla de la especie *Fragaria vesca* de la base de datos de miRbase (versión 22.1, Kozomara & Griffiths-Jones, 2014), dado que no se encuentra disponible por el momento una lista de miARNs correspondientes a *Fragaria x ananassa*. La información se almacenó en formato fasta para posteriormente indexarla usando bowtie-build y luego emplear el programa bowtie y utilizarla como referencia para los alineamientos.

#### 3.2.3.2 - Mapeo al genoma de referencia, anotación y cuantificación de miARNs

Luego del pre-procesado, se alinearon los archivos en formato fastq al genoma de referencia (*Fragaria x ananassa* versión 1.0, cv. *Camarosa*) siguiendo un procedimiento análogo al mencionado en el punto 2.3.2 de Materiales y métodos del Capítulo 2. Posteriormente, con la herramienta fastx\_collapser v. 0.0.14 (Gordon, 2011), utilizando las opciones por *default*, se colapsaron y cuantificaron las *reads* repetidas de los archivos resultantes. Se alineó cada uno de estos ficheros colapsados a la lista de miARNs de igual manera a la indicada en el punto 2.3.2 de Materiales y métodos del capítulo de tomate. Para conservar sólo aquellos miARNs que tuvieron alineamientos perfectos, se utilizó la herramienta Samtools v 1.5 (Heng Li et al., 2009), siguiendo el procedimiento detallado en el punto 2.3.2 de Materiales y métodos del Capítulo 2. La información obtenida se organizó como tabla indicando los niveles de expresión normalizados en *reads* por millón (RPM) de cada miARN para cada réplica de cada estadio de madurez.

### 3.2.4 - Análisis de expresión diferencial de miARNs

Para evaluar la expresión diferencial de cada miARN a lo largo de la maduración de los frutos se utilizó el paquete DESeq2 1.20.0 (Love et al., 2014) dentro del software Rstudio 1.1.456 (Allaire, 2012), utilizando la versión 3.6.0 de R. Se realizaron comparaciones entre los distintos estadios de madurez y se consideró como diferencia estadísticamente significativa a aquellas comparaciones con un valor de p-ajustado (FDR)  $\leq 0,05$ .

### 3.2.5 - Detección de loci en fase

Los archivos fastq pre-procesados correspondientes a réplicas biológicas de un mismo estadio fueron combinados en un solo archivo ya que una mayor cantidad de lecturas beneficia la identificación de *clusters* de ARNs pequeños en fase. Luego se utilizó el programa Shortstack 3.8.3 (Axtell, 2013), de acuerdo a lo indicado en el punto 2.5 de Materiales y métodos del Capítulo 2.

### 3.2.6 - Mapeo de datos de RNA-seq al genoma de referencia

Los datos de secuenciación de *RNA-seq* obtenidos de bibliotecas correspondientes a frutos de *Fragaria x ananassa* cv. Hongjia en estadios verde, intermedio y rojo fueron descargados de la base de datos SRA (Tabla 3.1) y las secuencias se filtraron por calidad como se indica en el punto 2.2. Para alinearlas al genoma de referencia se utilizó el programa Hisat2 2.1.0 (D. Kim et al., 2015) con parámetros por default y la opción "--no-softclip".

### 3.2.7 - Caracterización funcional mediante análisis de ontología génica

Para caracterizar los genes blanco de los *phasiARNs* detectados se realizó un análisis de clasificación funcional por Ontología Génica (GO) con la base de datos de Uniprot y luego se hizo un análisis de enriquecimiento de términos de ontología génica (*Gene Ontology Enrichment*) mediante el test de Fisher, utilizando el software TopGO 2.38.1 (Alexa & Rahnenführer, 2007) dentro del software Rstudio 1.1.456 (Allaire, 2012), con la versión 3.6.0 de R. Se analizaron en simultáneo los genes blanco correspondientes a todos los estadios de cada cultivar para las 3 categorías de ontología: función molecular (MF), componente celular (CC) y proceso biológico (BP). Se utilizó el algoritmo "classic" y se estableció como línea de corte un valor p ajustado (FDR)  $\leq 0,05$ .

### 3.2.8 - Validación de genes blanco con bibliotecas de degradoma

Los genes blanco de los miARNs analizados y de los phasiARNs detectados fueron validados bioinformáticamente con una biblioteca de degradoma de *Fragaria vesca* del cultivar *Yellow Wonder 5AF7* compuesta por un pool de aquenios y receptáculos de frutos de varios estadios (Tabla 3.1) y el software Cleaveland v4.4 (Addo-Quaye et al., 2009), conservando sólo los resultados con valor  $p \leq 0,1$  y tipo de degradoma 0 o 1. Los datos de secuenciación correspondientes a bibliotecas de degradoma fueron pre-procesados como se explica en el punto 2.2 (Pre-procesamiento de datos de NGS). Luego se convirtió el formato fastq a fasta con la herramienta FASTQ-to-FASTA v. 0.0.14 (Gordon, 2010) y las réplicas fueron concatenadas.

Se consideró como posibles blancos de miARNs y phasiARNs al conjunto de transcritos codificados en el genoma de *Fragaria x ananassa* y para esto se descargó la versión 1.2 del archivo de anotaciones correspondiente al genoma utilizado de la base de datos *Genome Database for Rosaceae* (<https://www.rosaceae.org/>). Además fue necesario reemplazar espacios y otros caracteres en la línea correspondiente al identificador de secuencia como se detalló en el Capítulo 2.

Tabla 3.1: Números de acceso de los datos de NGS utilizados

Estadio	Cultivar	Small RNA-seq	RNA-seq	Degradoma
Verde	Hongjia	SRR6995973 <sup>[1]</sup>	SRR6995967 <sup>[1]</sup>	-
		SRR6995974 <sup>[1]</sup>	SRR6995968 <sup>[1]</sup>	-
Intermedio	Hongjia	-	SRR6995969 <sup>[1]</sup>	-
		-	SRR6995970 <sup>[1]</sup>	-
Rojo	Hongjia	SRR6995975 <sup>[1]</sup>	SRR6995971 <sup>[1]</sup>	-
		SRR6995976 <sup>[1]</sup>	SRR6995972 <sup>[1]</sup>	-
Mezcla de estadios y tejidos	Yellow Wonder 5AF7	-	-	SRR1586429 <sup>[2]</sup>

## 3.3 - Resultados

### 3.3.1 - Análisis de microARNs

Al momento de iniciar este análisis existían publicaciones que reportaban análisis genómicos de ARNs pequeños en frutilla. Sin embargo, el enfoque llevado a cabo como parte de esta tesis aporta una perspectiva complementaria y específicamente orientada a analizar miARNs y phasiARNs durante el proceso de maduración de frutos de frutilla. Una caracterización similar realizada en plantas de la especie salvaje de frutillas, *Fragaria vesca* (Feng et al., 2019), utilizó bibliotecas de varios tejidos para el análisis de miARNs y phasiARNs, pero solo incluía frutos del estadio inmaduro. Otro grupo analizó los miARNs de *Fragaria x ananassa*, particularmente en relación con la senescencia de frutos, y dado que en ese momento no estaba aún disponible el genoma de esta especie, el análisis se realizó utilizando como referencia el genoma de *F. vesca* (X. Xu et al., 2013). Teniendo en cuenta estos antecedentes, realizamos este análisis para estudiar a nivel genómico la presencia y los cambios en la acumulación de miARNs durante el proceso de maduración de frutos y por otro lado, a identificar los *loci* generadores de phasiARNs durante este importante proceso, orientando el análisis especialmente a la identificación de genes regulados que contribuyan al desarrollo de propiedades organolépticas. En ambos casos se continuó con un análisis de identificación y validación de los transcritos regulados por estos ARNs pequeños, usando datos provenientes de secuenciación de bibliotecas de degradoma (Fig. 3.4). Teniendo en cuenta que al momento de iniciarse este análisis no existían estudios de ARNs pequeños en frutos de frutilla donde se use como referencia el nuevo genoma de *Fragaria x ananassa* disponible, estos resultados nos permitieron identificar por primera vez aquellos *loci PHAS* presentes en frutos de la variedad de frutilla comercial y analizar la presencia y acumulación de miARNs y phasiARNs durante la maduración de frutos (Fig. 3.4).

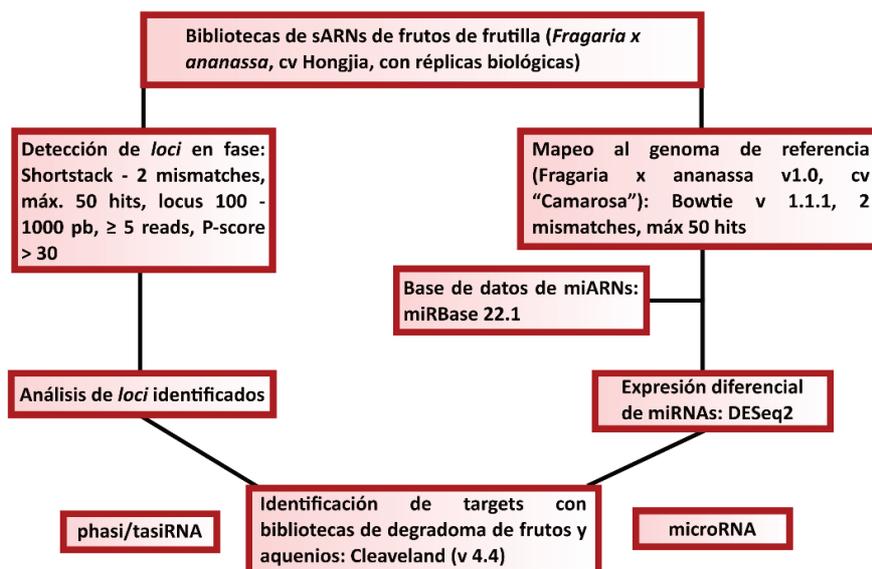


Figura 3.4: Procedimiento seguido para el procesamiento de los datos de secuenciación de ARNs pequeños de frutilla.

### 3.3.1.1 - Pre-procesado de los datos de secuenciación de *small RNA-seq*

En primer lugar se realizó el análisis de calidad de los datos de secuenciación correspondientes a 2 estadios de maduración (inmaduro y maduro, Tabla 3.1, Fig. 3.3) con dos réplicas biológicas cada uno. Los reportes del programa FastQC v 0.11.5 correspondientes a cada biblioteca, indicaron que la calidad por secuencia de cada biblioteca era superior a 30, el promedio de calidad por secuencia estaba entre 32 y 38 y además, el porcentaje de bases no asignadas (N) fue despreciable en todos los casos, por lo que se decidió continuar con el análisis sin realizar filtrados por calidad de los datos de secuenciación. Se identificó a la secuencia “AGATCGGA” como el inicio de la porción correspondiente a los adaptadores usados en la preparación de las bibliotecas de *small RNA-seq* y se eliminó de los datos de secuenciación.

En la Tabla 3.2 se puede observar que luego de eliminar adaptadores se retuvo más del 70% de las secuencias originales en las bibliotecas analizadas, y tras filtrar las secuencias contaminantes, siguiendo el procedimiento detallado en Materiales y Métodos, el porcentaje de *reads* limpias en todas las bibliotecas se mantuvo entre el 73,6% y el 78,1% de las secuencias sin adaptadores (Tabla 3.2). Luego de estos procedimientos, la cantidad de *reads* limpias remanentes para utilizar en los análisis posteriores es adecuada para el objetivo que se pretendía llevar a cabo (Fig. 3.4).

Tabla 3.2: Pre-procesado y mapeo de bibliotecas de *small RNA-seq* de frutilla

Muestra	Reads crudas	Reads tras eliminar adaptadores (%)	Reads limpias (%)*	Reads alineadas al genoma (%)**
Verde - R1	58522635	42689631 (73)	31651760 (74,1)	21101258 (66,7)
Verde - R2	61532075	43797529 (71,1)	34189436 (78,1)	20131523 (58,9)
Rojo - R1	57276783	41671898 (72,8)	30830367 (74)	21090814 (68,4)
Rojo - R2	55275383	40001945 (72,4)	29425584 (73,6)	20696931 (70,3)

\* Porcentaje respecto del número de *reads* sin adaptadores

\*\* Porcentaje respecto del número de *reads* limpias

### 3.3.1.2 - Expresión de miARNs durante la maduración de frutos de frutilla

Para estudiar a nivel genómico cuáles son los miARNs presentes durante la maduración de frutos de frutilla, se alinearon todas las secuencias de ARNs pequeños provenientes de datos de secuenciación de *small RNA-seq* de frutos de frutilla (*Fragaria x ananassa* cv. Hongjia) correspondientes a 2 estadios de maduración (inmaduro y maduro, Tabla 3.1) con dos réplicas biológicas cada uno a la versión 1.0 del genoma de referencia de *Fragaria x ananassa* cv. *Camarosa*, descargado de la base de datos de rosáceas *Genomic Database for Rosaceae* (<https://www.rosaceae.org/>, Tabla 3.2).

Debido a que la información de miRBase no contaba con un archivo de anotaciones de miARNs (.gff3) para frutilla, con el fin de anotar y cuantificar los miARNs, se alinearon las *reads* presentes en el genoma a la lista completa de miARNs de *Fragaria vesca* disponible en miRBase. Además, dado el interés particular en relación a la presente tesis doctoral, se añadieron al análisis las secuencias de miR858 y miR828 de 6 especies (*Arabidopsis thaliana*, *Malus domestica*, *Cucumis melo*, *Prunus persica*, *Picea abies* y *Camelina sativa*), ya que ambos miARNs estaban ausentes en miRBase, pero se ha reportado que estos miARNs se encuentran ampliamente conservados en plantas. Estos miARNs están involucrados en el metabolismo de antocianinas a través de la regulación de factores de transcripción de la familia MYB (Guan et al., 2014; Tirumalai et al., 2019; Xia et al., 2012). Para realizar la anotación y cuantificación de los miARNs

contenidos en la lista de miARNs confeccionada, se procedió como se detalla en el punto 3.2.3 de Materiales y Métodos.

A través de este análisis se determinó que de un total de 65 familias de miARNs, al menos un miembro de 28 de ellas se expresa en frutos durante el proceso de maduración (Fig. 3.5A). Dado que se usó la información de secuencias de miARNs de *F. vesca*, pero detectamos estos miARNs en la especie *Fragaria x ananassa*, de aquí en adelante nos referiremos a estos miARNs con el prefijo “*fan*”, ya que determinamos que también están presentes en esta especie. Posteriormente, para cada uno de los miARNs detectados en frutos, se realizó un análisis de identificación y validación de genes target usando bibliotecas de degradoma, que se resumen en la Fig. 3.5B-C (ver más adelante).

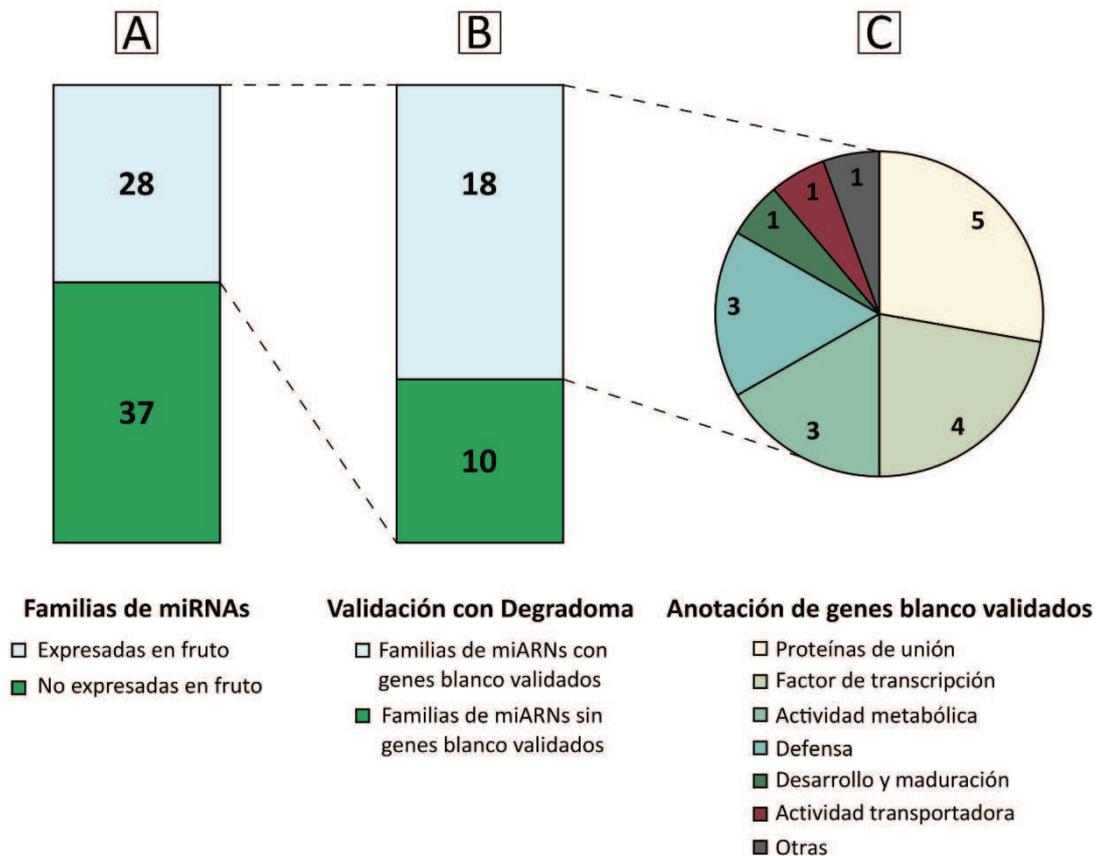


Fig. 3.5. Análisis de miARNs en frutilla. (A) Total de familias analizadas con alguno de sus miembros expresado en al menos uno de los estadios de maduración analizados o sin expresión. (B) Total de familias expresadas en frutos con o sin genes blanco validados. (C) Distribución de los procesos biológicos en los que participan los genes blanco de miARNs validados en frutos.

Se observan diferentes perfiles de expresión para los miembros de las 28 familias de miARNs expresadas en frutos, incluyendo miARNs expresados a niveles similares

entre los estadios verde (inmaduro) y rojo (maduro), y también otros con expresión diferencial entre estos dos estadios. Dentro de estos últimos se detectaron 14 miARNs diferencialmente expresados, que incluyen sólo 3 familias conservadas entre diversas especies: fan-miR159, fan-miR167 y fan-miR845, cuyos genes blanco son un factor de transcripción de la familia *MYB*, un factor de respuesta a auxinas, y retrotransposones específicos vinculados a la reproducción y al desarrollo de semillas, respectivamente (Fig. 3.6, Tabla 3.3; Borges et al., 2018; Csukasi et al., 2012; X. Xu et al., 2015). Dentro de la familia miR159, se detectó expresión diferencial de fan-miR159a y fan-miR159b. El primero muestra mayores niveles de expresión en el estadio rojo, mientras que el segundo manifiesta el comportamiento inverso (Fig. 3.5). Por su parte, se detectó a fan-miR167d en mayor medida en el estadio verde, mientras que fan-miR845 evidenció el comportamiento inverso (Fig. 3.5). En los 4 casos, se detectaron diferencias significativas entre los dos estadios estudiados.

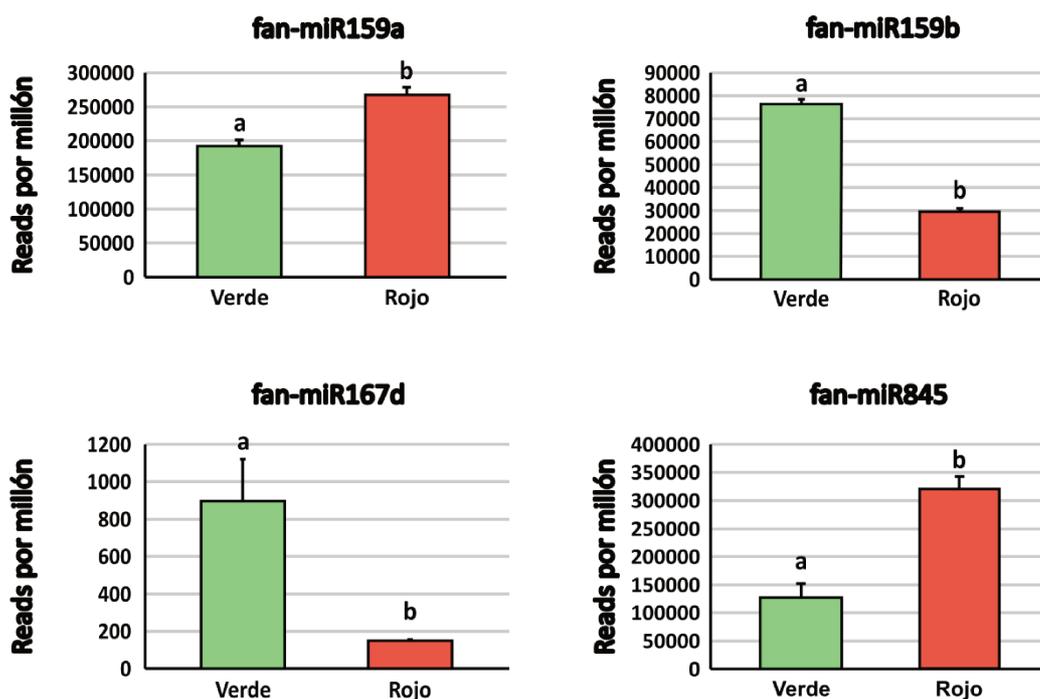


Fig. 3.6: Niveles de expresión de los miARNs fan-miR159a, fan-miR159b, fan-miR167d y fan-miR845 en frutos de frutilla. Las diferencias estadísticamente significativas se indican con letras sobre las barras de error.

Se determinó también la presencia de 3 miARNs de la familia de miR858 en las muestras de frutos analizadas (Fig. 3.7A). En la Fig. 3.7A se muestra un alineamiento de las secuencias de los miARN de la familia miR858 de *A. thaliana*, *P. persica*, *M. domestica*, *C. sativa*, *C. melo* y *P. abies*. Se detectó la presencia de miembros de la familia

de miR858 en *Fragaria x ananassa* con idéntica secuencia a los reportados para *M. domestica*, *P. persica* y miR858b de *A. thaliana*. Dado que los miARNs de la familia de miR858 no estaban reportados en la base de datos miRbase, nuestro análisis permite evidenciar su existencia en frutos de *Fragaria x ananassa* y procedimos a denominarlos de manera arbitraria fan-miR858a, fan-miR858b y fan-miR858c, respectivamente (Fig. 3.7A). Se ha demostrado en otras especies que esta familia de miARNs regula la expresión de factores de transcripción del tipo MYB, y nuestros resultados indican que la familia de miR858 y sus transcriptos regulados están también conservados en *Fragaria x ananassa* (Tabla 3.3; Sharma et al., 2016). El alineamiento de las secuencias de miR858 de las especies utilizadas en este análisis muestra el grado de conservación de este miARN, con un porcentaje de identidad del 80% entre ellos. El análisis de expresión en los estadios verde y rojo indican que fan-miR858a, fan-miR858b y fan-miR858c presentan niveles de acumulación sin diferencias significativas en ambos estadios de madurez (Fig. 3.7B).

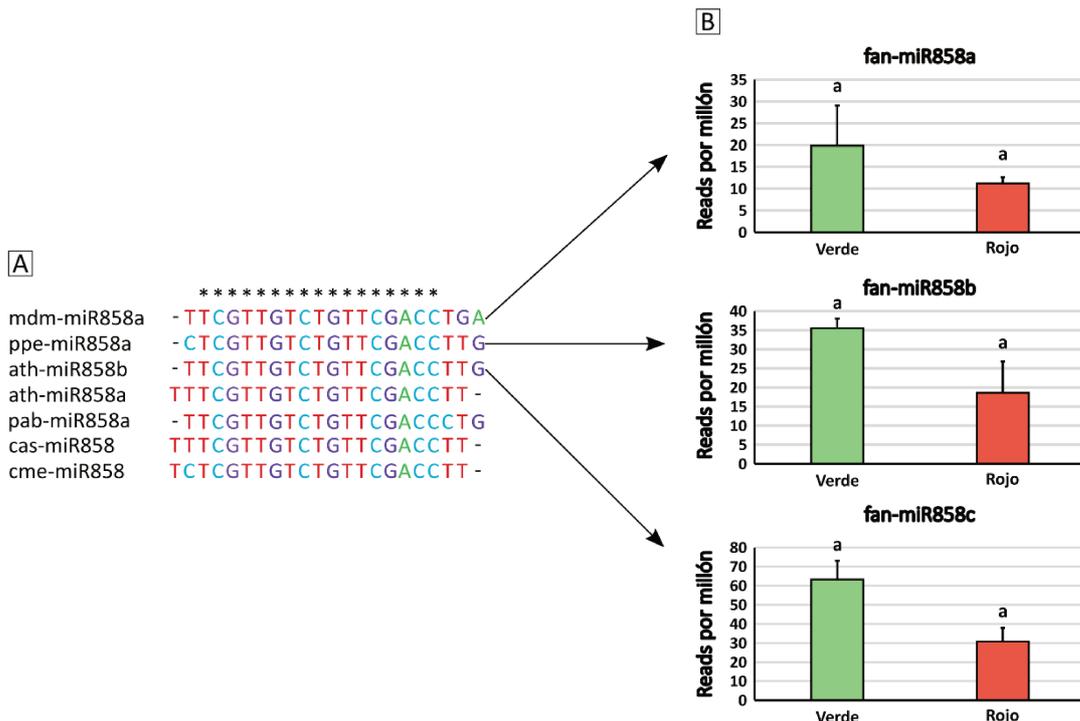


Fig. 3.7: (A) Alineamiento de las secuencias de miR858 utilizadas en este trabajo. (B) Perfil de expresión de fan-miR858a, fan-miR858b y fan-miR858c en frutos de distintos estadios de madurez.

También detectamos 11 miARNs no conservados con expresión diferencial entre los estadios inmaduro y maduro en *Fragaria x ananassa*. Hasta el momento, estos miARNs se habían reportado como específicos de *Fragaria vesca*, y a raíz de este estudio

denominamos fan-miR2109, fan-miR11285, fan-miR11286, fan-miR11299, fan-miR11288c, fan-miR11288e, fan-miR1511, fan-miR845, fan-miR11291, fan-miR11284 y fan-miR11312. Ninguno de estos miARNs regula la expresión de genes relacionados con la maduración de los frutos o con la adquisición de propiedades organolépticas, por lo que se decidió no profundizar su estudio (ver sección siguiente, Tabla 3.3).

### 3.3.1.3 - Identificación y validación de genes regulados por miARNs durante la maduración

Para estudiar los procesos biológicos regulados por miARNs durante la maduración de frutos de *Fragaria x ananassa*, se hizo un análisis de identificación y validación de los genes *target* usando el programa Cleaveland v4.4 (Addo-Quaye et al., 2009) y una biblioteca de degradoma correspondiente a un pool de ARN de aquenios y receptáculos de fruto inmaduro de la especie *Fragaria vesca* (Tabla 3.1). Se seleccionaron para análisis posteriores aquellos resultados cuyo *p-value* fue menor a 0,1 con gráficos *T-plot* de categoría 0 y 1, y además se realizó una inspección visual de todos los archivos *T-plots* generados como resultado de este análisis. Esto permitió conocer los procesos biológicos regulados por los miARNs presentes en frutos de frutilla, validándose los genes blanco de 18 familias de miARNs (Fig. 3.5B; Tabla 3.3).

De acuerdo a su anotación, se clasificaron los genes blanco según el proceso biológico del que participan (Fig. 3.5C, Tabla 3.3). La mayoría de los genes blanco regulados por miARNs en frutos corresponden a proteínas de unión y también se encontraron genes que codifican para factores de transcripción, proteínas relacionadas con la defensa de la planta, para proteínas con actividad metabólica, enzimas vinculadas al desarrollo y la maduración de frutos, y proteínas con actividad transportadora (Fig. 3.5C, Tabla 3.3). De acuerdo a los genes regulados por estos miARNs y en base a la bibliografía existente, dos de estas familias de miARNs se vinculan con distintos aspectos del desarrollo y la maduración de los frutos: fan-miR159 y fan-miR167 (Tabla 3.3), lo que los vuelve interesantes por su participación en la regulación de algunos puntos clave de estos procesos.

Tabla 3.3: Validación de genes blanco de familias de miARNs expresados en frutos de frutilla

	Familias de miARNs presentes en frutos	Expresión diferencial	Gen blanco
Factores de transcripción	fan-miR858	Sí	GAMYB-like protein
	fan-miR159	Sí	GAMYB-like protein
	fan-miR171	No	GRAS Transcription factor
	fan-miR396	No	Growth-Regulating Factor
Defensa	fan-miR482	No	Disease resistance protein (CC-NBS-LRR class)
	fan-miR2109	Sí	<i>NB-LRR</i> Disease resistance protein
	fan-miR11285	Sí	Pentatricopeptide repeat-containing protein
Actividad catalítica	fan-miR11299	Sí	Protein kinase 3
	fan-miR162	No	Gibberellin 2-beta-dioxygenase 2
	fan-miR168	No	Protein ARGONAUTE 1-like
Proteínas de unión	fan-miR11288	No	F-box/kelch-repeat protein
	fan-miR11293	No	MF: nucleic acid binding
	fan-miR11315	No	F-box protein CPR30-like
	fan-miR1511	Si	PHD finger protein MALE MEIOCYTE DEATH 1
	fan-miR394	No	F-box protein 6
Desarrollo y maduración	fan-miR167	Sí	Auxin Response Factor 6
Actividad transportadora	fan-miR11286	Sí	WAT1-related protein
Otras	fan-miR11289	No	Uncharacterized protein

Como se mencionó anteriormente, las familias de miARNs miR159 y miR858 regulan miembros de la familia de factores de transcripción MYB en diversas especies (Millar et al., 2019). Estos factores de transcripción están implicados en la formación de metabolitos secundarios muy importantes, que influyen en atributos organolépticos del fruto, como los flavonoides en la coloración o los fenilpropanoides relacionados con la adquisición de aroma, y su importancia en frutos de frutilla ha sido ampliamente

estudiada (Medina-Puche et al., 2015; H. Wang et al., 2020). Nuestros análisis de validación de genes target para fan-miR159 y fan-miR858 indican que también en esta especie estos miARNs regulan la acumulación de transcritos que codifican para factores de transcripción de la familia MYB (Tabla 3.3). Por su parte, nuestro análisis también identificó a la familia fan-miR167 como regulador de Factores de Respuesta a Auxinas (ARF, por sus siglas en inglés), genes implicados en la señalización de estas hormonas que son cruciales para el desarrollo de frutos de frutilla (Wang et al., 2019).

Luego analizamos la expresión de los genes de *Fragaria x ananassa* regulados por miARNs de interés para esta tesis. Para esto se utilizaron datos de secuenciación de *RNA-seq* del mismo conjunto de frutos (*Fragaria x ananassa* cv. Hongjia, Tabla 3.1), a los que se suma información de un estadio de madurez intermedio entre los analizados por *small RNA-seq*, los cuales fueron alineados al genoma de referencia de *Fragaria x ananassa*. En la Figura 3.8 se grafican los niveles de expresión durante la maduración de los genes blanco analizados, los cuales se nombran de acuerdo a reportes previos: *MYB-like protein J* (blanco de fan-miR159a), *MYB5-like* (blanco de fan-miR858a), *MY13* (blanco de fan-miR858b), *MYB9* (blanco de fan-miR858c) y *ARF6* (blanco de fan-miR167d). Los niveles de expresión de todos los genes analizados que codifican para factores de transcripción de la familia *MYB* se mantienen constantes durante la maduración, mientras que la expresión de *ARF6* es elevada y similar en los estadios verde e intermedio, pero significativamente menor en el estadio rojo (Fig. 3.8). Nuestros resultados indican que la regulación por parte de miR858 de factores de transcripción de la familia *MYB* se encuentra conservada también en frutos de frutilla, lo cual resulta de interés ya que esta regulación tendría incidencia en el desarrollo de color de los frutos.

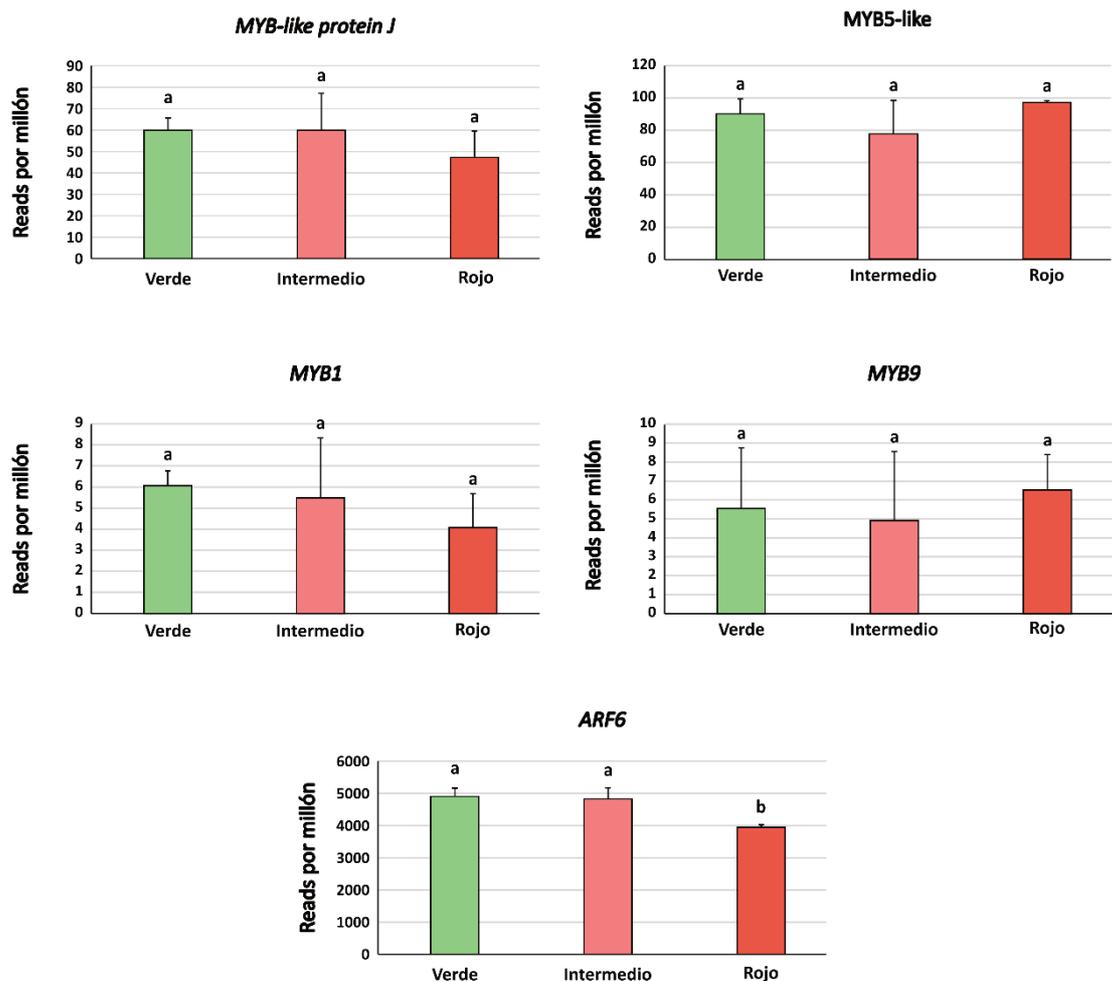


Fig. 3.8: Perfil de expresión de los genes MYB-like protein J, blanco de fan-miR159a; MYB5-like, blanco de fan-miR858a; MYB1, blanco de fan-miR858b; MYB9, blanco de fan-miR858c, y ARF6, blanco de fan-miR167d en frutos de distintos estadios de madurez.

### 3.3.2 - Análisis de descubrimiento y caracterización de loci generadores de phasiARNs

Con el fin de caracterizar a los phasiARNs presentes durante la maduración de frutos de frutilla, se realizó un análisis utilizando la herramienta Shortstack 3.8.3. Se detectaron 42 *loci* presentes en frutos de frutilla (Tabla 3.4) que no han sido descritos hasta el momento. Se analizaron en detalle aquellos *loci* que tenían un *P-score* mayor o igual a 30 y una longitud entre 100 y 1000 pb., de acuerdo a procedimientos similares reportados previamente (Das et al., 2020; Feng et al., 2019; Nakamura et al., 2019).

**Tabla 3.4: Descripción de los *loci* en fase detectados en frutos de *Fragaria x ananassa*.**

<b>Locus</b>	<b>Ubicación genómica</b>	<b>Modelo génico</b>	<b>Gen asociado</b>
<b>Locus 1</b>	Fvb1-1:12118468-12119350	Ninguno	-
<b>Locus 2</b>	Fvb1-2:18161316-18162285	maker-Fvb1-2-augustus-gene-181.39-mRNA-1	phosphatidate cytidyltransferase family protein
<b>Locus 3</b>	Fvb1-2:7676774-7677921	maker-Fvb1-2-augustus-gene-76.41-mRNA-1	proteasome alpha subunit D2
<b>Locus 4</b>	Fvb1-3:2312085-2312719	maker-Fvb1-3-snap-gene-23.72-mRNA-1	uncharacterized protein
<b>Locus 5</b>	Fvb1-3:3341713-3342717	maker-Fvb1-3-augustus-gene-33.50-mRNA-1	pentatricopeptide repeat-containing protein
<b>Locus 6</b>	Fvb1-3:5991289-5992259	maker-Fvb1-3-snap-gene-60.54	uncharacterized protein
<b>Locus 7</b>	Fvb1-4:3033958-3035198	Ninguno	-
<b>Locus 8</b>	Fvb1-4:4013480-4014095	maker-Fvb1-4-augustus-gene-40.35-mRNA-1	disease resistance protein (TIR-NBS-LRR class)
<b>Locus 9</b>	Fvb2-1:23411151-23411587	maker-Fvb2-1-augustus-gene-234.53-mRNA-1	E3 ubiquitin-protein ligase RNF126-like
<b>Locus 10</b>	Fvb2-2:10640025-10640504	Ninguno	-
<b>Locus 11</b>	Fvb2-2:22319519-22320740	maker-Fvb2-2-snap-gene-223.61-mRNA-1	RNAseH-like domain
<b>Locus 12</b>	Fvb2-2:788437-789300	maker-Fvb2-2-snap-gene-8.68	peptidyl-prolyl cis-trans isomerase FKBP62-like
<b>Locus 13</b>	Fvb2-3:11617174-11617729	maker-Fvb2-3-snap-gene-116.47-mRNA-1	U2 spliceosomal RNA
<b>Locus 14</b>	Fvb2-4:12280000-12280366	Ninguno	-
<b>Locus 15</b>	Fvb3-1:20782601-20783253	maker-Fvb3-1-augustus-gene-207.32	uncharacterized protein
<b>Locus 16</b>	Fvb3-1:26624382-26625362	maker-Fvb3-1-augustus-gene-266.30	DNA-damage-repair/toleration protein DRT111
<b>Locus 17</b>	Fvb3-2:4683514-4683935	maker-Fvb3-2-augustus-gene-46.28	Disease resistance protein (TIR-NBS-LRR class) family
<b>Locus 18</b>	Fvb3-2:5279709-5280685	maker-Fvb3-2-snap-gene-52.59	uncharacterized protein
<b>Locus 19</b>	Fvb3-3:27786692-27787045	Ninguno	-
<b>Locus 20</b>	Fvb3-3:9617540-9618137	maker-Fvb3-3-augustus-gene-96.32-mRNA-1	uncharacterized protein
<b>Locus 21</b>	Fvb3-3:9647392-9647937	maker-Fvb3-3-augustus-gene-96.35	uncharacterized protein
<b>Locus 22</b>	Fvb3-4:19476582-19477239	maker-Fvb3-4-augustus-gene-194.36	uncharacterized protein
<b>Locus 23</b>	Fvb3-4:19478425-19479175	maker-Fvb3-4-augustus-gene-194.33	uncharacterized protein
<b>Locus 24</b>	Fvb3-4:24414664-24415320	maker-Fvb3-4-augustus-gene-244.39-mRNA-1	disease resistance protein (TIR-NBS-LRR class)

<b>Locus 25</b>	Fvb4-3:13317103-13317605	maker-Fvb4-3-snap-gene-133.59-mRNA-1	Mono-/di-acylglycerol lipase
<b>Locus 26</b>	Fvb4-3:29853432-29853860	snap_masked-Fvb4-3-processed-gene-298.33	sodium-coupled neutral amino acid transporter 5-like
<b>Locus 27</b>	Fvb4-3:325935-326623	snap_masked-Fvb4-3-processed-gene-3.31-mRNA-1	uncharacterized protein
<b>Locus 28</b>	Fvb4-3:4271315-4272659	maker-Fvb4-3-augustus-gene-42.68	importin subunit alpha-1-like
<b>Locus 29</b>	Fvb4-4:531006-531993	Ninguno	-
<b>Locus 30</b>	Fvb5-1:4640051-4640887	snap_masked-Fvb5-1-processed-gene-46.29-mRNA-1	putative transmembrane protein
<b>Locus 31</b>	Fvb6-1:31946705-31948696	maker-Fvb6-1-augustus-gene-319.52-mRNA-1	uncharacterized protein
<b>Locus 32</b>	Fvb6-1:32301504-32302105	Ninguno	-
<b>Locus 33</b>	Fvb6-1:33070778-33071617	maker-Fvb6-1-snap-gene-330.76-mRNA-1	cytokinin oxidase 7
<b>Locus 34</b>	Fvb6-1:33075972-33076557	maker-Fvb6-1-augustus-gene-330.64	putative cytokinin dehydrogenase
<b>Locus 35</b>	Fvb6-2:30652833-30653548	maker-Fvb6-2-snap-gene-306.67-mRNA-1	Pentatricopeptide repeat (PPR) superfamily protein
<b>Locus 36</b>	Fvb6-3:3082842-3083715	Ninguno	-
<b>Locus 37</b>	Fvb6-3:38164181-38165164	maker-Fvb6-3-snap-gene-381.45	uncharacterized protein
<b>Locus 38</b>	Fvb6-3:41276618-41277553	maker-Fvb6-3-augustus-gene-412.63-mRNA-1	disease resistance protein (TIR-NBS-LRR class)
<b>Locus 39</b>	Fvb6-3:41397094-41397864	maker-Fvb6-3-snap-gene-414.61	uncharacterized protein
<b>Locus 40</b>	Fvb6-3:765650-766309	maker-Fvb6-3-augustus-gene-7.32-mRNA-1	auxin signaling F-box 3
<b>Locus 41</b>	Fvb6-4:33571488-33572266	maker-Fvb6-4-augustus-gene-335.71-mRNA-1	disease resistance protein RGA3 (LOC101302580), mRNA
<b>Locus 42</b>	Fvb7-3:16989118-16989848	Ninguno	-

De los 42 *loci* detectados, 29 coinciden en su ubicación genómica con genes que codifican para proteínas, anotados en la versión 1.0 del genoma de referencia y 13 no comparten ubicación con ningún gen, por lo que podrían considerarse *loci* TAS (Tabla 3.4, resaltados en rojo). Se vio que 12 genes están anotados como “proteína no caracterizada” y a pesar de realizar búsquedas por homología utilizando BLAST en diversas bases de datos, no se logró asociar una posible función para estos genes.

### 3.3.3 - Caracterización de transcritos regulados por ARNs pequeños

Se extrajo la secuencia de cada uno de los ARNs pequeños detectados por Shortstack tras el alineamiento, y de forma análoga a lo realizado para el análisis de

muestras de tomate, se les asignó un nombre de acuerdo al *locus* de origen y a la hebra a partir de la cual cada siARN se produce: S para hebra “sentido” y AS para hebra “anti sentido”. De esta forma, los nombres tienen el formato “*Locus\_hebra\_orden*”, por ejemplo: *L1\_S\_1*.

