UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

DOCTORADO EN CIENCIAS BIOMÉDICAS
CENTRO DE CIENCIAS GENÓMICAS

IDENTIFICACIÓN Y CARACTERIZACIÓN DE RNAS PEQUEÑOS EN FRIJOL Y EN SUS RAÍCES TRANSGÉNICAS

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
PABLO PELÁEZ HERNÁNDEZ

TUTOR:
DR. FEDERICO SÁNCHEZ RODRÍGUEZ
CENTRO DE CIENCIAS GENÓMICAS

COMITÉ TUTOR:
DR. JOSÉ LUIS REYES TABOADA
DR. CHRISTIAN SOHLENKAMP
CENTRO DE CIENCIAS GENÓMICAS

CUERNAVACA, MORELOS, NOVIEMBRE, 2014

A mi familia

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México.

Al Dr. Federico Sánchez Rodríguez por orientar y aconsejar este trabajo.

A la M.B. Georgina Estrada por su apoyo y colaboración.

A los integrantes del comité tutor: Dr. José Luis Reyes Taboada, Dr. Christian Sohlenkamp y Dr. Miguel Ángel Ramírez.

Al comité revisor de este escrito.

A los coautores del trabajo.

Al Dr. Oswaldo Valdés López, a la Dra. Alejandra Alicia Covarrubias y a la Dra. Georgina Hernández por sus sugerencias y apoyo.

A los miembros de los laboratorios del Dr. Federico Sánchez y de la M.C. Carmen Quinto.

Este trabajo fue financiado por CONACYT y PAPIIT.

CONTENIDO

Página

RESUMEN

Los RNAs pequeños (sRNAs) constituyen un grupo importante de reguladores de la expresión génica que controlan diferentes procesos biológicos en la mayoría de eucariontes. En plantas, participan en procesos como son: el desarrollo, el metabolismo, las respuestas a estrés, las respuestas de defensa (ej. la regulación de material genético foráneo) y el mantenimiento de la integridad del genoma, entre otros. A pesar de su importancia, los estudios sobre la población de RNAs pequeños en frijol, planta de interés agrícola, son escasos. Incluso, no existen reportes sobre el papel de estos reguladores en raíces transgénicas de frijol (raíces pilosas), órgano ampliamente utilizado para estudios de genómica funcional generado a partir de la inserción de un DNA transferido (T-DNA) proveniente de la bacteria *Agrobacterium rhizogenes*. En esta tesis doctoral, se identificaron y caracterizaron los RNAs pequeños, principalmente a los microRNAs (miRNAs), de diferentes órganos de frijol utilizando secuenciación masiva. También, se analizaron RNAs pequeños en callos y en raíces transgénicas de frijol inducidas por *A. rhizogenes* para determinar la participación e impacto del silenciamiento por RNA como mecanismo de defensa durante esta interacción particular planta-bacteria. Estos análisis permitieron la identificación de 114 microRNAs conservados pertenecientes a 33 familias. Además, se detectaron precursores de microRNAs tipo tallo-asa y genes blanco de diversos microRNAs conservados por medio de herramientas bioinformáticas. Se analizaron los niveles de expresión de las familias de microRNAs conservadas en los diferentes órganos. Familias de microRNAs menos conservados e isoformas específicas de frijol fueron también caracterizadas. Nuevos microRNAs fueron encontrados y sus posibles precursores fueron predichos. También se identificaron nuevos genes blanco de microRNAs nuevos y conservados. En raíces transgénicas, se detectaron diversos y abundantes RNAs pequeños provenientes de genes del T-DNA de *A. rhizogenes* (ArT-sRNAs). Estos RNAs pequeños inducen el silenciamiento de genes del T-DNA a nivel post-transcripcional. A partir de los transcritos de los genes del T-DNA se generan RNAs pequeños con un patrón de fase característico de RNAs pequeños secundarios. Se observó además que la producción de ArT-sRNAs no reduce la acumulación de microRNAs en raíces transgénicas en comparación con raíces silvestres. También se determinó que algunos ArT-sRNAs pueden moverse de las raíces transgénicas hacia hojas no transformadas en plantas compuestas de frijol. Curiosamente, analizando el degradoma de raíces transgénicas, se identificaron diversos genes de frijol que igualmente podrían ser regulados por ArT-sRNAs. Por otro lado, en comparación con la cantidad de ArT-sRNAs detectados en raíces transgénicas, la escasa detección de ArT-sRNAs en callos inducidos por *A. rhizogenes* sugiere que, como sucede en los tumores inducidos por *A. tumefaciens*, las rutas de silenciamiento no regulan a los genes del T-DNA en este tipo de tejido. En general, los resultados presentados en esta tesis contribuyen a nuestro entendimiento sobre los procesos regulados por RNAs pequeños en frijol y en su herramienta de estudio.

ABSTRACT

Small RNAs (sRNAs) constitute an important group of gene expression regulators that control different biological processes in most eukaryotes. In plants, they play roles in processes like: development, metabolism, stress responses, defense responses (e.g. regulation of foreign genetic material), nodulation and maintenance of genome integrity, among others. Despite its importance, studies concerning small RNA populations in common bean, valuable agricultural crop, are limited. Also, there are no reports about the roles played by these regulators in transgenic roots of common bean (hairy roots), organ generated from the insertion of a transferred DNA (T-DNA) of the *Agrobacterium rhizogenes* bacteria and widely used for functional genomics studies. In this PhD thesis, small RNAs, mainly microRNAs, were identified and characterized in different organs of common bean by high-throughput sequencing. Besides, we analyzed small RNAs in calli and hairy roots to determine the role of the RNA silencing mechanism in defense during this particular plant-bacteria interaction. These analyses allowed the identification of 114 conserved microRNAs belonging to 33 families. In addition, microRNA stem-loop precursors and target genes of diverse microRNAs were detected using bioinformatics tools. Expression levels of microRNA conserved families were analyzed in the different organs. Less conserved microRNA families and species-specific common bean miRNA isoforms were also characterized. Novel microRNAs were found and their potential precursors were predicted. Novel targets of novel and conserved microRNAs were identified. In hairy roots, diverse and abundant small RNAs originated from genes of the T-DNA of *A. rhizogenes* were detected (ArT-sRNAs). These small RNAs induced silencing of T-DNA genes at the post-transcriptional level. Small RNAs with a phased pattern characteristic of secondary small RNAs were generated from the transcripts of T-DNA genes. In addition, it was observed also that the production of ArT-sRNAs does not reduce the accumulation of microRNAs in hairy roots compared with wild type roots. Interestingly, degradome analysis of hairy roots revealed that diverse genes of common bean could be regulated by ArT-sRNAs. On the other hand, in contrast with the amount of ArT-sRNAs detected in hairy roots, the poor amount of ArT-sRNAs detected in calli induced by *A. rhizogenes* suggests that RNA silencing pathways does not regulate the T-DNA genes in this kind of tissues as occurs in tumors induced by *A. tumefaciens*. In general, the results presented in this thesis increased our understanding about the small RNA regulated processes in common bean and in the tool used for its study.

1. Introducción

1.1 El descubrimiento del silenciamiento génico y los RNAs pequeños

El silenciamiento génico llevado a cabo por RNAs pequeños es considerado uno de los descubrimientos más importantes en biología molecular de las últimas décadas. Se ha convertido en una herramienta poderosa usada ampliamente para silenciar la expresión de genes y realizar estudios de genética reversa. El descubrimiento y caracterización del silenciamiento génico por RNAs pequeños sucedió a finales de 1980 y durante la década de los 90s. Este descubrimiento representa la suma de esfuerzos de diversos investigadores trabajando en diferentes organismos modelo como plantas, gusanos, moscas, hongos, algas y células de mamífero. Desde el principio, las plantas y las interacciones planta-microbio han jugado un papel esencial en nuestro entendimiento del fenómeno.

La noción del RNA como una molécula con capacidad de regular genes está presente en modelos de regulación génica desde principios de la década de los sesenta (Jacob & Monod, 1961; Britten & Davidson, 1969). Desde finales de 1970 y principios de 1980, ya existía la evidencia de que la formación de dúplex de RNA-DNA o RNA-RNA podía afectar la función y expresión de secuencias específicas de RNA (Taniguchi & Weissmann, 1978; Echkardt & Luhrmann, 1979; Jayaraman et al., 1981; Tomizawa & Itoh, 1982; Tomizawa et al., 1981; Simons & Kleckner, 1983). Por ejemplo, en 1978 Stephenson y Zamecnik mostraron que oligonucleótidos sintéticos, complementarios a secuencias del *virus del sarcoma de Rous* (RSV), reducían la infección e inhibían la traducción del RNA viral *in vitro*. Sin embargo, los primeros estudios que propiciaron el descubrimiento del silenciamiento génico por RNAs pequeños de doble cadena son aquellos que utilizaron una estrategia de tipo antisentido para inhibir la expresión de genes. En 1984, dos trabajos reportaron que el uso de plásmidos diseñados para producir un RNA antisentido complementario a un determinado gen constituía una estrategia alternativa para inhibir su actividad. Ambos estudios confirmaron que el RNA antisentido diseñado para unirse al transcrito de un gen blanco reducía la acumulación del RNA mensajero y de la proteína (Izant & Weintraub, 1984; Mizuno et al., 1984). Varios estudios posteriores confirmaron la utilidad de esta estrategia en diferentes líneas de células y organismos

modelo (Rosenberg et al., 1985; Harland & Weintraub, 1985; Melton, 1985). Durante este tiempo, la inhibición de la expresión génica era explicada como un evento de reducción de la traducción o como un evento de degradación por la doble cadena de RNA formada. Mientras tanto, debido a una serie de estudios sobre plantas transgénicas resistentes a virus, varios investigadores se dieron cuenta de que plantas transgénicas expresando genes virales se volvían tolerantes a la infección. A este proceso se le denominó resistencia derivada de patógenos (PDR) y, desde 1985, se propuso que se podía inhibir la infección de un virus si se insertaba un fragmento de material genético proveniente del mismo en el genoma de la planta. En 1993, en base a observaciones relacionadas con plantas transformadas con material genético viral, se propuso un modelo para explicar la PDR que incluía la supresión, y posterior eliminación, de RNA viral en el citoplasma (Lindbo et al., 1993).

En 1990, después de que se reportó la resistencia derivada por patógenos, dos grupos de investigación publicaron observaciones cruciales que condujeron a su explicación y, más importante, al descubrimiento de RNAs pequeños de interferencia (siRNAs; Napoli et al., 1990; van der Krol et al., 1990). Mientras ambos grupos trabajaban con flores de petunias, observaron que la generación de plantas transgénicas sobreexpresando el transgene de la chalcona sintasa (*CHS*), al contrario de incrementar el fenotipo deseado, presentaron una supresión coordinada (co-supresión) del gen introducido y el gen endógeno. Este resultado inesperado, condujo a los investigadores a reflexionar sobre la explicación que se había propuesto sobre el silenciamiento de genes utilizando estrategias de tipo antisentido. La reevaluación de la explicación del silenciamiento génico fue posteriormente alentada por otros ejemplos donde se detectó co-supresión y, principalmente, por el estudio de Guo y Kemphues (1995) donde se reportó que la introducción de un RNA en sentido correspondiente a un determinado gen también induce su silenciamiento. Para elucidar los papeles que tienen las cadenas de RNAs de tipo sentido o antisentido en el silenciamiento génico, en 1998 Fire y Mello purificaron RNAs de cadena sencilla sentido y antisentido para silenciar el gen *unc-22* de *C. elegans* y probar su eficiencia en comparación con RNAs de doble cadena también diseñados para propiciar su silenciamiento (Fire et al., 1998). Su trabajo sienta las bases para el entendimiento del mecanismo. Ellos encontraron que el RNA de cadena sencilla es mucho menos efectivo

para silenciar el gen que el RNA de doble cadena (dsRNA), mostrando que en los tipos de silenciamiento previamente observados, la formación de RNA de doble cadena resulta esencial. También, Fire y Mello (1998) reportaron varias observaciones que permitieron comprender el mecanismo de silenciamiento por RNA como son: el movimiento de RNA, el transporte de RNA, la especificidad del silenciamiento, la eficiencia del silenciamiento y sus posibles modos de acción tanto a nivel post-transcripcional como transcripcional. En la misma década de los noventa, Voinnet y Baulcombe confirmaron que existe una señal de silenciamiento que puede moverse y actuar a nivel sistémico (Voinnet & Baulcombe, 1997). Asimismo, caracterizaron al sistema de silenciamiento como un mecanismo de defensa que es atacado por proteínas virales para afectar su actividad (Voinnet et al., 1998; Voinnet et al., 1999). De acuerdo al entendimiento que se tenía sobre la función del RNA de doble cadena en el silenciamiento génico, Hamilton y Baulcombe (1999) buscaron la cadena de RNA antisentido que en cierto momento se esperaba que se separara de su cadena complementaria para silenciar el RNA mensajero de un gen; sin embargo, ellos no lograron detectar la cadena antisentido completa sino que encontraron pequeños fragmentos de RNAs de la misma. Estudios posteriores probaron que los RNAs de doble cadena son procesados en RNAs pequeños de doble cadena que inducen el silenciamiento génico. De esta manera, comenzó la identificación y caracterización de RNAs pequeños. Actualmente, diferentes tipos de RNAs pequeños, rutas de biogénesis y modos de acción han sido caracterizados en diferentes organismos y tejidos celulares.

A pesar de lo mucho que queda por aprender acerca de los procesos moleculares del silenciamiento por RNA, el conocimiento que tenemos sobre el mecanismo y función del silenciamiento por RNA, propiciado por los descubrimientos pioneros antes mencionados, nos indica que este mecanismo es uno de los mecanismos de control génico más conservados y fundamentales en los organismos eucariontes. El descubrimiento de este mecanismo ha cambiado la forma de estudiar la regulación génica y el control del desarrollo en plantas y animales.

1.2 Las rutas de silenciamiento de RNAs pequeños en plantas

La mayoría de las rutas de silenciamiento de RNAs pequeños en plantas han sido caracterizadas en la planta modelo *Arabidopsis thaliana*. En base a los estudios hechos en esta planta modelo, y recientemente en otras plantas, se ha hecho evidente la gran diversificación funcional que han tenido las rutas de silenciamiento de RNAs pequeños durante la evolución en estos organismos (Chapman & Carrington, 2007). Por ejemplo, *A. thaliana* contiene cuatro diferentes ribonucleasas Dicer-like, encargadas de generar RNAs pequeños, mientras que los mamíferos sólo tienen una (Baulcombe, 2004). Además de tener un mayor número de determinados componentes del sistema de silenciamiento en relación con otros organismos, las plantas también poseen algunos componentes únicos (Herr et al., 2005; Onodera et al., 2005).

1.2.1 Componentes del mecanismo de silenciamiento

Las diferentes clases de RNAs pequeños propician el silenciamiento post-transcripcional (PTGS) o transcripcional (TGS) de diferentes elementos genéticos. En el silenciamiento post-transcripcional se lleva a cabo el corte del RNA blanco o se reprime su traducción, mientras que el silenciamiento transcripcional implica la metilación del DNA o la modificación de histonas (Saze et al., 2012). Los diferentes tipos de RNAs pequeños son producidos a través de diferentes rutas que comparten diversas familias de proteínas esenciales como son: las proteínas Dicer-like (DCLs), las RNA polimerasas dependientes de RNA (RDRs), la metiltransferasa de RNA HUA ENHANCER 1 (HEN1) y las proteínas Argonauta (AGOs). Estos elementos, junto con los precursores de RNA de doble cadena, son capaces de producir RNAs pequeños de tamaños específicos para regular la expresión génica.

Las ribonucleasas Dicer-like de la familia de las RNAsas tipo III procesan RNAs largos de doble cadena en dúplex de RNAs pequeños con 2 nucleótidos (nt) sobresaliendo en los extremos 3´ y con monofosfatos en los extremos 5´. Todas las plantas estudiadas tienen por lo menos cuatro proteínas Dicer-like base. Las proteínas Dicer-like en plantas contienen los dominios: DEAD-box, helicase-C, DUF283, PIWI/ARGONAUTE/ZWILLE

(PAZ), RNase-III y el dominio de unión a RNA de doble cadena (dsRBD). Estos dominios comparten similitudes en sus estructuras con sus contrapartes en animales y les proporcionan la capacidad de específicamente procesar distintos sustratos (Margis et al., 2006). La proteína Dicer-like 1 (DCL1) participa preferentemente en la producción de microRNAs de 21-nt. Contiene los dominios DExD/H-box (para RNA helicasa), DUF283, PAZ, dos dominios RNase-III continuos y dos dominios dsRBD. Los motivos de RNA helicasa, dsRBD y el sitio catalítico de RNase-III son esenciales para la producción de microRNAs (Kurihara & Watanabe, 2004; Mlotshwa et al., 2005). Uno de los dsRBD favorece la localización de DCL1 en el núcleo, lugar donde lleva a cabo el procesamiento del RNA. DCL1 se autorregula negativamente por los microRNAs miR162 y miR838 (Rajagopalan et al., 2006; Xie et al., 2003). La mayoría de los estudios de pérdida de función relacionados con esta proteína se realizan con mutantes que contienen alelos hipomórficos debido a que su completa eliminación es letal para el embrión (Vaucheret et al., 2004). Las otras tres enzimas Dicer-like DCL2, DCL3 y DCL4 se encargan de generar RNAs pequeños de 22, 24 y 21 nucleótidos, respectivamente, a partir de RNAs largos de doble cadena, con la mayoría de sus bases apareadas, generados por las RDRs (Henderson et al., 2006). La capacidad y especificidad de generar RNAs pequeños de diferentes tamaños, por parte de estas proteínas, se propone que radica en la distancia que separa al dominio PAZ y a los dominios catalíticos. Los RNAs pequeños de 22-nt, que produce la enzima DCL2, participan en el silenciamiento transcripcional y post-transcripcional (Jauvion et al., 2012). El corte realizado por RNAs pequeños de 22-nt induce la producción de RNAs pequeños secundarios al promover la síntesis de RNA de doble cadena por las RDRs (Fei et al., 2013). Por otra parte, los RNAs pequeños de 24-nt y DCL3 participan en el silenciamiento transcripcional, regulando principalmente transposones y secuencias repetidas (Pontes et al., 2006). La enzima DCL4 participa en el silenciamiento post-transcripcional de elementos genéticos endógenos y exógenos como genes, transposones o RNA viral. Las proteínas DCL2, DCL3 y DCL4 tienen distintas afinidades por RNAs largos de doble cadena que se propone depende de su localización subcelular y de la ruta de biosíntesis. Se ha observado, que la proteína DCL4 tiene afinidad por dsRNA sin importar la base del extremo 5´ fosforilado, mientras que DCL3 preferencialmente procesa dsRNAs de menor tamaño con una adenina o citosina fosforiladas en el extremo 5´

(Nagano et al., 2014). La eliminación de dos o más de estas proteínas conlleva importantes alteraciones en la biología de las plantas (Henderson et al., 2006).

Las RNA polimerasas dependientes de RNA son componentes esenciales del mecanismo de silenciamiento por RNA en plantas. Producen RNA de doble cadena a partir de RNA de cadena sencilla. Este tipo de polimerasas sorprendentemente están presentes en diversos organismos pero no en la mosca *Drosophila melanogaster* ni en los mamíferos estudiados hasta el momento (Zong et al., 2009). Los RNAs pequeños pueden promover la actividad de las RDRs al actuar como cebadores o al establecer el inicio de la síntesis del RNA de doble cadena donde se llevó a cabo el corte por las AGOs. Las RDRs RDR1, RDR2 y RDR6 de *Arabidopsis* conforman el clado RDRα. Estas RDRs comparten el motivo catalítico DLDGD, sin embargo, tienen funciones diferentes que van de la mano de las diferentes DCLs en las diferentes rutas de silenciamiento (Pumplin & Voinnet, 2013). Plantas mutantes en alguna de estas RDRs puede modificar la actividad de las otras. Las mutantes en *RDR6* y *RDR2* presentan alteraciones en el desarrollo reproductivo, principalmente en el desarrollo del gametofito femenino y en la fecundación (Olmedo-Monfil et al., 2010; Willmant et al., 2011). Las restantes tres RDRs de *Arabidopsis* RDR3, RDR4 y RDR5 conforman el clado RDRγ. Este clado contiene el motivo catalítico DFDGD. Su participación en el mecanismo de silenciamiento no ha sido explorada (Zong et al., 2009; Willmant et al., 2011).

La metiltransferasa HEN1 coloca grupos metilo en el 2´-OH del nucleótido 3´ terminal de cada cadena de RNA perteneciente a los dúplex producidos por las DCLs. La metilación protege a los RNAs de la degradación (Li et al., 2005; Yu et al., 2005). Esta proteína consta de cinco dominios, de los cuales cuatro interactúan directamente con los RNAs pequeños. La metilación de las dos cadenas ocurre de manera separada. En mutantes *hen1*, los microRNAs presentan una inusual heterogeneidad en su tamaño (Li et al., 2005; Yu et al., 2005). Dependiendo de la familia de microRNAs, en mutantes *hen1*, los microRNAs se encuentran con menos bases (truncados) o con nucleótidos de uracilo agregados (una cola) en el extremo 3´. Estas modificaciones en el extremo 3´ de los RNAs pequeños es dependiente de AGO1 (Zhai et al, 2013). Plantas sin la función de HEN1 presentan una reducción en el tamaño de sus órganos, alteraciones en la forma de la roseta de las hojas y esterilidad. Interesantemente, existe una competencia entre microRNAs y

RNAs pequeños de 24-nt por la actividad de esta metiltransferasa (Yu et al., 2010). Aunque se ha detectado la proteína en el núcleo y en el citoplasma, no se conoce exactamente el lugar subcelular preciso donde se lleva a cabo la metilación de los RNAs pequeños. (Xie et al., 2004).

Las proteínas Argonautas son los principales causantes del silenciamiento por RNA. La planta modelo *Arabidopsis* presenta en su genoma 10 proteínas AGO diferentes agrupadas en tres clados: AGO1, AGO5 y AGO10; AGO2, AGO3 y AGO7; AGO4, AGO6, AGO8 y AGO9 (Mallory & Vaucheret, 2010). Contienen cuatro dominios básicos: un dominio N-terminal variable y los dominios PAZ, MID y PIWI. Presentan una estructura en forma de dos lóbulos con un canal central para la unión del RNA pequeño (Wang et al., 2009). Un bucle que recubre la bolsa de unión a RNAs pequeños en el dominio MID reconoce el nucleótido 5´ de los RNAs, mientras que el dominio PAZ se une al final del extremo 3´ (Frank et al., 2012). El dominio PIWI tiene una estructura tipo RNasa que presenta la actividad de endonucleasa. Las proteínas AGO1, AGO2, AGO7, AGO10 y AGO4 tienen la capacidad de llevar a cabo el corte de RNA. El tamaño de los RNAs pequeños y el tipo de nucleótido en el extremo 5´ influyen para su incorporación en las diferentes AGOs (Mi et al., 2008; Montgomery et al., 2008; Zhu et al., 2011b). AGO1, AGO2, AGO5, AGO7 y AGO10 incorporan RNAs pequeños de 21 y 22 nucleótidos. AGO4, AGO6 y AGO9 se asocian principalmente con RNAs de 24 nucléotidos. La proteína AGO7 interacciona con el microRNA miR390, mientras que AGO10 lo hace con los microRNAs miR165 y miR166. En cuanto a la preferencia por incorporar RNAs pequeños con determinado nucleótido en el extremo 5´, AGO1 se une a aquellos que tienen uridina. AGO2, AGO4, AGO6 y AGO9 se adhieren principalmente a los que comienzan con adenosina y AGO5 con los que tienen citosina (Mi et al., 2008).

1.2.2 Tipos de RNAs pequeños

Los microRNAs son un grupo importante de reguladores de la expresión génica de aproximadamente 21-nt de largo. En la biogénesis canónica, los miRNAs se generan de estructuras tipo tallo-asa con apareamiento imperfecto que son formadas a partir de transcritos primarios de cadena sencilla (pri-miRNAs). Estos pri-miRNAs son generados

por la polimerasa II (PolII) (Lee et al., 2004). En el núcleo, los pri-miRNAs son procesados en precursores de miRNAs (pre-miRNAs) preferentemente por DCL1, la proteína con dedos de zinc SERRATE (SE), la proteína de unión a dsRNA HYPONASTIC LEAVES1 (HYL1), la proteína de unión a dsRNA 2 (DRB2) y por las proteínas de unión a la caperuza (cap) CBP80/ABH1 y CBP20 (Hammond et al., 2000; Kim, 2005; Lobbes et al., 2006; Kim et al., 2008; Eamens et al., 2012). La acumulación de miRNAs requiere de la proteína DAWDLE (DDL) que estabiliza los pri-miRNAs (Yu et al., 2008). Los pre-miRNAs son posteriormente procesados por el complejo DCL1-HYL1 para producir dúplex de miRNAs. Estos dúplex son después metilados por HEN1 y exportados al citoplasma por una proteína homóloga de la exportina-5 en plantas llamada HASTY (HST; Yu et al., 2005; Park et al., 2005). En el citoplasma, en la mayoría de los casos, una de las cadenas de cada dúplex de RNAs pequeños, es finalmente incorporada al complejo de silenciamiento inducido por RNA (RISC) para dirigir el silenciamiento génico. AGO1 y AGO10 son los principales acarreadores de miRNAs en el RISC (Baumberger & Baulcombe, 2005). Recientemente, la proteína de unión a RNA TOUGH se reportó que se une a pri-miRNAs y pre-miRNAs. Esta proteína también contribuye a favorecer la interacción entre los pri-miRNAs y HYL1 y, por lo tanto, podría estar regulando la abundancia de microRNAs a través de promover la eficiencia en el corte de DCL1 y/o reclutar pri-miRNAs. También, una nueva proteína identificada rica en prolinas denominada SICKLE (SIC) ha sido implicada en la biogénesis de microRNAs (Zhan et al., 2012).

Los Trans-acting siRNAs (Ta-siRNAs) son RNAs pequeños secundarios específicos de plantas que se generan a partir de RNAs de doble cadena que se sintetizaron como resultado del corte de un transcrito *TAS* (trans-acting siRNAs) por parte de un RNA pequeño o microRNA (Allen et al., 2005; Chen et al., 2007). En general, la biogénesis de Ta-siRNAs involucra el corte de un transcrito por algún RNA pequeño asociado a una proteína AGO, la generación de una dsRNA por RDRs y el procesamiento de la dsRNA por DCL4. Debido a que DCL4 comienza a procesar la cadena desde un punto determinado, los Ta-siRNAs se generan en fase (consecutivamente). Últimamente, los Ta-siRNAs han sido incluidos dentro de una categoría más amplia denominada RNAs pequeños interferentes secundarios en fase (phasiRNAs), ya que ciertos RNAs pequeños

generados de igual manera que los Ta-siRNAs actúan en *cis*. Las cuatro familias de transcritos *TAS* más estudiados en *Arabidopsis* requieren para su biogénesis de RDR6, DCL4, la proteína supresor del silenciamiento génico 3 (SGS3) y la proteína de unión a dsRNA 4 (DRB4) (Howell et al., 2007; Adenot et al., 2006). Estas cuatro familias necesitan microRNAs para su regulación (Fei et al., 2013). AGO7 es la proteína que participa en el reclutamiento de miR390 para realizar el corte de la familia *TAS3* (Montgomery et al., 2008). La biogénesis de Ta-siRNAs provenientes de *TAS3* requiere de dos sitios de unión para miR390 (Axtell et al., 2006). Al contrario de los Ta-siRNAs de *TAS3*, la biogénesis de Ta-siRNAs generados de las familias *TAS1*, *TAS2*, y *TAS4* requieren solamente de un sitio de unión para microRNAs. Recientemente, se reportó que un Ta-siRNA generado a partir de *TAS1c* producía un segundo corte en los transcritos *TAS2* y *TAS1a, TAS1b* y *TAS1c* (Rajeswaran et al., 2012). Los Ta-siRNAs promueven la metilación de DNA por medio de la RNA polimerasa V (PolV), RDR6, SGS3, AGO4 y AGO6.

Los RNAs pequeños de interferencia generados por transcritos antisentido naturales (nat-siRNAs) se generan a partir de dsRNAs que se forman por la complementariedad de bases que existe en este tipo de transcritos. Los nat-siRNAs primarios originados de los transcritos con complementariedad de bases pueden inducir la producción de nat-siRNAs secundarios que amplifican la señal de silenciamiento (Borsani et al., 2005). La mayoría de la información respecto a los elementos necesarios de la ruta de biogénesis de nat-siRNAs proviene de estudios sobre nat-siRNAs específicos identificados en determinados momentos del desarrollo o en condiciones de estrés (Borsani et al., 2005; Katiyar-Agarwal et al., 2006; Ron et al., 2010). La biogénesis de los nat-siRNAs inducidos por sal o bacteria, así como también el nat-siRNA con funciones reproductivas en *Arabidopsis*, requiere de DCL1y/o DCL2, HYL1 y HEN1. También, estos RNAs pequeños dependen para la amplificación de la señal de silenciamiento de SGS3, de la RNA polimerasa IV (PolIV), de RDR6 (en el caso del nat-siRNA inducido por estrés) y de RDR2 (en el caso del nat-siRNA relacionado con el funcionamiento del esperma durante la doble fecundación). En años recientes, se han realizado estudios importantes para caracterizar la población de nat-siRNAs y proveer un panorama general sobre su biogénesis (Jin et al., 2008; Zhang et al., 2012a). Zhang y colaboradores (2012a) realizaron un análisis de nat-

siRNAs en todo el genoma de *Arabidopsis thaliana* y arroz. Lo autores encontraron que los transcritos que generan nat-siRNAs son procesados por DCL1 y/o DCL3 para generar RNAs pequeños de 20 a 22 nt y de 23 a 28 nt respectivamente. También, encontraron que los nat-siRNAs dependientes de DCL3 requieren para su biogénesis de RDR2 y PolV, mientras que sólo unos cuantos nat-siRNAs dependientes de DCL1 dependen de RDRs y de PolIV. Además, observaron que los nat-siRNAs de 20 a 22 nt son principalmente reclutados por AGO2 y los de 23 a 26 nt por AGO1 y AGO4. De acuerdo a sus resultados, los autores proponen que la dependencia de DCL2 de los nat-siRNAs de 24-nt que se inducen con sal tiene que ser reexaminada, al igual que la posible participación de los nat-siRNAs en la metilación de DNA y la modificación de histonas.

Los RNAs pequeños de interferencia heterocromáticos (Hc-siRNAs) son RNAs pequeños de 24-nt involucrados en la metilacion *de novo* de citosinas, la desmetilación de DNA y la modificación de histonas (Kanno et al., 2011). Es la clase de RNAs pequeños más abundante y diversa en plantas (Henderson et al., 2006). Los Hc-siRNAs son producidos de muchas regiones diferentes del genoma incluyendo secuencias repetitivas, elementos transponibles y secuencias únicas de genes (Kasschau et al., 2007). En ocasiones, estos RNAs son llamados p4-siRNAs porque la mayoría son producidos por medio de PolIV (Zhang et al., 2007). La metilación de DNA es el proceso mejor caracterizado que llevan a cabo los Hc-siRNAs. Este proceso se realiza a través de la ruta de metilación de DNA dirigida por RNA (RdRM) que está formada por una gran cantidad de factores (Kanno et al., 2011; Matzke & Mosher, 2014). La biogénesis de Hc-siRNAs también requiere de la polimerasa II que recluta a PolIV y PolV en loci específicos (Zheng et al., 2009). Para su producción, los Hc-siRNAs también requieren de los siguientes componentes: RDR2, DRB2, DRB4, DCL3, HEN1 y la proteína CLASSY (CLSY1) (Pélissier et al., 2011; Lu et al., 2006; Xie et al., 2004; Smith et al., 2007). Los Hc-siRNAs son incorporados principalmente en AGO4, pero también pueden ser cargados en AGO6 y AGO9 (Havecker et al., 2010). A pesar de que los Hc-siRNAs son los principales RNAs pequeños en promover la metilación por DNA, existen datos que sugieren que otros RNAs pequeños, como nat-siRNAs, podrían llevarla a cabo también (Zhang et al., 2012a). Incluso, otros componentes de las rutas relacionadas al silenciamiento génico post-transcripcional se ha demostrado que participan en la metilación de DNA (Teixeira et al.,

2009). Este es el caso para la ruta nueva de metilación de DNA dirigida por RNAs pequeños interferentes NERD (Pontier et al., 2012). En esta ruta, los componentes involucrados en el silenciamiento post-transcripcional como RDR1, RDR6 y AGO2 participan en la metilación de DNA. Inesperadamente, en esta ruta, los RNAs pequeños de 21-nt promueven la metilación de DNA (Pointer et al., 2012).

En plantas, los transposones pueden también dar lugar a la generación de otro tipo de RNAs pequeños denominados RNAs pequeños interferentes epigeneticamente activados (easiRNAs). Los easiRNAs son RNAs pequeños de 21 nt que se producen principalmente en plantas mutantes en *ddm1* (*DECREASED DNA METHYLATION*) y en *met1* (*DNA METHYLTRANSFERASE 1*). También se ha observado que se generan en el núcleo vegetativo de los granos del pollen o en cultivos celulares desdiferenciados (Slotkin et al., 2009; Tanurdzic et al., 2008). Los easiRNAs se generan a partir de RNA de doble cadena formado de los transcritos de transposones. En la biogénesis de los easiRNAs pueden participar RDR6, DCL4, DCL2, AGO1 o AGO2 (Sarazin & Voinnet, 2014). Recientemente, se identificaron easiRNAs que son generados de manera similar a algunos RNAs pequeños secundarios como los Ta-siRNAs; es decir, que su biogénesis es propiciada por microRNAs incorporados en AGO7 o AGO1 (Creasey et al., 2014). Por esta razón, DCL1 también forma parte de la biogénesis de easiRNAs. La principal función de los easiRNAs es la de regular, a nivel post-transcripcional, los transposones reactivados en estadios particulares de desarrollo o estrés (Sarazin & Voinnet, 2014).

1.3 El papel regulatorio de los RNAs pequeños en plantas

Se ha descrito que el mecanismo de silenciamiento por RNA genera RNAs pequeños a partir de material genético endógeno o exógeno. Los RNAs pequeños endógenos participan en el desarrollo, la fisiología, las respuestas de defensa y el mantenimiento de la integridad del genoma de las plantas (Jones-Rhodes et al., 2006). Los RNAs pequeños generados a partir de transgenes o material genético viral son componentes esenciales de la inmunidad en las plantas (Katiyar-Agarwal & Jin, 2010). Una gran cantidad de RNAs pequeños endógenos regulan blancos que codifican factores de transcripción (Fahlgren et al., 2007). La identificación de los genes que son regulados

por RNAs pequeños es un paso crucial para conocer su función. A diferencia de lo que ocurre en animales, el grado tan elevado de complementariedad que existe entre los RNAs pequeños y sus blancos en plantas ha permitido predecir blancos de manera confiable (Jones-Rhodes et al., 2006).

1.3.1 RNAs pequeños endógenos y sus funciones

Los microRNAs son considerados reguladores maestros debido a que muchos de sus blancos son proteínas reguladoras (Fahlgren et al., 2007). Se han encontrado miRNAs que se encuentran conservados en todas las plantas o solamente en un determinado grupo de plantas. También, se han reportado miRNAs que se encuentran solamente en una especie. Cerca de la mitad de los microRNAs conservados tienen como blancos factores de transcripción (Fahlgren et al., 2007). Las familias de microRNAs que están presentes en todas las plantas analizadas son: miR394, miR172, miR169, miR167, miR164, miR396, miR393, miR168, miR156, miR159, miR319, miR160, miR166, miR171, miR390 (Cuperus et al., 2011; Montes et al., 2014). Varios de estos miRNAs controlan el desarrollo de las plantas y su morfología a través de la regulación de la proliferación y la diferenciación celular (Sun, 2012). En *Arabidopsis*, miR156 tiene como blancos 11 de los 17 genes *SPL* (SQUAMOSA PROMOTER BINDING PROTEIN-LIKE) y diversos genes *SPL* que son regulados por este miRNA promueven el cambio de la fase vegetativa y la transición floral (Wang et al., 2008; Wu & Poethig 2006). La sobreexpresión de miR156 prolonga la etapa juvenil y retrasa la floración (Schwab et al., 2005). MiR159 tiene como blancos los genes *MYB33*, *MYB65* y *MYB101*, que regulan la floración y la fertilidad masculina. Estos genes se relacionan con *GAMYB*, un gen que activa genes de respuesta a la hormona giberelina. La sobreexpresión de miR159 genera esterilidad masculina y retraso en la floración en días cortos (Reyes & Chua, 2007; Achard et al., 2004, Millar & Gubler, 2005). La familia miR319 tiene como blancos a diversos genes *TCP* (TEOSINTE BRANCHED-CYCLOIDEA/PROLIFERATING CELL FACTORS) involucrados en rutas de señalización incluyendo la proliferación y la diferenciación celular. Una baja actividad de los genes *TCP* como consecuencia de la acumulación de miR319 intensifica la proliferación celular y altera el correcto desarrollo de las hojas (Palatnik et al., 2003). La

ruta de señalización de auxina se encuentra altamente regulada por RNAs pequeños. La familia miR160 tiene como blancos tres factores de respuesta a auxina *ARF10*, *ARF16* y *ARF17*. La expresión de las versiones insensibles a la regulación por miR160 de *ARF10*, *ARF16* y *ARF17* ocasiona defectos de desarrollo tanto en la parte aérea de la planta como en su raíz (Liu et al. 2007, Mallory et al. 2005, Wang et al. 2005). Los blancos PHABULOSA (*PHB*), PHAVOLUTA (*PHV*) y REVOLUTA (*REV*) de la familia miR166 son factores de transcripción que promueven la identidad abaxial de órganos laterales (Emery et al., 2003; Mallory et al., 2004). La acumulación de miR166 en el lado abaxial del primordio de la hoja promueve la concentración de sus blancos en el lado adaxial, donde su actividad define las características de esta zona (Kidner & Martienssen, 2004). El microRNA miR390 participa en la generación de Ta-siRNAs a partir de la regulación del transcrito no codificante *TAS3* (Allen et al., 2005). Algunos Ta-siRNAs generados del TAS3 regulan dos factores de respuesta a auxinas: *ARF3* y *ARF4*. La regulación de estos genes contribuye al establecimiento de la identidad abaxial y la expresión de caracteres adultos en las hojas (Adenot et al., 2006; Fahlgren et al., 2006, Garcia et al., 2006; Hunter et al., 2006). La familia miR171 regula a SCARECROW-LIKE6-II (*SCL6-II*), a *SCL6-III* y a *SCL6-IV*, tres factores de transcripción miembros de la familia GRAS (*GAI*, *RGA*, y *SCR*) (Wang et al., 2010a). La sobreexpresión de miR171, o la triple mutante de sus blancos, altera diversos procesos de desarrollo en la planta como son: la ramificación, la altura, la acumulación de la clorofila, la elongación de la raíz, la estructura de la flor y la forma de la hoja.

Los RNAs pequeños secundarios como los Tasi-RNAs y los phasiRNAs pueden regular genes e iniciar la producción de más RNAs pequeños y formar cascadas de regulación (Fei et al., 2013). Son pocos los blancos confirmados de Tasi-RNAs y los estudios sobre su papel. Al igual que el microRNA miR390 regula a *TAS3*, miR173 regula a tres miembros de la familia TAS1 produciendo Tasi-RNAs que regulan a proteínas PPR (pentatricopeptide repeat). Estas proteínas son responsables de la correcta expresión de genes de cloroplastos y mitocondrias al participar en la edición y traducción de los transcritos (Vazquez et al., 2004; Peragine et al., 2004). Los microRNAs miR161 y miR400 regulan genes *PPR* que producen a su vez phasiRNAs que regulan a otros genes *PPR*. Los Tasi-RNAs generados de *TAS4* por el corte de miR828 regulan factores de

transcripción *MYB* que regulan la ruta de biosíntesis de antocianinas (Hsieh et al., 2009). Un microRNA miembro de la familia miR393 produce RNAs pequeños secundarios al regular al menos dos transcritos del grupo de receptores de auxina *TIR/AFB2* (TAAR). Estos RNAs pequeños secundarios, llamados siTAARs, regulan en *cis* los transcritos del grupo de receptores de auxina y en *trans* a los transcritos *At5g24650* y a *LPAT5* (lisofosfatidil aciltransferasa 5). Se ha observado que los siTAARs juegan un papel importante en el desarrollo de la hoja (Si-Ammour et al., 2011).

Los nat-siRNAs también juegan papeles importantes en el desarrollo y respuestas a estrés en plantas (Borsani et al., 2005; Ron et al., 2010). Los nat-siRNAs que se generan de la complementariedad de bases entre el transcrito *SRO5* (SIMILAR TO RCD ONE 5) y *P5CDH* (DELTA1-PYRROLINE-5-CARBOXYLATE DEHYDROGENASE) participan en la regulación de la tolerancia al estrés salino en *Arabidopsis*. El estrés salino induce la acumulación de especies reactivas de oxigeno (ROS) que promueven la expresión de *SRO5*. La expresión de *SRO5* y la formación de una doble cadena de RNA con *P5CDH*, induce la formación de nat-siRNAs de 24 nt. Estos nat-siRNAs promueven la formación de nat-siRNAs de 21 nt en fase, produciendo un incremento en el silenciamiento de *P5CDH*. El silenciamiento de *P5CDH* reduce la degradación de prolina durante el estrés y promueve la acumulación de ROS para señalizar otras respuestas (Borsani et al., 2005). Existe un nat-siRNA específico de células espermáticas que se forma a partir de la unión del transcrito *KPL* (*KOKOPELLI*) y *ARI14* (*ADRIADNE14*). En ausencia de *KPL*, el transcrito de *ARI14* en las células espermáticas se acumula y la fecundación se ve afectada (Ron et al., 2010).

Los RNAs pequeños de interferencia heterocromáticos juegan un papel muy importante en la regulación epigenética a través de la metilación del DNA y la modificación de residuos específicos en las histonas. Se estima que alrededor del 30% del DNA de todo el genoma de *Arabidopsis* es metilado por la actividad de los Hc-siRNAs (Simon & Meyers, 2011). Estos RNAs pequeños pueden inducir dominancia alélica en tejidos específicos en procesos como la auto-polinización (Tarutani et al., 2010). Otro papel en el que participan estos RNAs pequeños es en paramutación, un fenómeno que consiste en la transferencia epigenética de información de un alelo a otro (Chandler, 2007).

También, los Hc-siRNAs generados en un tejido pueden efectuar metilación de DNA en otro tejido lejano (Dunoyer et al., 2010).

Los easiRNAs se considera que tienen un papel importante en proteger al genoma de los transposones (Creasey et al., 2014). En cuanto a la función de los easiRNAs, vale la pena mencionar el trabajo de Slotkin et al. (2009) sobre el silenciamiento, por parte de estos RNAs pequeños, de transposones en el pollen. Slotkin y colaboradores observaron que los transposones son reactivados solamente en el núcleo vegetativo del pollen (que acompaña a las células espermáticas). Sin embargo, easiRNAs de retrotransposones pertenecientes a la familia *Athila* son generados y acumulados en el pollen y en las células espermáticas, lo que sugiere que los easiRNAs que se acumulan en las células espermáticas son generados de transposones del núcleo vegetativo. Por estas razones, se presume que la identidad de los transposones en los gametos es revelada a través de la reprogramación epigenética que ocurre en el núcleo vegetativo (Slotkin et al., 2009).

1.3.2 El silenciamiento génico en respuestas de defensa

El silenciamiento por RNA es una estrategia importante de inmunidad innata utilizada por las plantas para combatir patógenos incluyendo gusanos, hongos, protistas, bacterias y virus (Katiyar-Agarwal & Jin, 2010). La plasticidad en el modo de acción por parte de las rutas de silenciamiento de RNAs pequeños se refleja durante interacciones planta-patógeno. El papel del silenciamiento por RNA en respuestas de defensa se ha estudiado principalmente durante infecciones virales y bacterianas (Anexo 9.2). Para lidiar con este tipo de defensa, los patógenos han desarrollado mecanismos sofisticados tanto de evasión como de contraataque (Pumplin & Voinnet, 2013). Principalmente, el silenciamiento a nivel post-transcripcional y transcripcional parecen un sistema hecho a la medida para defenderse de los ataques virales por la forma en que éstos patógenos necesitan de las funciones celulares del hospedero para completar su ciclo de vida (Mandadi & Scholthof, 2013). Una vez que los virus introducen su material genético en forma de DNA o RNA, el mecanismo de silenciamiento por RNA se induce para controlar la replicación del virus (Pumplin & Voinnet, 2013). Al igual que ocurre para la regulación de elementos genéticos endógenos, el RNA de doble cadena de origen viral se procesa para

formar RNAs pequeños de origen viral (vsiRNAs) por las proteínas DCLs. Estos vsiRNAs son incorporados en las proteínas AGOs para silenciar el RNA o DNA viral. Las RDRs juegan un papel importante en la defensa al generar un silenciamiento sistémico por medio de la producción de RNAs pequeños secundarios. Muchas plantas resistentes a virus se han desarrollado en base a este conocimiento (Duan et al., 2012b). Para defenderse del mecanismo, los virus tienden a inhibir la capacidad del hospedero para silenciar su material genético utilizando una clase diversa de proteínas denominada supresores virales del silenciamiento por RNA (VSRs). Este mecanismo también suele controlar material genético foráneo como pueden ser transgenes (Dunoyer et al., 2005; Deleris et al., 2006; Morel et al., 2002; Mourrain et al., 2000). Incluso, juega un papel importante durante la infección por *Agrobacterium tumefaciens*, donde se transfiere DNA bacteriano hacia células de la planta (Dunoyer et al., 2006).

1.3.2.1 RNAs pequeños endógenos e inmunidad en plantas

Los RNAs pequeños endógenos montan una respuesta de defensa contra patógenos al modular genes involucrados en rutas que participan en la inmunidad en las plantas (Peláez & Sanchez, 2013). Diversos estudios han mostrado que los RNAs pequeños en plantas están directamente involucrados en las respuestas a ataques bacterianos. Los primeros RNAs pequeños identificados como participantes en la respuesta inmune en las plantas durante infecciones bacterianas fueron los microRNAs (Navarro et al., 2006). Plantas expuestas a bacterias patogénicas muestran cambios en la acumulación de microRNAs, particularmente aquellos involucrados en las rutas de señalización de auxina (Fahlgren et al., 2007; Jagadeeswaran et al., 2009; Zhang et al., 2011b). Los RNAs pequeños que responden a virus han sido analizados durante muy diversas interacciones planta-virus (Moissiard et al., 2007; Peláez & Sanchez, 2013). Sin embargo, en contraste con las infecciones de patógenos como las bacterias, la evidencia directa sobre el papel particular de los microRNAs en las infecciones virales es limitada debido a la perturbación que estos patógenos llevan a cabo sobre su biogénesis y su función. En muchos casos, los VSRs acumulan globalmente microRNAs al interaccionar directamente con componentes

de la ruta de biogénesis que pueden afectar su producción, función o estabilización. (Kasschau et al., 2003; Chen et al., 2004; Takeda et al., 2005; Bortolamiol et al., 2007; Csorba et al., 2007; Vogler et al.,2007; Azevedo et al., 2010; Varallyay et al., 2010; Shivaprasad et al.,2012). Se ha observado que, tanto infecciones virales como infecciones bacterianas, pueden alterar la transcripción de los miRNAs (Bazzini et al., 2009). Proteínas virales o RNAs satélite, no necesariamente VSRs, también pueden afectar la acumulación de los microRNAs (Bazzini et al., 2009; Feng et al., 2012). Muchas de las alteraciones de los miRNAs en infecciones virales se han correlacionado con síntomas de las plantas (Kasschau et al., 2003; Dunoyer et al., 2004; Moissiard et al., 2007; Jay et al., 2011). También se han reportado cambios en la acumulación de Ta-siRNAs, que se han correlacionado con cambios fenotípicos, debido a la acción de VSRs que afectan la acumulación de microRNAs como miR390 o alteran el funcionamiento de RDR6 o DCL4 (Moissiard et al., 2007; Wang et al., 2010b; Yifhar et al., 2012). De manera similar, las infecciones virales pueden modificar la producción de Hc-siRNAs y alterar la ruta de metilación de DNA mediada por RNA, promoviendo la reactivación de transposones y la transcripción de genes silenciados que pueden impactar negativamente las respuestas de defensa de la planta (Raja et al., 2008; Azevedo et al., 2010; Dowen et al., 2012). A pesar de que los virus tienden a causar alteraciones en la acumulación de los RNAs pequeños endógenos, algunas familias de RNAs pequeños están directamente involucradas en la inmunidad antiviral. Es interesante, que varios RNAs pequeños endógenos participan en las respuestas de defensa al regular genes de resistencia (genes *R*) o genes relacionados con la patogénesis (genes *PR*).

La percepción de la flagelina 22 (flg22) en *Arabidopsis* restringe la invasión de la bacteria patógena *Pseudomonas syringae*. Sin embargo, no se ha descrito el mecanismo que pudiera estar involucrado en este tipo de resistencia (Navarro et al., 2006). Analizando perfiles de expresión génica en plántulas tratadas con flg22, Navarro et al. (2006) observó que la acumulación de los transcritos de tres receptores de auxina (*TIR, AFB2* y *AFB3*), todos blancos de miR393, se reprimían en el tratamiento. Este resultado sugería que miR393 tenía un papel en regular las respuestas de defensa. Experimentos posteriores confirmaron que miR393 tiene un papel importante en las respuestas de defensa antibacterianas y que la represión de la señalización de auxina constituye una respuesta de

defensa hacia infecciones bacterianas (Navarro et al., 2006). La acumulación de miR393 es inducida también durante la infección de *A. tumefaciens* por cepas armadas o desarmadas (Pruss et al., 2008). Igualmente, se ha observado que el RNA pequeño complementario a miR393, es decir la cadena opuesta o "microRNA estrella", es reclutado por AGO2 y regula blancos que participan en la inmunidad en plantas (Zhang et al., 2011b).

Recientemente, se ha reportado que un grupo de familias de microRNAs de leguminosas y tomate tienen un papel importante en las respuestas de defensa de la planta al regular muchos genes *R* (Zhai et al., 2001; Shivaprasad et al., 2012). Las familias conservadas con este papel son principalmente tres: miR1507, miR2109 y la superfamilia miR482-miR2118. Muchos de los blancos *R* identificados en *Medicago truncatula*, soya y tomate, producen RNAs pequeños en fase como consecuencia del microRNA que los regula inicialmente. En base a la acumulación constitutiva de los miembros de las familias de estos microRNAs, se considera que la mayoría de sus genes *R* blancos se encuentran silenciados en ausencia de agentes patógenos. En este sentido, Shivaprasad et al. (2012) encontraron que el silenciamiento de algunos de estos genes se disminuye durante infecciones virales y bacterianas. Por este motivo, se ha propuesto que la regulación contínua de genes involucrados en defensa por parte de RNAs pequeños endógenos podría constituir un mecanismo inusual de contraataque cuando los supresores del silenciamiento de RNA de patógenos actúan y afectan el silenciamiento (Shivaprasad et al., 2012).

1.3.2.2 Componentes del mecanismo de silenciamiento en defensa

Diversos factores que participan en la biogénesis de RNAs pequeños contribuyen a la resistencia de las plantas contra el ataque de patógenos. Por ejemplo, componentes de las rutas de silenciamiento que son parte del PTGS forman el mecanismo de defensa antiviral (Peláez & Sanchez, 2013). Estudios funcionales han asignado una función singular o redundante a estos componentes durante la inmunidad antiviral (Blevins et al., 2006; Moissiard et al., 2007). La importancia de algunos de estos elementos en la regulación de las respuestas de defensa antivirales se refleja en la actividad que tienen muchos supresores virales para perturbar su función. Análisis de mutantes en componentes de las rutas de silenciamiento en *Arabidopsis*, ha proporcionado diferentes pistas sobre la

función de éstos en la inmunidad de las plantas. En mutantes *dcl1* y *hen1*, se incrementó el crecimiento de la bacteria *P. syringae* Pst DC3000 *hrcC*$^-$ (Navarro et al., 2008). El papel de DCL1 y HEN1 en la inmunidad de las plantas también ha sido analizado durante infecciones de *Agrobacterium tumefaciens* (Dunoyer et al., 2006). En plantas mutantes *dcl1* y *hen1* no hubo crecimiento de tumores a causa de la infección por *A. tumefaciens*, lo que sugiere que la inducción del tumor por *A. tumefaciens* probablemente requiere de un adecuado funcionamiento de los microRNAs. En cuanto a la respuesta antiviral, las cuatro DCLs de *Arabidopsis* cumplen una función esencial. (Blevins et al., 2006; Deleris et al., 2006; Moissiard and Voinnet, 2006; Qu et al., 2008; Garcia-Ruiz et al., 2010). Las DCLs pueden actuar de forma redundante y jerárquica. DCL4 es considerada la más importante DCL y la primera en actuar ante diversas interacciones virales (Blevins et al., 2006; Bouché et al., 2006; Deleris et al., 2006; Qu et al., 2008; Garcia-Ruiz et al., 2010). En este sentido, los supresores virales pueden afectar directa o indirectamente la actividad de DCL4 para evitar esta primera línea de defensa (Deleris et al., 2006; Qu et al., 2008). Mutantes *dcl4* alteran la respuesta inmune local y sistémica (Garcia-Ruiz et al., 2010). Los vsiRNAs generados por DCL4 son necesarios y suficientes para controlar virus que carecen de supresores (Garcia-Ruiz et al., 2010). La enzima DCL2 cumple importantes funciones durante las respuestas de defensa antivirales también. DCL2 es el principal respaldo de DCL4. En ciertas infecciones virales y tejidos, DCL2 puede conferir inmunidad antiviral en mutantes *dcl4* (Xie et al., 2004; Bouché et al., 2006; Deleris et al., 2006; Qu et al., 2008; Garcia-Ruiz et al., 2010). Plantas mutantes en *dcl2* y *dcl4* acumulan transcritos virales en gran cantidad. Aunque se considera que DCL3 tiene un papel menor en la inmunidad antiviral, se ha observado un incremento en los niveles de RNA viral en plantas mutantes en *dcl3* (Qu et al., 2008). El papel de DCL3 en la inmunidad antiviral se centra en contra de virus de DNA probablemente al inducir metilación del DNA viral. En el caso de DCL1, ésta enzima actúa como un regulador negativo de DCL4 y DCL3 en infecciones virales (Qu et al., 2008; Azevedo et al., 2010). Plantas mutantes *hen1* son más susceptibles a infecciones virales (Boutet et al., 2003; Zhang et al., 2012b).