A cada una de las secuencias correspondientes a un ARN pequeño, predichas en el paso anterior, se las sometió a un análisis de búsqueda y validación de *targets* usando datos de degradoma de *F. vesca* (Tabla 3.1) y el software Cleaveland 4.4. Se detectaron 142 genes blanco regulados por phasiARNs provenientes de frutos en estadio inmaduro y 134 regulados por ARNs pequeños derivados de frutos en estadio maduro, de los cuales 82 transcriptos fueron validados como genes blanco de phasiARNs presentes en ambos estadios de maduración.

#### 3.3.3.1 - Caracterización de transcriptos target en base a etiquetas de ontología génica

Para caracterizar los procesos biológicos en los que participan los genes validados como targets de los phasiARN detectados se realizó un análisis de clasificación funcional por Ontología Génica (GO), obteniendo estas etiquetas para cada gen a partir de la base de datos de Uniprot (Bateman et al., 2021). Luego se hizo un análisis de enriquecimiento de términos de ontología génica (*Gene Ontology Enrichment*) utilizando el test de Fisher, dentro del software TopGO (Alexa & Rahnenführer, 2007). Se detectaron 22 términos enriquecidos (FDR < 0.05), según la categoría Procesos Biológicos (BP) y 7 según la categoría Funciones Moleculares (MF) (Fig. 3.9). Dentro de BP los principales términos enriquecidos fueron respuesta celular a estímulos, transducción de señales, señalización, comunicación celular y respuesta a estímulos. En cuanto a MF, los 7 términos mencionados están relacionados a la actividad de unión de nucleótidos y nucleósidos. En la Figura 3.9 se grafican los términos de GO enriquecidos, resaltándose en rojo para cada categoría aquellos que presentaron un FDR < 0.05.

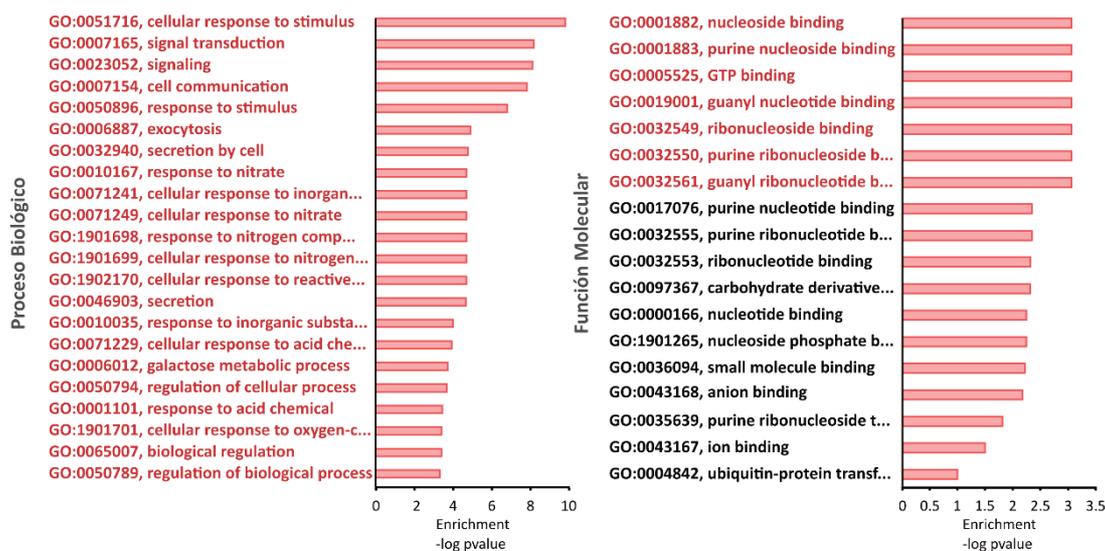


Fig. 3.9: Enriquecimiento de términos de ontología génica para las categorías “Procesos Biológicos”, “Función Molecular”.

Al analizar en detalle los genes incluidos en los términos enriquecidos significativamente para cada etiqueta GO, se vio que en la categoría MF, los genes codifican para proteínas GTPasas de la superfamilia Ras, implicadas en el transporte intracelular de proteínas y vesículas (Rutherford & Moore, 2002; Yun et al., 2019) y para la categoría BP se encontraron genes principalmente vinculados a la resistencia a estrés biótico y factores de transcripción con dominio NAC, típicamente involucrados en la respuesta a distintos tipos de estrés abiótico (Nakashima et al., 2012).

### 3.3.3.2 - Validación de genes blanco de interés

Con el fin de identificar transcritos validados como *targets* de phasiARNs que estén relacionados con la maduración de frutos, con el desarrollo de propiedades organolépticas de los mismos o con otros atributos de interés comercial, se realizó un análisis visual de los gráficos T-plot generados por el programa Cleaveland, mencionados en el apartado 3.3. Además, para completar el análisis, se procedió a comparar los niveles de expresión de los genes blanco validados utilizando las bibliotecas de *RNA-seq* mencionadas anteriormente (Tabla 3.1), usando el paquete estadístico DESeq2.

A partir de este análisis se identificaron 3 genes que codifican para proteínas con funciones biológicas interesantes para los objetivos propuestos en esta tesis doctoral. En la Fig. 3.10A se muestra la validación del corte de cada transcripto usando bibliotecas de degradoma. Se observa que en todos los casos los picos de los gráficos *T-plot* fueron

catalogados como “categoría 0”, indicando que en las bibliotecas de degradoma existe una mayor acumulación de *reads* que coinciden con los sitios de corte predichos para cada phasiARN, en comparación a la acumulación de *reads* que mapean en otras regiones de estos transcritos. Los niveles de expresión de cada gen durante la maduración se representan en la Fig. 3.10B y los niveles de expresión del correspondiente phasiARN se grafican en la Fig. 3.10C.

La enzima Xiloglucano-endo-transglicosidasa/hidrolasa 6 (Fa-XET6) está implicada en el metabolismo del xiloglucano, un componente muy abundante de la pared celular vegetal. La acción de esta y otras enzimas modifican la solubilización y despolimerización de componentes de la pared celular, alterando así la firmeza del fruto (Castro et al., 2021). Se detectó un incremento significativo en los niveles de expresión para los estadios intermedio y maduro respecto al estadio inicial (Fig. 3.10B), mientras que los niveles de L12\_S\_5, el phasiARN implicado en su regulación, se mantienen constantes en los dos estadios de maduración analizados (Fig. 3.10C).

Se encontró en este análisis que el gen codificante para una enzima Endoglucano-1,3-beta-glucosidasa (Fa- $\beta$ -1,3-glucanasa), relacionada con la defensa contra patógenos debido a su capacidad de degradar la pared de ciertos hongos (Beffa et al., 1993) y en menor medida con el ablandamiento de frutos (Choudhury et al., 2009; Peumans et al., 2000), es regulado por uno de estos phasiARNs detectados. Sus niveles de expresión son bajos para los estadios inmaduro e intermedio, pero aumentan significativamente hacia el final de la maduración (Fig. 3.10B). Por su parte, L31\_AS\_30 mantiene su expresión constante y elevada en ambos estadios (Fig. 3.10C), por lo que en conjunto los niveles de expresión de L31\_AS\_30 y de su transcripto regulado, se correlacionan con lo esperado para este tipo de regulación a nivel post-transcripcional.

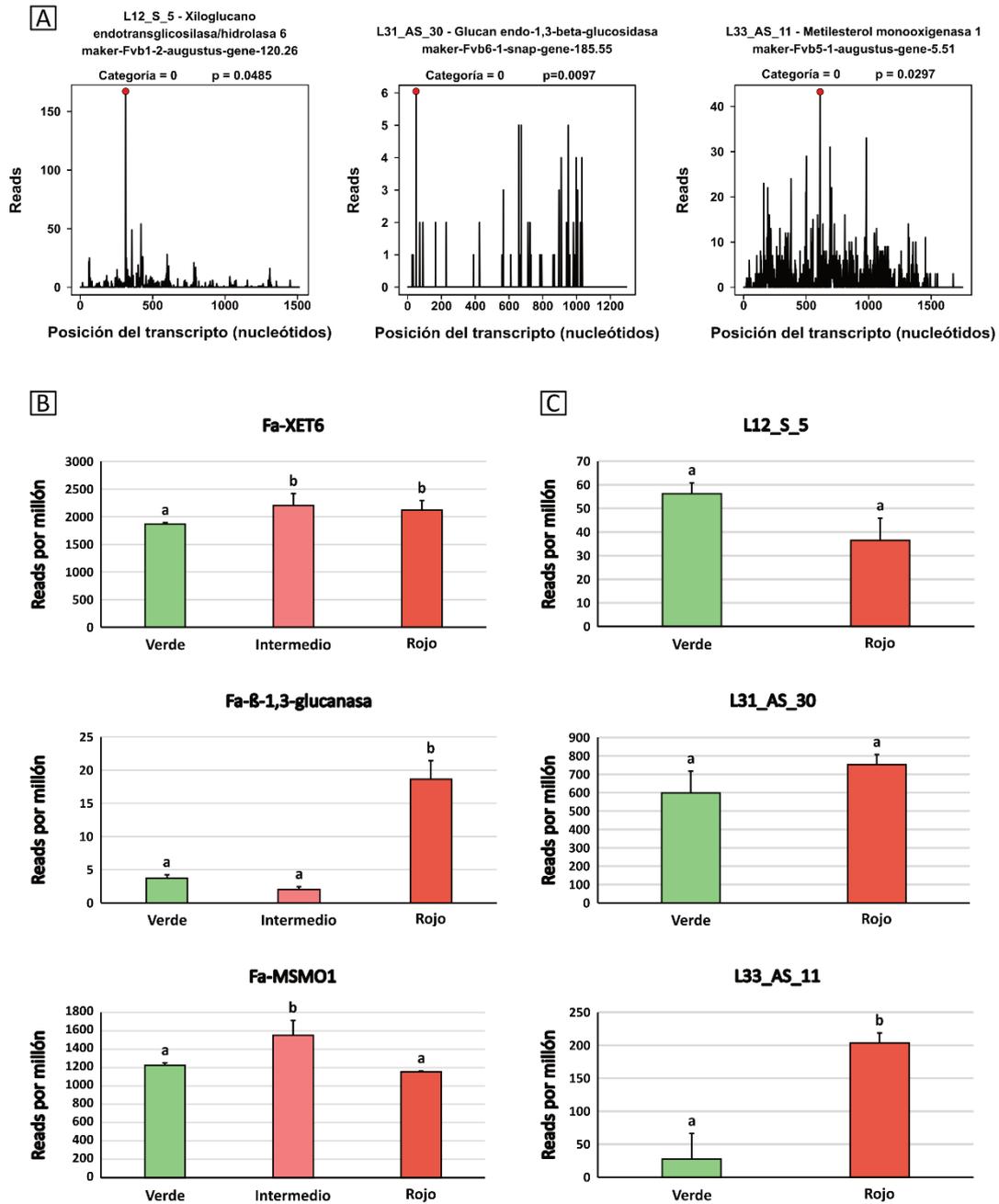


Fig. 3.10: (A) Validación del corte de los transcritos xiloglucano endotransglucosilasa/hidrolasa 6, glucano endo-1, 3-β-glucosidasa y metilesterol monooxigenasa 1 mediados por L12\_S\_5, L31\_AS\_30 y L33\_AS\_11, respectivamente. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcripto en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según Cleaveland. (B) Perfiles de expresión de los transcritos Fa-XET6, Fa-β-1,3-glucanasa y Fa-MSMO1 en frutos de distintos estadios de madurez. (C) Perfil de expresión de los phasiARNs L12\_S\_5, L31\_AS\_30 y L33\_AS\_11 en frutos de distintos estadios de madurez.

Por último, el gen que codifica para la proteína Metilesterol Monooxigenasa 1 (Fa-MSMO1) también fue validado como target del phasiARN L33\_AS\_11. La enzima Fa-MSMO1 está involucrada en la síntesis de esteroides, que a su vez están asociados a la reducción de niveles de colesterol en humanos (Valitova et al., 2016), lo que resulta importante para los consumidores y constituye otro aspecto interesante para continuar estudiando la regulación de su expresión a nivel post-transcripcional. Su expresión es mayor en el estadio intermedio con respecto al inmaduro y al maduro (Fig. 3.10B), mientras que L33\_AS\_11 evidencia niveles de expresión significativamente más altos en frutos maduros (Fig. 3.10C).

## 3.4 - Discusión

### 3.4.1 - En frutos de frutilla se expresan 28 familias de miARNs

Los datos analizados correspondientes a 2 estadios diferentes de maduración de frutos de frutilla evidencian la expresión de miembros de 28 familias de miARNs de frutilla. Si bien existía un reporte previo donde se estudiaron miARNs en frutos de *Fragaria x ananassa* solo en estadio de maduración rosa (He Li et al., 2013), nuestro análisis aporta un enfoque más amplio que permite verificar la conservación entre las especies *Fragaria vesca* y *Fragaria x ananassa* de 17 familias de miARNs ampliamente conservados entre especies y de 11 familias que serían específicas estas dos especies del género *Fragaria*.

Se validaron genes blanco para miembros de 18 familias de miARNs expresadas en frutos, a pesar de que se detectó la existencia de miembros de 28 familias de miARNs (Fig. 3.5). Es importante destacar que el análisis usando el software Cleaveland permite detectar regulación mediada por ARNs pequeños que actúan a través del corte de sus transcritos blanco, pero no permite la validación de regulación de la expresión por inhibición de la traducción, lo que podría explicar estos resultados. Sin embargo, otros escenarios son posibles y es necesario realizar análisis más profundos para establecer los motivos por los cuales existe acumulación de ciertos miARNs, pero no se detectan cortes en los genes blanco predichos. Este análisis estableció que la mayoría de los miARNs regula la expresión de factores de transcripción, proteínas de unión, con actividad metabólica y de defensa (Fig. 3.5B y 3.5C, Tabla 3.3). Liang y colaboradores

estudiaron la relación entre miARNs de frutos sanos y frutos infectados con *Botrytis cinerea*, y si bien no validaron los genes blanco con bibliotecas de degradoma, sus predicciones arrojan como resultado que los principales blancos de miARNs serían factores de transcripción y otros genes con actividades defensivas, en concordancia con lo evidenciado en este trabajo de tesis (Liang et al., 2018). Otro trabajo estudió los miARNs en múltiples tejidos de *Fragaria vesca* sin hacer hincapié en frutos (Feng et al., 2019). En este caso, sí se validaron los genes blanco con bibliotecas de degradoma de diversos tejidos, y éstos coinciden con los reportados en esta tesis doctoral para frutos de *Fragaria x ananassa*, aunque no detectaron la presencia de ningún miembro de las familias fan-miR858, fan-11299, fan-1511, fan-11286 y fan-11289 ni validaron sus genes regulados (Tabla 3.3). En conjunto, los estudios anteriores y el nuestro permiten establecer que existe conservación de varios módulos regulatorios que involucran miARNs entre la especie comercial de frutilla y la especie salvaje, aunque también existen diferencias entre los miARNs detectados dentro de la especie comercial, que posiblemente radican en la utilización de cultivares diferentes.

### 3.4.2 - miARNs reguladores del desarrollo de propiedades organolépticas de frutos de frutilla

En relación al interés central de esta tesis doctoral, incluimos en el análisis a dos miARNs, miR828 y miR858, que por su relación con la regulación de factores de transcripción de la familia MYB, resultan candidatos atractivos por su participación en la biosíntesis de antocianinas responsables del color característico de los frutos de frutilla (Tirumalai et al., 2019; F. Yang et al., 2013; B. Zhang et al., 2020). Nuestro análisis identificó la presencia de 3 miembros de la familia de miR858 durante la maduración de frutos (Fig. 3.5, Fig. 3.7B) y se validó la regulación por corte de los transcritos de 4 genes de la familia MYB (Fig. 3.8). Hasta el momento no existen reportes que hayan estudiado en detalle la influencia de este módulo regulatorio en el desarrollo de color en frutillas, lo cual destaca un área de sumo interés para profundizar, dado que uno de los atractivos centrales para los consumidores de los frutos de esta especie, es la presencia de un color rojo uniforme e intenso al momento de la comercialización (Bhat et al., 2015).

Nuestro análisis no detectó la presencia de miR828 durante la maduración de frutilla, el cual es responsable del corte inicial en transcritos TAS4, que luego dan lugar

a tasiRNAs que regulan también la expresión de factores de transcripción de la familia MYB (Rock, 2013). Sin embargo, además de miR858, identificamos la presencia de miembros de otras dos familias que regularían la expresión de genes que codifican para proteínas implicadas en el desarrollo de las propiedades organolépticas de frutos: fan-miR159 y fan-miR167. De manera análoga a lo expresado más arriba sobre el papel de miR858 en la regulación de la biosíntesis de antocianinas, la regulación de factores de transcripción MYB por parte de miR159 resulta también de interés para esta tesis doctoral. A diferencia de miR858, para el módulo regulatorio miR159-MYB, existe un trabajo previo donde se estudió la relación entre fan-miR159a y fan-miR159b con factores de transcripción MYB durante el desarrollo del receptáculo y la maduración del fruto (Csukasi et al., 2012). Llamativamente, Los niveles de expresión reportados no coinciden con los identificados en esta tesis, donde fan-miR159a disminuye con el transcurso de la maduración y fan-miR159b presenta el comportamiento inverso (Fig. 3.6). Además, estos autores reportan mayor expresión de estos miARNs en los receptáculos que en los aquenios. Como gen blanco identificaron un factor de transcripción de la familia GAMYB y sus niveles de expresión son inversos a los de sus miARNs reguladores.

En este trabajo de tesis utilizamos bibliotecas de degradoma y el software Cleaveland para identificar los genes blanco de cada miARN. Este programa reporta generalmente varios transcritos factibles de ser regulados por cada miARN, por lo que se debe discernir manualmente cuáles son los mejores resultados. Luego de este filtrado, conservamos 4 genes MYB regulados por las familias fan-miR159a y fan-miR858, que mantienen su expresión a niveles relativamente constantes durante la maduración (Fig. 3.8). La familia de factores de transcripción MYB es muy numerosa, está presente en todos los eucariotas y sus funciones son muy diversas, entre las que se pueden destacar desarrollo celular, respuesta a hormonas y regulación de varios procesos metabólicos (Ambawat et al., 2013). Debido a esto, se vuelve muy complejo el estudio de regulaciones post-transcripcionales de esta familia, por lo que posiblemente los niveles de acumulación observados para los módulos miR858/MYB en conjunto, reflejen también otros tipos de regulación además de la regulación por corte de transcripto mediada por miR858.

Se ha vinculado a los factores de transcripción del tipo MYB con la adquisición del color rojo característico de los frutos maduros, a través de su papel regulatorio en la biosíntesis de flavonoides (H. Wang et al., 2020). Además, las proteínas MYB participan en la producción de compuestos volátiles derivados de la vía de fenilpropanoides, implicados en el establecimiento del sabor y aroma de los frutos (Medina-Puche et al., 2015). En particular, se vio que el gen de frutilla *FaMYB1* regula negativamente la acumulación de antocianinas y flavonoles, y su sobreexpresión produjo decoloración en flores de *Nicotiana tabacum* (Aharoni et al., 2001). Llamativamente, en nuestro análisis no se detectaron diferencias significativas en los niveles de expresión en los distintos estadios de maduración, mientras que Aharoni y col. 2001 observaron un aumento de expresión en frutos maduros. *FaMYB9* está implicado en la biosíntesis de compuestos volátiles de 6 carbonos: su silenciamiento retrasa el desarrollo y la maduración del fruto, mientras que su sobreexpresión los acelera e incrementa la cantidad de sustancias aromáticas (Lu et al., 2020). Resulta interesante resaltar que, al igual que para *FaMYB1*, no se detectó expresión diferencial en los diferentes estadios, a pesar de que Lu y col. 2020 detectaron menores niveles de transcripto en estadios maduros. Por su parte, *FaMYB5* está también asociado a la pigmentación (Y. Luo et al., 2020) y se reportó que un factor *MYB5-like* en granada (*Punica granatum*) participa activamente en los primeros pasos de la producción de flavonoides en frutos inmaduros (Arlotta et al., 2020). En esta tesis detectamos expresión de 4 genes MYB validados como blancos de 2 familias de miARNs: fan-miR159 y fan-miR858. Estos genes son *FaMYB1*, *FaMYB5-like*, *FaMYB9* y *FaMYB-like protein j*. Se destaca que la familia miR858 no había sido descrita hasta el momento en frutilla, por lo que estos microARNs y sus transcriptos blanco son candidatos atractivos para seguir estudiando en detalle y posiblemente mejorar atributos comerciales como color y aroma en los frutos de esta especie.

La familia fan-miR167 ya ha sido asociada con la regulación de los genes ARF (Feng et al., 2019; J. Han et al., 2014; X. Xu et al., 2013), lo que coincide con lo detectado en este trabajo (Fig. 3.6, Fig. 3.8, Tabla 3.3). Las auxinas son cruciales para el desarrollo y la maduración de los frutos de frutilla, donde son producidas por los aquenios y provocan el incremento de tamaño del receptáculo en estadios tempranos; además, se sabe que si los aquenios son removidos, el receptáculo deja de crecer (Nitsch, 1950). Se

conocen bien las vías de señalización de auxinas a nivel post-transcripcional (Weijers & Wagner, 2016) y su regulación por parte de distintos tipos de ARNs pequeños: miARNs como miR167 y miR160, y tasiARNs derivados del locus *TAS3* denominados tasiARFs (S. B. Li et al., 2016). Particularmente en frutilla, se ha vinculado a *FaARF6* al incremento de tamaño del receptáculo y además se vio un aumento considerable de su expresión en este tejido maduro (Estrada-Johnson et al., 2017). En esta tesis se observó el comportamiento contrario, con una disminución significativa en la expresión de este gen hacia el final de la maduración (Fig. 3.8), efecto que puede deberse a que la biblioteca no está constituida solamente por datos de secuenciación del receptáculo, sino que también incluía aquenios. Sorprendentemente, nuestro análisis determinó que existen altos niveles de expresión de fan-miR167d en el estadio inmaduro y que la expresión disminuye conforme progresa la maduración, en lugar de la relación inversamente proporcional esperada (Fig. 3.6). Esto podría estar relacionado justamente a la importancia de las auxinas en las etapas iniciales de este proceso, que necesitaría de una regulación mucho más precisa y explicaría la elevada cantidad de fan-miR167d en ese momento.

### 3.4.3 - Se identificaron 42 loci productores de phasiARNs en frutos de frutilla

Dos estudios previos analizaron datos de secuenciación de *Fragaria vesca* correspondientes a diferentes tejidos, y sus autores reportaron la detección de *loci* PHAS en esta especie (Feng et al., 2019; Xia et al., 2015). En esta tesis, caracterizamos exclusivamente los *loci* PHAS de frutos de la especie comercial de frutilla, donde vimos que 29 de los 42 *loci* detectados superponen su ubicación en el genoma con genes codificantes anotados en la versión 1.0 del genoma de *Fragaria x ananassa*, mientras que 13 de ellos no comparten ubicación con ningún gen, por lo que podrían considerarse como *loci* TAS (Tabla 3.4, donde los *loci* TAS están resaltados en rojo). Además, 12 genes están anotados como “proteína no caracterizada”, algo que puede estar relacionado con que la secuenciación de este genoma es muy reciente o con el mecanismo seguido por los investigadores para realizar la anotación funcional de las proteínas, tal vez limitado a comparación con pocas especies.

A diferencia de los análisis de *loci* PHAS realizados en *F. vesca* donde se identificó que los phasiARNs estudiados participan en la respuesta a estrés (Feng et al., 2019) y en la regulación de genes de la familia F-BOX (Xia et al., 2015), en el análisis de enriquecimiento de términos de ontología génica realizado en este trabajo de tesis, observamos que los genes blanco de los phasiARNs detectados durante la maduración participan en diversos procesos fisiológicos, sin identificarse una única función principal (Fig. 3.9). Al igual que en estudios previos, se detectaron genes vinculados a la resistencia a estrés biótico y abiótico (Feng et al., 2019), por lo que sería interesante analizar en el futuro este tema, ya que es un aspecto muy importante para el desarrollo del cultivo y la comercialización de sus frutos.

#### **3.4.4 - Se validaron genes regulados por phasiARNs con potencial importancia comercial**

De los 194 genes blanco validados entre ambos estadios de maduración de frutos, sobresalen 3 por su relación con atributos de importancia comercial: Xiloglucano Endotransglucosidasa/hidrolasa (Fa-XET6), Metilesterol Monooxigenasa 1 (Fa-MSMO1) y Endoglucano-1,3-beta-Glucosidasa (Fa- $\beta$ -1,3-glucanasa), (Fig. 3.10). Durante la maduración de los frutos se producen una serie de modificaciones bioquímicas altamente reguladas. Una de ellas es el ablandamiento, producto del desensamblaje de la pared celular (Brummell, 2006), provocado por una serie de reacciones enzimáticas catalizadas por poligalacturonasas, xiloglucanasas, expansinas, pectato liasas y  $\beta$ -galactosidasas (Brummell & Harpster, 2001; Cosgrove, 1998; Y. Han et al., 2015; Marín-Rodríguez et al., 2002; Paniagua et al., 2016). De acuerdo a lo observado en la Fig. 3.10A, la enzima Fa-XET6 sería blanco de regulación de uno de los phasiARN detectados. Los resultados de un estudio realizado en kiwi y manzana sugieren que las XETs participan activamente en la regulación de la integridad de la pared celular, demostrando un posible rol en el ablandamiento de los frutos luego de cosechados. (Atkinson et al., 2009). Los datos transcriptómicos indican que su expresión es mayor en estadios tardíos, lo que permite suponer una mayor actividad en frutos maduros (Fig. 3.10B). Estas cualidades llevan a pensar en posibles aplicaciones comerciales mediante modificaciones biotecnológicas, teniendo como objetivo principal aumentar la vida

poscosecha de los cultivos. Para ello se necesitaría profundizar en estudios de la regulación de esta proteína *in vivo*.

Dejando de lado a las propiedades organolépticas de los frutos, se destaca que se detectaron 2 phasiARNs que tienen como blanco a un gen relacionado con la síntesis de esteroides (Metilesterol Monooxigenasa 1-1) y otro con la defensa frente a patógenos (la isoforma ácida de la Glucano endo-1,3-beta-glucosidasa). Esto es importante debido a que se trata de 2 aspectos cruciales desde el punto de vista comercial. Se ha reportado que la ingesta de fitoesteroides en la dieta favorece la disminución del nivel de LDL (Katan et al., 2003), lo que vuelve a estos compuestos atractivos para la obtención de cultivares mejorados con propiedades benéficas para la salud. De forma similar, la producción de variedades resistentes a plagas es siempre deseable, pues mejor capacidad defensiva implica disminución en las pérdidas de cosecha. Se estima que las enfermedades vegetales provocadas por hongos pueden llegar a generar pérdidas de alrededor de 200.000 millones de dólares por año a nivel mundial (Shuping & Eloff, 2017). Las enzimas del tipo 1,3-beta-glucanasas tienen la capacidad de degradar calosa, un polisacárido presente en la pared celular vegetal y fúngica. Por ello es que estas proteínas son enmarcadas en el grupo de “proteínas relacionadas a patogénesis” o PR (por sus siglas en inglés, *pathogenesis-related*), a pesar de que también están involucradas en procesos como dormancia y germinación de semillas (Leubner-Metzger, 2003), y se ha sugerido su participación en el ablandamiento de frutos durante la maduración (Choudhury et al., 2009; Peumans et al., 2000). Existen reportes sobre la efectividad antifúngica de esta familia de enzimas (D. Singh et al., 2014) como también de su actividad antibacteriana (S. C. Chen et al., 2006), aunque por el momento no hay trabajos al respecto en frutilla. Conocer más acerca de su regulación mediada por siARNs sería un paso inicial para la futura obtención de variedades mejoradas genéticamente.

### 3.5 - Conclusiones

- Se detectó expresión de 28 familias de miARNs en frutos de *Fragaria x ananassa*, y se validaron genes blanco para miembros de 18 familias.

- Se encontraron 3 miARNs de la familia de miR858 expresados en frutos, no reportados previamente en esta especie. Fueron llamados fan-miR858a, fan-miR858b y fan-miR858c.
- Se identificaron 42 *loci* productores de ARNs pequeños en fase, de los cuales 29 corresponderían a *loci* PHAS y 13 a *loci* TAS.
- Dentro de las etiquetas GO enriquecidas se detectaron genes de defensa frente a diversos tipos de estrés y genes relacionados con procesos metabólicos diversos.
- Se validaron genes blanco de phasiARNs con importancia comercial: uno relacionado a la firmeza de los frutos, Fa-XET6; uno relacionado con el contenido nutricional, Fa-MSMO1 y otro relacionado con la respuesta a patógenos (Fa- $\beta$ -1,3-glucanasa).

Capítulo 4: ARNs pequeños en  
naranja

## 4.1 - Introducción

### 4.1.1 - Género Citrus

El género *Citrus*, perteneciente a la familia Rutaceae, está compuesto por un elevado número de especies de gran importancia comercial, entre ellas limón (*Citrus limon*), lima (*Citrus × aurantiifolia*), mandarina (*Citrus reticulata*) y naranja (*Citrus sinensis*). Se cree que los cítricos son nativos del sudeste de Asia (Gmitter & Hu, 1990) y que se cultivan desde hace aproximadamente 4000 años (Q. Xu et al., 2013). Existen varios sistemas de clasificación de especies de cítricos pero los dos más utilizados son el de Swingle (Swingle, 1967) y el de Tanaka (Tanaka, 1954), basados en características morfológicas y bioquímicas. El primero reconoce 16 especies mientras que el segundo, 162.

El naranjo es un árbol perenne de porte mediano, que puede medir entre 5 y 15 metros de altura. Posee hojas de forma oval, flores blancas muy fragantes y su fruto es la naranja (Daurelio, 2009). Esta especie es un híbrido producto de la cruce del pomelo (*Citrus grandis*) y la mandarina (*Citrus reticulata*) (Nicolosi et al., 2000; Q. Xu et al., 2013). Se conocen y comercializan diversas variedades de naranjas en todo el mundo, que varían principalmente en su época de cosecha, tamaño, acidez y porcentaje de jugo. Estos cultivares se clasifican principalmente como naranjas redondas (*Valencia, Hamlin*), navelinas (vulgarmente conocidas como “naranjas de ombligo”), naranjas de pulpa roja (*sanguina, tarocco*) y naranjas dulces (Kimball, 1999).

### 4.1.2 - Características generales

#### 4.1.2.1 - Aspectos nutricionales, maduración y propiedades organolépticas

Los cítricos, y en particular las naranjas, son excelentes fuentes de vitamina C de fácil acceso (G. A. Wu et al., 2014). Sus principales hidratos de carbono son azúcares simples, como glucosa, fructosa y sacarosa, que a su vez contribuyen a su sabor dulce. También contienen fibras, principalmente pectina, lignina, hemicelulosa y celulosa, beneficiosas para el sistema digestivo humano (Anderson et al., 2009). Además son ricas en potasio, cuya ingesta está asociada a la disminución de la presión arterial y del riesgo de sufrir problemas coronarios (D'Elia et al., 2011).

Conforme progresa la maduración, se observan marcados cambios en el tamaño y la coloración de los frutos. Suele establecerse una división entre estadios inmaduros, donde los frutos son verdes y su tamaño aumenta paulatinamente, hasta el estadio rompiente, momento en el que el fruto alcanza su tamaño final y comienza a aparecer el color naranja (Fig. 4.1). Posteriormente, en los estadios maduros, se obtiene el tono final y se desarrollan los atributos sensoriales característicos (J. Wang et al., 2016). Durante su desarrollo, las naranjas sufren división celular, acumulación de agua, expansión, cambios en su pared celular y modificación de ácidos orgánicos, sólidos solubles y compuestos volátiles (K. Yu et al., 2012). Estos aspectos contribuyen al desarrollo de sus propiedades organolépticas, cruciales para que el producto sea del agrado del consumidor. Un estudio llevado a cabo en Estados Unidos concluyó que para los consumidores de frutos frescos de esta especie son muy importantes el sabor (particularmente la dulzura), la frescura, la apariencia y la cantidad de jugo, y en menor medida influyen a su aceptación la cantidad de semillas, el tamaño del fruto y la facilidad para pelarlo (Z. Gao et al., 2011).

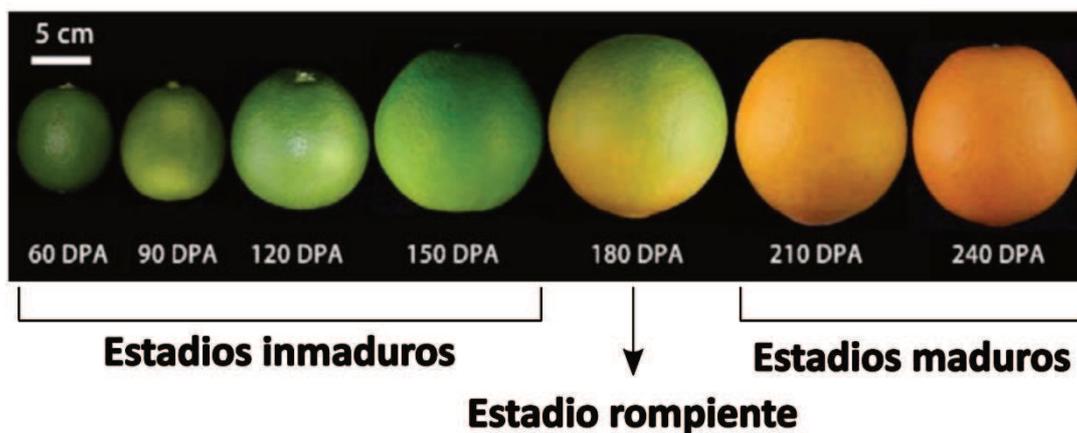


Fig. 4.1: Estadios de maduración de frutos de naranja, indicando los días post antesis (DPA) de cada fruto. Adaptado de Wang et al. (2016).

El naranjo es un fruto no climatérico, por lo que no pueden seguir madurando una vez fuera de la planta, ni siquiera con la presencia de etileno exógeno. A pesar de esto las naranjas muestran algunos signos de respuesta a la aplicación de etileno o de análogos químicos, como el etefón (Goldschmidt, 1998). En cítricos se vio que el etileno favorece la pigmentación de la cáscara y la síntesis de carotenoides (Stewart & Wheaton,

1972), además de acelerar el metabolismo de la pared celular en frutos y hojas, favoreciendo el fenómeno de abscisión (Goren et al., 1984). Los carotenoides son los principales pigmentos implicados en la coloración tanto de la cáscara como de la pulpa de los cítricos, y ambos atributos son muy importantes para los consumidores de frutos frescos (Lado et al., 2019). Esta familia de compuestos isoprenoides también tiene interesantes características nutricionales, ya que algunos de sus miembros son precursores de la vitamina A y presentan actividad antioxidante (Alquezar et al., 2008).

Los frutos de naranja presentan algunas características especiales en comparación con los frutos de tomate y frutilla analizados en los capítulos anteriores. Debido a que la recolección de naranjas se hace de forma manual, es importante tomar las precauciones necesarias durante este proceso para evitar que la calidad del producto se vea alterada. La fruta puede sufrir golpes por abscisión natural, descuido o durante su transporte, lo que puede derivar en grandes pérdidas económicas si el daño es severo. Por otro lado, los cítricos están cubiertos por una capa cerosa que los protege de ciertos microorganismos, evita la pérdida de agua por evaporación y proporciona un brillo particular que los vuelve estéticamente agradables (Shahid, 2007).

#### 4.1.2.2 - Producción e importancia económica

La naranja es producida y consumida en todo el mundo, tanto como fruta fresca como en productos derivados, principalmente jugo. Es una actividad económica muy importante, que reporta cifras de alrededor de 10.000 millones de dólares anuales (Q. Xu et al., 2013). De acuerdo a datos del año 2018, el principal productor a nivel mundial es Brasil, con aproximadamente el 22% del total mundial, seguido por China e India, con 12.2% y 11.1% del total, respectivamente. Estados Unidos y México ocupan el cuarto y quinto puesto, mientras que España se encuentra en el sexto lugar y es el principal productor de Europa. Nuestro país se ubica en la posición 16 a escala global, con una producción que supera el millón de toneladas, siendo el segundo productor más importante de Sudamérica y aportando el 1,3% del total general (*Food and Agriculture Organization*, 2018).

De acuerdo a datos de la Federación Argentina del Citrus, las provincias de Entre Ríos y Corrientes son las principales productoras de naranjas de nuestro país, con 35,7% y 30% del total, respectivamente, seguidas por las provincias del Noroeste argentino,

que aportan el 28% (encabezadas por Jujuy, responsable de alrededor del 16% nacional, Fig. 4.2), (Federcitrus, 2018).

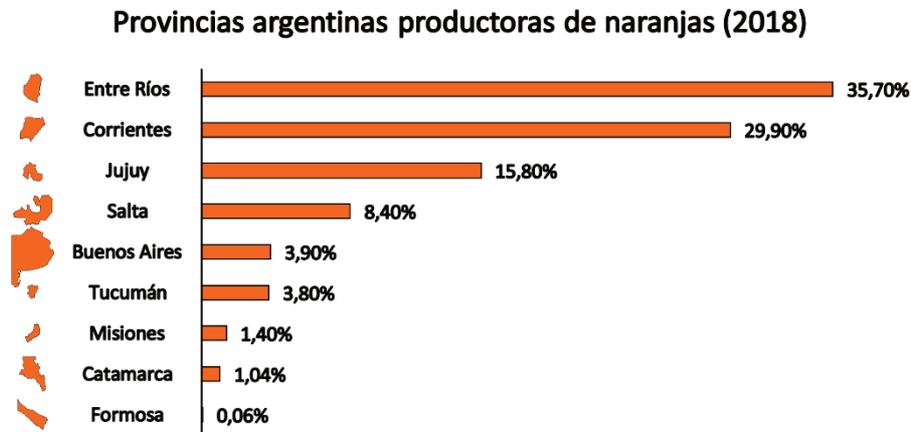


Fig. 4.2: Principales provincias exportadoras de naranja del año 2018

#### 4.1.3 - Genomas de referencia y ARNs pequeños de *Citrus sinensis*

*Citrus sinensis* es una especie diploide, con 9 cromosomas, un tamaño aproximado de 367 Mpb y aproximadamente 29.500 genes que codifican para proteínas funcionales (Arumuganathan & Earle, 1991; Q. Xu et al., 2013). Actualmente están disponibles los genomas de dos cultivares diferentes de naranja: “Ridge Pineapple”, versión 1.1 (G. A. Wu et al., 2014) y “Valencia”, versión 2 (Q. Xu et al., 2013). El primero está distribuido en 12.574 *scaffolds*, mientras que el segundo se encuentra en un estado más avanzado en cuanto a la organización y anotación de la información, donde los datos de secuenciación están organizados en los 9 cromosomas de la especie, más un cromosoma extra que aloja aquellas regiones que aún no pudieron ser ubicadas en cromosomas.

Existen diversos trabajos recientes donde se analizan miARNs en naranjo. Se han hecho búsquedas de miARNs y otros ARNs pequeños en diversos tejidos, como frutos, hojas y flores, con la intención de ampliar los conocimientos sobre sus vínculos con el proceso de maduración de frutos, pero sin profundizar en la relación de estas moléculas reguladoras con la adquisición de propiedades cualitativas de interés comercial de los frutos (Y. Liu et al., 2014). Otro campo de interés en esta especie es la resistencia frente a diversos patógenos y cómo los ARNs pequeños participan en actividades defensivas. Se ha visto participación de miARNs frente a infección del virus de la psorosis de los

cítricos (Marmisolle et al., 2020) e incluso se vio una marcada inducción de ARNs pequeños en plantas con la enfermedad Huanglongbing, posiblemente la más delicada para los cítricos actualmente (Hongwei Zhao et al., 2013).

En este trabajo de tesis se analizaron datos de secuenciación masiva obtenidos de bibliotecas de ARNs pequeños de frutos de naranjo del estadio de maduración “rompiente” (Fig. 4.1) con el fin de detectar ARNs pequeños que participen en la adquisición de propiedades organolépticas de frutos. Además, se utilizó una biblioteca de degradoma de frutos de *C. sinensis* para validar los resultados de predicción de genes target.

## 4.2 - Materiales y métodos

Los comandos utilizados para los procedimientos descritos en esta sección se detallan en el Anexo I.

### 4.2.1 - Datos de secuenciación masiva

Se realizaron búsquedas bibliográficas en *Pubmed* (<https://pubmed.ncbi.nlm.nih.gov/>) de publicaciones científicas que incluyeran el uso y producción de datos de secuenciación masiva (NGS) correspondientes a ARNs pequeños (*small RNA-seq*) de frutos de naranjo y que cuenten con datos provenientes de réplicas biológicas. También se hicieron búsquedas de datos de secuenciación masiva en la base de datos *Sequence Read Archive* (SRA, <https://www.ncbi.nlm.nih.gov/sra>), para identificar datos de secuenciación que no estén asociados a una publicación, al igual que en los capítulos anteriores. Se utilizó la herramienta *fastq-dump* del paquete *SRA-toolkit* (<http://ncbi.github.io/sra-tools/>) para convertir el formato SRA original a formato *fastq* y de esta manera se descargaron los conjuntos de datos de secuenciación de bibliotecas de *small RNA-seq* y de degradomas de frutos de naranjo *C. sinensis* cv. *Fengjie* y cv. *Anliu*, respectivamente (Tabla 4.1).

### 4.2.2 - Pre-procesamiento de datos de NGS

Se evaluó la calidad de secuenciación de los archivos *fastq* descargados utilizando la herramienta FastQC v 0.11.5 (Andrews, 2010). Para cortar la secuencia correspondiente a los adaptadores de secuenciación se utilizó el programa Fastx clipper

v. 0.0.14 (Gordon, 2011), modificando los parámetros por *default* como se indicó en capítulos anteriores.

#### 4.2.2.1 - Filtrado de secuencias contaminantes

Con el fin de eliminar las secuencias de ARN ribosomales (ARNr), ARN de transferencia (ARNt), ARN nucleares pequeños (snARN) y ARN nucleolares pequeños (snoARN), que no son de interés para el análisis, se descargó la base de datos de Rfam 14.0 (Griffiths-Jones et al., 2003) de la que sólo se conservaron las secuencias pertenecientes a *Citrus sinensis*, modificándose U por T, como se indicó anteriormente. Para eliminar las secuencias contaminantes se utilizaron los archivos *fastq* sometidos al pre-procesado detallado en el punto anterior y el programa Bowtie v 1.1.1 (Langmead, 2010) para descartar aquellas que alinean sobre el conjunto de secuencias ARNr/ARNt/sn y snoARN de naranjo, de manera análoga a lo indicado en los capítulos 2 y 3.

#### 4.2.3 - Análisis de miARNs de naranjo

##### 4.2.3.1 - Descarga de la información correspondiente a miARNs de naranjo

Se descargó la información correspondiente a miARNs de naranjo de la base de datos de miRbase (versión 22.1) (Kozomara & Griffiths-Jones, 2014). La información se almacenó en formato *fasta* para indexarla posteriormente usando *bowtie-build* para emplear el programa *bowtie*, y utilizarla como referencia para los alineamientos.

##### 4.2.3.2 - Mapeo al genoma de referencia, anotación y cuantificación de miARNs

Luego del pre-procesado, se alinearon los archivos en formato *fastq* al genoma de referencia (*Citrus sinensis* cv. *Valencia*, versión 2, <http://citrus.hzau.edu.cn/orange/>) siguiendo un procedimiento análogo al mencionado en el punto 2.3.2 de Materiales y métodos del Capítulo 2. Posteriormente, con la herramienta *fastx\_collapser* v. 0.0.14 (Gordon, 2011), utilizando las opciones por *default*, se colapsaron y cuantificaron las *reads* repetidas de los archivos resultantes. Se alineó cada uno de estos ficheros colapsados a la lista de miARNs de igual manera a la indicada en el punto 2.3.2 de Materiales y métodos del Capítulo 2. Para conservar sólo aquellos miARNs que tuvieron alineamientos perfectos, se utilizó la herramienta *Samtools* v 1.5 (Heng Li et al., 2009), siguiendo el procedimiento detallado en capítulos anteriores. La información obtenida

se organizó como tabla indicando los niveles de expresión normalizados en *reads* por millón (RPM) de cada miARN para cada réplica del estadio de madurez analizado.