Las proteínas AGO también tienen un papel importante en las respuestas de defensa ante patógenos. La mutante en *ago1-27* de *Arabidopsis* muestra una reducción en la susceptibilidad al hongo patógeno *Botrytis cinerea* (Weiberg et al., 2013). Mutantes en

ago1-27 presentan una disminución en la cantidad de calosa depositada en la pared celular y son más susceptibles a *P. syringae* y a virus. (Li et al., 2010b; Morel et al., 2002). AGO1 es capaz de reclutar vsiRNAs e inducir el corte de RNAs virales (Azevedo et al., 2010; Garcia et al., 2012) y es una proteína frecuentemente afectada por supresores virales para inhibir su actividad de corte (Bortolamiol et al., 2007; Csorba et al., 2007; Azevedo et al., 2010). La proteína AGO2, al igual que AGO1, tiene papeles importantes en las respuestas de defensa contra virus. Mutantes en *ago2-1* son hipersensibles a ciertos tipos de virus (Harvey et al., 2011). Cuando la actividad de AGO1 se ve afectada por supresores virales, se considera que AGO2 puede suplir sus funciones y constituir una segunda línea de defensa. Otras AGOs con funciones antivirales son: AGO5, AGO7 y AGO4 (Qu et al., 2008; Raja et al., 2008; Takeda et al., 2008; Bhattacharjee et al., 2009; Duan et al., 2012a; Hamera et al., 2012). AGO7 participa en la remoción de RNA viral con estructuras secundarias sencillas (Qu et al., 2008). Por otro lado, AGO4 es blanco de supresores probablemente para interrumpir sus funciones en la metilación de DNA viral, ya que se ha observado que existe silenciamiento transcripcional de DNA viral en el núcleo (Raja et al., 2008; Bhattacharjee et al., 2009; Duan et al., 2012; Hamera et al., 2012).

Para montar una adecuada respuesta de defensa sistémica antiviral, la cantidad de vsiRNAs (primarios) es amplificada a través de la producción de vsiRNAs secundarios por las RDRs (Pumplin & Voinnet, 2013). Por esta razón, las polimerasas RDR1, RDR2 y RDR6 han sido señaladas como componentes cruciales en las respuestas de defensa antivirales en diferentes especies de plantas (Mourrain et al., 2000; Yu et al., 2003; Diaz-Pendon et al., 2007; Donaire et al., 2008; Garcia-Ruiz et al., 2010; Wang et al., 2010b). Aunque en ciertos casos, algunas mutantes de RDRs no muestran ningún cambio en la susceptibilidad a virus que tienen supresores virales (Diaz-Pendon et al., 2007; Donaire et al., 2008). En la infección por *Agrobacterium tumefaciens*, las mutantes en *rdr6* presentan mayor susceptibilidad a esta bacteria (Dunoyer et al., 2006).

1.3.2.3 RNAs pequeños de material genético foráneo

El RNA viral de doble cadena es usado por las plantas para producir vsiRNAs de aproximadamente 20 a 24 nucleótidos de largo utilizando diversos componentes de las

rutas de silenciamiento para silenciar el material genético viral (Deleris et al., 2006). Las plantas, también pueden utilizar el T-DNA de *A. tumefaciens* como molde para la producción de RNAs pequeños provenientes de material genético de tipo bacteriano (Dunoyer et al., 2006). Es interesante la evidencia que existe respecto a la regulación que pueden llevar a cabo los vsiRNAs sobre genes endógenos del hospedero. A pesar de que los RNAs pequeños generados de RNA viral se identificaron hace varios años, todavía no se ha caracterizado su biogénesis completamente (Hamilton and Baulcombe, 1999).

La principal fuente del RNA viral de doble cadena que sirve como templado para generar vsiRNAs ha sido ampliamente discutida (Molnár et al., 2005). El dsRNA viral que es procesado en vsiRNAs en virus con genoma de RNA se postuló que su origen es principalmente a partir de intermediarios de dsRNA que son necesarios para la replicación del genoma. Sin embargo, evidencia reciente sugiere que precursores de RNA de cadena sencilla altamente estructurados y RNAs de doble cadena generados por las RDRs son importantes fuentes de vsiRNAs (Molnár et al., 2005; Donaire et al., 2008; Wang et al., 2010b). En el caso de virus con un genoma de DNA, la mayoría de los vsiRNAs son probablemente generados de unidades transcripcionales, aunque la transcripción bidireccional podría ser otra fuente de producción de vsiRNAs (Moissiard and Voinnet, 2006). En los genomas virales existen zonas más activas y específicas que producen vsiRNAs (Wang et al., 2010b).

De acuerdo con las propiedades bioquímicas particulares que cada DCL posee, los vsiRNAs tienen diferentes tamaños (Deleris et al., 2006). La clase de vsiRNAs de 21-nt es usualmente la más abundante en plantas de *Arabidopsis* infectadas con diferentes virus de RNA de cadena positiva debido al papel primario que tiene DCL4. En ausencia de DCL4, la clase de vsiRNAs de 22-nt, dependiente de DCL2, es la más abundante. La cantidad de vsiRNAs de 22-nt es una fracción pequeña del total cuando está presente una DCL4 activa (Blevins et al., 2006; Bouché et al., 2006; Deleris et al., 2006; Diaz-Pendon et al., 2007; Garcia-Ruiz et al., 2010). DCL3 genera vsiRNAs de 24-nt principalmente en mutantes *dcl4/dcl2* durante infecciones de virus de RNA. En infecciones de virus de DNA los vsiRNAs de 24-nt tienden a acumularse (Blevins et al., 2006; Bouché et al., 2006; Deleris et al., 2006; Qu et al., 2008; Raja et al., 2008; Garcia-Ruiz et al., 2010). Aún no es claro el papel de DCL1 en el procesamiento de RNA viral de doble cadena.

Diversos estudios han propuesto que las RDRs tienen papeles importantes en la producción de vsiRNAs. Las dobles cadenas de RNAs generada por RDRs son significativamente usadas para generar vsiRNAs secundarios que asisten a las respuestas de defensa antivirales y que son requeridas para montar una respuesta sistémica. La relevancia de su participación en la generación de vsiRNAs ha sido apoyada por análisis de secuenciación masiva de bibliotecas de RNAs pequeños provenientes de plantas infectadas. Principalmente RDR1 y RDR6 juegan un papel significativo en la biogénesis de vsiRNAs. Mutantes en *rdr1* y *rdr6* presentan menor cantidad de vsiRNAs (Qu et al., 2008).

Las proteínas AGO también participan en la biogénesis de vsiRNAs. Las proteínas AGO pueden cortar RNA viral para inducir la producción de vsiRNAs secundarios (Wang et al., 2011a). VsiRNAs de diferentes virus se han encontrado en inmunoprecipitados de AGO1, AGO5 y AGO2 (Zhang et al., 2006; Azevedo et al., 2010; Wang et al., 2011a; Takeda et al., 2008). Se ha reportado evidencia directa del corte inducido por una proteína AGO en complejo con un vsiRNA (Pantaleo et al., 2007; Szittya et al., 2010; Zhu et al., 2011a). Sin embargo, se cree que no todos los vsiRNAs producidos durante una interacción planta-virus son incorporados en las AGOs. Hoy en día, se sabe que la preferencia en el reclutamiento de los RNAs pequeños por las AGOs con determinado nucleótido en el extremo 5´ se mantiene para los vsiRNAs (Mi et al., 2008; Wang et al., 2011a).

Como se mencionó previamente, los vsiRNAs también regulan genes del hospedero que pueden tener un impacto en la infección. Análisis *in silico* de predicción de blancos han propuesto muchos blancos que potencialmente podrían ser regulados por vsiRNAs. De manera interesante, el mensajero de *At1g76950*, que tiene funciones de unión a diversos sustratos fue validado como un transcrito regulado por un vsiRNA (Moissiard and Voinnet, 2006). También, por ensayos de tipo 5´ RACE, dos transcritos que codifican un factor específico de poliadenilación y corte (*CPSF30*) y una proteína desconocida similar a una proteína alfa asociada a un canal translocador (*TRAPα*) se han validado (Qi et al., 2009). Sorprendentemente, dos grupos reportaron de manera simultanea que los sintomas amarillentos inducidos en *Nicotiana tabacum* por el RNA del satélite Y (Y-Sat) del virus del mosaico del pepino (CMV) es consecuencia del silenciamiento por un RNA pequeño

generado del Y-sat del mensajero del gen *CHL1* involucrado en la biosíntesis de la clorofila (Shimura et al., 2011; Smith et al., 2011). En el caso de interacciones planta-hongo, se ha detectado que RNAs pequeños generados en el hongo *Botrytis cinerea* pueden ser introducidos en la planta para regular genes de la planta que contribuyen con la inmunidad hacia este tipo de patógenos (Weiberg et al., 2013).

En bacterias, la bacteria patogénica *Agrobacterium tumefaciens* es conocida por introducir un T-DNA que se puede integrar en el genoma de las plantas. El T-DNA codifica genes que disparan la formación de un tumor que produce ciertas compuestos llamados opinas (Britton et al., 2008). Estos compuestos son utilizados por la bacteria como nutrientes. Considerando los efectos del silenciamiento por RNA sobre material genético foráneo, se pensó que este mecanismo podría jugar un papel importante durante la interacción particular entre plantas y esta bacteria (Voinnet et al., 2003). De manera interesante, se encontraron RNAs pequeños provenientes de dos genes del T-DNA en hojas de *N. benthamiana* después de tres días de haberse infectado con *A. tumefaciens* (Dunoyer et al., 2006). No obstante, de maner inesperada, no se detectaron RNAs pequeños de estos genes en el tumor inducido por la bacteria. La mayoría de los RNAs pequeños de genes del T-DNA que se detectaron en hojas eran de 21-nt, lo que sugiere un papel importante de DCL4 en la producción de éstos.

1.3.2.4 Supresores del mecanismo de silenciamiento por RNA

Las interacciones planta-microbio son sofisticadas y dinámicas, involucrando el continuo mejoramiento de complejas estrategias de defensa y contraataque por ambas partes. Varios microbios introducen proteínas efectoras dentro de las células de las plantas para suprimir la inmunidad. Se han detectado tanto supresores virales como bacterianos que afectan a proteínas de las rutas de silenciamiento de RNAs pequeños, a RNAs de doble cadena largos, a RNAs pequeños y a la metilación de DNA para modificar la biogénesis, maduración o funcionamiento tanto de RNAs pequeños del hospedero como de los generados a partir de material genético de patógenos (Peláez & Sanchez, 2013). Los VSRs forman un grupo diverso que está ampliamente distribuido entre los virus. En contraste,

son pocos los supresores bacterianos del silenciamiento por RNA (BSRs) que se han detectado (Navarro et al., 2008).

Recientemente, la caracterización a nivel molecular de la proteína 6b de *A. tumefaciens* de la región del T-DNA sugiere que esta proteína podría funcionar como un supresor del silenciamiento por RNA (Wang et al., 2011b). La proteína 6b interacciona con AGO1 *in vivo* e *in vitro*. Las plantas de *Arabidopsis* que sobreexpresan esta proteína tienen una disminución en la acumulación de microRNAs. Además, estas plantas que sobreexpresan a la proteína 6b comparten características fenotípicas con las mutantes en *ago1-27* y con plantas que sobreexpresan un supresor viral (Wang et al., 2011b). Aunque esta proteína interactúa con otras proteínas de la planta involucradas en proliferación celular, parece evidente que también juega un papel en la supresión de las rutas de silenciamiento. Se considera que la supresión del mecanismo de silenciamiento por RNA observado en tumores es el resultado de cambios en los niveles de las fitohormonas a causa de la transformación. La proteína 6b podría estar contribuyendo en mantener el estado de supresión del silenciamiento de los genes del T-DNA en tumores.

Desde la identificación del primer supresor viral, muchas proteínas que inhiben el silenciamiento por RNA durante las interacciones planta-virus han sido descritas. En muchos casos, estas proteínas tienen otras funciones a parte de la de suprimir las rutas de silenciamiento y, por lo general, no tienen secuencia o estructuras similares (Omarov & Scholthof, 2012). Dos estrategias son principalmente utilizadas por los supresores virales para inhibir el silenciamiento por RNA. Una de estas estrategias involucra la unión directa del supresor a RNAs de doble cadena largos o RNAs pequeños para evitar que los vsiRNAs sean estabilizados e incorporados en alguna proteína AGO. La segunda estrategia más frecuente empleada por los VSRs para interrumpir la correcta función del complejo de silenciamiento inducido por RNA implica la unión directa del supresor con componentes del mecanismo como AGO1. Se han identificado también supresores que emplean estrategias diferentes y particulares para alterar el silenciamiento por RNA como mecanismo de defensa (Pumplin & Voinnet, 2013; Peláez & Sanchez, 2013).

1.4 El frijol, RNAs pequeños y su transformación genética

El frijol común (*Phaseolus vulgaris*) es una leguminosa que pertenece a la subfamilia *Faboideae*, una de las tres subfamilias en las que se divide la familia *Leguminosae o Fabaceae* (Doyle & Luckow, 2003). Esta especie de leguminosa de grano es una de las leguminosas más consumidas a nivel mundial. El frijol tiene una gran importancia económica y nutricional en diversos países de América Latina y África debido a que representa una de las principales fuentes de proteínas, vitaminas y minerales de su dieta diaria (Broughton et al., 2003). Los estudios sobre RNAs pequeños en frijol han sido escasos (Arenas-Huertero et al., 2009; Sunkar & Jagadeeswaran, 2008; Valdés-López et al., 2008; Valdés-López et al., 2010; Contreras-Cubas et al., 2012; Peláez et al., 2012; Naya et al., 2014). La falta de producción de plantas de frijol genéticamente modificadas, de una manera sencilla, reproducible y eficiente, ha sido una limitante para su estudio y mejoramiento. Esta leguminosa no es susceptible de ser transformada establemente como ocurre con algunas otras plantas modelo como *A. thaliana* con el uso de la bacteria *A. tumefaciens*. Sin embargo, mediante el empleo de la bacteria *A. rhizogenes*, se puede llevar a cabo la producción de raíces transgénicas de frijol.

1.4.1 RNAs pequeños de frijol.

Desde que los primeros microRNAs fueron reportados en *Arabidopsis* en 2002, la identificación de microRNAs en diversas especies de plantas ha venido creciendo exponencialmente (Kozomara & Griffiths-Jones, 2013). Se han reportado 7057 loci de microRNAs correspondientes a 73 especies de plantas en miRBase (v.21). La mayoría de estos miRNAs reportados provienen de solo algunas plantas modelo con son: *Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*, *Glycine max* y *Medicago truncatula* (Kozomara & Griffiths-Jones, 2013). Por otro lado, hay solamente 10 microRNAs maduros de *Phaseolus vulgaris* reportados en miRBase (v.21) (Kozomara & Griffiths-Jones, 2013). La primera identificación de microRNAs en *P. vulgaris* se llevó a cabo utilizando una estrategia *in silico* (Sunkar & Jagadeeswaran, 2008). Posteriormente, otros microRNAs fueron identificados por Arenas-Huertero et al. (2009) en diferentes órganos y

condiciones de crecimiento utilizando el método tradicional de secuenciación de Sanger. El primer microRNA maduro de frijol en ser clonado y caracterizado fue pvu-miR399a (Valdés-López et al., 2008). Valdés-López et al. (2010) analizó la expresión de 68 microRNAs bajo diferentes condiciones de estrés en hojas, raíces y nódulos de frijol basándose en una estrategia de hibridación utilizando macroarreglos. El macroarreglo contenía sondas para 9 microRNAs identificados de frijol, 24 microRNAs conservados en otras leguminosas y 35 encontrados en soya. También se han llevado a cabo algunos estudios relacionados con el papel específico de unos cuantos microRNAs de frijol con características particulares sin profundizar en el conocimiento de la población (Contreras-Cubas et al., 2012; Naya et al., 2014).

1.4.2 Las raíces inducidas por *Agrobacterium rhizogenes* y sus aplicaciones

Agrobacterium rhizogenes es una bacteria gram negativa perteneciente al género *Agrobacterium*. Esta bacteria comparte con *Agrobacterium tumefaciens* la capacidad de transferir DNA dentro de las células de las plantas; sin embargo, en vez de inducir la formación de tumores, induce la formación de un tipo de raíces denominadas raíces pilosas (Hairy roots) (Britton et al., 2008). Durante el proceso infeccioso un T-DNA del plásmido inductor de raíces (Ri) se integra en el genoma de las células de la planta. Los genes codificados en el T-DNA inducen la proliferación anormal de las células por medio de cambios en las rutas de transducción de señales que afectan a las fitohormonas (Britton et al., 2008). La capacidad que tienen *A. tumefaciens* y *A. rhizogenes* para realizar transferencias de DNA entre reinos ha sido ampliamente explotada para transformar plantas. Las cepas de *A. tumefaciens* con plásmidos inductores de tumores (Ti) que carecen de los oncogenes del T-DNA, y por lo tanto no inducen el crecimiento de tumores, han sido utilizados para realizar ingeniería genética (Georgiev et al., 2012). Las raíces pilosas que tienen el T-DNA del plásmido Ri de *A. rhizogenes* se han convertido también en una herramienta muy utilizada en ciencia básica y en biotecnología (Georgiev et al., 2012). Son de gran utilidad para realizar estudios de genómica funcional (ej. pérdida de función de genes por silenciamiento) de plantas que muchas veces no se pueden transformar por otro medios. Estas raíces también son ampliamente utilizadas para la producción masiva de

metabolitos secundarios y proteínas heterólogas. Incluso, recientemente se aplican en procesos de fitorremediación (Georgiev et al., 2012). Este tipo de raíces surge a partir de células totipotenciales de callo que subsecuentemente se diferencian en raíces muy ramificadas y agravitropicas (Britton et al., 2008). Plantas completas regeneradas a partir de raíces pilosas presentan un fenotipo particular (llamado fenotipo Hairy root) que consiste en perdida de la dominancia apical, disminución de la fertilidad, hojas arrugadas, órganos anormales, retraso en el crecimiento e internodos cortos (Tepfer, 1984). Diversos estudios se han propuesto determinar la función precisa de los genes del T-DNA de *A. rhizogenes* en la infección y en plantas con el fenotipo Hairy root. En muchos casos, la bioquímica y el modo de acción preciso de los oncogenes no se ha esclarecido (Britton et al., 2008).

2. Objetivos

A pesar de que los RNAs pequeños como los microRNAs juegan un papel importante en diversos procesos biológicos de las plantas, éstos han sido escasamente identificados y caracterizados en una planta de gran interés como es el frijol. Así mismo, no se ha evaluado el impacto del silenciamiento por RNA como mecanismo de defensa en las raíces inducidas por *A. rhizogenes*, que son ampliamente utilizadas para estudios de genómica funcional en frijol y otras especies de plantas. Por estas razones, se plantearon los siguientes objetivos:

2.1 Objetivo General

Identificar y caracterizar poblaciones de RNAs pequeños provenientes de diferentes órganos de plantas de frijol, incluyendo raíces inducidas por *A. rhizogenes*.

2.2 Objetivos Particulares

- Identificar microRNAs conservados, nuevos, isoformas y precursores.
- Identificar genes blanco de microRNAs.

- Monitorear los niveles de expresión de microRNAs en órganos de frijol.
- Identificar y caracterizar los RNAs pequeños de raíces inducidas por *A. rhizogenes*.
- Analizar la regulación de genes del T-DNA y del frijol por RNAs pequeños provenientes de genes del T-DNA de *Agrobacterium rhizogenes* (ArT-sRNAs).

3. Resultados

3.1 MicroRNAs de frijol

Los primeros estudios para identificar microRNAs de plantas utilizaron el método tradicional de secuenciación tipo Sanger. A pesar de la utilidad de este método en la investigación científica, este tipo de secuenciación tiene varias limitaciones con respecto a la identificación de microRNAs (Morin et al., 2008; Moxon et al., 2008). Por ejemplo, no se pueden detectar los microRNAs poco abundantes. La introducción de nuevas tecnologías de secuenciación masiva han incrementado el número de microRNAs identificados, permitiendo realizar análisis de las poblaciones de RNAs pequeños a una escala global. Para llevar a cabo la identificación de microRNAs nuevos y conservados en el frijol, se generaron y secuenciaron de forma masiva cuatro bibliotecas de RNAs pequeños provenientes de hojas, raíces, plántulas y flores (Anexo 9.1). En conjunto, la secuenciación generó más de ochenta millones de secuencias crudas (lecturas; Tabla 1). Con el uso de herramientas bioinformáticas se realizó la identificación y caracterización de la población de microRNAs de frijol. La distribución del tamaño de las secuencias de RNAs pequeños mostró que las especies más diversas y abundantes son aquellas de entre 21 y 24 nucleótidos de largo, tamaños típicos de productos generados por las enzimas DCLs (Figura 1).

3.1.1 Identificación de microRNAs conservados e isoformas

Para identificar microRNAs conservados de frijol en las diferentes bibliotecas, se extrajeron las secuencias únicas de microRNAs maduros de plantas de la base de datos miRBase (Versión 16). Las herramientas BLASTN y SSAHA2 (Sequence Search and

Alignment by Hashing Algorithm) se utilizaron de manera separada para alinear las secuencias de las bibliotecas con las secuencias de microRNAs maduros reportados previamente, buscando exactamente el mismo tamaño y secuencia nucleotídica (Altschul et al., 1990; Ning et al., 2001). Ambas herramientas de alineamiento encontraron el mismo número de microRNAs. No se consideraron aquellos microRNAs que se encontraron en solo una biblioteca o que tenían un total de 15 lecturas absolutas o menos considerando las cuatro librerías.

	Leaves	**Roots**	**Seedlings**	**Flowers**
Total raw reads	16869046	20464127	17188077	27937376
High-quality reads[a]	11797480	16722005	16846110	20238107
Unique sequence tags	3372753	4187414	4015702	3453543
Total Rfam matching sequences	757788	5653553	5307793	9726833
Unique Rfam matching sequences	51227	60121	224240	96726
Perfect miRNA matching sequences[b]	265161	307003	582159	1415225
Total unfiltered isoform sequences	798835	287027	467691	228556
Total miRNA isoform sequences	744348	200630	335274	60167

Tabla 1. Resumen del análisis de los datos de la secuenciación de RNAs pequeños. [a]Secuencias después de remover el adaptador y pasar filtros de calidad y tamaño. [b]Secuencias totales de microRNAs conservados (Peláez et al., 2012).

En total, se identificaron 109 microRNAs conservados pertenecientes a 29 familias de frijol. Catorce de éstos son microRNAs maduros estrella. Diez fueron detectados en sólo dos bibliotecas, veinticinco en tres y setenta y cuatro en las cuatro bibliotecas. Veinticinco familias detectadas son muy conservadas entre las especies de plantas (Figura 2). Nueve familias conservadas identificadas en el frijol, incluyendo miR156, miR159, miR319, miR160, miR166, miR171, miR408, miR390 y miR395, estaban presentes en el ancestro común de todas las embriofitas (Cuperus et al., 2011). También se identificaron en frijol la familia miR396, que se encuentra en todas las traqueofitas, y las familias miR397 y miR398 que están presentes en las plantas con semilla. Otras nueve familias identificadas están presentes en todos los linajes de las plantas con flores: miR162,

miR164, miR167, miR168, miR169, miR172, miR393, miR394 y miR399. La familia miR2111, que ha sido identificada solamente en eudicotiledóneas, y la familia miR2118, que se ha encontrado en monocotiledóneas y en leguminosas, se identificaron también en frijol (Cuperus et al., 2011). Finalmente, dentro de las familias altamente conservadas que se detectaron en frijol, miR157, muy relacionada a miR156, está presente en *Brassicaceae*, *Solanaceae*, *Malvaceae* y *Fabaceae* (Cuperus et al., 2011).

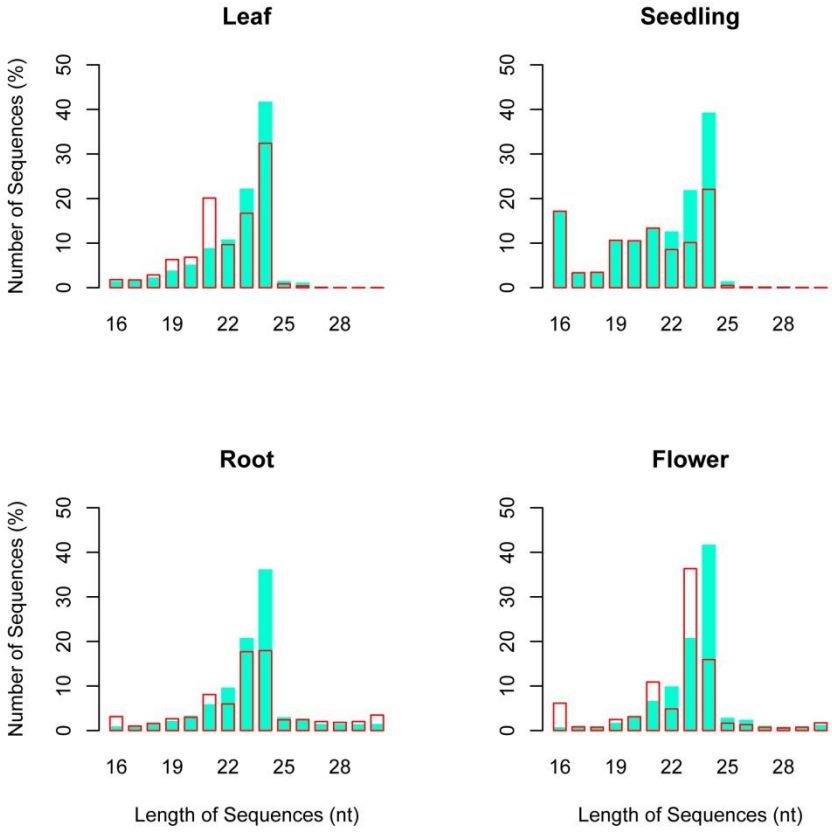


Figura 1. Distribución del largo de las secuencias de sRNAs de *P. vulgaris*. Distribución de los tamaños en las librerías de RNAs pequeños. Porcentaje promedio (eje-Y) de secuencias unicas (barras azules) y secuencias redundantes (barras rojas) de 16 a 30 nt (eje-X) de largo para cada una de las cuatro librerias secuenciadas.

Otras cuatro familias menos conservadas que fueron identificadas se han reportado principalmente en soya (*Glycine max*). Este es el caso de miR1511, que ha sido previamente reportado en soya, y que estudios previos indicaban su presencia también en frijol (Arenas-Huertero et al., 2009; Subramanian et al., 2008; Joshi et al., 2010; Valdés-

López et al., 2010). En este trabajo confirmamos que miR1511 tiene la misma secuencia tanto en soya como en frijol, y que es uno de los más abundantes microRNAs en estas leguminosas. Las familias miR1514 y miR2119 han sido reportadas exclusivamente en leguminosas como *Glycine max* y *Medicago truncatula*. El microRNA miR1515 que se encontró en frijol se ha reportado solamente en dos especies: *G. max* y *C. sinensis*.

Al principio se creyó que los RNAs pequeños como los microRNAs tenían una sola secuencia específica con tamaño delimitado. La identificación subsecuente de más microRNAs reveló que hay variación en el procesamiento de los precursores de los microRNAs. La base de datos de miRBase (Versión 16) contiene microRNAs maduros de plantas de 17 (.08%), 18 (.13%), 19 (.88%), 20 (9.96%), 21 (67.06%), 22 (13.85%), 23 (1.10%), y 24 (6.91%) nucleótidos de tamaño. Ebhardt et al. (2010) demostró que un quinto de los miRNAs anotados de *A. thaliana* tienen una isoforma estable con uno o dos nucleótidos de más. También, diversos estudios han mostrado que la función biológica de las isoformas puede diferir de la función de sus microRNAs previamente reportados (Mi et al., 2008; Vaucheret et al., 2009).

Con el propósito de identificar isoformas y variantes de miRNAs específicas de *P. vulgaris*, todas las secuencias de RNAs pequeños restantes de los análisis previos fueron alineadas contra la miRBase permitiendo no más de dos "mismatches" y/o dos nucleótidos de diferencia en tamaño. Se seleccionaron las variantes abundantes, probablemente funcionales, que tuvieran más de la mitad de lecturas que sus miRNAs de referencia. Las isoformas fueron clasificadas como variantes en tamaño, variantes no conservadas o variantes conservadas. Siete variantes basadas solamente en el tamaño fueron detectadas. El grupo de variantes no conservadas contiene nueve variantes correspondientes a ocho familias de miRNAs. Considerando el número de lecturas, y que homologos de miRNAs maduros presentan entre uno o dos "mismatches", otras tres familias de microRNAs fueron identificadas en frijol: miR1510, miR479 y miR2199. Las diez variantes de miRNAs, clasificadas en el grupo de variantes conservadas, constituyen nuevos candidatos a miRNAs de familias previamente identificadas.

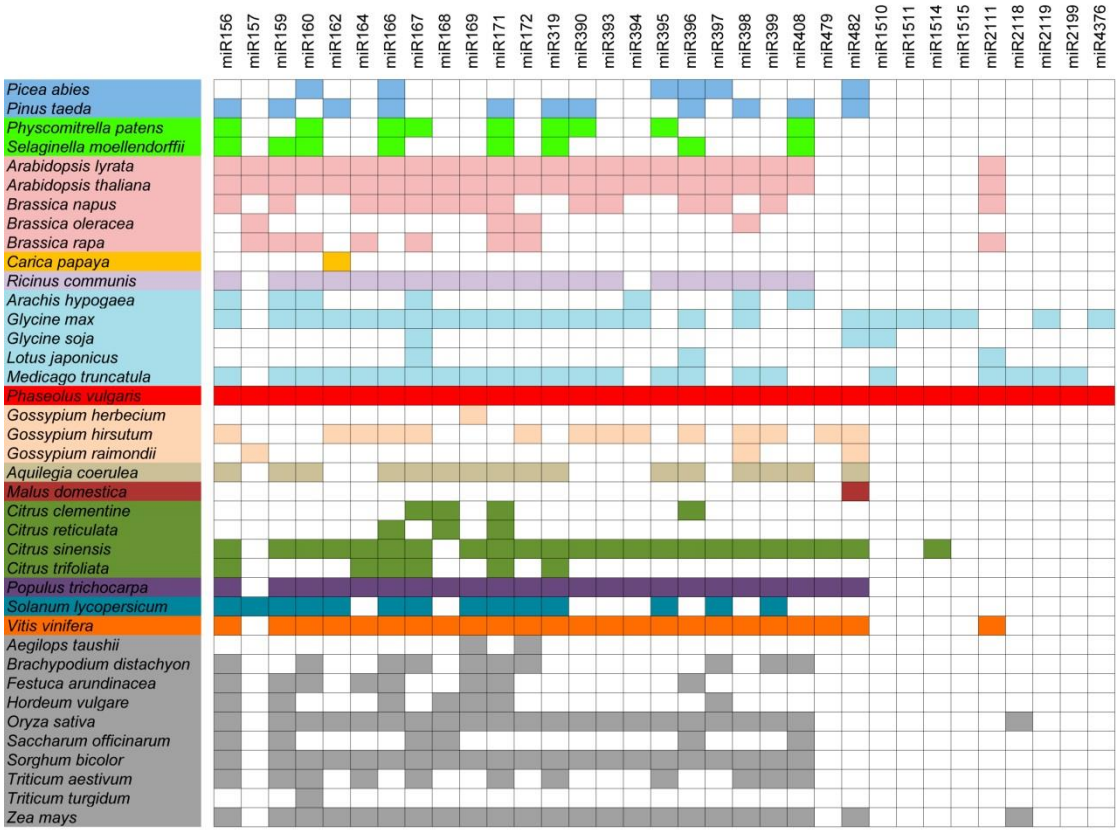


Figura 2. Familias de microRNAs conservadas en frijol y entre especies. Especies relacionadas se muestran con el mismo color. Los cuadros coloreados indican que se ha identificado esa familia de microRNA en esa especie.

3.1.2 Identificación de precursores

Las plantas tienen una población más diversa de RNAs pequeños que los animales principalmente por la particularidad de sus RNA polimerasas (Haag & Pikaard, 2011). Un criterio que apoya la anotación de un RNA pequeño como un microRNA es la identificación de su precursor tipo tallo-asa del cual se genera (Meyers et al., 2008). La búsqueda del potencial precursor de un miRNA involucra el análisis de la estructura secundaria de secuencias genómicas o EST (Expressed Sequence Tag) que coinciden perfectamente con la secuencia del miRNA. A pesar de que los precursores de microRNAs de plantas son más variables en tamaño y estructura que los de animales, varias características se encuentran conservadas entre los precursores de plantas. Para identificar precursores de frijol, los miRNAs identificados se alinearon contra todas las

secuencias de ESTs y GSSs (Genomic Survey Sequences) de NCBI (National Center for Biotechnology Information) en busca de alineamientos perfectos. Los ESTs y GSSs candidatos fueron sometidos a un estudio de estructura secundaria utilizando el programa *mfold* (Zucker, 2003). Solamente las estructuras con la menor energía libre fueron seleccionadas como ha sido descrito previamente (Reinhart et al., 2002; Meyers et al., 2008; Kozomara & Griffiths-Jones, 2013). El análisis de la estructura secundaria de las secuencias candidatas derivó en la identificación de once nuevos precursores pertenecientes a ocho familias conservadas. El precursor identificado de la familia miR171 fue el precursor de una variante. Las otras familias de miRNAs con un precursor identificado fueron: miR166, miR167, miR156, miR157, miR398, miR408 y miR168. Tomando en cuenta los precursores previamente reportados para frijol, un total de 39 microRNAs reportados en este trabajo pueden generarse por el procesamiento de los precursores identificados (Arenas-Huertero et al., 2009). Cinco secuencias de miRNAs que aparecieron en una sola librería o tenían menos de 15 lecturas alinean perfectamente con su respectivo precursor. De igual manera, se identificaron precursores para tres miRNAs maduros estrella.

3.1.3 Identificación de nuevos microRNAs

Numerosos estudios han descrito nuevos microRNAs en diferentes especies de plantas utilizando los resultados obtenidos de la secuenciación de bibliotecas de RNAs pequeños. Sin embargo, por lo general, estos estudios hacen uso de un genoma de referencia que permite la localización y análisis de los precursores. En este estudio, las lecturas de la secuenciación que quedaron, después de la identificación de miRNAs conocidos y sus isoformas, se emplearon para buscar en la colección de ESTs y secuencias genómicas de NCBI y de PlantGDB nuevos microRNAs utilizando el software miRDeep (Friedlander et al., 2008). Este programa busca aquellos RNAs pequeños presentes en las bibliotecas que puedan alinearse con una secuencia de referencia. Las estructuras secundarias de las secuencias seleccionadas son analizadas para identificar aquellas que puedan generar estructuras de tipo tallo-asa típicas de los precursores de microRNAs. Además, los RNAs pequeños tienen que localizarse en la región del tallo del RNA

plegado, y en caso de que se presente, un puntaje más alto es asignado a ese nuevo miRNA si también existen lecturas correspondientes al microRNA maduro estrella.

El análisis para la identificación de nuevos microRNAs se realizó con 95030, 74759, 87956 y 65741 secuencias únicas correspondientes a las bibliotecas de hoja, raíz, plántula y flor respectivamente. El número total de ESTs y secuencias genómicas utilizado fue de 124894. Un total de 29 candidatos para nuevos precursores fueron identificados. Cuatro de los precursores candidatos tienen lecturas para un microRNA maduro y uno estrella. El resto de los precursores tiene lecturas únicamente para el miRNA maduro predicho. El reducido número de ESTs y de secuencias genómicas disponibles para frijol al momento del estudio, limitó el número de candidatos obtenidos por el análisis. La disponibilidad de secuencias genómicas disponibles recientemente permitirán realizar una búsqueda exhaustiva de los miRNAs particulares de esta especie.

3.1.4 Identificación de genes blanco

Los microRNAs de plantas tienen por lo general un mayor grado de complementariedad con sus blancos en comparación con las interacciones blanco-miRNA de animales (Wang et al., 2004). La predicción de blancos de miRNAs en plantas se basa en este alto grado de complementariedad que se comparte entre los blancos y los miRNAs (Jones-Rhoades & Bartel, 2004). Muchos dúplex de mRNA:miRNA de plantas presentan nucleótidos apareados en la región 5´ del microRNA. Sin embargo, algunas interacciones de microRNAs y blancos con nucleótidos no apareados en esta región han sido identificadas en plantas. Por ejemplo, miR398a y su blanco *CSD2* (Copper/Zinc Superoxide Dismutase 2) en *A. thailana*. En este estudio, se usó el programa RNAhybrid para la predicción de blancos de microRNAs de frijol como fue descrito por Alves et al. (2009). El programa RNAhybrid predice el híbrido miRNA:mRNA más favorable energéticamente de acuerdo a las preferencias del usuario (Krüger et al., 2006). Todos los ESTs del género *Phaseolus* (NCBI) fueron evaluados como blancos de los microRNAs previamente identificados en los anteriores análisis (Allen et al., 2005). La identificación del blanco de miR390, el transcrito no codificante *TAS3*, se realizó de manera separada. Una vez que los híbridos se obtuvieron, los ESTs se alinearon contra secuencias de plantas

de la base de datos UniProtKB (UniProt Knowledgebase) utilizando BLASTX para su anotación.

El análisis de predicción de blancos identificó 194 ESTs conocidos como blancos establecidos de miRNAs en plantas. Treinta y siete ESTs habían sido reportados previamente por Arenas-Huertero et al. (2009). Las familias conservadas de miRNAs para las que se identificaron blancos fueron: miR156/miR157, miR160, miR164, miR167, miR168, miR169, miR171, miR172, miR319, miR393, miR395, miR397, miR398 y miR408. El número de ESTs encontrado para cada familia de genes es: *SBP*(7), *ARF*(18), *NAC*(1), *AGO1*(1), *NFY*(14), *SCL*(2), *AP2*(8), *TCP*(6), *TIR/F-box-AFB*(5), ATP sulfurilasa(14), lacasas(1), *COX/SOD*(66) y plastocianinas(11). Las familias miR156 y miR157 comparten las mismas secuencias blanco que corresponden a la familia de genes *SBP* (Squamosa Binding Proteins). Las energías libres mínimas (MFEs) para los dúplex conservados de miRNA:mRNA oscilan de entre -35.0 a -51.2 kcal/mol. Finalmente, dentro de los blancos conservados de miRNAs, se identificaron tres ESTs homologos al *TAS3* de *A. thaliana*, blanco de la familia miR390. Además, se identificaron 325 ESTs como blancos candidatos de los miRNAs conservados, miRNAs maduros estrella e isoformas. El análisis de la predicción de blancos realizado para los nuevos miRNAs candidatos generó blancos candidatos para 10 nuevos miRNAs representados por 177 ESTs.

3.1.5 Análisis de expresión de microRNAs en órganos de frijol

Los patrones específicos de la expresión de microRNAs se ha propuesto que son consecuencia de los elementos regulatorios específicos en los promotores de los genes de miRNAs o de su coordinado movimiento a través de las células de las plantas (Megraw et al., 2006; Parizotto et al., 2004; Válóczi et al., 2006; Benkovicks & Timmermans, 2014). La expresión de algunos genes de miRNAs conservados y los patrones de acumulación de los miRNAs maduros en diferentes órganos o en diferentes etapas del desarrollo, son esenciales para una correcta diferenciación celular y regulación del desarrollo de los órganos. La comprensión de los patrones de expresión de microRNAs en los órganos de las plantas es necesaria para discernir las rutas reguladas por estos RNAs pequeños. La frecuencia de miRNAs detectados por secuenciación masiva permite realizar una

estimación relativa de sus niveles de expresión. Moldovan et al. (2010), comparando las frecuencias de miRNAs de bibliotecas de raíces de *Arabidopsis* contra lecturas de hoja y plantas completas, encontró que la mayoría de las familias de miRNAs tienen una expresión particular en cada órgano.

Para explorar la expresión de los miRNAs de *P. vulgaris* en los diferentes órganos, se utilizó el paquete de R/Bioconductor DESeq (Anderson & Huber, 2010). El paquete de software estadístico DESeq es capaz de probar si existe expresión diferencial con base en la información obtenida de ensayos de secuenciación masiva. Este paquete ha sido previamente utilizado para analizar la expresión diferencial de microRNAs (Hackenberg et al., 2011; Dhabi et al., 2011). Para explorar la expresión en los órganos de las secuencias de miRNAs como familias, las frecuencias de todos los miembros de cada familia se sumaron para cada librería. Los miRNAs maduros estrella no se consideraron. Los datos fueron transformados con el paquete DESeq y se realizó un análisis de agrupamiento jerárquico (HCA).

Las familias de miRNAs se agruparon en dos grupos grandes: el grupo con menos miRNAs consiste en familias de miRNAs que fueron abundantemente expresadas en las cuatro bibliotecas, y el otro grupo, contiene a familias con expresión diferencial diversa entre los órganos (Figura 3). Los patrones de acumulación de las familias en hojas y raíces fueron más similares entre ellos que con las otras dos muestras. Las familias más abundantes del análisis, especialmente en plántulas, fueron miR159, miR319, miR396, miR166, miR408 y miR482. La familia miR159 se ha descrito como una familia muy abundante en todos los órganos de las plantas. Es importante mencionar que la expresión de miR319 fue abundante en todos los órganos excepto en hoja. De manera similar a miR319, en este estudio, miR396 fue menos abundante en hoja que en las otras tres tipos de muestras. Otras familias bastante abundantes fueron: miR1510, miR164, miR167, miR169, miR157 y miR156. Es importante mencionar, que miR1510, un microRNA que se ha detectado solamente en leguminosas, se encontró en las cuatro bibliotecas.

En el grupo de familias con un patrón de expresión más diverso, las familias miR2119, miR2111 y miR479 fueron acumuladas principalmente en un órgano. La familia miR21119 fue detectada específicamente en raíces. Este patrón de expresión correlaciona con estudios previos de tipo *northern blot* (Arenas-Huertero et al., 2009). La expresión de

miR2111 se observó principalmente en raíces y miR399 solo se expresó ligeramente en hojas. Ambos miRNAs se inducen en condiciones con deficiencia de fósforo (P) (Valdés-López et al., 2008; Pant et al., 2009). La familia miR479 se expresó preferentemente en plántulas. Por otra parte, se observaron familias interesantes que fueron poco expresadas en un órgano en particular. Por ejemplo, miR168 fue pobremente expresado en plántulas, miR1515 en hojas, miR2199 en raíces, miR394 en hojas, miR4375 en raíces y miR397 en flores. Finalmente, el análisis de agrupamiento jerárquico mostró que miR395 se acumuló pobremente en las cuatro librerías de RNAs pequeños, lo que concuerda con su inducción bajo condiciones limitantes de sulfato (Kawashima et al., 2009).

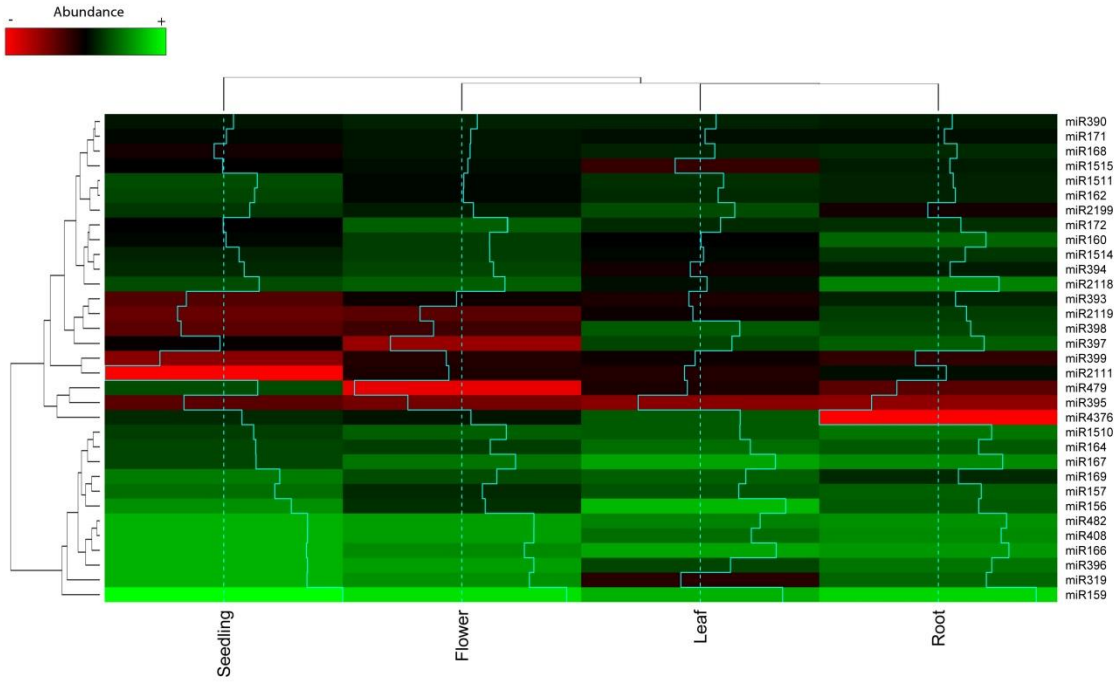


Figura 3. Análisis de agrupamiento jerárquico de las familias de microRNAs de *P. vulgaris*. Las frecuencias de familias conservadas fueron utilizadas para analizar su expresion relativa en hojas, plantulas, raices y flores. El cluster jerárquico se llevo a cabo con 33 familias conservadas (derecha). El color de la escala muestra familias abundantes (verde) y no abundantes (rojo) de acuerdo a la transformación de los datos. Gráficas de densidad se incluyeron en azul.

3.2 RNAs pequeños en raíces transgénicas de frijol

Los transgenes incorporados por la transformación realizada con *A. tumefaciens* frecuentemente son sujetos al silenciamiento por RNA. Voinnet et al. (2003) sugirió que el silenciamiento de transgenes podría ser una respuesta de defensa típica dirigida contra los oncogenes del T-DNA en bacterias virulentas de igual manera a como sucede en las infecciones virales. Poco después, Dunoyer et al. (2006) mostraron el papel defensivo que tiene el silenciamiento por RNA en infecciones virulentas de *A. tumefaciens*. Los autores concluyeron que el estado de supresión de las rutas de silenciamiento por RNA, que observaron en los tumores inducidos por *A. tumefaciens*, podría deberse a un evento intrínseco a la desdiferenciación y/o proliferación de las células. Las raíces inducidas por *A. rhizogenes* (Hairy roots) constituyen un tejido diferenciado que se ha establecido como una herramienta poderosa para realizar análisis de pérdida de función a través del silenciamiento génico en frijol y otras especies. Nosotros pensamos que en este tipo de raíces, a diferencia de lo que ocurre en tumores inducidos por *A. tumefaciens*, las rutas de silenciamiento no pueden encontrarse suprimidas y, como sucede en infecciones virales, las respuestas de defensa llevadas a cabo por el mecanismo de silenciamiento por RNA en contra del T-DNA de esta bacteria deben tener un efecto en este sistema biológico. Con el uso de técnicas de análisis de expresión y de secuenciación masiva analizamos la participación del mecanismo de silenciamiento por RNA en el control de los genes del T-DNA de *A. rhizogenes* y su posible impacto sobre los microRNAs y algunos genes del frijol (Anexo 9.3).

3.2.1 RNAs pequeños del T-DNA de *A. rhizogenes* (ArT-sRNAs)

Se generaron y secuenciaron masivamente bibliotecas de RNAs pequeños de raíces pilosas y callos para explorar si hay actividad del mecanismo de silenciamiento por RNA a lo largo del T-DNA de *A. rhizogenes*. La secuenciación de los RNAs pequeños produjo 12,396,242 y 27,272,479 secuencias totales crudas para las muestras de raíces pilosas y callos respectivamente. La distribución del largo de las secuencias de RNAs pequeños mostró que en la bilbioteca de raíces pilosas las clases de 21-nt (17,9%) y de 24-nt (19%)

son similares en abundancia. La clase de 21-nt fue la segunda más diversa en raíces pilosas y la cuarta en callos (Figura 4). El alineamiento de las secuencias contra el T-DNA de *A. rhizogenes* reveló la presencia de varias secuencias distintas y abundantes de RNAs pequeños derivados del T-DNA de *A. rhizogenes* (ArT-sRNAs) en raíces pilosas (Figura 5). Se identificaron 3,176 ArT-sRNAs distintos correspondientes a 17,000 transcritos por millón (TPM). En contraste, muy pocos ArT-sRNAs se detectaron en callos. Solamente se encontraron 563 secuencias únicas correspondientes a una abundancia de 90 TPM. La mayoría de ArT-sRNAs detectados en raíces pilosas alinean con unidades transcripcionales pertenecientes a los genes del T-DNA como son: *ORF2*, *ORF8*, *rolA*, *rolB*, *ORF13*, *ORF14* y la opina *CUS*. El oncogén *rolA* es la mayor fuente de ArT-sRNAs. Es importante destacar que prácticamente no se detectaron ArT-sRNAs provenientes del oncogén *rolC* en la librería de raíces pilosas. Además, solamente 134 secuencias únicas de ArT-sRNAs se comparten entre las dos bibliotecas. En el trabajo de Dunoyer et al. (2006), se observó que los microRNAs se acumulaban en tumores inducidos por *A. tumefaciens*, que presentan un estado de supresión de las rutas de silenciamiento, de manera similar a como se acumulaban en hojas y tallos. Solamente la acumulación en los tumores de los microRNAs estaba ligeramente disminuida. Por lo que en tumores, a pesar de la particularidad del estado de supresión observado, también se producen microRNAs como en tejidos diferenciados. Por esta razón, se monitoreó la acumulación de microRNAs en raíces pilosas y en callos inducidos por *A. rhizogenes*. Se observó que la acumulación de miRNAs en raíces pilosas no se encuentra ni ligeramente disminuida como en tumores inducidos por *A. tumefaciens* en comparación con la acumulación de raíces silvestres (Figura 6a). De manera similiar a lo que ocurre en tumores inducidos por *A. tumefaciens*, casi todos los miRNAs analizados tienen una menor acumulación en callos que en raíces pilosas (Figura6b). Inesperadamente, el microRNA miR319 fue el miRNA más abundante en callos. Debido a que los RNAs pequeños pueden moverse de una célula a otra y recorrer grandes distancias, estudiamos el movimiento de los ArT-sRNAs de las raíces pilosas hacia hojas no transformadas en raíces compuestas de frijol que no tienen raíz silvestre (Brosnan & Voinnet, 2011). Se generó una librería de RNAs pequeños de hojas de frijol de raíces compuestas para identificar ArT-sRNAs en este tejido. De manera sorprendente,

detectamos ArT-sRNAs en este tejido. Las secuencias detectadas comprenden ArT-sRNAs de 18, 19, 20, e incluso, 30 nucleótidos (Figura 7).

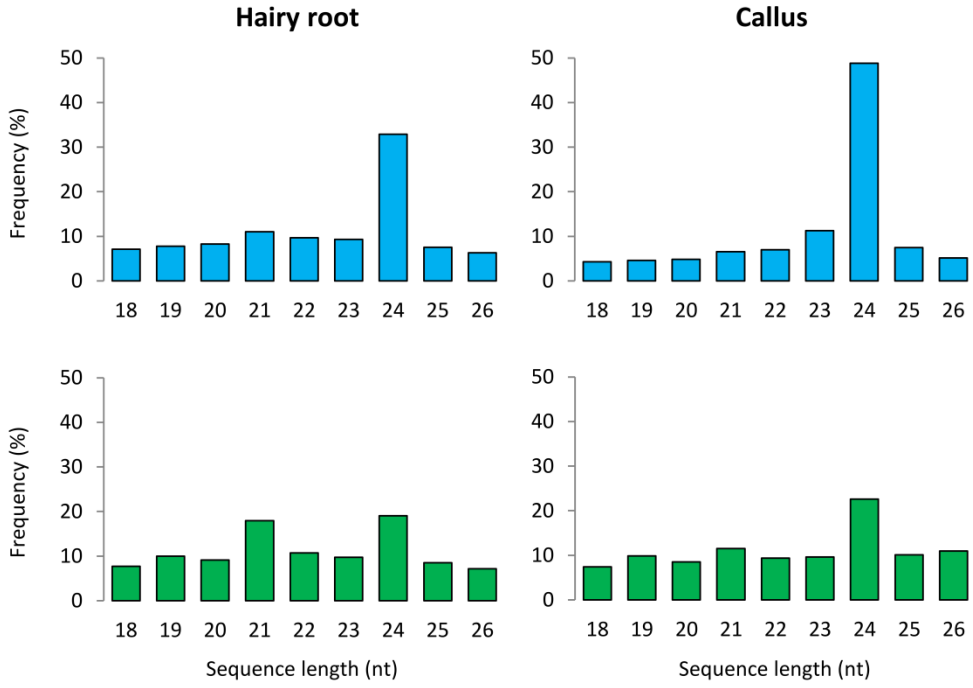


Figura 4. Distribución del largo de las secuencias de sRNAs de las librerías de callos y raíces pilosas. Porcentaje promedio (eje-Y) de secuencias unicas (barras azules) y secuencias redundantes (barras verdes) de 18 a 26 nt (eje-X) de largo para cada una de las librerías secuenciadas.

3.2.2 Silenciamiento génico post-transcripcional de genes del T-DNA productores de RNAs pequeños en fase.

Los ArT-sRNAs son RNAs pequeños principalmente de 21-nt. En raíces pilosas, 52.2% de los ArT-sRNAs, que representan un 67.9% de las secuencias totales que alinearon con el T-DNA, son de 21-nt de largo. Únicamente el 5% de las secuencias totales corresponden a la clase de 24-nt. La clase de 22-nt de ArT-sRNAs fue la segunda más diversa (18%) y abundante (19.7%). Un patrón con similitudes jerárquicas se observó para los ArT-sRNAs detectados en callos (Figura 8a). Los ArT-sRNAs se produjeron de las cadenas sentido y antisentido del T-DNA. Las razones de los ArT-sRNAs generados entre ambas cadenas fueron cercanas a uno. Cincuenta y dos por ciento de los ArT-sRNAs

encontrados en raíces pilosas alinearon con la cadena sentido, mientras que el cuarenta y ocho por ciento alineó con la cadena antisentido (Figura 8b). Analizando genes particulares codificados en la cadena sentido o antisentido, no se encontró preferencia por alguna de las cadenas en la producción de ArT-sRNAs con respecto a la posición del gen.

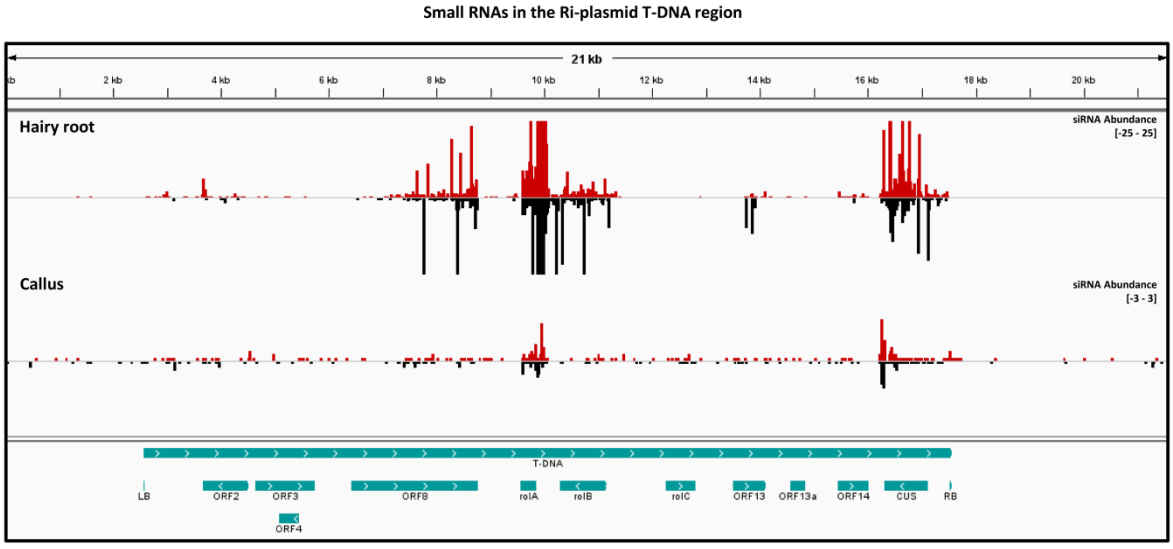


Figura 5. Abundancia y distribución de los RNAs pequeños generados de la región del T-DNA de *A. rhizogenes*. ArT-sRNAs que alinean con el T-DNA de *A. rhizogenes* (azul verde) en la cadena sentido (rojo) y en la cadena antisentido (negro) se muestran para las bibliotecas de raíces pilosas (parte superior) y callos (parte inferior). La escala de la abundancia (derecha) difiere para las dos bibliotecas. Se muestra en la parte inferior (azul verde) los bordes del T-DNA y las regiones codificantes.

Por otra parte, RNAs pequeños de 21-nt en fase, característicos de loci productores de phasiRNAs, se detectaron para los genes *ORF8*, *rolA*, *rolB*, *ORF13*, *ORF14* y *CUS* (Tabla 2). Genes como *rolA*, *ORF8*, *ORF13* y *ORF14* presentan solo una zona de agrupamiento de RNAs pequeños con una fase predominante, mientras que *rolB* y *CUS* tienen dos zonas de agrupamiento que no tienen fase predominante. Algunos RNAs pequeños en fase del transcrito de *rolA* se encuentran entre los más abundantes de los ArT-sRNAs detectados. Estructuras de tipo tallo-asa se observaron en las estructuras secundarias de los transcritos del T-DNA, lo que podría indicar que algunos de éstos también puedan ser directamente procesados por las DCLs. El análisis de la acumulación de los transcritos del T-DNA, así como del transcrito de *AGO1* que se induce en infecciones virales, mostró que *rolA* fue el transcrito más acumulado en callos y el menos en raíces pilosas. De acuerdo al análisis de acumulación diferencial de los transcritos entre

callos y raíces pilosas, solamente los transcritos *rolB* y *CUS* se encontraron significativamente incrementados en su expresión en raíces pilosas (Figura 9).

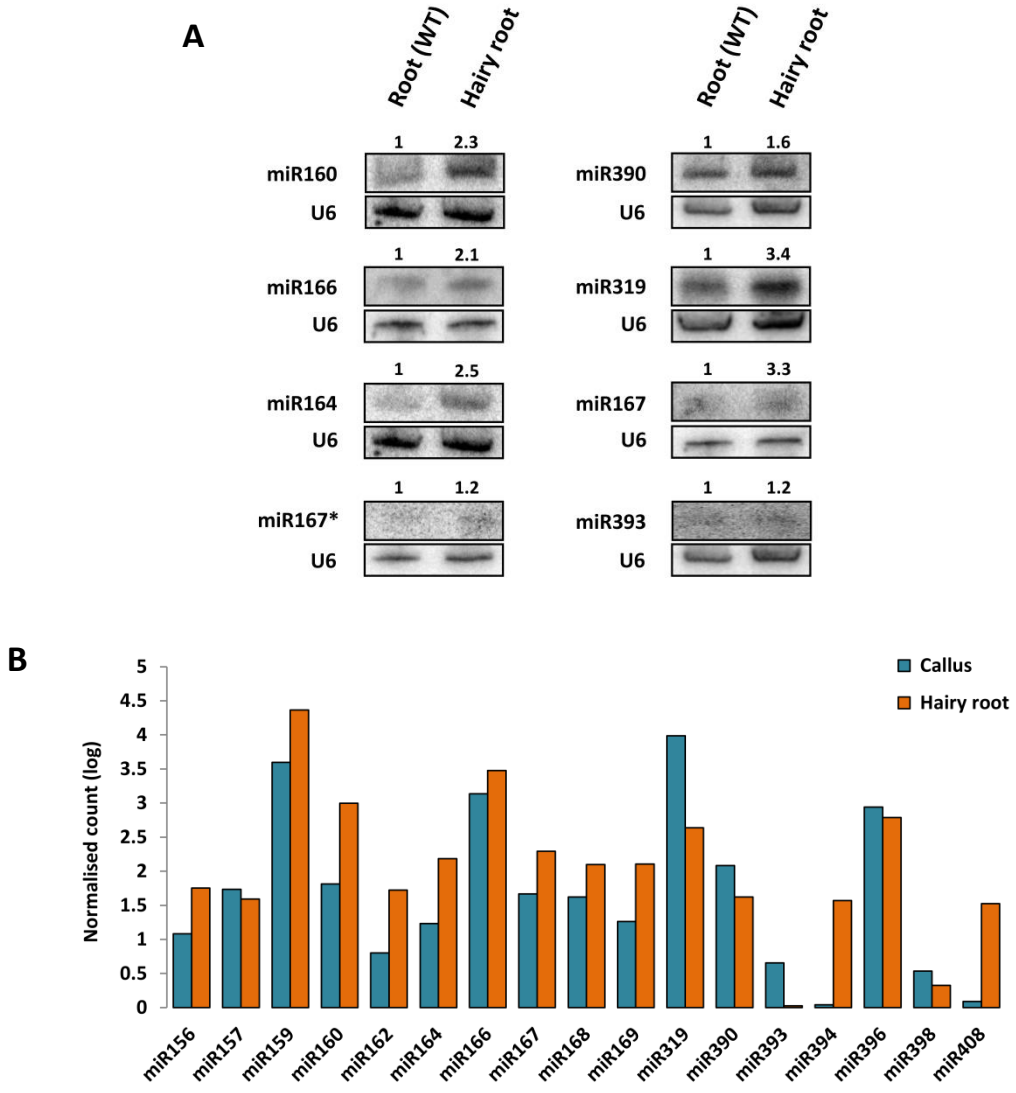


Figura 6. Acumulación de microRNAs en raíces pilosas y callos. (A) Análisis de northern blot para microRNAs en raíces pilosas y raíces silvestres de frijol. El snRNA U6 fue usado como control. La razón de la acumulación entre raíz silvestre/raíz pilosa fue calculada (arriba de cada gel). El miembro miR167d con 22-nt (*) de largo también fue detectado. (B) Análisis de la acumulación diferencial de microRNAs de las bibliotecas de RNAs pequeños de callos y de raíces pilosas. Las frecuencias de las lecturas fueron normalizadas (TPM) y representadas en forma logarítmica (eje-Y).

3.2.3 Análisis de los cortes inducidos por ArT-sRNAs sobre el T-DNA

Cuarenta por ciento de los ArT-sRNAs de 21-nt de largo tienen un uracilo en el extremo 5´ terminal. RNAs pequeños con esta característica pueden ser reclutados por AGO1. Para obtener información respecto al corte de transcritos del T-DNA por la asociación de los ArT-sRNAs con la proteínas AGO, una biblioteca de tipo degradoma (también llamada parallel analysis of RNA ends, PARE) de raíces pilosas fue secuenciada masivamente y analizada (German et al., 2009). La secuenciación del degradoma de raíces pilosas generó 19,064,598 secuencias totales crudas. 446 secuencias distintas alinearon con el T-DNA de *A. rhizogenes*. Se detectaron secuencias del degradoma correspondientes a todas las regiones codificantes del T-DNA con excepción del *ORF13*. La mayoría de las lecturas provienen de la secuencia de los genes *rolA* y *CUS*. El análisis para la identificación de interacciones entre ArT-sRNAs y secuencias del degradoma reveló 180 interacciones. Estas interacciones constituyen 73 posiciones de corte únicas. De las interacciones detectadas, treinta y cuatro y veinte por ciento de los RNAs pequeños fueron de 21 y 22 nucleótidos de largo respectivamente. De acuerdo a la abundancia de las lecturas del degradoma, las interacciones se clasifican en cuatro categorías (Folkes et al., 2012). La categoría 0 se define como una señal que presenta más de una secuencia cruda en la posición de corte. La abundancia en la posición es igual a la máxima en el transcrito y solo hay una máxima. La categoría 1 es igual a la categoría 0 en todos los aspectos excepto que existe más de una abundancia máxima en el transcrito. Esto implica que hay dos o más señales en el transcrito con la misma abundancia. La categoría 2 se define como una señal que tiene más de una secuencia cruda en la posición. La abundancia en la posición es menor a la máxima pero mayor a la abundancia media en el transcrito. La categoría 3 se define como una señal que tiene más de una secuencia cruda en la posición y la abundancia en la posición es menor o igual al valor de la media para el transcrito. La categoría 4 se define como una señal donde solo hay una secuencia cruda en la posición. De acuerdo a estas clasificaciones, cinco interacciones cayeron en la categoría 0, 116 en la categoría 2 y 59 en la categoría 4. Casi todas las interacciones encontradas, alrededor de 95.5%, indican sitios de corte para el transcrito de *rolA*, principalmente de la región 3´ no traducida.

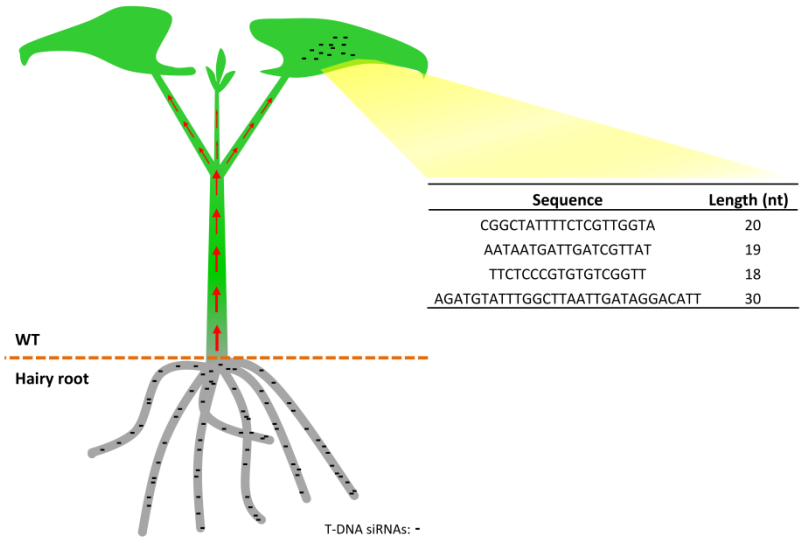


Figura 7. Movimiento de ArT-sRNAs. Esquema del movimiento de ArT-sRNAs (líneas negras) en plantas compuestas de las raíces pilosas (gris) a las hojas no transformadas (verde). Secuencias con su respectivo tamaño son mostradas (derecha).

Curiosamente, uno de los ArT-sRNAs que se agrupó en la categoría 0, y por lo tanto corta el fragmento más abundante, fue el ArT-sRNA más abundante de la biblioteca. Este ArT-sRNA de 22-nt de largo con una frecuencia absoluta de 3,304 en la librería tiene otra posición de corte que incluye dos nucleótidos no apareados. La posición de corte de este ArT-sRNA en el transcrito de *rolA* concuerda el inicio del registro de fase que se detectó en *rolA* (Figura 10). Además, la secuencia de este ArT-sRNA está altamente conservada entre los transcritos de rolA pertenecientes a diferentes plasmidos Ri de diferentes cepas: pRi8196, pRi1724, pRiA4 y pRi2659.

3.2.4 Silenciamiento de genes del hospedero por ArT-sRNAs

Los ArT-sRNAs pueden presentar el apareamiento de bases requerido con secuencias génicas del hospedero para regularlas. La predicción de blancos entre ArT-sRNAs y transcritos de frijol, nos permitió identificar 399 interacciones únicas entre ArT-sRNAs y transcritos de frijol utilizando parámetros rigurosos para la evaluación de la complementariedad. Las interacciones detectadas se componen de 299 transcritos distintos y 326 ArT-sRNAs distintos. Para obtener mayor evidencia sobre el posible corte de transcritos endógenos llevado a cabo por ArT-sRNAs, se analizó la biblioteca de degradoma. Se detectaron 402 interacciones entre ArT-sRNAs y transcritos del hospedero. El número total de transcritos únicos predichos de ser cortados por 346 ArT-sRNAs únicos es 228. La mayoría de los ArT-sRNAs que componen estas interacciones son de 21-nt (53.4%). Tomando en cuenta todas las interacciones, 58 caen en la categoría 0, 50 en la categoría 1, 155 en la categoría 2, 3 en la categoría 3 y 136 en la categoría 4. Los blancos identificados comprenden un diverso grupo de familias de proteínas. Se asignaron clasificaciones de tipo Gene Ontology (GO) a noventa y ocho de los blancos únicos. Los tres grupos más abundantes dentro de la categoría de componente celular fueron: componentes citoplasmáticos (17.2%), núcleo (17.2%) y componentes intracelulares (16.7%). Dentro de la categoría de procesos biológicos, procesos celulares (28.2%), procesos metabólicos (20%) y respuestas a estrés (9.6%) fueron los grupos más abundantes. En la categoría de función molecular, los grupos más abundantes correspondieron a unión (16.4%), actividad enzimática (14.23%), actividad de transferasa (12.2%; Figura 11). Con base en la abundancia del fragmento donde se registró el corte y la frecuencia del ArT-sRNA que lo corta, seis interacciones fueron seleccionadas como interacciones de candidatos fuertes. Los blancos de estas interacciones incluyen proteínas involucradas en defensa y en ciclo celular (Figura 12).

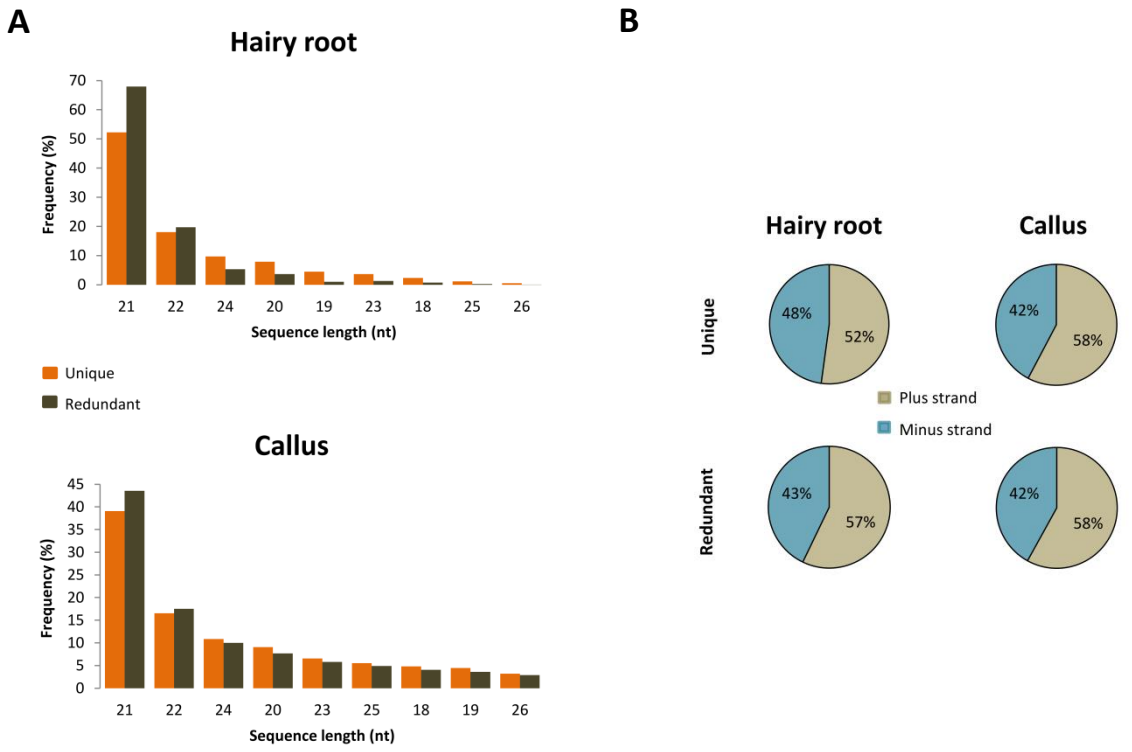


Figura 8. Tamaño y cadena de origen de los ArT-sRNAs. (A) Distribución del largo de las secuencias (eje-X) de ArT-sRNAs encontrados en la biblioteca de raíces pilosas y en la de callos. La abundancia de las secuencias se muestra en términos de porcentaje (eje-Y). (B) Porcentaje de secuencias únicas y redundantes que alinean con la cadena más (gris) y la cadena menos (azul) del T-DNA para cada una de las librerías.

4. Discusión

La identificación de RNAs pequeños en especies con o sin secuencia completa del genoma se ha revolucionado por las nuevas metodologías de secuenciación masiva. A pesar de que los experimentos realizados con datos de secuenciación masiva se han convertido en la principal fuente que respalda la anotación de microRNAs, diferentes variaciones técnicas y sesgos han sido reportados para este tipo de técnicas (Kozomara & Griffiths-Jones, 2013). Hacen falta estudios para elucidar el número de moléculas secuenciadas a partir de una biblioteca de RNAs pequeños para llevar a cabo estudios más precisos sobre RNAs pequeños. En términos de lecturas, las bibliotecas de RNAs pequeños secuenciadas en este estudio generaron un gran número de secuencias crudas (Tabla 1). Sin embargo, el número total de lecturas de miRNAs identificados, incluyendo isoformas, constituye solamente el 0.047% de todas las secuencias crudas. Más estudios son

necesarios para determinar la proporción en una muestra de RNAs pequeños secuenciados como tRNAs, rRNAs, snoRNAs y snRNAs.

PhasiRNA-generating loci from T-DNA

Min. Abundance (1)ł

Region*	Gene	Strand$^{\Sigma}$	Distinct Seqs.‡	Distinct Phased Seqs.$^{\wedge}$	P-value$^{£}$
10251-10502	rolB	W	61	8	3.01E-03
11074-11325	rolB	W	38	6	4.55E-03
13711-13962	ORF13	C	25	5	3.50E-03
15731-15982	ORF14	C	13	5	1.22E-04
16404-16655	CUS	W	139	14	4.45E-04
17053-17304	CUS	W	50	7	4.11E-03
8660-8911	ORF8	C	30	7	1.49E-04
9731-9982	rolA	C	226	17	2.27E-03
	TAS3	C	58	9	3.79E-04

Min. Abundance (2)ł

Region*	Gene	Strand$^{\Sigma}$	Distinct Seqs.‡	Distinct Phased Seqs.$^{\wedge}$	P-value$^{£}$
10737-10988	rolB	W	25	5	3.50E-03
11074-11325	rolB	W	15	4	3.15E-03
16404-16655	CUS	W	80	11	1.82E-04
17293-17544	CUS	W	14	4	2.38E-03
8723-8974	ORF8	C	3	3	8.20E-05
9582-9833	rolA	C	91	12	1.12E-04
	TAS3	C	30	9	1.05E-06

Tabla 2. Genes del T-DNA generan phasiRNAs. Identificación de phasiRNAs sobre los genes del T-DNA de *A. rhizogenes* con el método de Chen et al., 2007 tomando en cuenta ArT-sRNAs con abundancia de uno (tabla superior) o mayor a uno (tabla inferior; ł). Se muestra la region donde se localizan los phasiRNAs en pares de bases (*; número de acceso GenBank - EF433766), si la cadena donde se localizó la fase es Watson o Crick (Σ), el número de secuencias que alinean en la región de la fase (ǂ), el número de ArT-sRNAs en fase (ʌ) y el P-value(£).

La mayoría de las familias de microRNAs identificadas en este estudio en *P. vulgaris* se encuentran conservadas en *Fabaceae*, particularmente en las dos plantas más estudiadas *M. truncatula* y *G. max*. Algunas familias conservadas como son miR397 y miR408 no habían sido reportadas para *G. max* o *M. truncatula* en la base de datos miRBase. También, algunos miembros de familias conservadas, identificados en este estudio, no habían sido reportados en leguminosas. Por ejemplo un miembro de la familia miR397, altamente abundante, con la misma secuencia que ath-miR397. Es interesante que familias detectadas en soya como miR1511, miR1514 y miR1515 no hayan sido detectadas

en otras plantas aparte de frijol. En este sentido, otras familias reportadas específicamente en soya (gma-miR1524, gma-miR1532, gma-miR1526, gma-miR1516, gma-miR1513 y gma-miR1508) fueron detectadas por Valdés-López et al. (2010) utilizando una estrategia de hibridación (macroarreglos) bajo diferentes condiciones de estrés abiótico en hojas, raíces y nódulos de frijol. En este estudio no se detectaron a estas familias con excepción de dos lecturas para gma-miR1508. En el análisis de Valdés-López et al. (2010) se detectó expresión para los microRNAs miR1524, miR1526, miR1532 y miR1508 en plantas de frijol bajo condiciones nutricionales adecuadas y bajos condiciones estresantes. Las familias miR1513 y miR1516 se detectaron solamente bajo condiciones de estrés. La familia miR170, para la cual Valdés-López et al. (2010) detectó expresión en frijol, tampoco se encontró en este estudio, probablemente debido a la similitud que existe entre esta familia y miR171.

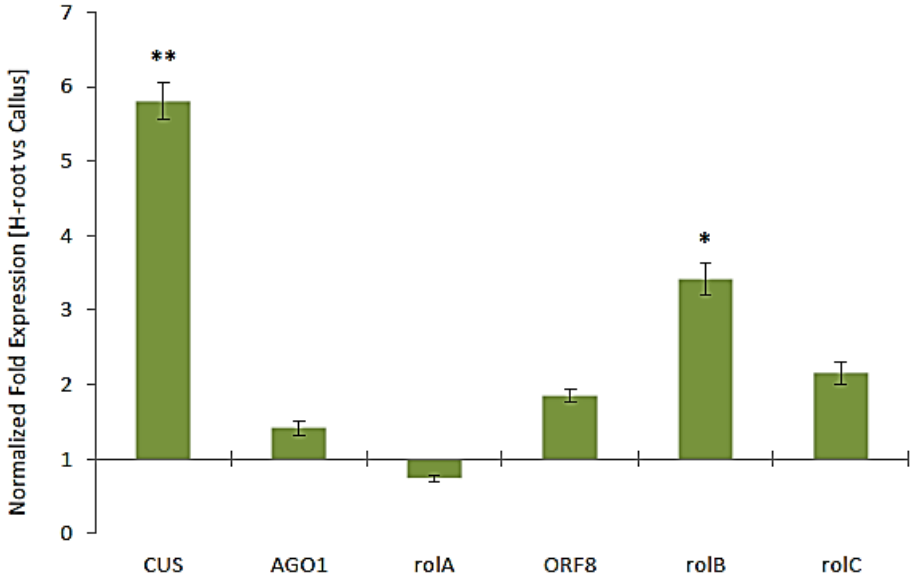


Figura 9. Analisis de PCR tiempo-real de transcritos en raíces pilosas y callos. La cuantificación se basa en el valor de Ct (cycle threshold) y la expresión de los genes se normalizó con el gen de referencia *EF1* (elongation factor 1). Los valores de expresión de los genes de la muestra de callos se usaron como fondo para mostrar la diferencia de expresión (n=3; *P<0.05; **P<0.005).

Las isoformas de miRNAs son consideradas productos de las imprecisiones de las DCLs en el proceso de corte de los precursores. Variantes de mucho menor tamaño o poco frecuentes son consideradas en muchos casos productos de degradación o errores de

secuenciación. Aquí, las secuencias de RNAs pequeños de las bibliotecas se clasificaron como isoformas si eran similares a un microRNA de referencia y si eran abundantes. De este análisis cuatro familias más de microRNAs fueron identificadas: miR1510, miR2199, miR4376 y miR479. Todas las variantes que se encontraron para estas cuatro familias fueron muy abundantes en las cuatro librerías. Es muy probable que las familias miR1510 y miR2199 formen parte de la población de microRNAs de *P. vulgaris* debido a su conservación en otras especies y a los resultados de los análisis de expresión previamente hechos (Arenas-Huertero et al., 2009; Valdés-López et al., 2010). Otra familia que es altamente probable que se encuentre presente en frijol, es la familia miR4376. La variante encontrada aquí para esta familia fue muy abundante y tiene solamente un nucleótido menos en relación con el miRNA de referencia de soya. Análisis bioinformáticos y experimentales se requieren para validar la variante identificada en frijol, correspondiente a la familia miR479, que tiene dos nucleótidos menos y un nucleótido diferente en comparación con el miRNA de referencia csi-miRNA479. Este es el caso también para otras familias identificadas que se encuentran menos conservadas como: miR858, miR2597, miR894 y miR1310. Así mismo, es indispensable determinar con futuros análisis hasta qué grado los miRNAs nuevos que se identificaron se confirman como genuinos microRNAs específicos de frijol y si están involucrados en procesos biológicos que definen esta especie.

Actualmente, enfoques de hibridación con microarreglos, análisis con PCR tiempo real, y tecnologías de secuenciación masiva son métodos ampliamente utilizados para conocer el perfil de microRNAs. Se desconoce todavía la función de diversos microRNAs en diferentes órganos, en diferentes condiciones de estrés o en diferentes etapas de desarrollo. Los estudios de perfiles de expresión de microRNAs aportan información preliminar para analizar la función de los microRNAs en determinadas condiciones. En este trabajo la expresión de microRNAs fue analizada para determinar diferencias significativas, en la expresión de ciertas familias de microRNAs, entre diferentes órganos de frijol. Las réplicas biológicas son esenciales para determinar si las diferencias observadas en un estudio son causadas por las condiciones analizadas y no por variaciones experimentales. Por este motivo, se realizó un análisis de agrupamiento. El paquete DESeq permite al usuario trabajar sin replicas con la advertencia de que la prueba perderá solidez.

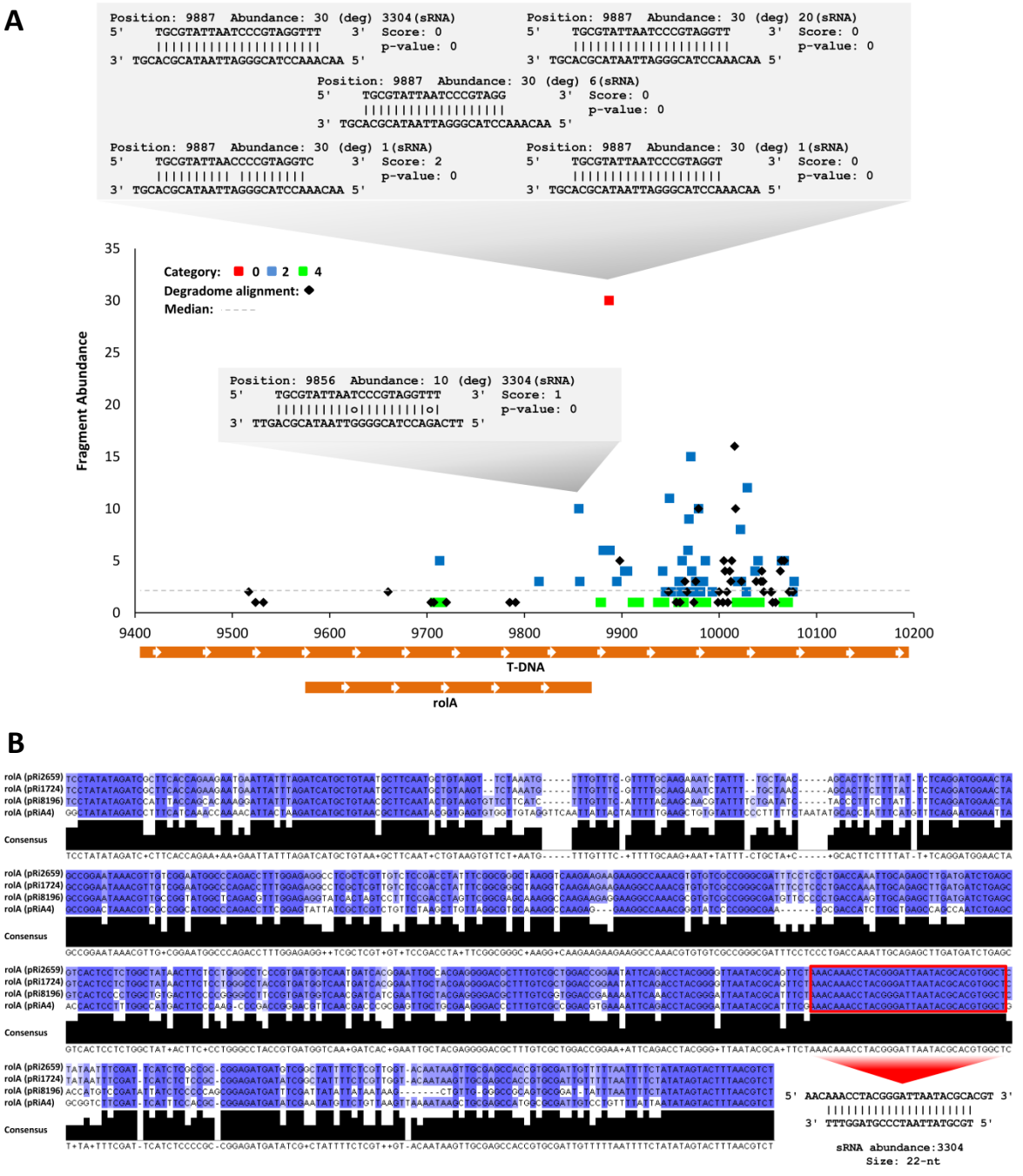


Figura 10. Análisis del corte de *rolA* por ArT-sRNAs. (A) Grafica de blanco (t-plot) de rolA. En recuadros se muestran los alineamientos para el ArT-sRNA más abundante de 22-nt y sus variantes. También se señala el sitio de corte predicho de acuerdo a las secuencias del degradoma (eje-Y). En la región del T-DNA (eje-X) se muestra la región codificante de *rolA* (recuadro inferior). (B) Alineamiento multiple de las secuencias de rolA de diferentes plasmidos Ri. El ArT-sRNA de 22-nt que podría generar la fase observada en *rolA* está altamente conservado entre las secuencias (recuadro y flecha roja).

Los experimentos diseñados para explorar la expresión de microRNAs y mRNAs son sujetos a diferentes sesgos técnicos y biológicos. El proceso diferencial de los precursores, los dúplex de RNAs pequeños y, en general, las propiedades bioquímicas de los microRNAs representan un reto para los métodos de análisis de expresión empleados para analizar mRNAs. La mayoría de las familias conservadas que se identificaron se expresan en todos los órganos en diferentes grados. El siguiente paso será determinar cuáles de los microRNAs conservados juegan un papel esencial en leguminosas para cierta etapa del desarrollo o respuestas a estrés. Probablemente, microRNAs de familias menos conservadas como miR1510, miR1511, miR1514, miR1515, miR2118, miR2199 juegan un papel importante en procesos característicos de leguminosas, como se ha demostrado recientemente para las familias miR482 y miR1515 durante la nodulación de soya (Li et al., 2010a).

En este trabajo, también exploramos el papel del silenciamiento por RNA en la regulación de los genes del T-DNA de *A. rhizogenes* en raíces pilosas de frijol. La patogenicidad de *Agrobacterium rhizogenes* se caracteriza por la inserción de un T-DNA en el genoma de las plantas. Por este motivo, representa un sistema poderoso para transformar plantas. También es capaz de transformar plantas recalcitrantes en las que *A. tumefaciens* no es exitoso. Las raíces inducidas por esta bacteria son utilizadas para realizar estudios de genómica funcional en frijol y otras especies recalcitrantes, incluyendo estudios sobre RNAs pequeños. En contraste con *A. tumefaciens*, *A. rhizogenes* es utilizado frecuentemente con el plásmido inductor de raíces "armado", portando los oncogenes del T-DNA. Por años, los oncogenes de *A. rhizogenes* han sido considerados moduladores del crecimiento neoplásico y la diferenciación de las células (Britton et al., 2008). Recientemente, también han sido caracterizados como moduladores del metabolismo secundario. La caracterización bioquímica y molecular de los genes del T-DNA de *A. rhizogenes* no ha sido del todo clara. En varios estudios, se han reportado resultados diferentes respecto a sus funciones, lo que sugiere que el surgimiento de un nuevo meristemo y la diferenciación posterior de las células transformadas requiere de complejos mecanismos celulares (Britton et al., 2008). A pesar de que el mecanismo de RNA es conocido por regular transgenes, ácidos nucleicos virales y genes del T-DNA de *A.*

tumefaciens, su participación en la regulación de genes del T-DNA de *A. rhizogenes* no había sido explorada.

En este estudio, se encontró que varios genes del T-DNA de A. rhizogenes son regulados por un mecanismo de silenciamiento mediado por RNAs pequeños en raíces pilosas. La producción de RNAs pequeños en raíces pilosas no se encuentra inhibida en comparación con lo que sucede en los tumores inducidos por *A. tumefaciens*. La cantidad de lecturas detectadas para varios ArT-sRNAs es similar a la de algunos microRNAs abundantes encontrados en diversos órganos de frijol (Peláez et al., 2012). Los pocos ArT-sRNAs detectados en callos, junto con la reducida acumulación de la mayoría de los miRNAs que se observó, sugieren que los callos también presentan un estado de supresión de las rutas de silenciamiento similar al observado en tumores. Los pocos ArT-sRNAs que son compartidos entre las bibliotecas de callos y raíces pilosas puede ser una consecuencia de las diferentes fuentes de RNA de doble cadena usadas como moldes para su producción. Por ejemplo, el T-DNA podría ser silenciado previo a su integración en el genoma de la planta. También, pensamos que la acumulación diferencial de miRNAs observada entre raíces silvestres y raíces pilosas de frijol podría implicar diferencias fenotípicas entre estos dos tipos de raíces. De manera sorpresiva, se encontró que miR319 fue el microRNA más abundante en la librería de callos. Usualmente, en todas las otras bibliotecas de RNAs pequeños secuenciadas de frijol miR159 fue el más abundante. La familia miR319 tiene como blancos a los genes *TCP* (TEOSINTE BRANCHED-CYCLOIDEA/PROLIFERATING CELL FACTORS) que están involucrados en proliferación y diferenciación celular. La disminución en la actividad de los genes *TCP* ocasionada por una mayor acumulación de miR319 promueve la proliferación celular. Nosotros pensamos que este microRNA es importante en la formación del callo.

De los transcritos del T-DNA que producen ArT-sRNAs, el transcrito rolA se destaca como la mayor fuente de ArT-sRNAs. Curiosamente, no se detectaron ArT-sRNAs pertenecientes al intron de la región 5´ no traducida de *rolA*. De acuerdo con todos los transcritos del T-DNA analizados, la acumulación de *rolA* difiere entre callos y raíces pilosas. La función precisa de la proteína rolA no se conoce. La poca expresión de *rolA* se ha reportado también en raíces y hojas. Plantas de tabaco que expresan *rolA* presentan hojas arrugadas, internodos cortos y flores anormales (Schmülling et al., 1988; Carneiro &

Vilaine, 1993). La expresión local de *rolA* en haces vasculares reduce el tamaño de las células del parenchyma de alrededor causando el fenotipo de hojas arrugadas. Por este motivo, y por estudios de injertos que involucran a *rolA*, se ha propuesto que rola genera un factor difusible (Britton et al., 2008). Nosotros postulamos que los ArT-sRNAs generados de *rolA* podrían ser buenos candidatos para constituir este factor difusible. También, otra posibilidad para los ArT-sRNAs, de *rolA* o en general, de indirectamente afectar el fenotipo de las plantas podría ser a través de la alteración de los procesos regulados por RNAs endógenos como se ha descrito durante el silenciamiento post-transcripcional de material genético foráneo (Martinez de Alba et al., 2011). Además, resulta interesante que no se detectaron ArT-sRNAs para el gen rolC en hairy roots, uno de los genes del T-DNA ampliamente estudiados también. Previos reportes sobre la expresión fuerte y estable de *rolC* en raíces de plantas transformadas junto con la nula detección de ArT-sRNAs generados de este gen, sugieren que *rolC* no es silenciado en raíces (Britton et al., 2008).

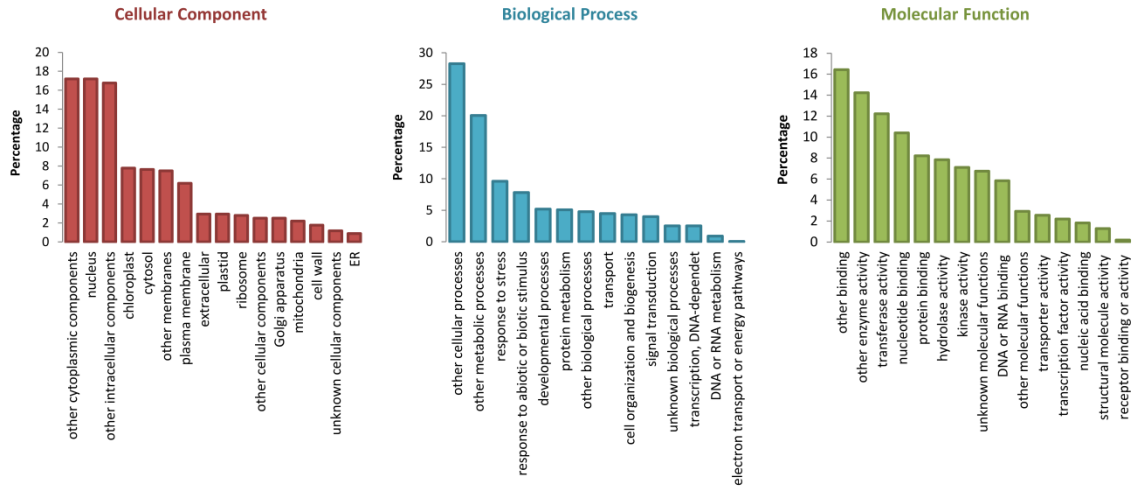


Figura 11. Categorización de blancos candidatos regulados por ArT-sRNAs. Los transcritos de frijol que son candidatos blancos de ArT-sRNAs se categorizaron de acuerdo al Gene Ontology y la proporción de cada categoría es mostrada (eje-Y).

Los RNAs pequeños son señales del silenciamiento sistémico capaces de moverse de una célula a otra célula. El silenciamiento sistémico ocurre durante el silenciamiento post-transcripcional de virus o transgenes (Brosnan & Voinnet, 2011). El movimiento a

larga distancia de los RNAs pequeños se ha observado que se lleva a cabo predominantemente de la parte aérea a la raíz. En este estudio, nosotros detectamos ArT-sRNAs en hojas no transformadas de plantas compuestas. Estos ArT-sRNAs fueron pocos y pobremente abundantes, sin embargo, este resultado nos sugiere fuertemente que RNAs pequeños generados en las raíces pilosas podrían disparar el silenciamiento sistémico a lo largo de la planta. En este sentido, un débil silenciamiento del gen reportero de la Beta-glucuronidasa (*GUS*) se ha observado en hojas de una línea transgénica de *Lotus japonicus* que tiene raíces pilosas expresando un tallo-asa de RNA complementario a este gen (Kumagai & Kouchi, 2003). Además, es interesante destacar que se detectó un ArT-sRNA con un tamaño de 30 nt que es capaz de moverse. Los RNAs pequeños que se han reportado que se pueden mover de una célula a otra han sido de menor tamaño (Brosnan & Voinnet, 2011).

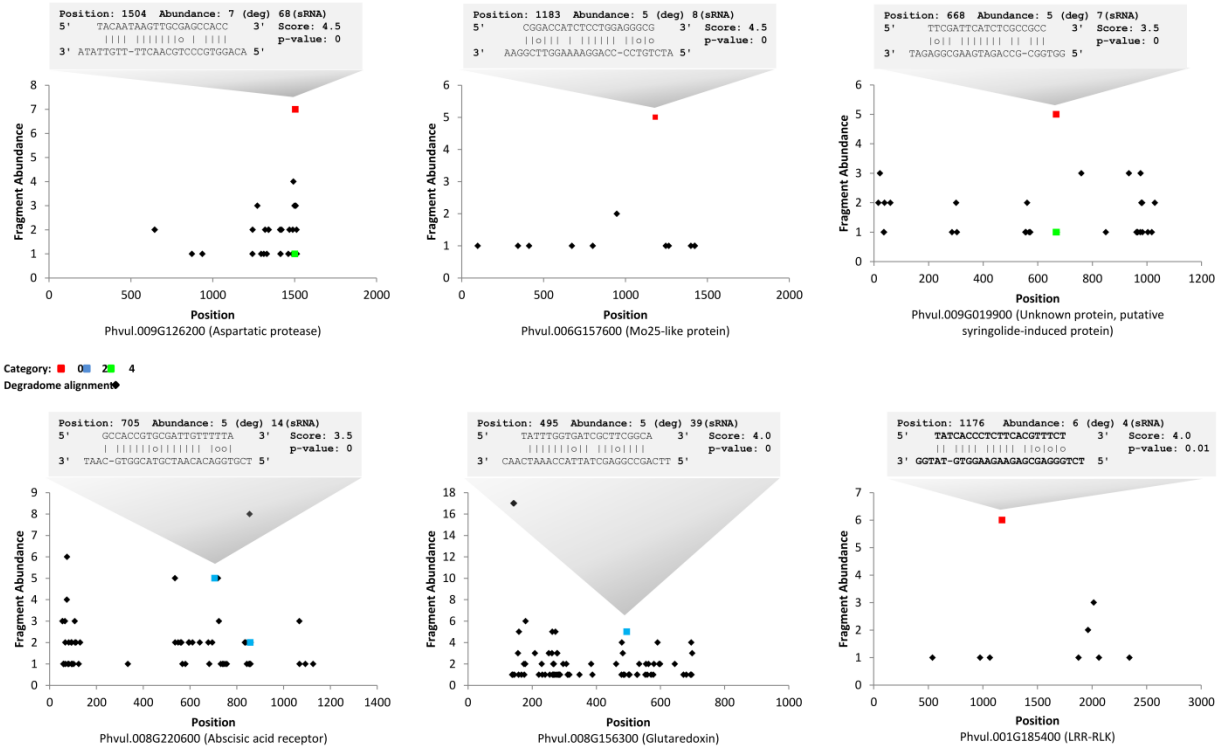


Figure 12. Gráficas blanco de transcritos de frijol. La interacción de los transcritos blanco y los reguladores ArT-sRNAs se indica arriba de cada gráfica. El eje-X muestra la posición de las señales de corte y el eje-Y la intensidad de las señales basándose en la información del degradoma.

La detección principalmente de ArT-sRNAs de 21-nt y 22-nt de tamaño nos indica que, al igual que sucede en las interacciones virales (principalmente de virus de RNA), el silenciamiento de los genes del T-DNA es a nivel post-transcripcional. También, nos indica que las principales enzimas DCLs que procesan los transcritos de los genes del T-DNA son las más activas en la defensa antiviral: DCL4 y DCL2. El encontrar que varios de estos genes generan phasiRNAs de 21-nt muestra una participación activa de una o más de las RDRs y, nuevamente, de la dependencia de DCL4. Así mismo, que es muy probable que los ArT-sRNAs puedan ser reclutados por las proteínas AGO para llevar a cabo una regulación en *cis* e inducir la amplificación de la señal de silenciamiento. Debido a las estructuras tipo tallo-asa que encontramos en las estructuras secundarias de los transcritos de los genes del T-DNA no se descarta que podrían existir diversos templados para la producción de ArT-sRNAs como ocurre con los vsiRNAs.

El hecho de que se hayan encontrado muchos ArT-sRNAs de 21-nt de largo con un uracilo en el extremo 5´ nos sugiere, que al igual que sucede en infecciones virales, AGO1 podría tener un papel importante en la regulación de los genes del T-DNA. El análisis de las secuencias de degradoma nos indica que ArT-sRNAs pueden ser reclutados por las proteínas AGO para regular los genes del T-DNA. De la misma manera que ha sido reportado para otros RNAs pequeños de otras especies, la mayoría de interacciones que identificamos con microRNAs conservados de frijol utilizando el programa PAREsnip son interacciones validadas experimentalmente (Folkes et al., 2012). Por este motivo, la cantidad de interacciones detectadas entre ArT-sRNAs y transcritos de frijol apunta a que los ArT-sRNAs podrían impactar directamente en el fenotipo de las raíces pilosas a través de la regulación de sus transcritos. Análisis experimentales son necesarios para confirmar esta regulación.

Creemos que el análisis realizado en esta tesis sobre el silenciamiento por RNA como mecanismo de defensa en contra de los genes del T-DNA de *A. rhizogenes* incide directamente en nuestro entendimiento y uso de las raíces pilosas. Por ejemplo, construcciones introducidas en estas raíces deben tomar en cuenta la población de RNAs pequeños generadas del T-DNA para evitar su silenciamiento. También, habría que evaluar si la transgénesis que ocurre, sujeta a silenciamiento, favorece o reduce la eficiencia de las construcciones utilizadas para silenciar genes. Pensamos que las funciones asignadas a

diversos genes del T-DNA en estudios previos, tienen que ser reevaluadas de acuerdo a lo que se observó en este estudio. Nuestros datos sugieren que probablemente más genes del T-DNA de *A. tumefaciens* podrían ser regulados al inicio de la infección. Será interesante evaluar si la regulación del T-DNA por RNAs pequeños favorece el surgimiento de las raíces pilosas o hasta que ocurre la diferenciación celular el silenciamiento por RNA regula los genes del T-DNA.

5. Conclusiones

Se identificaron 109 microRNAs conservados pertenecientes a 29 familias en frijol utilizando secuenciación masiva. Además, veintiséis isoformas de miRNAs fueron caracterizadas. El análisis sobre isoformas identificó cuatro familias abundantes más de microRNAs. Once nuevos precursores tallo-asa pertenecientes a ocho familias conservadas fueron determinados. Treinta y nueve microRNAs identificados puede explicarse por el procesamiento de precursores caracterizados de frijol. También, 29 nuevos microRNAs candidatos fueron predichos basados en secuencias de RNAs pequeños y predicciones de precursores. Se cuenta con secuencias de microRNAs estrella para cuatro de estos precursores. Con el análisis de la predicción de blancos se identificaron 157 nuevos ESTs de frijol como genes blanco establecidos de microRNAs. Se propusieron nuevos blancos candidatos para frijol. Se identificaron familias de microRNAs que se expresan diferencialmente en hojas, raíces, plántulas y flores de frijol. Estos análisis aportan una visión global de microRNAs nuevos y conservados en *Phaseolus vulgaris*.

Analizando el silenciamiento por RNA de genes del T-DNA de *A. rhizogenes* en raíces pilosas de frijol, se encontró que las rutas de silenciamiento no están suprimidas en raíces pilosas. Mostramos que el silenciamiento por RNA ejerce un papel de defensa en contra de varios genes del T-DNA como en infecciones virales. Se detectaron distintos RNAs pequeños generados del T-DNA pertenecientes a diversos oncogenes y al gen de una opina sintasa. Nuestros datos sugireren que muchos de los genes del T-DNA constituyen loci productores de phasiRNAs dependientes de DCL4 y que son sujetos a silenciamiento post-transcripcional por RNAs pequeños de 21-nt y 22-nt de largo. Detectamos que la acumulación de miRNAs en raíces pilosas no se encuentra

moderadamente reducida como sucede en tumores inducidos por *A. tumefaciens*. También encontramos que los ArT-sRNAs se mueven hacia la parte aérea no transformada y que estos RNAs pequeños generados a partir de DNA bacteriano pueden tener como blancos a genes endógenos de frijol. Los bajos niveles detectados de ArT-sRNAs en callos inducidos por *A. rhizogenes* sugiere que poseen un estado de supresión de las rutas de silenciamiento similar al observado en tumores. Los resultados que se presentan en esta tesis respecto a las respuestas de defensa llevadas a cabo por RNAs pequeños para regular genes del T-DNA en raíces pilosas, sugieren que el mecanismo de silenciamiento por RNA tiene un papel importante en eventos distintivos de esta interacción planta-bacteria como son la proliferación y la diferenciación celular.

6. Perspectivas

Los resultados presentados en esta tesis sientan un precedente para otros estudios sobre RNAs pequeños de frijol y otras especies de plantas que se estudien por medio de las raíces pilosas, así como aquellos que involucren el uso de *A. rhizogenes*. Un análisis más completo para caracterizar la población de microRNAs de frijol será posible con las secuencias genómicas completas de diferentes cultivares de frijol ahora disponibles. También la disponibilidad de transcriptomas de frijol generados de varios tejidos y condiciones de crecimiento permitirán identificar precursores y blancos de microRNAs de manera más eficiente. Varios miembros de las familias de microRNAs conservadas que se identificaron en este estudio y que son poco abundantes tendrán que ser evaluados experimentalmente para determinar si tienen un impacto en la regulación de genes blanco. De igual manera, varias isoformas, microRNAs nuevos y blancos nuevos, detectados en este trabajo, requieren mayor evidencia para ratificar su caracterización y evaluar su funcionalidad. Será importante verificar los resultados sobre los microRNAs identificados en frijol y su conservación con otras plantas, tomando en cuenta estudios recientes sobre los RNAs pequeños mal anotados en miRBase o los microRNAs identificados en plantas con mayor cantidad de datos (Montes et al., 2014; Kozomara & Griffiths-Jones, 2013; Hansen et al., 2011). Así mismo, será interesante determinar la causa y consecuencia de la acumulación diferencial de los microRNAs entre órganos de frijol. También, será

interesante comparar los patrones de expresión de los microRNAs de frijol con otro tipo de plantas modelo como *Arabidopsis*. En este sentido, muchas preguntas interesantes quedan por contestar respecto a la acumulación de microRNAs y RNAs pequeños generados del T-DNA en raíces pilosas y callos inducidos por *A. rhizogenes*. Por ejemplo, analizar si la ausencia o sobreexpresión de miR319 tiene un efecto en la formación del callo. También es necesario realizar experimentos relacionados con la identificación de blancos para determinar de manera contundente y precisa los transcritos de frijol regulados por ArT-sRNAs. Finalmente, evaluar el impacto que podría tener la regulación del T-DNA de *A. rhizogenes* por el mecanismo de silenciamiento en el uso de las raíces pilosas para el estudio de RNAs pequeños.