#### 4.2.4 - Detección de loci en fase

Los archivos fastq pre-procesados correspondientes a réplicas biológicas de un mismo estadio fueron combinados en un solo archivo ya que una mayor cantidad de lecturas beneficia la identificación de *clusters* de ARNs pequeños en fase. Luego se utilizó el programa Shortstack 3.8.3 (Axtell, 2013), de acuerdo a lo indicado en el punto 2.5 de Materiales y métodos del Capítulo 2.

#### 4.2.5 - Validación de genes blanco con bibliotecas de degradoma

Los genes blanco de los miARNs analizados y de los phasiARNs detectados fueron validados bioinformáticamente con una biblioteca de degradoma de frutos de *Citrus sinensis* del cultivar *Anliu* tomados 170 días luego de la floración (Tabla 4.1) y el software Cleaveland v4.4 (Addo-Quaye et al., 2009), conservando sólo los resultados con valor  $p \leq 0,1$  y tipo de degradoma 0 o 1. Los datos de secuenciación correspondientes a bibliotecas de degradoma fueron pre-procesados como se indicó anteriormente (Pre-procesamiento de datos de NGS). Luego se convirtió el formato fastq a fasta con la herramienta FASTQ-to-FASTA v. 0.0.14 (Gordon, 2011).

Se consideró como posibles blancos de miARNs y phasiARNs al conjunto de transcritos codificados en el genoma de *Citrus sinensis* y para esto se descargó la versión 2 del archivo de anotaciones correspondiente al genoma utilizado de la base de datos *Citrus sinensis Annotation Project* (<http://citrus.hzau.edu.cn/orange/>). Además fue necesario reemplazar espacios y otros caracteres en la línea correspondiente al identificador de secuencia, como se realizó en los capítulos anteriores.

**Tabla 4.1: Números de acceso de los datos de secuenciación analizados**

Número de Acceso	Tejido	Biblioteca
SRR3774377 <sup>[1]</sup>	Pulpa_R1	<i>Small RNA-seq</i>
SRR3774378 <sup>[1]</sup>	Pulpa_R1	<i>Small RNA-seq</i>
SRR848997 <sup>[2]</sup>	Pulpa - mezcla	<i>Degradoma</i>

Referencias: [1] Wu, J. et al., 2016, [2] Liu, Y. et al., 2014.

## 4.3 - Resultados

### 4.3.1 - Análisis de miARNs

Al momento de iniciar este análisis existían algunos reportes en los que se habían realizado análisis genómicos de ARNs pequeños. Sin embargo, estos reportes no profundizaron en la relación entre miARNs, phasiARNs, maduración y desarrollo de atributos de interés comercial en frutos de esta especie. Por ejemplo, un estudio comparó los miARNs para un solo estadio de maduración de frutos entre un cultivar *wild type* y el correspondiente mutante con maduración tardía, con la intención de identificar miARNs con expresión diferencial entre ambos. A su vez, también llevaron adelante un análisis de descubrimiento de *loci* PHAS, pero no realizaron validación de genes blanco para los phasiARNs detectados (J. Wu et al., 2016). Otro grupo analizó miARNs en hojas, flores y frutos de naranjo, haciendo hincapié en su relación con el desarrollo vegetal global, sin análisis particular del proceso de maduración de frutos (Y. Liu et al., 2014). Teniendo en cuenta estos antecedentes, realizamos este análisis orientado a identificar aquellos miARNs y phasiARNs presentes en frutos del estadio “rompiente”, y en ambos casos se continuó con un análisis de identificación y validación de los transcriptos regulados por estos ARNs pequeños, con especial énfasis en la identificación de genes relacionados con el tema de interés para esta tesis (Fig. 4.3).

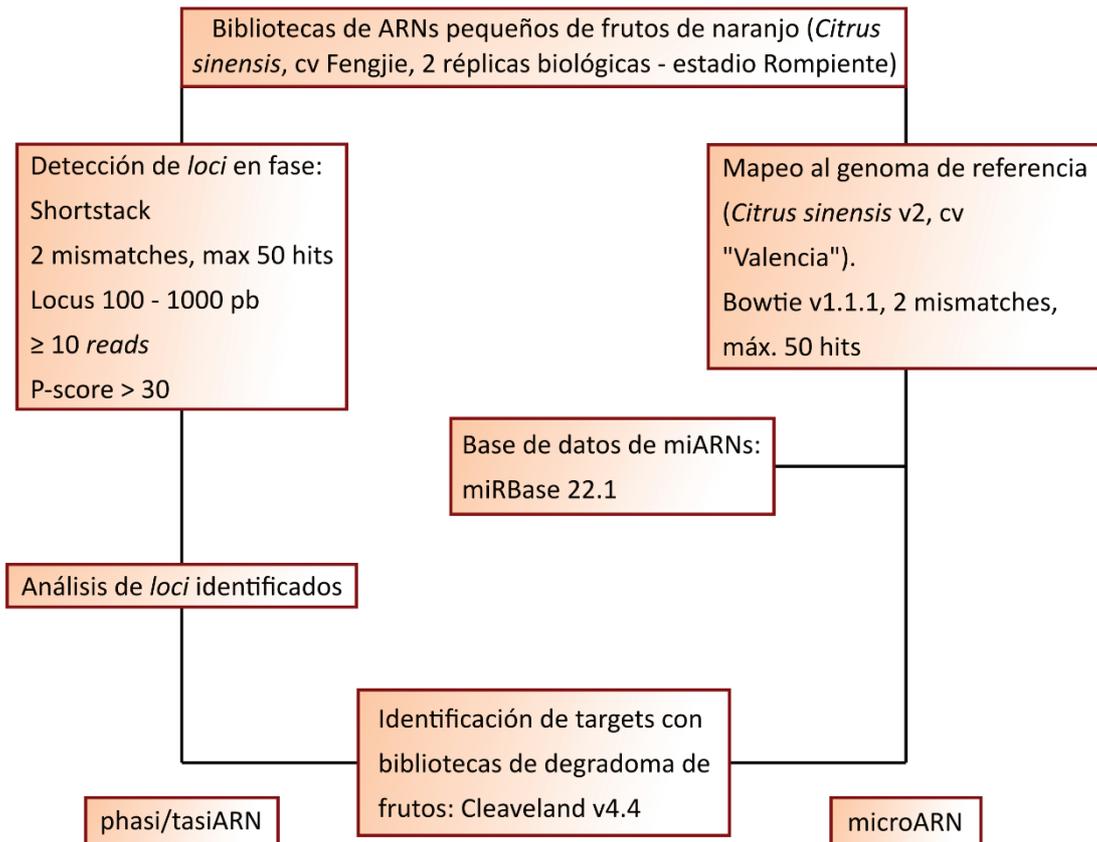


Fig. 4.3: Procedimiento seguido para el procesamiento de los datos de secuenciación de ARNs pequeños de naranjo.

#### 4.3.1.1 - Pre-procesado de los datos de secuenciación de small RNA-seq

En primer lugar se realizó el análisis de calidad de los datos de secuenciación (Tabla 4.1, Figura 4.3). Los reportes del programa FastQC v 0.11.5 correspondientes a cada biblioteca, indicaron que la calidad por secuencia de cada biblioteca era superior a 28, el promedio de calidad por secuencia estaba entre 32 y 38 y además, el porcentaje de bases no asignadas (N) fue despreciable en todos los casos, por lo que se decidió continuar con el análisis sin realizar filtrados por calidad de los datos de secuenciación. Se identificó a la secuencia "TCGTATGC" como el inicio de la porción correspondiente a los adaptadores usados en la preparación de las bibliotecas de *small RNA-seq* y se eliminó de los datos de secuenciación.

En la Tabla 4.2 se puede observar que luego de eliminar adaptadores se retuvo más del 98% de las secuencias originales en las bibliotecas analizadas, y tras filtrar las secuencias contaminantes, el porcentaje de *reads* limpias en todas las bibliotecas se

mantuvo entre el 89,5% y el 94,7% de las secuencias sin adaptadores. Luego de estos procedimientos, la cantidad de *reads* limpias remanentes para utilizar en los análisis posteriores fue adecuada para el objetivo que se pretendía llevar a cabo (Fig. 4.3). Sin embargo, es necesario mencionar que el porcentaje de *reads* alineadas al genoma de referencia es menor a los porcentajes observados típicamente.

Tabla 4.2: Pre-procesado y mapeo de bibliotecas de *small RNA-seq* de naranjo

Cv. Fengjie	<i>Reads</i> crudas	<i>Reads</i> tras eliminar adaptadores (%)	<i>Reads</i> tras filtrado por calidad (%)	<i>Reads</i> limpias (%)*	<i>Reads</i> alineadas al genoma (%)**
Rompiente R1	31146038	30926738 (99,3)	-	29278788 (94,7)	16605401 (56,7)
Rompiente R2	23716892	23468087 (98,9)	21754990 (92,7)	20994071 (89,5)	11364436 (54,1)

\* Porcentaje respecto del número de *reads* sin adaptadores

\*\* Porcentaje respecto del número de *reads* limpias

#### 4.3.1.2 - Expresión de miARNs en frutos de naranjo

Para identificar los miARNs presentes en frutos de naranjo y analizar si existen genes regulados por los mismos que estén relacionados con el desarrollo de las propiedades organolépticas de estos frutos, se alinearon todas las secuencias de ARNs pequeños remanentes, luego del pre-procesado (Tabla 4.2) a la versión 2 del genoma de referencia de *Citrus sinensis* cv. Valencia (<http://citrus.hzau.edu.cn/orange/>). Luego de la anotación y cuantificación de los miARNs disponibles en la base de datos miRbase, se determinó que de un total de 80 familias analizadas, al menos un miembro de 29 de ellas se expresa en la muestra de frutos analizada (Fig. 4.4A). Posteriormente, para cada uno de los miARNs detectados en estos frutos, se realizó un análisis de identificación y validación de genes target usando bibliotecas de degradoma, que se resumen en la Fig. 4.4B-C, y se describen más adelante. Dentro de las 29 familias de miARNs que se detectaron en la muestra analizada, solo 6 corresponden a familias de miARNs ampliamente conservadas en especies vegetales: *csi-miR156*, *csi-miR159*, *csi-miR167*, *csi-miR396*, *csi-miR482* y *csi-miR858*, mientras que las 23 restantes corresponden a miARNs específicos de esta especie.

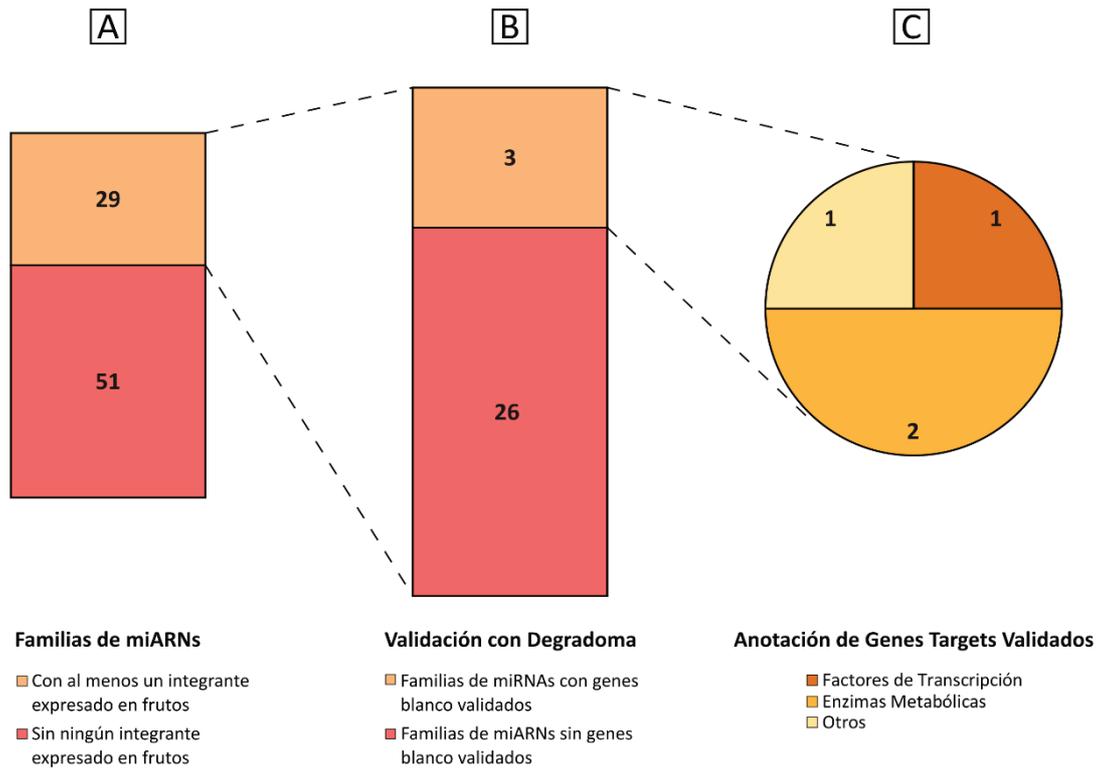


Fig. 4.4. Análisis de miARNs en *C. sinensis*. (A) Total de familias analizadas con alguno de sus miembros expresado en frutos o sin expresión. (B) Total de familias expresadas en frutos con o sin genes blanco validados. (C) Distribución de los procesos biológicos en los que participan los genes blanco de miARNs validados en frutos.

#### 4.3.1.3 - Identificación y validación de genes regulados por miARNs en frutos

Para estudiar los procesos biológicos regulados por los miARNs detectados en los frutos de *Citrus sinensis* analizados, se hizo un análisis de identificación y validación de los genes *target* análogo a los realizados en los Capítulos 2 y 3. Esto permitió conocer los procesos biológicos regulados por los miARNs presentes en frutos de naranja, validándose los genes blanco de 3 familias de miARNs (Fig. 4.4B; Tabla 4.3).

De acuerdo a su anotación, se clasificaron los genes blanco según el proceso biológico del que participan (Fig. 4.4C, Tabla 4.3). Dos de ellos corresponden a enzimas involucradas en actividades metabólicas: *csi-miR12107* regula a una proteína del tipo serina proteasa, cuya función implica la degradación de proteínas a péptidos simples (Alici & Arabaci, 2018) y un miembro de la familia *csi-miR156* regula la expresión del gen codificante para una proteína inhibidora de cisteína proteasas, las cuales han sido vinculadas a múltiples funciones fisiológicas, como senescencia, germinación de semillas y roles defensivos (Martínez et al., 2012). La familia *csi-miR156* regula, además, un factor

de transcripción de la familia *Squamosa Promoter Binding-like Protein*, que como se mencionó previamente, están relacionados al desarrollo temprano de los frutos, la floración, la distribución de tricomas y la síntesis de antocianinas en otras especies como tomate y *Arabidopsis*, pero está poco caracterizado en frutos de naranjo (Gou et al., 2011; G. F. F. E. Silva et al., 2014; J. W. Wang et al., 2009; N. Yu et al., 2010). Finalmente, como blanco de regulación de la familia *csi-miR159* se detectó a un transcripto que codifica para una proteína implicada en la reparación del ADN, a diferencia de lo estudiado en otras especies, donde los miARNs de esta familia regulan la expresión de factores de transcripción del tipo MYB (E. M. da Silva et al., 2017; Karlova et al., 2014). No se validaron genes target para las otras familias de miARNs conservadas que se mencionaron en el apartado anterior, a pesar de que *miR167*, *miR396*, *miR482* y *miR858* regulan transcriptos ampliamente distribuidos en las especies vegetales. Es sorprendente que solo se validaron targets para 3 de las 28 familias detectadas en las muestras de frutos analizadas.

**Tabla 4.3: Validación de genes blanco de familias de miARNs expresados en frutos de naranjo**

	<b>Familias de miARNs presentes en frutos</b>	<b>Gen blanco</b>
<b>Factores de transcripción</b>	<i>csi-miR156</i>	<i>Squamosa Promoter-binding-like protein</i>
<b>Actividad metabólica</b>	<i>csi-miR12107</i>	<i>Cucumisin - Xylem serine proteinase 1</i>
	<i>csi-miR156</i>	<i>Cysteine proteinase inhibitor 12</i>
<b>Otros</b>	<i>csi-miR159</i>	<i>DNA-damage-repair/toleration protein</i>

#### 4.3.2 - Descubrimiento y caracterización de loci generadores de phasiARNs

Con el fin de caracterizar a los phasiARNs presentes en las muestras de frutos analizadas, se realizó un análisis utilizando la herramienta Shortstack 3.8.3. Se detectaron 63 *loci* Phas (Tabla 4.4), 36 de ellos no han sido descritos hasta el momento, mientras que los restantes fueron reportados en un estudio previo (J. Wu et al., 2016), aunque en el estudio mencionado se utilizó otra herramienta bioinformática para la detección de phasiARNs y también analizaron bibliotecas de una variedad mutante, pero no se caracterizaron funcionalmente. En este trabajo de tesis se analizaron en detalle

aquellos *loci* que tenían un *P-score* mayor o igual a 30 y una longitud entre 100 y 1000 pb, manteniendo la metodología aplicada para los Capítulos 2 y 3 (Das et al., 2020; Feng et al., 2019; Nakamura et al., 2019).

**Tabla 4.4: Descripción de los *loci* en fase detectados en frutos de *Citrus sinensis*.**

<b>Locus</b>	<b>Ubicación genómica</b>	<b>Modelo génico</b>	<b>Gen asociado</b>
<b>Locus 1</b>	chr7:20107602-20108113	Cs7g22460	-
<b>Locus 2</b>	chrUn:69051387-69052223	Ninguno	-
<b>Locus 3</b>	chr7:25303300-25303972	Cs7g25130	-
<b>Locus 4</b>	chrUn:51117050-51117722	Ninguno	-
<b>Locus 5</b>	chr1:28371850-28372664	Cs1g26280	Putative senescence-associated protein
<b>Locus 6</b>	chr7:20286173-20286715	Cs7g22555	-
<b>Locus 7</b>	chr7:25320557-25320669	Ninguno	-
<b>Locus 8</b>	chr7:25339711-25340574	Ninguno	-
<b>Locus 9</b>	chr9:4267832-4268829	Cs9g06846	-
<b>Locus 10</b>	chr7:23022350-23022865	Cs7g24050	Vacuolar protein sorting-associated protein 36
<b>Locus 11</b>	chrUn:63868028-63868871	orange1.1t04123	Helicase domain and pentatricopeptide repeat-containing protein
<b>Locus 12</b>	chr7:25301538-25302117	Cs7g25130	-
<b>Locus 13</b>	chr1:11276435-11277242	Ninguno	-
<b>Locus 14</b>	chrUn:42232007-42232799	Ninguno	-
<b>Locus 15</b>	chr7:23020450-23020725	Cs7g24050	Vacuolar protein sorting-associated protein 36
<b>Locus 16</b>	chrUn:57334673-57335466	Ninguno	-
<b>Locus 17</b>	chr1:16120650-16121024	Cs1g13010	Probable disease resistance protein At5g63
<b>Locus 18</b>	chrUn:74022499-74023022	Ninguno	-
<b>Locus 19</b>	chr6:3972456-3972673	Cs6g03520	Endoribonuclease Dicer homolog 2
<b>Locus 20</b>	chr1:11167177-11168164	Cs1g09575	-
<b>Locus 21</b>	chr8:8670404-8670671	Ninguno	-
<b>Locus 22</b>	chrUn:82531557-82532136	orange1.1t05093	NAC domain-containing protein 69
<b>Locus 23</b>	chr1:11250699-11251334	Cs1g09660	NAC domain-containing protein 69
<b>Locus 24</b>	chr7:25340701-25341264	Ninguno	-
<b>Locus 25</b>	chrUn:50985550-50985856	Ninguno	-

<b>Locus 26</b>	chr5:21417876-21418224	Cs5g20240	Putative uncharacterized protein OSJNBb---6J12.7
<b>Locus 27</b>	chr1:22935004-22935961	Cs1g19720	Putative uncharacterized protein Sb-4g-364
<b>Locus 28</b>	chrUn:6602232-6602833	orange1.1t00557	-
<b>Locus 29</b>	chrUn:8394602-8394910	orange1.1t00706	TMV resistance protein N
<b>Locus 30</b>	chr3:17851136-17852115	Cs3g13390	Leucine-rich repeat containing protein 2C putative
<b>Locus 31</b>	chr6:3233191-3233498	Ninguno	-
<b>Locus 32</b>	chr2:1335828-1336034	Cs2g03070	Soluble starch synthase 3 2C chloroplastic/amyloplastic
<b>Locus 33</b>	chrUn:22189460-22189753	orange1.1t01308	Putative uncharacterized protein Sb26s--1- (Fragment)
<b>Locus 34</b>	chr2:11088012-11088584	Cs2g14260	-
<b>Locus 35</b>	chr5:24678673-24679155	Cs5g22170	Leucine-rich repeat receptor-like serine/threonine/tyrosine-protein kinase SOBIR1
<b>Locus 36</b>	chr1:15351815-15352557	Cs1g12310	Hyoscyamine 6-dioxygenase 3B Naringenin 2C2-oxoglutarate 3-dioxygenase
<b>Locus 37</b>	chr6:3971585-3971992	Cs6g03520	Endoribonuclease Dicer homolog 2
<b>Locus 38</b>	chrUn:20441425-20441757	Ninguno	-
<b>Locus 39</b>	chr2:17414276-17414943	Ninguno	-
<b>Locus 40</b>	chr3:14638202-14638480	Ninguno	-
<b>Locus 41</b>	chr7:27450809-27451623	Cs7g26750	Leucine-rich repeat containing protein 2C putative
<b>Locus 42</b>	chr1:19980453-19980923	Cs1g16680	Subtilisin-like protease
<b>Locus 43</b>	chr7:20104808-20105610	Ninguno	-
<b>Locus 44</b>	chr3:5918-6570	Ninguno	-
<b>Locus 45</b>	chr8:2908081-2908361	Ninguno	-
<b>Locus 46</b>	chr7:20079750-20080470	Cs7g22420	NAC domain-containing protein 4
<b>Locus 47</b>	chr3:9266239-9266868	Cs3g06650	Disease resistance protein RFL1 2C putative
<b>Locus 48</b>	chr7:27404228-27404906	Cs7g26730	TMV resistance protein N
<b>Locus 49</b>	chrUn:63890817-63891588	orange1.1t04124	Helicase domain and pentatricopeptide repeat-containing protein
<b>Locus 50</b>	chr1:15214400-15214832	Cs1g12210	LRR and NB-ARC domain-containing disease resistance protein
<b>Locus 51</b>	chr8:6332278-6332566	Cs8g08160	Helicase domain and pentatricopeptide repeat-containing protein

<b>Locus 52</b>	chr1:16034351-16034462	Ninguno	-
<b>Locus 53</b>	chr3:332451-333251	Ninguno	-
<b>Locus 54</b>	chr1:15193601-15193897	Cs1g12190	Probable disease resistance protein At5g63
<b>Locus 55</b>	chr1:15489980-15490214	Ninguno	-
<b>Locus 56</b>	chr7:26748977-26749373	Ninguno	-
<b>Locus 57</b>	chr7:27449905-27450313	Cs7g26750	Leucine-rich repeat containing protein 2C putative
<b>Locus 58</b>	chr8:5420424-5421245	Ninguno	-
<b>Locus 59</b>	chr8:6815490-6816377	Cs8g08570	Putative pentatricopeptide repeat-containing protein At1g127
<b>Locus 60</b>	chrUn:63874117-63874667	Ninguno	-
<b>Locus 61</b>	chrUn:85050783-85051109	orange1.1t05262	Putative plant disease resistant protein
<b>Locus 62</b>	chr8:6799501-6799990	Cs8g08549	-
<b>Locus 63</b>	chrUn:57455023-57455180	Ninguno	-

De los 63 *loci* detectados, 38 coinciden en su ubicación genómica con genes codificantes anotados en la versión 2 del genoma de referencia y 25 no comparten ubicación con ningún gen, por lo que podrían considerarse *loci* TAS (Tabla 4.4, resaltados en rojo). Además, 8 de los modelos génicos coincidentes con *loci* PHAS no tienen una anotación funcional para el gen asociado, por lo que se realizaron búsquedas por homología utilizando BLAST usando varias bases de datos contra distintas especies, pero no fue posible vincularlos con una potencial función.

#### 4.3.2.1 - Detección y validación de genes blanco

Posteriormente se identificaron los phasiARNs producidos a partir de estos 63 *loci*, se extrajo su secuencia a partir de los datos resultantes del alineamiento realizado por Shortstack y se les asignó un nombre, de manera análoga a lo realizado en los capítulos anteriores.

Para estudiar la importancia biológica de los siARNs identificados, se realizó un análisis de búsqueda y validación de genes blanco, análogo al realizado y descrito previamente para miARNs. La selección de aquellos genes blanco cuyos gráficos de *T-plot* pertenecen a las categorías 0 y 1 asignadas por Cleaveland y tienen un *p-value*  $\leq 0,1$  resultó en un total de 26 genes validados. De acuerdo al análisis detallado de las anotaciones funcionales de cada uno de los genes validados como targets de los

phasiARNs provenientes de los 63 *loci PHAS* identificados en este trabajo, se determinó que regulan una amplia variedad de procesos biológicos, los cuales se clasificaron en las categorías representadas en la Figura 4.5. De esta manera, se determinó que 6 de ellos codifican para enzimas no metabólicas, siendo esta la categoría mayoritaria, seguida por enzimas metabólicas y proteínas de unión, con 5 genes blanco cada una (Fig. 4.5). La mayoría de estos phasiARNs pertenecen al *locus 2*, y uno de ellos (L2\_S\_91) regula dos genes que codifican para enzimas del tipo Metilesterol Monooxigenasa, con propiedades beneficiosas para la salud humana, tal como se mencionó en el Capítulo 3.

### Distribución de genes blanco por función

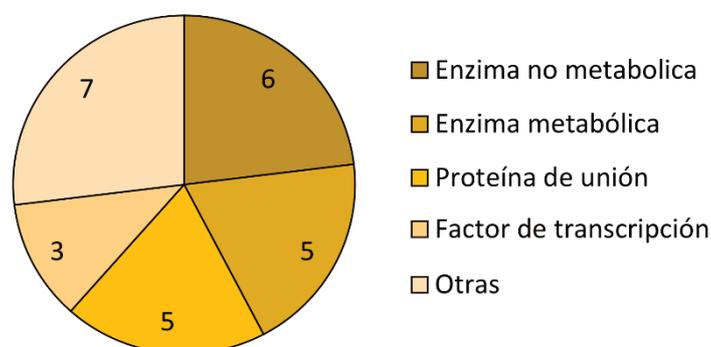


Fig. 4.5: Distribución de los procesos biológicos en los que participan los genes blanco validados para los phasiARNs detectados.

#### 4.3.2.2 - Genes relacionados a la adquisición de propiedades organolépticas regulados por phasiARNs

Con el objetivo de hallar transcritos validados como targets de phasiARNs implicados en la maduración de frutos o el desarrollo de propiedades organolépticas de los mismos, se analizaron de forma visual los gráficos T-plot producidos por el programa Cleaveland y se realizaron búsquedas bibliográficas sobre la función de los transcritos validados. De esta manera, se identificaron 2 genes que codifican para proteínas con funciones biológicas de interés para esta tesis doctoral. La enzima Cinamil Alcohol Deshidrogenasa, regulada por el phasiARN L10\_S\_9 (Fig. 4.6A, izquierda) está involucrada en la síntesis de lignina en frutos, afectando directamente su firmeza, ya que su expresión es más alta en frutos con mayor dureza, mientras que disminuye considerablemente en frutos blandos (Gabotti et al., 2015); mientras que el transcrito regulado por L36\_AS\_16 codifica para la enzima Naringenina 2-oxoglutarato 3-dioxigenasa (Fig. 4.6A, derecha). Esta enzima cataliza la  $\beta$ -hidroxilación de flavononas a

dihidroflavonoles, que son intermediarios en la biosíntesis de flavonoles, antocianidinas y catequinas (Lucini et al., 2018).

Como se muestra en la Figura 4.6A, la validación del corte para estos transcritos pertenecen a la “categoría 0” asignada por CleaveLand, indicando que la mayor acumulación de *reads* de degradoma que alinean sobre estos transcritos coincide con los sitios de corte predichos para los phasiARNs L10\_S\_9 y L36\_AS\_16. El análisis de los niveles de expresión para cada phasiARN indica que L10\_S\_9 es más abundante en comparación con L36\_AS\_16 (Fig. 4.6B), aunque en ambos casos la cantidad de *reads* acumuladas es suficiente para ejercer una función biológica.

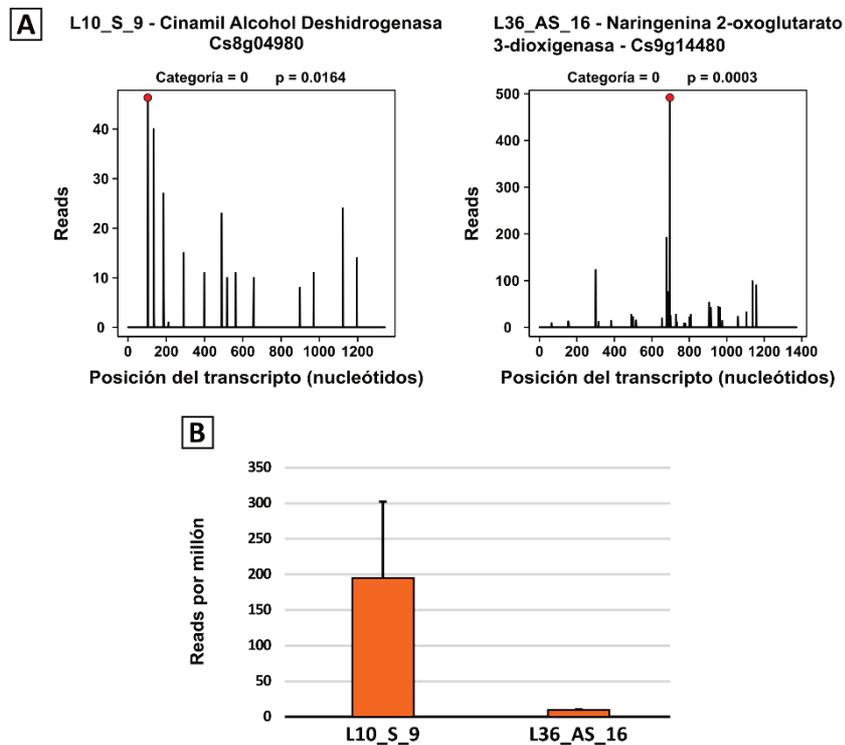


Fig. 4.6: (A) Validación del corte de los transcritos Cinamil Alcohol Deshidrogenasa y Naringenina 2-Oxoglutarato 3-Dioxigenasa mediados por L10\_S\_9 y L39\_AS\_16, respectivamente. El punto rojo indica las *reads* normalizadas que coinciden con la posición de corte del transcrito en la biblioteca de degradoma. Se indica categoría del pico y *p-value* según *CleaveLand*. (B) Perfil de expresión de los phasiARNs L10\_S\_9 y L39\_AS\_16 en frutos del estadio “rompiente”.

## 4.4 - Discusión

### 4.4.1 - En frutos de naranjo se expresan 29 familias de miARNs

Se ha analizado previamente a los miARNs de manera comparativa en diferentes tejidos de naranjo (Y. Liu et al., 2014), pero no hay hasta el momento estudios que los analicen en diversos estadios de maduración de frutos y evalúen su relación con la adquisición de propiedades organolépticas. Debido a la escasa disponibilidad de datos de secuenciación para esta especie, en este trabajo no fue posible realizar un análisis durante todo el proceso de maduración de frutos de naranja, sino que el análisis se vio limitado a un único estadio de maduración denominado “*breaker*” o “rompiente”, correspondiente a la aparición de la coloración naranja en los frutos (Fig. 4.1). Si bien nuestro análisis detectó la expresión de 29 familias de miARNs en este estadio (Fig. 4.4A), solo se validaron 4 genes blanco de regulación de algunos de estos miARNs: dos de ellos codifican para enzimas metabólicas, uno para un factor de transcripción del tipo SPL y el último para una proteína vinculada en la reparación del ADN (Fig. 4.4C, Tabla 4.3). Sorprendentemente, se detectó como gen blanco de regulación de un miembro de la familia *csi-miR156* a una proteína del tipo inhibidora de cisteína proteasa. Esta familia de miARNs regula en diversas especies la expresión de factores de transcripción SPL (Salinas et al., 2012; G. Wu & Poethig, 2006), por lo que este resultado constituye una novedad en la temática. Algo similar se observó para un miembro de la familia *csi-miR159*, cuyo blanco detectado fue un gen codificante para una proteína de reparación de ADN (Tabla 4.3). Esta familia de miARNs también está conservada entre especies y su blanco de regulación son los factores de transcripción del tipo MYB (Millar et al., 2019). Dado que ninguno de los genes blanco de regulación de miARNs detectados regulan genes relacionados con la adquisición de atributos de interés comercial para estos frutos, no profundizamos su análisis.

Para la mayoría de las familias detectadas de miARNs no se obtuvieron validaciones de genes blanco fidedignas. Esto puede estar relacionado con la baja expresión de estos genes en el estadio disponible para realizar estos análisis, y posiblemente el número sería mayor si se analizaran datos de secuenciación para frutos inmaduros y maduros.

#### 4.4.2 - Se identificaron 63 loci productores de phasiARNs en frutos de naranjo

En esta especie, se ha estudiado la relación de los ARNs pequeños en el proceso de embriogénesis somática (X. Wu et al., 2015) y se han buscado similitudes y diferencias entre frutos *wild type* y frutos mutantes con maduración tardía (J. Wu et al., 2016). En ambos trabajos se analiza la función de los genes PHAS, y se vio que muchos phasiARNs se originan a partir de una ubicación genómica que coincide con la de genes que dan lugar a proteínas relacionadas con funciones defensivas, resultados que concuerdan con lo reportado en este trabajo de tesis (Tabla 4.4). Al investigar los genes blanco regulados por los phasiARNs producidos, Wu J. y colaboradores (2015) se limitan al análisis de ontología génica, donde corroboran que los phasiARNs de interés para su análisis estaban implicados en muerte celular y resistencia a estrés, mientras que Wu X. y col. (2015) utilizan bibliotecas de degradoma provenientes de tejidos embrionarios para evaluar genes blanco, aunque no profundizan su estudio. De esta manera, el análisis con biblioteca de degradoma de frutos para cada phasiARN es una novedad de nuestro estudio, a través del cual se validaron 26 genes como blanco de regulación de los phasiARNs producidos por estos *loci*, que participan en diversos procesos fisiológicos.

Resultaron de particular interés para este trabajo las enzimas Cinamil Alcohol Deshidrogenasa y Naringenina 2-Oxoglutarato 3-Dioxigenasa (Fig. 4.6). La lignina se origina a partir de la vía de fenilpropanoides, y es uno de los principales componentes de la pared celular vegetal, vinculada con la firmeza de los frutos y con la protección de las semillas (Dardick et al., 2010). La enzima Cinamil Alcohol Deshidrogenasa participa activamente en la producción de este compuesto, y de acuerdo a lo observado en la Figura 4.6A, sería blanco de regulación de un phasiARN detectado en frutos de naranjo en el estadio de maduración “rompiente”. Se ha relacionado también a esta proteína a la resistencia al estrés por bajas temperaturas en durazno, donde se vio que sus niveles de expresión son menores en condiciones térmicas adversas (C. Zhang et al., 2010). Además se vio que en cultivares particularmente firmes de frutillas, la expresión de este gen se ve incrementada (Salentijn et al., 2003), por lo que es un candidato interesante pensando en la obtención de frutos con este aspecto mejorado.

Se detectó regulación mediada por phasiARNs para un gen codificante para Naringenina 2-Oxoglutarato 3-Dioxigenasa, otra enzima participante de la vía de fenilpropanoides. Esta proteína está implicada en la transformación de naringenina a dihidroflavonoles, y sus niveles son mayores conforme progresa la maduración de los frutos (A. Kumar et al., 2015). Se logró silenciar la expresión de este gen en frutilla y se comprobó que el contenido de antocianinas, flavonoles y glucósidos disminuyó marcadamente, y además se observó una disminución drástica en la coloración de los frutos (Fei Jiang et al., 2013). Esta información indica que se trata de una enzima crucial para el desarrollo de propiedades organolépticas de frutos y la vuelve una candidata interesante para estudiar en profundidad.

La secuenciación del genoma de naranjo se realizó hace relativamente poco tiempo y la comunidad científica especializada en esta especie es menor en comparación a las otras dos especies analizadas en esta tesis doctoral. Esto contribuye a que las anotaciones estructurales y funcionales de los genes de naranjo, en muchos casos necesitan ser mejoradas, lo cual típicamente se logra con la intervención de la comunidad científica. En el análisis realizado en esta tesis doctoral, esto resultó en dificultades adicionales para la ejecución de los estudios propuestos, sumado a la falta de información en general sobre las funciones de los ARNs pequeños en cítricos. A pesar de esto, nuestro análisis contribuyó al conocimiento de estos temas en naranjo, con un enfoque hacia el tema central de esta tesis que hasta el momento no había sido explorado.

## 4.5 - Conclusiones

- Se detectó expresión de 29 familias de miARNs en frutos de *Citrus sinensis*, y se validaron genes blanco para miembros de 3 familias.
- Se identificaron 63 *loci* productores de ARNs pequeños en fase, de los cuales 38 corresponderían a *loci* PHAS y 25 a *loci* TAS.
- Se validaron genes blanco de phasiARNs relacionados con propiedades organolépticas y desarrollo vegetal: Cinamil Alcohol Deshidrogenasa y Naringenina 2-Oxoglutarato 3-Dioxigenasa.

Capítulo 5: Biosíntesis de ARNs  
pequeños y el proceso de abscisión en  
naranja

## 5.1 - Introducción

### 5.1.1 - Componentes de las rutas de biogénesis y de acción de ARNs pequeños en naranjo

La biogénesis y acción de los ARNs pequeños en plantas depende principalmente de miembros de las familias ARGONAUTE (AGO), DICER-LIKE (DCL) y RNA-DEPENDENT RNA POLYMERASE (RDR). Las proteínas RDR son capaces de sintetizar ARN doble hebra (dsARN) usando ARN como molde y las proteínas DCL son responsables de cortar el dsARN en fragmentos de 21 a 24 nucleótidos gracias a su actividad del tipo RNAsaIII. Por su parte, el ARN pequeño proporciona la especificidad de acción del complejo RISC que contiene al factor AGO y es capaz de ejercer su acción tanto a nivel transcripcional como post-transcripcional (Q. Fei et al., 2013; Kapoor et al., 2008).

En plantas, las proteínas RDR contienen un único dominio conservado llamado ARN polimerasa dependiente de ARN (RdRP) (Bai et al., 2012; Wassenegger & Krczal, 2006), mientras que hay 6 dominios presentes en las proteínas DCL: DEXDc, Helicase-C, RNA-binding, PAZ, RNasella, RNasellb (RIBOc) y Dominio de Unión a ARN doble cadena (dsRB), aunque uno o más pueden estar ausentes (Margis et al., 2006). Dentro de las proteínas AGO, hay cuatro dominios principales, conocidos como N-Terminal, PAZ, Mid y PIWI (Hutvagner & Simard, 2008).

A raíz de la creciente disponibilidad de genomas vegetales en los últimos años, estas familias génicas han sido caracterizadas en diversas especies, lo cual ha contribuido a identificar la gran variabilidad existente en el número de miembros de cada familia y en la variable organización genómica de estos genes. De esta manera, se reportó un total de 20 genes codificantes para proteínas AGO, DCL y RDR en *Arabidopsis*, 28 genes en tomate y maíz (Bai et al., 2012; Qian et al., 2011), 32 en arroz (Kapoor et al., 2008) y 22 en uva y pimiento (Qin et al., 2018; Hualin Zhao et al., 2015). Además de su participación en la biogénesis y funcionalidad de ARNs pequeños, se ha relacionado a estas proteínas con procesos de respuesta frente a virus y diferentes tipos de estrés.

Esta información no estaba disponible para la especie *Citrus sinensis* al iniciar esta tesis doctoral, por lo que se llevó a cabo el estudio de identificación y caracterización de las familias AGO, DCL y RDR a nivel genómico, y se estudió la

participación de algunos de los miembros de estas familias y de otros genes de copia única en el proceso de abscisión de frutos, el cual se describe a continuación.

### 5.1.2 - El proceso de abscisión

La abscisión es el mecanismo mediante el cual las plantas pueden desprender órganos, como hojas o frutos. Tiene lugar en una capa de células específica, conocida como zona de abscisión (Merelo et al., 2017) y se lo divide en 4 etapas principales: desarrollo del tejido de la zona de abscisión, obtención de la capacidad de respuesta a estímulos de abscisión, percepción de estas señales con activación del fenómeno y finalmente, diferenciación de una capa protectora en la zona de abscisión del lado de la planta (Patterson, 2001). Desde un punto de vista evolutivo, es un proceso beneficioso ya que favorece la separación de órganos dañados o senescentes y la dispersión de semillas; sin embargo, desde un punto de vista productivo, la abscisión de los frutos presenta un gran impacto en el rendimiento, generando fuertes pérdidas y es un proceso particularmente importante en la producción de cítricos ya que la calidad de los frutos desprendidos por abscisión no es adecuada para su comercialización (Estornell et al., 2013; Merelo et al., 2017).

#### 5.1.2.1 - Regulación a nivel molecular del proceso de abscisión

A nivel celular, durante la abscisión los órganos sufren cambios en la expresión de genes relacionados con la integridad de la pared celular y con procesos biológicos relacionados a organelas con membranas, la señalización y biosíntesis de etileno y auxinas y respuesta a estrés (Nakano et al., 2012). Se ha propuesto que la inducción de la abscisión depende de una relación muy compleja entre la concentración de hormonas y otras sustancias que modifican la capacidad de los tejidos de responder a ellas (Gil-Amado & Gomez-Jimenez, 2013), y si bien existen estudios genómicos de esta temática en varias especies, el mecanismo no está aún completamente dilucidado (C. Cheng et al., 2015; Gil-Amado & Gomez-Jimenez, 2013; Hong Zhu et al., 2011). A pesar de que los cítricos son especies no climatéricas, la mayoría son sensibles a la inducción de la abscisión por agregado de etileno exógeno. Por el contrario, las auxinas retrasan este fenómeno, pero el mecanismo no está completamente esclarecido (Mesejo et al., 2021).

Por otro lado, se han observado cambios en los niveles de expresión de genes, pero existen escasos estudios sobre la participación de moléculas regulatorias como los ARNs pequeños. Un estudio previo analizó miARNs exclusivamente en los pedicelos de tomates tratados con etileno para inducir la abscisión de las hojas y detectaron variaciones en la expresión en genes relacionados con la unión de membranas y sus correspondientes miARNs reguladores respecto a las plantas sin tratar (T. Xu et al., 2015). Por su parte, estudios orientados a identificar cambios en la expresión de genes involucrados en este proceso, independientemente de que sean regulados por ARNs pequeños, identificaron variaciones en genes que codifican para enzimas relacionadas con el desensamblaje de la pared celular, como celulasas, glucanasas, pectato liasas y  $\beta$ -galactosidasas, entre otras, que aumentan su expresión al momento de producirse la abscisión (Estornell et al., 2013), y también se observaron cambios en la expresión de genes que codifican para proteínas relacionadas al transporte de metabolitos y nutrientes (Gil-Amado & Gomez-Jimenez, 2013).

Como parte de este trabajo de tesis, luego de llevar adelante la identificación y anotación de los componentes de las diversas rutas de biogénesis y acción de ARNs pequeños en el genoma de *Citrus sinensis*, se analizó cuáles de estas rutas participan en el proceso de abscisión a través del análisis de cambios de expresión génica usando datos provenientes de microarreglos de ADNc (Tadeo y col., comunicación personal, no publicado). Nuestro trabajo estableció que existe una relación entre la ruta de RdDM y el fenómeno de abscisión en *Citrus sinensis*, y se detalla a continuación.