7. Bibliografía

- Achard P, Herr A, Baulcombe DC, Harberd NP. 2004. Modulation of floral development by a gibberellin-regulated microRNA. Development (Cambridge, England) 131: 3357–3365.
- Adenot X, Elmayan T, Lauressergues D, Boutet S, Bouché N, Gasciolli V, Vaucheret H. 2006. DRB4-Dependent TAS3 trans-Acting siRNAs Control Leaf Morphology through AGO7. Current Biology 16: 927–932.
- Allen E, Xie Z, Gustafson AM, Carrington JC. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121: 207–221.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of molecular biology 215: 403–410.
- Alves-Junior L, Niemeier S, Hauenschild A, Rehmsmeier M, Merkle T. 2009. Comprehensive prediction of novel microRNA targets in Arabidopsis thaliana. Nucleic Acids Research 37: 4010–4021.
- Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome biology 11: R106.
- Arenas-Huertero C, Pérez B, Rabanal F, Blanco-Melo D, De La Rosa C, Estrada-Navarrete G, Sanchez F, Covarrubias AA, Reyes JL. 2009. Conserved and novel

miRNAs in the legume Phaseolus vulgaris in response to stress. Plant Molecular Biology 70: 385–401.

- Axtell MJ, Jan C, Rajagopalan R, Bartel DP. 2006. A Two-Hit Trigger for siRNA Biogenesis in Plants. Cell 127: 565–577.
- Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, Braun L, Bergdoll M, Hakimi MA, Lagrange T, et al. 2010. Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes & development 24: 904–915.
- Baulcombe D. 2004. RNA silencing in plants. Nature 431: 356–363.
- Baumberger N, Baulcombe DC. 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: 11928–11933.
- Bazzini AA, Almasia NI, Manacorda CA, Mongelli VC, Conti G, Maroniche GA, Rodriguez MC, Distéfano AJ, Hopp HE, del Vas M, et al. 2009. Virus infection elevates transcriptional activity of miR164a promoter in plants. BMC plant biology 9: 152.
- Benkovics AH, Timmermans MC. 2014. Developmental patterning by gradients of mobile small RNAs. Curr Opin Genet Dev 27: 83-91.
- Bhattacharjee S, Zamora A, Azhar MT, Sacco MA, Lambert LH, Moffett P. 2009. Virus resistance induced by NB-LRR proteins involves Argonaute4-dependent translational control. Plant Journal 58: 940–951.
- Blevins T, Rajeswaran R, Shivaprasad P V., Beknazariants D, Si-Ammour A, Park HS, Vazquez F, Robertson D, Meins F, Hohn T, et al. 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Research 34: 6233–6246.
- Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell 123: 1279–1291.

- Bortolamiol D, Pazhouhandeh M, Marrocco K, Genschik P, Ziegler-Graff V. 2007. The Polerovirus F Box Protein P0 Targets ARGONAUTE1 to Suppress RNA Silencing. Current Biology 17: 1615–1621.
- Bouché N, Lauressergues D, Gasciolli V, Vaucheret H. 2006. An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. The EMBO journal 25: 3347–3356.
- Boutet S, Vazquez F, Liu J, Béclin C, Fagard M, Gratias A, Morel JB, Crété P, Chen X, Vaucheret H. 2003. Arabidopsis HEN1: A genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Current Biology 13: 843–848.
- Britten RJ, Davidson EH. 1969. Gene regulation for higher cells: a theory. Science 165: 349–357.
- Britton MT, Escobar MA, Dandekar AM. 2008. The Oncogenes of Agrobacterium tumefaciens and Agrobacterium rhizogenes. In Agrobacterium: from biology to biotechnology (T. Tzfira and V. Citovsky, eds.) Springer: 523-563.
- Brosnan CA, Voinnet O. 2011. Cell-to-cell and long-distance siRNA movement in plants: Mechanisms and biological implications. Current Opinion in Plant Biology 14: 580–587.
- Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. 2003. Beans (Phaseolus spp.) - Model food legumes. Plant and Soil.55–128.
- Chandler VL. 2007. Paramutation: From Maize to Mice. Cell 128: 641–645.
- Chapman EJ, Carrington JC. 2007. Specialization and evolution of endogenous small RNA pathways. Nature reviews. Genetics 8: 884–896.
- Chen H-M, Li Y-H, Wu S-H. 2007. Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 104: 3318–3323.
- Chen J, Li WX, Xie D, Peng JR, Ding SW. 2004. Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microrna in host gene expression. The Plant cell 16: 1302–1313.

- Carneiro M, Vilaine F. 1993. Differential expression of the rolA plant oncogene and its effect on tobacco development. Plant Journal 3: 785-792.
- Contreras-Cubas C, Rabanal FA, Arenas-Huertero C, Ortiz MA, Covarrubias AA, Reyes JL. 2012. The Phaseolus vulgaris miR159a precursor encodes a second differentially expressed microRNA. Plant Molecular Biology 80: 103–115.
- Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA. 2014. miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. Nature 508: 411–415.
- Csorba T, Bovi A, Dalmay T, Burgyán J. 2007. The p122 subunit of Tobacco Mosaic Virus replicase is a potent silencing suppressor and compromises both small interfering RNA- and microRNA-mediated pathways. Journal of virology 81: 11768–11780.
- Cuperus JT, Fahlgren N, Carrington JC. 2011. Evolution and functional diversification of MIRNA genes. The Plant cell 23: 431–442.
- Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. 2006. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science (New York, N.Y.) 313: 68–71.
- Dhahbi JM, Atamna H, Boffelli D, Magis W, Spindler SR, Martin DIK. 2011. Deep sequencing reveals novel micrornas and regulation of microRNA expression during cell senescence. PLoS ONE 6.
- Diaz-Pendon JA, Li F, Li W-X, Ding S-W. 2007. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. The Plant cell 19: 2053–2063.
- Donaire L, Barajas D, Martínez-García B, Martínez-Priego L, Pagán I, Llave C. 2008. Structural and genetic requirements for the biogenesis of tobacco rattle virus-derived small interfering RNAs. Journal of virology 82: 5167–5177.
- Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. From the Cover: PNAS Plus: Widespread dynamic DNA methylation in response to biotic stress. Proceedings of the National Academy of Sciences 109: E2183–E2191.

- Doyle JJ, Luckow MA. 2003. The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiology 131: 900-910.
- Duan C-G, Fang Y-Y, Zhou B-J, Zhao J-H, Hou W-N, Zhu H, Ding S-W, Guo H-S. 2012a. Suppression of Arabidopsis ARGONAUTE1-Mediated Slicing, Transgene-Induced RNA Silencing, and DNA Methylation by Distinct Domains of the Cucumber mosaic virus 2b Protein. The Plant Cell 24: 259–274.
- Duan C-G, Wang C-H, Guo H-S. 2012b. Application of RNA silencing to plant disease resistance. Silence 3: 5.
- Dunoyer P, Brosnan CA, Schott G, Wang Y, Jay F, Alioua A, Himber C, Voinnet O. 2010. An endogenous, systemic RNAi pathway in plants. The EMBO journal 29: 1699–1712.
- Dunoyer P, Himber C, Voinnet O. 2005. DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. Nature genetics 37: 1356–1360.
- Dunoyer P, Himber C, Voinnet O. 2006. Induction, suppression and requirement of RNA silencing pathways in virulent Agrobacterium tumefaciens infections. Nature genetics 38: 258–263.
- Dunoyer P, Lecellier C-H, Parizotto EA, Himber C, Voinnet O. 2004. Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. The Plant cell 16: 1235–1250.
- Eamens AL, Kim KW, Curtin SJ, Waterhouse PM. 2012. DRB2 is required for microRNA biogenesis in Arabidopsis thaliana. PLoS ONE 7.
- Ebhardt HA, Fedynak A, Fahlman RP. 2010. Naturally occurring variations in sequence length creates microRNA isoforms that differ in argonaute effector complex specificity. Silence 1: 12.
- Echkardt H, Luhrmann R. 1979. Blocking of the initiation of protein biosynthesis by a pentanucleotide complementary to the 3’ end of Escherichia coli 16s rRNA. J Biol Chem 254: 11185-1188.
- Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL. 2003. Radial Patterning of Arabidopsis Shoots by Class III HD-ZIP and KANADI Genes. Current Biology 13: 1768–1774.

- Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, et al. 2007. High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of MIRNA genes. PLoS ONE 2.
- Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, Carrington JC. 2006. Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. Current Biology 16: 939–944.
- Fei Q, Xia R, Meyers BC. 2013. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. The Plant cell 25: 2400–15.
- Feng J, Lai L, Lin R, Jin C, Chen J. 2012. Differential effects of Cucumber mosaic virus satellite RNAs in the perturbation of microRNA-regulated gene expression in tomato. Molecular Biology Reports 39: 775–784.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391: 806–811.
- Folkes L, Moxon S, Woolfenden HC, Stocks MB, Szittya G, Dalmay T, Moulton V. 2012. PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. Nucleic Acids Research 40(13): e103.
- Frank F, Hauver J, Sonenberg N, Nagar B. 2012. Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. The EMBO Journal 31: 3588–3595.
- Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. 2008. Discovering microRNAs from deep sequencing data using miRDeep. Nature biotechnology 26: 407–415.
- Garcia D, Collier SA, Byrne ME, Martienssen RA. 2006. Specification of Leaf Polarity in Arabidopsis via the trans-Acting siRNA Pathway. Current Biology 16: 933–938.
- Garcia D, Garcia S, Pontier D, Marchais A, Renou JP, Lagrange T, Voinnet O. 2012. Ago Hook and RNA Helicase Motifs Underpin Dual Roles for SDE3 in

Antiviral Defense and Silencing of Nonconserved Intergenic Regions. Molecular Cell 48: 109–120.

- Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, Carrington JC. 2010. Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. The Plant cell 22: 481–496.
- Georgiev MI, Agostini E, Ludwig-Müller J, Xu J. 2012. Genetically transformed roots: From plant disease to biotechnological resource. Trends in Biotechnology 30: 528–537.
- German MA, Luo S, Schroth G, Meyers BC, Green PJ. 2009. Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. Nature Protocols 4:356–362.
- Guo S, Kemphues KJ. 1995. par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell 81: 611–620.
- Hackenberg M, Rodríguez-Ezpeleta N, Aransay AM. 2011. MiRanalyzer: An update on the detection and analysis of microRNAs in high-throughput sequencing experiments. Nucleic Acids Research 39.
- Haag JR, Pikaard CS. 2011. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol. 14: 483–492.
- Hamera S, Song X, Su L, Chen X, Fang R. 2012. Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. Plant Journal 69: 104–115.
- Hamilton AJ, Baulcombe DC. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science (New York, N.Y.) 286: 950–952.
- Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404: 293–296.
- Hansen TB, Kjems J, Bramsen JB. 2011. Enhancing miRNA annotation confidence in miRBase by continuous cross dataset analysis. RNA Biology 8: 378–383.

- Harland R, Weintraub H. 1985. Translation of mRNA injected into Xenopus oocytes is specifically inhibited by antisense RNA. The Journal of cell biology 101: 1094–1099.
- Harvey JJW, Lewsey MG, Patel K, Westwood J, Heimstädt S, Carr JP, Baulcombe DC. 2011. An antiviral defense role of AGO2 in plants. PLoS ONE 6.
- Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC. 2010. The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. The Plant cell 22: 321–334.
- Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE. 2006. Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nature genetics 38: 721–725.
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. 2005. RNA polymerase IV directs silencing of endogenous DNA. Science (New York, N.Y.) 308: 118–120.
- Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, Kasschau KD, Carrington JC. 2007. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. The Plant cell 19: 926–942.
- Hsieh L-C, Lin S-I, Shih AC-C, Chen J-W, Lin W-Y, Tseng C-Y, Li W-H, Chiou T-J. 2009. Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. Plant physiology 151: 2120–2132.
- Hunter C, Willmann MR, Wu G, Yoshikawa M, de la Luz Gutiérrez-Nava M, Poethig SR. 2006. Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. Development (Cambridge, England) 133: 2973–2981.
- Izant JG, Weintraub H. 1984. Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. Cell 36: 1007–1015.
- Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3: 318–356.

- Jagadeeswaran G, Saini A, Sunkar R. 2009. Biotic and abiotic stress down-regulate miR398 expression in Arabidopsis. Planta 229: 1009–1014.
- Jauvion V, Rivard M, Bouteiller N, Elmayan T, Vaucheret H. 2012. RDR2 partially antagonizes the production of RDR6-dependent siRNA in sense transgene-mediated PTGS. PLoS ONE 7.
- Jay F, Wang Y, Yu A, Taconnat L, Pelletier S, Colot V, Renou JP, Voinnet O. 2011. Misregulation of AUXIN RESPONSE FACTOR 8 underlies the developmental abnormalities caused by three distinct viral silencing suppressors in Arabidopsis. PLoS Pathogens 7.
- Jayaraman K, McParland K, Miller P, Ts'o PO. 1981. Selective inhibition of Escherichia coli protein synthesis and growth by nonionic oligonucleotides complementary to the 3' end of 16S rRNA. Proceedings of the National Academy of Sciences of the United States of America 78: 1537–1541.
- Jin H, Vacic V, Girke T, Lonardi S, Zhu J-K. 2008. Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. BMC molecular biology 9: 6.
- Jones-Rhoades MW, Bartel DP. 2004. Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. Molecular Cell 14: 787–799.
- Jones-Rhoades MW, Bartel DP, Bartel B. 2006. MicroRNAS and their regulatory roles in plants. Annual review of plant biology 57: 19–53.
- Joshi T, Yan Z, Libault M, Jeong D-H, Park S, Green PJ, Sherrier DJ, Farmer A, May G, Meyers BC, et al. 2010. Prediction of novel miRNAs and associated target genes in Glycine max. BMC bioinformatics 11 Suppl 1: S14.
- Kanno T, Habu Y. 2011. SiRNA-mediated chromatin maintenance and its function in Arabidopsis thaliana. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1809: 444–451.
- Kasschau KD, Carrington JC. 1998. A counterdefensive strategy of plant viruses: Suppression of posttranscriptional gene silencing. Cell. 95:461–470.
- Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biology 5: 0479–0493.

- Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC. 2003. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. Developmental Cell 4: 205–217.
- Katiyar-Agarwal S, Jin H. 2010. Role of small RNAs in host-microbe interactions. Annual review of phytopathology 48: 225–246.
- Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Zhu J-K, Staskawicz BJ, Jin H. 2006. A pathogen-inducible endogenous siRNA in plant immunity. Proceedings of the National Academy of Sciences of the United States of America 103: 18002–18007.
- Kawashima CG, Yoshimoto N, Maruyama-Nakashita A, Tsuchiya YN, Saito K, Takahashi H, Dalmay T. 2009. Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. The Plant journal : for cell and molecular biology 57: 313–321.
- Kidner CA, Martienssen RA. 2004. Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. Nature 428: 81-84.
- Kim VN. 2005. MicroRNA biogenesis: coordinated cropping and dicing. Nature reviews. Molecular cell biology 6: 376–385.
- Kim S, Yang JY, Xu J, Jang IC, Prigge MJ, Chua NH. 2008. Two cap-binding proteins CBP20 and CBP80 are involved in processing primary microRNAs. Plant and Cell Physiology 49: 1634–1644.
- Kozomara A, Griffiths-Jones S. 2013. MiRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Research 42: D68-73.
- Krüger J, Rehmsmeier M. 2006. RNAhybrid: MicroRNA target prediction easy, fast and flexible. Nucleic Acids Research 34.
- Van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR. 1990. Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. The Plant cell 2: 291–299.
- Kumagai H, Kouchi H. 2003. Gene silencing by expression of hairpin RNA in Lotus japonicus roots and root nodules. Molecular plant-microbe interactions : MPMI 16: 663–668.

- Kurihara Y, Watanabe Y. 2004. Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proceedings of the National Academy of Sciences of the United States of America 101: 12753–12758.
- Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, Kim VN. 2004. MicroRNA genes are transcribed by RNA polymerase II. The EMBO journal 23: 4051–4060.
- Li H, Deng Y, Wu T, Subramanian S, Yu O. 2010a. Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. Plant physiology 153: 1759–1770.
- Li J, Yang Z, Yu B, Liu J, Chen X. 2005. Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. Current Biology 15: 1501–1507.
- Li Y, Zhang Q, Zhang J, Wu L, Qi Y, Zhou J-M. 2010b. Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity. Plant physiology 152: 2222–2231.
- Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG. 1993. Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. The Plant cell 5: 1749–1759.
- Liu PP, Montgomery TA, Fahlgren N, Kasschau KD, Nonogaki H, Carrington JC. 2007. Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. Plant Journal 52: 133–146.
- Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J. 2006. SERRATE: a new player on the plant microRNA scene. EMBO reports 7: 1052–1058.
- Lu C, Kulkarni K, Souret FF, MuthuValliappan R, Tej SS, Poethig RS, Henderson IR, Jacobsen SE, Wang W, Green PJ, et al. 2006. MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. Genome research 16: 1276–1288.
- Mallory AC, Bartel DP, Bartel B. 2005. MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. The Plant cell 17: 1360–1375.

- Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, Bartel DP. 2004. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. The EMBO journal 23: 3356–3364.
- Mallory A, Vaucheret H. 2010. Form, function, and regulation of ARGONAUTE proteins. The Plant cell 22: 3879–3889.
- Mandadi KK, Scholthof KB. 2013. Plant immune responses against viruses: how does a virus cause disease? Plant Cell 25: 1489–1505.
- Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM. 2006. The evolution and diversification of Dicers in plants. FEBS Letters 580: 2442–2450.
- Martínez De Alba AE, Jauvion V, Mallory AC, Bouteiller N, Vaucheret H. 2011. The miRNA pathway limits AGO1 availability during siRNA-mediated PTGS defense against exogenous RNA. Nucleic Acids Research 39: 9339–9344.
- Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nature Review Genetics 15: 394-408.
- Megraw M, Baev V, Rusinov V, Jensen ST, Kalantidis K, Hatzigeorgiou AG. 2006. MicroRNA promoter element discovery in Arabidopsis. RNA (New York, N.Y.) 12: 1612–1619.
- Melton DA. 1985. Injected anti-sense RNAs specifically block messenger RNA translation in vivo. Proceedings of the National Academy of Sciences of the United States of America 82: 144–148.
- Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, et al. 2008. Criteria for annotation of plant MicroRNAs. Plant Cell 20: 3186–3190.
- Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, et al. 2008. Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5´ Terminal Nucleotide. Cell 133: 116–127.
- Millar AA, Gubler F. 2005. The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. The Plant cell 17: 705–721.

- Mizuno T, Chou MY, Inouye M. 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proceedings of the National Academy of Sciences of the United States of America 81: 1966–1970.
- Mlotshwa S, Schauer SE, Smith TH, Mallory AC, Herr JM, Roth B, Merchant DS, Ray A, Bowman LH, Vance VB. 2005. Ectopic DICER-LIKE1 expression in P1/HC-Pro Arabidopsis rescues phenotypic anomalies but not defects in microRNA and silencing pathways. The Plant cell 17: 2873–2885.
- Moissiard G, Parizotto EA, Himber C, Voinnet O. 2007. Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. RNA (New York, N.Y.) 13: 1268–1278.
- Moissiard G, Voinnet O. 2006. RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. Proceedings of the National Academy of Sciences of the United States of America 103: 19593–19598.
- Moldovan D, Spriggs A, Yang J, Pogson BJ, Dennis ES, Wilson IW. 2010. Hypoxia-responsive microRNAs and trans-acting small interfering RNAs in Arabidopsis. Journal of Experimental Botany 61: 165–177.
- Molnár A, Csorba T, Lakatos L, Várallyay E, Lacomme C, Burgyán J. 2005. Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. Journal of virology 79: 7812–7818.
- Montes RA, de Fátima Rosas-Cárdenas F, De Paoli E, Accerbi M, Rymarquis LA, Mahalingam G, Marsch-Martínez N, Meyers BC, Green PJ, de Folter S. 2014. Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. Nature Communications 5:3722.
- Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC. 2008. Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. Cell 133: 128–141.

- Morel J-B, Godon C, Mourrain P, Béclin C, Boutet S, Feuerbach F, Proux F, Vaucheret H. 2002. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. The Plant cell 14: 629–639.
- Morin RD, Connor MDO, Griffith M, Kuchenbauer F, Delaney A, Prabhu A, Zhao Y, Mcdonald H, Zeng T, Hirst M, et al. 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Research: 610–621.
- Mourrain P, Béclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, et al. 2000. Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell 101: 533–542.
- Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T. 2008. Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome research 18: 1602–1609.
- Nagano H, Fukudome A, Hiraguri A, Moriyama H, Fukuhara T. 2014. Distinct substrate specificities of Arabidopsis DCL3 and DCL4. Nucleic Acids Research 42: 1845–1856.
- 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. The Plant cell 2: 279–289.
- Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JDG. 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science (New York, N.Y.) 312: 436–439.
- Navarro L, Jay F, Nomura K, He SY, Voinnet O. 2008. Suppression of the microRNA pathway by bacterial effector proteins. Science (New York, N.Y.) 321: 964–967.
- Naya L, Paul S, Valdés-López O, Mendoza-Soto AB, Nova-Franco B, Sosa-Valencia G, Reyes JL, Hernández G. 2014. Regulation of copper homeostasis and biotic interactions by microRNA 398b in common bean. PLoS ONE 9.

- Ning Z, Cox AJ, Mullikin JC. 2001. SSAHA: a fast search method for large DNA databases. Genome research 11: 1725–1729.
- Olmedo-Monfil V, Durán-Figueroa N, Arteaga-Vázquez M, Demesa-Arévalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen RA, Vielle-Calzada J-P. 2010. Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature 464: 628–632.
- Omarov RT, Scholthof HB. 2012. Biological chemistry of virus-encoded suppressors of RNA silencing: an overview. Methods Molecular Biology. 894: 39–56.
- Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation- dependent heterochromatin formation. Cell 120: 613–622.
- Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D. 2003. Control of leaf morphogenesis by microRNAs. Nature 425: 257–263.
- Pant BD, Musialak-Lange M, Nuc P, May P, Buhtz A, Kehr J, Walther D, Scheible W-R. 2009. Identification of nutrient-responsive Arabidopsis and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing. Plant physiology 150: 1541–1555.
- Pantaleo V, Szittya G, Burgyán J. 2007. Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. Journal of virology 81: 3797–3806.
- Parizotto EA, Dunoyer P, Rahm N, Himber C, Voinnet O. 2004. In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA. Genes & development 18: 2237–2242.
- Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. 2005. Nuclear processing and export of microRNAs in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 102: 3691–3696.
- Peláez P, Sanchez F. 2013. Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences. Frontiers in plant science 4: 343.

- Peláez P, Trejo MS, Iniguez LP, Estrada-Navarrete G, Covarrubias AA, Reyes JL, Sanchez F. 2012. Identification and characterization of microRNAs in Phaseolus vulgaris by high-throughput sequencing. BMC Genomics 13: 83.
- Pélissier T, Clavel M, Chaparro C, Pouch-Pélissier M-N, Vaucheret H, Deragon J-M. 2011. Double-stranded RNA binding proteins DRB2 and DRB4 have an antagonistic impact on polymerase IV-dependent siRNA levels in Arabidopsis. RNA (New York, N.Y.) 17: 1502–1510.
- Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS. 2004. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes & development 18: 2368–2379.
- Pontes O, Li CF, Nunes PC, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS. 2006. The Arabidopsis Chromatin-Modifying Nuclear siRNA Pathway Involves a Nucleolar RNA Processing Center. Cell 126: 79–92.
- Pontier D, Picart C, Roudier F, Garcia D, Lahmy S, Azevedo J, Alart E, Laudié M, Karlowski WM, Cooke R, et al. 2012. NERD, a Plant-Specific GW Protein, Defines an Additional RNAi-Dependent Chromatin-Based Pathway in Arabidopsis. Molecular Cell 48: 121–132.
- Pruss GJ, Nester EW, Vance V. 2008. Infiltration with Agrobacterium tumefaciens induces host defense and development-dependent responses in the infiltrated zone. Molecular plant-microbe interactions : MPMI 21: 1528–1538.
- Pumplin N, Voinnet O. 2013. RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nature reviews. Microbiology 11: 745–60.
- Qi X, Bao FS, Xie Z. 2009. Small RNA deep sequencing reveals role for Arabidopsis thaliana RNA-dependent RNA polymerases in viral siRNA biogenesis. PLoS ONE 4.
- Qu F, Ye X, Morris TJ. 2008. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proceedings of the National Academy of Sciences of the United States of America 105: 14732–14737.

- Raja P, Sanville BC, Buchmann RC, Bisaro DM. 2008. Viral genome methylation as an epigenetic defense against geminiviruses. Journal of virology 82: 8997–9007.
- Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. 2006. A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. Genes & development 20: 3407–3425.
- Rajeswaran R, Aregger M, Zvereva AS, Borah BK, Gubaeva EG, Pooggin MM. 2012. Sequencing of RDR6-dependent double-stranded RNAs reveals novel features of plant siRNA biogenesis. Nucleic Acids Research 40: 6241–6254.
- Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. 2002. MicroRNAs in plants. Genes & development 16: 1616–1626.
- Reyes JL, Chua NH. 2007. ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. Plant Journal 49: 592–606.
- Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. 2010. Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis. Genes & development 24: 1010–1021.
- Rosenberg UB, Preiss A, Seifert E, Jackle H, Knipple DC. 1985. Production of phenocopies by Kruppel antisense RNA injection into Drosophila embryos. Nature 313: 703-704.
- Sarazin A, Voinnet O. 2014. Exploring new models of easiRNA biogenesis. Nature Genetics 46: 530-1.
- Saze H, Tsugane K, Kanno T, Nishimura T. 2012. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. Plant Cell Physiol 53: 766–784.
- Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D. 2005. Specific effects of microRNAs on the plant transcriptome. Developmental Cell 8: 517–527.
- Schmülling T, Schell J, Spena A. 1988. Single genes from Agrobacterium rhizogenes influence plant development. Embo Journal 7: 2621-2629.
- Shimura H, Pantaleo V, Ishihara T, Myojo N, Inaba J ichi, Sueda K, Burgyán J, Masuta C. 2011. A Viral Satellite RNA Induces Yellow Symptoms on Tobacco By

Targeting a Gene Involved in Chlorophyll Biosynthesis Using the RNA Silencing Machinery. PLoS Pathogens 7.

- Shivaprasad P V., Chen H-M, Patel K, Bond DM, Santos BACM, Baulcombe DC. 2012. A MicroRNA Superfamily Regulates Nucleotide Binding Site-Leucine-Rich Repeats and Other mRNAs. The Plant Cell 24: 859–874.
- Si-Ammour A, Windels D, Arn-Bouldoires E, Kutter C, Ailhas J, Meins F, Vazquez F. 2011. miR393 and Secondary siRNAs Regulate Expression of the TIR1/AFB2 Auxin Receptor Clade and Auxin-Related Development of Arabidopsis Leaves. PLANT PHYSIOLOGY 157: 683–691.
- Simon SA, Meyers BC. 2011. Small RNA-mediated epigenetic modifications in plants. Current Opinion in Plant Biology 14: 148–155.
- Simons RW, Kleckner N. 1983. Translational control of IS10 transposition. Cell 34: 683–691.
- Slotkin RK, Slotkin RK, Vaughn M, Vaughn M, Borges F, Borges F, Tanurdzić M, Tanurdzić M, Becker JD, Becker JD, et al. 2009. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell 136: 461–72.
- Smith NA, Eamens AL, Wang MB. 2011. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. PLoS Pathogens 7.
- Smith LM, Pontes O, Searle I, Yelina N, Yousafzai FK, Herr AJ, Pikaard CS, Baulcombe DC. 2007. An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in Arabidopsis. The Plant cell 19: 1507–1521.
- Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu J-K, Yu O. 2008. Novel and nodulation-regulated microRNAs in soybean roots. BMC genomics 9: 160.
- Sun G. 2012. MicroRNAs and their diverse functions in plants. Plant Molecular Biology 80: 17–36.
- Sunkar R, Jagadeeswaran G. 2008. In silico identification of conserved microRNAs in large number of diverse plant species. BMC plant biology 8: 37.
- Szittya G, Moxon S, Pantaleo V, Toth G, Rusholme Pilcher RL, Moulton V, Burgyan J, Dalmay T. 2010. Structural and functional analysis of viral siRNAs. PLoS pathogens 6: e1000838.

- Takeda A, Tsukuda M, Mizumoto H, Okamoto K, Kaido M, Mise K, Okuno T. 2005. A plant RNA virus suppresses RNA silencing through viral RNA replication. The EMBO journal 24: 3147–3157.
- Taniguchi T, Weissmann C. 1978. Inhibition of Qβ RNA 70s ribosome initiation complex formation by an oligonucleotide complementary to the 3' terminal region of E. coli 16s ribosomal RNA. Nature 275: 770-772.
- Tanurdzic M, Vaughn MW, Jiang H, Lee T-J, Slotkin RK, Sosinski B, Thompson WF, Doerge RW, Martienssen RA. 2008. Epigenomic consequences of immortalized plant cell suspension culture. PLoS Biology 6: 2880–2895.
- Tarutani Y, Shiba H, Iwano M, Kakizaki T, Suzuki G, Watanabe M, Isogai A, Takayama S. 2010. Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility. Nature 466: 983–986.
- Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, Ciaudo C, Cruaud C, Poulain J, Berdasco M, Fraga MF, et al. 2009. A role for RNAi in the selective correction of DNA methylation defects. Science (New York, N.Y.) 323: 1600–1604.
- Tepfer D. 1984. Transformation of several species of higher plants by Agrobacterium rhizogenes: sexual transmission of the transformed genotype and phenotype. Cell 37: 959–967.
- Tomizawa JI, Itoh T. 1982. The importance of RNA secondary structure in CoIE1 primer formation. Cell 31: 575–583.
- Tomizawa J, Itoh T, Selzer G, Som T. 1981. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. Proceedings of the National Academy of Sciences of the United States of America 78: 1421–1425.
- Valdés-López O, Arenas-Huertero C, Ramírez M, Girard L, Sánchez F, Vance CP, Luis Reyes J, Hernández G. 2008. Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean roots. Plant, Cell and Environment 31: 1834–1843.
- Valdés-López O, Yang SS, Aparicio-Fabre R, Graham PH, Reyes JL, Vance CP, Hernández G. 2010. MicroRNA expression profile in common bean (Phaseolus

vulgaris) under nutrient deficiency stresses and manganese toxicity. New Phytologist 187: 805–818.

- Válóczi A, Várallyay É, Kauppinen S, Burgyán J, Havelda Z. 2006. Spatio-temporal accumulation of microRNAs is highly coordinated in developing plant tissues. Plant Journal 47: 140–151.
- Várallyay E, Válóczi A, Agyi A, Burgyán J, Havelda Z. 2010. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. The EMBO journal 29: 3507–3519.
- Vaucheret H. 2009. AGO1 homeostasis involves differential production of 21-nt and 22-nt miR168 species by MIR168a and MIR168b. PLoS ONE 4.
- Vaucheret H, Vazquez F, Crété P, Bartel DP. 2004. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes & development 18: 1187–1197.
- Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gasciolli V, Mallory AC, Hilbert JL, Bartel DP, Crété P. 2004. Endogenous trans-acting siRNAs regulate the accumulation of arabidopsis mRNAs. Molecular Cell 16: 69–79.
- Vogler H, Akbergenov R, Shivaprasad P V, Dang V, Fasler M, Kwon M-O, Zhanybekova S, Hohn T, Heinlein M. 2007. Modification of small RNAs associated with suppression of RNA silencing by tobamovirus replicase protein. Journal of virology 81: 10379–10388.
- Voinnet O, Baulcombe DC. 1997. Systemic signaling in gene silencing. Nature 389, 553.
- Voinnet O, Pinto YM, Baulcombe DC. 1999. Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. Proceeding of the National Academy of Sciences of the United States of America 96: 14147-14152.
- Voinnet O, Rivas S, Mestre P, Baulcombe D. 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant Journal 33: 949–956.

- Voinnet O, Vain P, Angell S, Baulcombe DC. 1998. Systemic spread of sequence-specific transgene RNA degradation is initiated by localised introduction of ectopic promoterless DNA. Cell 95: 177–187.
- Wang X-B, Jovel J, Udomporn P, Wang Y, Wu Q, Li W-X, Gasciolli V, Vaucheret H, Ding S-W. 2011a. The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative argonautes in Arabidopsis thaliana. The Plant cell 23: 1625–1638.
- Wang L, Mai YX, Zhang YC, Luo Q, Yang HQ. 2010a. MicroRNA171c-targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in arabidopsis. Molecular Plant 3: 794–806.
- Wang H-W, Noland C, Siridechadilok B, Taylor DW, Ma E, Felderer K, Doudna JA, Nogales E. 2009. Structural insights into RNA processing by the human RISC-loading complex. Nature structural & molecular biology 16: 1148–1153.
- Wang X-J, Reyes JL, Chua N-H, Gaasterland T. 2004. Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. Genome biology 5: R65.
- Wang J-W, Schwab R, Czech B, Mica E, Weigel D. 2008. Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in Arabidopsis thaliana. The Plant cell 20: 1231–1243.
- Wang M, Soyano T, Machida S, Yang J-Y, Jung C, Chua N-H, Yuan YA. 2011b. Molecular insights into plant cell proliferation disturbance by Agrobacterium protein 6b. Genes & development 25: 64–76.
- Wang J-W, Wang L-J, Mao Y-B, Cai W-J, Xue H-W, Chen X-Y. 2005. Control of root cap formation by MicroRNA-targeted auxin response factors in Arabidopsis. The Plant cell 17: 2204–2216.
- Wang X-B, Wu Q, Ito T, Cillo F, Li W-X, Chen X, Yu J-L, Ding S-W. 2010b. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 107: 484–489.

- Weiberg A, Wang M, Lin F-M, Zhao H, Zhang Z, Kaloshian I, Huang H-D, Jin H. 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. Science (New York, N.Y.) 342: 118–23.
- Willmann MR, Endres MW, Cook RT, Gregory BD. 2011. The functions of RNA-dependent RNA Polymerases in Arabidopsis. The Arabidopsis Book 9: e0146.
- Wu G, Poethig RS. 2006. Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. Development (Cambridge, England) 133: 3539–3547.
- Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. 2004. Genetic and functional diversification of small RNA pathways in plants. PLoS Biology 2.
- Xie Z, Kasschau KD, Carrington JC. 2003. Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. Current Biology 13: 784–789.
- Yifhar T, Pekker I, Peled D, Friedlander G, Pistunov A, Sabban M, Wachsman G, Alvarez JP, 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: 3575–3589.
- Yu B, Bi L, Zhai J, Agarwal M, Li S, Wu Q, Ding SW, Meyers BC, Vaucheret H, Chen X. 2010. siRNAs compete with miRNAs for methylation by HEN1 in Arabidopsis. Nucleic Acids Research 38: 5844–5850.
- Yu B, Bi L, Zheng B, Ji L, Chevalier D, Agarwal M, Ramachandran V, Li W, Lagrange T, Walker JC, et al. 2008. The FHA domain proteins DAWDLE in Arabidopsis and SNIP1 in humans act in small RNA biogenesis. Proceedings of the National Academy of Sciences of the United States of America 105: 10073–10078.
- Yu D, Fan B, MacFarlane SA, Chen Z. 2003. Analysis of the involvement of an inducible Arabidopsis RNA-dependent RNA polymerase in antiviral defense. Molecular plant-microbe interactions : MPMI 16: 206–216.
- Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X. 2005. Methylation as a crucial step in plant microRNA biogenesis. Science (New York, N.Y.) 307: 932–935.

- Zamecnik PC, Stephenson ML. 1978. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proceedings of the National Academy of Sciences of the United States of America 75: 280–284.
- Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA, et al. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. Genes & Development 25: 2540–2553.
- Zhai J, Zhao Y, Simon SA, Huang S, Petsch K, et al. 2013. Plant microRNAs display differential 3' truncation and tailing modifications that are ARGONAUTE1 dependent and conserved across species. Plant Cell 25: 2417–2428.
- Zhan X, Wang B, Li H, Liu R, Kalia RK, Zhu J-K, Chinnusamy V. 2012. Arabidopsis proline-rich protein important for development and abiotic stress tolerance is involved in microRNA biogenesis. Proceedings of the National Academy of Sciences 109: 18198–18203.
- Zhang W, Gao S, Zhou X, Chellappan P, Chen Z, Zhou X, Zhang X, Fromuth N, Coutino G, Coffey M, et al. 2011a. Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. Plant Molecular Biology 75: 93–105.
- Zhang X, Henderson IR, Lu C, Green PJ, Jacobsen SE. 2007. Role of RNA polymerase IV in plant small RNA metabolism. Proceedings of the National Academy of Sciences of the United States of America 104: 4536–4541.
- Zhang X, Xia J, Lii Y, Barrera-Figueroa B, Zhou X, Gao S, Lu L, Niu D, Liang W, Chen Z, et al. 2012a. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. Genome Biology 13: R20.
- Zhang X, Yuan Y-R, Pei Y, Lin S-S, Tuschl T, Patel DJ, Chua N-H. 2006. Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. Genes & development 20: 3255–3268.
- Zhang X, Zhang X, Singh J, Li D, Qu F. 2012b. Temperature-Dependent Survival of Turnip Crinkle Virus-Infected Arabidopsis Plants Relies on an RNA Silencing-Based Defense That Requires DCL2, AGO2, and HEN1. Journal of Virology 86: 6847–6854.

- Zhang X, Zhao H, Gao S, Wang WC, Katiyar-Agarwal S, Huang H Da, Raikhel N, Jin H. 2011b. Arabidopsis Argonaute 2 Regulates Innate Immunity via miRNA393*-Mediated Silencing of a Golgi-Localized SNARE Gene, MEMB12. Molecular Cell 42: 356–366.
- Zheng B, Wang Z, Li S, Yu B, Liu J-Y, Chen X. 2009. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. Genes & development 23: 2850–2860.
- Zhu H, Duan C-G, Hou W-N, Du Q-S, Lv D-Q, Fang R-X, Guo H-S. 2011a. Satellite RNA-Derived Small Interfering RNA satsiR-12 Targeting the 3' Untranslated Region of Cucumber Mosaic Virus Triggers Viral RNAs for Degradation. Journal of Virology 85: 13384–13397.
- Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, Barefoot A, Dickman M, Zhang X. 2011b. Arabidopsis argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell 145: 242–256.
- 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: 29–39.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31: 3406–3415.

8. Abreviaturas y siglas

AGOs	Argonaute proteins	Proteinas argonauta
ArT-sRNAs	*A. Rhizogenes* T-DNA sRNAs	sRNAs del T-DNA de *A. rhizogenes*
BSRs	Bacterial suppresors of RNA silencing	Supresores bacterianos del silenciamiento por RNA
DCLs	Dicer-like proteins	Proteinas Dicer-like
DNA	Deoxyribonucleic acid	Ácido desoxirribonucleico
dsRNA	double stranded RNA	RNA de doble cadena
EST	Expressed Sequence Tag	Marcador de secuencia expresada

GSS	Genomic Survey Sequences	Secuencias reconocidas de genoma
Hc-siRNAs	Heterochromatic small interfering RNAs	RNAs pequeños de interferencia heterocromáticos
MFEs	Minimum free energies	energías libres mínimas
miRNAs	MicroRNAs	MicroRNAs
nat-siRNAs	Natural antisense transcript-derived small interfering RNAs	RNAs pequeños de interferencia generados por transcritos antisentido naturales
PARE	Parallel analysis of RNA ends	Analisis paralelo de extremos de RNA
PDR	Pathogen-derived resistance	Resistencia derivada de patógenos
phasiRNAs	Phased, secondary, small interfering RNAs	RNAs pequeños interferentes secundarios en fase
PR	Pathogen-related	Relcionados a patógenos
Pre-miRNAs	MiRNA precursor	Precursor de miRNA
Pri-miRNAs	Primary miRNA transcripts	Transcritos de miRNAs primarios
PTGS	Post- transcriptional gene silencing	Silenciamiento génico post-transcripcional
R	Resistance	Resistencia
RdRM	RNA-directed DNA methylation	Metilación de DNA dirigida por RNA
RDRs	RNA-dependent RNA polymerases	RNA polimerasas dependendientes de RNA
Ri	Root-inducing	Inductor de raíces
RNA	Ribonucleic acid	Ácido ribonucleico
siRNA	Small interfering RNA	ARN pequeño de interferencia
siTAARs	TAAR transcript- derived small interfering RNAs	RNAs pequeños de interferencia generados de transcritos TAAR
snRNA	Small nuclear RNA	RNA pequeño nuclear

sRNAs	Small RNAs	RNAs pequeños
Ta-siRNAs	trans-acting small interfering RNAs	RNAs pequeños interferentes que actúan en trans
T-DNA	Transferred DNA	DNA transferido
TGS	Transcriptional gene silencing	Silenciamiento génico transcripcional
Ti	Tumor-inducing	Inductor de raíces
TPM	Transcripts per million	Transcritos por millon
vsiRNAs	Virus-derived small interfering RNAs	RNAs pequeños interferentes generados de virus
VSRs	Viral suppressors of RNA silencing	Supresores virales del silenciamiento por RNA

9. Anexos

9.1 Peláez et al., 2012

BMC Genomics

RESEARCH ARTICLE Open Access

Identification and characterization of microRNAs in *Phaseolus vulgaris* by high-throughput sequencing

Pablo Peláez, Minerva S Trejo, Luis P Iñiguez, Georgina Estrada-Navarrete, Alejandra A Covarrubias, José L Reyes and Federico Sanchez*

Abstract

Background: MicroRNAs (miRNAs) are endogenously encoded small RNAs that post-transcriptionally regulate gene expression. MiRNAs play essential roles in almost all plant biological processes. Currently, few miRNAs have been identified in the model food legume *Phaseolus vulgaris* (common bean). Recent advances in next generation sequencing technologies have allowed the identification of conserved and novel miRNAs in many plant species. Here, we used Illumina's sequencing by synthesis (SBS) technology to identify and characterize the miRNA population of *Phaseolus vulgaris*.

Results: Small RNA libraries were generated from roots, flowers, leaves, and seedlings of *P. vulgaris*. Based on similarity to previously reported plant miRNAs,114 miRNAs belonging to 33 conserved miRNA families were identified. Stem-loop precursors and target gene sequences for several conserved common bean miRNAs were determined from publicly available databases. Less conserved miRNA families and species-specific common bean miRNA isoforms were also characterized. Moreover, novel miRNAs based on the small RNAs were found and their potential precursors were predicted. In addition, new target candidates for novel and conserved miRNAs were proposed. Finally, we studied organ-specific miRNA family expression levels through miRNA read frequencies.

Conclusions: This work represents the first massive-scale RNA sequencing study performed in *Phaseolus vulgaris* to identify and characterize its miRNA population. It significantly increases the number of miRNAs, precursors, and targets identified in this agronomically important species. The miRNA expression analysis provides a foundation for understanding common bean miRNA organ-specific expression patterns. The present study offers an expanded picture of *P. vulgaris* miRNAs in relation to those of other legumes.

Background

Small non-coding RNAs (sncRNAs) have been recognized as an important class of gene expression regulators. MicroRNAs (miRNAs) are ~21 nucleotide (nt) sncRNAs that regulate a multitude of biological processes in plants, including development, metabolism, stress responses, defense against pathogens, and maintenance of genome integrity [1]. MiRNAs direct cleavage or translational inhibition of a specific messenger RNA (mRNA) based on base-pairing complementarity between the miRNA and the target mRNA. MiRNAs are derived from imperfectly matched stem-loop structures that are formed from single-stranded primary miRNA transcripts (pri-miRNAs). MiRNA genes are RNA polymerase II transcription units that give rise to pri-miRNAs [2]. Pri-miRNAs are preferentially processed in the nucleus by an RNaseIII-type enzyme DICER-LIKE1 (DCL1) to release the precursor miRNAs (pre-miRNAs) [3]. Most plant pre-miRNAs produce a single miRNA/miRNA* duplex; the exceptions are some miR159 and miR319 loci [4]. These small RNA duplexes are subsequently 2'-*O*-methylated by the nuclear HUA ENHANCER 1 (HEN1) protein, preventing miRNA turnover, and are exported to the cytoplasm by the plant homolog of exportin-5 HASTY (HST) [5-7]. Finally, one of the strands of each duplex is incorporated into the RNA-

* Correspondence: federico@ibt.unam.mx
Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Induced Silencing Complex (RISC) to guide gene silencing [8]. RISCs contain a member of the ARGONAUTE (AGO) protein family to direct the endonucleolytic cleavage of target RNAs. AGO1 is the predominant carrier of plant miRNAs [9]. Thus, miRNAs are used by RISCs as templates for recognizing complementary mRNA to regulate gene expression.

Since the first plant miRNAs were reported in *Arabidopsis thaliana* in 2002, there has been an exponential growth of identified miRNAs in a diverse number of plant species [10]. The miRNA database miRBase (release 16, Sept 2010) contains 15,172 microRNA loci from 142 species. Plant miRNA loci belonging to 43 species correspond to 22.4% of the total mature miRNA entries. The plant species with the most miRNA loci identified are *Oryza sativa*, *Medicago truncatula*, *Arabidopsis thaliana*, *Physcomitrella patens*, *Glycine max*, *Sorghum bicolor*, *Populus trichocarpa*, and *Zea mays*. These 8 widely studied plant model species contribute 65.8% of the plant mature miRNA entries. On the other hand, 15 plant species, including *Phaseolus vulgaris*, have fewer than 10 reported miRNA loci in miRBase.

Identification of *P. vulgaris* microRNAs was first performed using an *in silico* approach [11]. The eight miRNA loci reported for *P. vulgaris* in miRBase were identified by Arenas-Huertero et al. (2009) in different organs and growth conditions, although the first mature miRNA sequence ever cloned and characterized for *P. vulgaris* was pvu-miR399a [12]. This group of eight miRNAs in miRBase comprises miR2118, miR159a, miR1514a, miR482, miR2119, miR166a, miR319c, and miR399a. Besides these miRNAs and their stem-loop precursors, another 19 mature miRNA sequences have been identified [13]. In addition, Valdés-López et al. (2010) analyzed the expression of 68 miRNAs under several abiotic stress conditions in leaves, roots and nodules of *P. vulgaris* based on a hybridization approach using miRNA macroarrays. The macroarrays contained probes for 9 miRNAs previously reported for common bean, 24 conserved miRNAs also identified in other legumes and 35 miRNAs found in soybean [11,13,14].

The first studies to identify plant miRNAs employed the traditional Sanger sequencing method. This was also the case for the study performed by Arenas-Huertero et al. (2009) to identify miRNA sequences in *Phaseolus vulgaris*. Despite the Sanger sequencing method's usefulness and importance in scientific research, it has several limitations with regard to miRNA identification [15,16]. For example, low-abundance miRNAs, such as many non-conserved miRNAs, are inefficiently detected by this small-scale sequencing method. Introduction of high-throughput next-generation sequencing (NGS) technologies has increased the number of miRNAs identified by allowing small RNA population analysis on a massive scale (RNA-seq). Deep-sequencing technologies have identified a larger number of novel miRNAs due to their ability to generate millions of reads with a determined length [17]. Moreover, the population of short sequence reads produced allows the identification of differential processing of stem-loop precursors by DICER-LIKE1, and mature miRNA isoforms within and between species. Likewise, relative abundance associated with small RNAs estimated by these technologies has permitted a higher level of confidence for miRNA annotation.

Legumes such as common bean are valuable crops for worldwide consumption because they are rich in protein and constitute a high calorie food source. Involvement of miRNAs in biological processes like nutrient balance, development, reproduction and plant-microbe interaction, makes its research crucial for improvement of staple crops. Up to now, identification and characterization of *Phaseolus vulgaris* miRNAs have been very limited. In this study, Illumina's sequencing by synthesis (SBS) technology was used to examine conserved, novel and species-specific common bean miRNAs comprehensively based on small RNA libraries generated from leaves, roots, seedlings and flowers. A detailed analysis of mature, mature-star (the complementary strand of the mature miRNA), isoforms and novel common bean miRNAs found in different organs is reported. Stem-loop precursors and target sequences for several common bean miRNAs were also determined from public databases. Overall, this work serves to extend our knowledge of the common bean miRNA population and to spotlight new miRNA variants found in different organs. It facilitates comparisons between common bean miRNAs and those found in other legumes and plants.

Results

Small RNA sequencing analysis

In order to identify novel and conserved miRNAs in common bean, we generated four small RNA libraries from leaves (LL), roots (RL), seedlings (SL) and flowers (FL) using the Genome Analyzer II and the Illumina Cluster Station (Illumina Inc, USA). Small RNA sequencing yielded more than eighty million raw sequence reads (Table 1). After removing low-quality sequences, adapters, and small sequences (< 16 nt long), 79.5% of the raw reads were left. These remaining sequences represent 3,372,753 (LL), 4,187,414 (RL), 4,015,702 (SL), and 3,453,543 (FL) unique sequence tags. The small RNA length distribution (16-30 nt) of each library showed that the most abundant and diverse species were those 21-24 nt in length, a typical size range for Dicer-derived products (Figure 1). The 24 nt class was the most diverse in all four libraries, followed by the 21 nt class, which was the second most abundant class in

Table 1 Summary of small RNA sequencing data analysis

	Leaves	Roots	Seedlings	Flowers
Total raw reads	16869046	20464127	17188077	27937376
High-quality reads[a]	11797480	16722005	16846110	20238107
Unique sequence tags	3372753	4187414	4015702	3453543
Total Rfam matching sequences	757788	5653553	5307793	9726833
Unique Rfam matching sequences	51227	60121	224240	96726
Perfect miRNA matching sequences[b]	265161	307003	582159	1415225
Total unfiltered isoform sequences	798835	287027	467691	228556
Total miRNA isoform sequences	744348	200630	335274	60167

[a]Sequences after quality, adapter and size filters. [b]Total sequences from conserved miRNAs.

the LL library (20%), and the third most abundant class in the SL (13.4%), RL (8%), and FL (10.8%) libraries. Filtering sequences for removal of non-coding RNAs such as tRNA, rRNA, snoRNA or snRNA was performed using the RNA families database Rfam because of the limited availability of *P. vulgaris* genome sequences. After removing miRNA sequences from Rfam, the remainder of the database (Rfam 10.0) was used to eliminate fragments corresponding to non-coding RNAs. Alignment against the four small RNA libraries was carried out using BLASTN and only perfect matches were removed from the libraries (Table 1).

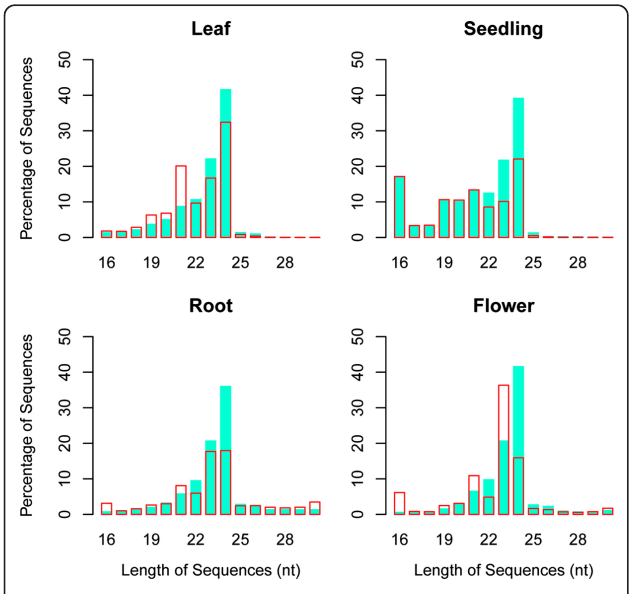


Figure 1 Sequence length distribution of *P. vulgaris* sRNAs. Size distribution in small RNA libraries. Average percentage (Y-axis) of unique (filled blue bars) and redundant (non-filled red bars) sequences of 16-30 nt length (X-axis) for each of the four sequenced libraries.

Identification of conserved miRNAs

To identify conserved miRNAs in the common bean small RNA libraries, unique mature plant miRNA sequences were extracted from miRBase (Release 16). BLASTN and SSAHA2 (Sequence Search and Alignment by Hashing Algorithm) tools were used separately for alignment of small RNA datasets against mature miRNA sequences in search of previously reported miRNAs with exactly the same size and nucleotide composition [18,19]. Both alignment tools found the same number of hits. MiRNAs that were detected in just one library or that totalled fewer than 15 absolute appearances in all libraries combined were removed (Additional file 1).

In total, 109 conserved miRNAs (Table 2) belonging to 29 families were identified in common bean. Fourteen of the conserved miRNAs were mature-star miRNAs (Table 3). Ten were detected in just two libraries, twenty five in three libraries, and seventy four in all four libraries. Twenty five of the families detected are highly conserved among plant species (Figure 2). Nine conserved families identified in common bean, including miR156, miR159, miR319, miR160, miR166, miR171, miR408, miR390 and miR395, were present in the common ancestor of all embryophytes [20]. Additionally, the miR396 family identified in all tracheophytes, and a pair of miRNA families detected in seed plants, miR397 and miR398, were present in *P. vulgaris*. Another nine identified are present in all angiosperm lineages: miR162, miR164, miR167, miR168, miR169, miR172, miR393, miR394, and miR399. The miR2111 family, which has been identified only in eudicot species, and the miR2118 family which has been found in monocots and in *Fabaceae*, were also identified in common bean. Finally, among the deeply conserved miRNA families detected, miR157, closely related to miR156, is present in *Brassicaceae*, *Solanaceae, Malvaceae* and *Fabaceae*.

Four other, less conserved miRNA families that were identified have been reported mainly in legumes, specifically in *Glycine max*. This is the case for miR1511, which had been previously detected only in soybean, although recent *P. vulgaris* miRNA studies pointed out its presence in common bean [13,14,21,22]. Here, it was confirmed that soybean and common bean share exactly the same sequence for this miRNA and that it is one of the most abundant miRNAs in these legumes (Table 2). The miR1514 and miR2119 families have been reported exclusively in legumes such as *Glycine max*, *Medicago truncatula* and *Phaseolus vulgaris*. Surprisingly, miR1515 was reported only in two species: *G. max* and *Citrus Sinensis*.

The *P. vulgaris* miRNAs with the most reads in all four libraries (ath-miR159a, mtr-miR319, ath-miR156a, ath-miR166a, zma-miR166c*, ath-miR396a, osa-miR167d, pvu-miR2118 and ath-miR167d) generally

Table 2 Conserved miRNAs from common bean

miRNA family	Reference miRNA	Sequence (5'- 3')	Length (nt)	Reads				
				LL	FL	RL	SL	Total
156	ahy-miR156a	UGACAGAAGAGAGAGAGCAC	20	56	15	8	41	120
	ahy-miR156c	UUGACAGAAGAGAGAGAGCAC	21	94	3512	3	317	3926
	ath-miR156a	UGACAGAAGAGAGUGAGCAC	20	82529	63	3705	14478	100775
	ath-miR156g	CGACAGAAGAGAGUGAGCAC	20	74	0	3	10	87
	bna-miR156a	UGACAGAAGAGAGUGAGCACA	21	112	21	2	26	161
	ptc-miR156k	UGACAGAAGAGAGGGAGCAC	20	47	0	6	2	55
	vvi-miR156e	UGACAGAGGAGAGUGAGCAC	20	30	0	1	1	32
	zma-miR156k	UGACAGAAGAGAGCGAGCAC	20	183	0	14	10	207
157	ath-miR157a	UUGACAGAAGAUAGAGAGCAC	21	3114	2787	4370	4839	15110
	ath-miR157d	UGACAGAAGAUAGAGAGCAC	20	361	30	26	89	506
159	aqc-miR159	UUUGGACUGAAGGGAGCUCUA	21	33	312	43	89	477
	pvu-miR159a.1	UUUGGAUUGAAGGGAGCUCUA	21	67007	1082688	180739	453692	1784126
	ath-miR159b	UUUGGAUUGAAGGGAGCUCUU	21	3	438	135	178	754
	gma-miR159c	AUUGGAGUGAAGGGAGCUCCG	21	0	0	1044	25	1069
	gma-miR159d	AGCUGCUUAGCUAUGGAUCCC	21	78	148	202	240	668
	osa-miR159a.1	UUUGGAUUGAAGGGAGCUCUG	21	37	224	31	148	440
	osa-miR159f	CUUGGAUUGAAGGGAGCUCUA	21	57	666	298	182	1203
	pta-miR159a	UUGGAUUGAAGGGAGCUCCA	20	0	3	0	25	28
160	ath-miR160a	UGCCUGGCUCCCUGUAUGCCA	21	136	5160	5474	104	10874
	bdi-miR160	UGCCUGGCUCCCUGUAUGCC	20	9	212	225	5	451
	osa-miR160f	UGCCUGGCUCCCUGAAUGCCA	21	9	3	240	2	254
162	ath-miR162a	UCGAUAAACCUCUGCAUCCAG	21	597	583	517	1192	2889
	zma-miR162	UCGAUAAACCUCUGCAUCCA	20	14	22	15	30	81
164	ath-miR164a	UGGAGAAGCAGGGCACGUGCA	21	7245	5621	3628	1229	17723
	ath-miR164c	UGGAGAAGCAGGGCACGUGCG	21	49	71	91	122	333
166	gma-miR166a	UCGGACCAGGCUUCAUUCCCC	21	5333	53276	20481	15163	94253
	ctr-miR166	UCGGACCAGGCUUCAUUCCCCC	22	2	18	10	10	40
	osa-miR166e	UCGAACCAGGCUUCAUUCCCC	21	0	17	8	4	29
	vvi-miR166a	UCGGACCAGGCUUCAUUCC	19	93	205	166	382	846
	zma-miR166h	UCGGACCAGGCUUCAUUCCC	20	140	769	352	630	1891
167	ath-miR167a	UGAAGCUGCCAGCAUGAUCUA	21	661	2883	3251	90	6885
	ath-miR167d	UGAAGCUGCCAGCAUGAUCUGG	22	22003	1443	5196	626	29268
	bna-miR167a	UGAAGCUGCCAGCAUGAUCUAA	22	2	0	17	0	19
	ccl-miR167a	UGAAGCUGCCAGCAUGAUCUGA	22	2267	8164	2727	276	13434
	osa-miR167d	UGAAGCUGCCAGCAUGAUCUG	21	15021	20039	6209	378	41647
	ptc-miR167f	UGAAGCUGCCAGCAUGAUCUU	21	36	252	146	5	439
168	ath-miR168a	UCGCUUGGUGCAGGUCGGGAA	21	382	991	605	43	2021
169	ath-miR169a	CAGCCAAGGAUGACUUGCCGA	21	0	11	0	44	55
	ath-miR169b	CAGCCAAGGAUGACUUGCCGG	21	1339	8784	154	6351	16628
	gma-miR169d	UGAGCCAAGGAUGACUUGCCGGU	23	0	4	11	0	15
	gma-miR169e	AGCCAAGGAUGACUUGCCGG	20	133	47	62	316	558
	mtr-miR169c	CAGCCAAGGGUGAUUUGCCGG	21	3916	166	78	2	4162
	mtr-miR169d	AAGCCAAGGAUGACUUGCCGG	21	126	863	389	4	1382
	osa-miR169e	UAGCCAAGGAUGACUUGCCGG	21	1	29	1	16	47
171	ath-miR171b	UUGAGCCGUGCCAAUAUCACG	21	119	790	0	37	946
	sly-miR171d	UUGAGCCGCGCCAAUAUCAC	20	2	1	67	7	77
	zma-miR171b	UUGAGCCGUGCCAAUAUCAC	20	28	40	13	2	83

Table 2 Conserved miRNAs from common bean *(Continued)*

	zma-miR171f	UUGAGCCGUGCCAAUAUCACA	21	62	245	139	68	514
172	ath-miR172a	AGAAUCUUGAUGAUGCUGCAU	21	43	38	24	1	106
	ptc-miR172g	GGAAUCUUGAUGAUGCUGCAG	21	353	18985	200	51	19589
	sbi-miR172b	GGAAUCUUGAUGAUGCUGCA	20	1	193	7	0	201
	zma-miR172a	AGAAUCUUGAUGAUGCUGCA	20	344	565	622	36	1567
319	ath-miR319a	UUGGACUGAAGGGAGCUCCCU	21	0	1672	192	320	2184
	mtr-miR319	UUGGACUGAAGGGAGCUCCC	20	21	66392	4936	40730	112079
	osa-miR319a	UUGGACUGAAGGGUGCUCCC	20	0	19	5	3	27
	ppt-miR319a	CUUGGACUGAAGGGAGCUCC	20	1	1622	14	106	1743
	ppt-miR319c	CUUGGACUGAAGGGAGCUCCC	21	16	11934	643	3082	15675
	pta-miR319	UUGGACUGAAGGGAGCUCC	19	0	1917	201	1139	3257
	ptc-miR319e	UUGGACUGAAGGGAGCUCCU	20	0	22	58	45	125
	vvi-miR319e	UUUGGACUGAAGGGAGCUCCU	21	0	55	2	1	58
	vvi-miR319g	UUGGACUGAAGGGAGCUCCCA	21	0	285	23	53	361
390	ath-miR390a	AAGCUCAGGAGGGAUAGCGCC	21	299	1264	153	63	1779
	gma-miR390b	AAGCUCAGGAGGGAUAGCACC	21	46	25	193	124	388
393	ath-miR393a	UCCAAAGGGAUCGCAUUGAUCC	22	51	295	441	9	796
	osa-miR393	UCCAAAGGGAUCGCAUUGAUC	21	13	65	101	0	179
394	ath-miR394a	UUGGCAUUCUGUCCACCUCC	20	71	7473	339	482	8365
	vvi-miR394a	UUGGCAUUCUGUCCACCUCCAU	22	0	112	7	0	119
395	tae-miR395b	UGAAGUGUUUGGGGGAACUC	20	5	25	4	8	42
396	ath-miR396a	UUCCACAGCUUUCUUGAACUG	21	760	44655	10489	601	56505
	ath-miR396b	UUCCACAGCUUUCUUGAACUU	21	1	185	33	16	235
	gma-miR396d	AAGAAAGCUGUGGGAGAAUAUGGC	24	0	61	22	77	160
	gma-miR396e	UUCCACAGCUUUCUUGAACUGU	22	0	29	6	0	35
	ptc-miR396f	UUCCACGGCUUUCUUGAACUG	21	0	11	4	1	16
	vvi-miR396a	UUCCACAGCUUUCUUGAACUA	21	0	23	2	1	26
	vvi-miR396b	UUCCACAGCUUUCUUGAACU	20	742	19924	2108	823	23597
397	ath-miR397a	UCAUUGAGUGCAGCGUUGAUG	21	2057	13	5276	70	7416
	zma-miR397a	UCAUUGAGCGCAGCGUUGAUG	21	5	0	23	0	28
398	ahy-miR398	UGUGUUCUCAGGUCACCCCU	20	0	25	350	0	375
	osa-miR398b	UGUGUUCUCAGGUCGCCCCUG	21	3760	57	1717	7	5541
399	ath-miR399a	UGCCAAAGGAGAUUUGCCCUG	21	15	12	4	0	31
	ath-miR399b	UGCCAAAGGAGAGUUGCCCUG	21	35	121	22	2	180
	osa-miR399e	UGCCAAAGGAGAUUUGCCCAG	21	49	45	3	1	98
408	ath-miR408	AUGCACUGCCUCUUCCCUGGC	21	2083	3718	9050	167	15018
	bdi-miR408	AUGCACUGCCUCUUCCCUGG	20	5	25	67	0	97
	osa-miR408	CUGCACUGCCUCUUCCCUGGC	21	3	5	17	0	25
	ppt-miR408b	UGCACUGCCUCUUCCCUGGCU	21	0	5	55	0	60
482	gso-miR482a	UCUUCCCUACACCUCCCAUAC	21	7	318	85	2	412
	pvu-miR482	UCUUCCCAAUUCCGCCCAUUCC	22	5	291	54	0	350
1511	gma-miR1511	AACCAGGCUCUGAUACCAUG	20	978	625	449	1529	3581
1514	pvu-miR1514a	UUCAUUUUGAAAAUAGGCAUUG	22	188	5481	1282	311	7262
1515	gma-miR1515	UCAUUUUGCGUGCAAUGAUCUG	22	27	853	341	86	1307
2111	ath-miR2111a	UAAUCUGCAUCCUGAGGUUUA	21	48	211	255	0	514
2118	mtr-miR2118	UUACCGAUUCCACCCAUUCCUA	22	8	7	1	1	17
	pvu-miR2118	UUGCCGAUUCCACCCAUUCCUA	22	238	17300	13707	1771	33016
2119	gma-miR2119	UCAAAGGGAGUUGUAGGGGAA	21	85	42	1448	6	1581

Conserved miRNAs in common bean identified in leaves (LL), roots (RL), seedlings (SL) and flowers (FL).

Table 3 Conserved mature-star miRNAs from common bean

miRNA family	Reference miRNA	Sequence (5′- 3′)	Length (nt)	Reads				
				LL	FL	RL	SL	Total
166	aly-miR166b*	GGACUGUUGUCUGGCUCGAGG	21	233	0	5	3	241
	aly-miR166g*	GGAAUGUUGUUUGGCUCGAGG	21	38	143	186	257	624
	zma-miR166a*	GGAAUGUUGUCUGGCUCGGGG	21	26	3	11	7	47
	gma-miR166a-5p	GGAAUGUUGUCUGGCUCGAGG	21	36866	3839	6253	23902	70860
	zma-miR166m*	GGAAUGUUGGCUGGCUCGAGG	21	55	655	50	3363	4123
167	ahy-miR167-3p	AGAUCAUGUGGCAGUUUCACC	21	155	16	84	11	266
168	aly-miR168a*	CCCGCCUUGCAUCAACUGAAU	21	80	22	8	3	113
171	aly-MIR171a*	UAUUGGCCUGGUUCACUCAGA	21	179	710	1	543	1433
172	aly-miR172c*	GGAGCAUCAUCAAGAUUCACA	21	10	352	19	1	382
390	aly-miR390a*	CGCUAUCCAUCCUGAGUUUCA	21	149	524	69	10	752
	gma-miR390a-3p	CGCUAUCCAUCCUGAGUUUC	20	14	3	0	2	19
396	aly-miR396a*	GUUCAAUAAAGCUGUGGGAAG	21	267	306	3931	791	5295
482	gma-miR482a-5p	AGAAUUUGUGGGAAUGGGCUGA	22	0	17	1	1	19
	pvu-miR482*	GGAAUGGGCUGAUUGGGAAGCA	22	1185	6	440	10	1641

Conserved mature-star miRNAs in common bean identified in leaves (LL), roots (RL), seedlings (SL) and flowers (FL).

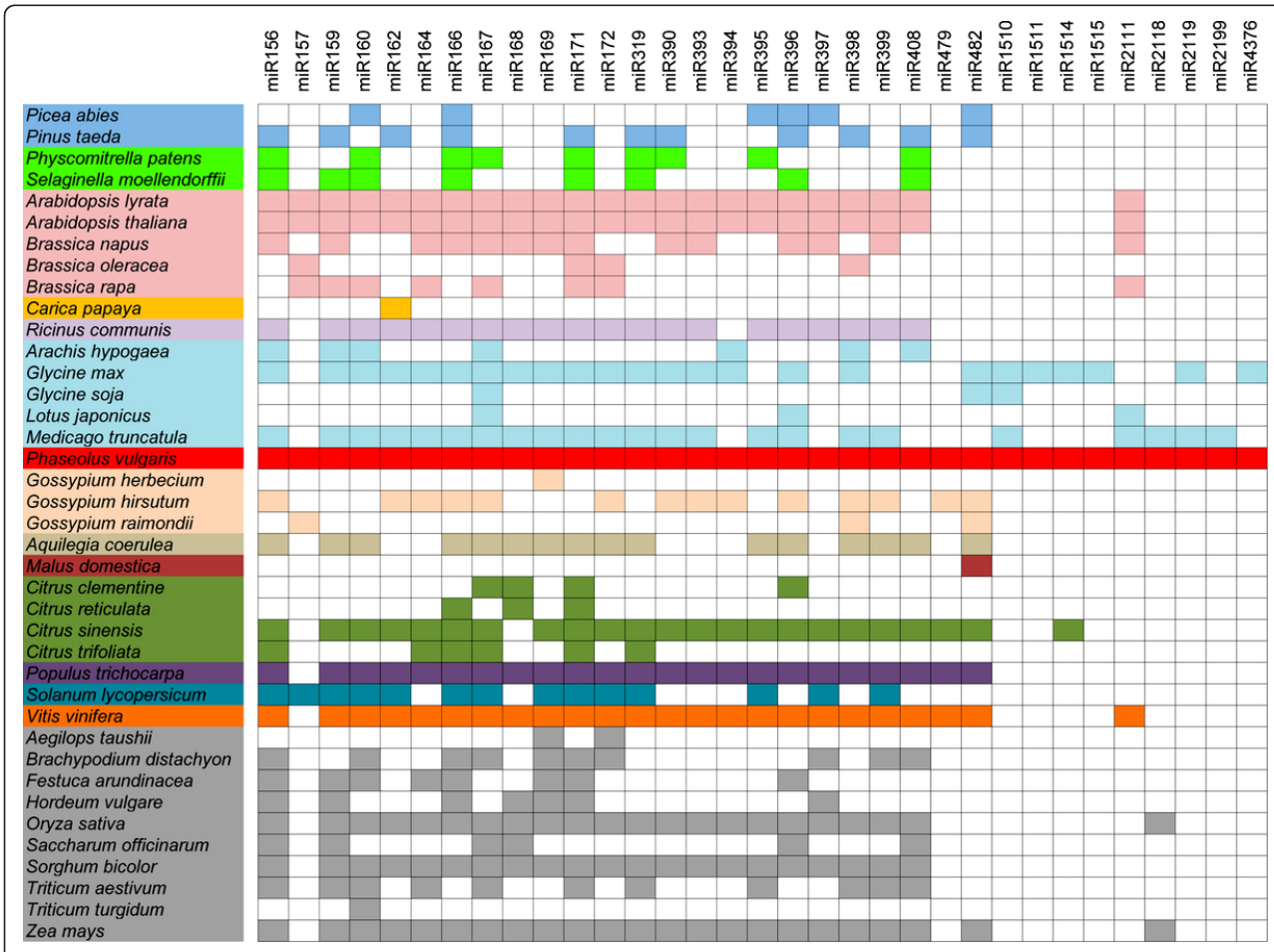


Figure 2 Conserved miRNA families in *P. vulgaris* and across-species. Evolutionary conservation of thirty-three miRNA families identified in *P. vulgaris* within plant species reported in miRBase (v.16). Closely related species are shown in the same color. Colored boxes denote the presence of the indicated miRNA family.

corresponded to the most abundant miRNAs detected in legumes like *M. truncatula*, *G. max*, and in other plants like *Vitis vinifera* [23-25]. It is worth noting that among these ten most abundant miRNAs there is a mature-star miRNA gma-miR166a-5p that has also been highly detected by both deep sequencing and microarray expression analysis in the shoot apical meristem (SAM) of soybean [26]. Also, it has been demonstrated that gma-miR166a and its mature-star miRNA have different expression patterns in the SAM, suggesting that they may play different roles in regulating leaf development [26]. The high number of reads detected for other mature-star sequences found in this study may also imply particular regulatory roles of gene expression conditioned by these small RNAs since inhibitory activity of mature-star species has been reported in animals and plants (Table 3) [27,28].

The miRNA families comprising the most miRNAs in common bean are miR156, miR159, miR166, miR167, miR169, miR319 and miR396. These highly conserved families, which generally have the most abundant miRNAs, have variable numbers of loci among the model legumes *G. max* and *M. truncatula* (Table 4) (miRBase, release 16). According to mature miRNA sequence differences, the predicted numbers of loci for *P. vulgaris* miRNA families have a higher Pearson's correlation coefficient with the numbers of loci reported for soybean (0.78) than with those identified in *M. truncatula* (0.21).

Table 4 Predicted number of loci in common bean miRNA families

miRNA family	Number of loci			miRNA family	Number of loci		
	Pvu	*Gma*	*Mtr*		*Pvu*	*Gma*	*Mtr*
156	6	7	9	396	5	5	2
157	1	0	0	397	2	0	0
159	6	4	1	398	1	2	3
160	2	1	5	399	3	0	17
162	1	1	1	408	2	0	0
164	2	1	4	479	1	0	0
166	2	2	8	482	3	2	0
167	5	7	1	1510	1	2	2
168	1	1	1	1511	1	1	0
169	6	5	17	1514	1	2	0
171	3	3	7	1515	1	1	0
172	2	3	1	2111	1	0	19
319	4	3	2	2118	2	0	1
390	3	3	1	2119	1	1	1
393	1	3	2	2199	2	0	1
394	1	2	0	4376	1	1	0
395	1	0	18				

Abbreviations stand for: *Pvu*, *Phaseolus vulgaris*; *Gma*, *Glycine max*; *Mtr*, *Medicago truncatula*.

Identification of miRNA isoforms

Small non-coding RNAs such as microRNAs were initially thought to have a specific sequence of a defined length. Identification of more miRNAs from different species has revealed that there is variation in pre-miRNA processing. The miRBase (release16) dataset for plant miRNAs consists of small sequences of 17 (.08%), 18 (.13%), 19 (.88%), 20 (9.96%), 21 (67.06%), 22 (13.85%), 23 (1.10%), and 24 (6.91%) nucleotides in length. Ebhardt *et al.* (2010) demonstrated that a fifth of the annotated *Arabidopsis thaliana* miRNAs (miRBase, release 14) have a stable miRNA isoform of one or two nucleotides longer [29]. Also, several studies have revealed that the biological function of miRNA isoforms may differ from the function of their previously reported miRNAs due to differential associations with AGO proteins [30,31]. The legumes *M. truncatula* and *G. max*, which account for the majority of entries for the *Fabaceae* family in miRBase, present mature miRNA sequences ranging from 20-22 and 19-24 nucleotides in length, respectively.

With the aim of identifying *P. vulgaris* miRNA length isoforms and species-specific miRNA variants, all small RNA sequences remaining from previous analyses were aligned against miRBase, allowing at most two mismatches and/or two nucleotides in length difference. The total number of variants found for each library (Table 1) was subjected to a filter that consisted of choosing variants that had a total number of reads 50% greater than the number of total reads of their reference miRNA previously reported, so that low-abundance and probable non-functional variants were discarded. MiRNA isoforms were classified as length variants, non-conserved miRNA variants, or conserved miRNA variants (Table 5). Seven variants based only on length were detected. In all four *P. vulgaris* libraries, the 21-nt variant of gma-miR1511 is clearly predominant. In the case of the variant for gso-miR482a, it was highly expressed in the flower (FL) and root (RL) libraries. Unexpectedly, a variant for miRNA mtr-miR171c, a miRNA that was absent in all libraries, was detected mainly in flower. The high abundance of pvu-miR482* in common bean compared with other plant species was previously corroborated by northern blot [13]. As Table 5 shows, the high number of reads detected for this mature-star miRNA in all organs was produced by the 21-nt variant and not by the one reported in miRBase. As mentioned before, some miR159 and miR319 loci produce two miRNA/miRNA* duplexes. This was the case for the *P. vulgaris* miR159a stem-loop. The variant for pvu-miR159a.2 that is found closer to the loop is

Table 5 MiRNA isoforms from *P.vulgaris*

Variant group	Reference miRNA	Sequence (5′- 3′)	Length (nt)	Reads				
				LL	FL	RL	SL	Total
Length variants	gma-miR1511	AACCAGGCTCTGATACCATG	20	978	625	449	1529	3581
	pvu-isomiR 1511	AACCAGGCTCTGATACCATGA	21	14902	13725	7707	17239	53573
	gso-miR482a	TCTTCCCTACACCTCCCATAC	21	7	318	85	2	412
	pvu-isomiR 482a	TCTTCCCTACACCTCCCATACC	22	369	17616	5564	305	23854
	mtr-miR171c	TGATTGAGCCGTGCCAATATT	21	0	0	0	0	0
	pvu-isomiR 171a	TGATTGAGCCGTGCCAATA	19	115	1390	60	57	1622
	pvu-miR482*	GGAATGGGCTGATTGGGAAGCA	22	1185	6	440	10	1641
	pvu-isomiR 482*	GGAATGGGCTGATTGGGAAGC	21	714044	3806	173153	1	1207244
	pvu-miR159a.2	CTTCCATATCTGGGGAGCTTC	21	1	0	0	0	1
	pvu-isomiR 159a	CTTCCATATCTGGGGAGCT	19	13	263	157	23	456
	gma-miR4376	TACGCAGGAGAGATGACGCTGT	22	1	1	0	0	2
	pvu-isomiR 4376	TACGCAGGAGAGATGACGCTG	21	3841	1090	0	398	5329
	mtr-miR171b	TGATTGAGCCGCGTCAATATC	21	0	0	0	0	0
	pvu-isomiR171b	TCTGATTGAGCCGCGTCAATA	21	101	1	79	10	191
Non conserved variants	ath-miR858	TTTCGTTGTCTGTTCGACCTT	21	0	0	0	0	0
	pvu-isomiR 858	CTCGTTGTCTGTTCGACCTTG	21	37	351	0	26	414
	csi-miR479	TGTGATATTGGTTCGGCTCATC	22	0	0	0	0	0
	pvu-isomiR479	TGTGATATTGGTTTGGCTCA	20	58	1	1588	11	1658
	gma-miR1510a-3p	TTGTTGTTTTACCTATTCCACC	22	0	0	0	0	0
	pvu-isomiR1510a	TGTTGTTTTTCCTATTCCACC	21	463	877	902	28	2270
	gma-miR1510b	TGTTGTTTTACCTATTCCACC	21	0	0	4	0	4
	pvu-isomiR1510b	TTGTTTTTCCTATTCCACCAA	21	3313	17893	8413	222	29841
	mtr-miR2199	TGATACACTAGCACGGATCAC	21	8	0	0	0	8
	pvu-isomiR 2199a	TGATACACTAGCACGGGTCAC	21	16	1323	799	59	2197
	pvu-isomiR 2199b	TGATACACTAGTACGGATCAC	21	2586	0	0	5	2591
	mtr-miR2597	TTTGGTACTTCGTCGATTTGA	21	0	0	0	0	0
	pvu-isomiR2597	TTTGGTACTTCCTTGATTTGA	21	0	22	253	0	275
	ppt-miR894	CGTTTCACGTCGGGTTCACC	20	1	3	1	0	5
	pvu-isomiR894	CGTTTCACGTCAGGTTCACCA	21	15	6	0	0	21
	pta-miR1310	GGCATCGGGGGCGTAACGCCCCT	23	0	0	0	0	0
	pvu-isomiR1310	GGCATCGGGGGCGCAACGCCC	21	33	5	0	0	38
Conserved variants	ctr-miR166a	TCGGACCAGGCTTCATTCCCCC	22	2	18	10	10	40
	pvu-isomiR166a	TCGGACCAGGCTTCCTTCCCC	21	113	371	241	63	788
	osa-miR156l	CGACAGAAGAGAGTGAGCATA	21	0	0	0	0	0
	pvu-isomiR 156a	TGACAGAAGAGAGTGAGCA	19	2682	7	334	222	3245
	mtr-miR164d	TGGAGAAGCAGGGCACATGCT	21	0	0	0	0	0
	pvu-isomiR164a	TGGAGAAGCAGGACACATGC	20	58	64	535	12	669
	vvi-miR156e	TGACAGAGGAGAGTGAGCAC	20	30	0	1	1	32
	pvu-isomiR 156b	TGACAGACGAGAGTGAGCAC	20	261	0	2	2	265
	csi-miR393	ATCCAAAGGGATCGCATTGATC	22	1	0	0	0	1
	pvu-isomiR393	TTCCAAAGGGATCGCATTGA	20	911	581	667	4	2163
	ctr-miR171	TTGAGCCGCGTCAATATCTCC	21	0	0	0	0	0
	pvu-isomiR171c	TTGAGCCGCGTCAATATCTCA	21	30	151	129	58	368
	mtr-miR169p	TGAGCCAGGATGGCTTGCCGG	21	0	0	0	0	0
	pvu-isomiR169a	TGAGCCGGGATGGCTTGCCGG	21	2	228	26	97	353

Table 5 MiRNA isoforms from *P.vulgaris* (Continued)

ptc-miR396f	TTCCACGGCTTTCTTGAACTG	21	0	11	4	1	16
pvu-isomiR396a	TTCCACCGCTTTCTTGAACTG	21	11	10	3	0	24
zma-miR398a	TGTGTTCTCAGGTCGCCCCCG	21	2	0	2	0	4
pvu-isomiR398a	TGTGTTCTCAGGCCGCCCCTG	21	21	0	18	0	39
sly-miR171c	TATTGGTGCGGTTCAATGAGA	21	0	0	0	0	0
pvu-isomiR171d	TATTGGTCCGGTTCAATGAGA	21	353	386	0	192	931

Phaseolus vulgaris miRNA isoforms were classified as length variants, non-conserved miRNA variants, or conserved miRNA variants. Variants were named with the prefix pvu-isomiR followed by the miRNA family number of the reference miRNA. Underlined bases denote differences in length and/or composition of bases between variants and reference miRNAs.

mainly a 19 nt miRNA, and its number of reads does not equate to the abundance of pvu-miR159a.1. The length variant gma-miR4376 increases the number of *P. vulgaris* miRNA families identified in this study because of its high number of reads. The miR4376 family had been identified only in soybean (Figure 2), and was registered in miRBase as a 22 nt miRNA [21]. The group of non-conserved miRNA variants contains nine variants from eight non-conserved miRNA families (Table 5). Considering the number of reads and that mature-miRNA homologs may present one or two mismatches, another three common bean miRNA families were identified: miR1510, miR479, and miR2199. For the miR1510 family, two variants were identified that can be produced from the same locus. This family has been reported only in soybean and *M. truncatula*, and was proposed to be conserved in common bean due to a 20 nt isoform [13,14,23]. The variant of csi-miR479, a miRNA detected in *Citrus sinensis*, *Vitis vinifera*, *Populus trichocarpa* and *Gossypium hirsutum* was particularly abundant in roots [25,32-34]. Interestingly, in this study two variants with the same length were detected for mtr-miR2199, one that was detected in all four libraries (A17 > G) and the other (C12 > T) expressed mostly in roots. A previous study in *M. truncatula* where this family was detected, reported the mature miRNA as the variant A17 > G and the genome derived-hairpin as A17, which accounts for its annotation in miRBase with an adenine in position 17 [35]. The miRNA sequence conserved in *Lotus japonicus* was the variant A17 > G also found in *P. vulgaris* [35]. The ten miRNA variants classified in the conserved miRNA variants group constitute candidates for new miRNAs of previously identified miRNA families (Table 5). The miRNA candidates with the most reads were the variants of osa-miR156l, csi-miR393, sly-miR171c, and mtr-miR169p.

Identification of stem-loop precursors

Plants have a more diverse population of sncRNAs than do animals mainly because of particular RNA polymerases. A criterion that supports miRNA annotation is the identification of a stem-loop precursor from which the duplex of miRNA/miRNA* is excised. A search for potential miRNA precursors involves secondary structure prediction analysis of genomic DNA or EST (Expressed Sequence Tag) sequences that match perfectly with determined miRNAs. Although plant miRNA stem-loops are more variable in length and structure than are animal pre-miRNAs, several characteristics are conserved among plant precursors. To identify common bean stem-loop precursors, miRNAs from Table 2, Table 3, Additional file 1, and Table 5 were aligned against all *P. vulgaris* ESTs and GSSs (Genomic Survey Sequences) from NCBI in search of perfect alignments. EST and GSS candidates were subjected to secondary structure prediction analysis using the *mfold* program with its default parameters [36]. Only the lowest energy structures were selected as described by Reinhart et al. (2002).

Secondary structure prediction analysis of *P. vulgaris* sequences resulted in the identification of eleven new stem-loop precursors belonging to eight conserved miRNA families (Figure 3 for an example and Additional file 2 for complete set). The precursor for the miR171 family was the precursor for the variant of ctr-miR171 (pvu-isomiR171c). The other miRNA families with a precursor identified were: miR166, miR167, miR156, miR157, miR398, miR408, and miR168. Taking into account previously reported *P. vulgaris* precursors, a total of 39 miRNAs can be predicted by the processing of common bean precursors, including pvu-isomiR171c (Additional file 2) [13]. The miRNAs ath-miR319c, osa-miR156k, gma-miR156f, ath-miR398a, and gma-miR482a-3p, discarded because they appeared in just one library or gave rise to fewer than 15 reads in all libraries, aligned perfectly with their respective precursors. Pvu-miR482*, aly-miR168a* and gma-miR166a-5p were also found in the precursor sequences. Therefore, common bean isoforms pvu-isomiR159a and pvu-isomiR482* were most likely derived also from these precursors.

Identification of novel microRNAs

From previous analyses, *P. vulgaris* miRNAs identical to those already present in miRBase or close homologs

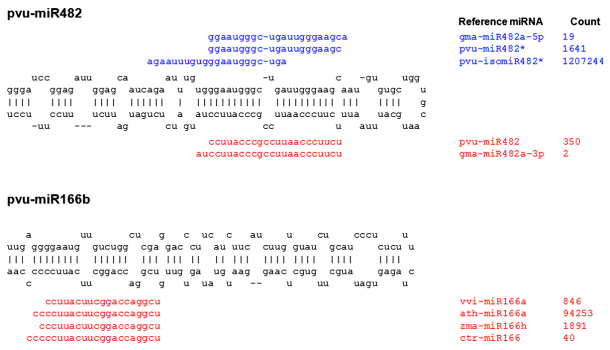


Figure 3 Differential processing of *P. vulgaris* pre-miRNAs. Stem-loop precursors of pvu-miR482 (acc.: MI0010702) and novel pvu-miR166b (gi: 171606821) pre-miRNAs were aligned against possible derived mature (red) and mature-star (blue) miRNA sequences. Count data number represents the total number of reads found in all four libraries (right column).

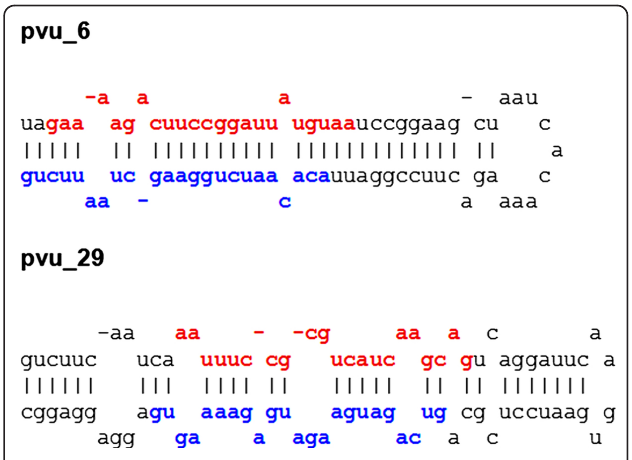


Figure 4 Predicted stem-loop precursors of two novel miRNAs. Secondary structure prediction for candidate novel miRNAs pvu_6 and pvu_29 were carried out with *mfold*. Mature (red) and mature-star (blue) miRNAs correspond to predicted mature and mature star miRNA sequences from miRDeep.

were identified. Numerous studies have described novel miRNAs in several plant species based on small RNA high-throughput sequencing results. However, this analysis is usually performed based on a reference genome that allows for mapping of a given small RNA to a genomic location and retrieval of adjoining sequence to help with secondary structure prediction of a miRNA precursor. In the present study, the sequencing reads remaining after removing known miRNAs were employed to scan the collection of ESTs and genomic sequences available from NCBI and PlantGDB in search of potential miRNAs using the miRDeep software [37]. This program searches for those small RNAs present in the sequencing reads that can be mapped to a given reference sequence. Selected sequences are searched for their potential RNA secondary structure to identify those that can be folded as hairpin structures typical of plant miRNA precursors. Additionally, the mapped small RNAs should be located in the stem region of the folded RNA, and a higher probability score is awarded to a given candidate if there are sequencing reads corresponding to the predicted miRNA* region.

The analysis for identification of novel miRNAs was carried out with 95030, 74759, 87956 and 65741 unique sequences from the leaf, root, seedling and flower small RNA libraries, respectively, and a total of 124894 ESTs and genomic sequences. A total of 29 candidates for new miRNA precursors were identified (Figure 4 for examples and Additional file 3 for complete set of candidate miRNAs). Four of the candidate precursors included reads for a mature miRNA and a putative miRNA* and the other 25 had only reads consistent with the predicted mature miRNA. The small number of ESTs and amount of genomic data available for *P. vulgaris* limits the number of candidates recovered by this analysis. The forthcoming availability of a complete genome should reveal a more comprehensive picture of the miRNA genes present in the common bean genome.

Identification of miRNA targets in the genus *Phaseolus*

Plant miRNAs typically have higher sequence complementarity with their target mRNAs compared with miRNA-target interactions in animals [38]. MiRNA target prediction in plants is based on this high degree of complementarity between miRNAs and targets [39]. Many plant miRNA:mRNA duplexes exhibit paired nucleotides in the microRNA 5' region. However, some conserved plant miRNA-target interactions with unpaired nucleotides in the 'seed' region have been identified, for instance miR398a-CSD2 and miR396a-GRL7/GRL8/GRL9 in *A. thaliana* [39]. In this study, the RNAhybrid program as described by Alves et al. (2009) for prediction of *Phaseolus* miRNA targets was used (See methods) [40]. RNAhybrid predicts the energetically most favorable miRNA:mRNA hybrids according to user preferences [41]. All *Phaseolus* ESTs (NCBI) were evaluated as targets of miRNAs from Table 2, Table 3, and Table 5. Identification analysis for *P. vulgaris* miR390 non-coding transcript target TAS3 was done separately (See methods) [42]. Once miRNA:mRNA hybrids were obtained, ESTs were aligned against plant sequences of the UniProt Knowledgebase (UniProtKB) (release 2011_01) using BLASTX for annotation.

Target prediction analysis identified 194 ESTs annotated as established target gene families in plants (Additional file 4). Thirty seven ESTs were previously reported by Arenas-Huertero et al. (2009) and most are ESTs with small minimum free energies (MFEs) (data

not shown). The conserved miRNA families for which conserved targets were found are: miR156/miR157, miR160, miR164, miR167, miR168, miR169, miR171, miR172, miR319, miR393, miR395, miR397, miR398 and miR408. The number of ESTs found for each gene family are: SBP(7), ARF(18), NAC(1), AGO1(1), NFY (14), SCL(2), AP2(8), TCP(6), TIR/F-box-AFB(5), ATP sulfurylases(14), Laccases(1), COX/SOD(66) and Plastocyanins(11). The miR156 and miR157 families share the same target EST sequences for Squamosa Binding Proteins (SBP). The MFEs for the conserved miRNA:mRNA hybrids range from -35.0 to -51.2 kcal/mol. Finally, among conserved miRNA targets, three ESTs homologous to AtTAS3, a target of miR390 family were identified (Additional file 4). In addition, 325 ESTs as putative targets for *P. vulgaris* conserved miRNAs, mature-star sequences and miRNA isoforms were identified (Additional file 5). Target prediction analysis for novel candidate miRNAs generated candidate targets for ten novel miRNAs represented by 177 ESTs (Additional file 6).

Organ-specific expression of *P. vulgaris* miRNA families

Specific miRNA expression patterns are proposed to be a consequence of tissue-specific, cell-specific and/or stress-specific regulatory elements in promoters of plant microRNA (*MIRNA*) genes [43-45]. The expression of some conserved *MIRNA* genes and the relative abundance patterns of mature miRNAs in different organs or in different developmental stages are essential for proper cell differentiation and organ developmental regulation. Modification of *MIRNA* gene expression leads to severe developmental defects. Understanding expression patterns of microRNAs in plant organs is necessary to discern miRNA-mediated regulatory pathways. The frequency of miRNAs detected by high throughout sequencing methods serves for relative miRNA expression estimation [46]. Moldovan et al. (2009), comparing *Arabidopsis* root miRNA frequencies against leaf and whole plant reads represented in the Arabidopsis Small RNA Project database (ASRP), found that most miRNA families have organ-specific expression [47,48].

To explore organ-specific miRNA expression in *Phaseolus vulgaris*, the open-source R/Bioconductor software package DESeq was employed [49,50]. The DESeq package, supported by a model based on negative binomial distribution, estimates variance-mean dependence in count data from high-throughput sequencing assays and tests for differential expression. The miRanalyzer tool bases its differential expression module on the DESeq package and has been used to test for miRNA differential expression [51,52]. To explore organ-specific expression for miRNA families as a whole, frequencies for miRNAs of the same family were pooled together for each library and employed as input data. The five miRNAs that aligned perfectly with precursor sequences and were discarded because they appeared in just one library or totalled fewer than 15 reads in all libraries, as well as those newly identified from the isoform analysis, were also considered. Mature-star miRNAs were discarded. Count data was transformed with the DESeq package and hierarchical cluster analysis (HCA) was performed (Figure 5, see methods).

MiRNA families were clustered in two major groups: the group with fewer miRNAs consists of miRNA families that were abundantly expressed in all four libraries, and the other group contains miRNA families with differential expression patterns among organs. The accumulation patterns of miRNA families in leaf and root organs were more similar to each other than to the other two samples. The most abundant families of the analysis, especially in seedlings, were miR159, miR319, miR396, miR166, miR408, and miR482. The miR159 family has previously been reported as an abundant and widespread family in all plant organs. It is important to mention that the expression of miR319 was abundant in all organs except for leaf (Figure 5). Similar to miR319, in this study miR396 was less abundant in leaf than in the other three organs. Other quite abundant miRNA families were: miR1510, miR164, miR167, miR169, miR157 and miR156. Interestingly miR1510, a miRNA family detected only in legumes, was found in all four libraries.

In the group of families with a particular expression pattern, miR2119, miR2111, and miR479 were highly accumulated specifically in one organ. MiR2119 was detected especially in roots. This expression pattern correlates with previous northern blot analysis of roots, leaves, immature embryos and dry embryos from *P. vulgaris* [13]. Expression of miR2111 was observed mainly in roots and miR399 was only slightly expressed in leaves. Both miRNAs are induced under phosphorus (P) limiting conditions [12,53]. The miR479 family was preferentially expressed in seedlings. On the other hand, there were some interesting families that were poorly expressed in one particular organ. For example, miR168 was poorly expressed in seedlings, miR1515 in leaves, miR2199 in roots, miR394 in leaves, miR4375 in roots, and miR397 in flowers. Finally, HCA showed that miR395 was poorly expressed in all four small RNA libraries, consistent with its induction upon low sulphate conditions [54].

Discussion

Identification of microRNAs in species with or without fully sequenced genomes has been revolutionized by high-throughput sequencing methods. High-throughput sequencing miRNA analysis in *Phaseolus vulgaris* will facilitate more particular and specific bean miRNA

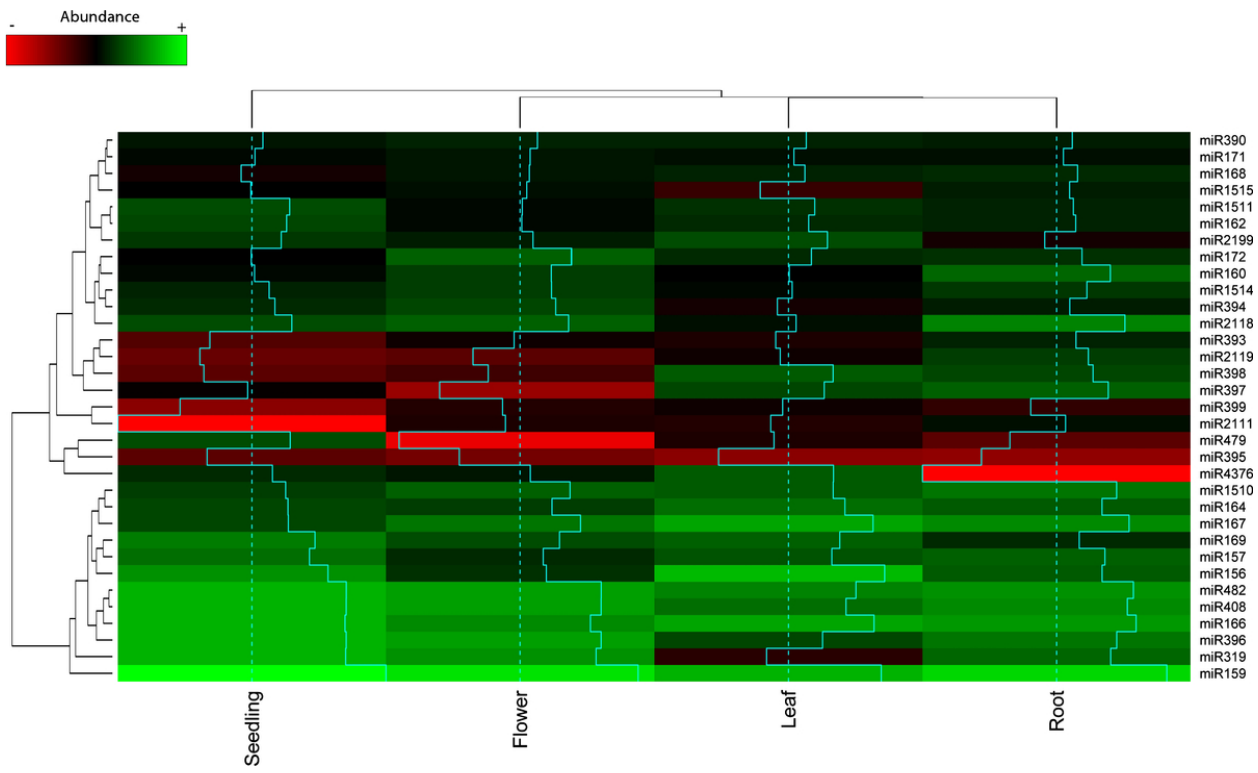


Figure 5 Hierarchical cluster analysis (HCA) of *P. vulgaris* miRNA families. Frequencies of conserved miRNA families were used to analyze relative miRNA family expression in leaf, seedling, root and flower. Hierarchical cluster analysis was performed on thirty-three conserved miRNA families (right). The color scale denotes abundant (green) and non-abundant (red) miRNA families according to transformed count data. Density plots are included in blue lines.

studies as well as legume sncRNAs research in this rapidly growing field. Although deep-sequencing experiments have become the major source supporting microRNA annotations, technical variations, transcript length bias and mapping bias have been reported with this approach for transcriptome analysis of coding RNAs (mRNAs) [17,55,56]. Studies to elucidate the number of miRNA molecules sequenced from a small RNA library are still needed for more accurate small RNA profiling studies. In terms of reads, the small RNA libraries sequenced here yielded a larger number of total raw reads to work with than did many other studies in which plant miRNAs have been identified; however, the total number of reads of identified miRNAs, including non-abundant and miRNA isoforms, constitutes only 0.047% of total raw reads. Further studies are needed to understand to what extent the 24-nt class representing siRNA populations, usually the most abundant and diverse class of sncRNAs sequenced in small RNA libraries, masks miRNA populations. Also, coverage analyses with fully sequenced genomes are needed to elucidate sequenced sample proportions of small non-coding RNAs such as tRNA, rRNA, snoRNA or snRNA.

Most miRNA families identified in this study in *P. vulgaris* are evolutionarily conserved in *Fabaceae*, particularly in the best-studied plants *M. truncatula* and *G. max*. MiRNA family miR157 has not been reported in these two legumes, probably because of the similarity of its sequence to that of miR156. Other miRNA families not reported in miRBase in *G. max* and *M. truncatula* are miR397 and miR408. The only *Fabaceae* where miR408 has been previously found was *Arachis hypogaea*. One miRNA of the bean miR397 family with the same sequence as ath-miR397 was detected herein with 7416 reads in total, so it is likely to be present in other legumes as well. MicroRNA miR2199 has been reported in miRBase only for *M. truncatula* and the variant found here for this miRNA in all four libraries of *P. vulgaris* was previously reported as miRS1 by Arenas-Huertero et al. (2009) because mtr-miR2199 was not yet reported. As well as in *Lotus japonicus*, the A17- > G (miRS1) variant was found in peanuts (*Arachis hypogaea L.*) [35,57].

It is interesting that miRNA families such as miR1511, miR1514 and miR1515 first detected in soybean, have not been found in other legumes besides *P. vulgaris*. In this regard, other miRNA families reported specifically

in soybean (gma-miR1524, gma-miR1532, gma-miR1526, gma-miR1516 and gma-miR1513 and gma-miR1508) were detected by Valdés-López et al. (2010) using a hybridization approach (macroarrays) under several abiotic stress conditions in leaves, roots and nodules of *P. vulgaris*. These miRNAs from soybean were not detected by high-throughput sequencing, with the exception of two reads for gma-miR1508. In the Valdés-López et al. (2010) analysis, expression was detected for miRNAs miR1524, miR1526, miR1532 and miR1508 in common bean plants grown under sufficient nutrient conditions or stressed conditions. Family miRNAs gma-miR1513 and gma-miR1516 were detected only under stressed conditions. In addition, expression was detected in the macroarrays for an oligonucleotide probe designated as pvu-miR1509. This probe was based on a variant detected by Arenas-Huertero et al. (2009) that was not detected here because it has more than two mismatches with the reference miRNA gma-miR1509. The miR170 family, for which Valdés-López et al. (2010) detected expression in *P. vulgaris*, was neither found in this study. MiR170 is very similar to miR171, so it is possible that the macroarray probes hybridized with the same miRNA family.

MiRNA variants are considered to be a consequence of inaccuracies in Dicer pre-miRNA processing. Smaller variants with missing bases and low frequencies are viewed as degradation products or pyrophosphate sequencing errors. Herein, small RNA sequences from libraries were classified as miRNA isoforms only if they were similar to a reference miRNA reported in miRBase and had a significantly greater number of reads compared to those found for the reference miRNA in all four organs. From this isoform identification analysis, four more miRNA families (miR1510, miR2199, miR4376 and miR479) were added to the total number of conserved miRNA families identified in common bean. All of the variants found here for these four families were highly abundant in all four libraries. It is probable that miR1510 and miR2199 miRNA families are part of the *P. vulgaris* miRNA population, based on the two variants identified for each of these two families, along with their conservation in other species (see results) and previously experimental expression analyses [13,22].

Another family highly likely to be present in common bean is the miR4376 family. The variant found here for miR4376 was very abundant in common bean and has only one nucleotide missing relative to its reference miRNA found in soybean. On the other hand, further experimental and genomic sequences analyses are still needed to validate the variant identified in common bean for the miR479 family, which has two nucleotides missing and one mismatch compared with its reference miRNA csi-miRNA479. That is also the case for other less conserved miRNA families identified based on the isoform analysis, like miR858, miR2597, miR894 and miR1310. Most of the reference miRNAs for these less conserved miRNA family variants were poorly or not detected in the libraries.

In plant and animal microRNAs, 3' heterogeneity is the most frequent length variation. Most of the variants identified from the length variant group exhibit 3' heterogeneity. Sequence length heterogeneity for plant microRNAs has been proven to be essential for correct plant development and environmental responses. It is known that miR168 in *Arabidopsis thaliana* is produced in length variants of 21 and 22 nucleotides [31]. A decrease in abundance of the 21 nt variant reduces miR168 homeostasis and leads to developmental defects. The large number of reads detected for some variants in *P. vulgaris* suggests significant regulatory roles like detected for miR168 in *A. thaliana*. Using the pre-miRNA of pvu-miR482 (Figure 4) as an example, three variants generated for "mature-star" miRNAs (5p) and two variants generated for the mature miRNAs (3p) were observed. Considering the number of reads found for these variants, mature-star sequence variants were by far greater in abundance than their corresponding mature sequence variants. In particularl two mature-star sequence length variants, one previously reported in miRBase as pvu-miR482* of 22 nucleotides and the other here denoted as pvu-isomiR482* of 21 nucleotides, had 1641 and 1207244 total reads respectively. It is important to take into account that sequences for reference miRNA gso-miR482a and its highly abundant variant pvu-isomiR482a were detected, and that these variants necessarily have to be excised from another stem-loop precursor not yet identified. For this reason, it is possible that the variants pvu-miR482* and pvu-isomiR482* were generated from different loci, as actually happens for the two miR168 variants in *A. thaliana*. In addition to miRNAs previously reported in other plants and those present in closely related species such as soybean or *Medicago*, common bean may encode species-specific miRNAs. To address this question, sequencing reads were explored using the miRDeep algorithm [37]. Those candidates were favoured where in addition to a mature miRNA present in the stem region of the hairpin precursor, there was also evidence of a miRNA* sequence recovered. Next, candidate miRNAs having a plausible stem-loop precursor but without any miRNA* sequences in the libraries were obtained. It will be interesting to see to what extent these candidates can be confirmed as genuine miRNAs and whether they are involved in biological processes specific to *P. vulgaris*. As in other examples shown here, the availability of a fully-sequenced genome will provide a more complete picture of novel miRNAs.

Currently, microarray hybridization approaches, real time quantitative PCR (RT-qPCR) analyses and high-throughput sequencing technologies are widely used microRNA profiling methods. MiRNA regulatory functions in different organs, different stress conditions, and different developmental stages are still unknown. MiRNA profiling studies are important first approximations to analyze miRNA functions according to their different expression patterns. In this work miRNA expression was analyzed in order to find important differences in miRNA family expression levels within common bean organs. Biological replicates are essential to determine if differences observed are caused by conditions and not just by experimental variations. Because of this, cluster analysis was employed. The DESeq package allows users to work without replicates with the caveat that the test will lose strength. It assumes that most transcripts will have approximately the same levels within replicates under the different conditions, and that the estimated variance should not change too much.

Experiments designed to explore miRNA and mRNA expression are subjected to many different technical and biological biases. Nevertheless, differential processing of stem-loop precursors, small RNA duplexes, and, in general, miRNA biochemical properties challenge expression analysis methods usually used to analyze mRNAs. The majority of conserved miRNA families identified were expressed in all organs at different levels. Conserved miRNA families seem to be transcribed in almost all plant organs. Which of those conserved miRNAs have an essential role in legumes for certain developmental stages or stress responses is not completely understood. Perhaps less conserved miRNAs families that are highly abundant in legumes such as miR1510, miR1511, miR1514, miR1515, miR2118 and miR2199, play essential roles in characteristic processes of legumes such as nodulation, as altered expression of miR482 and miR1515 has been proven to increase soybean nodulation [58].

The present study contributes, together with previous common bean miRNA studies, to characterizing the *P. vulgaris* miRNA population. It represents a unique analysis of *P. vulgaris* miRNAs performed with high-throughput next-generation DNA sequencing (NGS) technologies. Shortly, full genome sequence and transcriptome datasets from different *P. vulgaris* cultivars will be available. The analysis of novel common bean genome sequence information will benefit from the tools presented here to expand important small RNA research needed in this critical worldwide crop.