## 5.2 - Materiales y métodos

### 5.2.1 - Identificación de los genes AGO, DCL y RDR en naranjo

Se descargó de Phytozome v12.1 (<https://phytozome.jgi.doe.gov>) la versión v1.1 del genoma de *Citrus sinensis* y sus secuencias de proteínas y la versión V2 del genoma y proteínas de “*Citrus sinensis annotation Project*” (CSAP, <http://citrus.hzau.edu.cn/orange/>). Las secuencias de aminoácidos de las proteínas DCL, AGO y RDR de *Arabidopsis thaliana* se obtuvieron de TAIR (<https://www.arabidopsis.org/>) y se las utilizó para buscar genes homólogos de naranjo con la herramienta online BLASTP de Phytozome. Se analizaron los dominios

conservados en las proteínas de naranja utilizando la herramienta Simple Modular Architecture Research Tool (SMART, <http://smart.embl-heidelberg.de/>) y la base de datos Pfam (31.0, <https://pfam.xfam.org/>). Se calculó el punto isoeléctrico y el peso molecular de las proteínas usando Compute pI/Mw ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). Las estructuras génicas se diseñaron con GSDS 2.0 (Hu et al., 2015).

### 5.2.2 - Análisis filogenético y localización cromosomal

Se alineó a las proteínas candidatas de *Citrus sinensis* con las proteínas AGO, RDR y DCL de *Arabidopsis thaliana*, *Solanum lycopersicum*, *Populus trichocarpa* y *Oryza sativa* utilizando MEGA 6.0 (Tamura et al., 2013). Para estimar un árbol filogenético se utilizó el algoritmo MUSCLE (con los parámetros *gap open*, -2.9; *gap extend*, 0; *hidrophobicity multiplier*, 1.2; *clustering method*, UPGMB; Edgar, 2004). Se construyó un árbol filogenético para cada familia de proteínas usando el método *Neighbor-Joining* con el test *bootstrap* repetido 1000 veces. Se hizo una representación de ubicación cromosomal de acuerdo a las anotaciones de CSAP, usando la herramienta online *MapGene2Chromosome V2* ([http://mg2c.iask.in/mg2c\\_v2.0/](http://mg2c.iask.in/mg2c_v2.0/)).

### 5.2.3 - Análisis de expresión de genes DCL, AGO y RDR en tejidos de plantas de naranja

Se usaron bibliotecas de *RNA-seq* disponibles públicamente para analizar los niveles de expresión de estos genes en diferentes tejidos. Los números de acceso se detallan en la Tabla 5.1. Luego de realizar el análisis de calidad de los datos crudos usando FastQC (Andrews, 2010), se eliminaron las secuencias de RNA ribosomal obtenidos de la base de datos Rfam 13.0 (<http://rfam.xfam.org/>), usando Bowtie 1.1.1 y permitiendo 2 mismatches (Langmead, 2010). Las *reads* limpias se mapearon al genoma de referencia de *Citrus sinensis* de CSAP o de Phytozome para el caso de CsRDR1c y CsRDR6b, que no están anotadas en la versión del genoma de CSAP. Se emplearon los programas Hisat2, versión 2.1.0 (D. Kim et al., 2015), y hisat2-build para crear los correspondientes genomas indexados. Luego, se utilizó featureCounts 1.5.3 (Liao et al., 2014) para obtener las *reads* específicas de cada gen, empleando los modelos de CSAP (*csi.gene.models.gff3*) o de Phytozome para CsRDR1c y CsRDR6b

(Csinensis\_154\_v1.1.gene.gff3). Los tejidos analizados se compararon todos contra todos de a pares, y se seleccionó como genes expresados diferencialmente a aquellos con un *fold change* de 2 y un *p-value* ajustado < 0.05 en al menos una de las comparaciones, usando DESeq2 versión 1.20.0 (Love et al., 2014). El gráfico estilo *heatmap* para representar esta información se hizo con el paquete de R *Ggplot2*, dentro del software RStudio 1.1.456 (Allaire, 2012; Villanueva & Chen, 2019).

**Tabla 5.1: Números de acceso de las bibliotecas de RNA-seq utilizadas**

Número de Acceso	Tejido	Biblioteca
SRR4050292	Raíz_R1	<i>RNA-seq</i>
SRR4050293	Raíz_R2	<i>RNA-seq</i>
SRR6451222	Pulpa_R1	<i>RNA-seq</i>
SRR6451225	Pulpa_R2	<i>RNA-seq</i>
SRR6451221	Pulpa_R3	<i>RNA-seq</i>
SRR6448479	Hoja_R1	<i>RNA-seq</i>
SRR6448480	Hoja_R2	<i>RNA-seq</i>
SRR4072700	Embrión_R1	<i>RNA-seq</i>
SRR4076115	Embrión_R2	<i>RNA-seq</i>
SRR2025104	Piel_R1	<i>RNA-seq</i>
SRR2025105	Piel_R2	<i>RNA-seq</i>

#### 5.2.4 - Análisis de datos de expresión de genes DCL, AGO y RDR de naranja durante el proceso de abscisión

Para el análisis de expresión génica en la zona de abscisión, se usaron los datos obtenidos de Merelo et al. (2017). En este trabajo analizaron la expresión génica en frutos de naranjos Washington Navel luego de 0, 12 y 24 horas de tratamiento con etileno, usado para promover la abscisión. Las muestras de ARN de la zona de abscisión y de la corteza de los frutos se obtuvieron por LCM. Para el análisis de expresión génica se utilizó un microarreglo de ADNc con 21.081 genes putativos de citrus (Merelo et al., 2017). Se consideró como diferencialmente expresados a los genes con un *p-value* ajustado < 0.05 y un *fold change* de  $\pm 2$ .

### 5.2.5 - Análisis cuantitativo de RT-PCR

Los tejidos de la zona de abscisión y corteza de los frutos fueron seccionados de frutas de naranja de la variedad Salustiana luego de 0, 12 y 24 horas de tratamiento con etileno, como se describió previamente (Merelo et al., 2017). Los frutos fueron colocados en dos desecadores herméticamente sellados, uno para frutos control y otro para frutos tratados en una atmósfera con etileno en una concentración de 10  $\mu\text{L/L}$ , generado *in situ* a partir de la mezcla de etefón al 48% (Gleba) con NaOH al 40%. El ARN total se extrajo utilizando el reactivo Quick-zol (Kalium Technologies) siguiendo las instrucciones del fabricante y se lo trató con DNasa I (Promega). Se sintetizó ADNc a partir de 1000 ng de ARN por muestra usando transcriptasa reversa EasyScript (Trans) de acuerdo al protocolo del fabricante. Se diseñaron *primers* específicos para los genes (Tabla 5.2) para usar con TransStart Tip Green qPCR SuperMix (Trans). Se determinó la especificidad de todos los productos de amplificación con análisis de curva de disociación. Se calcularon los valores de cuantificación relativa basados en tres réplicas biológicas y tres réplicas técnicas con el método  $2^{-\Delta\text{Ct}}$ , usando la expresión de Ubiquitina (CsUBI) como control de normalización. Sólo se necesitó corrección de eficiencia para los *primers* de CsSGS3 (que se calculó en 88%), como se describió previamente (Pfaffl, 2001).

**Tabla 5.2: Primers utilizados para RT-PCR**

Nombre	ID de Phytozome	Nombre del Primer	Secuencia del Primer (5' - 3')
CsSSH1	orange1.1g026004m	FW_CsSSH1	GAATGTGTCAAGGTCAACGTT G
		RV_CsSSH1	TTGGCAGCCTTCTGTGTCGT
CsSGS3	orange1.1g006189m	FW_CsSGS3	AGGCATTGGGAGGAAGAGGT
		RV_CsSGS3	TTTCTAACATCCTTACATCTCAG
CsNRPD2	orange1.1g001050m	FW_CsNRPD2	GCTACTCCCTCTGTTGATGC
		RV_CsNRPD2	CGAACCATTTACCAGTTCCGG
CsDCL1	orange1.1g000174m	FW_CsDCL1	CTGTTTGCGTAGAAATTGGCCT
		RV_CsDCL1	TTTTGACACTAGGCATTGGCTC
CsAGO4a	orange1.1g002449m	FW_CsAGO4a	GATGAGCTGCAGGAAGTAGTG
		RV_CsAGO4a	GCAGATGTCATCCCACCCTG
CsUBI	orange1.1g032515m	FW_CsUbi	TGTTTGTAGGGCTATAATCGCT
		RV_CsUbi	CGTCAATCACCCCTTCTTTGG

## 5.3 - Resultados

### 5.3.1 - Identificación y análisis in silico de los genes AGO, RDR y DCL en naranjo

Para identificar las familias génicas AGO, RDR y DCL en naranjo, se reunió información de genes previamente caracterizados en Arabidopsis, arroz, álamo y tomate (*Additional file 1*; Sabbione et al., 2019). Empleando las secuencias de Arabidopsis como entradas en los análisis usando la herramienta BLAST, se identificaron 5 genes que codifican para proteínas DCL (CsDCLs), 7 para proteínas RDR (CsRDRs) y 13 para proteínas AGO (CsAGOs) en las bases de datos de Phytozome y de “*Citrus Sinensis Annotation Project*” (CSAP), (Tabla 5.3). Aunque la mayoría de estos genes estuvieron presentes en ambas bases de datos, se detectaron varias inconsistencias en sus anotaciones estructurales y funcionales entre las dos versiones del genoma, como también una mala anotación para algunos de estos genes. En general, una mejor anotación estructural consistente con los datos de *RNA-seq* se observó en la base de datos de CSAP, excepto para los genes orange1.1g001771m (CsRDR3) y

orange1.1g002204m (CsAGO5a), que están mejor anotados en Phytozome (Tabla 5.3). Para trabajar con los modelos más exactos, se mejoró la anotación de 5 genes usando información de ambas bases de datos, como también información de *RNA-seq* de diferentes tejidos (Tabla 5.3). Por ejemplo, el modelo anotado en CSAP para CsAGO5c carece de los dominios Mid y PIWI; pero al compararlo con los datos de *RNA-seq*, se detectaron todos los dominios típicos de las proteínas AGO en el nuevo modelo anotado (Fig. 5.1, *Additional file 2*; Sabbione et al., 2019). Las características de todos los genes identificados en este estudio se detallan en la Tabla 5.3, incluyendo los números de identificación de estos genes (*gene ID*) en ambas bases de datos, longitud de los marcos abiertos de lectura, longitud de las proteínas, punto isoeléctrico (IP) y peso molecular (Mw), mientras que el análisis de estructura génica y las secuencias actualizadas de las proteínas para las anotaciones corregidas se detallan en Sabbione y col., 2019, junto con los ID de Pfam y SMART para todos los dominios identificados en las proteínas CsAGO, CsRDR y CsDCL y sus términos de ontología génica (GOs) asociados (*Additional files 2, 3 y 4*; Sabbione et al., 2019).

Los 7 genes de RDR encontrados en el genoma de naranja codifican proteínas con el dominio RdRP. El número de aminoácidos de los miembros de esta familia varía entre 969 y 1197 (Tabla 5.3). Se mejoró la anotación del gen orange1.1g048783m/Cs1g14730 (CsRDR6b) usando datos de *RNA-seq* de bibliotecas de hoja, raíz, embrión, pulpa y cáscara (Tabla 5.1, Tabla 5.3). Se observó expresión en todas las bibliotecas en la región anotada como tercer intrón para este gen, por lo tanto la estructura génica se actualizó. La región de la secuencia actualizada codifica para el dominio RdDP de esta proteína, que se predice correctamente al utilizar SMART luego de la re-anotación (Fig. 5.1), mientras que la versión anterior del gen mostraba un dominio RdDP incompleto.

La familia CsDCL está constituida por 5 miembros para los que las regiones codificantes varían entre 4191 y 5892 pb, y codifican para proteínas entre 1396 y 1963 aminoácidos (Tabla 5.3). Los análisis de Pfam y SMART revelaron que todas las proteínas CsDCLs mostraron 6 dominios conservados: DEXDc, Helicase C, RNA-binding, PAZ y dos dominios RIBOc consecutivos (RNaseIIIa y RNaseIIIb). Además de estos, CsDCL1 y CsDCL4 presentaron dos dominios dsRB después de los dominios RIBOc (Fig. 5.1).

**Tabla 5.3: Caracterización de los genes AGO, DCL y RDR en *Citrus sinensis***

	Nombre del gen	ID de Phytozome	ID de CSAP	Ubicación	Proteína			
					Largo de ORF (pb)	Largo total	PI	MW (Da)
DCL	CsDCL1	orange1.1g000174m	orange1.1t00584	ChrUn: 6828329–6,838,566 (- strand)	5892	1963	5.96	219,934
	CsDCL2a	orange1.1g000607m	Cs6g03520	Chr6: 3969936–3,982,992 (- strand)	4191	1396	7.65	158,488
	CsDCL2b	orange1.1g003062m	Cs6g03500	Chr6: 3910063–3,922,849 (- strand)	4206	1401	6.48	158,770
	CsDCL3	orange1.1g000379m	Cs4g06370	Chr4: 3914566–3,925,649 (+ strand)	4959	1652	7.95	184,339
	CsDCL4	orange1.1g000380m	Cs4g01340	Chr4: 255555–267,608 (+ strand)	4902	1633	6.47	183,256
RDR	CsRDR1a	orange1.1g002586m	Cs2g17570	Chr2: 14327852–14,333,995 (+ strand)	3399	1132	7.13	129,642
	CsRDR1b	orange1.1g003789m	Cs5g14110	Chr5: 11659292–11,663,430 (- strand)	3249	1082	6.04	122,608
	CsRDR1c	orange1.1g035741m	-	Scaffold00168–67,689 – 71,577 (- strand)	2910	969	8.58	110,376
	CsRDR2	orange1.1g001183m	Cs5g05170	Chr5: 3069393–3,074,020 (+ strand)	3396	1131	6.59	128,958
	CsRDR3 <sup>1</sup>	orange1.1g001771m	Cs4g15260	Chr4: 14278137–14,289,339 (+ strand)	3048	1015	6.8	115,671
	CsRDR6a	orange1.1g041430m	Cs7g05350	Chr7: 2849984–2,855,672 (+ strand)	3594	1197	6.13	136,437
	CsRDR6b <sup>2</sup>	orange1.1g048783m	Cs1g14730	Chr1: 18022840–18,027,238 (+ strand)	3501	1166	6.01	132,490
	CsAGO1	orange1.1g001466m	Cs5g16710	Chr5: 15851375–15,859,311 (- strand)	3222	1073	9.38	118,338

	CsAGO2a	orange1.1g012649m	Cs2g10760	Chr2: 8058870–8,063,724 (+ strand)	2946	981	9.29	110,071
	CsAGO2b	-	Cs2g10770	Chr2: 8075960–8,080,634 (+ strand)	2946	981	9.26	110,304
	CsAGO4a	orange1.1g002449m	Cs2g29070	Chr2: 28612071–28,619,504 (- strand)	2763	920	8.98	102,998
	CsAGO4b	orange1.1g002636m	Cs3g06860	Chr3: 9639528–9,647,482 (+ strand)	2697	898	9.43	100,950
	CsAGO5a <sup>1</sup>	orange1.1g002204m	Cs7g17940	Chr7: 13746350–13,751,613 (- strand)	2865	954	9.27	106,773
AGO	CsAGO5b <sup>2</sup>	orange1.1g003630m	Cs7g17970	Chr7: 13774473–13,781,129 (- strand)	2847	948	9.50	105,299
	CsAGO5c <sup>2</sup>	-	Cs7g17930	Chr7: 13732122–13,738,824 (- strand)	3021	1006	9.28	112,004
	CsAGO5d	-	Cs6g14430	Chr6: 15643144–15,650,122 (- strand)	2688	895	9.23	99,978
	CsAGO6	orange1.1g002661m	Cs2g20520	Chr2: 17331361–17,339,897 (- strand)	2688	895	9.41	100,500
	CsAGO6-like <sup>2</sup>	orange1.1g048669m	Cs6g16080	Chr6: 17010270–17,016,001 (- strand)	2814	937	9.71	105,659
	CsAGO7	orange1.1g001684m	Cs7g03360	Chr7: 1453333–1,457,453 (- strand)	3093	1030	9.24	117,036
	CsAGO10	orange1.1g001954m	Cs9g07740	Chr9: 4853982–4,862,787 (- strand)	2979	992	9.34	111,516

Las 13 proteínas AGO identificadas tienen los dominios N-terminal, L1, PAZ y PIWI, mientras que el dominio L2 está presente en 12 de estas proteínas, pero falta en orange1.1g003630m/Cs7g17970 (CsAGO5b). Además, el dominio Mid está presente en todas las proteínas, excepto en orange1.1g001684m/Cs7g03360 (CsAGO7) y orange1.1g001466m/Cs5g16710 (CsAGO1) es la única que comienza con un dominio rico en Glicina (Gly-Rich; Fig. 5.1).

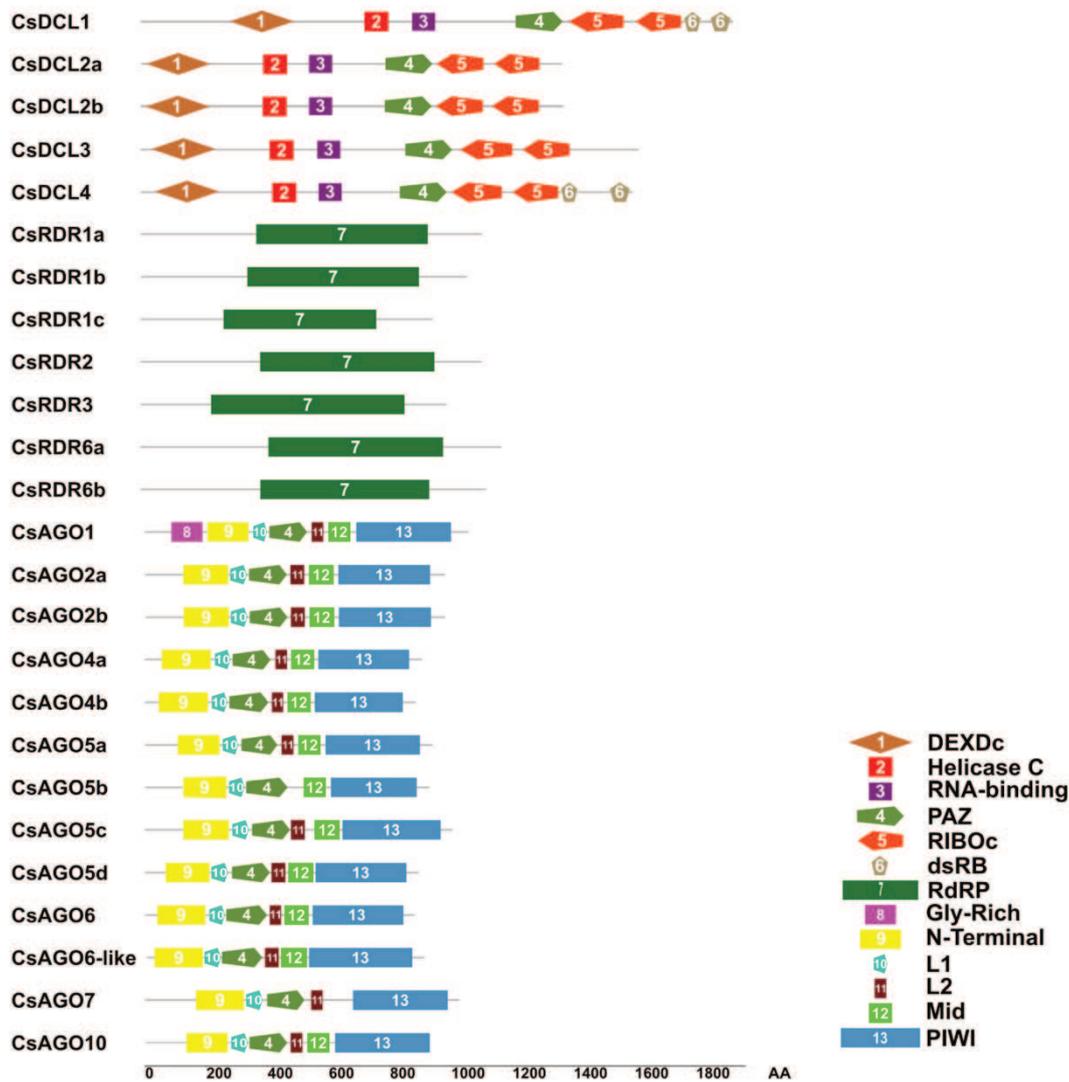


Fig. 5.1: Dominios presentes en las familias de proteínas AGO, DCL y RDR de *C. sinensis*. 1: DEXDc; 2: Helicase C; 3: RNA-binding; 4: PAZ; 5: RIBOc; 6: dsRB; 7: RdRP; 8: Gly-Rich; 9: N-terminal; 10: L1; 11: L2; 12: Mid; 13: PIWI. La representación está en escala de aminoácidos (AA).

Además de los dominios conservados presentes en estas proteínas, también se examinó la presencia de motivos conservados en estas familias. En *Arabidopsis*, las proteínas AGO tienen motivos DDH/H o DDD/H en el dominio PIWI, que son responsables de su actividad endonucleasa (Baumberger & Baulcombe, 2005; Carbonell et al., 2012; Rivas et al., 2005). Se encontró el motivo DDH/H en CsAGO1, CsAGO7 y CsAGO10, y el motivo DDD/H en CsAGO2a (Tabla 5.4). También se identificó el motivo DDH/P para CsAGO4a y CsAGO4b; un motivo DDH/S en CsAGO6 y los motivos DDY/H y DDY/P para CsAGO5a y CsAGO5b, respectivamente. Por último, se observó un motivo D-H/T en CsAGO6-like protein (Tabla 5.4), que por lo tanto aparentemente carecería de actividad endonucleasa.

Se detectó el motivo conservado DECH en todas las proteínas CsDCL, que es característico de esta familia proteica de plantas (Y. Lee et al., 2003; C. Liu et al., 2012) y se encontró el motivo DLDGD, que típicamente está presente en el dominio catalítico de proteínas RDR (J. Y. Cao et al., 2016; Wassenegger & Krczal, 2006) en 5 de las 7 proteínas CsRDR (Tabla 5.2).

**Tabla 5.4: Motivos conservados en las familias de proteínas analizadas de *Citrus sinensis***

	<b>Motivo conservado de AGO: DDH/H</b>		<b>Motivo conservado de DCL: DECH</b>		<b>Motivo conservado de RDR: DLDGD</b>	
<b>Motivos presentes</b>	<b>CsAGO1</b>	DDH/H	<b>CsDCL1</b>	DECH	<b>CsRDR1a</b>	DLDGD
	<b>CsAGO2a</b>	DDD/H	<b>CsDCL2a</b>	DECH	<b>CsRDR1b</b>	DLDGD
	<b>CsAGO4a</b>	DDH/P	<b>CsDCL2b</b>	DECH	<b>CsRDR1c</b>	DQDGE
	<b>CsAGO4b</b>	DDH/P	<b>CsDCL3</b>	DECH	<b>CsRDR2</b>	DLDGD
	<b>CsAGO5a</b>	DDY/H	<b>CsDCL4</b>	DECH	<b>CsRDR3</b>	DFDGD
	<b>CsAGO5b</b>	DDY/P			<b>CsRDR6a</b>	DLDGD
	<b>CsAGO6</b>	DDH/S			<b>CsRDR6b</b>	DLDGD
	<b>CsAGO6-like</b>	D-H/T				
	<b>CsAGO7</b>	DDH/H				
	<b>CsAGO10</b>	DDH/H				

### 5.3.2 - Análisis filogenético para identificar los ortólogos putativos de DCL, RDR y AGO en naranja

Se utilizaron análisis filogenéticos para identificar los genes de ortólogos putativos de las familias AGO, RDR y DCL en naranja. Se construyó un árbol filogenético sin raíz con el método *neighbor-joining* para las tres familias estudiadas usando secuencias proteicas completas de Arabidopsis, tomate, álamo y arroz (Fig. 5.2). La familia AGO se separó en 6 clados: AGO1, AGO10, AGO5, AGO7, AGO2/3 y AGO4/6/8/9 (Fig. 5.2A). Sólo se identificó un gen en el genoma de *C. sinensis* en cada uno de los clados que contenían miembros de AGO1, AGO10 y AGO7 en el resto de las especies analizadas, que fueron llamados CsAGO1, CsAGO10 y CsAGO7, respectivamente. Se identificaron dos miembros de la familia AGO en el clado AGO2/3 y se los llamó CsAGO2a y CsAGO2b, estando el último sólo presente en el genoma de CSAP, compartiendo un

93% de identidad de aminoácidos entre ellos. A cuatro proteínas de *C. sinensis* agrupadas en el clado AGO5, y que compartían entre 52 y 87% de identidad de aminoácidos, se las llamó CsAGO5a, CsAGO5b, CsAGO5c y CsAGO5d, destacando que las dos últimas sólo están presentes en el genoma de CSAP. Además, se encontraron cuatro miembros de la familia AGO de naranja agrupados en el clado AGO4/6/8/9; de ellos, se nombraron dos como CsAGO4a y CsAGO4b, mientras que a un tercero se lo llamó CsAGO6, debido a su proximidad con AtAGO6 (Fig. 5.2A). Finalmente, no se pudo agrupar una proteína AGO en el clado AGO4/6/8/9 con otras proteínas caracterizadas en dicho clado y se la llamó CsAGO6-like debido a su proximidad con el grupo AGO6 (Tabla 5.1; Fig. 5.2A).

Se distribuyó a las 5 proteínas CsDCL en cuatro clados: DCL1, DCL2, DCL3 y DCL4 (Fig. 5.2B). Se agrupó a dos proteínas altamente similares con 90% de identidad de aminoácidos en el clado DCL2, y por lo tanto fueron llamadas CsDCL2a y CsDCL2b, respectivamente. Los tres clados restantes contienen una proteína CsDCL cada uno, que consecuentemente se llamaron CsDCL1, CsDCL3 y CsDCL4 (Tabla 5.1, Fig. 5.2B).

Por último, el árbol derivado de las secuencias de CsRDRs también consiste en cuatro clados: RDR1, RDR2, RDR6 y RDR3/4/5. El clado RDR1 contiene tres proteínas: CsRDR1a, CsRDR1b y, sólo presente en el genoma de Phytozome, CsRDR1c (Tabla 5.1, Fig. 5.2C). Dos proteínas fueron ubicadas en el clado RDR6, CsRDR6a y CsRDR6b, ambas similares en longitud, con 1197 y 1145 aminoácidos, respectivamente, y con 80% de su secuencia de aminoácidos compartida. Finalmente, sólo se ubicó una proteína en cada uno de los clados RDR2 y RDR3/4/5, y se las llamó CsRDR2 y CsRDR3, respectivamente (Tabla 5.1, Fig. 5.2C).

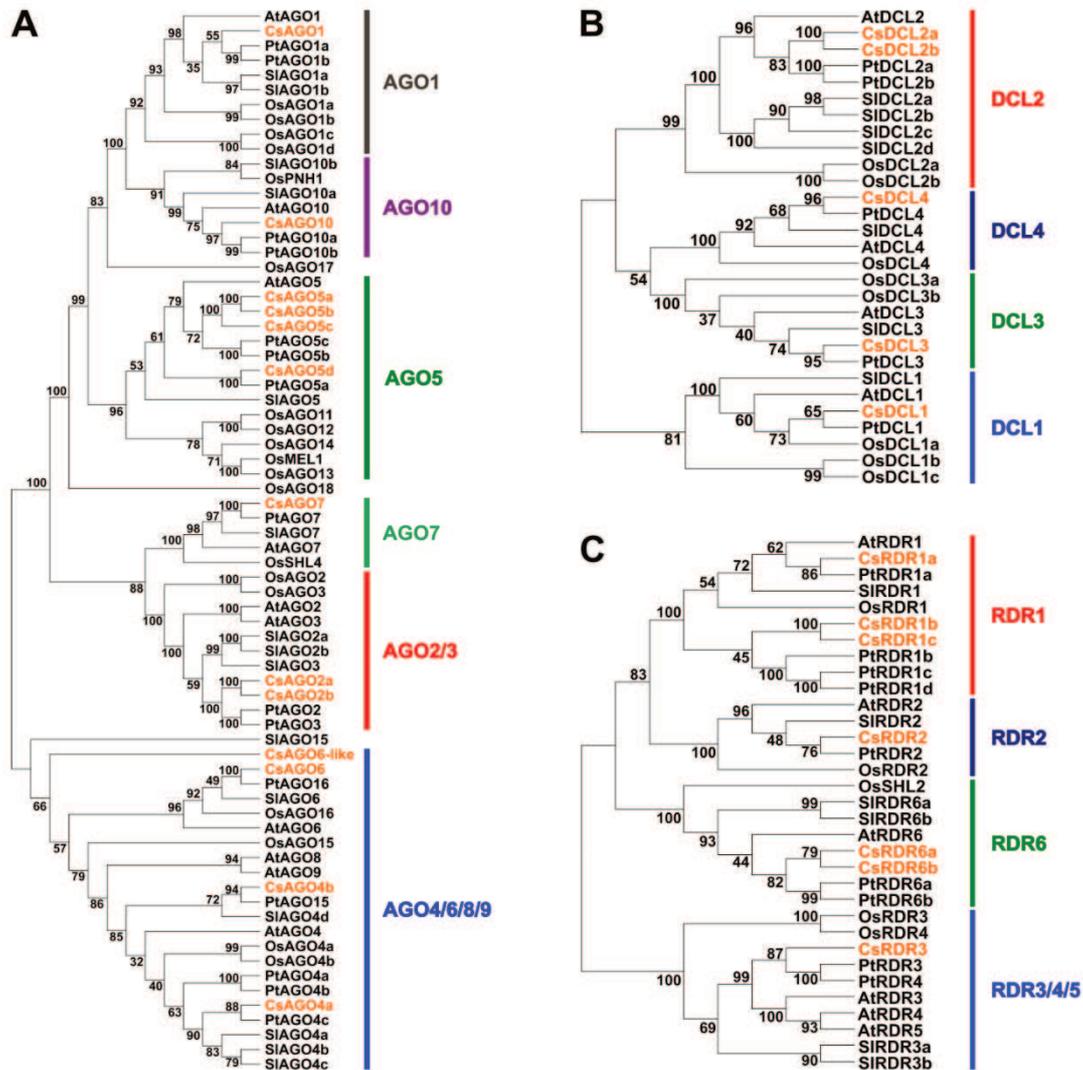


Fig. 5.2: Análisis filogenético de los genes DCL, AGO y RDR de naranjo, Arabidopsis, álamo, tomate y arroz. Se generaron árboles sin raíz con el método *Neighbor-Joining* a partir de alineamientos múltiples de secuencias de proteínas (a) AGO, (b) DCL y (c) RDR usando el algoritmo *Muscle* en MEGA 6.0. Se indican los valores de *bootstrap* de 1000 réplicas. Las familias génicas están divididas por clados, resaltados con diferentes colores. Los genes de *C. sinensis* están coloreados en naranja.

### 5.3.3 - Ubicación cromosómica de los genes DCL, RDR y AGO de naranjo

Luego, se analizó la distribución cromosómica de los genes DCL, RDR y AGO en naranjo utilizando la base de datos de CSAP, que provee ubicación cromosómica (Fig. 5.3). A los genes correspondientes a 4 CsDCL, 5 CsRDR y 13 proteínas AGO se los distribuyó en todos los cromosomas, excepto el cromosoma 8, que no albergó ningún miembro de las familias analizadas. Sólo un gen de la familia AGO está codificado en los cromosomas 3 y 9, CsAGO4b y CsAGO10 respectivamente. Los parálogos CsDCL2a y CsDCL2b aparecen muy cerca en el cromosoma 6, así como CsAGO2a y CsAGO2b en el

cromosoma 2 y CsAGO5a, CsAGO5b y CsAGO5c en el cromosoma 7 (Fig. 5.3; Tabla 5.1). Por el contrario, CsRDR1a y CsRDR1b se ubican en el cromosoma 2 y 5, respectivamente. Los cromosomas 2 y 7 contienen 5 miembros de las familias analizadas cada uno, mientras que los cromosomas 4 y 5, albergan a 3 de ellos. Sólo CsRDR1c no fue anotado en esta base de datos, por lo tanto no está representado en la Fig. 5.3 (Tabla 5.1), mientras que CsDCL1 está ubicado en el “Cromosoma desconocido” (ChrUN, Fig. 5.3; Tabla 5.1), que contiene las secuencias genómicas de *C. sinensis* que todavía no pudieron ser asignadas a ninguno de los cromosomas.

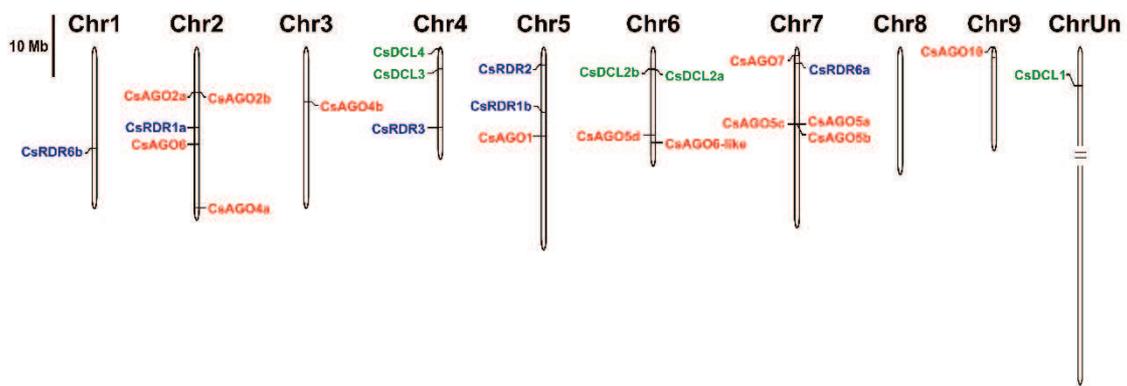


Fig. 5.3: Distribución genómica de los genes CsDCL, CsRDR y CsAGO. Los cromosomas están representados en la escala indicada a la izquierda de la figura. Las diferentes familias génicas están representadas con diferentes colores. Se quitaron 30 mb de “ChrUn” para ajustarlo a la figura.

### 5.3.4 - Patrones de expresión comunes y específicos de genes AGO, DCL y RDR en diferentes tejidos

Para caracterizar los patrones de expresión de diferentes miembros de estas familias, se analizaron datos de *RNA-seq* de 5 tejidos diferentes: hoja, raíz, pulpa y embrión (los números de acceso para los datos crudos se detallan en la Tabla 5.1). Se analizó la expresión en todos los tejidos y se representaron los niveles de expresión de cada gen en un gráfico estilo *heatmap* para cada familia (Fig. 5.4). CsAGO1 y CsAGO4a mostraron los niveles de expresión más altos dentro de la familia AGO y ambos se acumulan a niveles mucho mayores que los del resto de los genes AGO analizados. Para apreciar mejor las diferencias de expresión, en la Fig. 5.4D se muestra el mismo gráfico que la Figura 5.4A, sin estos 2 genes. En embrión se evidenció la expresión más alta para la mayoría de los genes AGO en comparación con el resto de los tejidos analizados, excepto para CsAGO2a, CsAGO2b, CsAGO5a y CsAGO5c, para los que los niveles de

expresión en raíces son levemente más altos (Fig. 5.4A). Por el contrario, se detectó una expresión débil en la mayoría de los genes en pulpa, piel y hoja, mientras que CsAGO2b, CsAGO5b y CsAGO6-like presentaron niveles de expresión muy bajos en todos los tejidos (Fig. 5.4A).

Todos los genes que codifican para proteínas CsDCL están expresados en los 5 tejidos analizados: CsDCL2a y CsDCL2b muestran los niveles de expresión más bajos dentro de esta familia, mientras que CsDCL1 tiene los más altos, especialmente en hoja y pulpa. Por otro lado, CsDCL3 y CsDCL4 presentan expresión intermedia, evidenciando la máxima acumulación en pulpa y embrión, respectivamente (Fig. 5.4B).

En cuanto a la familia RDR, CsRDR1c no fue detectado en ninguno de los tejidos analizados (Fig. 5.4C), CsRDR1b se acumula débilmente en pulpa, cáscara, hoja y raíz y CsRDR1a es el gen más abundante de esta familia en las muestras analizadas. CsRDR2 está altamente expresado en embrión y CsRDR6a se acumula en raíces preferentemente, mientras que CsRDR6b está débilmente expresado en todos los tejidos y muestra acumulación preferencial en pulpa (Fig. 5.4c).

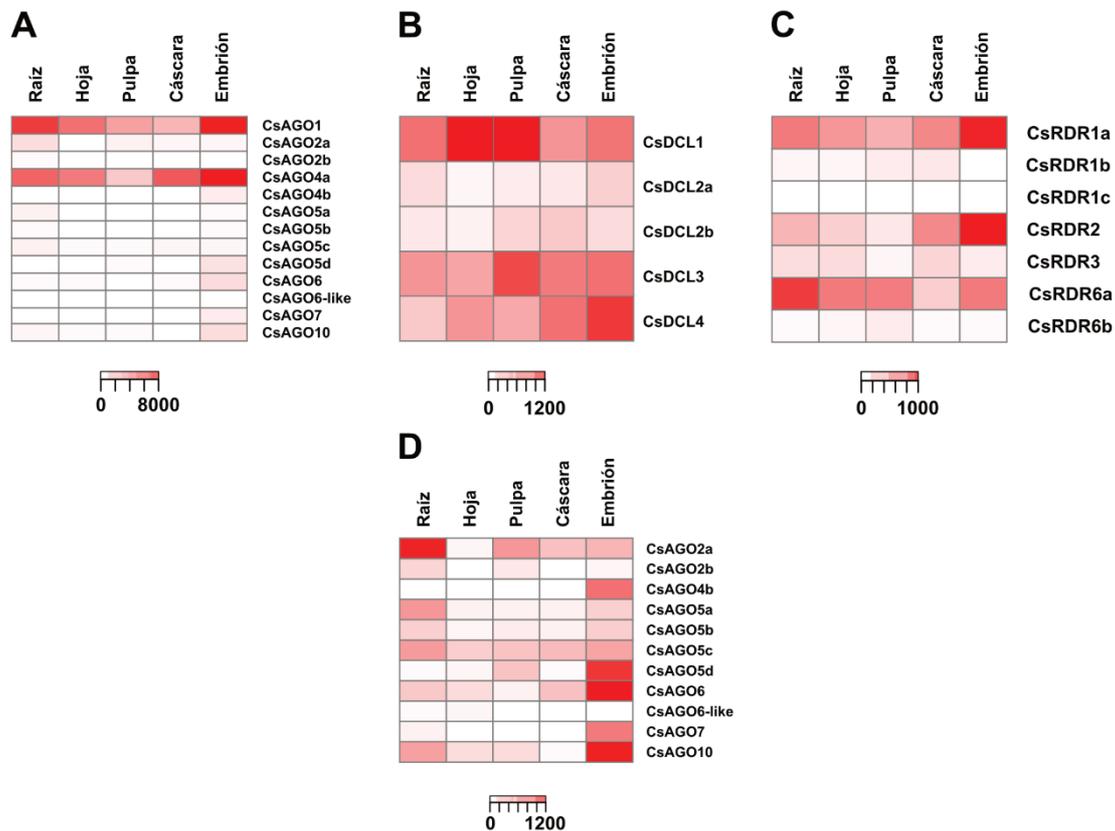


Fig. 5.4: Niveles de expresión de genes (A) CsAGO, (B) CsDCL y (C) CsRDR. Los tejidos analizados fueron raíz, hoja, pulpa, cáscara y embrión. (D) Niveles de expresión de genes CsAGO sin la información de expresión correspondiente a CsAGO1 y CsAGO4. Se representa la expresión normalizada por DESeq2 para cada familia génica con diferente intensidad de color, de blanco a rojo.

### 5.3.5 - La expresión de genes involucrados en RdDM y vías de tasiARN está regulada negativamente durante la abscisión de los frutos de naranjo.

Para entender mejor el rol de las vías de silenciamiento por ARN durante el proceso de abscisión, se identificaron todas las sondas correspondientes a factores relacionados con ARNs pequeños presentes en el microarreglo de ADNc utilizado por Merelo y colaboradores (2017), incluyendo los nuevos genes AGO, RDR Y DCL caracterizados junto con otros genes que codifican para factores de copia simple involucrados en la biogénesis y acción de ARNs pequeños (Fig. 5.5; Tabla 5.5).

Debido a que los datos obtenidos por análisis de expresión usando microarreglos son típicamente validados usando la técnica de qRT-PCR, se replicó el experimento realizado por Merelo y col. (2017), usando frutos de la variedad Salustiana y se realizó un tratamiento con etileno exógeno para inducir el proceso de abscisión, para luego obtener muestras equivalentes a las utilizadas para el análisis con microarreglos (Fig. 5.6).

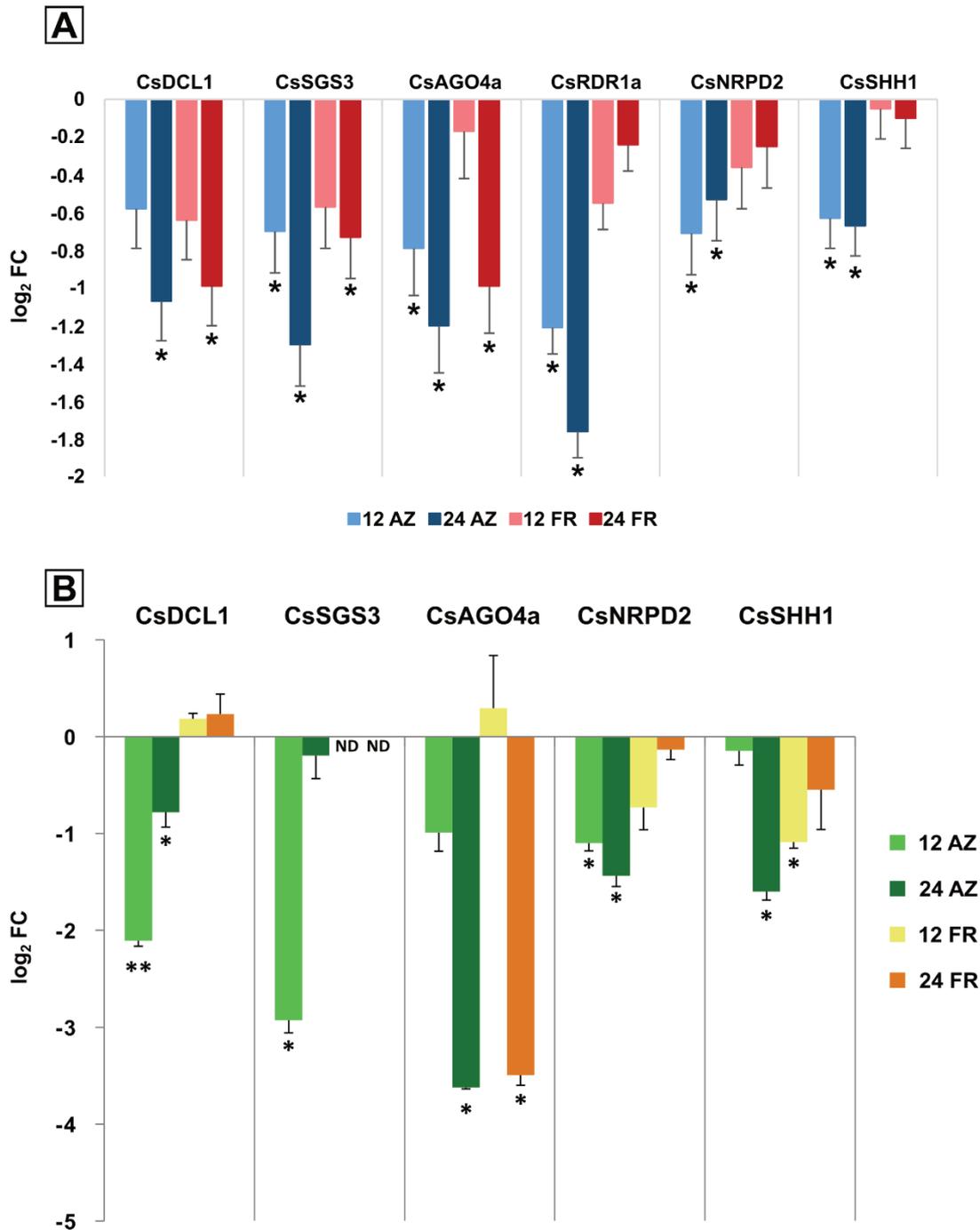


Fig. 5.5: Expresión relativa de genes involucrados en el silenciamiento de ARN. (A) Comparaciones entre la zona de abscisión (AZ) y la corteza del fruto (FR) luego de 12 y 24 horas del tratamiento con etileno. Los valores se representan relativos a la expresión en AZ o FR sin tratamiento con etileno (0 horas). La expresión diferencial significativa se indica con un asterisco (\* =  $q \leq 0.05$  y un valor de contraste de  $\log_2$  FC de  $\pm 0.5$ ). (B) Niveles de expresión de cada transcripto determinados por qRT-PCR (media  $\pm$  DE;  $n = 3$ ) comparada a la expresión de cada gen en AZ o FR sin tratamiento con etileno (0 h) y usando CsUBI como control de normalización (\*  $p < 0.05$ ; \*\*  $p$ -value  $< 0.001$ ).