Conclusions

109 miRNAs belonging to 29 conserved families in *Phaseolus vulgaris* were identified using high-throughput sequencing. In addition, twenty six miRNA isoforms were characterized. MiRNA variant analysis identified four highly abundant miRNA families. Eleven new stem-loop precursors belonging to eight conserved miRNA families were determined. Thirty nine miRNAs identified can be explained by processing of characterized common bean precursors. In addition, twenty nine novel miRNA candidates were predicted based on small RNA reads and precursor prediction. Evidence for miRNA* sequences for four of these precursors was found. Target prediction analysis identified 157 new *Phaseolus* ESTs as established miRNA target genes in plants. Candidate targets for miRNA families derived from *Phaseolus* ESTs were proposed. The common bean miRNA families identified were differentially expressed in leaves, roots, seedlings and flowers. This work provides an important global view of conserved and novel *Phaseolus vulgaris* miRNAs, their precursors and their targets.

Methods

Plant material and growth conditions

Phaseolus vulgaris L. cv. Negro Jamapa and cv. Pinto Villa were used in this study. Common bean (cv. Negro Jamapa) seeds were surface sterilized by an initial treatment with 100% ethanol for 1 min, rinsed with sterile water, and treated with 20% sodium hypochlorite for 5 min. Sterilized seeds were transferred to sterile trays containing wet paper towels. Trays were covered with foil and held at 28°C for 72 hours until seed germination. Sprouts then were transferred to small plastic pots containing vermiculite watered with B&D solution supplemented with 8 mM KNO_3 [59]. Incubation was performed in a chamber with 16 h of light and 8 h of dark at 28°C. Plants were watered every third day until organ collection. Roots (15 d old) and flower buds (35-40 d old) were collected in liquid N_2 and stored at - 80°C.

P. vulgaris cv. Pinto Villa seeds were surface-sterilized in 50% (v/v) sodium hypochlorite and 0.5% (v/v) Tween-20 for 3 min, and rinsed with distilled water for 10 min. Seeds were transferred to trays containing wet paper towels. Whole seedlings 1-4 days old were collected in liquid N_2 and stored at -80°C. For leaf collection, 4-day-old seedlings were transferred from trays to plastic pots containing 40% vermiculite, 30% Metro-Mix soil and 30% Agrolite. Once the first trifolium appeared, plants were kept well-watered. A pool of leaves from 10 and 20 days old was harvested for RNA purification.

RNA isolation and small RNA library sequencing

Total RNA was isolated from frozen roots, seedlings, flower buds and leaves using the Trizol reagent according to manufacturer's instructions (Invitrogen, Carlsbad, CA). The RNA Integrity Number (RIN) was larger than

seven and the 28S/18S ratio was larger than 1.6 for all sample organs. Ten micrograms of each sample (roots, flower buds, seedlings, and leaves) were prepared for Deep Sequencing following Illumina's Small RNA alternative sample preparation protocol v1.5. Small RNA fragments ranging from 18-30 nt were selected for the construction of the small RNA libraries. Complementary DNA libraries were separately Single Read-sequenced using the Genome Analyzer IIx (GAIIx)(36 bp) and the Illumina Cluster Station (Illumina Inc, USA) at the Instituto de Biotecnología (Universidad Nacional Autónoma de México).

Small RNA sequencing analysis

Raw reads from the *Illumina Pipeline 1.4* for the four small RNA libraries were cleaned of sequence adapters, low quality tags and small sequences (< 16 nt long). Quality analysis per cycle was performed for each library. Later, sequences were converted to FASTA format grouped in unique sequence tags with their respective frequencies. The R/Bioconductor software package ShortRead (version 1.10.4) was used for RNA sequence length distribution analysis [60]. Mature miRNA sequences from miRBase (release 16) were removed from the RNA families database (Rfam 10.0). Libraries were aligned against miRNA cleaned Rfam database using BLASTN. Exact matches identified with a Perl script for parsing blast results were then removed from libraries.

Identification of conserved, isoform, and novel miRNAs

Mature and mature-star miRNA sequences from plants were extracted from miRBase (release 16). Sequences were grouped into unique sequences with a reference miRNA identifier. Small RNA libraries cleaned with the Rfam database were aligned against the unique mature and mature-star miRNA sequences using BLASTN and SSAHA2 (version 2.5.3). Alignment results were processed to obtain small RNA sequences that corresponded exactly in size and nucleotide composition to reported plant miRNA sequences. MiRNAs that were detected in just one library or that totalled fewer than 15 absolute appearances in all libraries were separated. Correlation coefficient analysis was done with the number of loci reported in miRBase (v.16) for *Medicago truncatula* and *Glycine max*, and the number of possible loci for *P. vulgaris* based on mature miRNA differences. With the purpose of identifying miRNA isoforms, sequences left from Rfam and miRBase filters that presented a total frequency higher than 10 in all libraries were BLASTN aligned against miRBase (mature and mature-star plant miRNA sequences) allowing at most two mismatches and/or two different nucleotides in length. The total number of variants found for each library was subjected to an abundance filter which consisted of choosing variants that had a total number of reads 50% greater than the total reads of their reference miRNA previously reported. If no reference miRNA for a variant was previously detected in all libraries, the variant with the highest frequency was considered. Isoforms detected for conserved miRNAs that belong to a miRNA family were analyzed to discard the possibility of being simply another miRNA of the same family. MiRNA isoforms were classified as length variants, non-conserved miRNA variants or conserved miRNA variants. Identification of novel miRNAs was performed with sequencing reads that remained after known miRNAs and miRNA isoforms were removed. EST and genomic sequences available from NCBI and PlantGDB were used to search for potential miRNAs using the miRDeep software [37]. Novel miRNA candidates were further aligned against nucleotide and protein NCBI databases to discard possible RNA degradation products.

Identification of stem-loop precursors and targets

To identify stem-loop precursors, *P. vulgaris* ESTs and GSSs http://www.ncbi.nlm.nih.gov/projects/dbEST/; http://www.ncbi.nlm.nih.gov/projects/dbGSS/from NCBI were aligned against mature miRNAs, mature-star sequences and miRNA isoforms (Table 2, Table 3, Table 5 and Additional file 1) [61]. Sequence candidates with 100% coverage and identity were tested with the secondary structure program *mfold* for pre-miRNA structures [36]. Only the lowest energy structures generated for each sequence candidate were analyzed. Target prediction analysis was performed for all ESTs of the genus *Phaseolus* from NCBI. EST sequences of the genus *Phaseolus* were confirmed to be targets of mature miRNAs, mature-star sequences and miRNA isoforms (Table 2, Table 3, and Table 5) using the RNAhybrid program (version 2.1) as described by Alves et al. (2009). Potential EST targets were then aligned against plant sequences of the UniProt Knowledgebase (UniProtKB) (release 2011_01) using BLASTX for annotation. Identification analysis for common bean miR390 non-coding transcript target TAS3 was done separately. The AtTAS3 (locus: AT3G17185.1) sequence from the Arabidopsis Information Resource (TAIR) was used for BLASTN alignment against *Phaseolus* ESTs. Later, miR390 conserved binding sites were analyzed on EST candidates. Only novel mature and mature-star miRNA sequences detected in the libraries were used for target prediction analysis.

Organ-specific expression analysis

As first step for organ-specific expression analysis, frequencies of miRNAs of the same miRNA family were added. MiRNAs considered for this analysis were the

conserved miRNAs (Table 2), the five miRNAs from Additional file 1 that aligned with common bean stem-loop precursors, and the new four miRNA families proposed from the miRNA isoform study. Mature-star sequences were not considered. To test for miRNA differential expression the open-source R/Bioconductor software package DESeq was used [49]. The effective library size for all small RNA library data was estimated with the *estimateSizefactors* function. This function is used to normalize frequencies of transcripts in relation to library sizes. Then, *estimateVarianceFunctions* was used to predict the variance from the mean for each organ sample. Within these functions, the *method = "blind"* parameter for comparisons without replicates was employed. Eventually the count data was transformed with the *getVarianceStabilizedData* function such that its variance becomes independent of the mean producing a homoscedastic version of the data. Heatmap was performed using the tool heatmap.2 from the gplopts package of open source R software. Hierarchical cluster analysis was performed with the *hclust* function from the stats package.

Additional material

Additional file 1: Conserved miRNAs detected only in one library and/or with fewer than 15 reads. MiRNAs in common bean identified in leaves (LL), roots (RL), seedlings (SL) and flowers (FL) detected in just one library and/or with fewer than 15 total reads.

Additional file 2: Stem-loop miRNA precursors in common bean. EST fragments (new conserved pre-miRNAs) and fragments of stem-loop sequences reported in miRBase (v. 16) for *P. vulgaris* where aligned against miRNA sequences (blue sequences). Count data number represents the total number of reads in all four libraries. Count data number represents the total number of reads in all four libraries.

Additional file 3: Novel miRNAs detected in *Phaseolus vulgaris*. Novel miRNAs in common bean identified in leaves (LL), roots (RL), seedlings (SL) and flowers (FL) using miRDeep.

Additional file 4: Predicted conserved targets in *Phaseolus*. Predicted conserved miRNA targets based on ESTs of the genus *Phaseolus*. GenBank accession numbers are used for EST identification. Calculated MFEs (kcal/mol) using RNAhybrid are shown. MiRNA families 156 and 157 share the same predicted conserved targets. MicroRNA sequences in 3'-5' sense were used to represent miRNA:target pairing. Crosses (x) and asterisks (*) denote one-nucleotide wobbles and mismatches, respectively. G:U base pairing is not considered a mismatch (|)

Additional file 5: Predicted Putative targets in *Phaseolus*. Predicted putative miRNA targets based on ESTs of the genus *Phaseolus*. GenBank accession numbers are used for EST identification. Calculated MFEs (kcal/mol) using RNAhybrid are shown. MicroRNA sequences in 3'-5' sense were used to represent miRNA:target pairing. Crosses (x) and asterisks (*) denote one-nucleotide wobbles and mismatches, respectively. G:U base pairing is not considered a mismatch (|)

Additional file 6: Putative targets for novel miRNAs identified in *Phaseolus*. Predicted putative targets for novel miRNAs in common bean based on EST of the genus *Phaseolus*. MicroRNA sequences in 3'-5' sense were used to represent miRNA:target pairing. Crosses (x) and asterisks (*) denote one-nucleotide wobbles and mismatches, respectively. G:U base pairing is not considered a mismatch (|)

Acknowledgements
We thank the Unidad Universitaria de Secuenciación Masiva de DNA (UUSMD) of the Universidad Nacional Autónoma de México (UNAM) for sequencing services. We also are grateful to the Instituto de Biotecnología-UNAM for giving us access to its computer cluster. Financial support for this research was provided by the International Atomic Energy Agency (research contract No. 16601/R0), Consejo Nacional de Ciencia y Tecnología 83324 and Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica IN214909-3 grants.

Authors' contributions
JL and PP conceived of and coordinated the study. PP wrote the manuscript. PP, MS and LP carried out the bioinformatic analyses. GE and JL carried out plant growth, RNA extraction and preparation, and helped drafting the manuscript. JL, AA and FE participated in its design, helped to draft the manuscript and provided intellectual suggestions. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 14 October 2011 Accepted: 6 March 2012
Published: 6 March 2012

References

1. Mallory AC, Vaucheret H: **Functions of microRNAs and related small RNAs in plants.** *Nat Genet* 2006, **38(Suppl)**:S31-S36.
2. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN: **MicroRNA genes are transcribed by RNA polymerase II.** *EMBO J* 2004, **23(20)**:4051-4060.
3. Kim VN: **MicroRNA biogenesis: coordinated cropping and dicing.** *Nat Rev Mol Cell Biol* 2005, **6(5)**:376-385.
4. Bologna NG, Mateos JL, Bresso EG, Palatnik JF: **A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159.** *EMBO J* 2009, **28(23)**:3646-3656.
5. Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X: **Methylation as a crucial step in plant microRNA biogenesis.** *Science* 2005, **307(5711)**:932-935.
6. Li J, Yang Z, Yu B, Liu J, Chen X: **Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*.** *Curr Biol* 2005, **15(16)**:1501-1507.
7. Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS: **Nuclear processing and export of microRNAs in *Arabidopsis*.** *Proc Natl Acad Sci* 2005, **102(10)**:3691-3696.
8. Hammond SM, Bernstein E, Beach D, Hannon GJ: **An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells.** *Nature* 2000, **404(6775)**:293-296.
9. Baumberger N, Baulcombe DC: ***Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs.** *Proc Natl Acad Sci* 2005, **102(33)**:11928-11933.
10. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP: **MicroRNAs in plants.** *Genes Dev* 2002, **16(13)**:1616-1626.
11. Sunkar R, Jagadeeswaran G: **In silico identification of conserved microRNAs in large number of diverse plant species.** *BMC Plant Biol* 2008, **8**:37.
12. Valdés-López O, Arenas-Huertero C, Ramírez M, Girard L, Sánchez F, Vance CP, Luis Reyes J, Hernández G: **Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean roots.** *Plant Cell Environ* 2008, **31(12)**:1834-1843.
13. Arenas-Huertero C, Pérez B, Rabanal F, Blanco-Melo D, De la Rosa C, Estrada-Navarrete G, Sanchez F, Covarrubias AA, Reyes JL: **Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress.** *Plant Mol Biol* 2009, **70(4)**:385-401.
14. Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu JK, Yu O: **Novel and nodulation-regulated microRNAs in soybean roots.** *BMC Genomics* 2008, **9**:160.
15. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA: **Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells.** *Genome Res* 2008, **18(4)**:610-621.

16. Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T: **Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening.** *Genome Res* 2008, **18(10)**:1602-1609.
17. Kozomara A, Griffiths-Jones S: **miRBase: integrating microRNA annotation and deep-sequencing data.** *Nucleic Acids Res* 2011, , **39 Database:** D152-7.
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215(3)**:403-410.
19. Ning Z, Cox AJ, Mullikin JC: **SSAHA: a fast search method for large DNA databases.** *Genome Res* 2001, **11(10)**:1725-1729.
20. Cuperus JT, Fahlgren N, Carrington JC: **Evolution and functional diversification of MIRNA genes.** *Plant Cell* 2011, **23(2)**:431-442.
21. Joshi T, Yan Z, Libault M, Jeong DH, Park S, Green PJ, Sherrier DJ, Farmer A, May G, Meyers BC, Xu D, Stacey G: **Prediction of novel miRNAs and associated target genes in *Glycine max*.** *BMC Bioinformatics* 2010, **11(Suppl 1)**:S14.
22. Valdés-López O, Yang SS, Aparicio-Fabre R, Graham PH, Reyes JL, Vance CP, Hernández G: **MicroRNA expression profile in common bean (*Phaseolus vulgaris*) under nutrient deficiency stresses and manganese toxicity.** *New Phytol* 2010, **187(3)**:805-818.
23. Szittya G, Moxon S, Santos DM, Jing R, Fevereiro MP, Moulton V, Dalmay T: **High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families.** *BMC Genomics* 2008, **9**:593.
24. Song QX, Liu YF, Hu XY, Zhang WK, Ma B, Chen SY, Zhang JS: **Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing.** *BMC Plant Biol* 2011, **11**:5.
25. Mica E, Piccolo V, Delledonne M, Ferrarini A, Pezzotti M, Casati C, Del Fabbro C, Valle G, Policriti A, Morgante M, Pesole G, Pè ME, Horner DS: **Correction: High throughput approaches reveal splicing of primary microRNA transcripts and tissue specific expression of mature microRNAs in *Vitis vinifer*.** *BMC Genomics* 2010, **11**:109.
26. Wong CE, Zhao YT, Wang XJ, Croft L, Wang ZH, Haerizadeh F, Mattick JS, Singh MB, Carroll BJ, Bhalla PL: **MicroRNAs in the shoot apical meristem of soybean.** *J Exp Bot* 2011, **62(8)**:2495-2506.
27. Okamura K, Phillips MD, Tyler DM, Duan H, Chou YT, Lai EC: **The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution.** *Nat Struct Mol Biol* 2008, **15(4)**:354-363.
28. Devers EA, Branscheid A, May P, Krajinski F: **Stars and Symbiosis: MicroRNA- and MicroRNA*-Mediated Transcript Cleavage Involved in Arbuscular Mycorrhizal Symbiosis.** *Plant Physiol* 2011, **156(4)**:1990-2010.
29. Ebhardt HA, Fedynak A, Fahlman RP: **Naturally occurring variations in sequence length creates microRNA isoforms that differ in argonaute effector complex specificity.** *Silence* 2010, **1(1)**:12.
30. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y: **Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide.** *Cell* 2008, **133(1)**:116-127.
31. Vaucheret H: **AGO1 homeostasis involves differential production of 21-nt and 22-nt miR168 species by MIR168a and MIR168b.** *PLoS One* 2009, **4(7)**: e6442.
32. Xu Q, Liu Y, Zhu A, Wu X, Ye J, Yu K, Guo W, Deng X: **Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type.** *BMC Genomics* 2010, **11**:246.
33. Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL: **Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*.** *Plant Cell* 2005, **17**:2186-2203.
34. Pang M, Woodward AW, Agarwal V, Guan X, Ha M, Ramachandran V, Chen X, Triplett BA, Stelly DM, Chen ZJ: **Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber development in allotetraploid cotton (*Gossypium hirsutum L.*).** *Genome Biol* 2009, **10**:R122.
35. Jagadeeswaran G, Zheng Y, Li YF, Shukla LI, Matts J, Hoyt P, Macmil SL, Wiley GB, Roe BA, Zhang W, Sunkar R: **Cloning and characterization of small RNAs from *Medicago truncatula* reveals four novel legume-specific microRNA families.** *New Phytol* 2009, **184(1)**:85-98.
36. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction.** *Nucleic Acids Res* 2003, **31**:3406-3415.
37. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N: **Discovering microRNAs from deep sequencing data using miRDeep.** *Nat Biotechnol* 2008, **26**:407-415.
38. Wang XJ, Reyes JL, Chua NH, Gaasterland T: **Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets.** *Genome Biol* 2004, **5(9)**:R65.
39. Jones-Rhoades MW, Bartel DP: **Computational identification of plant microRNAs and their targets, including a stress-induced miRNA.** *Mol Cell* 2004, **14(6)**:787-799.
40. Alves L Jr, Niemeier S, Hauenschild A, Rehmsmeier M, Merkle T: **Comprehensive prediction of novel microRNA targets in *Arabidopsis thaliana*.** *Nucleic Acids Res* 2009, **37(12)**:4010-4021.
41. Krüger J, Rehmsmeier M: **RNAhybrid: microRNA target prediction easy, fast and flexible.** *Nucleic Acids Res* 2006, , **34 Web Server:** W451-W454.
42. Allen E, Xie Z, Gustafson AM, Carrington JC: **MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants.** *Cell* 2005, **121(2)**:207-221.
43. Megraw M, Baev V, Rusinov V, Jensen ST, Kalantidis K, Hatzigeorgiou AG: **MicroRNA promoter element discovery in *Arabidopsis*.** *RNA* 2006, **12(9)**:1612-1619.
44. Parizotto EA, Dunoyer P, Rahm N, Himber C, Voinnet O: **In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA.** *Genes Dev* 2004, **18(18)**:2237-2242.
45. Válóczi A, Várallyay E, Kauppinen S, Burgyán J, Havelda Z: **Spatio-temporal accumulation of microRNAs is highly coordinated in developing plant tissues.** *Plant J* 2006, **47(1)**:140-151.
46. 't Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX, Boer JM, van Ommen GJ, den Dunnen JT: **Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms.** *Nucleic Acids Res* 2008, **36(21)**:e141.
47. Gustafson AM, Allen E, Givan S, Smith D, Carrington JC, Kasschau KD: **ASRP: the *Arabidopsis* Small RNA Project Database.** *Nucleic Acids Res* 2005, , **33 Database:** D637-D640.
48. Moldovan D, Spriggs A, Yang J, Pogson BJ, Dennis ES, Wilson IW: **Hypoxia-responsive microRNAs and trans-acting small interfering RNAs in *Arabidopsis*.** *J Exp Bot* 2010, **61(1)**:165-177.
49. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome Biol* 2010, **11(10)**:R106.
50. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5(10)**: R80.
51. Hackenberg M, Rodríguez-Ezpeleta N, Aransay AM: **miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments.** *Nucleic Acids Res* 2011, , **39 Web Server:** W132-W138.
52. Dhahbi JM, Atamna H, Boffelli D, Magis W, Spindler SR, Martin DI: **Deep sequencing reveals novel microRNAs and regulation of microRNA expression during cell senescence.** *PLoS One* 2011, **6(5)**:e20509.
53. Pant BD, Musialak-Lange M, Nuc P, May P, Buhtz A, Kehr J, Walther D, Scheible WR: **Identification of nutrient-responsive *Arabidopsis* and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing.** *Plant Physiol* 2009, **150(3)**:1541-1555.
54. Kawashima CG, Yoshimoto N, Maruyama-Nakashita A, Tsuchiya YN, Saito K, Takahashi H, Dalmay T: **Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types.** *Plant J* 2009, **52(2)**:313-321.
55. Frith MC, Wan R, Horton P: **Incorporating sequence quality data into alignment improves DNA read mapping.** *Nucleic Acids Res* 2010, **38(7)**: e100.
56. Oshlack A, Wakefield MJ: **Transcript length bias in RNA-seq data confounds systems biology.** *Biol Direct* 2009, **4**:14.
57. Zhao CZ, Xia H, Frazier TP, Yao YY, Bi YP, Li AQ, Li MJ, Li CS, Zhang BH, Wang XJ: **Deep sequencing identifies novel and conserved microRNAs in peanuts (Arachis *hypogaea L.*).** *BMC Plant Biol* 2010, **10**:3.
58. Li H, Deng Y, Wu T, Subramanian S, Yu O: **Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation.** *Plant Physiol* 2010, **153(4)**:1759-1770.
59. Broughton WJ, Dilworth MJ: **Control of leghaemoglobin synthesis in snake beans.** *Biochem J* 1971, **125(4)**:1075-1080.

60. Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, Gentleman R: **ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data.** *Bioinformatics* 2009, **25(19)**:2607-2608.
61. **The National Center for Biotecnological information.** , [http://www.ncbi.nlm.nih.gov/projects/dbEST][http://www.ncbi.nlm.nih.gov/projects/dbGSS].

doi:10.1186/1471-2164-13-83
Cite this article as: Peláez *et al.*: Identification and characterization of microRNAs in *Phaseolus vulgaris* by high-throughput sequencing. *BMC Genomics* 2012 **13**:83.

9.2 Peláez et al., 2013

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 05 September 2013
doi: 10.3389/fpls.2013.00343

Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences

Pablo Peláez and Federico Sanchez*

Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

Edited by:
Vitaly Citovsky, State University of New York at Stony Brook, USA

Reviewed by:
Michael Goodin, University of Kentucky, USA
Herman B. Scholthof, Texas A&M University, USA

****Correspondence:***
Federico Sanchez, Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, 62210 Cuernavaca, Morelos, México
e-mail: federico@ibt.unam.mx

Small non-coding RNAs constitute an important class of gene expression regulators that control different biological processes in most eukaryotes. In plants, several small RNA (sRNA) silencing pathways have evolved to produce a wide range of small RNAs with specialized functions. Evidence for the diverse mode of action of the small RNA pathways has been highlighted during plant–microbe interactions. Host sRNAs and small RNA silencing pathways have been recognized as essential components of plant immunity. One way plants respond and defend against pathogen infections is through the small RNA silencing immune system. To deal with plant defense responses, pathogens have evolved sophisticated mechanisms to avoid and counterattack this defense strategy. The relevance of the small RNA-mediated plant defense responses during viral infections has been well-established. Recent evidence points out its importance also during plant–bacteria interactions. Herein, this review discusses recent findings, similarities and differences about the small RNA-mediated arms race between plants and these two groups of microbes, including the small RNA silencing pathway components that contribute to plant immune responses, the pathogen-responsive endogenous sRNAs and the pathogen-delivered effector proteins.

Keywords: RNA silencing, immunity, susceptibility, plant pathogens, small RNAs

INTRODUCTION

Immune responses against pathogens are present in most multicellular organisms. Plants and animals have evolved complex and effective innate immune systems to protect themselves from invading microorganisms (Ausubel, 2005). Plants are capable to fight pathogens through a multilayered innate immune system that accomplishes immune memory and self-tolerance (Spoel and Dong, 2012). The first line of defense against phytopathogens occurs when transmembrane pattern recognition receptors (PRRs) recognize microbial- or pathogen-associated molecular patterns (MAMPs or PAMPs) to trigger a general defense response referred to as PAMP-triggered immunity (PTI). Successful microbes deliver effector proteins into the host cells to suppress PTI resulting in effector-triggered susceptibility (ETS). In turn, many plant species have subsequently evolved a type of immunity triggered by resistance (R) proteins that responds to pathogen effector proteins and overcomes PTI arrest. This type of immunity is called effector-triggered immunity (ETI), and results in disease resistance, usually in consequence of a hypersensitive response (HR) at the infection site. In this co-evolutionary context, natural selection drives pathogens and plants to diversify their effector and resistance genes, respectively (Dangl and Jones, 2001; Jones and Dangl, 2006; Ronald and Beutler, 2010).

Host-encoded small RNAs and small RNA silencing pathway proteins impact factors involved in PTI and ETI (Jin, 2008; Padmanabhan et al., 2009). The small RNA classes work in plant defense responses by causing either post-transcriptional gene silencing (PTGS) or transcriptional gene silencing (TGS) to a set of host or pathogen genes (Baulcombe, 2004). The PTGS process mediated by sRNAs leads to messenger RNA (mRNA) cleavage or translational repression, while the TGS mechanism triggered by sRNAs regulates DNA methylation and histone modifications (Schramke and Allshire, 2004). The small RNA classes that have been identified in plants are: microRNAs (miRNAs), *trans*-acting small interfering RNAs (ta-siRNAs), natural antisense transcript-derived small interfering RNAs (nat-siRNAs), repeat-associated small interfering RNAs (ra-siRNAs) or heterochromatic small interfering RNAs (hc-siRNAs), and long small interfering RNAs (lsiRNAs; Chen, 2009; Vazquez et al., 2010). These regulatory small RNAs are produced through different pathways that share several conserved protein families like: the RNA-dependent RNA polymerases (RDRs), the double-stranded RNA-binding proteins (DRBPs), the Dicer-like proteins (DCLs), the small RNA methyltransferase (HEN1), and the Argonaute proteins (AGOs; Chapman and Carrington, 2007). In general, this group of proteins, together with double-stranded RNA (dsRNA) templates, is able to produce small RNAs with specific sizes to rearrange gene expression and to mount a response during plant–microbe interactions. In this review, the focus is on the differences and the similarities of the plant innate immunity involving host small RNAs, sRNA silencing pathway components, silencing suppressors and microbe-derived sRNAs that have been shown to play a specific role in the interaction of plants with bacteria and viruses.

SMALL RNA-MEDIATED DEFENSE RESPONSES IN PLANT–MICROBE INTERACTIONS

RNA silencing is an important innate immunity approach used by plants to counteract pathogens including nematodes, fungi, and

protists (Katiyar-Agarwal and Jin, 2010). Small RNA-mediated antiviral immunity in plants was observed since the beginnings of RNA silencing molecular characterization (Lindbo et al., 1993; Ratcliff et al., 1997; Wang et al., 2011a). PTGS constitutes a very important line of defense against viruses mainly because these are obligate intracellular pathogens whose life cycle depends on the host cellular functions. Once plant viruses introduce their genetic material of DNA or RNA in either single-stranded or double-stranded form, PTGS and/or TGS are induced to control virus replication and spreading throughout the plant. Similar to endogenous RNA silencing regulation, viral double-stranded RNA triggers the formation of virus-derived small interfering RNAs (vsiRNAs) by DCLs. vsiRNAs are then loaded into AGOs to direct viral DNA or RNA silencing. An important amplification step that promotes production of secondary vsiRNAs and supports the systemic silencing involves the activity of the RDRs. In a certain way, antiviral immunity accomplished through viral RNA silencing could be seen as the exploitation of the host life cycle dependency of viruses. Many biotechnological approaches have taken advantage of this natural occurring antiviral strategy to engineer virus-resistant plants. Interestingly, as usually occurs in plant–viruses interactions, viruses have developed a counter defense strategy based on inhibiting host sRNA-mediated antiviral responses with a diverse class of proteins denominated viral suppressors of RNA silencing (VSRs). An efficient viral suppressor may constitute the difference between an infected or immune plant. Increasing evidence suggests that VSRs may have a strong impact on small RNA silencing pathways and therefore in host endogenous sRNA-regulated processes. In fact, many symptoms of plants during viral diseases are attributed to changes in the regulation of endogenous genes caused by VSRs. Also, some vsiRNAs have been reported to target host genes that may be involved in plant defense responses.

Although RNA silencing seems to be a custom-made defense mechanism against viruses, it also stands out as an important mechanism that bacteria have to overcome in order to cause disease in plants. Host small RNAs have been observed to respond in either viral or bacterial infections to promote plant disease resistance. Many small RNA biogenesis factors have been implicated in regulating plant defense responses to pathogenic bacteria as well. Similar to vsiRNAs, dsRNA from genes of the transfer DNA (T-DNA) of *Agrobacterium tumefaciens* triggers the production of bacteria-derived siRNAs. Similar to viruses, bacteria secrete suppressors of RNA silencing to counteract plants RNA silencing defense responses. Most of the roles of RNA silencing in antibacterial immunity have been described in plant interactions with *Pseudomonas syringae* and *Agrobacterium tumefaciens*. In this section, current research on plant small RNA-mediated immunity against viruses and bacteria, including relevant pathogenic elements triggering plant defense is reviewed. Common similarities and particular differences related to antibacterial and antiviral small RNA-mediated immunity are highlighted.

HOST SMALL RNAs AND PLANT IMMUNITY AGAINST BACTERIAL AND VIRAL INFECTIONS

Several studies have shown that plant small RNAs are directly involved in bacterial disease responses. The first sRNAs identified to participate in plant immunity during bacterial infections were miRNAs. Plants tested with pathogenic bacteria showed several changes in miRNA accumulation, particularly in auxin signaling related miRNAs (Fahlgren et al., 2007; Jagadeeswaran et al., 2009; Zhang et al., 2011a). Virus-responsive sRNAs have been analyzed during very diverse plant–virus interactions (Kasschau et al., 2003; Chapman et al., 2004; Chen et al., 2004; Dunoyer et al., 2004; Chellappan et al., 2005; Takeda et al., 2005; Zhang et al., 2006; Bazzini et al., 2007, 2009; Bortolamiol et al., 2007; Csorba et al., 2007; Lim et al., 2007; Moissiard et al., 2007; Vogler et al., 2007; He et al., 2008; Azevedo et al., 2010; Naqvi et al., 2010; Perez-Quintero et al., 2010; Varallyay et al., 2010; Amin et al., 2011a,b; Du et al., 2011; Hu et al., 2011; Jay et al., 2011; Lang et al., 2011; Feng et al., 2012; Guo et al., 2012; Li et al., 2012; Pacheco et al., 2012; Shivaprasad et al., 2012; Yifhar et al., 2012). Nevertheless, in contrast to bacterial infections, direct evidence for the specific role of endogenous small RNAs in plant antiviral immunity has been limited due to disturbance of the biogenesis and the function of plant sRNAs by viruses. In several cases, VSRs globally induce or accumulate miRNAs through interaction with components of the miRNA biogenesis pathway that may affect miRNA production, function, or stabilization (Kasschau et al., 2003; Chen et al., 2004; Takeda et al., 2005; Bortolamiol et al., 2007; Csorba et al., 2007; Vogler et al., 2007; Azevedo et al., 2010; Varallyay et al., 2010; Shivaprasad et al., 2012). Similar to a bacterial suppressor that induced transcriptional repression of miR393, viral infections may also alter miRNAs at the transcriptional level (Bazzini et al., 2009). In addition, other viral proteins or satellite RNAs, not necessarily VSRs, may disturb miRNA accumulation (Bazzini et al., 2007; Feng et al., 2012). Several of these disturbances in sRNA accumulation have been correlated with plant symptoms of viral diseases, particularly for alterations of miRNA-target interactions involving the miR156, miR171, miR167, miR390, and miR173 families (Kasschau et al., 2003; Dunoyer et al., 2004; Moissiard et al., 2007; Jay et al., 2011). As a consequence of changes in miRNA accumulation and disturbance of RDR6 or DCL4 by VSRs, changes in accumulation of some miRNA-derived tasiRNAs have been also correlated with phenotypic changes during viral infections (Moissiard et al., 2007; Wang et al., 2010; Yifhar et al., 2012). Furthermore, viral infections can modify hc-siRNAs production and alter the RNA-directed DNA methylation (RdDM) pathway leading to reactivation of transposons and transcription of silenced genes that could negatively impact plant defense responses (Raja et al., 2008; Azevedo et al., 2010; Dowen et al., 2012). Despite the fact that viruses cause broad effects on hosts small RNAs, some specific miRNA families seem to be directly involved in antiviral immunity. Interestingly, several host small RNAs are involved in antibacterial or in antiviral immunity because they regulate R or pathogenesis-related (PR) genes (**Table 1**).

Perception of the PAMP flagellin by *Arabidopsis* was reported to restrict *P. syringae* invasion; however, no mechanism was known that could be involved in triggering this resistance. Analyzing gene expression profilings of seedlings challenged with flg22, Navarro et al. (2006) observed that the accumulation of three auxin receptor transcripts (*TIR1*, *AFB2*, *AFB3*), targets of miR393, was repressed upon treatment with flg22. However, one of the auxin receptors *AFB1* was not affected possibly due to a slightly

Table 1 | Plant small RNAs involved in plant immunity.

Small RNA	Target(s)	Host(s)	Pathogen(s)	Reference(s)
miR393	TIR1, AFB2, AFB3, AFB1	*Arabidopsis*	Bacteria	Navarro et al. (2006)
miR160	ARF10, ARF16, ARF17	*Arabidopsis*	Bacteria	Li et al. (2010b)
miR398	CSD1, CSD2, COX5	*Arabidopsis*	Bacteria	Jagadeeswaran et al. (2009); Li et al. (2010b)
miR773	DMT2	*Arabidopsis*	Bacteria	Li et al. (2010b)
miR168	AGO1	*Arabidopsis* and *N. benthamiana*	Viruses	Bortolamiol et al. (2007); Varallyay et al. (2010)
miR162	DCL1	*Arabidopsis*	Viruses	Zhang et al. (2006); Azevedo et al. (2010)
miR482/ miR2118	R genes	*N. benthamiana*	Viruses and Bacteria	Zhai et al. (2011); Shivaprasad et al. (2012)
miR158	PPR gene	*Brassica napus* and *Brassica rapa*	Viruses	He et al. (2008)
miR1885	TIR-NBS-LRR gene	*Brassica napus* and *Brassica rapa*	Viruses	He et al. (2008)
miR393*	MEMB12	*Arabidopsis* and *N. benthamiana*	Bacteria	Zhang et al. (2011b)
nta-miR6019	Receptor N	*N. tabacum*	Viruses	Li et al. (2012)
nta-miR6020	Receptor N	*N. tabacum*	Viruses	Li et al. (2012)
nat-siRNAATGB2	PPRL	*Arabidopsis*	Bacteria	Katiyar-Agarwal et al. (2006)
AtlsiRNA-1	AtRAP	*Arabidopsis*	Bacteria	Katiyar-Agarwal et al., 2007)

**Mature star miRNA.*

Host small RNAs with a role in antibacterial and/or antiviral immunity. The plant hosts where the small RNAs came from and/or where functional studies to characterize its roles in plant defenses were performed are specified.

different miR393-binding site. This result suggested a role for miR393 in regulating defense responses during *P. syringae* infection. Further experiments confirmed this role for miR393a since its overexpression enhanced plant bacteria resistance and as consequence reduced virulent *P. syringae* pv. tomato (Pst) DC3000 *in planta* growth (Navarro et al., 2006). Besides, expression of mi393a was induced in flg22-stimulated *Arabidopsis* seedlings. When the miR393-resistance auxin receptor *AFB1*-Myc was overexpressed in a *tir1-1* background, it resulted in enhanced disease susceptibility. Interestingly, both virulent Pst DC3000 and avirulent Pst DC3000 carrying the type III effector protein avrRpt2 showed similar growth under the *AFB1*-Myc-overexpressing plants. Accumulation of miR393 was also reported to be induced during infiltration with *Agrobacterium tumefaciens*, the plant pathogen that causes crown gall disease (tumor formation) by transferring bacterial DNA to the plant genome. Both, disarmed and oncogenic strains induced miR393 in an early stage of infection. Interestingly, the flg22 of *Agrobacterium tumefaciens* that is completely inactive to the receptor kinase FLAGELLIN INSENSITIVE2 (*FLS2*) as a ligand, maintained unaltered miR393 expression levels (Pruss et al., 2008). Together, these results suggest that miR393a is clearly involved in ETI and, most importantly, that repression of auxin signaling constitutes a plants defense response to bacterial infection.

Following high-throughput sequencing analyses have supported the upregulation of miR393 during plant–bacteria interactions (Fahlgren et al., 2007; Li et al., 2010b; Zhang et al., 2011b). In addition to miR393, the small RNAs miR160 and miR167, regulators of auxin response factors (ARFs) and also members of the auxin signaling pathway were induced after non-pathogenic Pst DC3000 *hrcC*$^{-}$ inoculation and flg22 treatment. Surprisingly, treatment with flg22 did not show a significant reduction of miR167 targets *ARF8* and *ARF6*, in contrast to the downregulation of miR160 targets *ARF10*, *ARF16*, *ARF17*. Plants overexpressing miR160a that were treated with flg22 and *hrcC*$^{-}$ mutant bacteria increased callose deposition. Nevertheless, in miR160 overexpressing plants, resistance to Pst DC3000 bacteria proliferation was not affected (Li et al., 2010b).

As well as increasing accumulation of hormone-related miRNAs contributes to plant defense, downregulation of certain miRNAs impact plant immunity (Sunkar et al., 2006; Jagadeeswaran et al., 2009; Li et al., 2010b). For instance, miR398, whose targets are two copper superoxide dismutases (*CSD1*,*CSD2*) and a cytochrome c oxidase subunit V (*COX5*), is reduced only in plants challenged with avirulent strains such as Pst DC3000 *avrRpm1* and Pst DC3000 *avrRpt2* (Jagadeeswaran et al., 2009). Accumulation of miR398 is also changed by abiotic and biotic stresses such as salinity, increased light, increased Cu^{2+} and Fe^{3+}, ozone stress, and flg22 treatment (Sunkar et al., 2006; Jagadeeswaran et al., 2009; Li et al., 2010b). During biotic and abiotic stresses plants induce early and rapid accumulation of reactive oxygen species (ROS) in the infection zone. Superoxide dismutases (SOD) enzymes process superoxide into oxygen and hydrogen peroxide and therefore regulate ROS. Expression of miR398 is reduced in oxidative stress, promoting accumulation of *CSD1* and *CSD2* (Sunkar et al., 2006). In accordance with expression analyses in different stresses, overexpression of miR398 reduced callose deposition after flg22 treatment and Pst DC3000 *hrC*$^{-}$ infection. Moreover, these transgenic plants were more susceptible to virulent and avirulent strains of *P. syringae* as consequence of *CSD1*, *CSD2* and *COX5* gene silencing (Li et al., 2010b). These last observations confirmed a link between the miR398 family and miRNA-mediated plant defense responses.

Another reported miRNA that is involved in PTI is miR773 (Li et al., 2010b). This miRNA targets the mRNA coding for the DNA methyltransferase 2 (*DMT2*). RNAi-mediated gene silencing of *DMT2*, and a different DNA methyltransferase (*DMT1*), results in reduced tumor formation during *Agrobacterium* infection (Crane and Gelvin, 2007). Interestingly, *MET1* is downregulated in response to biotic and salicylic acid (SA) stresses (Dowen et al., 2012). Deep sequencing analysis of AGO1-immunoprecipitated small RNAs revealed a reduction in miR773 accumulation after flg22 treatment. In accordance, accumulation of the miRNA target *DMT2* was induced in response to flg22 treatment. Transgenic plants overexpressing miR773 showed reduced *MET2* mRNA levels, reduced callose deposition and enhanced susceptibility to the Pst DC3000 and Pst DC3000 *hrC*$^{-}$ strains. Reduction of miR398 and miR773 upon biotic stress exemplifies a negatively regulation of PTI (Li et al., 2010b).

Maintenance of AGO1 homeostasis in plants depends on the *AGO1*-mediated stabilization of miR168 and on the miR168-mediated cleavage of *AGO1* mRNA guided by the same AGO1 protein (Rhoades et al., 2002). Viral infections in plants induce the accumulation of both the *AGO1* mRNA and the miR168 mature miRNA (Bortolamiol et al., 2007; Varallyay et al., 2010). Induction of *AGO1* mRNA during viral infections is considered a plant defense response, while accumulation of miR168 is a viral counterstrategy. Increasing *AGO1* mRNA is a recurrent plant defense reaction toward viral infections because AGO1 protein guides vsiRNAs against viral RNAs. Interestingly, induction of miR168 was spatially correlated with *Tombusvirus* accumulation carrying the p19 suppressor. The specific role of miR168 in plant–virus interactions was further supported when the viral suppressor p19 was removed from this virus and no accumulation of miR168 was observed. Failure of viruses to induce miR168 accumulation indeed promotes accumulation of AGO1 protein and results in a stronger antiviral response (Varallyay et al., 2010).

The miRNA miR162 is also a miRNA that regulates an important component of the miRNA biogenesis pathway so, it is therefore not surprising that this miRNA constitutes a relevant sRNA in antiviral responses as well (Zhang et al., 2006; Azevedo et al., 2010). The mRNA of *DCL1* is regulated by miR162 loaded into AGO1 (Xie et al., 2003). In plant–virus interactions, several VSRs impaired AGO1 activity enhancing *DCL1* mRNA accumulation (Xie et al., 2003; Zhang et al., 2006; Azevedo et al., 2010). Unexpectedly, the accumulation of *DCL1* in *Arabidopsis* plants infected with the *Turnip crinkle virus* (TCV; carrying the P38 viral suppressor) favors reduction of DCL4 and DCL3 two important DCLs in charge of producing vsiRNAs (Azevedo et al., 2010). Although miR162 accumulation in plants infected with the TCV was reduced, it was determined that transcriptional enhancement of *DCL1* is mainly a consequence of inhibition of AGO1 activity. In this sense, the antiviral role of miR162 is coupled with disturbance of AGO1 activity and involves regulation of DCLs. It would be interesting to determine if miR838, other miRNA that regulates *DCL1*, could also be related to antiviral immunity (Rajagopalan et al., 2006).

Ultimately, a group of miRNA families have been reported in legumes and tomato to be directly involved in ETI by regulating many R genes of the NBS-LRR class (Zhai et al., 2011; Shivaprasad et al., 2012). The NBS-LRR proteins activate plant defense responses through recognition of microbe effectors to initiate usually a race-specific ETI. The conserved miRNA families reported to regulate many of these genes are mainly three: miR1507, miR2109, and the superfamily miR482–miR2118 (due to similar sequence members). Most of the members of these families have been detected in legumes. Several of the R target genes identified in these studies produced phased siRNAs as consequence of the 22-nt size of the miRNA regulator members. In some cases, the phased loci (NBS-LRR genes) producing secondary siRNAs presented the "two-hit" model of two miRNA target sites in the mRNA. Identification of phased R loci in *M. truncatula*, soybean, and tomato showed variability in the number of regulated NBS-LRR genes by these miRNA families between species. Based on the constitutive accumulation of some members of these families in tomato and legumes, it is considered that most of the reported miRNA-regulated NBS-LRR targets are silenced in absence of a pathogenic agent (Zhai et al., 2011; Shivaprasad et al., 2012). In this regard, Shivaprasad et al. (2012) found that miRNA-mediated silencing of two disease resistance mRNAs by the miR482–miR2118 superfamily was decreased in plants infected with *P. syringae* DC3000. Similar to bacterial infections, leaves of tomato inoculated with the TCV, the *Cucumber mosaic virus* (CMV) and the *Tobacco rattle virus* (TRV) presented reduced accumulation of miR482. In accordance, two miR482 NBS-LRR mRNA targets were induced in infected plants, especially in plants infected with TCV. A secondary siRNA, product of one of these phased resistance loci that targets a defense-related mRNA and forms a sRNA regulatory cascade, was also suppressed during viral and bacterial infections. Taking into account these observations, they proposed that miRNA-regulated R genes might participate in an uncommon non-race-specific immunity mechanism when this small RNA regulation is blocked by pathogen-encoded suppressors of RNA silencing to release the defense resistance targets. Although further experiments are necessary to extensively validate this hypothesis, previous observations suggest this could be a possible defense mechanism. For example, Li et al. (2012) established other cases for miRNA regulation of innate immune receptors (NBS-LRR genes) in tobacco. Likewise, they showed that increased miRNA repression of a specific R gene attenuates resistance to the *Tobacco mosaic virus* (TMV) in *Nicotiana benthamiana*. Additionally, it was previously reported that overexpression of miR482 miRNA resulted in hypernodulated soybeans, that miR482 is induced after *Bradyrhizobium japonicum* inoculation, and that miR1507 accumulates upon rhizobia infection in the roots of a supernodulated mutant (Li et al., 2010a).

Other small RNAs that are involved in ETI during virus infections are miR158 and miR1885. Inoculation assays performed in *Brassica napus* and *Brassica rapa* with the *Turnip mosaic virus* (TuMV) revealed an enhanced accumulation of miR158 and miR1885 (He et al., 2008). Surprisingly, these miRNAs are generated from the same precursor. The sRNA miR1885 was predicted to target TIR-NBS-LRR class disease-resistant transcripts in *Brassica*, while miR158 targets mRNAs of pentatricopeptide repeat (PPR) containing proteins in *Arabidopsis*. However, further studies are still needed to determine the precise role of these miRNAs

in antiviral immunity besides the identity of their targets because it is well-known that the viral suppressor of TuMV HC-Pro increased miRNA accumulation and inhibits miRNA cleavage function (Kasschau et al., 2003).

In the beginning of miRNAs characterization, it was conceived that only one of the strands of the sRNA duplexes was functional and therefore was selected to be loaded into the RNA-induced silencing complex (RISC) to regulate gene expression. Now, several studies have confirmed that the opposite strand (star or passenger strand) of a defined miRNA could also be functional in plants and animals. In relation with small RNA-mediated plant defense responses, the star strand miR393* has also been shown to play a role in plant immunity (Zhang et al., 2011b). In *Arabidopsis*, AGO2 is strongly induced after *P. syringae* pv. tomato contact. One of the star strands abundantly loaded into AGO2 after Pst *avrRpt2* treatment was miR393*. Among three predicted targets for this miRNA, one target experimentally validated for this miRNA, and predicted to have a function in vesicle transport, is *MEMB12*, a Golgi-localized SNARE protein. The knockout mutant of *memb12* exhibit enhanced resistance to avirulent and virulent strains of *P. syringae*. The plant secretory machinery has been proposed as a critical mechanism in plant–microbe interactions due to its capacity to secrete antimicrobial proteins and several biomolecules. During loss-of-function of MEMB12, it was observed that the major secreted antimicrobial pathogenesis-related protein PR1 was strongly secreted. This result suggests that enhanced resistance conferred by the mutant *memb12* is a consequence of the accumulation and secretion of PR proteins. As expected, overexpression of miR393* also showed enhanced disease resistance to Pst *avrRpt2* and also presented increased accumulation and secretion of the PR1 protein (Zhang et al., 2011b).

Lately, two miRNAs named nta-miR6019 and nta-miR6020 were reported to target transcripts of the TIR-NB-LRR immune receptor *N* in tobacco. The immune receptor *N* was the first virus-related resistance gene identified. It confers resistance to the TMV. Expression of these two miRNAs in *N. Benthamiana* was found to reduce *N*-mediated resistance to the TMV. Interestingly, the 22-nt miRNA nta-miR6019 triggers production of secondary siRNAs from the *N* gene similar to the miR482 miRNA family. In this study, eight more miRNA families from tobacco, tomato, and potato that may target R genes were identified, including the miR482 family (Li et al., 2012).

Just as miRNAs contribute to plant immunity, other host siRNAs also promote gene expression reorganization during plant defense responses. These small RNAs are induced in response to pathogenic bacteria and are implicated in triggering plant disease resistance. As mentioned before, the type of siRNAs that have been described to participate in plant immune responses are the nat-siRNA nat-siRNAATGB2 and the bacteria-induced long siRNAs (mainly AtlsiRNA-1; Katiyar-Agarwal et al., 2006, 2007).

Previous to the discovery of nat-siRNAATGB2, siRNAs were known to be involved in antiviral defense mechanisms in plants and animals. Even so, there was a lack of information regarding siRNA-mediated regulation of antibacterial defenses in plants. In this context, finding of nat-siRNAATGB2 revealed, for the first time, the importance of siRNAs and, particularly nat-siRNAs, in controlling antibacterial immunity in plants (Katiyar-Agarwal et al., 2006). This siRNA is generated from the overlapping region of two NATs: *ATGB2* (Rab2-like small GTP-binding protein gene) and *PPRL* (pentatricopeptide repeat-like gene). It targets the 3′ UTR region of *PPRL*. In plants infected with Pst (*avrRpt2*), nat-siRNAATGB2 strongly and specifically accumulates. Surprisingly, this siRNA has a role in disease resistance against pathogenic bacteria by silencing the *PPRL* gene which negatively regulates the coiled-coil NBS-LRR type R protein RPS2. Induction of this siRNA also depends on the disease resistance gene *RPS2* and on the *NDR1* gene that is required for RPS2-specified resistance. Plants overexpressing a resistant version of *PPRL* (without UTR) against nat-siRNAATGB2 that were inoculated with Pst *avrRpto* showed delayed HR, reduced level of cell death and enhanced pathogen growth. All these results suggest that nat-siRNAATGB2 plays a positive role in disease resistance through regulation of *PPRL*.

In an effort to identify other small RNAs specifically induced in a pathogen interaction (*P. syringae*), Katiyar-Agarwal et al. (2007) discovered a class of small RNAs (lsiRNAs). Of the six lsiRNAs discovered, five are induced in response to Pst (*avrRpt2*) infection. The most functionally characterized lsiRNA is AtlsiRNA-1. This siRNA is generated from the overlapping region of the *SRRLK* (putative leucine-rich repeat receptor-like protein kinase) and *atRAP* (an expressed protein that contains a putative RNA-binding domain) natural antisense transcripts. AtlsiRNA-1 is complementary to the 3′ UTR of the antisense gene *AtRAP* and therefore regulates its expression. In mutant *atrap* plants less growth of both virulent and avirulent Pst was observed, suggesting a negative regulatory role for this gene in plant resistance responses. Based on these results, AtlsiRNA-1 may promote resistance against Pst *avrRpt2* infection because of the particular regulation of its target. Almost all the discovered AtlsiRNAs still have to be confirmed to play a role in plant defense responses. Nonetheless, taking into account the cases exposed here of siRNAs related to plant immunity during bacterial infections, together with the growing knowledge generated around small RNAs, for example Pol IV-dependent siRNAs (p4-siRNAs), more siRNAs could also be involved in regulating bacteria stress responses.

ANTIVIRAL AND ANTIBACTERIAL ROLES OF THE SMALL RNA BIOGENESIS FACTORS

Previously described sRNAs that play important roles in bacterial and viral infections are produced by diverse small RNA silencing pathways. Some small RNA biogenesis factors are directly involved in plant immunity against pathogens (**Table 2**). Consequently, the sRNA silencing pathway components may be altered during bacterial infection to precisely affect production of the sRNAs that contribute to plant resistance. Besides, many components of the small RNA silencing pathways involved in PTGS form the antiviral RNA silencing defense mechanism. Functional studies have assigned a concise or a redundant function in antiviral immunity to various elements of the sRNA silencing pathways. The importance of some of these elements in regulating antiviral plant defenses is reflected in the attempt of many viral suppressors to disturb their activity. Coincidentally, some of these small RNA biogenesis factors have a dual role in plant defense responses against both bacterial and viral infections.

Table 2 | Small RNA biogenesis factors involved in plant defense.

Protein	Bacteria	Viruses	Reference(s)
DCL1	✓	✓	Dunoyer et al. (2006); Katiyar-Agarwal et al. (2006, 2007); Navarro et al. (2008); Qu et al. (2008); Azevedo et al. (2010); Li et al. (2010b)
HEN1	✓	✓	Boutet et al. (2003); Blevins et al. (2006); Katiyar-Agarwal et al. (2006, 2007); Navarro et al. (2008); Jamous et al. (2011); Zhang et al. (2012)
RDR6	✓	✓	Mourrain et al. (2000); Muangsan et al. (2004); Qu et al. (2005, 2008); Schwach et al. (2005); Adenot et al. (2006); Dunoyer et al. (2006); Katiyar-Agarwal et al. (2006, 2007); Donaire et al. (2008); Garcia-Ruiz et al. (2010); Wang et al. (2010, 2011b)
AGO1	✓	✓	Morel et al. (2002); Mi et al. (2008); Qu et al. (2008); Azevedo et al. (2010); Li et al. (2010b); Wang et al. (2011b); Garcia et al. (2012)
AGO2	✓	✓	Harvey et al. (2011); Jaubert et al. (2011); Scholthof et al. (2011); Wang et al. (2011b); Zhang et al. (2011b, 2012); Carbonell et al. (2012)
AGO4	✓	✓	Agorio and Vera (2007); Raja et al. (2008); Bhattacharjee et al. (2009); Duan et al. (2012); Hamera et al. (2012)
AGO7	✓	✓	Katiyar-Agarwal et al. (2007); Qu et al. (2008)
Pol V	✓		López et al. (2011)
AGO5		✓	Takeda et al. (2008)
DCL4	✓	✓	Blevins et al. (2006); Bouché et al. (2006); Deleris et al. (2006); Katiyar-Agarwal et al. (2007); Qu et al. (2008); Garcia-Ruiz et al. (2010)
DCL3		✓	Qu et al. (2008); Raja et al. (2008); Azevedo et al. (2010)
DCL2		✓	Xie et al. (2004); Bouché et al. (2006); Deleris et al. (2006); Qu et al. (2008); Garcia-Ruiz et al. (2010); Zhang et al. (2012)
RDR1		✓	Xie et al. (2001); Yu et al. (2003); Donaire et al. (2008); Garcia-Ruiz et al. (2010); He et al. (2010); Wang et al. (2010)
RDR2		✓	Donaire et al. (2008); Garcia-Ruiz et al. (2010)
SDE3		✓	Dalmay et al. (2001); Garcia et al. (2012)
SDE5		✓	Hernandez-Pinzon et al. (2007)
SGS3	✓	✓	Muangsan et al. (2004); Adenot et al. (2006); Katiyar-Agarwal et al. (2006, 2007)
HYL1	✓		Katiyar-Agarwal et al. (2006, 2007)

Small RNA biogenesis factors involved in plant immunity against viruses and bacteria. Check mark (✓) denote its participation in plant defense responses during viral or bacterial interactions.

Analyses performed in *Arabidopsis thaliana* of mutants corresponding to the sRNA biogenesis factors have provided different clues about their function in plant immunity. DCL1 and HEN1 have been observed to play important roles in PTI and ETI. Growth of Pst DC3000 *hrcC*$^{-}$, a strain capable to trigger PTI but unable to fight it, was enhanced in *dcl1-9* and *hen1-1* mutants. Also, in these mutants, the *P. syringae* pv. phaseolicola (Psp) strain NPS3121, which does not infect *Arabidopsis*, the non-pathogenic *P. fluorescens* Pf-5 and the *E. coli* W3110 strains showed enhanced growth (Navarro et al., 2008). Besides, *dcl1-9* mutants pretreated with flg22 failed to increase resistance to Pst DC3000 and to induce callose deposition (Li et al., 2010b). The impact of DCL1 and HEN1 in plant immunity also has been tested during *Agrobacterium tumefaciens* infections. Roots and stems of *dcl1* and *hen1* mutants were immune to infection. This means that no tumor growth was observed and that tumor induction by *Agrobacterium tumefaciens* probably requires miRNA adequate functioning. Conversely, *rdr6* mutants were reported to be more susceptible to *Agrobacterium tumefaciens* (Dunoyer et al., 2006). DCL1, HYPONASTIC LEAVES 1 (HYL1), HEN1, SUPPRESSOR OF GENE SILENCING 3 (SGS3), and RDR6 are involved also in plant defense responses due to their role in the biogenesis of the NAT-siRNA nat-siRNAATGB2 and the bacteria-induced lsiRNA AtlsiRNA-1 (Katiyar-Agarwal et al., 2006, 2007). The accumulation of AtlsiRNA-1 was also reduced in a *dcl4-2* mutant so, DCL4 is considered to function in antibacterial plant responses (Katiyar-Agarwal et al., 2007).

As expected, the AGO proteins, the most important components of the RISC, participate in plant resistance to bacteria as well. AGO1 is required in the seedling growth inhibition process determined by flg22 treatment. Also, as a result of analyzing the expression of two PAMP-response genes, it was proposed that AGO1 is necessary for flg22-induced gene expression. The *ago1-25* and the *ago1-27* mutants, like the *dcl1-9* mutants, showed reduced callose deposition and failed to increase resistance to Pst DC3000 (Li et al., 2010b). Besides, AGO1 contributes to antibacterial responses by loading miR393 into the RISC (Mi et al., 2008). AGO7 also plays an antibacterial function through its participation in the biogenesis of AtlsiRNA-1 (Katiyar-Agarwal

et al., 2007). An *ago7* mutant resulted in increased susceptibility to Pst (*avrRpt2*; Li et al., 2010b). The argonaute protein AGO2 provides antibacterial resistance by carrying miR393b* to regulate exocytosis of antimicrobial PR proteins. AGO2 is highly induced by Pst to function in innate immune responses. Furthermore, the *ago2-1* mutant displayed enhanced susceptibility to Pst (*avrRpt2* and *EV*; Zhang et al., 2011b).

Small RNA biogenesis factors that are involved in the biogenesis of hc-siRNAs and in RdDM are also essential for antibacterial resistance. The *ago4-2* mutant in *Arabidopsis thaliana* showed enhanced susceptibility to the pathogenic bacteria Pst DC3000, to the avirulent bacteria Pst DC3000 (*avrRpm1*) and to the non-host pathogen *P. syringae* pv. tabaci (Agorio and Vera, 2007). Interestingly, loss-of-function of other components of the RdDM pathway working upstream and downstream of AGO4 did not affected susceptibility to Pst DC3000. Moreover, pol V mutants showed enhanced disease resistance against *P. syringae* DC3000 and enhanced SA-mediated defense responses (López et al., 2011). These results suggest that the RdDM pathway may regulate antibacterial immune responses in plants.

Argonaute proteins in *Arabidopsis* play important roles in plant resistance to viruses as well (Morel et al., 2002; Zhang et al., 2006, 2012; Pantaleo et al., 2007; Qu et al., 2008; Raja et al., 2008; Takeda et al., 2008; Bhattacharjee et al., 2009; Azevedo et al., 2010; Harvey et al., 2011; Jaubert et al., 2011; Scholthof et al., 2011; Wang et al., 2011b; Carbonell et al., 2012; Duan et al., 2012; Garcia et al., 2012; Hamera et al., 2012). AGO1 is a key component in this response. AGO1 is able to load vsiRNAs and target viral RNAs (Azevedo et al., 2010; Garcia et al., 2012). It is a protein highly targeted by VSRs to inhibit its cleavage activity or promote its degradation (Bortolamiol et al., 2007; Csorba et al., 2007; Azevedo et al., 2010). As expected, susceptibility to viruses was increased in *ago1* mutants (Morel et al., 2002). Also, AGO1 participates in the removal of viral RNAs, mainly those with more compact secondary structures (Qu et al., 2008). In addition, as previously mentioned, AGO1 constitutes a master regulator of a complex network that involves DCLs and AGO2 (regulated by miR403 in an AGO1-dependent manner; Azevedo et al., 2010). Interestingly, AGO1 and AGO2 act together in a non-redundant way against the CMV lacking the 2b suppressor downstream the biogenesis of secondary vsiRNAs (Wang et al., 2011b). Actually, important antiviral roles for AGO2 have been lately described. Mutants of *ago2* were hyper-susceptible to the TCV and to the CMV (Harvey et al., 2011). During infections with these two viruses, wild plants showed an increase in AGO2 protein. In view of the interference of AGO1 activity caused by the viral suppressors of these two viruses, it is considered that AGO2 may constitute a second layer of defense against viruses. The VSR of the TCV also interact directly with AGO2 (Zhang et al., 2012). The catalytic activity of AGO2 was essential for local and systemic antiviral resistance against the TuMV (Carbonell et al., 2012). In addition, AGO2 resulted to be a key element for non-host resistance toward the *Potato virus X* (PVX; Jaubert et al., 2011). Interestingly, PVX which commonly does not infect *Arabidopsis,* was able to infect the plant in presence of the *Pepper ringspot virus* (PepRSV) that carries a viral suppressor that probably targets AGO2. In *N. benthamiana*, AGO2 plays an important antiviral role against the *Tomato bushy stunt virus* (TBSV; Scholthof et al., 2011). Besides AGO1 and AGO2, antiviral activity has been assigned also to AGO5, AGO7, and AGO4 (Qu et al., 2008; Raja et al., 2008; Takeda et al., 2008; Bhattacharjee et al., 2009; Duan et al., 2012; Hamera et al., 2012). The AGO7 protein participates in the removal of viral RNAs with less structured secondary structure (Qu et al., 2008). In the case of AGO4, viral suppressors may affect its activities to suppress its antiviral roles involving viral DNA methylation (Raja et al., 2008; Bhattacharjee et al., 2009; Duan et al., 2012; Hamera et al., 2012).

All four *Arabidopsis* DCLs perform an antiviral activity against different kind of viruses (Blevins et al., 2006; Deleris et al., 2006; Moissiard and Voinnet, 2006; Qu et al., 2008; Garcia-Ruiz et al., 2010). Their functions result essential for antiviral defense responses. DCLs may act in a hierarchical and redundant way (Blevins et al., 2006; Deleris et al., 2006; Moissiard et al., 2007). DCL4 is considered the most important DCL enzyme and the first to act against viruses in diverse interactions (Blevins et al., 2006; Bouché et al., 2006; Deleris et al., 2006; Qu et al., 2008; Garcia-Ruiz et al., 2010). In this sense, viral suppressors may affect directly or indirectly DCL4 activity to avoid this first layer of defense (Deleris et al., 2006; Qu et al., 2008). Mutants of *dcl4* altered local and systemic antiviral immunity (Garcia-Ruiz et al., 2010). DCL4 also promotes secondary siRNAs production via transitivity (Moissiard et al., 2007). DCL4-dependent virus-derived siRNAs have been described as necessary and sufficient to control a virus without its VSR (Garcia-Ruiz et al., 2010). DCL2 plays diverse roles in antiviral defense responses as well. DCL2 is the major backup during reduction of DCL4 activity by viruses. In certain viral infections and tissues, DCL2 may accomplish antiviral immunity in *dcl4* mutants (Xie et al., 2004; Bouché et al., 2006; Deleris et al., 2006; Qu et al., 2008; Garcia-Ruiz et al., 2010). Plants involving *DCL4* and *DCL2* mutants strongly overaccumulated viral RNAs. DCL2 was also required for systemic antiviral immunity in inflorescences and transitivity (Bouché et al., 2006; Deleris et al., 2006). Crucial roles in regulating viral infection against the TCV were assigned for DCL2. Higher temperature upregulates DCL2 activity and allows resistance to this virus (Zhang et al., 2012). Viral RNA levels were increased in *dcl3* mutants, albeit being considered to have a minor role in antiviral immunity against RNA viruses (Qu et al., 2008). Moderately enhanced susceptibility in *dcl3* plants challenged with the *Beet curly top virus* (BCTV) and the *Cabbage leaf curl virus* (CaLCuV) was observed (Raja et al., 2008). Mostly, DCL3 has antiviral roles against DNA viruses and presumably by inducing viral DNA methylation. Finally, as mentioned before, *DCL1* has an indirect antiviral immunity role acting as a negative regulator of DCL4 and DCL3 (Qu et al., 2008; Azevedo et al., 2010). Nevertheless, DCL1 also acts as a positive regulator in the production of virus-derived siRNAs (Blevins et al., 2006; Moissiard and Voinnet, 2006).

The DRBPs interact with DCLs to produce small RNAs. For this reason, DRBPs were analyzed to determine its possible roles in antiviral immunity (Curtin et al., 2008). The only DRB protein found to be important for antiviral defenses is DRB4. The DRB4 protein interacts with DCL4 *in vivo* to facilitate production of *trans*-acting siRNAs. In viral infections, the viral suppressor P6 of the *Cauliflower mosaic virus* (CaMV) was shown to interact with DRB4 to accomplish its replication. Furthermore, plants expressing P6 induce similar symptoms as the *drb4* mutant (Haas

et al., 2008). In plants infected with the TCV, *drb4* mutants showed increased viral RNA levels (Qu et al., 2008). In addition, the role of DRB4 has been explored in *Arabidopsis* plants infected with the *Turnip yellow mosaic virus* (TYMV). *DRB4* was reported to be induced upon infection with this virus and it was observed that this protein controls viral coat protein accumulation. Besides, DRB4 was found to interact *in vitro* with a RNA translational enhancer of the TYMV, suggesting a role for DRB4 in repressing viral RNAs at a translational level (Jakubiec et al., 2012).

The small RNA methyltransferase HEN1 is another factor required for PTGS with antiviral functions. Mutant plants in *HEN1* were more susceptible to CMV and TCV infection and significantly accumulated their viral RNAs (Boutet et al., 2003; Zhang et al., 2012). Furthermore, HEN1 was found to participate in VIGS and in spreading RNA silencing in new growth through methylation of viral siRNAs. Interestingly, methylation of viral siRNAs by HEN1 was affected during the *Oilseed rape mosaic virus* (ORMV) infection, suggesting the presence of a viral suppressor with this activity in ORMV (Blevins et al., 2006). In fact, the *Zucchini yellow mosaic virus* (ZYMV) HC-Pro suppressor inhibits the activity of HEN1 *in vitro* (Jamous et al., 2011).

In *Arabidopsis*, the SILENCING DEFECTIVE 3 (*SDE3*) gene was found to disturb PTGS. SDE3 encodes a RNA helicase-like protein. Infection of *sde3* plants with CMV increased accumulation of viral RNA and severity of symptoms (Dalmay et al., 2001). The SDE3 protein has RNA helicase and AGO-binding functions that are essential for silencing a green fluorescent protein (GFP)-tagged PVX. Further, SDE3 act downstream of RDR6, and together with AGO1 and AGO2, promotes the production of secondary siRNAs (Garcia et al., 2012). In addition, the SILENCING DEFECTIVE 5 (*SDE5*) gene, a putative homolog of a human mRNA export factor, is involved in antiviral immunity. Mutants of *sde5* were hypersusceptible to the CMV but not to the TuMV (Hernandez-Pinzon et al., 2007).

The *SGS3* has also a role in antiviral defense. The *sgs3* mutant plants exhibited enhanced susceptibility to CMV and retarded viral-induced symptoms. An hypomorphic *sgs3* mutant also overaccumulated CMV RNA (Adenot et al., 2006). Susceptibility to TuMV or to the *Turnip vein-clearing virus* (TVCV) infections was not altered in *sgs3* plants (Mourrain et al., 2000). Moreover, *sgs3* mutants increased severity of symptoms and viral DNA levels when using a DNA VIGS vector derived from the CaLCuV. Also, DNA VIGS required SGS3 (Muangsan et al., 2004).

To mount an adequate systemic antiviral defense response the primary vsiRNAs pool is amplified through the production of secondary vsiRNAs by the RDRs. For that reason, RDR1, RDR2, and RDR6/SGS2/SDE1 have been highlighted as very important and crucial factors in antiviral plant defense responses in different plant species (Mourrain et al., 2000; Yu et al., 2003; Muangsan et al., 2004; Yang et al., 2004; Qu et al., 2005, 2008; Schwach et al., 2005; Adenot et al., 2006; Diaz-Pendon et al., 2007; Donaire et al., 2008; Garcia-Ruiz et al., 2010; He et al., 2010; Wang et al., 2010, 2011b). Although in certain cases, some RDRs mutants have no effects on the susceptibility to viruses carrying VSRs (Diaz-Pendon et al., 2007; Donaire et al., 2008). RDR1 was induced upon infection with TMV in *Arabidopsis* and in tobacco. *Arabidopsis rdr1* plants infected with TMV showed increased levels of viral RNAs and enhanced susceptibility locally and systemically (Xie et al., 2001; Yu et al., 2003). Also, *rdr1* plants accumulated higher levels of viral RNA from a suppressor defective CMV strain (Wang et al., 2010). Tobacco plants expressing antisense RNA for *RDR1* also exhibit enhanced susceptibility to TMV and PVX (Xie et al., 2001). *RDR1* was induced in maize when infected with the *Sugarcane mosaic virus* (SCMV). Silenced *RDR1* maize plants were more susceptible to the infection of SCMV and accumulated more viral RNA (He et al., 2010). Interestingly, RDR1 was induced by SA (Xie et al., 2001; Yu et al., 2003). Double or triple mutants of RDRs, necessarily including *RDR1,* altered suppressor defective TuMV and CMV infections (Garcia-Ruiz et al., 2010; Wang et al., 2010). In the case of RDR2, it was showed that it contributes together with RDR1 and RDR6/SGS2/SDE1 to deal with TRV and suppressor defective TuMV infection (Donaire et al., 2008; Garcia-Ruiz et al., 2010). On the other hand, *rdr6/sgs2/sde1* plants exhibited enhanced susceptibility to CMV and to TCV (Mourrain et al., 2000; Qu et al., 2008). Also, this RDR was important for DNA and RNA VIGS (Muangsan et al., 2004). Hypomorphic *rdr6* mutants also over-accumulate CMV RNA (Adenot et al., 2006). *RDR6*-silenced *N. benthamiana* plants resulted to be hypersusceptible to different viruses and temperature-dependent (Qu et al., 2005; Schwach et al., 2005). Furthermore, RDR6/SGS2/SDE1 participated in limiting systemic infection of a suppressor defective TuMV (Garcia-Ruiz et al., 2010). Likewise, RDR6 contributed together with RDR1 in reducing viral RNAs of a suppressor defective CMV strain as well (Wang et al., 2010). However, resistance to a different suppressor defective CMV strain was attributed mainly to RDR6 (Wang et al., 2011b). RDR6 also participates with RDR1 and RDR2 in antiviral defenses against TRV (Donaire et al., 2008).

BIOGENESIS AND ROLES OF PATHOGEN-DERIVED siRNAs

Viral double-stranded RNAs are used by plants to produce vsiRNAs of ∼20 to 24 nucleotides in length using several components of the silencing pathways to guide silencing of viruses genetic material. As previously mentioned, plants may also use transferred DNA from bacteria as template to produce bacteria-derived siRNAs (Dunoyer et al., 2006). Processing of viral dsRNAs by Dicer-like enzymes is not sufficient to block virus replication. Interestingly, several studies have indicated that vsiRNAs may regulate host gene expression. Although the identification of sRNAs derived from viral RNA in infected plant cells came along with the discovery of sRNAs in PTGS, a complete understanding of their biogenesis and their roles for the different type of viruses is still a challenge (Hamilton and Baulcombe, 1999).

The principal source of the viral dsRNA that serves as template to generate vsiRNAs has been extensively discussed. The viral dsRNA that is processed into vsiRNAs for viruses with a RNA genome was thought to originate mainly from the dsRNA intermediates that were needed for their genome replication; however, recent evidence suggests that highly structured single-stranded viral RNA precursors and perfect dsRNAs generated by host RDRs are important sources for vsiRNAs (Molnár et al., 2005; Donaire et al., 2008; Wang et al., 2010). In the case of viruses with a DNA genome, most of vsiRNAs are probably derived from transcriptional units, although bidirectional transcription is considered an option for vsiRNAs production (Moissiard and Voinnet, 2006).

Furthermore, viral genomes have active and specific virus-derived siRNAs regions (hot-spots; Wang et al., 2010).

In accordance with the particular biochemical properties that each of the DCLs possess, vsiRNAs are generated into determined size classes. The 21-nt class of vsiRNAs is usually the most abundant in *Arabidopsis* plants infected with several (+) strand RNA viruses because of the primary role DCL4 has in antiviral defense. In absence of DCL4, the DCL2-dependent 22-nt class of vsiRNAs is the most abundant. The amount of 22-nt vsiRNAs when an active DCL4 is present constitutes only a small portion of the total vsiRNAs (Blevins et al., 2006; Bouché et al., 2006; Deleris et al., 2006; Diaz-Pendon et al., 2007; Garcia-Ruiz et al., 2010). In TCV infections carrying its DCL4-targeting suppressor, the 22-nt class size is the predominant size for vsiRNAs in the infected leaves, although 20, 21, and 22-nt vsiRNAs were found in systemic leaves. The 20-nt class size of vsiRNAs is explained as partial degradation products of DCL4 (Qu et al., 2008). DCL3 generates 24-nt vsiRNAs mainly in *dcl4/dcl2* double mutants during RNA virus infections. DCL3-dependent 24-nt vsiRNAs tend to accumulate in DNA virus infections and are actively important vsiRNAs (Blevins et al., 2006; Bouché et al., 2006; Deleris et al., 2006; Qu et al., 2008; Raja et al., 2008; Garcia-Ruiz et al., 2010). In the case of *DCL1*, as mentioned previously, it is considered to have an indirect role in the biogenesis of vsiRNAs. Whether DCL1 is able to process specific viral dsRNA structures or not has not been completely demonstrated. Low levels of 21-nt vsiRNAs have been detected in *dcl4 dcl2 dcl3* triple mutants infected with TuMV and CMV. Interestingly, 21-nt vsiRNAs from CMV were even detected in *dcl1 dcl2 dcl3 dcl4* plants, suggesting that other mechanism or RNase type III enzyme(s) could be involved (Bouché et al., 2006; Garcia-Ruiz et al., 2010).

The dsRNA-binding protein DRB4 participates in the biogenesis of the vsiRNAs derived from the *Tomato spotted wilt virus* (TSWV) and TYMV. The *drb4* plants infected with these two viruses showed reduced accumulation of 21-nt vsiRNAs (Curtin et al., 2008; Jakubiec et al., 2012). Nonetheless, the DRB4 roles during the biogenesis of vsiRNAs have been questioned because *drb4* plants infected with TCV showed only a small reduction in the level of 21-nt vsiRNAs and increased viral RNA levels (Qu et al., 2008). Similar to other endogenous sRNAs, vsiRNAs are subjected to methylation by HEN1 for their stabilization. Methylation of vsiRNAs is also important for spreading VIGS systemically (Blevins et al., 2006).

Diverse studies have proposed important roles for RDRs in vsiRNA biogenesis as well. The dsRNA generated by RDRs are significantly used to generate secondary vsiRNAs that assist antiviral defense responses and are required for systemic antiviral immunity. RDRs may use primary vsiRNAs as primers or aberrant viral RNA sequences for dsRNA synthesis. The relevance of their role in the biogenesis of vsiRNAs has been recently supported by high-throughput sequencing analyses of small RNA libraries from infected plant tissues. RDR1 and RDR6 play a direct role in the biogenesis of vsiRNAs. The amount of TMV-derived vsiRNAs was reduced in *rdr1* and *rdr6 Arabidopsis* plants. Small RNA deep sequencing analysis showed that TMV-derived vsiRNAs mainly depended on RDR1 (Qi et al., 2009). A significant reduction in the accumulation of vsiRNAs was also observed in *rdr6* plants infected with three TCV mutants (Qu et al., 2008). The production of vsiRNAs from a CMV 2b-deletion mutant was observed to depend mainly on RDR1 activity as well (Diaz-Pendon et al., 2007). Later, analysis with a similar mutation in the 2b gene of CMV showed a cooperative role between RDR1 and RDR6 in the biogenesis of vsiRNAs (Wang et al., 2010). Recently, RDR6 was proposed to be the predominant RDR involved in silencing a different mutant of this suppressor through the production of secondary vsiRNAs (Wang et al., 2011b). The accumulation of vsiRNAs derived from TuMV was significantly reduced in *rdr1* mutant plants. RDR6 participates in the biogenesis of vsiRNAs during suppressor defective TuMV infection (Garcia-Ruiz et al., 2010). Interestingly, no changes in the accumulation of TRV vsiRNAs were detected in single *RDR* mutants (Deleris et al., 2006; Donaire et al., 2008).

The AGO proteins are also relevant members in the biogenesis of vsiRNAs. AGO proteins may cleave viral RNA templates to induce the production of secondary vsiRNAs (Wang et al., 2011b). Viral siRNAs derived from the CMV, the TYMV and the TCV viruses were found in immunoprecipitates of AGO1 (Zhang et al., 2006; Azevedo et al., 2010; Wang et al., 2011b). In addition, CMV-derived vsiRNAs were found in immunoprecipitates of AGO5 and AGO2 as well (Takeda et al., 2008; Wang et al., 2011b). Direct evidence of an AGO-induced viral RNA cleavage mediated by a specific vsiRNA or a satRNA-derived small interfering RNA (satsiRNA) has been reported (Pantaleo et al., 2007; Szittya et al., 2010; Zhu et al., 2011). However, it is uncertain if all the vsiRNAs produced during a plant–virus interaction are loaded into a particular AGO protein. Nowadays, it is known that the 5′ terminal nucleotide preference in loading a small RNA for certain AGOs is conserved also for vsiRNAs (Mi et al., 2008; Wang et al., 2011b).

It was previously mentioned that plants infected with viruses may present a wide range of disease symptoms that have been correlated with disturbances in endogenous sRNA-regulated target genes (Kasschau et al., 2003; Moissiard et al., 2007). As expected, vsiRNAs also regulate host gene targets that may have an impact in viral infections. *In silico* target-prediction analyses have proposed many host genes that could be potentially regulated by vsiRNAs. More than 100 *Arabidopsis* transcripts were found to be potentially targeted by CaMV-derived vsiRNAs. Interestingly, the mRNA At1g76950, that has a Ran GTPase binding, chromatin binding, and zinc ion binding functions, was validated as a directly vsiRNA-regulated transcript (Moissiard and Voinnet, 2006). Moreover, bioinformatic analyses revealed that many host transcripts could be also potentially targeted by TMV-derived siRNAs. For this set of putative host targets, the cleavage of two transcripts was validated by 5′ RACE assays. The two transcripts encode for a cleavage and polyadenylation specific factor (CPSF30) and an unknown protein similar to the translocon-associated protein alpha (TRAP α; Qi et al., 2009). Surprisingly, two studies reported simultaneously that the yellowing symptoms induced in *N. tabacum* by the CMV Y-satellite RNA (Y-Sat) is the consequence of the chlorophyll biosynthetic gene (*CHL1* mRNA) downregulation mediated by Y-Sat-derived siRNAs. The *CHL1* mRNA from *N. tabacum* was found to possess a 22-nt sequence site complementary to the Y-Sat. Interestingly, it was also reported that other two *Nicotiana* species that do not exhibit yellowing symptoms when infected with the CMV Y-Sat are due to a single nucleotide polymorphism presented in the *CHL1* mRNA sequence (Shimura et al., 2011; Smith et al.,

2011). Finally, using sRNA and degradome data, a recent study performed in *Vitis vinifera* showed that several host transcripts were subjected to cleavage by vsiRNAs of the *Grapevine fleck virus* (GFkV) and the *Grapevine rupestris stem pitting-associated virus* (GRSPaV; Miozzi et al., 2013).

In bacteria, the widely studied pathogenic *Agrobacterium tumefaciens* is well-known for introducing a T-DNA that integrates into the genome of plants. The T-DNA encodes genes that trigger the formation of a callus that produces certain compounds called opines. These compounds are used by the bacteria as nutrients. Considering the effects of RNA silencing against foreign genetic material, it was thought that this mechanism may played an important role during this particular bacteria–plant interaction. Small RNAs from the tryptophan 2-monooxygenase (*iaaM*) oncogene and the agropine synthase (*ags*) gene were detected after 3 days of post-infiltration with *A. tumefaciens* in *N. benthamiana* leaves. Interestingly, like in several cases of viral RNA silencing, sRNAs were predominantly 21-nt long and originated from sense and antisense strands, suggesting an important role for DCL4 and RDR6 in the biogenesis of this bacteria-derived sRNAs.

BACTERIAL AND VIRAL SUPPRESSORS OF RNA SILENCING

Plant–microbe interactions are sophisticated and dynamic, involving the continuous improvement of complex defense and counter-defense strategies from both sides. Several microbes introduce effector proteins into plant cells in order to suppress PTI. The contribution of sRNA-mediated silencing in PTI and ETI suggested the existence of bacterial suppressors of RNA silencing (BSRs) and VSRs. Suppressors of RNA silencing may impact small RNA silencing pathway proteins, long double-stranded RNAs, small RNAs, DNA methylation, or sRNA-derived genes to modify the biogenesis, maturation, or function of endogenous and microbe-derived small RNAs. VSRs constitute a diverse group that is widely distributed among viruses. In contrast, only few BSRs have been identified; however, VSRs and BSRs share common strategies like AGO1 disturbance (**Figure 1**).

To test if bacterial effectors evolved to suppress plants miRNAs, Navarro et al. (2008) analyzed modifications in miRNA transcription, biogenesis, or activity favored by this group of proteins. Plants treated with Pst DC3000, compared with those treated with Pst DC3000 *hrcC*$^{-}$, presented reduced accumulation of the PAMP-responsive miRNA precursors pri-miR393a/b and pri-miR396b. The PAMP-insensitive pri-miRNAs pri-miR166a and pri-miR173 were unaltered. These results suggested that some bacteria effectors may suppress PAMP activation of *At-miR393a* and *At-miR393b* transcription. Analyzing *Arabidopsis* plants transiently transformed with different effector proteins, they found that the protein AvrPtoB, a protein with E3-ubiquitin ligase activity, function as a specific bacterial suppressor causing transcriptional repression of the At-miR393a and the At-miR393b precursors. Along with the AvrPtoB suppressor, a different effector protein called AvrPto was identified. In this case, the AvrPto suppressor caused reduction in miR393, miR171, and miR173 accumulation. In contrast to AvrPtoB, no changes in the transcription rate for these three pri-miRNAs were observed to be caused by AvrPto, indicating that AvrPto may affect miRNA biogenesis or stability. Finally, the protein HopT1-1, classified as BSRs, was observed to be involved in suppressing miRNA activity through AGO1 disruption. HopT1-1 apparently interferes with AGO1 affecting miRNA activity related to transcript degradation and translational repression (Navarro et al., 2008). Exactly how these bacterial suppressors act at the molecular level to alter miRNA transcription, biogenesis, and activity still has to be determined.

Lately, the characterization at the molecular level of the *Agrobacterium tumefaciens* 6b protein from the T-DNA region of the Ti plasmid suggests that this protein may function as an RNA silencing suppressor (Wang et al., 2011a). The 6b protein interacts with AGO1 and SE *in vivo* and *in vitro*. *Arabidopsis* plants overexpressing the 6b protein presented reduced accumulation of miRNAs by targeting AGO1 and SE. Besides, plants overexpressing the 6b protein plants shared similar morphological phenotype with *ago1-27* and *se-1* mutants, and with plants overexpressing the 2b RNA silencing suppressor (Wang et al., 2011a). Although the protein 6b can interact with other proteins in the nucleus, it seems evident that this protein plays a role in RNA silencing suppression. Although it is considered that suppression of RNA silencing pathways in tumors may be a consequence of phytohormones produced as consequence of transformation, it will be interesting to determine if the protein 6b contributes to the RNA silencing suppression state observed against the T-DNA genes in tumors.

Since the identification of the first VSRs, many proteins that inhibit RNA silencing during plant–virus interactions have been identified. In many cases, these proteins have other functions besides suppressing RNA silencing and usually do not share sequence or structural similarities. Two major approaches are commonly used by VSRs to inhibit RNA silencing. One of these strategies involves direct binding of VSRs to long dsRNAs and small RNAs to avoid vsiRNAs from being stabilized or loaded into AGO proteins. VSRs like B2 (*Flock house virus*), NS3 (*Influenza A virus*), 2b (*Tomato aspermy virus* and CMV), P14 (*Pothos latent virus*), and P38 (TCV) bind size-independent dsRNAs (Jiang et al., 2012; Omarov and Scholthof, 2012). The P19 viral suppressor of tombusviruses, a widely studied protein, preferentially binds to dsRNA of 19 base pairs long; however, this suppressor could also bind sRNAs of different sizes such as DCL4-dependent 21-nt siRNAs (Scholthof, 2006). Structural analyses showed that P19 is able to measure and select small RNAs in a homodimer conformation (Silhavy et al., 2002; Vargason et al., 2003; Dunoyer et al., 2010). Likewise to P19, several suppressors including P21 (*Beet yellows virus*), P15 (*Peanut clump virus*), γB (*Barley stripe mosaic virus*), HC-Pro (*Tobacco etch virus*), P122 (TMV), NS3 (*Rice hoja blanca virus*), Pns10 (*Rice dwarf virus*), and the tospoviral NSs proteins bind mostly size-dependent double-stranded sRNAs having 2-nt 3′ overhangs (Jiang et al., 2012; Omarov and Scholthof, 2012). The P1b suppressor from the *Cucumber vein yellowing virus* (CVYV) interacts with similar affinity to double-stranded sRNAs with a phosphoryl group or a free OH at their 5′ ends and to duplexes with 2-nt 3′ overhangs or blunt-ends (Valli et al., 2011). Interestingly, the AC4 protein of the *African cassava mosaic virus* (ACMV) binds only single-stranded miRNAs and siRNAs (Chellappan et al., 2005). The second major strategy employed by VSRs to arrest the assembly of functional RISCs is carried out through the direct binding of VSRs with components of the RISC, for instance AGO1. The two viral suppressors P38 (TCV)

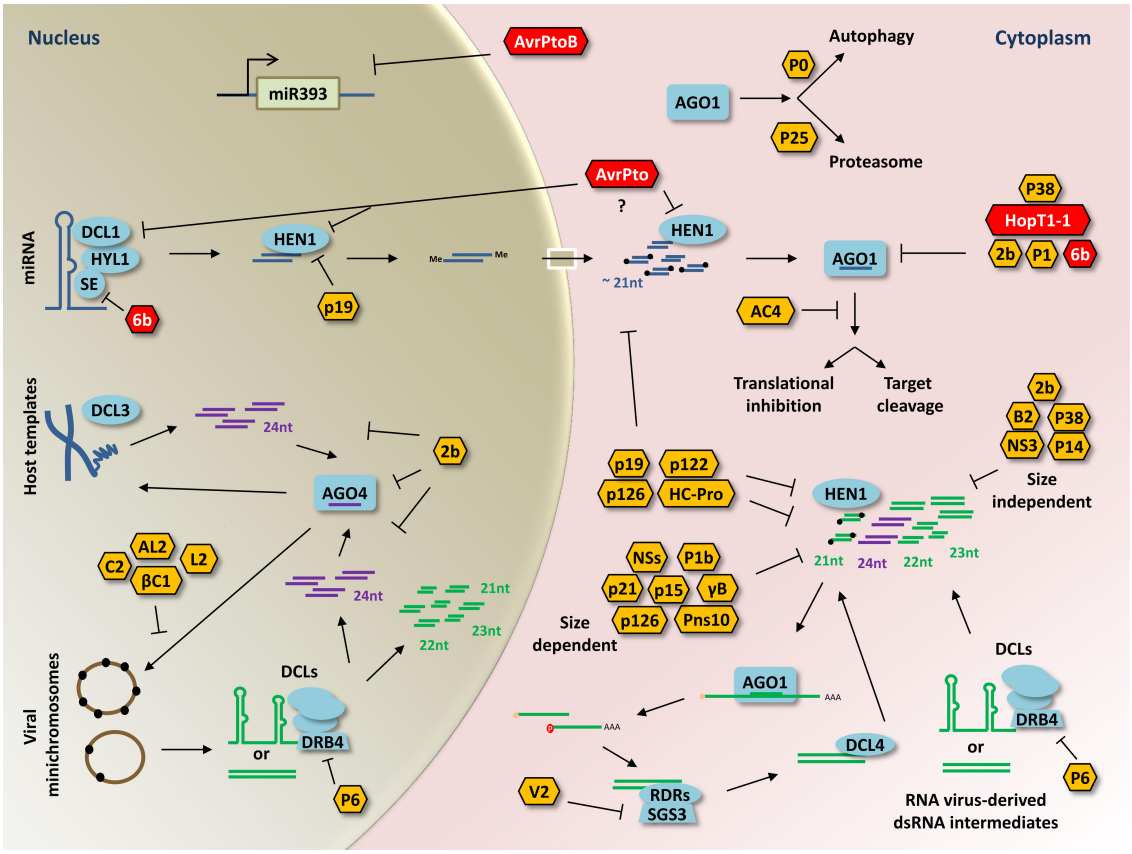


FIGURE 1 | Representative model of viral and bacterial RNA silencing suppressor functions in plants. Double-stranded RNA intermediates are processed by DCLs to generate small RNA duplexes of different sizes. Single-stranded small RNAs are then loaded into AGO proteins to guide RNA silencing through RNA cleavage, translational repression, or DNA methylation. Most of the BSRs discussed in this review (red hexagons) target sRNA silencing pathway components involved in endogenous sRNA-regulated process such as AGO1, SE, DCL1, or HEN1. In contrast, the BSR AvrPtoB causes transcriptional repression of the At-miR393a and the At-miR393b precursors in the nucleus (gray background). The specific target of AvrPto has not been determined but it may target DCL1 or HEN1 because of its effects on miRNA accumulation (question mark). In antiviral RNA silencing, VSRs (yellow hexagons) disturb sRNA silencing pathway components as well. Several of these targeted host-factors participate in the production, stability, and function of vsiRNAs generated from DNA or RNA virus-derived dsRNA intermediates (green structures). The viral suppressor V2 interacts with the SGS3 factor and avoids sRNA amplification. The P6 protein affects production of DCL4-dependent sRNAs by interacting with DRB4. Some VSRs (P122, HC-Pro, P19, P126) have been reported to inhibit 3′ methylation of sRNAs probably by sequestering double-stranded sRNAs and/or arresting HEN1 methylation. Similar to BSRs, viral suppressors (P38, 2b, P1) interact directly with AGO1. The AC4 protein suppress PTGS by binding only single-stranded sRNAs. For DNA viruses (pararetroviruses or geminiviruses) that accumulate as minichromosomes in the nucleus (brown circles), methylation-mediated transcriptional gene silencing is affected by several viral suppressors (βC1, C2, AL2, L2, 2b). Another major strategy employed by VSRs consists on sequestering sRNAs with a specific (size-dependent) or unspecific (size-independent) nucleotide length. Furthermore, P0 and P25 promote AGO1 degradation through autophagy and proteasome-dependent degradation, respectively. Small RNA binding and disturbance or inhibition of small RNA silencing factors are indicated (perpendicular lines).

and P1 [*Sweet potato mild mottle virus* (SPMMV)] interact with the AGO1 protein using GW/WG motifs commonly employed by some endogenous RNA silencing components to assemble functional RISCs. Point mutations in the GW residues of P38 resulted in resistant plants against the TCV infection. Interestingly, the AGO1-binding activity of the P1 protein of SPMMV allows this suppressor to inhibit mature assembled AGO1-containing RISCs and their *de novo* assembly (Azevedo et al., 2010; Giner et al., 2010). The 2b suppressor of CMV was the first suppressor identified to directly bind AGO proteins. The interaction of 2b with AGO1 blocks AGO1 cleavage activity and occurs through one surface of the Piwi-Argonaute-Zwille (PAZ)-containing module and a part of the P-element-induced wimpy testis (PIWI) domain (Zhang et al., 2006). In addition, the 2b protein interacts also with AGO4 altering the RdDM pathway. Further analysis from 2b immunoprecipitates revealed the great binding affinity this suppressor has for 24-nt repeat-associated sRNAs (Hamera et al., 2012). Surprisingly, the *Polerovirus* protein P0 that encodes an F-box protein induces AGO1 degradation. The P0 suppressor does not interact directly with AGO1, but instead interacts with the S-phase kinase-related proteins ASK1 and ASK2, two components of the SCF (SKP1/Cullin1/F-box/RBX) E3 ubiquitin ligase complex. For this reason, it was thought that P0 promoted AGO1 ubiquitination and degradation via the ubiquitin proteasome system (UPS); however, a recent report showed that P0-induced degradation of AGO1 occurs through the autophagy pathway (Pazhouhandeh

et al., 2006; Baumberger et al., 2007; Derrien et al., 2012). Another viral suppressor that triggers AGO1 degradation is the P25 protein encoded by the PVX. In contrast with P0, this suppressor indeed promotes proteasome-dependent degradation of AGO1. Also, P25 may interact with AGO2, AGO3, and AGO4 (Chiu et al., 2010).

Specific viral suppressors that employ different and particular strategies to avoid RNA silencing have been identified as well. For example, the V2 protein of TYLCV is a RNA silencing suppressor that interacts directly with SGS3 and binds dsRNA. Disruption of V2–SGS3 interaction with a point mutation version of V2 stops RNA silencing suppression. These results suggested that V2 affects the interaction of SGS3–RDR6 to avoid small RNA amplification (Glick et al., 2008). Notably, the dsRNA-specific class 1 RNA endoribonuclease III (RNase3) of the *Sweet potato chlorotic stunt virus* (SPCSV) acts as a silencing suppressor using a particular approach that involves its endonuclease activity. This RNase was shown to cleave small RNAs of 21, 22, and 24-nt long into unfunctional small RNAs of 14 base pair size (Cuellar et al., 2009). The viral translational *trans*-activator protein P6 of the CaMV also abolishes RNA silencing. The P6 protein interacts with the dsRNA-binding protein DRB4, which interacts with DCL4 to produce 21-nt siRNAs. This means that P6 suppress DCL4-dependent vsiRNAs production (Haas et al., 2008). Other VSRs including P122, HC-Pro, P19, P126 have been reported to inhibit 3′ methylation of small RNA duplexes probably by sequestering double-stranded small RNAs and blocking HEN1 methylation (Jiang et al., 2012; Omarov and Scholthof, 2012). Furthermore, several VSRs from DNA viruses have been shown to suppress or alter the PTGS mechanism that regulates DNA methylation and histone adjustments. The AL2 protein of *Tomato golden mosaic virus* (TGMV) and the L2 protein of BCTV suppress RNA silencing by inhibiting the adenosine kinase (ADK) activity, which plays relevant roles in adenosine salvage and methyl cycle maintenance. Inactivation of ADK resulted in suppression of RNA silencing as occurred with the incorporation of the two geminivirus proteins. These two proteins can reverse TGS of a GFP transgene introduced in *N. benthamiana*. In addition, the AL2 and L2 proteins were found to cause ectopic expression of an endogenous loci silenced by methylation and a global reduction in cytosine methylation (Wang et al., 2005; Buchmann et al., 2009). Another viral suppressor that affects methylation modifications is the C2 protein of the *Beet severe curly top virus* (BSCTV). The C2 protein interacts with the *S*-adenosyl-methionine decarboxylase 1 (SAMDC1) attenuating its proteasome-dependent degradation. The SAMDC1 protein participates in the polyamine biosynthesis, but is also important for SAM/dcSAM balance and transmethylation (Zhang et al., 2011c). The betasatellite of the *Tomato yellow leaf curl China virus* (TYLCCNB) encodes a protein called βC1 that acts as a suppressor of methylation-mediated TGS. The βC1 protein interacts with the *S*-adenosylhomocysteine hydrolase (SAHH) enzyme that is involved in the methyl cycle and therefore plays a role in TGS. The expression of the βC1 protein decreases cytosine methylation of the viral and host genomes. Also, this protein was shown to reverse TGS applied to a transgene and an endogenous locus (Yang et al., 2011). As previously mentioned, many VSRs with small RNA-binding activities and/or capable of interacting with important sRNA silencing pathway components, may modify the levels of endogenous sRNAs like miRNAs and ta-siRNAs (Shimura and Pantaleo, 2011).

CONCLUDING REMARKS

The small RNA-mediated plant defense responses have emerged as relevant components of the innate immune system. Increasing evidence has highlighted the warfare that takes place between plants and microbes around the RNA silencing system. In this review, the recent findings, similarities and differences related to the RNA-mediated arms race between plants and two important group of microbes such as bacteria and viruses were discussed. In general, evident biological differences between these two groups of microbes are reflected in big differences regarding small RNA-mediated antiviral and antibacterial immunity; however, there are also specific similarities in plants defense responses through RNA silencing against bacteria and viruses. Likewise, similar strategies have been identified related to the microbes counter defense responses against RNA silencing. Bacteria-responsive miRNAs with potential roles in regulating bacterial immunity have been reported for several plant–bacteria strains interactions, flg22 treatment and even during symbiotic nitrogen fixing bacteria inoculation and nodule development (Subramanian et al., 2008; Lelandais-Brière et al., 2009; Wang et al., 2009; De Luis et al., 2012; Reynoso et al., 2012). Most of these miRNAs are involved in hormone signaling pathways and ETI. In viral infections, few miRNAs have been identified to play a direct role in antiviral immunity because viruses usually affect the RNA silencing system to circumvent this kind of plant defense response. Additionally, viral infections induce expression of novel phased miRNAs from conserved miRNA precursors (Du et al., 2011). Although, sRNA-regulated genes involved in ETI during bacterial infections are well-documented, it still remains to be evaluated to what extent host small RNAs participate in antiviral immunity. In this regard, several miRNAs have been proposed to be good candidates to directly act as vsiRNAs regulating viral RNAs (Perez-Quintero et al., 2010). Comparing the antiviral and the antibacterial roles of the small RNA biogenesis factors may shed light on the complex modes of regulation these proteins have to confer plants disease resistance. The study of VSRs and BSRs along with their targets may help to decipher redundancy in the activity of several RNA silencing components during plant–microbe interactions. Further studies related to this growing field will define more precisely the global small RNA-mediated plant defense responses induced by bacteria and viruses. We expect that understanding small RNA responses to viral and bacterial infections will provide novel means to generate disease-resistant plants.

ACKNOWLEDGMENTS

We thank Dr. Alejandra A. Covarrubias and Dr. Oswaldo Valdés for reading the manuscript and fruitful discussions. We apologize to those colleagues whose research we were not able to discuss due to space limitations. Pablo Peláez was supported by a fellowship from CONACYT and this work is part of his Ph.D. thesis at the PhD program in Biomedical Sciences (UNAM). The Dr. Federico Sanchez laboratory is currently funded by CONACYT 17744 and DGAPA IN106012 grants.

REFERENCES

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouche, N., Gasciolli, V., et al. (2006). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Curr. Biol.* 16, 927–932. doi: 10.1016/j.cub.2006.03.035

Agorio, A., and Vera, P. (2007). ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in *Arabidopsis*. *Plant Cell* 19, 3778–3790. doi: 10.1105/tpc.107.054494

Amin, I., Patil, B. L., Briddon, R. W., Mansoor, S., and Fauquet, C. M. (2011a). Comparison of phenotypes produced in response to transient expression of genes encoded by four distinct begomoviruses in *Nicotiana benthamiana* and their correlation with the levels of developmental miRNAs. *Virol. J.* 8, 238. doi: 10.1186/1743-422X-8-238

Amin, I., Patil, B. L., Briddon, R. W., Mansoor, S., and Fauquet, C. M. (2011b). A common set of developmental miRNAs are upregulated in *Nicotiana benthamiana* by diverse begomoviruses. *Virol. J.* 8, 143. doi: 10.1186/1743-422X-8-143

Ausubel, F. M. (2005). Are innate immune signaling pathways in plants and animals conserved? *Nat. Immunol.* 6, 973–979. doi: 10.1038/ni1253

Azevedo, J., Garcia, D., Pontier, D., Ohnesorge, S., Yu, A., Garcia, S., et al. (2010). Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. *Genes Dev.* 24, 904–915. doi: 10.1101/gad.1908710

Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356–363. doi: 10.1038/nature02874

Baumberger, N., Tsai, C. H., Lie, M., Havecker, E., and Baulcombe, D. C. (2007). The Polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. *Curr. Biol.* 17, 1609–1614. doi: 10.1016/j.cub.2007.08.039

Bazzini, A. A., Almasia, N. I., Manacorda, C. A., Mongelli, V. C., Conti, G., Maroniche, G. A., et al. (2009). Virus infection elevates transcriptional activity of miR164a promoter in plants. *BMC Plant Biol.* 9:152. doi: 10.1186/1471-2229-9-152

Bazzini, A. A., Hopp, H. E., Beachy, R. N., and Asurmendi, S. (2007). Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12157–12162. doi: 10.1073/pnas.0705114104

Bhattacharjee, S., Zamora, A., Azhar, M. T., Sacco, M. A., Lambert, L. H., and Moffett, P. (2009). Virus resistance induced by NB–LRR proteins involves Argonaute4-dependent translational control. *Plant J.* 58, 940–951. doi: 10.1111/j.1365-313X.2009.03832.x

Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Beknazariants, D., Si-Ammour, A., Park, H. S., et al. (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res.* 34, 6233–6246. doi: 10.1093/nar/gkl886

Bortolamiol, D., Pazhouhandeh, M., Marrocco, K., Genschik, P., and Ziegler-Graff, V. (2007). The Polerovirus F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. *Curr. Biol.* 17, 1615–1621. doi: 10.1016/j.cub.2007.07.061

Bouché, N., Lauressergues, D., Gasciolli, V., and Vaucheret, H. (2006). An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J.* 25, 3347–3356. doi: 10.1038/sj.emboj.7601217

Boutet, S., Vazquez, F., Liu, J., Beclin, C., Fagard, M., Gratias, A., et al. (2003). *Arabidopsis* HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. *Curr. Biol.* 13, 843–848. doi: 10.1016/S0960-9822(03)00293-8

Buchmann, R. C., Asad, S., Wolf, J. N., Mohannath, G., and Bisaro, D. M. (2009). Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. *J. Virol.* 83, 5005–5013. doi: 10.1128/JVI.01771-08

Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K. B., Montgomery, T. A., Nguyen, T., et al. (2012). Functional analysis of three *Arabidopsis* ARGONAUTES using slicer-defective mutants. *Plant Cell* 24, 3613–3629. doi: 10.1105/tpc.112.099945

Chapman, E. J., and Carrington, J. C. (2007). Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.* 8, 884–896. doi: 10.1038/nrg2179

Chapman, E. J., Prokhnevsky, A. I., Gopinath, K., Dolja, V. V., and Carrington, J. C. (2004). Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes Dev.* 18, 1179–1186. doi: 10.1101/gad.1201204

Chellappan, P., Vanitharani, R., and Fauquet, C. M. (2005). MicroRNA-binding viral protein interferes with *Arabidopsis* development. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10381–10386. doi: 10.1073/pnas.0504439102

Chen, J., Li, W. X., Xie, D., Peng, J. R., and Ding, S. W. (2004). Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microRNA in host gene expression. *Plant Cell* 16, 1302–1313. doi: 10.1105/tpc.018986

Chen, X. (2009). Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* 25, 21–44. doi: 10.1146/annurev.cellbio.042308.113417

Chiu, M. H., Chen, I. H., Baulcombe, D. C., and Tsai, C. H. (2010). The silencing suppressor P25 of Potato virus X interacts with Argonaute1 and mediates its degradation through the proteasome pathway. *Mol. Plant Pathol.* 11, 641–649. doi: 10.1111/j.1364-3703.2010.00634.x

Crane, Y. M., and Gelvin, S. B. (2007). RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15156–15161. doi: 10.1073/pnas.0706986104

Csorba, T., Bovi, A., Dalmay, T., and Burgyan, J. (2007). The p122 subunit of Tobacco Mosaic Virus replicase is a potent silencing suppressor and compromises both small interfering RNA- and miRNA-mediated pathways. *J. Virol.* 81, 11768–11780. doi: 10.1128/JVI.01230-07

Cuellar, W. J., Kreuze, J. F., Rajamäki, M. L., Cruzado, K. R., Untiveros, M., and Valkonen, J. P. (2009). Elimination of antiviral defense by viral RNase III. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10354–10358. doi: 10.1073/pnas.0806042106

Curtin, S. J., Watson, J. M., Smith, N. A., Eamens, A. L., Blanchard, C. L., and Waterhouse, P. M. (2008). The roles of plant dsRNA-binding proteins in RNAi-like pathways. *FEBS Lett.* 582, 2753–2760. doi: 10.1016/j.febslet.2008.07.004

Dalmay, T., Horsefield, R., Braunstein, T. H., and Baulcombe, D. C. (2001). SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in *Arabidopsis*. *EMBO J.* 20, 2069–2078. doi: 10.1093/emboj/20.8.2069

Dangl, J. L., and Jones, J. D. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411, 826–833. doi: 10.1038/35081161

Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K. D., Carrington, J. C., and Voinnet, O. (2006). Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313, 68–71. doi: 10.1126/science.1128214

De Luis, A., Markmann, K., Cognat, V., Holt, D. B., Charpentier, M., Parniske, M., et al. (2012). Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume *Lotus japonicus*. *Plant Physiol.* 160, 2137–2154. doi: 10.1104/pp.112.204883

Derrien, B., Baumberger, N., Schepetilnikov, M., Viotti, C., De Cillia, J., Ziegler-Graff, V., et al. (2012). Degradation of the antiviral component ARGONAUTE1 by the autophagy pathway. *Proc. Natl. Acad. Sci. U.S.A.* 109, 15942–15946. doi: 10.1073/pnas.1209487109

Diaz-Pendon, J. A., Li, F., Li, W. X., and Ding, S. W. (2007). Suppression of antiviral silencing by cucumber mosaic virus 2b protein in *Arabidopsis* is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell* 19, 2053–2063. doi: 10.1105/tpc.106.047449

Donaire, L., Barajas, D., Martinez-Garcia, B., Martinez-Priego, L., Pagan, I., and Llave, C. (2008). Structural and genetic requirements for the biogenesis of tobacco rattle virus-derived small interfering RNAs. *J. Virol.* 82, 5167–5177. doi: 10.1128/JVI.00272-08

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., et al. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2183–E2191. doi: 10.1073/pnas.1209329109

Du, P., Wu, J., Zhang, J., Zhao, S., Zheng, H., Gao, G., et al. (2011). Viral infection induces expression of novel phased microRNAs from conserved cellular microRNA precursors. *PLoS Pathog.* 7:e1002176. doi: 10.1371/journal.ppat.1002176

Duan, C. G., Fang, Y. Y., Zhou, B. J., Zhao, J. H., Hou, W. N., Zhu, H., et al. (2012). Suppression of *Arabidopsis* ARGONAUTE1-mediated slicing, transgene-induced RNA silencing, and DNA methylation by distinct domains of the cucumber mosaic virus 2b protein. *Plant Cell* 24, 259–274. doi: 10.1105/tpc.111.092718

Dunoyer, P., Himber, C., and Voinnet, O. (2006). Induction, suppression and requirement of RNA silencing pathways in virulent *Agrobacterium tumefaciens* infections. *Nat. Genet.* 38, 258–263. doi: 10.1038/ng1722

Dunoyer, P., Lecellier, C. H., Parizotto, E. A., Himber, C., and Voinnet, O.