**Tabla 5.5: Anotaciones correspondientes a las sondas de microarray utilizadas**

ID Microarray	ID Phytozome	Anotación	Expresión diferencial en datos de microarray
aKN0AAI3BB02FM1_c	orange1.1g000174m	CsDCL1	Sí
aCL5658Contig1	orange1.1g002586m	CsRDR1a	Sí
aCL6327Contig1	orange1.1g001771m	CsRDR3	No
aCL4320Contig1	orange1.1g001466m	CsAGO1	Sí
aCL5782Contig1	orange1.1g002449m	CsAGO4a	Sí
aC34102E06EF_c	orange1.1g003129m	CsAGO4a	Sí
aIC0AAA29DG08RM1_c	orange1.1g002636m	CsAGO4b	No
aIC0AAA58DD01RM1_c	orange1.1g002204m	CsAGO5a	No
aIC0AAA44CC08RM1_c	orange1.1g001684m	CsAGO7	No
aCL8362Contig1	orange1.1g005428m	CsAGO10	No
aC32104F12EF_c	orange1.1g001954m	CsAGO10	No
aCL1403Contig1	orange1.1g006189m	CsSGS3	Sí
aC06017E10SK_c	orange1.1g026004m	CsSHH1	Sí
aCL2100Contig1	orange1.1g001050m	CsNRPD2	Sí
aKN0AAP5YG14FM1_c	orange1.1g001050m	CsNRPD2	No



Fig. 5.6: Tratamiento con etileno de frutas de naranja Salustiana.

Inicialmente se analizaron datos de expresión de la zona de abscisión (AZ) en comparación con la corteza del fruto (FR), obtenidos usando LCM, luego de 12 o 24 h de tratamiento con etileno para inducir abscisión (Merelo et al., 2017). En base al análisis utilizando microarreglos, se detectó regulación negativa en los niveles de expresión de 3 miembros de las familias analizadas: CsDCL1, CsRDR1a y CsAGO4a, así como en 3 genes de copia simple: el ortólogo de SUPPRESSOR OF GENE SILENCING 3 (CsSGS3), un componente de la vía de biogénesis de tasiARN; el ortólogo de naranja de la subunidad 2 de PolIV y PolV (CsNRPD2) y el ortólogo de SAWADEE HOMEODOMAIN HOMOLOGUE 1 (CsSHH1), que está involucrado en la vía de RdDM (Tabla 5.5). La regulación negativa de estos genes también fue detectada por qRT-PCR, excepto para CsRDR1 que no fue evaluado usando este método (Fig. 5.5). De estos, CsRDR1a, CsSHH1 y CsNRPD2 están específicamente regulados negativamente en la AZ, pero no en FR, tanto a 12 como a 24 h después del tratamiento con etileno (Fig. 5.5). CsSGS3 y CsAGO4a están regulados negativamente en AZ a las 12 h, pero a las 24 h, también lo están en FR. Finalmente, CsDCL1 está regulado negativamente tanto en AZ como en FR luego de 24 h de tratamiento con etileno (Fig. 5.5). Posteriormente se realizó el análisis de expresión usando qRT-PCR, lo cual permitió confirmar algunos de estos cambios de expresión y se ratificó la tendencia general de regulación negativa de estos genes durante la abscisión (Fig. 5.5B,). Estos análisis muestran que se necesita una regulación negativa de varios factores involucrados en la biogénesis de ARNs pequeños durante la abscisión de frutos, lo que implica un papel aún sin explorar para los ARNs pequeños en este importante proceso.

## 5.4 - Discusión

El silenciamiento por ARN es un mecanismo molecular antiguo que supone la participación de ARNs pequeños actuando a diferentes niveles en la regulación de la expresión génica. Las plantas han desarrollado una amplia variedad de vías sobre diferentes combinaciones de RDR-DCL-AGO, proteínas que han demostrado participar en varios aspectos del desarrollo vegetal (Bologna & Voinnet, 2014; Borges & Martienssen, 2015). Por ejemplo, DCL1 y AGO1 están principalmente involucrados en la vía de miARNs, pero no se necesitan proteínas RDR para la biogénesis de miARNs (Bologna & Voinnet, 2014; Y. Yu et al., 2017). Algunos miARNs excepcionales han

desarrollado modos de acción especializados, como miR166 que se une exclusivamente a AGO10 para regular el desarrollo apical del tallo y la también exclusiva combinación miR390/AGO7 en Arabidopsis para regular los *AUXIN RESPONSE FACTORS 3 y 4 (ARF3 y ARF4)*, (Montgomery et al., 2008; Hongliang Zhu et al., 2011). RDR6, SGS3, AGO1 y DCL4 son los principales componentes de la vía de tasiARNs, responsable de la regulación negativa de ARNm blanco, a nivel post-transcripcional en múltiples especies de plantas estudiadas hasta ahora (Dotto et al., 2014; Fahlgren et al., 2006; Yifhar et al., 2012); mientras que RDR2, AGO4 y DCL3 participan en la vía RdDM, principalmente implicada en el silenciamiento de regiones repetitivas en el genoma (J. a Law & Jacobsen, 2010; W. Liu et al., 2018; F. Wang & Axtell, 2017).

En este trabajo se identificaron por primera vez a nivel genómico diferentes miembros de las familias AGO, RDR y DCL de naranjo, además de otros componentes de copia única que participan en rutas regulatorias que involucran ARNs pequeños.

#### **5.4.1 - La familia RDR en Citrus sinensis presenta características distintivas y patrones de expresión específicos para cada uno**

Las proteínas RDRs se caracterizan por tener un dominio conservado requerido para copiar ARN simple hebra en ARN doble hebra, que se detectó en todos los miembros de esta familia identificados en este trabajo (Fig. 5.1). En estas proteínas también está presente un motivo DLDGD conservado en el sitio catalítico, y fue detectado en 5 CsRDRs. Dentro de este motivo, la lisina suele variar, y fue el caso para CsRDR3, que presentó un motivo DFDGD, también detectado en ortólogos de otras especies, como Arabidopsis, arroz, pimiento y café (Fernandes-Brum et al., 2017; Qin et al., 2017; Wassenegger & Krczal, 2006), mientras que CsRDR1c presentó un motivo DQDGE, con dos sustituciones de aminoácidos, para la que aún resta evaluar su actividad funcional (*Additional file 6*; Sabbione et al., 2019). La función biológica de las proteínas RDR usualmente está ligada a la posterior acción de proteínas DCL específicas y se ha visto en genomas eucariotas, incluyendo plantas, hongos y animales invertebrados, pero no en vertebrados e insectos (Zong et al., 2009). Los diferentes miembros de la familia RDR muestran diversificación funcional que usualmente está conservada entre especies; muestran preferencia de tejidos, sugiriendo que podrían tener roles bajo circunstancias

específicas en el crecimiento y desarrollo de plantas o que su expresión podría estar inducida en respuesta a señales ambientales específicas.

#### 5.4.2 - Los miembros de la familia DCL en *Citrus sinensis* presentan diferentes patrones de expresión

Las endonucleasas DCL procesan dsARN en pequeños dúplex de ARN con 2 nucleótidos 3' sobresalientes. DCL1 es capaz de reconocer sustratos que corresponden a transcritos pre-miARN que pueden formar estructuras de tallo y burbuja (*stem-loop*) imperfectos, mientras que el resto de las DCLs encontradas en plantas son responsables de la producción de siARNs de 21, 22 y 24 nucleótidos (Fukudome & Fukuhara, 2017). Los genes DCL de plantas de un grupo monofilético aparecieron luego de la separación de plantas y animales pero antes de la divergencia entre mono y dicotiledóneas hace 150 millones de años (Henderson et al., 2006). Todos los genes identificados en este estudio contienen los dominios DEXDc, Helicase C, *RNA binding* y dos dominios RIBOc en tándem, mientras que CsDCL1 y CsDCL4 mostraron dos dominios dsRB adicionales (Fig. 5.1). Además, todos presentaron el característico motivo DECH (*Additional file 6*; Sabbione et al., 2019). El genoma de *C. sinensis* presenta un miembro de cada clado DCL1, DCL3 y DCL4, similar a *Arabidopsis* y tomate, pero sorprendentemente 2 miembros del clado DCL2, que también es el caso de las proteínas OsDCL2. Las proteínas CsDCL2, que se presume que son responsables de la generación de siARNs de 22-nt basados en sus equivalentes en *Arabidopsis*, parecen haber sufrido duplicación en tándem considerando la posición cercana de CsDCL2a y CsDCL2b en el cromosoma 6 y el 90% de identidad en su secuencia de aminoácidos (Fig. 5.2 y 5.3). Los diferentes patrones de expresión detectados para estos dos genes sugieren que adquirieron diferente regulación de su expresión y podrían tener papeles preferenciales en tejidos específicos. CsDCL1 mostró expresión en todos los tejidos analizados, lo que es consistente con su rol putativo en la biogénesis de miARN (Bologna & Voinnet, 2014; Y. Yu et al., 2017). Del mismo modo, CsDCL3 mostró la expresión más alta en pulpa y CsDCL4 en embrión, pero están expresados abundantemente en todas las muestras analizadas, lo que también es coherente con su papel ubicuo en RdDM y la biogénesis de tasiARNs, respectivamente (Borges & Martienssen, 2015).

### 5.4.3 - La familia AGO en *Citrus sinensis* sufrió eventos de duplicación en tándem durante la evolución

Las proteínas AGO son los principales efectores del silenciamiento de ARN, ya que poseen la actividad de corte necesaria para las vías regulatorias mediadas por ARNs pequeños. La familia AGO se expandió durante la evolución de las plantas, de algas verdes ancestrales unicelulares o multicelulares (ej. *Micromonas pusilla* y *Volvox carteri*), donde están presentes 3 genes AGO o menos, a 10 o más miembros en plantas con flores. La expansión de la familia AGO en plantas sugiere una diversificación funcional de las proteínas AGO presumiblemente debida a la expansión de vías regulatorias dirigidas por ARNs pequeños (H. Zhang et al., 2015). Las proteínas AGO de eucariotas contienen 4 dominios principales: un dominio N-terminal variable y los dominios altamente conservados PAZ, MID y PIWI (Bologna & Voinnet, 2014). En este trabajo se identificaron estos dominios en todas las proteínas AGO caracterizadas, con la excepción de CsAGO7 que carece del dominio MID (Fig. 5.1), que ha sido implicado en el ordenamiento de ARNs pequeños en diferentes AGOs (Frank et al., 2012). Se ha demostrado en *Arabidopsis* que AGO7 une exclusivamente a miR390 (Montgomery et al., 2008), pero las bases de esta asociación única entre AGO7 y miR390 no se conocen en su totalidad y aún no han sido estudiadas en *C. sinensis*. También se analizó la presencia de motivos característicos DDH/H y DDD/H en estas proteínas, que están presentes en las proteínas AGO de *Arabidopsis* y sus funciones han sido exhaustivamente estudiadas, estableciendo que son necesarios para la actividad de corte de las proteínas AGO (Arribas-Hernández et al., 2016; Baumberger & Baulcombe, 2005; Carbonell et al., 2012). Se detectaron algunas sustituciones de aminoácidos en otras especies como tomate y brasicáceas, incluyendo los motivos DDH/P, DDH/S, DDY/H y DDY/P detectados en las proteínas CsAGO identificadas en este estudio (Bai et al., 2012; J. Y. Cao et al., 2016). Sin embargo, CsAGO6-like no presentó este motivo, lo que sugiere que este miembro de la familia CsAGO podría carecer de actividad de corte (*additional file 6*; Sabbione et al., 2019).

Se identificaron copias únicas de CsAGO1, CsAGO6, CsAGO7 y CsAGO10, mientras que se detectaron 2 alelos altamente similares de CsAGO2 (93% de identidad de aminoácidos), posiblemente originados por duplicación en tándem basada en su

cercana ubicación en el cromosoma 2 (Fig. 5.3; Tabla 5.1). De forma similar, 3 parálogos de CsAGO5 se detectaron en el genoma de naranjo, en ubicaciones cercanas del cromosoma 7, lo que también sugiere que hubo eventos de duplicación que dieron lugar a estos alelos que comparten un 90% de identidad de aminoácidos (Fig. 5.3, Tabla 5.1). CsAGO5a y CsAGO5b tienen una expresión levemente más alta en raíces, mientras que CsAGO5b tiene baja expresión en todas las muestras analizadas, lo que sugiere nueva funcionalización del alelo tras eventos de duplicación (Fig. 5.4). CsAGO1 muestra alta acumulación en la mayoría de los tejidos como se esperaba por su papel putativo como efector de la regulación mediada por miARNs. CsAGO2 se expresa preferentemente en raíces, mientras que CsAGO4a, CsAGO5d, CsAGO6 y CsAGO10 se acumulan mayormente en embrión. Esto podría estar relacionado a una necesidad de CsAGO4 y CsAGO6, involucradas en la vía de RdDM durante el desarrollo embrionario, ya que esta vía es de particular importancia durante este proceso y otros, como gametogénesis y meiosis, donde los hc-siARNs son esenciales para silenciar las regiones repetitivas del genoma (Borges & Martienssen, 2015); así como un rol putativo de CsAGO10 análogo al demostrado de AtAGO10 en desarrollo embrionario (Moussian et al., 1998; M. R. Tucker et al., 2008). De forma similar a CsAGO5b, se observó muy baja expresión para CsAGO7 y CsAGO4b (Fig. 5.4; *Additional file 8*; Sabbione et al., 2019); lo que podría indicar un papel preferencial de los productos de estos genes en otros tejidos y/o etapas de desarrollo y/o durante exposición a condiciones ambientales específicas.

Es interesante observar que se detectó la co-expresión de factores involucrados en las mismas vías regulatorias en nuestros análisis de expresión. Por ejemplo, para biogénesis y acción de miARNs tanto CsDCL1 como CsAGO1 son necesarias, y se detectó co-expresión de estos genes (Fig. 5.4). De manera similar, RdDM requiere de la acción coordinada de CsRDR2, CsDCL3 y CsAGO4 y la vía de tasiARNs necesita CsRDR6, CsDCL4 y CsAGO1, y todos ellos mostraron co-expresión en los análisis realizados (Fig. 5.4). Sin embargo, también es importante tener en cuenta que la función biológica de los ARNs pequeños generados por estas vías es típicamente necesaria en la mayoría de las plantas durante el crecimiento y el desarrollo. Por ejemplo, los miARNs son reguladores clave de varios factores de transcripción involucrados en el desarrollo de hojas y frutos (X. Chen, 2005; José Ripoll et al., 2015; Karlova et al., 2013; Y. Yu et al., 2017), así como los tasiARNs derivados de TAS3, que regulan la respuesta a auxinas a través de la escisión

de los transcriptos ARF (Dotto et al., 2014; Fahlgren et al., 2006; Yifhar et al., 2012), y el silenciamiento génico a nivel transcripcional que también es regulado por componentes silenciadores de ARN (Djupedal & Ekwall, 2009; Nuthikattu et al., 2013; Simon & Meyers, 2011). Por lo tanto, diferentes combinaciones de estos factores están generalmente activas de forma ubicua. Es clave dilucidar de qué forma son necesarias las diferentes combinaciones de factores silenciadores de ARN para entender cómo la expansión de los miembros de las familias génicas AGO, RDR y DCL ha contribuido a su funcionalización única en diferentes especies. También es necesario obtener información de la localización subcelular de los diferentes miembros de estas familias, que se describió bien para *Arabidopsis* y tomate, por ejemplo, pero se sabe muy poco en otras especies (Poulsen et al., 2013; Xie et al., 2004).

#### 5.4.4 - La vía de RdDM está regulada negativamente durante la abscisión de frutos

En cítricos, la abscisión del fruto representa un alto porcentaje de las pérdidas de rendimiento anuales. Se sabe que las hormonas de crecimiento de plantas están profundamente involucradas en la abscisión, y de ellas, se cree que el etileno es su regulador natural (Jackson & Osborne, 1970). Este proceso ocurre específicamente en la zona de abscisión (AZ) de frutas y otros órganos, a través de cambios coordinados en la expresión génica y es acelerado por el tratamiento con etileno (Agustí et al., 2008, 2009; Merelo et al., 2017). El análisis de expresión génica en abscisión de fruto de naranjo realizado en este trabajo se orientó a analizar cambios en la expresión de genes relacionados al silenciamiento de ARN, en un sistema modelo utilizado con anterioridad para analizar mecanismos subyacentes al proceso de abscisión inducidos por tratamiento con etileno (Agustí et al., 2008, 2009; Merelo et al., 2017). Por eso, se comparó la expresión de genes AGO, RDR, DCL y genes de copia simple adicionales con homología a factores de ARNi conocidos en la zona de abscisión de las frutas (Fig. 5.5, *Additional file 10*; Sabbione et al., 2019). Se detectó una regulación negativa general de los factores de ARNi. En particular, se vieron bajos niveles de expresión para CsAGO4a, CsSHH1 y CsNRPD2 en la AZ. Estos genes participan en la vía RdDM, que es una importante vía epigenética mediada por ARNi en plantas. La vía RdDM está involucrada en el silenciamiento transcripcional de transposones y secuencias repetitivas y se basa

en maquinaria transcripcional específica que incluye las ARN polimerasas específicas de plantas Pol IV y Pol V. Los transcriptos de Pol IV son rápidamente procesados en dsARNs por RDR2, y subsecuentemente procesados en siARNs de 24 nucleótidos por DCL3 y exportados al citoplasma. En el citoplasma son principalmente incorporados a complejos que contienen a AGO4 e importados nuevamente al núcleo para marcar transcriptos nacientes generados por Pol V en los mismos *loci*, dando lugar a la metilación del ADN (Borges & Martienssen, 2015). El silenciamiento génico también conlleva la remodelación de la cromatina, típicamente a través de la metilación de la histona H3. El papel de SHH1 en *A. thaliana* es reconocer H3K9me2 (dimetilación de la lisina 9 en la histona H3), que recluta a Pol IV e inicia la biogénesis de siARNs para el mantenimiento del silenciamiento génico (J. A. Law et al., 2013). Además de la vía de RdDM dirigida por siARNs de 24-nt, se demostró en Arabidopsis que los tasiARNs de 21-nt también pueden dirigir el proceso de RdDM. En lugar de ser procesados por DCL4 y cargados en AGO1 para marcar el corte de los mARN en trans, estos tasiARNs son procesados por DCL1 y se cargan en AtAGO4 o AtAGO6, para dirigir la metilación de los *loci* TAS (L. Wu et al., 2012). En este sentido, la regulación negativa observada para CsDCL1 y CsSGS3 en FR o AZ podría contribuir a la inhibición general de la metilación del ADN durante el proceso de abscisión, como resultado de una menor expresión de estos componentes de la vía RdDM relacionada a genes TAS y a la biogénesis de tasiARN. Los roles de RDR1 y RDR6 usualmente son redundantes cuando están involucrados en defensa antiviral y ambas proteínas producen sustratos de DCL4 (Pumplin & Voinnet, 2013). Por lo tanto, se puede proponer un papel para CsRDR1 en el contexto de la biogénesis de tasiARNs y/o RdDM relacionada con *loci* TAS durante el proceso de abscisión, en lugar de CsRDR6 para la que no se observó expresión diferencial en las muestras analizadas. Estos cambios en la expresión de factores de silenciamiento por ARNi podrían llevar a modificaciones en la expresión de genes TAS y otras regiones genómicas reguladas por las vías de RdDM y de biogénesis de tasiARNs, presumiblemente involucradas en procesos celulares que participan en la abscisión de frutos, como el desensamblaje de la pared celular, aunque se requieren experimentos adicionales para lograr un mejor entendimiento de los procesos moleculares que ocurren durante la abscisión de frutos y su relación con los distintos tipos de ARNs pequeños (Merelo et al., 2017).

## 5.5 - Conclusiones

- Se identificaron y caracterizaron 13 genes AGO, 5 DCL y 7 RDR presentes en el genoma de naranjo.
- La mayoría de estos genes se expresan ubicuamente pero muestran diferentes niveles de expresión en diferentes tejidos.
- Se mejoró la anotación de 4 de estos genes (CsRDR6b, CsAGO5b, CsAGO5c y CsAGO6-like) usando datos de *RNA-seq*.
- Se determinó que CsDCL1, CsAGO4a, CsRDR1, CsSHH1, CsNRPD2 y CsAGO4a, mostraron menores niveles de expresión en la zona de abscisión del fruto de las muestras de naranja, sugiriendo que la vía de RdDM juega un papel importante en el proceso de abscisión de frutos.

## Capítulo 6: Discusión y conclusiones generales

## 6.1 - ARNs pequeños como reguladores de aspectos de interés comercial en frutos

Los frutos de las especies analizadas en este trabajo de tesis se encuentran entre los cultivos frutihortícolas más importantes para el consumo interno de nuestro país, por lo que es muy importante que no sufran deterioro en el proceso de producción y comercialización, y que desarrollen propiedades organolépticas óptimas para el agrado del consumidor. La selección de especies a analizar se basó en la posibilidad de obtener información sobre rutas regulatorias mediadas por ARNs pequeños en frutos con características organolépticas muy distintas, y también con diferentes patrones de preferencia por parte de los consumidores. De esta manera, los resultados obtenidos permiten abarcar un panorama amplio a la hora de estudiar la influencia de ARNs pequeños en el desarrollo de atributos cualitativos de importancia comercial y que pueden ser aprovechados en futuras instancias de mejoramiento a través de técnicas de ingeniería genética.

La calidad nutricional y el atractivo sensorial de los frutos están determinados en gran medida por el contenido de metabolitos secundarios, incluyendo los productos de la ruta de los fenilpropanoides y de los terpenoides. Al momento de iniciarse esta tesis doctoral, existían reportes que vinculaban dos tipos de ARNs pequeños con la regulación de la ruta de los fenilpropanoides: miRNAs y tasiRNAs. A través de la acción de miR159, miR858 y miR828 que da origen a tasiARNs procedentes del *locus* TAS4, estos ARNs pequeños regulan la expresión de factores de transcripción de la familia MYB (Rajagopalan et al., 2006; Xia et al., 2012). Esto estaría vinculado directamente con aspectos relacionados a diversos atributos de interés comercial, ya que estos factores de transcripción regulan en varios puntos a la vía de compuestos fenilpropanoides, vinculados a la protección a la luz UV, defensa frente a patógenos y desarrollo de pigmentos florales y compuestos aromáticos (Lin-Wang et al., 2010).

Los miARNs miR828 y miR858 se han identificado en diversos tejidos de *Arabidopsis* como tallos, flores y silicuas (A. Sharma et al., 2020; F. Yang et al., 2013), en fibras de algodón (Guan et al., 2014), en hojas y frutos de uva (Tirumalai et al., 2019), en hojas y frutos de tomate (X. Jia et al., 2015) y también en cáscara y pulpa de papa (Bonar

et al., 2018). Sin embargo, nuestros análisis no detectaron expresión de miR828 en ninguna de las bibliotecas de frutos analizadas, aunque se encontró a la familia miR858 en frutilla y en naranja, pero en esta última especie no se validó ningún target para *csi-miR858*. El papel que juega este miRNA en frutos de frutilla aún no se conoce en detalle, pero es probable que participe en el desarrollo de color de estos frutos, ya que el mismo está determinado por el contenido de antocianinas, y la biosíntesis de estos metabolitos secundarios depende de factores de transcripción de la familia MYB (Lin-Wang et al., 2010). La ausencia de miR858 en frutos de tomate y de targets en naranjo, es consistente con el hecho de que la acumulación de antocianinas no es muy importante en estos dos frutos, al igual que la ausencia de miR828. No obstante, sí logramos evidenciar la expresión de miR159 en las 3 especies estudiadas y del phasiARN L1\_S\_81 en tomate, cuyo gen blanco de regulación también es un factor de transcripción del tipo MYB, sugiriendo que la expresión de estos factores de transcripción también estaría regulada por phasiARNs, aunque resta comprender cuáles son los procesos biológicos sobre los que influyen estos módulos regulatorios. También resulta interesante que identificamos que el gen regulado por miR159 en frutos de naranja no es un factor de transcripción de esta familia, sino que en esta especie este miARN regularía un transcripto que codifica para *DNA-damage-repair/toleration protein* (Tabla 4.3), lo que constituye una diversificación de la red regulatoria de miR159 en esta especie.

Se destaca la detección de phasiARNs en frutos de frutilla y de tomate con la capacidad de regular la expresión de genes que codifican para enzimas de la pared celular, ya que en estas especies, el rápido ablandamiento de los frutos luego de cosechados representa un serio problema. En esta tesis identificamos a los phasiARNs L1\_S\_99 de tomate y L12\_S\_5 de frutilla, que regulan sendos transcriptos que codifican para una Poligalacturonasa y una Xiloglucano-endo-transglicosilasa, respectivamente. Hasta el momento no existían reportes que relacionen la regulación mediada por ARNs pequeños de la expresión de genes que codifican para enzimas de pared celular, abriendo nuevas posibilidades de estudio a través de modificación por ingeniería genética para mejorar la vida poscosecha de los frutos de estas especies.

Es notable que los análisis realizados con muestras de frutos de tomate permitieron la identificación de un gran número de transcriptos regulados por

phasiARNs que codifican para proteínas relacionadas con el desarrollo de propiedades organolépticas en esta especie (Tabla 2.5), mientras que en las otras dos especies se identificó un número menor. La caracterización en detalle de varios de los transcriptos de frutilla y de naranja se vio dificultada por la falta de buenas anotaciones funcionales para muchos de ellos. Es probable que futuras versiones de los genomas de frutilla y de naranja con mejoras en las anotaciones de sus transcriptos, permitan la identificación de otros genes regulados por ARNs pequeños con importancia en características de interés comercial para estos frutos.

## 6.2 - Regulación mediada por ARNs pequeños de otras cualidades de importancia comerciales

Desde un punto de vista comercial, es deseable que los frutos, además de tener propiedades organolépticas destacables, demuestren resistencia frente a los distintos tipos de estrés a los que pueden estar sometidos durante su producción y comercialización. Al analizar el conjunto de genes target de phasiARNs de frutilla, detectamos enriquecimiento en los términos de Ontología Génica para genes vinculados a estrés biótico y abiótico. A su vez, identificamos phasiARNs que regulan factores de transcripción del tipo GRAS en tomate, los cuales están relacionados con la defensa frente a diversos tipos de estrés (Niu et al., 2017). Por lo tanto, nuestros análisis también aportan información sobre la existencia de rutas regulatorias que involucran ARNs pequeños que podrían utilizarse para mejorar la respuesta de estos cultivos a estos tipos de estrés.

Por otro lado, también se destaca la caracterización de las principales familias involucradas en la biosíntesis y acción de ARNs pequeños en *C. sinensis* (Sabbione et al., 2019), lo cual permitió relacionar por primera vez a estas moléculas regulatorias con el proceso de abscisión y la ruta de RdDM en particular. Como se mencionó en el capítulo 5, el proceso de abscisión de cítricos es responsable de cuantiosas pérdidas económicas para este cultivo de gran importancia agronómica a nivel mundial, por lo que este trabajo constituye una puerta de entrada para profundizar en la regulación a nivel molecular del proceso de abscisión, a fin de desarrollar mecanismos que permitan controlar este proceso y minimizar las pérdidas de rendimiento.

## 6.3 - Perspectivas a futuro

La obtención de cultivares con contenidos de metabolitos secundarios optimizados desde el punto de vista nutricional es cada vez más explorada, en particular a través del estudio de las redes regulatorias que modulan el metabolismo de estos compuestos (Butelli et al., 2008; Patra et al., 2013). Son numerosas las posibles aplicaciones biotecnológicas que pueden lograrse con el conocimiento de las rutas regulatorias mediadas por ARNs pequeños que fueron identificadas en nuestro trabajo. Su estudio abre la puerta a la obtención de variedades con mejoras desde el punto de vista comercial, como cultivares con mayor firmeza, vida poscosecha incrementada o con ciertas propiedades organolépticas resaltadas. Por ejemplo, el uso de construcciones que sobreexpresen L12\_S\_5 (regulador de Fa-XET6) en frutillas, posibilitaría el desarrollo de frutos con una menor pérdida de firmeza en el período poscosecha, o tal vez podrían obtenerse frutos de esta especie con colores rojos más intensos al modificar los niveles de fan-miR858.

También resulta de interés analizar las cualidades organolépticas de frutos mutantes en el gen SIDCL2b, los cuales fueron generados por Wang y colaboradores (Zhengming Wang et al., 2018), ya que en nuestro trabajo identificamos 5 *loci* PHAS que comparten ubicación genómica con este gen. En particular, estos autores reportaron una línea mutante con una delección cercana al *locus 1*, por lo que podría tener influencia en la expresión de los phasiARNs producidos a partir de este *locus*. Estas plantas mutantes podrían ser utilizadas para evaluar el desarrollo y la maduración de frutos, ya que dicho reporte centró sus estudios en phasiARNs presentes en hojas de tomate, relacionados con la defensa ante infecciones virales.

Desde el punto de vista del mejoramiento en aroma y sabor, nuestra identificación del módulo regulatorio L36\_AS\_16/Naringenina 2-Oxoglutarato 3-Dioxigenasa en naranjo, puede ser una interesante opción para estudios orientados a la caracterización de la influencia de esta enzima en el desarrollo de estos atributos. En el caso de la planta de naranjo, donde las modificaciones usando ingeniería genética implican tiempos largos y técnicas complejas, sería interesante analizar posibles tecnologías que incluyan el uso de uno o un conjunto de ARNs pequeños en aplicaciones

exógenas sobre los frutos (Dalakouras et al., 2020; Dubrovina & Kiselev, 2019), ya sea mientras se encuentran aún en la planta o durante su poscosecha, teniendo en cuenta el marco regulatorio legal en todos los casos. Las aplicaciones de este tipo en frutos que aún están en la planta, también podrían implementarse para desarrollar aplicaciones basadas en la relación entre el proceso de abscisión y la ruta de RdDM. Esta temática tiene mucho potencial de estudio debido a la escasa información existente y a la importancia comercial del proceso de abscisión en esta especie. Aparece así como un desafío la implementación de estrategias que permitan retrasar la abscisión de frutos en especies de cítricos aplicando un enfoque molecular donde los ARNs pequeños sean protagonistas.

La información generada en esta tesis doctoral será un importante recurso para la planificación y ejecución de estudios experimentales orientados a la caracterización de los módulos regulatorios identificados aquí, para acercarlos a desarrollos tecnológicos que contribuyan a mejorar aspectos cualitativos de estos frutos, y posiblemente también de frutos de otras especies.

# Bibliografía

- Addo-Quaye, C., Miller, W., & Axtell, M. J. (2009). CleaveLand: A pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics*, 25(1), 130–131. <https://doi.org/10.1093/bioinformatics/btn604>
- Aguiar-Martínez, J. A., Poza-Carrión, C., & Cubas, P. (2007). Arabidopsis Branched1 acts as an integrator of branching signals within axillary buds. *Plant Cell*, 19(2), 458–472. <https://doi.org/10.1105/tpc.106.048934>
- Agustí, J., Merelo, P., Cercás, M., Tadeo, F. R., & Talán, M. (2009). Comparative transcriptional survey between laser-microdissected cells from laminar abscission zone and petiolar cortical tissue during ethylene-promoted abscission in citrus leaves. *BMC Plant Biology*, 9, 1–20. <https://doi.org/10.1186/1471-2229-9-127>
- Agustí, J., Merelo, P., Cercós, M., Tadeo, F. R., & Talón, M. (2008). Ethylene-induced differential gene expression during abscission of citrus leaves. *Journal of Experimental Botany*, 59(10), 2717–2733. <https://doi.org/10.1093/jxb/ern138>
- Aharoni, A., De Vos, C. H. R., Wein, M., Sun, Z., Greco, R., Kroon, A., Mol, J. N. M., & O'Connell, A. P. (2001). The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant Journal*, 28(3), 319–332. <https://doi.org/10.1046/j.1365-313X.2001.01154.x>
- Alam, T., & Goyal, G. K. (2007). Packaging and storage of tomato puree and paste. *Stewart Postharvest Review*, 3(5), 1–8. <https://doi.org/10.2212/spr.2007.5.1>
- Alexa, A., & Rahnenführer, J. (2007). Gene set enrichment analysis with topGO. *Bioconductor Improvements*, 27. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Gene+set+enrichment+analysis+with+topGO+O#0%5Cnftpt://mirrors.nic.funet.fi/bioconductor.org/2.7/bioc/vignettes/topGO/inst/doc/topGO.pdf>
- Alexander, L., & Grierson, D. (2002). Ethylene biosynthesis and action in tomato: A model for climacteric fruit ripening. *Journal of Experimental Botany*, 53(377), 2039–2055. <https://doi.org/10.1093/jxb/erf072>
- Alici, E. H., & Arabaci, G. (2018). A novel serine protease from strawberry (*Fragaria ananassa*): Purification and biochemical characterization. *International Journal of Biological Macromolecules*, 114, 1295–1304. <https://doi.org/10.1016/j.ijbiomac.2018.03.165>
- Allaire, J. (2012). *RStudio: integrated development environment for R*.
- Alquezar, B., Rodrigo, M. J., & Zacarías, L. (2008). Regulation of carotenoid biosynthesis during fruit maturation in the red-fleshed orange mutant Cara Cara. *Phytochemistry*, 69(10), 1997–2007. <https://doi.org/10.1016/j.phytochem.2008.04.020>
- Ambawat, S., Sharma, P., Yadav, N. R., & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: An overview. *Physiology and Molecular Biology of Plants*, 19(3), 307–321. <https://doi.org/10.1007/s12298-013-0179-1>
- Anderson, J. W., Baird, P., Davis, R. H., Ferreri, S., Knudtson, M., Koraym, A., Waters, V., & Williams, C. L. (2009). Health benefits of dietary fiber. *Nutrition Reviews*, 67(4), 188–205. <https://doi.org/10.1111/j.1753-4887.2009.00189.x>
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Website: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Arah, I. K., Ahorbo, G. K., Anku, E. K., Kumah, E. K., & Amaglo, H. (2016). Postharvest handling practices and treatment methods for tomato handlers in developing countries: a mini review. *Advances in Agriculture*, 2016, 1–8. <https://doi.org/10.1155/2016/6436945>
- Ariza, M. T., Soria, C., Medina, J. J., & Martínez-Ferri, E. (2011). Fruit misshapen in strawberry cultivars (*Fragaria × ananassa*) is related to achenes functionality. *Annals of Applied Biology*, 158(1), 130–138.
- Arlotta, C., Puglia, G. D., Genovese, C., Toscano, V., Karlova, R., Beekwilder, J., De Vos, R. C. H., & Raccuia, S. A. (2020). MYB5-like and bHLH influence flavonoid composition in pomegranate. *Plant Science*, 298, 110563. <https://doi.org/10.1016/j.plantsci.2020.110563>
- Arribas-Hernández, L., Marchais, A., Poulsen, C., Haase, B., Hauptmann, J., Benes, V., Meister, G., & Brodersen, P. (2016). The slicer activity of ARGONAUTE1 is required specifically for the phasing, not production, of trans-acting short interfering RNAs in arabidopsis. *Plant Cell*, 28(7), 1563–1580.

<https://doi.org/10.1105/tpc.16.00121>

Arruda, P., & Barreto, P. (2020). Lysine catabolism through the saccharopine pathway: enzymes and intermediates involved in plant responses to abiotic and biotic stress. *Frontiers in Plant Science*, *11*(May), 1–10. <https://doi.org/10.3389/fpls.2020.00587>

Arumuganathan, K., & Earle, E. D. (1991). Nuclear DNA content of some important plant species nuclear DNA content material and methods. *Plant Molecular Biology Reporter*, *9*(3), 208–218.

Atkinson, R. G., Johnston, S. L., Yauk, Y. K., Sharma, N. N., & Schröder, R. (2009). Analysis of xyloglucan endotransglucosylase/hydrolase (XTH) gene families in kiwifruit and apple. *Postharvest Biology and Technology*, *51*(2), 149–157. <https://doi.org/10.1016/j.postharvbio.2008.06.014>

Axtell, M. J. (2013). ShortStack: Comprehensive annotation and quantification of small RNA genes. *Rna*, *19*(6), 740–751. <https://doi.org/10.1261/rna.035279.112>

Axtell, M. J., Westholm, J. O., & Lai, E. C. (2011). Vive la différence: Biogenesis and evolution of microRNAs in plants and animals. *Genome Biology*, *12*(4), 1–13. <https://doi.org/10.1186/gb-2011-12-4-221>

Bai, M., Yang, G. S., Chen, W. T., Mao, Z. C., Kang, H. X., Chen, G. H., Yang, Y. H., & Xie, B. Y. (2012). Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene*, *501*(1), 52–62. <https://doi.org/10.1016/j.gene.2012.02.009>

Baldwin, E. A., Scott, J. W., Einstein, M. A., Malundo, T. M. M., Carr, B. T., Shewfelt, R. L., & Tandon, K. S. (1998). Relationship between sensory and instrumental analysis for tomato flavor. In *Journal of the American Society for Horticultural Science* (Vol. 123, Issue 5, pp. 906–915). <https://doi.org/10.21273/jashs.123.5.906>

Banerjee, J., Gantait, S., Sarkar, S., & Bhattacharyya, P. K. (2018). Transgenic research on tomato: problems, strategies, and achievements. In *Biotechnologies of Crop Improvement, Volume 2* (pp. 287–334). Springer.

Barringer, S. A., Bennett, M. A., & Bash, W. D. (1999). Effect of fruit maturity and nitrogen fertilizer levels on tomato peeling efficiency. *Journal of Vegetable Crop Production*, *5*(1), 3–11. [https://doi.org/10.1300/J068v05n01\\_02](https://doi.org/10.1300/J068v05n01_02)

Basson, C. E., Groenewald, J. H., Kossmann, J., Cronjé, C., & Bauer, R. (2010). Sugar and acid-related quality attributes and enzyme activities in strawberry fruits: Invertase is the main sucrose hydrolysing enzyme. *Food Chemistry*, *121*(4), 1156–1162. <https://doi.org/10.1016/j.foodchem.2010.01.064>

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Silva, A. Da, Denny, P., Dogan, T., Ebenezer, T. G., Fan, J., Castro, L. G., ... Zhang, J. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480–D489. <https://doi.org/10.1093/nar/gkaa1100>

Baumberger, N., & Baulcombe, D. C. (2005). Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(33), 11928–11933. <https://doi.org/10.1073/pnas.0505461102>

Beffa, R. S., Neuhaus, J. M., & Meins, F. (1993). Physiological compensation in antisense transformants: Specific induction of an “ersatz” glucan endo-1,3- $\beta$ -glucosidase in plants infected with necrotizing viruses. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(19), 8792–8796. <https://doi.org/10.1073/pnas.90.19.8792>

Bhat, R., Geppert, J., Funken, E., & Stamminger, R. (2015). Consumers perceptions and preference for strawberries—A case study from Germany. *International Journal of Fruit Science*, *15*(4), 405–424. <https://doi.org/10.1080/15538362.2015.1021408>

Bleidorn, C. (2016). Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity*, *14*(1), 1–8. <https://doi.org/10.1080/14772000.2015.1099575>

Bologna, N. G., & Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small rnas in *Arabidopsis*. *Annual Review of Plant Biology*, *65*(1), 473–503. <https://doi.org/10.1146/annurev-arplant-050213-035728>

Bonar, N., Liney, M., Zhang, R., Austin, C., Dessoly, J., Davidson, D., Stephens, J., McDougall, G., Taylor, M., & Bryan, G. J. (2018). Potato miR828 is associated with purple tuber skin and flesh color. *Frontiers in Plant Science*, *9*, 1742.

- Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nature Publishing Group, November*. <https://doi.org/10.1038/nrm4085>
- Borges, F., Parent, J. S., Van Ex, F., Wolff, P., Martínez, G., Köhler, C., & Martienssen, R. A. (2018). Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis. *Nature Genetics, 50*(2), 186–192. <https://doi.org/10.1038/s41588-017-0032-5>
- Bouzayen, M., Latché, A., Nath, P., & Pech, J.-C. (2010). Mechanism of fruit ripening. In *Plant developmental biology-Biotechnological perspectives* (pp. 319–339). Springer.
- Bramley, P. M. (1997). The regulation and genetic manipulation of carotenoid biosynthesis in tomato fruit. *Pure and Applied Chemistry, 69*(10), 2159–2162. <https://doi.org/10.1351/pac199769102159>
- Bramley, P. M. (2002). Regulation of carotenoid formation during tomato fruit ripening and development. *Journal of Experimental Botany, 53*(377), 2107–2113.
- Breitel, D. A., Chappell-Maor, L., Meir, S., Panizel, I., Puig, C. P., Hao, Y., Yifhar, T., Yasuor, H., Zouine, M., Bouzayen, M., Granell Richart, A., Rogachev, I., & Aharoni, A. (2016). AUXIN RESPONSE FACTOR 2 intersects hormonal signals in the regulation of tomato fruit ripening. *PLoS Genetics, 12*(3), 1–34. <https://doi.org/10.1371/journal.pgen.1005903>
- Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R., & Turner, S. R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell, 17*(8), 2281–2295. <https://doi.org/10.1105/tpc.105.031542>
- Brummell, D. A. (2006). Cell wall disassembly in ripening fruit. *Functional Plant Biology, 33*(2), 103–119. <https://doi.org/10.1071/FP05234>
- Brummell, D. A., & Harpster, M. H. (2001). Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Molecular Biology, 47*(1–2), 311–339. <https://doi.org/10.1023/A:1010656104304>
- Brummell, D. A., Harpster, M. H., Civello, P. M., Palys, J. M., Bennett, A. B., & Dunsmuir, P. (1999). Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening. *Plant Cell, 11*(11), 2203–2216. <https://doi.org/10.1105/tpc.11.11.2203>
- Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., Schijlen, E. G. W. M., Hall, R. D., Bovy, A. G., & Luo, J. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature Biotechnology, 26*(11), 1301–1308.
- Cantu, D., Vicente, A. R., Greve, L. C., Dewey, F. M., Bennett, A. B., Labavitch, J. M., & Powell, A. L. T. (2008). The intersection between cell wall disassembly, ripening, and fruit susceptibility to *Botrytis cinerea*. *Proceedings of the National Academy of Sciences, 105*(3), 859–864.
- Cao, D., Wang, J., Ju, Z., Liu, Q., Li, S., Tian, H., Fu, D., Zhu, H., Luo, Y., & Zhu, B. (2016). Regulations on growth and development in tomato cotyledon, flower and fruit via destruction of miR396 with short tandem target mimic. *Plant Science, 247*, 1–12. <https://doi.org/10.1016/j.plantsci.2016.02.012>
- Cao, J. Y., Xu, Y. P., Li, W., Li, S. S., Rahman, H., & Cai, X. Z. (2016). Genome-wide identification of dicer-like, argonaute, and RNA-dependent RNA polymerase gene families in brassica species and functional analyses of their arabidopsis homologs in resistance to *Sclerotinia sclerotiorum*. *Frontiers in Plant Science, 7*(OCTOBER2016), 1–17. <https://doi.org/10.3389/fpls.2016.01614>
- Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K. B., Montgomery, T. A., Nguyen, T., Cuperus, J. T., & Carrington, J. C. (2012). Functional analysis of three Arabidopsis argonautes using slicer-defective mutants. *Plant Cell, 24*(9), 3613–3629. <https://doi.org/10.1105/tpc.112.099945>
- Castro, R. I., Muñoz-Vera, M., Parra-Palma, C., Valenzuela-Riffo, F., Figueroa, C. R., & Morales-Quintana, L. (2021). Characterization of cell wall modification through thermogravimetric analysis during ripening of Chilean strawberry (*Fragaria chiloensis*) fruit. *Cellulose, 28*(8), 4611–4623. <https://doi.org/10.1007/s10570-021-03806-w>
- Causse, M., Buret, M., Robini, K., & Verschave, P. (2003). Inheritance of nutritional and sensory quality traits in fresh market tomato and relation to consumer preferences. *Journal of Food Science, 68*(7), 2342–2350.
- Causse, M., Desplat, N., Pascual, L., Le Paslier, M. C., Sauvage, C., Bauchet, G., Bérard, A., Bounon, R., Tchoumakov, M., Brunel, D., & Bouchet, J. P. (2013). Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genomics, 14*(1). <https://doi.org/10.1186/1471-2164->

- Causse, M., Friguier, C., Coiret, C., Lépicié, M., Navez, B., Lee, M., Holthuysen, N., Sinesio, F., Moneta, E., & Grandillo, S. (2010). Consumer preferences for fresh tomato at the european scale: a common segmentation on taste and firmness. *Journal of Food Science*, *75*(9). <https://doi.org/10.1111/j.1750-3841.2010.01841.x>
- Cedillo-Jimenez, C. A., Feregrino-Perez, A. A., Guevara-González, R. G., & Cruz-Hernández, A. (2020). MicroRNA regulation during the tomato fruit development and ripening: A review. *Scientia Horticulturae*, *270*(April), 109435. <https://doi.org/10.1016/j.scienta.2020.109435>
- Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. *Current Opinion in Plant Biology*, *27*, 207–216. <https://doi.org/10.1016/j.pbi.2015.08.003>
- Chen, S. C., Liu, A. R., & Zou, Z. R. (2006). Overexpression of glucanase gene and defensin gene in transgenic tomato enhances resistance to *Ralstonia solanacearum*. *Russian Journal of Plant Physiology*, *53*(5), 671–677. <https://doi.org/10.1134/S1021443706050116>
- Chen, W., Kong, J., Lai, T., Manning, K., Wu, C., Wang, Y., Qin, C., Li, B., Yu, Z., Zhang, X., He, M., Zhang, P., Gu, M., Yang, X., Mahammed, A., Li, C., Osman, T., Shi, N., Wang, H., ... Hong, Y. (2015). Tuning LeSPL-CNR expression by SlymiR157 affects tomato fruit ripening. *Scientific Reports*, *5*, 1–6. <https://doi.org/10.1038/srep07852>
- Chen, X. (2005). microRNA biogenesis and function in plants. *FEBS Letters*, *579*(26), 5923–5931. <https://doi.org/10.1016/j.febslet.2005.07.071>
- Chen, Y., McCarthy, D., Ritchie, M., Robinson, M., Smyth, G., & Hall, E. (2020). edgeR: differential analysis of sequence read count data User's Guide. *R Package, June*, 1–121.
- Cheng, C., Zhang, L., Yang, X., & Zhong, G. (2015). Profiling gene expression in citrus fruit calyx abscission zone (AZ-C) treated with ethylene. *Molecular Genetics and Genomics*, *290*(5), 1991–2006.
- Cheng, J., Niu, Q., Zhang, B., Chen, K., Yang, R., Zhu, J.-K., Zhang, Y., & Lang, Z. (2018). Downregulation of RdDM during strawberry fruit ripening. *Genome Biology*, *19*(1), 1–14.
- Choudhury, S. R., Roy, S., & Sengupta, D. N. (2009). Characterization of cultivar differences in  $\beta$ -1,3 glucanase gene expression, glucanase activity and fruit pulp softening rates during fruit ripening in three naturally occurring banana cultivars. *Plant Cell Reports*, *28*(11), 1641–1653. <https://doi.org/10.1007/s00299-009-0764-5>
- Christenhusz, M. J. M., & Byng, J. W. (2016). The number of known plants species in the world and its annual increase. *Phytotaxa*, *261*(3), 201–217. <https://doi.org/10.11646/phytotaxa.261.3.1>
- Cisneros, A. E., & Carbonell, A. (2020). Artificial small RNA-based silencing tools for antiviral resistance in plants. *Plants*, *9*(6), 669.
- Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschield, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M., & Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, *452*(7184), 215–219. <https://doi.org/10.1038/nature06745>
- Collins, F. S., Morgan, M., & Patrinos, A. (2003). The Human Genome Project: Lessons from large-scale biology. *Science*, *300*(5617), 286–290. <https://doi.org/10.1126/science.1084564>
- Committee, A. A. of C. C. F. (2001). The definition of dietary fiber: Report of the dietary fiber definition committee to the board of directors of the american association of cereal chemists. *Cereal Foods World*, *46*(3), 112–126.
- Cosgrove, D. J. (1998). Cell wall loosening by expansins. *Plant Physiology*, 333–339.
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*, *6*(11), 850–861. <https://doi.org/10.1038/nrm1746>
- Crane, P. R., Friis, E. M., & Pedersen, K. R. (1995). The origin and early diversification of angiosperms. *Nature*, *374*(6517), 27–33.
- Crookes, P. R., & Grierson, D. (1983). Ultrastructure of tomato fruit ripening and the role of polygalacturonase isoenzymes in cell wall degradation. *Plant Physiology*, *72*(4), 1088–1093. <https://doi.org/10.1104/pp.72.4.1088>
- Csukasi, F., Donaire, L., Casañal, A., Martínez-Priego, L., Botella, M. A., Medina-Escobar, N., Llave, C., & Valpuesta, V. (2012). Two strawberry miR159 family members display developmental-specific expression

- patterns in the fruit receptacle and cooperatively regulate Fa-GAMYB. *New Phytologist*, 195(1), 47–57. <https://doi.org/10.1111/j.1469-8137.2012.04134.x>
- Cui, J., You, C., & Chen, X. (2017). The evolution of microRNAs in plants. *Current Opinion in Plant Biology*, 35, 61–67. <https://doi.org/10.1016/j.pbi.2016.11.006>
- D'Elia, L., Barba, G., Cappuccio, F. P., & Strazzullo, P. (2011). Potassium intake, stroke, and cardiovascular disease: A meta-analysis of prospective studies. *Journal of the American College of Cardiology*, 57(10), 1210–1219. <https://doi.org/10.1016/j.jacc.2010.09.070>
- da Silva, E. M., Silva, G. F. F. e, Bidoia, D. B., da Silva Azevedo, M., de Jesus, F. A., Pino, L. E., Peres, L. E. P., Carrera, E., López-Díaz, I., & Nogueira, F. T. S. (2017). Micro RNA 159-targeted Sl GAMYB transcription factors are required for fruit set in tomato. *The Plant Journal*, 92(1), 95–109.
- Dalakouras, A., Wassenegger, M., Dadami, E., Ganopoulos, I., Pappas, M. L., & Papadopoulou, K. (2020). Genetically modified organism-free RNA interference: exogenous application of RNA molecules in plants. *Plant Physiology*, 182(1), 38–50.
- Danisman, S., van der Wal, F., Dhondt, S., Waites, R., de Folter, S., Bimbo, A., van Dijk, A. J., Muino, J. M., Cutri, L., Dornelas, M. C., Angenent, G. C., & Immink, R. G. H. (2012). Arabidopsis class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically. *Plant Physiology*, 159(4), 1511–1523. <https://doi.org/10.1104/pp.112.200303>
- Dardick, C. D., Callahan, A. M., Chiozzotto, R., Schaffer, R. J., Piagnani, M. C., & Scorza, R. (2010). Stone formation in peach fruit exhibits spatial coordination of the lignin and flavonoid pathways and similarity to Arabidopsis dehiscence. *BMC Biology*, 8, 1–17. <https://doi.org/10.1186/1741-7007-8-13>
- Das, S., Swetha, C., Pachamuthu, K., Nair, A., & Shivaprasad, P. V. (2020). Loss of function of *Oryza sativa* Argonaute 18 induces male sterility and reduction in phased small RNAs. *Plant Reproduction*, 33(1), 59–73. <https://doi.org/10.1007/s00497-020-00386-w>
- Daurelio, L.D. 2009. Tesis Doctoral. Facultad de Ciencias Bioquímicas y Farmacéuticas, UNR.
- De Jong, M., Wolters-Arts, M., Feron, R., Mariani, C., & Vriezen, W. H. (2009). The *Solanum lycopersicum* auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. *Plant Journal*, 57(1), 160–170. <https://doi.org/10.1111/j.1365-313X.2008.03671.x>
- De Vries, S., Kloesges, T., & Rose, L. E. (2015). Evolutionarily dynamic, but robust, targeting of resistance genes by the miR482/2118 gene family in the Solanaceae. *Genome Biology and Evolution*, 7(12), 3307–3321. <https://doi.org/10.1093/gbe/evv225>
- Deckers, T., Daemen, E., Lemmens, K., & Missotten, C. (1996). Influence of foliar applications of Mn during summer on the fruit quality of Jonagold. *III International Symposium on Mineral Nutrition of Deciduous Fruit Trees 448*, 467–474.
- DeYoung, B. J., & Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*, 7(12), 1243–1249. <https://doi.org/10.1038/ni1410>
- Díaz de León-Sánchez, F., Pelayo-Zaldívar, C., Rivera-Cabrera, F., Ponce-Valadez, M., Ávila-Alejandre, X., Fernández, F. J., Escalona-Buendía, H. B., & Pérez-Flores, L. J. (2009). Effect of refrigerated storage on aroma and alcohol dehydrogenase activity in tomato fruit. *Postharvest Biology and Technology*, 54(2), 93–100. <https://doi.org/10.1016/j.postharvbio.2009.07.003>
- Djupedal, I., & Ekwall, K. (2009). Epigenetics: Heterochromatin meets RNAi. *Cell Research*, 19(3), 282–295. <https://doi.org/10.1038/cr.2009.13>
- do Nascimento Nunes, M. C. (2009). Color atlas of postharvest quality of fruits and vegetables. John Wiley & Sons.
- Domínguez, E., Cuartero, J., & Heredia, A. (2011). An overview on plant cuticle biomechanics. *Plant Science*, 181(2), 77–84. <https://doi.org/10.1016/j.plantsci.2011.04.016>
- Dong, Z., Han, M. H., & Fedoroff, N. (2008). The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. *Proceedings of the National Academy of Sciences of the United States of America*, 105(29), 9970–9975. <https://doi.org/10.1073/pnas.0803356105>
- Dotto, M. C., Petsch, K. A., Aukerman, M. J., Beatty, M., Hammell, M., & Timmermans, M. C. P. (2014). Genome-wide analysis of leafbladeless1-regulated and phased small RNAs underscores the importance of the tas3 ta-

- siRNA pathway to maize development. *PLoS Genetics*, *10*(12). <https://doi.org/10.1371/journal.pgen.1004826>
- Dubrovina, A. S., & Kiselev, K. V. (2019). Exogenous RNAs for gene regulation and plant resistance. *International Journal of Molecular Sciences*, *20*(9), 2282.
- Echeverría, G., Graell, J., López, M. L., & Lara, I. (2004). Volatile production, quality and aroma-related enzyme activities during maturation of “Fuji” apples. *Postharvest Biology and Technology*, *31*(3), 217–227. <https://doi.org/10.1016/j.postharvbio.2003.09.003>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., Smith, R. D., Teresi, S. J., Nelson, A. D. L., Wai, C. M., Alger, E. I., Bird, K. A., Yocca, A. E., Pumplun, N., Ou, S., Ben-Zvi, G., Brodt, A., Baruch, K., Swale, T., ... Knapp, S. J. (2019). Origin and evolution of the octoploid strawberry genome. *Nature Genetics*, *51*(3), 541–547. <https://doi.org/10.1038/s41588-019-0356-4>
- Edger, P. P., VanBuren, R., Colle, M., Poorten, T. J., Wai, C. M., Niederhuth, C. E., Alger, E. I., Ou, S., Acharya, C. B., Wang, J., Callow, P., McKain, M. R., Shi, J., Collier, C., Xiong, Z., Mower, J. P., Slovin, J. P., Hytönen, T., Jiang, N., ... Knapp, S. J. (2018). Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *GigaScience*, *7*(2), 1–7. <https://doi.org/10.1093/gigascience/gix124>
- Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., & Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant Journal*, *49*(3), 414–427. <https://doi.org/10.1111/j.1365-313X.2006.02964.x>
- Estornell, L. H., Agustí, J., Merelo, P., Talón, M., & Tadeo, F. R. (2013). Elucidating mechanisms underlying organ abscission. *Plant Science*, *199*, 48–60.
- Estrada-Johnson, E., Csukasi, F., Pizarro, C. M., Vallarino, J. G., Kiryakova, Y., Vioque, A., Brumos, J., Medina-Escobar, N., Botella, M. A., Alonso, J. M., Fernie, A. R., Sánchez-Sevilla, J. F., Osorio, S., & Valpuesta, V. (2017). Transcriptomic analysis in strawberry fruits reveals active auxin biosynthesis and signaling in the ripe receptacle. *Frontiers in Plant Science*, *8*(May), 1–16. <https://doi.org/10.3389/fpls.2017.00889>
- Fahlgren, N., Montgomery, T. A., Howell, M. D., Allen, E., Dvorak, S. K., Alexander, A. L., & Carrington, J. C. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. *Current Biology*, *16*(9), 939–944. <https://doi.org/10.1016/j.cub.2006.03.065>
- Federcitrus, 2018. Federación Argentina del Citrus. La actividad cítrica argentina. Disponible en <https://www.federcitrus.org/estadisticas/#>. Último acceso: 16 de abril de 2020
- Fei, Q., Xia, R., & Meyers, B. C. (2013). Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *The Plant Cell*, *25*(7), 2400–2415. <https://doi.org/10.1105/tpc.113.114652>
- Fei, Qili, Li, P., Teng, C., & Meyers, B. C. (2015). Secondary siRNAs from Medicago NB-LRRs modulated via miRNA-target interactions and their abundances. *Plant Journal*, *83*(3), 451–465. <https://doi.org/10.1111/tpj.12900>
- Feng, L., Xia, R., & Liu, Y. (2019). Comprehensive characterization of miRNA and PHAS loci in the diploid strawberry (*Fragaria vesca*) Genome. *Horticultural Plant Journal*, *5*(6), 255–267. <https://doi.org/10.1016/j.hpj.2019.11.004>
- Fernandes-Brum, C. N., Rezende, P. M., Ribeiro, T. H. C., De Oliveira, R. R., De Sousa Cardoso, T. C., Do Amaral, L. R., De Souza Gomes, M., & Chalfun, A. (2017). A genome-wide analysis of the RNA-guided silencing pathway in coffee reveals insights into its regulatory mechanisms. *PLoS ONE*, *12*(4), 1–29. <https://doi.org/10.1371/journal.pone.0176333>
- Font Quer, P. (1953). *Diccionario de botánica*.
- Food and Agriculture Organization [FAO] (2014). Disponible en: <http://www.fao.org/faostat/en/#data/QC> . Último acceso: 15 de Abril de 2020.
- Forlani, S., Masiero, S., & Mizzotti, C. (2019). Fruit ripening: the role of hormones, cell wall modifications, and their relationship with pathogens. *Journal of Experimental Botany*, *70*(11), 2993–3006.
- Forney, C. F., Kalt, W., & Jordan, M. A. (2000). The composition of strawberry aroma is influenced by cultivar, maturity, and storage. *HortScience*, *35*(6), 1022–1026. <https://doi.org/10.21273/hortsci.35.6.1022>

Fragaria, en The Plant List, vers.1.1, 2013. Último acceso: octubre de 2021

Frank, F., Hauver, J., Sonenberg, N., & Nagar, B. (2012). Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. *EMBO Journal*, 31(17), 3588–3595. <https://doi.org/10.1038/emboj.2012.204>

Fraser, P. D., Truesdale, M. R., Bird, C. R., Schuch, W., & Bramley, P. M. (1994). Carotenoid biosynthesis during tomato fruit development. Evidence for tissue-specific gene expression. *Plant Physiology*, 105(1), 405–413. <https://doi.org/10.1104/pp.105.1.405>

Friedman, M. (2013). Anticarcinogenic, cardioprotective, and other health benefits of tomato compounds lycopene,  $\alpha$ -tomatine, and tomatidine in pure form and in fresh and processed tomatoes. *Journal of Agricultural and Food Chemistry*, 61(40), 9534–9550. <https://doi.org/10.1021/jf402654e>

Fry, S. C. (2004). Primary cell wall metabolism: Tracking the careers of wall polymers in living plant cells. *New Phytologist*, 161(3), 641–675. <https://doi.org/10.1111/j.1469-8137.2004.00980.x>

Fukudome, A., & Fukuhara, T. (2017). Plant dicer-like proteins: double-stranded RNA-cleaving enzymes for small RNA biogenesis. *J Plant Res*, 130, 33–44.

Gabotti, D., Negrini, N., Morgutti, S., Nocito, F. F., & Cocucci, M. (2015). Cinnamyl alcohol dehydrogenases in the mesocarp of ripening fruit of *Prunus persica* genotypes with different flesh characteristics: Changes in activity and protein and transcript levels. *Physiologia Plantarum*, 154(3), 329–348. <https://doi.org/10.1111/ppl.12319>

Gandikota, M., Birkenbihl, R. P., Höhmann, S., Cardon, G. H., Saedler, H., & Huijser, P. (2007). The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant Journal*, 49(4), 683–693. <https://doi.org/10.1111/j.1365-313X.2006.02983.x>

Gao, C., Ju, Z., Cao, D., Zhai, B., Qin, G., Zhu, H., Fu, D., Luo, Y., & Zhu, B. (2015). MicroRNA profiling analysis throughout tomato fruit development and ripening reveals potential regulatory role of RIN on microRNAs accumulation. *Plant Biotechnology Journal*, 13(3), 370–382. <https://doi.org/10.1111/pbi.12297>

Gao, Z., House, L. O., Gmitter, F. G., Valim, M. F., Plotto, A., & Baldwin, E. A. (2011). Consumer preferences for fresh citrus: Impacts of demographic and behavioral characteristics. *International Food and Agribusiness Management Review*, 14(1), 23–40. <https://doi.org/10.22004/ag.econ.100769>

García, M. G., Ontivero, M., Diaz Ricci, J. C., & Castagnaro, A. (2002). Morphological traits and high resolution RAPD markers for the identification of the main strawberry varieties cultivated in Argentina. *Plant Breeding*, 121(1), 76–80. <https://doi.org/10.1046/j.1439-0523.2002.00671.x>

Gautier, H., Diakou-Verdin, V., Bénard, C., Reich, M., Buret, M., Bourgaud, F., Poëssel, J. L., Caris-Veyrat, C., & Génard, M. (2008). How does tomato quality (sugar, acid, and nutritional quality) vary with ripening stage, temperature, and irradiance? *Journal of Agricultural and Food Chemistry*, 56(4), 1241–1250. <https://doi.org/10.1021/jf072196t>

German, M. A., Luo, S., Schroth, G., Meyers, B. C., & Green, P. J. (2009). Construction of parallel analysis of RNA ends (Pare) libraries for the study of cleaved mirna targets and the RNA degradome. *Nature Protocols*, 4(3), 356–362. <https://doi.org/10.1038/nprot.2009.8>

Gil-Amado, J. A., & Gomez-Jimenez, M. C. (2013). Transcriptome analysis of mature fruit abscission control in olive. *Plant and Cell Physiology*, 54(2), 244–269.

Girard, A. L., Mounet, F., Lemaire-Chamley, M., Gaillard, C., Elmorjani, K., Vivancos, J., Runavot, J. L., Quemener, B., Petit, J., Germain, V., Rothan, C., Marion, D., & Bakan, B. (2012). Tomato GDLS1 is required for cutin deposition in the fruit cuticle. *Plant Cell*, 24(7), 3106–3121. <https://doi.org/10.1105/tpc.112.101055>

Given, N. K., Venis, M. A., & Gierson, D. (1988). Hormonal regulation of ripening in the strawberry, a non-climacteric fruit. *Planta*, 174(3), 402–406. <https://doi.org/10.1007/BF00959527>

Gmitter, F. G., & Hu, X. (1990). The possible role of Yunnan, China, in the origin of contemporary citrus species (rutaceae). *Economic Botany*, 44(2), 267–277. <https://doi.org/10.1007/BF02860491>

Gol, N. B., Patel, P. R., & Rao, T. V. R. (2013). Improvement of quality and shelf-life of strawberries with edible coatings enriched with chitosan. *Postharvest Biology and Technology*, 85, 185–195. <https://doi.org/10.1016/j.postharvbio.2013.06.008>

- Goldschmidt, E. E. (1998). Ripening of citrus and other non-climacteric fruits: A role for ethylene. In *Acta Horticulturae* (Vol. 463, pp. 335–340). <https://doi.org/10.17660/ActaHortic.1998.463.42>
- Goodin, M. M., Zaitlin, D., Naidu, R. A., & Lommel, S. A. (2008). *Nicotiana benthamiana*: Its history and future as a model for plant-pathogen interactions. *Molecular Plant-Microbe Interactions*, *21*(8), 1015–1026. <https://doi.org/10.1094/MPMI-21-8-1015>
- Gordon, A. 2011. FASTX-Toolkit. Computer program distributed by the author, website [http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html).
- Goren, R., Huberman, M., & Zamski, E. (1984). *Antomical aspects of citrus abscission - effects of ethylene on leaf and fruit explants*. In Ethylene (pp. 241-254). Springer, Dordrecht.
- Gou, J. Y., Felippes, F. F., Liu, C. J., Weigel, D., & Wang, J. W. (2011). Negative Regulation of Anthocyanin Biosynthesis in Arabidopsis by a miR156-Targeted SPL Transcription Factor. *Plant Cell*, *23*(4), 1512–1522. <https://doi.org/10.1105/tpc.111.084525>
- Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., & Eddy, S. R. (2003). Rfam: An RNA family database. *Nucleic Acids Research*, *31*(1), 439–441. <https://doi.org/10.1093/nar/gkg006>
- Guan, X., Pang, M., Nah, G., Shi, X., Ye, W., Stelly, D. M., & Chen, Z. J. (2014). MiR828 and miR858 regulate homoeologous MYB2 gene functions in Arabidopsis trichome and cotton fibre development. *Nature Communications*, *5*(May 2013). <https://doi.org/10.1038/ncomms4050>
- Hall, J. N., Moore, S., Harper, S. B., & Lynch, J. W. (2009). Global variability in fruit and vegetable consumption. *American Journal of Preventive Medicine*, *36*(5), 402–409.
- Han, J., Li, A., Liu, H., Wen, X., Zhao, M., Korir, N. B., Korir, N. K., Wang, C., & Fang, J. (2014). Computational identification of microRNAs in the strawberry (*Fragaria×ananassa*) genome sequence and validation of their precise sequences by miR-RACE. *Gene*, *536*(1), 151–162. <https://doi.org/10.1016/j.gene.2013.11.023>
- Han, Y., Zhu, Q., Zhang, Z., Meng, K., Hou, Y., Ban, Q., Suo, J., & Rao, J. (2015). Analysis of xyloglucan endotransglycosylase/hydrolase (XTH) genes and diverse roles of isoenzymes during persimmon fruit development and postharvest softening. *PLoS ONE*, *10*(4), 1–20. <https://doi.org/10.1371/journal.pone.0123668>
- Hancock, J. F., Lansing, E., Finn, C. E., Luby, J. J., Dale, A., Ny, C., & Callow, P. W. (2010). Reconstruction of the Strawberry, *Fragaria ananassa*, Using Genotypes of *F. virginiana* and *F. chiloensis*. *45*(7), 1006–1013.
- Hao, Y., Hu, G., Breitel, D., Liu, M., Mila, I., Frasse, P., Fu, Y., Aharoni, A., Bouzayen, M., & Zouine, M. (2015). Auxin Response Factor SlARF2 is an essential component of the regulatory mechanism controlling fruit ripening in tomato. *PLoS Genetics*, *11*(12), 1–23. <https://doi.org/10.1371/journal.pgen.1005649>
- Härtl, K., Denton, A., Franz-Oberdorf, K., Hoffmann, T., Spornraft, M., Usadel, B., & Schwab, W. (2017). Early metabolic and transcriptional variations in fruit of natural white-fruited *Fragaria vesca* genotypes. *Scientific Reports*, *7*, 1–16. <https://doi.org/10.1038/srep45113>
- Henderson, I. R., Zhang, X., Lu, C., Johnson, L., Meyers, B. C., Green, P. J., & Jacobsen, S. E. (2006). Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature Genetics*, *38*(1), 721–725. <https://doi.org/10.1038/nrg3683>
- Heng Koh, T., & Melton, L. D. (2002). Ripening-related changes in cell wall polysaccharides of strawberry cortical and pith tissues. *Postharvest Biology and Technology*, *26*(1), 23–33. [https://doi.org/10.1016/S0925-5214\(02\)00003-0](https://doi.org/10.1016/S0925-5214(02)00003-0)
- Hidalgo Suarez, C. G., Guevara Burbano, M. E., Bucheli Guerrero, V. A., & Moreno Tovar, P. A. (2018). Bioinformatics software for genomic: a systematic review on GitHub. <https://doi.org/10.7287/peerj.preprints.27352>
- Hieber, A. D., Bugos, R. C., & Yamamoto, H. Y. (2000). Plant lipocalins: Violaxanthin de-epoxidase and zeaxanthin epoxidase. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, *1482*(1–2), 84–91. [https://doi.org/10.1016/S0167-4838\(00\)00141-2](https://doi.org/10.1016/S0167-4838(00)00141-2)
- Hosmani, P. S., Flores-Gonzalez, M., van de Geest, H., Maumus, F., Bakker, L. V., Schijlen, E., van Haarst, J., Cordewener, J., Sanchez-Perez, G., Peters, S., Fei, Z., Giovannoni, J. J., Mueller, L. A., & Saha, S. (2019). An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and optical maps. *BioRxiv*, *2012*, 1–21. <https://doi.org/10.1101/767764>

- Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., & Gao, G. (2015). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics*, *31*(8), 1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>
- Huang, W., Peng, S., Xian, Z., Lin, D., Hu, G., Yang, L., Ren, M., & Li, Z. (2017). Overexpression of a tomato miR171 target gene SIGRAS24 impacts multiple agronomical traits via regulating gibberellin and auxin homeostasis. *Plant Biotechnology Journal*, *15*(4), 472–488. <https://doi.org/10.1111/pbi.12646>
- Hui, Y. H., Chen, F., Nollet, L. M. L., Guiné, R. P. F., Martín-Belloso, O., Mínguez-Mosquera, M. I., Paliyath, G., Pessoa, F. L. P., Le Quéré, J.-L., & Sidhu, J. S. (2010). *Handbook of fruit and vegetable flavors*. John Wiley and Sons.
- Hutvagner, G., & Simard, M. J. (2008). Argonaute proteins: Key players in RNA silencing. *Nature Reviews Molecular Cell Biology*, *9*(1), 22–32. <https://doi.org/10.1038/nrm2321>
- Iglesias, D. J., Cercós, M., Colmenero-Flores, J. M., Naranjo, M. A., Ríos, G., Carrera, E., Ruiz-Rivero, O., Lliso, I., Morillon, R., & Tadeo, F. R. (2007). Physiology of citrus fruiting. *Brazilian Journal of Plant Physiology*, *19*, 333–362.
- Illera, A. E., Sanz, M. T., Trigueros, E., Beltrán, S., & Melgosa, R. (2018). Effect of high pressure carbon dioxide on tomato juice: Inactivation kinetics of pectin methylesterase and polygalacturonase and determination of other quality parameters. *Journal of Food Engineering*, *239*, 64–71. <https://doi.org/10.1016/j.jfoodeng.2018.06.027>
- Informe de la Food and Agriculture Organization (FAO, 2014). Disponible online en: <http://www.fao.org/faostat/en/#data/QC> . Último acceso: 18 de marzo de 2020
- Informe del Ministerio de Agroindustria, Presidencia de la Nación: Cadena del tomate para industrias, 2018. Disponible online en: [http://www.alimentosargentinos.gob.ar/HomeAlimentos/Cadenas%20de%20Valor%20de%20Alimentos%20%20Bebidas/informes/Resumen\\_Cadena\\_2018\\_TOMATE\\_INDUSTRIA.pdf](http://www.alimentosargentinos.gob.ar/HomeAlimentos/Cadenas%20de%20Valor%20de%20Alimentos%20%20Bebidas/informes/Resumen_Cadena_2018_TOMATE_INDUSTRIA.pdf)
- Informe del National Research Council. (1989). Diet and health: implications for reducing chronic disease risk. National Academies Press.
- Jackson, M. B., & Osborne, D. J. (1970). Ethylene, the natural regulator of leaf abscission. *Nature*, *225*, 1019–1022. <http://www.mendeley.com/research/discreteness-conductance-chnge-n-bimolecular-lipid-membrane-presence-certain-antibiotics/>
- Jia, H. F., Chai, Y. M., Li, C. L., Lu, D., Luo, J. J., Qin, L., & Shen, Y. Y. (2011). Abscisic acid plays an important role in the regulation of strawberry fruit ripening. *Plant Physiology*, *157*(1), 188–199. <https://doi.org/10.1104/pp.111.177311>
- Jia, X., Shen, J., Liu, H., Li, F., Ding, N., Gao, C., Pattanaik, S., Patra, B., Li, R., & Yuan, L. (2015). Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato. *Planta*, *242*(1), 283–293. <https://doi.org/10.1007/s00425-015-2305-5>
- Jian, W., Cao, H., Yuan, S., Liu, Y., Lu, J., Lu, W., Li, N., Wang, J., Zou, J., Tang, N., Xu, C., Cheng, Y., Gao, Y., Xi, W., Bouzayen, M., & Li, Z. (2019). SIMYB75, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits. *Horticulture Research*, *6*(1). <https://doi.org/10.1038/s41438-018-0098-y>
- Jiang, C. Z., Lu, F., Imsabai, W., Meir, S., & Reid, M. S. (2008). Silencing polygalacturonase expression inhibits tomato petiole abscission. *Journal of Experimental Botany*, *59*(4), 973–979. <https://doi.org/10.1093/jxb/ern023>
- Jiang, Fangling, Lopez, A., Jeon, S., de Freitas, S. T., Yu, Q., Wu, Z., Labavitch, J. M., Tian, S., Powell, A. L. T., & Mitcham, E. (2019). Disassembly of the fruit cell wall by the ripening-associated polygalacturonase and expansin influences tomato cracking. *Horticulture Research*, *6*(1). <https://doi.org/10.1038/s41438-018-0105-3>
- Jiang, Fei, Wang, J.-Y., Jia, H.-F., Jia, W.-S., Wang, H.-Q., & Xiao, M. (2013). RNAi-mediated silencing of the flavanone 3-hydroxylase gene and its effect on flavonoid biosynthesis in strawberry fruit. *Journal of Plant Growth Regulation*, *32*(1), 182–190.
- Jiang, Y., & Joyce, D. C. (2003). ABA effects on ethylene production, PAL activity, anthocyanin and phenolic contents of strawberry fruit. *Plant Growth Regulation*, *39*(2), 171–174. <https://doi.org/10.1023/A:1022539901044>
- Jiménez-Bermúdez, S., Redondo-Nevado, J., Muñoz-Blanco, J., Caballero, J. L., López-Aranda, J. M., Valpuesta,

- V., Pliego-Alfaro, F., Quesada, M. A., & Mercado, J. A. (2002). Manipulation of strawberry fruit softening by antisense expression of a pectate lyase gene. *Plant Physiology*, *128*(2), 751–759. <https://doi.org/10.1104/pp.010671>
- Jin, D., Wang, Y., Zhao, Y., & Chen, M. (2013). MicroRNAs and their cross-talks in plant development. *Journal of Genetics and Genomics*, *40*(4), 161–170. <https://doi.org/10.1016/j.jgg.2013.02.003>
- Jin, J., Liu, J., Wang, H., Wong, L., & Chua, N. H. (2013). PLncDB: Plant long non-coding RNA database. *Bioinformatics*, *29*(8), 1068–1071. <https://doi.org/10.1093/bioinformatics/btt107>
- José Ripoll, J., Bailey, L. J., Mai, Q. A., Wu, S. L., Hon, C. T., Chapman, E. J., Ditta, G. S., Estelle, M., & Yanofsky, M. F. (2015). MicroRNA regulation of fruit growth. *Nature Plants*, *1*(April), 1–9. <https://doi.org/10.1038/nplants.2015.36>
- Joubès, J., & Domergue, F. (2018). Biosynthesis of the plant cuticle. *Hydrocarbons, Oils and Lipids: Diversity, Origin, Chemistry and Fate*, 1–19. [https://doi.org/10.1007/978-3-319-54529-5\\_8-1](https://doi.org/10.1007/978-3-319-54529-5_8-1)
- Kakrana, A., Mathioni, S. M., Huang, K., Hammond, R., Vandivier, L., Patel, P., Arikiti, S., Shevchenko, O., Harkess, A. E., Kingham, B., Gregory, B. D., Leebens-Mack, J. H., & Meyers, B. C. (2018). Plant 24-nt reproductive phasiRNAs from intramolecular duplex mRNAs in diverse monocots. *Genome Research*, *28*(9), 1333–1344. <https://doi.org/10.1101/gr.228163.117>
- Kamthan, A., Chaudhuri, A., Kamthan, M., & Datta, A. (2015). Small RNAs in plants: recent development and application for crop improvement. *Frontiers in Plant Science*, *6*, 208.
- Kapoor, M., Arora, R., Lama, T., Nijhawan, A., Khurana, J. P., Tyagi, A. K., & Kapoor, S. (2008). Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics*, *9*, 1–17. <https://doi.org/10.1186/1471-2164-9-451>
- Karlova, R., Chapman, N., David, K., Angenent, G. C., Seymour, G. B., & De Maagd, R. A. (2014). Transcriptional control of fleshy fruit development and ripening. *Journal of Experimental Botany*, *65*(16), 4527–4541. <https://doi.org/10.1093/jxb/eru316>
- Karlova, R., Van Haarst, J. C., Maliepaard, C., Van De Geest, H., Bovy, A. G., Lammers, M., Angenent, G. C., & De Maagd, R. A. (2013). Identification of microRNA targets in tomato fruit development using high-throughput sequencing and degradome analysis. *Journal of Experimental Botany*, *64*(7), 1863–1878. <https://doi.org/10.1093/jxb/ert049>
- Katan, M. B., Grundy, S. M., Jones, P., Law, M., Miettinen, T., & Paoletti, R. (2003). Efficacy and safety of plant stanols and sterols in the management of blood cholesterol levels. *Mayo Clinic Proceedings*, *78*(8), 965–978. <https://doi.org/10.4065/78.8.965>
- Khaldun, A. B. M., Huang, W., Lv, H., Liao, S., Zeng, S., & Wang, Y. (2016). Comparative profiling of miRNAs and target gene identification in distant-grafting between tomato and lycium (goji berry). *Frontiers in Plant Science*, *7*(OCTOBER2016), 1–18. <https://doi.org/10.3389/fpls.2016.01475>
- Kidner, C. A., & Martienssen, R. A. (2005). The developmental role of microRNA in plants. *Current Opinion in Plant Biology*, *8*(1), 38–44. <https://doi.org/10.1016/j.pbi.2004.11.008>
- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. <https://doi.org/10.1038/nmeth.3317>
- Kim, J.-E., Oh, S.-K., Lee, J.-H., Lee, B.-M., & Jo, S.-H. (2014). Genome-wide SNP calling using next generation sequencing data in tomato. *Molecules and Cells*, *37*(1), 36.
- Kimball, D. A. (1999). Citrus processing: A complete guide. *Springer Science & Business Media*.
- Kimura, S., & Sinha, N. (2008). Tomato (*Solanum lycopersicum*): A model fruit-bearing crop. *Cold Spring Harbor Protocols*, *3*(11). <https://doi.org/10.1101/pdb.emo105>
- Klee, H. J. (2010). Improving the flavor of fresh fruits: Genomics, biochemistry, and biotechnology. *New Phytologist*, *187*(1), 44–56. <https://doi.org/10.1111/j.1469-8137.2010.03281.x>
- Kliewer, W. M. (1977). Influence of temperature, solar radiation and nitrogen on coloration and composition of Emperor grapes. *American Journal of Enology and Viticulture*, *28*(2), 96–103.
- Knapp, S., Bohs, L., Nee, M., & Spooner, D. M. (2004). Solanaceae - A model for linking genomics with

- biodiversity. *Comparative and Functional Genomics*, 5(3), 285–291. <https://doi.org/10.1002/cfg.393>
- Kobayashi, S., Ishimaru, M., Hiraoka, K., & Honda, C. (2002). Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis. *Planta*, 215(6), 924–933. <https://doi.org/10.1007/s00425-002-0830-5>
- Kosma, D. K., Parsons, E. P., Isaacson, T., Lü, S., Rose, J. K. C., & Jenks, M. A. (2010). Fruit cuticle lipid composition during development in tomato ripening mutants. *Physiologia Plantarum*, 139(1), 107–117. <https://doi.org/10.1111/j.1399-3054.2009.01342.x>
- Kozomara, A., & Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, 42(D1), 68–73. <https://doi.org/10.1093/nar/gkt1181>
- Kravchik, M., Stav, R., Belausov, E., & Arazi, T. (2019). Functional characterization of microRNA171 family in Tomato. *Plants*, 8(1). <https://doi.org/10.3390/plants8010010>
- Kumar, A., Singh, B., & Singh, K. (2015). Functional characterization of flavanone 3-hydroxylase gene from *Phyllanthus emblica* (L.). *Journal of Plant Biochemistry and Biotechnology*, 24(4), 453–460. <https://doi.org/10.1007/s13562-014-0296-0>
- Kumar, R., Tyagi, A. K., & Sharma, A. K. (2011). Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. *Molecular Genetics and Genomics*, 285(3), 245–260. <https://doi.org/10.1007/s00438-011-0602-7>
- Kwon, Y. R., Lee, H. J., Kim, K. H., Hong, S. W., Lee, S. J., & Lee, H. (2008). Ectopic expression of Expansin3 or Expansinβ1 causes enhanced hormone and salt stress sensitivity in Arabidopsis. *Biotechnology Letters*, 30(7), 1281–1288. <https://doi.org/10.1007/s10529-008-9678-5>
- Lacroix, M. (2007). The use of essential oils and bacteriocins as natural antimicrobial and antioxidant compounds. *Food*, 2, 181–192.
- Lado, J., Alós, E., Manzi, M., Cronje, P. J. R., Gómez-Cadenas, A., Rodrigo, M. J., & Zacarías, L. (2019). Light regulation of carotenoid biosynthesis in the peel of mandarin and sweet orange fruits. *Frontiers in Plant Science*, 10(October), 1–16. <https://doi.org/10.3389/fpls.2019.01288>
- Langmead, B. (2010). Aligning short sequencing reads with Bowtie. *Current Protocols in Bioinformatics*, SUPP.32, 1–14. <https://doi.org/10.1002/0471250953.bi1107s32>
- Lara, I., Heredia, A., & Domínguez, E. (2019). Shelf life potential and the fruit cuticle: The unexpected player. *Frontiers in Plant Science*, 10(June). <https://doi.org/10.3389/fpls.2019.00770>
- Law, J. A., Du, J., Hale, C. J., Feng, S., Krajewski, K., Palanca, A. M. S., Strahl, B. D., Patel, D. J., & Jacobsen, S. E. (2013). Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature*, 498(7454), 385–389. <https://doi.org/10.1038/nature12178>
- Law, J. a, & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.*, 11(3), 204–220. <https://doi.org/10.1038/nrg2719>.Establishing
- Lee, D.-K., Ahn, J. H., Song, S.-K., Choi, Y. Do, & Lee, J. S. (2003). Expression of an expansin gene is correlated with root elongation in soybean. *Plant Physiology*, 131(3), 985–997.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., & Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425(6956), 415–419. <https://doi.org/10.1038/nature01957>
- Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO Journal*, 23(20), 4051–4060. <https://doi.org/10.1038/sj.emboj.7600385>
- Leubner-Metzger, G. (2003). Functions and regulation of β-1,3-glucanases during seed germination, dormancy release and after-ripening. *Seed Science Research*, 13(1), 17–34. <https://doi.org/10.1079/ssr2002121>
- Li, C., Ng, C. K. Y., & Fan, L. M. (2015). MYB transcription factors; active players in abiotic stress signaling. *Environmental and Experimental Botany*, 114, 80–91. <https://doi.org/10.1016/j.envexpbot.2014.06.014>
- Li, D., Mou, W., Xia, R., Li, L., Zawora, C., Ying, T., Mao, L., Liu, Z., & Luo, Z. (2019). Integrated analysis of high-throughput sequencing data shows abscisic acid-responsive genes and miRNAs in strawberry receptacle fruit ripening. *Horticulture Research*, 6(1), 1–13. <https://doi.org/10.1038/s41438-018-0100-8>
- Li, He, Mao, W., Liu, W., Dai, H., Liu, Y., Ma, Y., & Zhang, Z. (2013). Deep sequencing discovery of novel and

- conserved microRNAs in wild type and a white-flesh mutant strawberry. *Planta*, 238(4), 695–713. <https://doi.org/10.1007/s00425-013-1917-x>
- Li, Heng, Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, M., Wang, X., Li, C., Li, H., Zhang, J., & Ye, Z. (2018). Silencing GRAS2 reduces fruit weight in tomato. *Journal of Integrative Plant Biology*, 60(6), 498–513. <https://doi.org/10.1111/jipb.12636>
- Li, R., Fu, D., Zhu, B., Luo, Y., & Zhu, H. (2018). CRISPR/Cas9-mediated mutagenesis of IncRNA1459 alters tomato fruit ripening. *Plant Journal*, 94(3), 513–524. <https://doi.org/10.1111/tbj.13872>
- Li, S. B., Xie, Z. Z., Hu, C. G., & Zhang, J. Z. (2016). A review of auxin response factors (ARFs) in plants. *Frontiers in Plant Science*, 7(FEB2016), 1–7. <https://doi.org/10.3389/fpls.2016.00047>
- Li, Xindi, Wang, Y., Chen, S., Tian, H., Fu, D., Zhu, B., Luo, Y., & Zhu, H. (2018). Lycopene is enriched in tomato fruit by CRISPR/Cas9-mediated multiplex genome editing. *Frontiers in Plant Science*, 9, 559.
- Li, Xuyan, Xie, X., Li, J., Cui, Y., Hou, Y., Zhai, L., Wang, X., Fu, Y., Liu, R., & Bian, S. (2017). Conservation and diversification of the miR166 family in soybean and potential roles of newly identified miR166s. *BMC Plant Biology*, 17(1), 1–18. <https://doi.org/10.1186/s12870-017-0983-9>
- Li, Z., Peng, R., Tian, Y., Han, H., Xu, J., & Yao, Q. (2016). Genome-wide identification and analysis of the MYB transcription factor superfamily in *Solanum lycopersicum*. *Plant and Cell Physiology*, 57(8), 1657–1677. <https://doi.org/10.1093/pcp/pcw091>
- Liang, Y., Guan, Y., Wang, S., Li, Y., Zhang, Z., & Li, H. (2018). Identification and characterization of known and novel microARNs in strawberry fruits induced by *Botrytis cinerea*. *Scientific Reports*, 8(1), 1–17. <https://doi.org/10.1038/s41598-018-29289-7>
- Liao, Y., Smyth, G. K., & Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Liener, I. E. (1994). Implications of antinutritional components in soybean foods. In *Critical Reviews in Food Science and Nutrition* (Vol. 34, Issue 1). <https://doi.org/10.1080/10408399409527649>
- Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T. K., Espley, R. V., Hellens, R. P., & Allan, A. C. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biology*, 10(1), 1–17.
- Liu, C., Axtell, M. J., & Fedoroff, N. V. (2012). The helicase and RNaseIII domains of Arabidopsis Dicer-Like1 modulate catalytic parameters during MicroRNA biogenesis. *Plant Physiology*, 159(2), 748–758. <https://doi.org/10.1104/pp.112.193508>
- Liu, Hangkong, Qian, M., Song, C., Li, J., Zhao, C., Li, G., Wang, A., & Han, M. (2018). Down-regulation of PpBGAL10 and PpBGAL16 delays fruit softening in peach by reducing polygalacturonase and pectin methylesterase activity. *Frontiers in Plant Science*, 9(July), 1–11. <https://doi.org/10.3389/fpls.2018.01015>
- Liu, Hongjun, Qin, C., Chen, Z., Zuo, T., Yang, X., Zhou, H., Xu, M., Cao, S., Shen, Y., Lin, H., He, X., Zhang, Y., Li, L., Ding, H., Lübberstedt, T., Zhang, Z., & Pan, G. (2014). Identification of miRNAs and their target genes in developing maize ears by combined small RNA and degradome sequencing. *BMC Genomics*, 15(1). <https://doi.org/10.1186/1471-2164-15-25>
- Liu, J., Osbourn, A., & Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Molecular Plant*, 8(5), 689–708. <https://doi.org/10.1016/j.molp.2015.03.012>
- Liu, M., Chen, Y., Chen, Y., Shin, J. H., Mila, I., Audran, C., Zouine, M., Pirrello, J., & Bouzayen, M. (2018). The tomato Ethylene Response Factor Sl-ERF.B3 integrates ethylene and auxin signaling via direct regulation of Sl-Aux/IAA27. *New Phytologist*, 219(2), 631–640. <https://doi.org/10.1111/nph.15165>
- Liu, M., Gomes, B. L., Mila, I., Purgatto, E., Peres, L. E. P., Frasse, P., Maza, E., Zouine, M., Roustan, J. P., Bouzayen, M., & Pirrello, J. (2016). Comprehensive profiling of ethylene response factor expression identifies ripening-associated ERF genes and their link to key regulators of fruit ripening in tomato. *Plant Physiology*, 170(3), 1732–1744. <https://doi.org/10.1104/pp.15.01859>
- Liu, R. H. (2013). Health-promoting components of fruits and vegetables in the diet. *Advances in Nutrition*, 4(3).

<https://doi.org/10.3945/an.112.003517>

Liu, W., Duttke, S. H., Hetzel, J., Groth, M., Feng, S., Gallego-Bartolome, J., Zhong, Z., Kuo, H. Y., Wang, Z., Zhai, J., Chory, J., & Jacobsen, S. E. (2018). RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase  $\nu$  transcripts in Arabidopsis. *Nature Plants*, 4(3), 181–188. <https://doi.org/10.1038/s41477-017-0100-y>

Liu, Y., Teng, C., Xia, R., & Meyers, B. C. (2020). PhasiRNAs in Plants: Their Biogenesis, genic sources, and roles in stress responses, development, and reproduction. *Plant Cell*, 32(10), 3059–3080. <https://doi.org/10.1105/tpc.20.00335>

Liu, Y., Wang, L., Chen, D., Wu, X., Huang, D., Chen, L., Li, L., Deng, X., & Xu, Q. (2014). Genome-wide comparison of microRNAs and their targeted transcripts among leaf, flower and fruit of sweet orange. *BMC Genomics*, 15(1), 1–15. <https://doi.org/10.1186/1471-2164-15-695>

López-Galiano, M. J., García-Robles, I., González-Hernández, A. I., Camañes, G., Vicedo, B., Real, M. D., & Rausell, C. (2019). Expression of miR159 is altered in tomato plants undergoing drought stress. *Plants*, 8(7), 1–11. <https://doi.org/10.3390/plants8070201>

Lopez-Gomollon, S., Mohorianu, I., Szitty, G., Moulton, V., & Dalmay, T. (2012). Diverse correlation patterns between microRNAs and their targets during tomato fruit development indicates different modes of microRNA actions. *Planta*, 236(6), 1875–1887. <https://doi.org/10.1007/s00425-012-1734-7>

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1–21. <https://doi.org/10.1186/s13059-014-0550-8>

Lu, H., Luo, Z., Wang, L., Liu, W., Li, D., Belwal, T., Xu, Y., & Li, L. (2020). FaMYB9 is involved in the regulation of C6 volatile biosynthesis in strawberry. *Plant Science*, 293(January). <https://doi.org/10.1016/j.plantsci.2020.110422>

Luan, W., Dai, Y., Li, X. Y., Wang, Y., Tao, X., Li, C. X., Mao, P., & Ma, X. R. (2020). Identification of tRFs and phasiRNAs in tomato (*Solanum lycopersicum*) and their responses to exogenous abscisic acid. *BMC Plant Biology*, 20(1), 1–14. <https://doi.org/10.1186/s12870-020-02528-8>

Luan, Y., Cui, J., Zhai, J., Li, J., Han, L., & Meng, J. (2015). High-throughput sequencing reveals differential expression of miRNAs in tomato inoculated with *Phytophthora infestans*. *Planta*, 241(6), 1405–1416. <https://doi.org/10.1007/s00425-015-2267-7>

Luan, Y., Wang, W., & Liu, P. (2014). Identification and functional analysis of novel and conserved microRNAs in tomato. *Molecular Biology Reports*, 41(8), 5385–5394. <https://doi.org/10.1007/s11033-014-3410-4>

Lucero Reinoso, Ariel y Quiroga, Guillermo - Argentina's potential for the processing tomato industry. The Tomato News Conference, Avignon, 2019. Disponible online en: <http://www.tomatonews.com/maj/phototheque/photos/PDF/Argentina-Avignon%202019.pdf>

Lucini, L., Baccolo, G., Roupheal, Y., Colla, G., Bavaresco, L., & Trevisan, M. (2018). Chitosan treatment elicited defence mechanisms, pentacyclic triterpenoids and stilbene accumulation in grape (*Vitis vinifera* L.) bunches. *Phytochemistry*, 156(August), 1–8. <https://doi.org/10.1016/j.phytochem.2018.08.011>

Luo, Q.-J., Mittal, A., Jia, F., & Rock, C. D. (2012). An autoregulatory feedback loop involving PAP1 and TAS4 in response to sugars in Arabidopsis. *Plant Molecular Biology*, 80(1), 117–129.

Luo, Y., Ge, C., Ling, Y., Mo, F., Yang, M., Jiang, L., Chen, Q., Lin, Y., Sun, B., Zhang, Y., Wang, Y., Li, M., Wang, X., & Tang, H. (2020). ABA and sucrose co-regulate strawberry fruit ripening and show inhibition of glycolysis. *Molecular Genetics and Genomics*, 295(2), 421–438. <https://doi.org/10.1007/s00438-019-01629-w>

Mahjoub, A., Hernould, M., Joubès, J., Decendit, A., Mars, M., Barrieu, F., Hamdi, S., & Delrot, S. (2009). Overexpression of a grapevine R2R3-MYB factor in tomato affects vegetative development, flower morphology and flavonoid and terpenoid metabolism. *Plant Physiology and Biochemistry*, 47(7), 551–561. <https://doi.org/10.1016/j.plaphy.2009.02.015>

Mallory, A. C., & Vaucheret, H. (2006). Functions of microRNAs and related small RNAs in plants. 38(June), 31–37. <https://doi.org/10.1038/ng1791>

Manríquez, D., El-Sharkawy, I., Flores, F. B., El-Yahyaoui, F., Regad, F., Bouzayen, M., Latché, A., & Pech, J. C. (2006). Two highly divergent alcohol dehydrogenases of melon exhibit fruit ripening-specific expression and distinct biochemical characteristics. *Plant Molecular Biology*, 61(4–5), 675–685.

<https://doi.org/10.1007/s11103-006-0040-9>

Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J., & Waterhouse, P. M. (2006). The evolution and diversification of Dicers in plants. *FEBS Letters*, *580*(10), 2442–2450. <https://doi.org/10.1016/j.febslet.2006.03.072>

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., ... Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), 376–380. <https://doi.org/10.1038/nature03959>

Marín-Rodríguez, M. C., Orchard, J., & Seymour, G. B. (2002). Pectate lyases, cell wall degradation and fruit softening. *Journal of Experimental Botany*, *53*(377), 2115–2119. <https://doi.org/10.1093/jxb/erf089>

Marmisolle, F. E., Arizmendi, A., Ribone, A., Rivarola, M., García, M. L., & Reyes, C. A. (2020). Up-regulation of microRNA targets correlates with symptom severity in *Citrus sinensis* plants infected with two different isolates of citrus psorosis virus. *Planta*, *251*(1), 1–11. <https://doi.org/10.1007/s00425-019-03294-0>

Marsh, K. B., Volz, R. K., Cashmore, W., & Reay, P. (1996). Fruit colour, leaf nitrogen level, and tree vigour in 'Fuji' apples. *New Zealand Journal of Crop and Horticultural Science*, *24*(4), 393–399. <https://doi.org/10.1080/01140671.1996.9513976>

Martínez, M., Cambra, I., González-Melendi, P., Santamaría, M. E., & Díaz, I. (2012). C1A cysteine-proteases and their inhibitors in plants. *Physiologia Plantarum*, *145*(1), 85–94.

Maxam, A. M., & Gilbert, W. (1992). A new method for sequencing DNA. 1977. *Biotechnology (Reading, Mass.)*, *24*(2), 99–103.

McAvoy, G. (2008). Cultivar selection in tomato and pepper production. *EDIS*, *2008*(1).

McQueen-Mason, S. J., & Cosgrove, D. J. (1995). Expansin mode of action on cell walls: Analysis of wall hydrolysis, stress relaxation, and binding. *Plant Physiology*, *107*(1), 87–100. <https://doi.org/10.1104/pp.107.1.87>

Medina-Puche, L., Cumplido-Laso, G., Amil-Ruiz, F., Hoffmann, T., Ring, L., Rodríguez-Franco, A., Caballero, J. L., Schwab, W., Muñoz-Blanco, J., & Blanco-Portales, R. (2014). MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria × ananassa* fruits. *Journal of Experimental Botany*, *65*(2), 401–417. <https://doi.org/10.1093/jxb/ert377>

Medina-Puche, L., Molina-Hidalgo, F. J., Boersma, M., Schuurink, R. C., López-Vidriero, I., Solano, R., Franco-Zorrilla, J. M., Caballero, J. L., Blanco-Portales, R., & Muñoz-Blanco, J. (2015). An R2R3-MYB transcription factor regulates eugenol production in ripe strawberry fruit receptacles. *Plant Physiology*, *168*(2), 598–614. <https://doi.org/10.1104/pp.114.252908>

Mercado, J. A., Quesada, M. A., Blanco-Portales, R., Pose, S., García-Gago, J. A., Jiménez-Bermúdez, S., Muñoz-Serrano, A., Caballero, J. L., Pliego-Alfaro, F., & Muñoz-Blanco, J. (2009). Antisense down-regulation of the FaPG1 gene reveals an unexpected central role for polygalacturonase in strawberry fruit softening1[W]. *Plant Physiology*, *150*(2), 1022–1032. <https://doi.org/10.1104/pp.109.138297>

Merelo, P., Agustí, J., Arbona, V., Costa, M. L., Estornell, L. H., Gómez-Cadenas, A., Coimbra, S., Gómez, M. D., Pérez-Amador, M. A., Domingo, C., Talón, M., & Tadeo, F. R. (2017). Cell wall remodeling in abscission zone cells during ethylene-promoted fruit abscission in citrus. *Frontiers in Plant Science*, *8*(February). <https://doi.org/10.3389/fpls.2017.00301>

Mesejo, C., Marzal, A., Martínez-Fuentes, A., Reig, C., & Agustí, M. (2021). On how auxin, ethylene and IDA-peptide relate during mature Citrus fruit abscission. *Scientia Horticulturae*, *278*, 109855.

Meyer, D. J. L. (2005). 7 Seed development and structure in floral crops. *Www. Ketabdownload. Com*, 117.

Meyers, K. J., Watkins, C. B., Pritts, M. P., & Liu, R. H. (2003). Antioxidant and antiproliferative activities of strawberries. *Journal of Agricultural and Food Chemistry*, *51*(23), 6887–6892. <https://doi.org/10.1021/jf034506n>

Millar, A. A., Lohe, A., & Wong, G. (2019). Biology and function of miR159 in plants. *Plants*, *8*(8). <https://doi.org/10.3390/plants8080255>

Minoia, S., Boualem, A., Marcel, F., Troadec, C., Quemener, B., Cellini, F., Petrozza, A., Vigouroux, J., Lahaye, M., Carriero, F., & Bendahmane, A. (2015). Induced mutations in tomato S1Exp1 alter cell wall metabolism and

- delay fruit softening. *Plant Science*, 242, 195–202. <https://doi.org/10.1016/j.plantsci.2015.07.001>
- Mohorianu, I., Schwach, F., Jing, R., Lopez-Gomollon, S., Moxon, S., Szittyá, G., Sorefan, K., Moulton, V., & Dalmay, T. (2011). Profiling of short RNAs during fleshy fruit development reveals stage-specific sRNAome expression patterns. *Plant Journal*, 67(2), 232–246. <https://doi.org/10.1111/j.1365-313X.2011.04586.x>
- Mol, J., Grotewold, E., & Koes, R. (1998). How genes paint flowers and seeds. *1385*(98).
- Montgomery, T. A., Howell, M. D., Cuperus, J. T., Li, D., Hansen, J. E., Alexander, A. L., Chapman, E. J., Fahlgren, N., Allen, E., & Carrington, J. C. (2008). Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 Trans-Acting siRNA formation. *Cell*, 133(1), 128–141. <https://doi.org/10.1016/j.cell.2008.02.033>
- Moussian, B., Schoof, H., Haecker, A., Jürgens, G., & Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. *EMBO Journal*, 17(6), 1799–1809. <https://doi.org/10.1093/emboj/17.6.1799>
- Moxon, S., Jing, R., Szittyá, G., Schwach, F., Rusholme Pilcher, R. L., Moulton, V., & Dalmay, T. (2008). Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Research*, 18(10), 1602–1609. <https://doi.org/10.1101/gr.080127.108>
- Muñoz-Bertomeu, J., Miedes, E., & Lorences, E. P. (2013). Expression of xyloglucan endotransglucosylase/hydrolase (XTH) genes and XET activity in ethylene treated apple and tomato fruits. *Journal of Plant Physiology*, 170(13), 1194–1201. <https://doi.org/10.1016/j.jplph.2013.03.015>
- Nag, A., & Jack, T. (2010). Sculpting the flower; the role of microRNAs in flower development. *Current Topics in Developmental Biology*, 91(C), 349–378. [https://doi.org/10.1016/S0070-2153\(10\)91012-0](https://doi.org/10.1016/S0070-2153(10)91012-0)
- Nag, A., King, S., & Jack, T. (2009). miR319a targeting of TCP4 is critical for petal growth and development in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 106(52), 22534–22539. <https://doi.org/10.1073/pnas.0908718106>
- Naika, S., Jeude, J. van L. de, Goffau, M. de, Hilmi, M., & Dam, B. Van. (2005). *Cultivation of tomato* (Issue 17).
- Nakamura, M., Köhler, C., & Hennig, L. (2019). Tissue-specific transposon-associated small RNAs in the gymnosperm tree, Norway spruce. *BMC Genomics*, 20(1), 1–10. <https://doi.org/10.1186/s12864-019-6385-7>
- Nakano, T., Kimbara, J., Fujisawa, M., Kitagawa, M., Ihashi, N., Maeda, H., Kasumi, T., & Ito, Y. (2012). MACROCALYX and JOINTLESS interact in the transcriptional regulation of tomato fruit abscission zone development. *Plant Physiology*, 158(1), 439–450.
- Nakashima, K., Takasaki, H., Mizoi, J., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2012). NAC transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1819(2), 97–103. <https://doi.org/10.1016/j.bbagr.2011.10.005>
- Napoli, C., Lemieux, C., & Jorgensen, R. (1990). Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell*, 2(4), 279–289. <https://doi.org/10.2307/3869076>
- Negri, A. S., Allegra, D., Simoni, L., Rusconi, F., Tonelli, C., Espen, L., & Galbiati, M. (2015). Comparative analysis of fruit aroma patterns in the domesticated wild strawberries “Profumata di Tortona” (*F. moschata*) and “Regina delle Valli” (*F. vesca*). *Frontiers in Plant Science*, 6(FEB), 1–13. <https://doi.org/10.3389/fpls.2015.00056>
- Nicolosi, E., Deng, Z. N., Gentile, A., La Malfa, S., Continella, G., & Tribulato, E. (2000). Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theoretical and Applied Genetics*, 100(8), 1155–1166.
- Nitsch, J. P. (1950). Growth and morphogenesis of the strawberry as related to auxin. *American Journal of Botany*, 37(3), 211–215. <https://doi.org/10.1002/j.1537-2197.1950.tb12183.x>
- Niu, Y., Zhao, T., Xu, X., & Li, J. (2017). Genome-wide identification and characterization of GRAS transcription factors in tomato (*Solanum lycopersicum*). *PeerJ*, 5, e3955. <https://doi.org/10.7717/peerj.3955>
- Noman, A., Fahad, S., Aqeel, M., Ali, U., Amanullah, Anwar, S., Baloch, S. K., & Zainab, M. (2017). miRNAs: Major modulators for crop growth and development under abiotic stresses. *Biotechnology Letters*, 39(5), 685–700. <https://doi.org/10.1007/s10529-017-2302-9>
- Nour, V., Trandafir, I., & Ionica, M. E. (2013). Antioxidant compounds, mineral content and antioxidant activity of several tomato cultivars grown in Southwestern Romania. *Notulae Botanicae Horti Agrobotanici Cluj-*

*Napoca*, 41(1), 136–142. <https://doi.org/10.15835/nbha4119026>

Nuthikattu, S., McCue, A. D., Panda, K., Fultz, D., DeFraia, C., Thomas, E. N., & Keith Slotkin, R. (2013). The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiology*, 162(1), 116–131. <https://doi.org/10.1104/pp.113.216481>

Oeller, P. W., Min-Wong, L. U., Taylor, L. P., Pike, D. A., & Theologis, A. (1991). Reversible inhibition of tomato fruit senescence by antisense RNA. *Science*, 254(5030), 437–439. <https://doi.org/10.1126/science.1925603>

Oltman, A. E., Yates, M. D., & Drake, M. A. (2016). Preference mapping of fresh tomatoes across 3 stages of consumption. *Journal of Food Science*, 81(6), S1495–S1505.

P. Darley, C., M. Forrester, A., & J. Mcqueen-Mason, S. (2001). The molecular basis of plant cell wall extension. *Plant Molecular Biology*, 47(1–2), 179–195. <https://doi.org/10.1023/A:1010687600670>

Pandurangaiyah, S., Ravishankar, K. V., Shivashankar, K. S., Sadashiva, A. T., Pillakenchappa, K., & Narayanan, S. K. (2016). Differential expression of carotenoid biosynthetic pathway genes in two contrasting tomato genotypes for lycopene content. *Journal of Biosciences*, 41(2), 257–264. <https://doi.org/10.1007/s12038-016-9602-4>

Paniagua, C., Blanco-Portales, R., Barceló-Muñoz, M., García-Gago, J. A., Waldron, K. W., Quesada, M. A., Muñoz-Blanco, J., & Mercado, J. A. (2016). Antisense down-regulation of the strawberry  $\beta$ -galactosidase gene Fa $\beta$ Gal4 increases cell wall galactose levels and reduces fruit softening. *Journal of Experimental Botany*, 67(3), 619–631. <https://doi.org/10.1093/jxb/erv462>

Pantaleão, L. C., & Ozanne, S. E. (2018). Small RNA sequencing: A technique for miRNA profiling. *Methods in Molecular Biology*, 1735, 321–330. [https://doi.org/10.1007/978-1-4939-7614-0\\_21](https://doi.org/10.1007/978-1-4939-7614-0_21)

Parapunova, V., Busscher, M., Busscher-Lange, J., Lammers, M., Karlova, R., Bovy, A. G., Angenent, G. C., & De Maagd, R. A. (2014). Identification, cloning and characterization of the tomato TCP transcription factor family. *BMC Plant Biology*, 14(1), 1–17. <https://doi.org/10.1186/1471-2229-14-157>

Park, P. J. (2009). ChIP-seq: Advantages and challenges of a maturing technology. *Nature Reviews Genetics*, 10(10), 669–680. <https://doi.org/10.1038/nrg2641>

Patra, B., Schluttenhofer, C., Wu, Y., Pattanaik, S., & Yuan, L. (2013). Transcriptional regulation of secondary metabolite biosynthesis in plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1829(11), 1236–1247.

Patterson, S. E. (2001). Cutting loose. Abscission and dehiscence in Arabidopsis. *Plant Physiology*, 126(2), 494–500.

Paul, V., Pandey, R., & Srivastava, G. C. (2012). The fading distinctions between classical patterns of ripening in climacteric and non-climacteric fruit and the ubiquity of ethylene-An overview. *Journal of Food Science and Technology*, 49(1), 1–21. <https://doi.org/10.1007/s13197-011-0293-4>

Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A., & Xiong, G. (2013). Hemicellulose biosynthesis. *Planta*, 238(4), 627–642. <https://doi.org/10.1007/s00425-013-1921-1>

Peled-Zehavi, H., Oliva, M., Xie, Q., Tzin, V., Oren-Shamir, M., Aharoni, A., & Galili, G. (2015). Metabolic engineering of the phenylpropanoid and its primary, precursor pathway to enhance the flavor of fruits and the aroma of flowers. *Bioengineering*, 2(4), 204–212. <https://doi.org/10.3390/bioengineering2040204>

Pérez, A. G., & Sanz, C. (2008). Formation of fruit flavour. In *Fruit and Vegetable Flavour: Recent Advances and Future Prospects*. Woodhead Publishing Limited. <https://doi.org/10.1533/9781845694296.2.41>

Perveen, R., Suleria, H. A. R., Anjum, F. M., Butt, M. S., Pasha, I., & Ahmad, S. (2015). Tomato (*Solanum lycopersicum*) carotenoids and lycopenes chemistry; metabolism, absorption, nutrition, and allied health claims—a comprehensive review. *Critical Reviews in Food Science and Nutrition*, 55(7), 919–929. <https://doi.org/10.1080/10408398.2012.657809>

Pesaresi, P., Mizzotti, C., Colombo, M., & Masiero, S. (2014). Genetic regulation and structural changes during tomato fruit development and ripening. *Frontiers in Plant Science*, 5(APR), 1–14. <https://doi.org/10.3389/fpls.2014.00124>

Petit, J., Bres, C., Just, D., Garcia, V., Mauxion, J. P., Marion, D., Bakan, B., Joubès, J., Domergue, F., & Rothan, C. (2014). Analyses of tomato fruit brightness mutants uncover both cutin-deficient and cutin-abundant mutants and a new hypomorphic allele of GDSL Lipase. *Plant Physiology*, 164(2), 888–906.

<https://doi.org/10.1104/pp.113.232645>

Peumans, W. J., Barre, A., Derycke, V., Rougé, P., Zhang, W., May, G. D., Delcour, J. A., Van Leuven, F., & Van Damme, E. J. M. (2000). Purification, characterization and structural analysis of an abundant  $\beta$ -1,3-glucanase from banana fruit. *European Journal of Biochemistry*, 267(4), 1188–1195. <https://doi.org/10.1046/j.1432-1327.2000.01117.x>

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29(9), 2001. <https://doi.org/10.1093/nar/29.9.e45>

Phan, N. T., Trinh, L. T., Rho, M.-Y., Park, T.-S., Kim, O.-R., Zhao, J., Kim, H.-M., & Sim, S.-C. (2019). Identification of loci associated with fruit traits using genome-wide single nucleotide polymorphisms in a core collection of tomato (*Solanum lycopersicum* L.). *Scientia Horticulturae*, 243, 567–574.

Pirrello, J., Jaimes-Miranda, F., Sanchez-Ballesta, M. T., Tournier, B., Khalil-Ahmad, Q., Regad, F., Latché, A., Pech, J. C., & Bouzayen, M. (2006). Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. *Plant and Cell Physiology*, 47(9), 1195–1205. <https://doi.org/10.1093/pcp/pcj084>

Pott, D. M., Osorio, S., & Vallarino, J. G. (2019). From central to specialized metabolism: An overview of some secondary compounds derived from the primary metabolism for their role in conferring nutritional and organoleptic characteristics to fruit. *Frontiers in Plant Science*, 10(June). <https://doi.org/10.3389/fpls.2019.00835>

Poulsen, C., Vaucheret, H., & Brodersen, P. (2013). Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell*, 25(1), 22–37. <https://doi.org/10.1105/tpc.112.105643>

Prange, R. K., & DeEll, J. R. (2019). Preharvest factors affecting postharvest quality of berry crops. *HortScience*, 30(4), 751B – 751. <https://doi.org/10.21273/hortsci.30.4.751b>

Prasanna, V., Prabha, T. N., & Tharanathan, R. N. (2007). Fruit ripening phenomena-an overview. *Critical Reviews in Food Science and Nutrition*, 47(1), 1–19. <https://doi.org/10.1080/10408390600976841>

Prat, S., Frommer, W. B., Höfgen, R., Keil, M., Koßmann, J., Köster-Töpfer, M., Liu, X. J., Müller, B., Pēna-Cortés, H., Rocha-Sosa, M., Sánchez-Serrano, J. J., Sonnewald, U., & Willmitzer, L. (1990). Gene expression during tuber development in potato plants. *FEBS Letters*, 268(2), 334–338. [https://doi.org/10.1016/0014-5793\(90\)81281-R](https://doi.org/10.1016/0014-5793(90)81281-R)

Prigigallo, M. I., Križnik, M., De Paola, D., Catalano, D., Gruden, K., Finetti-Sialer, M. M., & Cillo, F. (2019). Potato virus Y infection alters small RNA metabolism and immune response in tomato. *Viruses*, 11(12), 1–27. <https://doi.org/10.3390/v11121100>

Pumplin, N., & Voinnet, O. (2013). RNA silencing suppression by plant pathogens: Defence, counter-defence and counter-counter-defence. *Nature Reviews Microbiology*, 11(11), 745–760. <https://doi.org/10.1038/nrmicro3120>

Qian, Y., Cheng, Y., Cheng, X., Jiang, H., Zhu, S., & Cheng, B. (2011). Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Reports*, 30(7), 1347–1363. <https://doi.org/10.1007/s00299-011-1046-6>

Qin, L., Mo, N., Muhammad, T., & Liang, Y. (2018). Genome-wide analysis of DCL, AGO, and RDR gene families in pepper (*Capsicum Annuum* L.). *International Journal of Molecular Sciences*, 19(4). <https://doi.org/10.3390/ijms19041038>

Qin, L., Mo, N., Zhang, Y., Muhammad, T., Zhao, G., Zhang, Y., & Yan, L. (2017). CaRDR1, an RNA-Dependent RNA Polymerase Plays a Positive Role in Pepper Resistance against TMV. *Front Plant Sci*, 8, 1–13.

Quiroga, G. o. S., & Fraschina, A. A. (1997). Evaluation of sensory attributes and biochemical parameters in transgenic tomato fruit with reduced polygalacturonase activity. *Food Science and Technology International*, 3(2), 93–102. <https://doi.org/10.1177/108201329700300204>

Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes & Development*, 20(24), 3407–3425.

Ramsay, N. A., & Glover, B. J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Science*, 10(2), 63–70. <https://doi.org/10.1016/j.tplants.2004.12.011>

Reichel, M., Li, Y., Li, J., & Millar, A. A. (2015). Inhibiting plant micro RNA activity: molecular SPONGE s, target MIMIC s and STTM s all display variable efficacies against target micro RNA s. *Plant Biotechnology Journal*, 13(7), 915–926.

- Rick, C. M., & Holle, M. (1990). Andean *Lycopersicon esculentum* var. cerasiforme: genetic variation and its evolutionary significance. *Economic Botany*, 44(3 Supplement), 69–78. <https://doi.org/10.1007/BF02860476>
- Riederer, M., & Schreiber, L. (2001). Protecting against water loss: Analysis of the barrier properties of plant cuticles. *Journal of Experimental Botany*, 52(363), 2023–2032. <https://doi.org/10.1093/jexbot/52.363.2023>
- Rivas, F. V., Tolia, N. H., Song, J. J., Aragon, J. P., Liu, J., Hannon, G. J., & Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Structural and Molecular Biology*, 12(4), 340–349. <https://doi.org/10.1038/nsmb918>
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Rock, C. D. (2013). Trans-acting small interfering RNA4: Key to nutraceutical synthesis in grape development? *Trends in Plant Science*, 18(11), 601–610. <https://doi.org/10.1016/j.tplants.2013.07.006>
- Rodriguez, R. E., Ercoli, M. F., Debernardi, J. M., & Palatnik, J. F. (2016). Growth-Regulating Factors, A transcription factor family regulating more than just plant growth. In *Plant Transcription Factors: Evolutionary, Structural and Functional Aspects*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-800854-6.00017-8>
- Ronen, G., Carmel-Goren, L., Zamir, D., & Hirschberg, J. (2000). An alternative pathway to  $\beta$ -carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20), 11102–11107. <https://doi.org/10.1073/pnas.190177497>
- Rutherford, S., & Moore, I. (2002). The Arabidopsis Rab GTPase family: Another enigma variation. *Current Opinion in Plant Biology*, 5(6), 518–528. [https://doi.org/10.1016/S1369-5266\(02\)00307-2](https://doi.org/10.1016/S1369-5266(02)00307-2)
- Sabbione, A., Daurelio, L., Vegetti, A., Talón, M., Tadeo, F., & Dotto, M. (2019). Genome-wide analysis of AGO, DCL and RDR gene families reveals RNA-directed DNA methylation is involved in fruit abscission in *Citrus sinensis*. *BMC Plant Biology*, 19(1), 1–13.
- Sagar, M., Chervin, C., Mila, I., Hao, Y., Roustan, J. P., Benichou, M., Gibon, Y., Biais, B., Maury, P., Latché, A., Pech, J. C., Bouzayen, M., & Zouin, M. (2013). SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. *Plant Physiology*, 161(3), 1362–1374. <https://doi.org/10.1104/pp.113.213843>
- Salentijn, E. M. J., Aharoni, A., Schaart, J. G., Boone, M. J., & Krens, F. A. (2003). Differential gene expression analysis of strawberry cultivars that differ in fruit-firmness. *Physiologia Plantarum*, 118(4), 571–578. <https://doi.org/10.1034/j.1399-3054.2003.00138.x>
- Salinas, M., Xing, S., Höhmann, S., Berndtgen, R., & Huijser, P. (2012). Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. *Planta*, 235(6), 1171–1184. <https://doi.org/10.1007/s00425-011-1565-y>
- Sanger, F., Nicklen, S., & Coulson, A. (1977). DNA sequencing with chain-terminating. *Proc Natl Acad Sci USA*, 74(12), 5463–5467.
- Sanger, Frederick, Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M., & Smith, M. (1977). Nucleotide sequence of bacteriophage  $\phi$ X174 DNA. *Nature*, 265(5596), 687–695.
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., Kaneko, T., Nakamura, Y., Shibata, D., Aoki, K., Egholm, M., Knight, J., Bogden, R., Li, C., Shuang, Y., Xu, X., Pan, S., Cheng, S., Liu, X., ... Gianese, G. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485(7400), 635–641. <https://doi.org/10.1038/nature11119>
- Schommer, C., Debernardi, J. M., Bresso, E. G., Rodriguez, R. E., & Palatnik, J. F. (2014). Repression of cell proliferation by miR319-regulated TCP4. *Molecular Plant*, 7(10), 1533–1544. <https://doi.org/10.1093/mp/ssu084>
- Schuch, W., Kanczler, J., Robertson, D., Hobson, G., Tucker, G., Grierson, D., Bright, S., & Bird, C. (1991). Fruit quality characteristics of transgenic tomato fruit with altered polygalacturonase activity. *HortScience*, 26(12), 1517–1520.
- Sexton, R., Palmer, J. M., Whyte, N. A., & Littlejohns, S. (1997). Cellulase, fruit softening and abscission in red raspberry *Rubus idaeus* L. cv Glen Clova. *Annals of Botany*, 80(3), 371–376. <https://doi.org/10.1006/anbo.1997.0465>

- Seymour, G. B., Taylor, J. E., & Tucker, G. A. (2012). *Biochemistry of fruit ripening*. Springer Science & Business Media.
- Shahid, M. N. (2007). Effect of bee wax coating on the organoleptic changes in fruit of sweet orange (*Citrus sinensis* L.) cv." blood red". *Sarhad Journal of Agriculture*, 23(2), 411.
- Sharma, A., Badola, P. K., Bhatia, C., Sharma, D., & Trivedi, P. K. (2020). Primary transcript of miR858 encodes regulatory peptide and controls flavonoid biosynthesis and development in Arabidopsis. *Nature Plants*, 6(10), 1262–1274.
- Sharma, D., Tiwari, M., Pandey, A., Bhatia, C., Sharma, A., & Trivedi, P. K. (2016). MicroRNA858 is a potential regulator of phenylpropanoid pathway and plant development. *Plant Physiology*, 171(2), 944–959. <https://doi.org/10.1104/pp.15.01831>
- Sharma, M. K., Kumar, R., Solanke, A. U., Sharma, R., Tyagi, A. K., & Sharma, A. K. (2010). Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. *Molecular Genetics and Genomics*, 284(6), 455–475.
- Shuai, P., Su, Y., Liang, D., Zhang, Z., Xia, X., & Yin, W. (2016). Identification of phasiRNAs and their drought-responsiveness in *Populus trichocarpa*. *FEBS Letters*, 590(20), 3616–3627. <https://doi.org/10.1002/1873-3468.12419>
- Shuping, D. S. S., & Eloff, J. N. (2017). *The use of plants to protect plants and food against fungal pathogens : A review Shuping and Eloff, Afr J Tradit Complement Altern Med ., ( 2017 ) 14 ( 4 ) : 120-127. 14*, 120–127.
- Silva, G. F. F. E., Silva, E. M., Da Silva Azevedo, M., Guivin, M. A. C., Ramiro, D. A., Figueiredo, C. R., Carrer, H., Peres, L. E. P., & Nogueira, F. T. S. (2014). MicroRNA156-targeted SPL/SBP box transcription factors regulate tomato ovary and fruit development. *Plant Journal*, 78(4), 604–618. <https://doi.org/10.1111/tpj.12493>
- Simon, S. A., & Meyers, B. C. (2011). Small RNA-mediated epigenetic modifications in plants. *Current Opinion in Plant Biology*, 14(2), 148–155. <https://doi.org/10.1016/j.pbi.2010.11.007>
- Simons, T., Sivertsen, H., & Guinard, J.-X. (2018). Mapping the preferences of adult and child consumers for California-grown navel oranges. *HortScience*, 53(5), 661–668.
- Singh, A., Mohorianu, I., Green, D., Dalmay, T., Dasgupta, I., & Mukherjee, S. K. (2019). Artificially induced phased siRNAs promote virus resistance in transgenic plants. *Virology*, 537, 208–215.
- Singh, A., Saraf, S., Dasgupta, I., & Mukherjee, S. K. (2016). Identification and validation of a virus-inducible ta-siRNA-generating TAS4 locus in tomato. *Journal of Biosciences*, 41(1), 109–118. <https://doi.org/10.1007/s12038-016-9590-4>
- Singh, A., Taneja, J., Dasgupta, I., & Mukherjee, S. K. (2015). Development of plants resistant to tomato geminiviruses using artificial trans-acting small interfering RNA. *Molecular Plant Pathology*, 16(7), 724–734.
- Singh, D., Ambroise, A., Haicour, R., Sihachakr, D., & Rajam, M. V. (2014). Increased resistance to fungal wilts in transgenic eggplant expressing alfalfa glucanase gene. *Physiology and Molecular Biology of Plants*, 20(2), 143–150. <https://doi.org/10.1007/s12298-014-0225-7>
- Singh, G. (2019). *Plant systematics: an integrated approach*. CRC Press.
- Singh, K. K., & Reddy, B. S. (2006). Post-harvest physico-mechanical properties of orange peel and fruit. *Journal of Food Engineering*, 73(2), 112–120.
- Singh, R. K., Srivastava, S., Chidley, H. G., Nath, P., & Sane, V. A. (2018). Overexpression of mango alcohol dehydrogenase (MiADH1) mimics hypoxia in transgenic tomato and alters fruit flavor components. *Agri Gene*, 7(October 2017), 23–33. <https://doi.org/10.1016/j.aggene.2017.10.003>
- Slavin, J. L., & Lloyd, B. (2012). Health benefits of fruits and vegetables. *Advances in Nutrition*, 3(4), 506–516.
- Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., & Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*, 321(6071), 674–679.
- Sorrequieta, A., Ferraro, G., Boggio, S. B., & Valle, E. M. (2010). Free amino acid production during tomato fruit ripening: A focus on L-glutamate. *Amino Acids*, 38(5), 1523–1532. <https://doi.org/10.1007/s00726-009-0373-1>
- Sosa-Valencia, G., Palomar, M., Covarrubias, A. A., & Reyes, J. L. (2017). The legume miR1514a modulates a NAC transcription factor transcript to trigger phasiRNA formation in response to drought. *Journal of*

- Experimental Botany*, 68(8), 2013–2026. <https://doi.org/10.1093/jxb/erw380>
- Soule, J., & Grierson, W. (1986). Maturity and grade standards. *Fresh Citrus Fruits; Wardowski, WF, Nagy, S., Grierson, W., Eds*, 23–48.
- Specht, C. D., & Bartlett, M. E. (2009). Flower evolution: The origin and subsequent diversification of the angiosperm flower. *Annual Review of Ecology, Evolution, and Systematics*, 40, 217–243. <https://doi.org/10.1146/annurev.ecolsys.110308.120203>
- Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N. S., Loveys, B., & Schuch, W. (1998). Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. *Plant Physiology*, 117(3), 1047–1058. <https://doi.org/10.1104/pp.117.3.1047>
- Spelt, C., Quattrocchio, F., Mol, J., & Koes, R. (2002). ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. *Plant Cell*, 14(9), 2121–2135. <https://doi.org/10.1105/tpc.003772>
- Stewart, I., & Wheaton, T. A. (1972). Carotenoids in Citrus: Their Accumulation Induced by Ethylene. *Journal of Agricultural and Food Chemistry*, 20(2), 448–449. <https://doi.org/10.1021/jf60180a024>
- Sunkar, R., & Jagadeeswaran, G. (2008). In silico identification of conserved microRNAs in large number of diverse plant species. *BMC Plant Biology*, 8, 1–13. <https://doi.org/10.1186/1471-2229-8-37>
- Sunkar, R., & Zhu, J. (2007). Micro RNAs and short-interfering RNAs in plants. *Journal of Integrative Plant Biology*, 49(6), 817–826.
- Suresh, B. V., Roy, R., Sahu, K., Misra, G., & Chattopadhyay, D. (2014). Tomato genomic resources database: An integrated repository of useful tomato genomic information for basic and applied research. *PLoS ONE*, 9(1), 1–9. <https://doi.org/10.1371/journal.pone.0086387>
- Swingle, W. T. (1967). 05 - The botany of the citrus fruits and their wild relatives as a guide to their use in breeding. *Florida State Horticultural Society, October*, 156–164.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Tanaka, T. (1954). Species problem in citrus. *Japanese Society for the Promotion of Science*, 11(4), 1–107.
- Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), 178–192. <https://doi.org/10.1093/bib/bbs017>
- Tieman, D., Bliss, P., McIntyre, L. M., Blandon-Ubeda, A., Bies, D., Odabasi, A. Z., Rodríguez, G. R., Van Der Knaap, E., Taylor, M. G., Goulet, C., Mageroy, M. H., Snyder, D. J., Colquhoun, T., Moskowicz, H., Clark, D. G., Sims, C., Bartoshuk, L., & Klee, H. J. (2012). The chemical interactions underlying tomato flavor preferences. *Current Biology*, 22(11), 1035–1039. <https://doi.org/10.1016/j.cub.2012.04.016>
- Tirumalai, V., Swetha, C., Nair, A., Pandit, A., & Shivaprasad, P. V. (2019). miR828 and miR858 regulate VvMYB114 to promote anthocyanin and flavonol accumulation in grapes. *Journal of Experimental Botany*, 70(18), 4775–4792.
- Toll, D. (2014). Biosynthesis and biological functions of terpenoids in plants. *Advances in Biochemical Engineering/Biotechnology*, 123(July 2015), 127–141. <https://doi.org/10.1007/10>
- Treangen, T. J., & Salzberg, S. L. (2012). Repetitive DNA and next-generation sequencing: Computational challenges and solutions. *Nature Reviews Genetics*, 13(1), 36–46. <https://doi.org/10.1038/nrg3117>
- Trejo-Téllez, Libia, Gómez-Merino, F. (2004). Nutrient management in strawberry: Effects on yield, quality and plant health. *Strawberries: Cultivation, antioxidant properties and health benefits. November*, 1–8.
- Trudel, M. J., & Ozbun, J. L. (1970). Relationship between chlorophylls and carotenoids of ripening tomato fruit as influenced by potassium nutrition. *Journal of Experimental Botany*, 21(4), 881–886. <https://doi.org/10.1093/jxb/21.4.881>
- Tucker, M. L. (2014). Cell-wall metabolism and softening during ripening. *Fruit Ripening: Physiology, Signalling and Genomics*, 48–62.
- Tucker, M. R., Hinze, A., Tucker, E. J., Takada, S., Jürgens, G., & Laux, T. (2008). Vascular signalling mediated by

ZWILLE potentiates WUSCHEL function during shoot meristem stem cell development in the Arabidopsis embryo.

United States Department of Agriculture. October 1, 1991. United States Standards for Grades of Fresh Tomatoes, [https://www.ams.usda.gov/sites/default/files/media/Tomato\\_Standard%5B1%5D.pdf](https://www.ams.usda.gov/sites/default/files/media/Tomato_Standard%5B1%5D.pdf).

Valitova, J. N., Sulkarnayeva, A. G., & Minibayeva, F. V. (2016). Plant sterols: Diversity, biosynthesis, and physiological functions. *Biochemistry (Moscow)*, *81*(8), 819–834. <https://doi.org/10.1134/S0006297916080046>

Vallarino, J. G., de Abreu e Lima, F., Soria, C., Tong, H., Pott, D. M., Willmitzer, L., Fernie, A. R., Nikoloski, Z., & Osorio, S. (2018). Genetic diversity of strawberry germplasm using metabolomic biomarkers. *Scientific Reports*, *8*(1), 1–13. <https://doi.org/10.1038/s41598-018-32212-9>

Vallarino, J. G., Merchante, C., Sánchez-Sevilla, J. F., de Luis Balaguer, M. A., Pott, D. M., Ariza, M. T., Casanal, A., Posé, D., Vioque, A., & Amaya, I. (2020). Characterizing the involvement of FaMADS9 in the regulation of strawberry fruit receptacle development. *Plant Biotechnology Journal*, *18*(4), 929–943.

Van Der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N. M., & Stuitje, A. R. (1990). Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell*, *2*(4), 291–299. <https://doi.org/10.1105/tpc.2.4.291>

Vazquez, F. (2006). Arabidopsis endogenous small RNAs: highways and byways. *Trends in Plant Science*, *11*(9), 460–468. <https://doi.org/10.1016/j.tplants.2006.07.006>

Vazquez, F., Legrand, S., & Windels, D. (2010). The biosynthetic pathways and biological scopes of plant small RNAs. *Trends in Plant Science*, *15*(6), 337–345.

Villanueva, R. A. M., & Chen, Z. J. (2019). Ggplot2: elegant graphics for data analysis. Taylor & Francis.

Vincente, A. R., Manganaris, G. A., Ortiz, C. M., Sozzi, G. O., & Crisosto, C. H. (2014). Nutritional Quality of Fruits and Vegetables. *Postharvest Handling: A Systems Approach*, 69–122. <https://doi.org/10.1016/B978-0-12-408137-6.00005-3>

Viuda-Martos, M., Sanchez-Zapata, E., Sayas-Barberá, E., Sendra, E., Pérez-Álvarez, J. A., & Fernández-López, J. (2014). Tomato and tomato byproducts. human health benefits of lycopene and its application to meat products: A review. *Critical Reviews in Food Science and Nutrition*, *54*(8), 1032–1049. <https://doi.org/10.1080/10408398.2011.623799>

Vu, K. D., Hollingsworth, R. G., Leroux, E., Salmieri, S., & Lacroix, M. (2011). Development of edible bioactive coating based on modified chitosan for increasing the shelf life of strawberries. *Food Research International*, *44*(1), 198–203. <https://doi.org/10.1016/j.foodres.2010.10.037>

Waldron, K. W., Parker, M. L., & Smith, A. C. (2003). Plant cell walls and food quality. *Comprehensive Reviews in Food Science and Food Safety*, *2*(4), 128–146.

Wang, F., & Axtell, M. J. (2017). AGO4 is specifically required for heterochromatic siRNA accumulation at Pol V-dependent loci in Arabidopsis thaliana. *Plant Journal*, *90*(1), 37–47. <https://doi.org/10.1111/tpj.13463>

Wang, H., Zhang, H., Yang, Y., Li, M., Zhang, Y., Liu, J., Dong, J., Li, J., Butelli, E., Xue, Z., Wang, A., Wang, G., Martin, C., & Jin, W. (2020). The control of red colour by a family of MYB transcription factors in octoploid strawberry (*Fragaria × ananassa*) fruits. *Plant Biotechnology Journal*, *18*(5), 1169–1184. <https://doi.org/10.1111/pbi.13282>

Wang, J., Sun, L., Xie, L., He, Y., Luo, T., Sheng, L., Luo, Y., Zeng, Y., Xu, J., Deng, X., & Cheng, Y. (2016). Regulation of cuticle formation during fruit development and ripening in “Newhall” navel orange (*Citrus sinensis* Osbeck) revealed by transcriptomic and metabolomic profiling. *Plant Science*, *243*, 131–144. <https://doi.org/10.1016/j.plantsci.2015.12.010>

Wang, J. W., Czech, B., & Weigel, D. (2009). miR156-Regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. *Cell*, *138*(4), 738–749. <https://doi.org/10.1016/j.cell.2009.06.014>

Wang, L., & Wu, J. (2013). The essential role of jasmonic acid in plant-herbivore interactions - using the wild tobacco *Nicotiana attenuata* as a model. *Journal of Genetics and Genomics*, *40*(12), 597–606. <https://doi.org/10.1016/j.jgg.2013.10.001>

WANG, S. xi, SHI, F. yan, DONG, X. xiang, LI, Y. xiang, ZHANG, Z. hong, & LI, H. (2019). Genome-wide identification and expression analysis of auxin response factor (ARF) gene family in strawberry (*Fragaria vesca*).

- Journal of Integrative Agriculture*, 18(7), 1587–1603. [https://doi.org/10.1016/S2095-3119\(19\)62556-6](https://doi.org/10.1016/S2095-3119(19)62556-6)
- Wang, T., Deng, Z., Zhang, X., Wang, H., Wang, Y., Liu, X., Liu, S., Xu, F., Li, T., Fu, D., Zhu, B., Luo, Y., & Zhu, H. (2018). Tomato DCL2b is required for the biosynthesis of 22-nt small RNAs, the resulting secondary siRNAs, and the host defense against ToMV. *Horticulture Research*, 5(1). <https://doi.org/10.1038/s41438-018-0073-7>
- Wang, T., Zhang, H., & Zhu, H. (2019). CRISPR technology is revolutionizing the improvement of tomato and other fruit crops. *Horticulture Research*, 6(1), 1–13.
- Wang, W., & Luan, Y. (2015). The advance of tomato disease-related microRNAs. *Plant Cell Reports*, 34(7), 1089–1097. <https://doi.org/10.1007/s00299-015-1782-0>
- Wang, Y., Wang, Q., Gao, L., Zhu, B., Ju, Z., Luo, Y., & Zuo, J. (2017). Parsing the regulatory network between small RNAs and target genes in ethylene pathway in tomato. *Frontiers in Plant Science*, 8(April), 1–13. <https://doi.org/10.3389/fpls.2017.00527>
- Wang, Zhengming, Hardcastle, T. J., Pastor, A. C., Yip, W. H., Tang, S., & Baulcombe, D. C. (2018). A novel DCL2-dependent miRNA pathway in tomato affects susceptibility to RNA viruses. *Genes and Development*, 32(17–18), 1155–1160. <https://doi.org/10.1101/gad.313601.118>
- Wang, Zhong, Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics in Western Equatoria State. *Nature Reviews Genetics*, 10(1), 57.
- Wassenegger, M., & Krczal, G. (2006). Nomenclature and functions of RNA-directed RNA polymerases. *Trends in Plant Science*, 11(3), 142–151. <https://doi.org/10.1016/j.tplants.2006.01.003>
- Weijers, D., & Wagner, D. (2016). Transcriptional responses to the auxin hormone. *Annual Review of Plant Biology*, 67, 539–574. <https://doi.org/10.1146/annurev-arplant-043015-112122>
- White, P. J. (2002). Recent advances in fruit development and ripening: An overview. *Journal of Experimental Botany*, 53(377), 1995–2000. <https://doi.org/10.1093/jxb/erf105>
- Willats, W. G. T., McCartney, L., Mackie, W., & Knox, J. P. (2001). Pectin: cell biology and prospects for functional analysis. *Plant Molecular Biology*, 47(1), 9–27.
- Willmann, M. R., & Poethig, R. S. (2007). Conservation and evolution of miRNA regulatory programs in plant development. *Current Opinion in Plant Biology*, 10(5), 503–511. <https://doi.org/10.1016/j.pbi.2007.07.004>
- Willson, M. F. (1993). Dispersal mode, seed shadows, and colonization patterns. *Vegetatio*, 107–108(1), 261–280. <https://doi.org/10.1007/BF00052229>
- Wilusz, J. E., Sunwoo, H., & Spector, D. L. (2009). Long noncoding RNAs: Functional surprises from the RNA world. *Genes and Development*, 23(13), 1494–1504. <https://doi.org/10.1101/gad.1800909>
- Wu, F., Chen, Y., Tian, X., Zhu, X., & Jin, W. (2017). Genome-wide identification and characterization of phased small interfering RNA genes in response to *Botrytis cinerea* infection in *Solanum lycopersicum*. *Scientific Reports*, 7(1), 1–10. <https://doi.org/10.1038/s41598-017-02233-x>
- Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., Takita, M. A., Labadie, K., Poulain, J., Couloux, A., Jabbari, K., Cattonaro, F., Del Fabbro, C., Pinosio, S., Zuccolo, A., ... Rokhsar, D. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology*, 32(7), 656–662. <https://doi.org/10.1038/nbt.2906>
- Wu, G., & Poethig, R. S. (2006). Temporal regulation of shoot development in *Arabidopsis thaliana* by miRr156 and its target SPL3. *Development*, 133(18), 3539–3547. <https://doi.org/10.1242/dev.02521>
- Wu, J., Zheng, S., Feng, G., & Yi, H. (2016). Comparative analysis of mirnas and their target transcripts between a spontaneous late-ripening sweet orange mutant and its wild-type using small RNA and degradome sequencing. 7(September), 1–17. <https://doi.org/10.3389/fpls.2016.01416>
- Wu, L., Mao, L., & Qi, Y. (2012). Roles of DICER-LIKE and ARGONAUTE proteins in TAS-derived small interfering RNA-triggered DNA methylation. *Plant Physiology*, 160(2), 990–999. <https://doi.org/10.1104/pp.112.200279>
- Wu, X., Kou, S., Liu, Y., Fang, Y., Xu, Q., & Guo, W. (2015). Genomewide analysis of small RNA s in nonembryogenic and embryogenic tissues of citrus: micro RNA-and si RNA-mediated transcript cleavage involved in somatic embryogenesis. *Plant Biotechnology Journal*, 13(3), 383–394.
- Xia, R., Chen, C., Pokhrel, S., Ma, W., Huang, K., Patel, P., Wang, F., Xu, J., Liu, Z., Li, J., & Meyers, B. C. (2019).

- 24-nt reproductive phasiRNAs are broadly present in angiosperms. *Nature Communications*, *10*(1), 1–8. <https://doi.org/10.1038/s41467-019-08543-0>
- Xia, R., Xu, J., & Meyers, B. C. (2017). The emergence, evolution, and diversification of the miR390-TAS3-ARF pathway in land plants. *Plant Cell*, *29*(6), 1232–1247. <https://doi.org/10.1105/tpc.17.00185>
- Xia, R., Ye, S., Liu, Z., Meyers, B. C., & Liu, Z. (2015). Novel and recently evolved microARN clusters regulate expansive F-BOX gene networks through phased small interfering RNAs in wild diploid strawberry. *Plant Physiology*, *169*(1), 594–610. <https://doi.org/10.1104/pp.15.00253>
- Xia, R., Zhu, H., An, Y., Qiang, Beers, E. P., & Liu, Z. (2012). Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biology*, *13*(6), R47. <https://doi.org/10.1186/gb-2012-13-6-r47>
- Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E., & Carrington, J. C. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS Biology*, *2*(5), 642–652. <https://doi.org/10.1371/journal.pbio.0020104>
- Xu, Q., Chen, L. L., Ruan, X., Chen, D., Zhu, A., Chen, C., Bertrand, D., Jiao, W. B., Hao, B. H., Lyon, M. P., Chen, J., Gao, S., Xing, F., Lan, H., Chang, J. W., Ge, X., Lei, Y., Hu, Q., Miao, Y., ... Ruan, Y. (2013). The draft genome of sweet orange (*Citrus sinensis*). *Nature Genetics*, *45*(1), 59–66. <https://doi.org/10.1038/ng.2472>
- Xu, T., Wang, Y., Liu, X., Lv, S., Feng, C., Qi, M., & Li, T. (2015). Small RNA and degradome sequencing reveals microRNAs and their targets involved in tomato pedicel abscission. *Planta*, *242*(4), 963–984.
- Xu, X., Ma, X., Lei, H., Yin, L., Shi, X., & Song, H. (2015). MicroRNAs play an important role in the regulation of strawberry fruit senescence in low temperature. *Postharvest Biology and Technology*, *108*, 39–47. <https://doi.org/10.1016/j.postharvbio.2015.05.006>
- Xu, X., Yin, L., Ying, Q., Song, H., Xue, D., Lai, T., Xu, M., Shen, B., Wang, H., & Shi, X. (2013). High-throughput sequencing and degradome analysis identify mirnas and their targets involved in fruit senescence of *Fragaria ananassa*. *PLoS ONE*, *8*(8), e70959. <https://doi.org/10.1371/journal.pone.0070959>
- Xue, C., Guan, S. C., Chen, J. Q., Wen, C. J., Cai, J. F., & Chen, X. (2020). Genome wide identification and functional characterization of strawberry pectin methylesterases related to fruit softening. *BMC Plant Biology*, *20*(1), 1–17. <https://doi.org/10.1186/s12870-019-2225-9>
- Yang, F., Cai, J., Yang, Y., & Liu, Z. (2013). Overexpression of microRNA828 reduces anthocyanin accumulation in Arabidopsis. *Plant Cell, Tissue and Organ Culture*, *115*(2), 159–167. <https://doi.org/10.1007/s11240-013-0349-4>
- Yang, L., & Huang, H. (2014). Roles of small RNAs in plant disease resistance. *Journal of Integrative Plant Biology*, *56*(10), 962–970. <https://doi.org/10.1111/jipb.12200>
- Yang, X., & Li, L. (2012). Analyzing the microRNA transcriptome in plants using deep sequencing data. *Biology*, *1*(2), 297–310. <https://doi.org/10.3390/biology1020297>
- Yang, Y. Y., Shan, W., Kuang, J. F., Chen, J. Y., & Lu, W. J. (2020). Four HD-ZIPs are involved in banana fruit ripening by activating the transcription of ethylene biosynthetic and cell wall-modifying genes. *Plant Cell Reports*, *39*(3), 351–362. <https://doi.org/10.1007/s00299-019-02495-x>
- Ye, J., Hu, T., Yang, C., Li, H., Yang, M., Ijaz, R., Ye, Z., & Zhang, Y. (2015). Transcriptome profiling of tomato fruit development reveals transcription factors associated with ascorbic acid, carotenoid and flavonoid biosynthesis. *PLoS ONE*, *10*(7), 1–25. <https://doi.org/10.1371/journal.pone.0130885>
- Yeats, T. H., Martin, L. B. B., Viart, H. M. F., Isaacson, T., He, Y., Zhao, L., Matas, A. J., Buda, G. J., Domozych, D. S., Clausen, M. H., & Rose, J. K. C. (2012). The identification of cutin synthase: Formation of the plant polyester cutin. *Nature Chemical Biology*, *8*(7), 609–611. <https://doi.org/10.1038/nchembio.960>
- Yifhar, T., Pekker, I., Peled, D., Friedlander, G., Pistunov, A., Sabban, M., Wachsmann, G., Alvarez, J. P., Amsellem, Z., & Eshed, Y. (2012). Failure of the Tomato Trans-Acting Short Interfering RNA Program to Regulate AUXIN RESPONSE FACTOR3 and ARF4 Underlies the Wiry Leaf Syndrome. *The Plant Cell*, *24*(9), 3575–3589. <https://doi.org/10.1105/tpc.112.100222>
- Yoshida, Y., Koyama, N., & Tamura, H. (2002). Color and Anthocyanin Composition of Strawberry Fruit: Changes during Fruit Development and Differences among Cultivars, with Special Reference to the Occurrence of Pelargonidin 3-malonylglucoside. *Chemical Pharmaceutical Bulletin*, *71*(3), 355–361. <http://www.mendeley.com/research/geology-volcanic-history-eruptive-style-yakedake-volcano-group-central-japan/>

- Yu, K., Xu, Q., Da, X., Guo, F., Ding, Y., & Deng, X. (2012). Transcriptome changes during fruit development and ripening of sweet orange (*Citrus sinensis*). *BMC Genomics*, *13*(1). <https://doi.org/10.1186/1471-2164-13-10>
- Yu, N., Cai, W. J., Wang, S., Shan, C. M., Wang, L. J., & Chena, X. Y. (2010). Temporal control of trichome distribution by microRNA156-targeted SPL genes in *Arabidopsis thaliana*. *Plant Cell*, *22*(7), 2322–2335. <https://doi.org/10.1105/tpc.109.072579>
- Yu, Q., Liu, Y., Li, M., & Yu, B. (2017). Small RNA Biogenesis and Degradation in Plants. 107–127. [https://doi.org/10.1007/978-3-319-55520-1\\_6](https://doi.org/10.1007/978-3-319-55520-1_6)
- Yu, Y., Jia, T., & Chen, X. (2017). The ‘how’ and ‘where’ of plant microRNAs. *New Phytologist*, *216*(4), 1002–1017. <https://doi.org/10.1111/nph.14834>
- Yu, Z., Kang, B., He, X., Lv, S., Bai, Y., Ding, W., Chen, M., Cho, H. T., & Wu, P. (2011). Root hair-specific expansins modulate root hair elongation in rice. *Plant Journal*, *66*(5), 725–734. <https://doi.org/10.1111/j.1365-313X.2011.04533.x>
- Yun, J. S., Ha, S. C., Kim, S., Kim, Y. G., Kim, H., & Chang, J. H. (2019). Crystal structure of *Arabidopsis thaliana* RabA1a. *Journal of Integrative Plant Biology*, *61*(2), 93–109. <https://doi.org/10.1111/jipb.12700>
- Zhai, J., Jeong, D. H., de Paoli, E., Park, S., Rosen, B. D., Li, Y., González, A. J., Yan, Z., Kitto, S. L., Grusak, M. A., Jackson, S. A., Stacey, G., Cook, D. R., Green, P. J., Sherrier, D. J., & Meyers, B. C. (2011). MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes and Development*, *25*(23), 2540–2553. <https://doi.org/10.1101/gad.177527.111>
- Zhang, B., Yang, H. J., Yang, Y. Z., Zhu, Z. Z., Li, Y. N., Qu, D., & Zhao, Z. Y. (2020). mdm-miR828 participates in the feedback loop to regulate anthocyanin accumulation in apple peel. *Frontiers in Plant Science*, *11*(December). <https://doi.org/10.3389/fpls.2020.608109>
- Zhang, C., Ding, Z., Xu, X., Wang, Q., Qin, G., & Tian, S. (2010). Crucial roles of membrane stability and its related proteins in the tolerance of peach fruit to chilling injury. *Amino Acids*, *39*(1), 181–194. <https://doi.org/10.1007/s00726-009-0397-6>
- Zhang, H., Xia, R., Meyers, B. C., & Walbot, V. (2015). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Current Opinion in Plant Biology*, *27*, 84–90. <https://doi.org/10.1016/j.pbi.2015.06.011>
- Zhang, J., Chiodini, R., Badr, A., & Zhang, G. (2011). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, *38*(3), 95–109. <https://doi.org/10.1016/j.jgg.2011.02.003>
- Zhang, X., Zou, Z., Zhang, J., Zhang, Y., Han, Q., Hu, T., Xu, X., Liu, H., Li, H., & Ye, Z. (2011). Over-expression of sly-miR156a in tomato results in multiple vegetative and reproductive trait alterations and partial phenocopy of the sft mutant. *FEBS Letters*, *585*(2), 435–439. <https://doi.org/10.1016/j.febslet.2010.12.036>
- Zhang, Z., Wang, N., Jiang, S., Xu, H., Wang, Y., Wang, C., Li, M., Liu, J., Qu, C., Liu, W., Wu, S., Chen, X., & Chen, X. (2017). Analysis of the xyloglucan endotransglucosylase/hydrolase gene family during apple fruit ripening and softening. *Journal of Agricultural and Food Chemistry*, *65*(2), 429–434. <https://doi.org/10.1021/acs.jafc.6b04536>
- Zhao, Hongwei, Sun, R., Albrecht, U., Padmanabhan, C., Wang, A., Coffey, M. D., Girke, T., Wang, Z., Close, T. J., Roose, M., Yokomi, R. K., Folimonova, S., Vidalakis, G., Rouse, R., Bowman, K. D., & Jin, H. (2013). Small RNA profiling reveals phosphorus deficiency as a contributing factor in symptom expression for citrus Huanglongbing disease. *Molecular Plant*, *6*(2), 301–310. <https://doi.org/10.1093/mp/sst002>
- Zhao, Hualin, Zhao, K., Wang, J., Chen, X., Chen, Z., Cai, R., & Xiang, Y. (2015). Comprehensive analysis of dicer-like, argonaute, and rna-dependent rna polymerase gene families in grapevine (*Vitis Vinifera*). *Journal of Plant Growth Regulation*, *34*(1), 108–121. <https://doi.org/10.1007/s00344-014-9448-7>
- Zhong, S., Fei, Z., Chen, Y. R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., & Giovannoni, J. J. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology*, *31*(2), 154–159. <https://doi.org/10.1038/nbt.2462>
- Zhu, Hong, Dardick, C. D., Beers, E. P., Callanhan, A. M., Xia, R., & Yuan, R. (2011). Transcriptomics of shading-induced and NAA-induced abscission in apple (*Malus domestica*) reveals a shared pathway involving reduced photosynthesis, alterations in carbohydrate transport and signaling and hormone crosstalk. *BMC Plant Biology*, *11*(1), 1–20.
- Zhu, Hongliang, Hu, F., Wang, R., Zhou, X., Sze, S. H., Liou, L. W., Barefoot, A., Dickman, M., & Zhang, X. (2011).

Arabidopsis argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell*, 145(2), 242–256. <https://doi.org/10.1016/j.cell.2011.03.024>

Zhu, X., Leng, X., Sun, X., Mu, Q., Wang, B., Li, X., Wang, C., & Fang, J. (2015). Discovery of conservation and diversification of miR171 genes by phylogenetic analysis based on global genomes. *The Plant Genome*, 8(2). <https://doi.org/10.3835/plantgenome2014.10.0076>

Zong, J., Yao, X., Yin, J., Zhang, D., & Ma, H. (2009). Evolution of the RNA-dependent RNA polymerase (RdRP) genes: Duplications and possible losses before and after the divergence of major eukaryotic groups. *Gene*, 447(1), 29–39. <https://doi.org/10.1016/j.gene.2009.07.004>

Zuo, J. H., Wang, Y. X., Liu, H. P., Ma, Y. Z., Ju, Z., Zhai, B. Q., Fu, D. Q., Zhu, Y., Luo, Y. B., & Zhu, B. Z. (2011). MicroRNAs in tomato plants. *Science China Life Sciences*, 54(7), 599–605. <https://doi.org/10.1007/s11427-011-4188-4>

Zuo, J., Zhu, B., Fu, D., Zhu, Y., Ma, Y., Chi, L., Ju, Z., Wang, Y., Zhai, B., & Luo, Y. (2012). Sculpting the maturation, softening and ethylene pathway: The influences of microRNAs on tomato fruits. *BMC Genomics*, 13(1), 7. <https://doi.org/10.1186/1471-2164-13-7>

# Anexos

## Anexo I - Comandos bioinformáticos

### Descarga de bibliotecas de secuenciación masiva con Fastq-dump:

```
fastq-dump número_de_acceso
```

Los números de acceso se detallan en cada capítulo.

---

### Evaluación de calidad de secuenciación con FastqC:

```
/PATH/fastqc /PATH/input.fastq
```

---

### Corte de adaptadores con fastx clipper:

```
fastx_clipper -a ADAPTADOR -l N -c -v -M N -i /PATH/Input.fastq -o  
/PATH/Output.fastq
```

Los parámetros `-a`, `-l` y `-M` se especifican en cada capítulo.

---

### Filtrado por calidad de secuenciación fastq quality filter:

```
fastq_quality_filter -q N -p N -v -l /PATH/Input.fastq -o /PATH/Output.fastq
```

Los parámetros `-q` y `-p` se detallan en cada capítulo.

---

### Colapso y cuantificación de *reads* repetidas con fastx collapser:

```
fastx_collapser -v -i /PATH/Input.fastq -o /PATH/Output.fasta
```

---

### Creación de índices de referencia con Bowtie:

```
bowtie-build /PATH/genoma.fasta /PATH/Output
```

---

### Eliminación de secuencias contaminantes con Bowtie:

```
bowtie -v 2 --un /PATH/output_clean /PATH/índice_secuencias_contaminantes  
/PATH/input.fastq /PATH/output_mapeo_contaminantes
```

El archivo `output_mapeo_contaminantes` generado no se utiliza en análisis posteriores.

---

### Alineamiento al genoma de referencia con Bowtie:

```
bowtie -v 2 -m 50 -k 50 -S --al /PATH/Output.fastq /PATH/Índice_genoma_referencia  
/PATH/Input.fastq /PATH/Output.sam
```

---

### Detección de ARNs pequeños con Shortstack:

```
ShortStack --readfile /PATH/input.fastq --genomefile /PATH/genoma.fasta --outdir  
/PATH/output --ranmax none --mismatches 2 --nohp
```

---

### Validación de cortes de transcriptos mediados por ARNs pequeños con Cleaveland:

```
/PATH/CleaveLand4.pl -d /PATH/Degradome_density_file.fasta_dd -u  
/PATH/small_RNA_queries.fasta -n /PATH/transcriptoma.fasta -a -t -p 0.1 -c 2 -o  
/PATH/T_plots > /PATH/Reporte.txt
```

---

### Creación de índices de referencia con hisat2-build:

```
hisat2-build /PATH/genoma.fasta /PATH/Output
```

---

### Alineamiento al genoma de referencia con Hisat2:

```
hisat2 --no-unal --no-softclip -x /PATH/índice -S /PATH/Output.sam -l  
/PATH/Input_réplica1.fastq -2 /PATH/Input_réplica2.fastq 2>  
/PATH/Reporte_output.txt
```

---

### Extracción de reads totales utilizando featureCounts:

```
featureCounts -p -t gene -g ID -a /PATH/ anotaciones.gff3 -o /PATH/Counts.tab  
/PATH/input1.sam /PATH/input2.sam
```

---

### Expresión diferencial con EdgeR:

```
##setwd("/PATH/")  
library(edgeR)  
data <- as.matrix(read.table("tabla_comparacion_input.tab"))  
g <- c(0,0,1,1)  
d <- DGEList(counts=data,group=g,lib.size=c(tamaño_library1, tamaño_library2,  
tamaño_library3, tamaño_library4))  
d <- calcNormFactors(d)  
d <- estimateCommonDisp(d)  
d <- estimateTagwiseDisp(d)  
de.com <- exactTest(d)  
results <- topTags(de.com,n = length(data[,1]))  
write.table(as.matrix(results$table),file="Output.tab",sep="\t")  
##write.table(d$pseudo.counts, file="Output_Normalizadas.txt", row.names = TRUE,  
col.names = TRUE, sep = "\t" )
```

---

### Expresión diferencial con DESeq2:

```
##setwd("/PATH/")
countdata <- read.table("tabla_comparacion_input.tab ", header=TRUE, row.names=1)
countdata <- as.matrix(countdata)
head(countdata)
(condition <- factor(c(rep("Condition_1", 2), rep("Condition_2", 2))))
library(DESeq2)
(coldata <- data.frame(row.names=colnames(countdata), condition))
dds <- DESeqDataSetFromMatrix(countData=countdata, colData=coldata,
design=~condition)
##dds
dds <- DESeq(dds)
res <- results(dds)
table(res$padj<0.05)
res <- res[order(res$padj), ]
resdata <- merge(as.data.frame(res), as.data.frame(counts(dds, normalized=TRUE)),
by="row.names", sort=FALSE)
names(resdata)[1] <- "Gene"
head(resdata)
write.csv(resdata, file="Output.tab")
table_counts_normalized <- counts(dds, normalized=TRUE)
write.table(table_counts_normalized, "Output_Normalizadas.tab", sep="\t")
```

---

### Análisis de enriquecimiento de términos de Ontología Génica:

```
library("topGO")
options(scipen=999) # Para eliminar la notación científica
##setwd("/PATH/")
geneID2GO <- readMappings(file = "/PATH/Genome_with_GO.tab/")
geneUniverse <- names(geneID2GO)
genesOfInterest <- read.table("/PATH/Input.tab", header=FALSE)
genesOfInterest <- as.character(genesOfInterest$V1)
geneList <- factor(as.integer(geneUniverse %in% genesOfInterest))
names(geneList) <- geneUniverse
myGOdata <- new("topGOdata", description="My project",
```

```

ontology="BP", ##O "MF/CC"
allGenes=geneList, annot = annFUN.gene2GO, gene2GO = geneID2GO)
sg <- sigGenes(myGOdata)
str(sg)
numSigGenes(myGOdata)
resultFisher <- runTest(myGOdata, algorithm="classic", statistic="fisher")
resultFisher.table <- GenTable(myGOdata, resultFisher, topNodes =
length(resultFisher@score), numChar=100)
##View(resultFisher.table)
names(resultFisher.table)[6] <- "p.value"
idx <- grep("<", resultFisher.table$p.value)
min.pvalue <- as.numeric(strsplit(resultFisher.table$p.value[idx[1]], "< ")[[1]][2])
resultFisher.table$p.value[idx] <- min.pvalue
##View(resultFisher.table)
##resultFisher.table <- cbind(resultFisher.table, adj.P.val_BY =
p.adjust(resultFisher.table[, "p.value"], method="BY"))
resultFisher.table <- cbind(resultFisher.table, adj.P.val_BH =
p.adjust(resultFisher.table[, "p.value"], method="BH"))
resultFisher.table[which(resultFisher.table$adj.P.val<0.05),]
write.table(resultFisher.table, "/PATH/Output_padj.tab", sep = '\t')
showSigOfNodes(myGOdata, score(resultFisher), firstSigNodes = 5, useInfo = 'all')
printGraph(myGOdata, resultFisher, firstSigNodes = 5, fn.prefix = "tGO", useInfo = "all",
pdfSW = TRUE)

```

## Anexo II - Listado de miARNs utilizados en el análisis del capítulo 2

Nombre	Secuencia	Bibliografía	Expresión en frutos
sly-miR156a	UUGACAGAAGAUAGAGAGCAC	miRBase	No
sly-miR156b	UUGACAGAAGAUAGAGAGCAC	miRBase	No
sly-miR156c	UUGACAGAAGAUAGAGAGCAC	miRBase	No
sly-miR156d-3p	GCUCACUGCUCUAUCUGUCACC	miRBase	No
sly-miR156d-5p	UGACAGAAGAGAGUGAGCAC	miRBase	No
sly-miR156e-3p	GCUUACUCUCUAUCUGUCACC	miRBase	No
sly-miR156e-5p	UGAUAGAAGAGAGUGAGCAC	miRBase	No
sly-miR159	UUUGGAUUGAAGGGAGCUCUA	miRBase	Si
sly-miR160a	UGCCUGGCCUCCUGUAUGCCA	miRBase	No
sly-miR162	UCGAUAAACCUCUGCAUCCAG	miRBase	Si
sly-miR164a-3p	CAUGUGCCUGUUUUCCCAUC	miRBase	Si
sly-miR164a-5p	UGGAGAAGCAGGGCAGUGCA	miRBase	No
sly-miR164b-3p	CACGUGUUCUCCUUCUCCAAC	miRBase	No
sly-miR164b-5p	UGGAGAAGCAGGGCAGUGCA	miRBase	No
sly-miR166a	UCGGACCAGGCUUCAUUCCCC	miRBase	No
sly-miR166b	UCGGACCAGGCUUCAUUCCCC	miRBase	No
sly-miR166c-3p	UCGGACCAGGCUUCAUUCUC	miRBase	Si
sly-miR166c-5p	GGGAUGUUGUCUGGCUCGACA	miRBase	Si
sly-miR167a	UGAAGCUGCCAGCAUGAUCUA	miRBase	No
sly-miR167b-3p	AGGUCAUCUAGCAGCUUCAU	miRBase	No
sly-miR167b-5p	UAAAGCUGCCAGCAUGAUCUGG	miRBase	Si
sly-miR168a-3p	CCUGCCUUGCAUCAACUGAAU	miRBase	No
sly-miR168a-5p	UCGCUUGGUGCAGGUCGGGAC	miRBase	No
sly-miR168b-3p	CCCGCCUUGCAUCAACUGAAU	miRBase	Si
sly-miR168b-5p	UCGCUUGGUGCAGGUCGGGAC	miRBase	No
sly-miR169a	CAGCCAAGGAUGACUUGCCGG	miRBase	No
sly-miR169b	UAGCCAAGGAUGACUUGCCUG	miRBase	No
sly-miR169c	CAGCCAAGGAUGACUUGCCGA	miRBase	No
sly-miR169d	UAGCCAAGGAUGACUUGCCUA	miRBase	No
sly-miR169e-3p	UGGCAAGCAUCUUUGGCGACU	miRBase	Si
sly-miR169e-5p	UAGCCAAGGAUGACUUGCCUUU	miRBase	No
sly-miR171a	UGAUUGAGCCGUGCCAAUAUC	miRBase	No
sly-miR171b	UUGAGCCGUGCCAAUAUCACG	miRBase	No
sly-miR171c	UAUUGGUGCGGUUCAUGAGA	miRBase	No
sly-miR171d	UUGAGCCGCGCCAAUAUCAC	miRBase	No
sly-miR171e	UUGAGCCGCGUCAAAUAUCUCU	miRBase	Si
sly-miR172a	AGAAUCUUGAUGAUGCUGCAU	miRBase	No
sly-miR172b	AGAAUCUUGAUGAUGCUGCAU	miRBase	No
sly-miR1916	AUUUCACUUAGACACCUCAA	miRBase	No
sly-miR1917	AUUAAUAAAGAGUGCUAAAGU	miRBase	No
sly-miR1918	UGUUGGUGAGAGAUUCGAUUCUC	miRBase	No
sly-miR1919a	ACGAGAGUCAUCUGUGACAGG	miRBase	No

sly-miR1919b	ACGAGAGUCAUCUGUGACAGG	miRBase	No
sly-miR1919c-3p	ACGAGAGUCAUCUGUGACAGG	miRBase	No
sly-miR1919c-5p	UGUCGCAGAUGACUUUCGCCC	miRBase	Si
sly-miR319a	CUUGGACUGAAGGGAGCUCC	miRBase	No
sly-miR319b	UUGGACUGAAGGGAGCUCCU	miRBase	No
sly-miR319c-3p	UUGGACUGAAGGGAGCUCCU	miRBase	Si
sly-miR319c-5p	AGAGCUUCCUUCAGCCCACUC	miRBase	No
sly-miR390a-3p	CGCUAUCCAUCCUGAGUUUUA	miRBase	No
sly-miR390a-5p	AAGCUCAGGAGGGAUAGCACC	miRBase	Si
sly-miR390b-3p	CGCUAUCCAUCCUGAGUUUCA	miRBase	No
sly-miR390b-5p	AAGCUCAGGAGGGAUAGCGCC	miRBase	No
SlymiR393	UCCAAAGGGAUCGCAUUGAUC	miRBase	No
sly-miR394-3p	AGGUGGGCAUACUGUCAACA	miRBase	No
sly-miR394-5p	UUGGCAUUCUGUCCACCUCC	miRBase	No
sly-miR395a	CUGAAGUGUUUGGGGGAACUCC	miRBase	No
sly-miR395b	CUGAAGUGUUUGGGGGAACUCC	miRBase	No
sly-miR396a-3p	GUUCAAUAAAGCUGUGGGAAG	miRBase	Si
sly-miR396a-5p	UUCCACAGCUUUCUUGAACUG	miRBase	Si
sly-miR396b	UUCCACAGCUUUCUUGAACUU	miRBase	Si
sly-miR397	AUUGAGUGCAGCGUUGAUGA	miRBase	No
sly-miR399	UGCCAAAGGAGAGUUGCCCUA	miRBase	No
sly-miR403-3p	CUAGAUUCACGCACAAGCUCG	miRBase	Si
sly-miR403-5p	CGUUUGUGCGUGAAUCUAACA	miRBase	No
sly-miR4376	ACGCAGGAGAGAUGAUGCUGGA	miRBase	No
sly-miR477-3p	AGUUCUUGUAGGGUGAGACAAC	miRBase	Si
sly-miR477-5p	UGUCUCUCCCUCAAGGGCUCC	miRBase	No
sly-miR482a	UUUCCAAUCCACCCAUCCUA	miRBase	No
sly-miR482b	UCUUGCCUACACCGCCAUGCC	miRBase	Si
sly-miR482c	UCUUGCCAAUACCGCCAUCC	miRBase	Si
sly-miR482d-3p	UUUCCAAUCCACCCAUGCCAA	miRBase	No
sly-miR482d-5p	GGAGUGGGUGGGGAUGGAAAAA	miRBase	No
sly-miR482e-3p	UCUUCCUACUCCUCCAUACC	miRBase	Si
sly-miR482e-5p	UGUGGGUGGGUGGAAAGAUU	miRBase	No
sly-miR5300	UCCCCAGUCCAGGCAUCCAAC	miRBase	Si
sly-miR5302a	AAACGAGGUUUGUJACUUUGG	miRBase	No
sly-miR5302b-3p	UUUUCAACUAUAGCAUUAUUUU	miRBase	No
sly-miR5302b-5p	UGAAAUGCUAUAGUUGGAAAGU	miRBase	No
sly-miR5303	UUUUUGAAGAGUUCGAGCAAC	miRBase	No
sly-miR5304	UCA AUGCUACAUACUCAUCCC	miRBase	No
sly-miR6022	UGGAAGGGAGAAUAUCCAGGA	miRBase	Si
sly-miR6023	UCCAUGAAAGAGUUUUUGGAU	miRBase	Si
sly-miR6024	UUUUAGCAAGAGUUGUUUUACC	miRBase	Si
sly-miR6026	UUCUUGGCUAGAGUUGUAUUGC	miRBase	Si
sly-miR6027-3p	UGAAUCCUUCGGCUAUCCAUA	miRBase	No
sly-miR6027-5p	AUGGGUAGCACAAGGAUUAUG	miRBase	Si

sly-miR828	UCUUGCUCAAAUGAGUAUUCCA	miRBase	No
Sly-miR858	UUUCGUUGUCUGUUCGACCUU	miRBase	No
sly-miR9469-3p	AUUCGGUCUUCUUAUGUGGAC	miRBase	No
sly-miR9469-5p	CCACAUAAAGAAGACCGAAUUC	miRBase	No
sly-miR9470-3p	UUUGGCUCAUGGAUUUUAGC	miRBase	Si
sly-miR9470-5p	UGAAAUCCAUGAGCCUAAACU	miRBase	Si
sly-miR9471a-3p	UUGGCUGAGUGAGCAUCACGG	miRBase	Si
sly-miR9471a-5p	CAGGUGCUCACUCAGCUAAUA	miRBase	No
sly-miR9471b-3p	UUGGCUGAGUGAGCAUCACUG	miRBase	Si
sly-miR9471b-5p	GAGGUGCUCACUCAGCUAAUA	miRBase	No
sly-miR9472-3p	UUCACAAUCUCUGCUGAAAAA	miRBase	No
sly-miR9472-5p	UUUCAGUAGACGUUGUGAAUA	miRBase	Si
sly-miR9473-3p	AAACGAGUUCAGAUUUACAGC	miRBase	No
sly-miR9473-5p	UGGCUGUAAAUCUAAACUCGU	miRBase	No
sly-miR9474-3p	UUUUGUUCGCAGAUACUACAGU	miRBase	No
sly-miR9474-5p	UGUAGAAGUCAUGAAUAAAAUG	miRBase	Si
sly-miR9475-3p	CUACAAUGUAGAGAUUGUUUU	miRBase	No
sly-miR9475-5p	AACGAUCUCUACAUUGUAGGC	miRBase	Si
sly-miR9476-3p	AAAAAGAUGCAGGACUAGACC	miRBase	Si
sly-miR9476-5p	UCUAGUCCUGCAUCUUUUUUU	miRBase	No
sly-miR9477-3p	UUGGGAAAGGGAACAACUGAUAGU	miRBase	No
sly-miR9477-5p	UAUCCGUUGUUCUUUUUCCUACC	miRBase	No
sly-miR9478-3p	UUCGAUGACAUAUUUGAGCCU	miRBase	No
sly-miR9478-5p	GCUUAAAUAUGUAGAUCGAACU	miRBase	No
sly-miR9479-3p	GAGAAUGGUAGAGGGUCGGACC	miRBase	Si
sly-miR9479-5p	UCCAGUCCUCUACCCUUCUCCA	miRBase	No
mir827	UUAGAUGAACAUCAACAAACA	J. Zuo et al., 2012	No
miRZ1	AUUUUUGAAGAGUCCGAGCA	J. Zuo et al., 2012	No
miRZ2	UCAUUGCUACAUACUCAUCCC	J. Zuo et al., 2012	No
miRZ3	AGCUGCUGACCUAUGGAUCC	J. Zuo et al., 2012	Si
miRZ4	UCCUGCCUUGCAUCAACUGAAU	J. Zuo et al., 2012	Si
miRZ5	UAGGGUGUCGAGUUGAGGAGA	J. Zuo et al., 2012	Si
miRZ6	UGAAUCCUUCGGCUAUCCAUA	J. Zuo et al., 2012	No
miRZ7	UGUUUCUCGUGAAUCCUUCGGC	J. Zuo et al., 2012	Si
miRZ8	CUAUGAGAUAAAGUUAACGUG	J. Zuo et al., 2012	Si
miRZ9	ACAAGAUAGUGUCACGUAGGACA	J. Zuo et al., 2012	No
miR1863-1	AUGCUCUGAUACCAUGUUAGAUU	S. Sato et al., 2012	No
MiR1863-2	GCUCUGAUACCAUGUUAGAUUA	S. Sato et al., 2012	No
MiR1863-3	AAGCUCUGAUACCAUGUUAGAUU	S. Sato et al., 2012	No
MiR1863-4	UGCUCUGAUACCAUGUUAGAUUA	S. Sato et al., 2012	No
MiR1919_Sly-miR0813	UCUGUGACAGGAUUAUGGAAG	S. Sato et al., 2012	No
MiR3627_Sly-miR0211	UCGCAGGAGAGAUGGCACUUGC	S. Sato et al., 2012	No

MiR399_Sly-miR032_y_033	UGCCAAAGAAGAUUUGCCCCG	S. Sato et al., 2012	No
MiR4414_Sly-miR0312	UUGCUGCUGAAUCAUUGGCUC	S. Sato et al., 2012	No
MiR477_Sly-miR015	ACUCUCCCUCAAAGGCUUCUG	S. Sato et al., 2012	No
miR530	UGCAUUUGCACCUGCACCUUA	S. Sato et al., 2012	No
miR5301	UGUGGGUGGGGUGGAAAGAUU	Tomato Functional Genomics Database	No
Mir156_Sly-miR028	UUGAUAGAAGAUAGAGAGCAC	Tomato Genomic Resource Database	No
Mir156_Sly-miR076	UUGACAGAAGAGAGAGAGUAC	Tomato Genomic Resource Database	No
Mir156_Sly-miR077	UUGACAGAAGAGAGAAAGCAC	Tomato Genomic Resource Database	No
Mir156_Sly-miR084	UGACAGAAGAGCGUGAGCAC	Tomato Genomic Resource Database	No
Mir156_Sly-miR122	CUGGCAGAAGAGAGUGAGCA	Tomato Genomic Resource Database	Si
MiR160_Sly-miR123	UGCCUGGCUCCUGGAUGCCAA	Tomato Genomic Resource Database	No
Mir171_Sly-miR029	AAUUGAGCCGUGCCAAUAUCA	Tomato Genomic Resource Database	No
MiR172_Sly-miR055	GGAAUCUUGAUGAUGCUGCAG	Tomato Genomic Resource Database	No
MiR172_Sly-miR061	GGAAUCUUGAUGAUGCUGCAU	Tomato Genomic Resource Database	No
Mir172_Sly-miR101	UGAAUCUUGAUGAUGCUGCAU	Tomato Genomic Resource Database	No
MiR5303_Sly-miR011	UUUUUGAAGAGUAUGAGCAA	Tomato Genomic Resource Database	No
MiR5303_Sly-miR014_y_042	UUUUUGAAGGAUCCGAGCAAC	Tomato Genomic Resource Database	No
MiR5303_Sly-miR091	UCUUUGAAGAGUCCGAGCAGC	Tomato Genomic Resource Database	No
MiR5303_Sly-miR126	UUUUUGAAGAGUUUGGGCAAC	Tomato Genomic Resource Database	No
MiR5303_Sly-miR127	UUUUUGAAGAGUUUGAGCAA	Tomato Genomic Resource Database	No
MiR902_Sly-miR012	UAUGAUUUAGAUUCCGCAU	Tomato Genomic Resource Database	No
Sly_miR165/166	UCCGACCAGGCUUCAUCCCC	S. Moxon et al., 2008	No
miR157b	GAAAAGAAGAUAGAGAGCAC	Y. Luan et al., 2014	No
miR170a	UGAUUGAGCCGUGUCAUAUC	Y. Luan et al., 2014	No
miR2111	UAAUCUGCAUCCUGAGGUUUA	Y. Luan et al., 2014	No
miR2118b	UUACCAAUCCACCAUCCUA	Y. Luan et al., 2014	No
miR398	UGUGUUCUCAGGUCACCCCUU	Y. Luan et al., 2014	No
miR408	UGCACUGCCUCUCCCCUGGCU	R. Karlova et al., 2012	No

miR894	CGUUUCACGUCGGGUUCACCA	R. Karlova et al., 2012	No
--------	-----------------------	----------------------------	----