(2004). Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. *Plant Cell* 16, 1235–1250. doi: 10.1105/tpc.020719

Dunoyer, P., Schott, G., Himber, C., Meyer, D., Takeda, A., Carrington, J. C., et al. (2010). Small RNA duplexes function as mobile silencing signals between plant cells. *Science* 328, 912–916. doi: 10.1126/science.1185880

Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., et al. (2007). High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS ONE* 2:e219. doi: 10.1371/journal.pone.0000219

Feng, J., Lai, L., Lin, R., Jin, C., and Chen, J. (2012). Differential effects of Cucumber mosaic virus satellite RNAs in the perturbation of microRNA-regulated gene expression in tomato. *Mol. Biol. Rep.* 39, 775–784. doi: 10.1007/s11033-011-0798-y

Garcia, D., Garcia, S., Pontier, D., Marchais, A., Renou, J. P., Lagrange, T., et al. (2012). Ago hook and RNA helicase motifs underpin dual roles for SDE3 in antiviral defense and silencing of nonconserved intergenic regions. *Mol. Cell* 48, 109–120. doi: 10.1016/j.molcel.2012.07.028

Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J., et al. (2010). *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *Plant Cell* 22, 481–496. doi: 10.1105/tpc.109.073056

Giner, A., Lakatos, L., García-Chapa, M., López-Moya, J. J., and Burgyán, J. (2010). Viral protein inhibits RISC activity by Argonaute binding through conserved WG/GW motifs. *PLoS Pathog.* 6:e1000996. doi: 10.1371/journal.ppat.1000996

Glick, E., Zrachya, A., Levy, Y., Mett, A., Gidoni, D., Belausov, E., et al. (2008). Interaction with host SGS3 is required for suppression of RNA silencing by tomato yellow leaf curl virus V2 protein. *Proc. Natl. Acad. Sci. U.S.A.* 105, 157–161. doi: 10.1073/pnas.0709036105

Guo, W., Wu, G., Yan, F., Lu, Y., Zheng, H., Lin, L., et al. (2012). Identification of novel *Oryza sativa* miRNAs in deep sequencing-based small RNA libraries of rice infected with Rice stripe virus. *PLoS ONE* 7:e46443. doi: 10.1371/journal.pone.0046443

Haas, G., Azevedo, J., Moissiard, G., Geldreich, A., Himber, C., Bureau, M., et al. (2008). Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. *EMBO J.* 27, 2102–2112. doi: 10.1038/emboj.2008.129

Hamera, S., Song, X., Su, L., Chen, X., and Fang, R. (2012). Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. *Plant J.* 69, 104–115. doi: 10.1111/j.1365-313X.2011.04774.x

Hamilton, A. J., and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–952. doi: 10.1126/science.286.5441.950

Harvey, J. J. W., Lewsey, M. G., Patel, K., Westwood, J., Heimsta, S., Carr, J. P., et al. (2011). An antiviral defense role of AGO2 in plants. *PLoS ONE* 6:e14639. doi: 10.1371/journal.pone.0014639

He, J., Dong, Z., Jia, Z., Wang, J., and Wang, G. (2010). Isolation, expression and functional analysis of a putative RNA-dependent RNA polymerase gene from maize (*Zea mays* L.). *Mol. Biol. Rep.* 37, 865–874. doi: 10.1007/s11033-009-9692-2

He, X. F., Fang, Y. Y., Feng, L., and Guo, H. S. (2008). Characterization of conserved and novel microRNAs and their targets, including a TuMV-induced TIR-NBS-LRR class R gene-derived novel miRNA in *Brassica*. *FEBS Lett.* 582, 2445–2452. doi: 10.1016/j.febslet.2008.06.011

Hernandez-Pinzon, I., Yelina, N. E., Schwach, F., Studholme, D. J., Baulcombe, D., and Dalmay, T. (2007). SDE5, the putative homologue of a human mRNA export factor, is required for transgene silencing and accumulation of trans-acting endogenous siRNA. *Plant J.* 50, 140–148. doi: 10.1111/j.1365-313X.2007.03043.x

Hu, Q., Hollunder, J., Niehl, A., Kørner, C. J., Gereige, D., Windels, D., et al. (2011). Specific impact of tobamovirus infection on the *Arabidopsis* small RNA profile. *PLoS ONE* 6:e19549. doi: 10.1371/journal.pone.0019549

Jagadeeswaran, G., Saini, A., and Sunkar, R. (2009). Biotic and abiotic stress down-regulate miR398 expression in *Arabidopsis*. *Planta* 229, 1009–1014. doi: 10.1007/s00425-009-0889-3

Jakubiec, A., Yang, S. W., and Chua, N. H. (2012). *Arabidopsis* DRB4 protein in antiviral defense against Turnip yellow mosaic virus infection. *Plant J.* 69, 14–25. doi: 10.1111/j.1365-313X.2011.04765.x

Jamous, R. M., Boonrod, K., Fuellgrabe, M. W., Ali-Shtayeh, M. S., Krczal, G., and Wassenegger, M. (2011). The helper component-proteinase of the *Zucchini yellow mosaic virus* inhibits the Hua Enhancer 1 methyltransferase activity in vitro. *J. Gen. Virol.* 92, 2222–2226. doi: 10.1099/vir.0.031534-0

Jaubert, M., Bhattacharjee, S., Mello, A. F. S., Perry, K. L., and Moffett, P. (2011). ARGONAUTE2 mediates RNA-silencing antiviral defenses against Potato virus X in *Arabidopsis*. *Plant Physiol.* 156, 1556–1564. doi: 10.1104/pp.111.178012

Jay, F., Wang, Y., Yu, A., Taconnat, L., Pelletier, S., Colot, V., et al. (2011). Misregulation of AUXIN RESPONSE FACTOR 8 underlies the developmental abnormalities caused by three distinct viral silencing suppressors in *Arabidopsis*. *PLoS Pathog.* 7:e1002035. doi: 10.1371/journal.ppat.1002035

Jiang, L., Wei, C., and Li, Y. (2012). Viral suppression of RNA silencing. *Sci. China Life Sci.* 55, 109–118. doi: 10.1007/s11427-012-4279-x

Jin, H. (2008). Endogenous small RNAs and antibacterial immunity in plants. *FEBS Lett.* 582, 2679–2684. doi: 10.1016/j.febslet.2008.06.053

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Kasschau, K. D., Xie, Z., Allen, E., Llave, C., Chapman, E. J., Krizan, K. A., et al. (2003). P1/HC-Pro, a viral suppressor of RNA silencing interferes with *Arabidopsis* development and miRNA function. *Dev. Cell* 4, 205–217. doi: 10.1016/S1534-5807(03)00025-X

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., and Jin, H. (2007). A novel class of bacteria-induced small RNAs in *Arabidopsis*. *Genes Dev.* 21, 3123–3134. doi: 10.1101/gad.1595107

Katiyar-Agarwal, S., and Jin, H. (2010). Role of small RNAs in host–microbe interactions. *Annu. Rev. Phytopathol.* 48, 225–246. doi: 10.1146/annurev-phyto-073009-114457

Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A., Zhu, J. K., et al. (2006). A pathogen-inducible endogenous siRNA in plant immunity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18002–18007. doi: 10.1073/pnas.0608258103

Lang, Q., Jin, C., Lai, L., Feng, J., Chen, S., and Chen, J. (2011). Tobacco microRNAs prediction and their expression infected with Cucumber mosaic virus and Potato virus X. *Mol. Biol. Rep.* 38, 1523–1531. doi: 10.1007/s11033-010-0260-6

Lelandais-Brière, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., et al. (2009). Genome-wide *Medicago truncatula* smallRNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell* 21, 2780–2796. doi: 10.1105/tpc.109.068130

Li, F., Pignatta, D., Bendix, C., Brunkard, J. O., Cohn, M. M., Tung, J., et al. (2012). MicroRNA regulation of plant innate immune receptors. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1790–1795. doi: 10.1073/pnas.1118282109

Li, H., Deng, Y., Wu, T., Subramanian, S., and Yu, O. (2010a). Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol.* 153, 1759–1770. doi: 10.1104/pp.110.156950

Li, Y., Zhang, Q., Zhang, J., Wu, L., Qi, Y., and Zhou, J. M. (2010b). Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity. *Plant Physiol.* 152, 2222–2231. doi: 10.1104/pp.109.151803

Lim, H. S., Ko, T. S., Hobbs, H. A., Lambert, K. N., Yu, J. M., McCoppin, N. K., et al. (2007). *Soybean mosaic virus* helper component-protease alters leaf morphology and reduces seed production in transgenic soybean plants. *Phytopathology* 97, 366–372. doi: 10.1094/PHYTO-97-3-0366

Lindbo, J. A., Silva-Rosales, L., Proebsting, W. M., and Dougherty, W. G. (1993). Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. *Plant Cell* 5, 1749–1759.

López, A., Ramírez, V., García-Andrade, J., Flors, V., and Vera, P. (2011). The RNA silencing enzyme RNA polymerase v is required for plant immunity. *PLoS Genet.* 7:e1002434. doi: 10.1371/journal.pgen.1002434

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., et al. (2008). Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5′ terminal nucleotide. *Cell* 133, 116–127. doi: 10.1016/j.cell.2008.02.034

Miozzi, L., Gambino, G., Burgyan, J., and Pantaleo, V. (2013). Genome-wide identification of viral and host transcripts targeted by viral siRNAs in *Vitis vinifera*. *Mol. Plant Pathol.* 14, 30–43. doi: 10.1111/j.1364-3703.2012.00828.x

Moissiard, G., Parizotto, E. A., Himber, C., and Voinnet, O. (2007). Transitivity in *Arabidopsis* can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded

suppressor proteins. *RNA* 13, 1268–1278. doi: 10.1261/rna.541307

Moissiard, G., and Voinnet, O. (2006). RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four *Arabidopsis* Dicer-like proteins. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19593–19598. doi: 10.1073/pnas.0604627103

Molnár, A., Csorba, T., Lakatos, L., Várallyay, E., Lacomme, C., and Burgyán, J. (2005). Plant Virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. *J. Virol.* 79, 7812–7818. doi: 10.1128/JVI.79.12.7812-7818.2005

Morel, J. B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., et al. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* 14, 629–639. doi: 10.1105/tpc.010358

Mourrain, P., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., et al. (2000). *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* 101, 533–542. doi: 10.1016/S0092-8674(00)80863-6

Muangsan, N., Beclin, C., Vaucheret, H., and Robertson, D. (2004). Geminivirus VIGS of endogenous genes required SGS2/SDE1 and SGS3 and defines a new branch in the genetic pathway for silencing in plants. *Plant J.* 38, 1004–1014. doi: 10.1111/j.1365-313X.2004.02103.x

Naqvi, A. R., Haq, Q. M., and Mukherjee, S. K. (2010). MicroRNA profiling of tomato leaf curl New Delhi virus (tolcndv) infected tomato leaves indicates that deregulation of mir159/319 and mir172 might be linked with leaf curl disease. *Virol. J.* 7, 281. doi: 10.1186/1743-422X-7-281

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., et al. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signalling. *Science* 312, 436–439. doi: 10.1126/science.1126088

Navarro, L., Jay, F., Nomura, K., He, S. Y., and Voinnet, O. (2008). Suppression of the microRNA pathway by bacterial effector proteins. *Science* 321, 964–967. doi: 10.1126/science.1159505

Omarov, R. T., and Scholthof, H. B. (2012). Biological chemistry of virus-encoded RNA silencing suppressors: an overview. *Methods Mol. Biol.* 894, 39–56. doi: 10.1007/978-1-61779-882-5_3

Pacheco, R., García-Marcos, A., Barajas, D., Martiáñez, J., and Tenllado, F. (2012). PVX-potyvirus synergistic infections differentially alter microRNA accumulation in *Nicotiana benthamiana*. *Virus Res.* 165, 231–235. doi: 10.1016/j.virusres.2012.02.012

Padmanabhan, C., Zhang, X., and Jin, H. (2009). Host small RNAs are big contributors to plant innate immunity. *Curr. Opin. Plant Biol.* 12,465–472. doi: 10.1016/j.pbi.2009.06.005

Pantaleo, V., Szittya, G., and Burgyan, J. (2007). Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. *J. Virol.* 81, 3797–3806. doi: 10.1128/JVI.02383-06

Pazhouhandeh, M., Dieterle, M., Marrocco, K., Lechner, E., Berry, B., Brault, V., et al. (2006). F-box-like domain in the Polerovirus protein P0 is required for silencing suppressor function. *Proc. Natl. Acad. Sci. U.S.A.* 103, 1994–1999. doi: 10.1073/pnas.0510784103

Perez-Quintero, A. L., Neme, R., Zapata, A., and Lopez, C. (2010). Plant microRNAs and their role in defense against viruses: a bioinformatics approach. *BMC Plant Biol.* 10:138. doi: 10.1186/1471-2229-10-138

Pruss, G. J., Nester, E. W., and Vance, V. (2008). Infiltration with *Agrobacterium tumefaciens* induces host defense and development-dependent responses in the infiltrated zone. *Mol. Plant Microbe Interact.* 21, 1528–1538. doi: 10.1094/MPMI-21-12-1528

Qi, X., Bao, F. S., and Xie, Z. (2009). Small RNA deep sequencing reveals role for *Arabidopsis thaliana* RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS ONE* 4:e4971. doi: 10.1371/journal.pone.0004971

Qu, F., Ye, X., Hou, G., Sato, S., Clemente, T. E., and Morris, T. J. (2005). RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in *Nicotiana benthamiana*. *J. Virol.* 79, 15209–15217. doi: 10.1128/JVI.79.24.15209-15217.2005

Qu, F., Ye, X., and Morris, T. J. (2008). *Arabidopsis* DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14732–14737. doi: 10.1073/pnas.0805760105

Raja, P., Sanville, B. C., Buchmann, R. C., and Bisaro, D. M. (2008). Viral genome methylation as an epigenetic defense against geminiviruses. *J. Virol.* 82, 8997–9007. doi: 10.1128/JVI.00719-08

Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* 20, 3407–3425. doi: 10.1101/gad.1476406

Ratcliff, F., Harrison, B. D., and Baulcombe, D. C. (1997). A similarity between viral defense and gene silencing in plants. *Science* 276, 1558–1560. doi: 10.1126/science.276.5318.1558

Reynoso, M. A., Blanco, F. A., Bailey-Serres, J., Crespi, M., and Zanetti, M. E. (2012). Selective recruitment of mRNAs and miRNAs to polyribosomes in response to rhizobia infection in *Medicago truncatula*. *Plant J.* doi: 10.1111/tpj.12033

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell* 110, 513–520. doi: 10.1016/S0092-8674(02)00863-2

Ronald, P. C., and Beutler, B. (2010). Plant and animal sensors of conserved microbial signatures. *Science* 330, 1061–1064. doi: 10.1126/science.1189468

Scholthof, H. B. (2006). The *Tombusvirus*-encoded P19: from irrelevance to elegance. *Nat. Rev. Microbiol.* 4, 405–411. doi: 10.1038/nrmicro1395

Scholthof, H. B., Alvarado, V. Y., Vega-Arreguin, J. C., Ciomperlik, J., Odokonyero, D., Brosseau, C., et al. (2011). Identification of an ARGONAUTE for antiviral RNA silencing in *Nicotiana benthamiana*. *Plant Physiol.* 156, 1548–1555. doi: 10.1104/pp.111.178764

Schramke, V., and Allshire, R. (2004). Those interfering little RNAs! Silencing and eliminating chromatin. *Curr. Opin. Genet. Dev.* 14, 174–180. doi: 10.1016/j.gde.2004.02.006

Schwach, F., Vaistij, F. E., Jones, L., and Baulcombe, D. C. (2005). An RNA-dependent RNA polymerase prevents meristem invasion by potato virus X and is required for the activity but not the production of a systemic silencing signal. *Plant Physiol.* 138, 1842–1852. doi: 10.1104/pp.105.063537

Shimura, H., and Pantaleo, V. (2011). Viral induction and suppression of RNA silencing in plants. *Biochim. Biophys. Acta* 1809, 601–612. doi: 10.1016/j.bbagrm.2011.04.005

Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J., Sueda, K., et al. (2011). A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. *PLoS Pathog.* 7:e1002021. doi: 10.1371/journal.ppat.1002021

Shivaprasad, P. V., Chen, H. M., Patel, K., Bond, D. M., Santos, B. A., and Baulcombe, D. C. (2012). A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* 24, 859–874. doi: 10.1105/tpc.111.095380

Silhavy, D., Molnár, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M., et al. (2002). A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *EMBO J.* 21, 3070–3080. doi: 10.1093/emboj/cdf312

Smith, N. A., Eamens, A. L., and Wang, M. B. (2011). Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathog.* 7:e1002022. doi: 10.1371/journal.ppat.1002022

Spoel, S. H., and Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nat. Rev. Immunol.* 12, 89–100. doi: 10.1038/nri3141

Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W. B., Zhu, J. K., and Yu, O. (2008). Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics* 9:160. doi: 10.1186/1471-2164-9-160

Sunkar, R., Kapoor, A., and Zhu, J. K. (2006). Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by down regulation of miR398 and important for oxidative stress tolerance. *Plant Cell* 18, 2051–2065. doi: 10.1105/tpc.106.041673

Szittya, G., Moxon, S., Pantaleo, V., Toth, G., Rusholme, Pilcher, R. L., Moulton, V., et al. (2010). Structural and functional analysis of viral siRNAs. *PLoS Pathog.* 6:e1000838. doi: 10.1371/journal.ppat.1000838

Takeda, A., Iwasaki, S., Watanabe, T., Utsumi, M., and Watanabe, Y. (2008). The mechanism selecting the guide strand from small RNA duplexes is different among Argonaute proteins. *Plant Cell Physiol.* 49, 493–500. doi: 10.1093/pcp/pcn043

Takeda, A., Tsukuda, M., Mizumoto, H., Okamoto, K., Kaido, M., Mise, K., et al. (2005). A plant RNA virus suppresses RNA silencing through viral RNA replication. *EMBO J.* 24, 3147–3157. doi: 10.1038/sj.emboj.7600776

Valli, A., Oliveros, J. C., Molnar, A., Baulcombe, D., and García, J. A.

(2011). The specific binding to 21-nt double-stranded RNAs is crucial for the anti-silencing activity of Cucumber vein yellowing virus P1b and perturbs endogenous small RNA populations. *RNA* 17, 1148–1158. doi: 10.1261/rna.2510611

Varallyay, E., Valoczi, A., Agyi, A., Burgyan, J., and Havelda, Z. (2010). Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *EMBO J.* 29, 3507–3519. doi: 10.1038/emboj.2010.215

Vargason, J. M., Szittya, G., Burgyán, J., and Hall, T. M. (2003). Size selective recognition of siRNA by an RNA silencing suppressor. *Cell* 115, 799–811. doi: 10.1016/S0092-8674(03)00984-X

Vazquez, F., Legrand, S., and Windels, D. (2010). The biosynthetic pathways and biological scopes of plant small RNAs. *Trends Plant Sci.* 15, 337–345. doi: 10.1016/j.tplants.2010.04.001

Vogler, H., Akbergenov, R., Shivaprasad, P. V., Dang, V., Fasler, M., Kwon, M. O., et al. (2007). Modification of small RNAs associated with suppression of RNA silencing by tobamovirus replicase protein. *J. Virol.* 81, 10379–10388. doi: 10.1128/JVI.00727-07

Wang, H., Buckley, K. J., Yang, X., Buchmann, R. C., and Bisaro, D. M. (2005). Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. *J. Virol.* 79, 7410–7418. doi: 10.1128/JVI.79.12.7410-7418.2005

Wang, M., Soyano, T., Machida, S., Yang, J. Y., Jung, C., Chua, N. H., et al. (2011a). Molecular insights into plant cell proliferation disturbance by *Agrobacterium* protein 6b. *Genes Dev.* 25, 64–76. doi: 10.1101/gad.1985511

Wang, X. B., Jovel, J., Udomporn, P., Wang, Y., Wu, Q., Li, W. X., et al. (2011b). The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative Argonautes in *Arabidopsis thaliana*. *Plant Cell* 23, 1625–1638. doi: 10.1105/tpc.110.082305

Wang, X. B., Wu, Q., Ito, T., Cillo, F., Li, W. X., Chen, X., et al. (2010). RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 484–489. doi: 10.1073/pnas.0904086107

Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A., and Li, X. (2009). Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochem. Biophys. Res. Commun.* 378, 799–803. doi: 10.1016/j.bbrc.2008.11.140

Xie, Z., Fan, B., Chen, C., and Chen, Z. (2001). An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6516–6521. doi: 10.1073/pnas.111440998

Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., et al. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* 2:e104. doi: 10.1371/journal.pbio.0020104

Xie, Z., Kasschau, K. D., and Carrington, J. C. (2003). Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. *Curr. Biol.* 13, 784–789. doi: 10.1016/S0960-9822(03)00281-1

Yang, S. J., Carter, S. A., Cole, A. B., Cheng, N. H., and Nelson, R. S. (2004). A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by *Nicotiana benthamiana*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6297–6302. doi: 10.1073/pnas.0304346101

Yang, X., Xie, Y., Raja, P., Li, S., Wolf, J. N., Shen, Q., et al. (2011). Suppression of methylation-mediated transcriptional gene silencing by βC1-SAHH protein interaction during Geminivirus-Betasatellite Infection. *PLoS Pathog.* 7:e1002329. doi: 10.1371/journal.ppat.1002329

Yifhar, T., Pekker, I., Peled, D., Friedlander, G., Pistunov, A., Sabban, M., et al. (2012). Failure of the tomato trans-acting short interfering RNA program to regulate AUXIN RESPONSE FACTOR3 and ARF4 underlies the wiry leaf syndrome. *Plant Cell* 24, 3575–3589. doi: 10.1105/tpc.112.100222

Yu, D., Fan, B., MacFarlane, S. A., and Chen, Z. (2003). Analysis of the involvement of an inducible *Arabidopsis* RNA-dependent RNA polymerase in antiviral defense. *Mol. Plant Microbe Interact.* 16, 206–216. doi: 10.1094/MPMI.2003.16.3.206

Zhai, J., Jeong, D. H., De Paoli, E., Park, S., Rosen, B. D., Li, Y., et al. (2011). MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* 25, 2540–2553. doi: 10.1101/gad.177527.111

Zhang, W., Gao, S., Zhou, X., Chellappan, P., Chen, Z., Zhou, X., et al. (2011a). Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. *Plant Mol. Biol.* 75, 93–105. doi: 10.1007/s11103-010-9710-8

Zhang, X., Zhao, H., Gao, S., Wang, W. C., Katiyar-Agarwal, S., Huang, H. D., et al. (2011b). *Arabidopsis* Argonaute 2 regulates innate immunity via miRNA393*-mediated silencing of a golgi-localized SNARE gene, MEMB12. *Cell* 42, 356–366. doi: 10.1016/j.molcel.2011.04.010

Zhang, Z., Chen, H., Huang, X., Xia, R., Zhao, Q., Lai, J., et al. (2011c). BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in *Arabidopsis*. *Plant Cell* 23, 273–288. doi: 10.1105/tpc.110.081695

Zhang, X., Yuan, Y. R., Pei, Y., Lin, S. S., Tuschl, T., Patel, D. J., et al. (2006). Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis* Argonaute1 cleavage activity to counter plant defense. *Genes Dev.* 20, 3255–3268. doi: 10.1101/gad.1495506

Zhang, X., Zhang, X., Singh, J., Li, D., and Qu, F. (2012). Temperature-dependent survival of Turnip crinkle virus-infected *Arabidopsis* plants relies on an RNA silencing-based defense that requires dcl2, AGO2, and HEN1. *J. Virol.* 86, 6847–6854. doi: 10.1128/JVI.00497-12

Zhu, H., Duan, C. G., Hou, W. N., Du, Q. S., Lv, D. Q., Fang, R. X., et al. (2011). Satellite RNA-derived satsiR-12 targeting the 3′ UTR of Cucumber mosaic virus triggers viral RNAs for degradation. *J. Virol.* 85, 13384–13397. doi: 10.1128/JVI.05806-11

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 July 2013; accepted: 14 August 2013; published online: 05 September 2013.
*Citation: Peláez P and Sanchez F (2013) Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences. Front. Plant Sci. **4**:343. doi: 10.3389/fpls.2013.00343*
This article was submitted to Plant-Microbe Interaction, a section of the journal Frontiers in Plant Science.

9.3 Peláez et al., 2014

Title: RNA silencing targets genes from the T-DNA of *Agrobacterium rhizogenes* as in viral infections

Authors: Pablo Peláez[1], Alejandrina Hernández-López[1], Georgina Estrada-Navarrete[1] and Federico Sanchez[1*].

[1]Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México.

***Corresponding Author:** Federico Sanchez. Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, 62210, Cuernavaca, Morelos, México. Tel: +52 (777) 3291653 e-mail: federico@ibt.unam.mx

Running title: RNA silencing in hairy roots

Total word count: 4,873
Introduction word count: 1,215
Material and Methods word count: 1,164
Results word count: 1,188
Discussion word count: 1,210
Acknowledgments word count: 87
Figures: 9
Tables: 0

Summary (word count: 195)

- *Agrobacterium rhizogenes* is a pathogenic bacteria that causes hairy root disease by transferring bacterial DNA into plants. It is an essential tool for industry and research due to its capacity to produce genetically modified organisms. Here, we show that RNA silencing exerts a defensive role against several genes from the transferred DNA (T-DNA) of *A. rhizogenes* as in viral infections.
- High-throughput sequencing of small RNAs (sRNAs), degradome sequencing and gene expression techniques were used to analyze the defensive role carried out by the RNA silencing mechanism against the T-DNA in hairy roots.
- Distinct abundant *A. rhizogenes* T-DNA-derived small RNAs (ArT-sRNAs) belonging to several oncogenes were detected in hairy roots but not in callus. Oncogenes are post-transcriptionally regulated by ArT-sRNAs of 21- and 22-nucleotides. Many T-DNA encoded genes constitute DCL4-dependent phasiRNA-producing loci (PHAS loci). Interestingly, we found that ArT-siRNAs move to untransformed leaves. Also, degradome analysis revealed that ArT-sRNAs target host genes.
- In contrast to *Agrobacterium tumefaciens*-induced tumors and probably *A. rhizogenes*-induced callus, RNA silencing pathways are not suppressed in hairy roots. We suggest that the role of RNA silencing observed in this study has implications in our understanding and usage of this unique plant-bacteria interaction.

Key words: *Agrobacterium rhizogenes*, callus, hairy roots, immunity, RNA silencing, small RNAs, T-DNA, tumors.

Introduction

Agrobacterium rhizogenes and *Agrobacterium tumefaciens* are plant pathogenic bacteria capable of inducing uncontrolled cell proliferation and development of hairy roots and crown gall tumors, respectively. During the infection process a transferred DNA from the bacteria root inducing (Ri) or tumor-inducing (Ti) plasmids is integrated into the plant cells genome. The genes encoded in the T-DNA (oncogenes) lead to abnormal proliferation of cells through changes in signal transduction pathways and transcription factors that affect plant hormones. Transformed cells produce amino acid and sugar derivatives (opines) that are used by bacteria as nutrients and trigger conjugal transfer of T-DNA-carrying plasmids. The unique capability of these bacteria to perform cross-kingdom DNA transfer has been widely exploited to transform plants. *A. tumefaciens* strains with Ti plasmids lacking T-DNA oncogenes do not induce tumor growth and have been used as efficient delivery systems for genetic engineering. Hairy roots harboring the *A. rhizogenes* Ri plasmid T-DNA have become an important tool for basic research and biotechnology as well (Georgiev *et al.*, 2012). They emerge from totipotent callus cells that subsequently differentiate into highly branched ageotropic roots. Whole plants regenerated from hairy roots exhibit a particular phenotype (called hairy root phenotype) that comprises loss of apical dominance, wrinkled leaves, reduced fertility, shortened internodes, stunted growth and abnormal organs (Tepfer, 1984). Several studies have attempted to determine the precise function of the T-DNA genes from *A. rhizogenes* strains in the infection process and in the hairy root phenotype (Britton *et al.*, 2008).

RNA silencing constitutes a defense mechanism against foreign genetic material such as transgenes or viral nucleic acids (Hamilton & Baulcombe, 1999). Small RNAs of 21 to 24 nucleotides (nt) in length trigger post-transcriptional gene silencing (PTGS) or transcriptional gene silencing (TGS) of endogenous or foreign genetic elements (Baulcombe, 2004). Double-stranded RNA (dsRNA) is cleaved by Dicer-like proteins (DCLs) to form sRNAs. These sRNAs are then loaded into Argonaute proteins (AGOs) to direct RNA or DNA silencing via base pairing complementarity. Furthermore, RNA-dependent RNA polymerases (RDRs) produce secondary sRNAs that supports systemic spread of RNA silencing through long distances. Each of the four basic DCLs present in

plants generates sRNAs with a determined size. The size of sRNAs impacts their function (Chapman & Carrington, 2007). DCL1 is required for the biogenesis of 21-nt microRNAs (miRNAs) from imperfectly matched stem-loop structures that are formed from single-stranded transcripts (Park *et al.*, 2002). The biogenesis of 22-nt sRNAs is carried out through DCL2 activity (Gasciolli *et al.*, 2005). Small RNAs of 22-nt may trigger production of phasiRNAs, sRNAs derived of consecutive DCL cleavages from a particular nucleotide start point (Fei *et al.*, 2013). DCL3 generates 24-nt heterochromatic siRNAs (hc-siRNAs) involved in TGS for promoting DNA methylation and histone modifications (Matzke *et al.*, 2009). Finally, DCL4 produces phasiRNAs of 21-nt from RDR6-dependent dsRNA that can be loaded into AGO1 (Vazquez *et al.*, 2004).

Several small RNA biogenesis factors control virus replication and spreading throughout the plant. Virus contend small RNA-mediated plant defense responses with diverse viral suppressors of RNA silencing (VSRs) that may alter the biogenesis, maturation, or function of endogenous and virus-derived small interfering RNAs (vsiRNAs; Peláez & Sanchez, 2013). Viral dsRNA is cleaved mainly by DCL4 and DCL2 into vsiRNAs (Deleris *et al.*, 2006). DNA viruses may be controlled through DCL3-triggered DNA methylation (Blevins *et al.*, 2006). AGO1 is a very important factor in antiviral immunity. It is able to load vsiRNAs and target viral RNAs for further signal amplification (Azevedo *et al.*, 2010; Garcia *et al.*, 2012). RDRs, mostly RDR1 and RDR6, have important roles in the biogenesis of vsiRNAs and in the maintenance of the systemic antiviral immunity as well (Xie *et al.*, 2001; Mourrain *et al.*, 2000). RDRs use primary vsiRNAs or aberrant viral RNA sequences to favor production of secondary vsiRNAs. Interestingly, vsiRNAs regulate host genes like endogenous sRNAs and as fungal small RNAs (Moissiard & Voinnet, 2000; Qi *et al.*, 2009; Shimura *et al.*, 2011; Smith *et al.*, 2011; Weiberg *et al.*, 2013). Small RNA biogenesis factors such as DCL4, DCL2, AGO1 and RDR6 are required also for transgene-induced post-transcriptional gene silencing (Dunoyer *et al.*, 2005; Deleris *et al.*, 2006; Morel *et al.*, 2002; Mourrain *et al.*, 2000).

Transgenes incorporated by Agrobacterium-mediated transformation frequently are subjected to RNA silencing. Voinnet *et al.* (2003) suggested that transgene silencing could be a defense response typically conducted against T-DNA oncogenes of virulent bacteria similar to that in virus infections. Soon after, Dunoyer *et al.* (2006) reported the defensive

role of RNA silencing in virulent *A. tumefaciens* infections. Mutants of *dcl1* and plants deficient in miRNA functions resulted immune to *A. tumefaciens*. Accumulation and activity of miRNAs in tumors were only moderately altered, suggesting the importance of miRNA-regulated processes for tumor development. Besides, *Arabidopsis thaliana rdr6* mutants and viral suppressor expressing plants were more susceptible to the infection. Surprisingly, RNA silencing against T-DNA genes works at the beginning of the infection but is inhibited in tumors. Small RNAs of 21-nt in length from the tryptophan 2-monooxygenase (*iaaM*) and the agropine synthase (*ags*) T-DNA genes were detected at 3 days post-infiltration of virulent *A. tumefaciens* in leaves of *Nicotiana benthamiana*; however, no sRNAs were detected in tumors. The authors proposed that suppression of RNA silencing pathways in *A. tumefaciens*-induced tumors could be an event intrinsic to dedifferentiation and/or proliferation probably caused by phytohormones. In this regard, the 6b protein from the T-DNA region of the Ti plasmid was reported to interact with AGO1 and the SERRATE (SE) protein involved in the biogenesis of miRNAs. The reduced accumulation of miRNAs observed in plants overexpressing the 6b protein suggests that it could act as a suppressor of RNA silencing and may contribute to the RNA silencing suppression state observed in tumors (Wang *et al.*, 2011).

Hairy roots are differentiated roots that have been proven to be a powerful tool for loss-of-function analyses of genes via RNA silencing. We reasoned that *A. rhizogenes*-induced hairy roots should not have the RNA silencing pathways suppressed like *A. tumefaciens*-induced tumors and, as in viral infections, the active small RNA-mediated plant defense responses against the T-DNA genes must have an effect in this biological system. In accordance with this reasoning, we found indeed that RNA silencing pathways are not suppressed in hairy roots, so in this article the defensive role carried out by the RNA silencing mechanism against the T-DNA of *A. rhizogenes* was analyzed. Distinct abundant *A. rhizogenes* T-DNA-derived small RNAs belonging to several oncogenes and to the opine synthase gene were detected. Our data demonstrate that many T-DNA encoded genes constitute DCL4-dependent phasiRNA-producing loci subjected to PTGS triggered by dsRNA-derived ArT-sRNAs of 21- and 22-nt in length. Accumulation of miRNAs in hairy roots is not moderately reduced as in *A. tumefaciens*-induced tumors. We also found that ArT-sRNAs move to untransformed shoots and that sRNAs generated from bacteria's

genetic material target endogenous genes according to degradome sequencing analysis. The low levels of ArT-sRNAs detected in *A. rhizogenes*-induced hairy callus suggest a similar suppression state of the RNA silencing pathways as that observed in *A. tumefaciens*-induced tumors. Our results indicate that small RNA-mediated plant defense responses regulate T-DNA genes of the Ri plasmid of *A. rhizogenes* in hairy roots like transgenes and viruses, suggesting important roles played by RNA silencing in distinctive events of this plant-bacteria interaction such as cellular proliferation, differentiation and immunity.

Materials and Methods

Plant material

Hairy roots and callus were obtained from *Phaseolus vulgaris* L. cv. Negro Jamapa adjusting the procedure described by Estrada-Navarrete *et al.* (2007). Ten days after infection with *A. rhizogenes* K599, the primary root was removed and plants were grown in hydroponics (B&D solution supplemented with 8 mM KNO_3, pH 6.5) for eleven days in a chamber with 16 h of light and 8 h of dark at 28°C until hairy roots and leaves (mainly cotyledonary leaves) collection. For 6 days old callus induction and collection, germinated seeds grown in hydroponics for 2 days were infected in the stem. Common bean wild type roots were obtained from germinated seeds grown for 4 days in hydroponics. All samples were frozen in liquid nitrogen and stored at -80°C until RNA extraction.

RNA extraction and library sequencing

Total RNA was extracted from hairy roots, callus, and leaves using the Trizol reagent (Invitrogen). Ten micrograms of each sample were prepared for deep sequencing of small RNA libraries. Sequences ranging from 18-30 nt were purified for the construction of the libraries. Small RNA library of hairy roots was prepared following Illumina's Small RNA alternative sample preparation protocol v1.5. The small RNA library of callus and leaves was constructed according to the Illumina's TruSeq Small RNA Sample Prep Kit (indexes 1-12). Libraries were separately Single Read-sequenced using the Genome

Analyzer IIx (GAIIx) and the Illumina Cluster Station at the Instituto de Biotecnología (Universidad Nacional Autónoma de México). Degradome library construction for the hairy roots sample was performed as described by Ma *et al.* (2010) with some changes. Approximately 150 ng of poly(A)$^+$ RNA was used to anneal with biotinylated random primers. Strapavidin capture of RNA fragments was performed through biotinylated random primers. 5′ adaptor ligation was only performed to those RNAs containing 5′-monophosphates. The library was sequenced with the Illumina's Cluster Station and the GAIIx using the 5′ adapter only, resulting in the sequencing of the first 36 nt of the inserts that represent the 5′ ends of the original RNAs (LC Sciences).

Bioinformatics and data analysis

Sequence length distribution and alignment (PatMaN; Prüfer *et al.*, 2008) of small RNAs were performed using the UEA small RNA analysis toolkit (Version 2.5.0; Stocks *et al.*, 2012). Small RNAs (18-26 nt) were aligned to the *A. rhizogenes* plasmid Ri T-DNA region without mismatches (GenBank accession number: EF433766). Whole T-DNA alignments were seen with the Integrative Genomics Viewer (IGV; Thorvaldóttir *et al.*, 2013). Small RNA libraries from Peláez *et al.* (2012) were used as controls when looking for ArT-sRNAs in leaves. RNA secondary structures of the T-DNA loci were predicted with the Vienna RNA secondary structure server using the minimum free energy algorithm (MFE; Hofacker, 2003). *In silico* approach to predict *P. vulgaris* genes targeted by T-DNA derived sRNAs was performed using the plant small RNA target analysis server psRNATarget with a maximum expectation threshold of 0-2. *P. vulgaris* (Phytozome v9.0) spliced mRNA transcripts without alternative splice variants were used for target prediction analysis. Multiple alignment of *rolA* sequences from different Ri plasmids (pRi1724/gb: AP002086, pRi8196/gb:M60490, pRiA4/gb:X12579) was performed using ClustalW2 with default settings (Larkin *et al.*, 2007) and viewed with Jalview (Waterhouse *et al.*, 2009). Previously reported microRNA sequences from *P. vulgaris* were used for the identification of microRNAs in the small RNA libraries (Peláez *et al.*, 2012). Normalized read frequencies (TPM) of miRNAs from the same family were added for miRNA accumulation analysis.

Analysis of phasiRNAs

Two strategies were used for the identification of loci producing phasiRNAs of 21 nt in length (Chen *et al.*, 2007; Axtell, 2010). The algorithm developed by Chen *et al.* (2007) was used through the UEA small RNA analysis toolkit implementation. Only sRNAs derived from the T-DNA and the plasmid Ri T-DNA sequence (gb: EF433766) were used as input. Moreover, this algorithm was performed for the identification of phasing-generating loci from *P. vulgaris* used as controls (Table S1). Combined abundances of two and five bins were used in the methodology developed by Axtell (2010) to calculate the phase ratio of *rolB* and *CUS* loci respectively.

Northern blot analysis

Total RNA was extracted using Trizol reagent (Invitrogen). RNA (20ug) was separated by 15% PAGE/8 M urea/ 1x TBE buffer. Gels were electro-blotted to a Hybond-N^+ membrane (GE Healthcare) and UV cross-linked twice. Oligonucleotide probes (Table S2) were labelled using [γ-^{32}P]-ATP (Perkin-Elmer) and T4 PNK (New England Biolabs). Labeled probes were purified with the Quick spin-oligo columns (Roche) before hybridization. Hybridizations were performed at 42°C overnight in UltraHyb-oligo solution (Ambion). After hybridization, membranes were washed twice in 2 x SSC/0.1% SDS and exposed to the Phosphor Screen System (GE Healthcare). The screen was scanned in a Storm 860 Gel and Blot Imaging System (GE Healthcare). As loading control, an oligonucleotide probe complementary to the U6 small nuclear RNA was used. Signal intensities were quantified using the ImageQuant 5.2 software (Molecular Dynamics). The U6 signal intensity in each blot was used for normalization and calculation of expression ratios.

Degradome data analysis

Perfect alignments of degradome sequences to the T-DNA were performed using PatMaN (Prüfer *et al.*, 2008). The PAREsnip tool was used for the discovery of small RNA-guided cleavage sites with default parameters (Folkes *et al.*, 2012). The small RNAs derived from the T-DNA, the T-DNA sequence (gb: EF433766) and the degradome sequences were used as input for the characterization of sRNA-mediated cleavages on the *A. rhizogenes* plasmid Ri T-DNA region. Primary transcripts (Phytozome v9.0) were used as input for the discovery of *P. vulgaris* targets cleaved by T-DNA derived sRNAs. Host targets were discarded if other sRNA from the library was also predicted to cleave the transcript and presented a better hit in at least one of the parameters evaluated such as alignment score, category, reads abundance or p-value. *Arabidopsis* homolog targets were functionally grouped according to the Gene Ontology (GO) categories (Berardini *et al.*, 2004) using the TAIR GO Annotation Bulk Download and Analysis tool (http://www.arabidopsis.org/tools/bulk/go/index.jsp). Target plots (t-plots) were performed for does target-sRNA interactions with cleaved fragment frequencies greater or equal to five and sRNA frequencies greater or equal to four. MicroRNAs (Table S3) were used for analysis of validated target-sRNA interactions in the degradome library (Table S4).

Quantitative PCR

Quantification of transcripts accumulation from hairy roots and callus was performed using the iQ5 Real-Time PCR Detection System and the iQ5 Optical System Software (Bio-Rad). First strand cDNA was synthesized from DNA-free RNA with the RevertAid H Minus First Strand cDNA Synthesis kit (Fermentas). The qRT-PCR reactions were carried out in triplicate for three biological replicates with the Maxima SYBR green-fluorescein qPCR master mix (Fermentas). The gene-specific oligonucleotides belonging to the T-DNA genes were tested also on non-transformed tissues as controls (Table S5). The melting temperature was set at 60°C. Quantification was based on a cycle threshold (Ct) value and the genes expression levels were normalized with the elongation factor 1 (*EF1*) reference gene. Genes with differential transcript accumulation between hairy roots and callus were identified using the parametric Limma test with a Benjamini-Hochberg p-value

correction. Expression values of genes from callus were used as background to show differential expression.

Results

Small RNAs from the T-DNA of *Agrobacterium rhizogenes*

Small RNA libraries from hairy roots and callus were generated to explore RNA silencing activity across the T-DNA. Small RNA sequencing yielded 12,396,242 and 27,272,479 total raw reads for the hairy roots and callus samples, respectively. The length distribution of trimmed small RNA sequences showed that in hairy roots the 21-nt (17,9%) and the 24-nt (19%) classes were similar in abundance (Fig. 1a). The 21-nt class was the second most diverse in hairy roots and the fourth in callus. Sequence alignment to the T-DNA of *A. rhizogenes* revealed the presence of several distinct and abundant ArT-sRNAs in hairy roots (Fig. 1b). 3,176 distinct Art-sRNAs were identified corresponding to 17,000 transcripts per million (TPM). In contrast, very few ArT-sRNAs were detected in callus. Only 563 unique sequences corresponding to an abundance of 90 TPM were encountered (Fig. 1c; Table S6; De Paoli *et al.*, 2009). Most of the Art-sRNAs observed in hairy roots aligned with transcriptional units of the T-DNA belonging to genes such as *ORF2*, *ORF8*, *rolA*, *rolB*, *ORF13*, *ORF14* and the opine *CUS* (Fig. 1b; Fig. S1). The oncogene *rolA* constitutes the major source of Art-sRNAs. Interestingly, virtually no Art-sRNAs derived from the oncogene *rolC* were detected in the library of hairy roots. Also, nothing but 134 unique Art-sRNAs were shared between hairy roots and callus. Moreover, production of miRNAs in hairy roots was not inhibited (Fig. 2a). Most of all miRNAs analyzed were less accumulated in callus than in hairy roots (Fig. 2b). Unusually, microRNA miR319 was de most abundant miRNA in callus (Peláez *et al.*, 2012). Surprisingly, ArT-sRNAs were detected also in untransformed leaves of composite plants (Fig. 3).

Post-transcriptional gene silencing of the phasiRNA-producing T-DNA genes

Art-sRNAs were largely small RNAs of 21-nt in length (Fig. 4a). In hairy roots, 52,2% of the Art-sRNAs representing 67.9% of the total sequences aligned to the T-DNA were of 21-nt in length. Only 5% of the total sequences corresponded to the 24-nt class. The 22-nt class of Art-sRNAs was the second most diverse (18%) and abundant (19.7%). A similar hierarchical pattern of accumulation was observed for the Art-sRNAs detected in callus. Art-sRNAs were produced from the sense and antisense strands (Fig. 4b). The sense and antisense strand ratios of Art-sRNAs were closer to one. 52% of the Art-sRNAs found in hairy roots aligned to the sense strand while 48% aligned to the antisense strand. Taking a closer look to particular genes encoded in the sense or antisense strand no bias was observed for either one of the strands in the production of Art-sRNAs (Fig. 5). Furthermore, phased secondary 21-nt small interfering RNAs characteristic of PHAS loci were identified for *ORF8*, *rolA*, *rolB*, *ORF13*, *ORF14* and *CUS* genes (Fig. 6a,b,c). Genes like *rolA*, *ORF8*, *ORF13* and *ORF14* exhibited one sRNA cluster and one predominant phase, whereas *rolB* and *CUS* presented two clusters without one main phase. Some phased small RNAs of the *rolA* transcript were among the most abundant sRNAs detected. Stem-loop-like structures were observed in the RNA secondary structures of T-DNA transcripts which might be processed directly by DCLs (Fig. S2). Transcript accumulation analysis of the T-DNA genes, as well as *AGO1* transcript, showed that *rolA* was the most accumulated transcript in callus and the least accumulated in hairy roots. According to the differential accumulation analysis of transcripts between hairy roots and callus only the *rolB* and *CUS* transcripts were significantly up-regulated in hairy roots (Fig. 7).

ArT-sRNAs-mediated T-DNA cleavage

Forty percent of the ArT-sRNAs of 21-nt in length had a 5’ terminal uracil (Fig. S3). Small RNAs with this characteristic are mainly recruited by AGO1. To gain insight on cleavage of T-DNA transcripts through AGOs/ArT-sRNAs associations, a degradome (also called parallel analysis of RNA ends, PARE) library from hairy roots was deep-sequenced and analyzed. Degradome sequencing yielded 19,064,598 total raw reads. 446 distinct sequences aligned to the T-DNA of *A. rhizogenes*. Degradome sequences corresponding to all of the coding sequences of the T-DNA genes except for *ORF13* were detected (Fig. S4;

Fig. S5). Most of the reads detected had the sequence of the *rolA* and *CUS* genes. Discovery analysis of sRNA/target interactions using degradome and small RNA libraries generated from hairy roots exposed 180 interactions between ArT-sRNAs and the transcripts of the T-DNA genes (Table S7). These interactions constitute 73 unique cleavage positions. Thirty four and twenty percent of the sRNAs of the interactions were of 21 and 22 nucleotides in length, respectively. According to degradome read abundance, 5 interactions fell into category 0, 116 into category 2, and 59 into category 4. Almost all of the interactions found, around 95.5%, pointed out cleavage sites of the *rolA* transcript, mainly from the 3' untranslated region (Fig. 8a). Interestingly, one of the ArT-sRNAs that fell into category 0, and thus cleaved the most abundant fragment, was the most abundant ArT-sRNA in the library. This ArT-sRNA of 22-nt in length with 3,304 absolute read counts in the library presented another cleavage position including two mismatches. This cleavage position in the transcript of *rolA* synchronized with the phasing register of the detected *rolA* phase. In addition, the sequence of this ArT-sRNA was found to be highly conserved among *rolA* sequences of the pRi8196, the pRi1724, the pRiA4 and the pRi2659 Ri plasmids (Fig. 8b).

Host gene silencing through Art-sRNAs

Art-sRNAs may present the required complementarity base pairing with host gene sequences to regulate them. To search for complementarity base pairing between Art-sRNAs and host transcripts (*Phaseolus vulgaris*) a target prediction analysis was performed. We identified 399 unique Art-sRNAs/target interactions using a stringent cut-off threshold of the complementarity score (Table S8). The interactions detected were composed by 299 distinct transcripts and 326 distinct Art-sRNAs. To provide further evidence of ArT-sRNAs-mediated cleavage of host genes, we analyzed the degradome library obtained from hairy roots in pursuit of interactions. 402 interactions between host targets and ArT-sRNAs were detected (Table S9). The total number of unique host transcripts predicted to be cleaved by 346 unique ArT-sRNAs was 228. Most of the ArT-sRNAs that define these interactions were of 21-nt in length (53.4%). Taking into account all the interactions, 58 fell into category 0, 50 into category 1, 155 into category 2, 3 into

category 3, and 136 into category 4. The host targets identified comprises a diverse set of protein families (Table S10). Gene Ontology (GO) classifications were assigned to ninety eight of the unique host targets (Fig. 9a). The three most abundant groups within the cellular component category were cytoplasmic components (17.2%), nucleus (17.2%), and intracellular components (16.7%). Cellular processes (28.2%), metabolic processes (20%), and responses to stress (9.6%) were the most abundant groups observed from the biological process category. In the molecular function category, the most abundant groups corresponded to binding (16.4%), enzyme activity (14.23%), and transferase activity (12.2%). Furthermore, six target-sRNA interactions with abundant cleaved fragment and sRNA frequencies were selected for target-plot (t-plot) display (Fig. 9b). The targets of these interactions included proteins involved in plant immunity and cell cycle.

Discussion

Agrobacterium rhizogenes pathogenicity is characterized by the insertion of the T-DNA into the plants nuclear genome. For this reason, it represents a powerful system to transform plants. Also, it is capable to transform recalcitrant plants that even *A. tumefaciens* cannot. Recent progress on the understanding of the induction and regeneration of hairy roots has increased its usage in the production of recombinant proteins, metabolic engineering and phytoremediation. In contrast to *A. tumefaciens*, *A. rhizogenes* is frequently used with the root-inducing plasmid "armed", carrying the T-DNA oncogenes. For years, *A. rhizogenes* oncogenes have been considered modulators of neoplastic growth and differentiation. Recently, they have been also identified as triggers of secondary metabolism. The biochemical and molecular characterization of the T-DNA genes from *A. rhizogenes* is not completely understood. In several studies, conflicting results regarding their functions have been reported, suggesting that new meristem emergence and subsequent differentiation of transformed plant cells may be conducted through complex molecular mechanisms. In recent years, small RNAs have emerged as important regulators of gene expression during different biological processes. Although RNA silencing mechanism has been reported to regulate transgenes, viral nucleic acids and genes from the

T-DNA of *A. tumefaciens*, its regulatory activity against the T-DNA of *A. rhizogenes* in hairy roots had not been analyzed.

In this study, several genes from the T-DNA of *A. rhizogenes* were found to be regulated by the RNA silencing mechanism in hairy roots. Production of small RNAs in hairy roots was not inhibited contrary to what was reported in *A. tumefaciens*-induced tumors (Dunoyer *et al.*, 2006). The amount of reads detected for several ArT-sRNAs were similar to some abundant endogenous sRNAs detected in several common bean organs (Peláez *et al.*, 2012). Differences found between the total amount of reads corresponding to ArT-sRNAs detected in callus and in hairy roots resembled those between wild type and chalcone synthase-A (*CHS-A*) silenced petunias (De Paoli *et al.*, 2009). The few ArT-sRNAs detected in callus, together with the reduced accumulation of most miRNAs compared with hairy roots, suggests that callus also present a RNA silencing suppression state similar to tumors. The poor number of ArT-sRNAs shared between hairy roots and callus may be a consequence of different sources of dsRNA used as templates for their production. For example, the T-DNA perhaps is regulated by RNA silencing before its integration. Furthermore, differential accumulation of miRNAs observed between wild type roots and hairy roots might imply phenotypic differences in these two types of roots. Interestingly, miR319 was the most abundant miRNA family in the library of callus. Usually, in all the other sequenced small RNA libraries from *P. vulgaris* miR159 was the most abundant miRNA. The miR319 family targets TCP (TEOSINTE BRANCHED-CYCLOIDEA/PROLIFERATING CELL FACTORS) genes involved in hormone and signaling pathways, proliferation and differentiation. Reduced activity of TCP genes caused by the increase of miR319 accumulation intensifies cellular proliferation (Schommer *et al.*, 2012). We think that this miRNA is important for callus formation. On the other hand, in the transcript accumulation analysis, we expected all of the transcripts of the T-DNA to be significantly up-regulated in hairy roots compared with callus if they were not silenced because tumors only contain a small amount of transformed cells.

Among the T-DNA transcripts producing ArT-sRNAs, the *rolA* transcript stands out as a major source of small RNAs. Curiously, no ArT-sRNAs belonging to the 5' UTR intron of *rolA* were identified. In accordance with all the T-DNA transcripts analyzed, accumulation of *rolA* drastically differs between hairy roots and callus. The precise

function of the rolA protein is still unknown. Poor expression of *rolA* has been reported in roots and in leaves. Tobacco plants expressing *rolA* present wrinkled leaves, shortened internodes and abnormal flowers (Carneiro & Vilaine, 1993). Local expression of *rolA* in the vascular bundles reduces the size of the surrounding parenchyma cells causing the wrinkled leaf phenotype. For this reason, and reciprocal grafting studies involving *rolA*, it has been proposed that *rolA* generates a diffusible factor (Guivarc'h *et al.*, 1996; Britton *et al.*, 2008). We think that ArT-sRNAs derived from *rolA* are good candidates to constitute this diffusible factor. Also, another possibility for ArT-sRNAs, from *rolA* or in general, to indirectly impact on plants phenotype could be through altering endogenous sRNA-regulated processes as described during PTGS of foreign genetic material (Martínez de Alba *et al.*, 2011). Moreover, it was intriguing that no ArT-sRNAs were related to the *rolC* gene in hairy roots, one of the T-DNA genes widely studied as well. Strong and stable expression of *rolC* in roots of transformed plants previously reported together with the null detection of ArT-sRNAs generated from this gene, suggests that *rolC* is not subjected to silencing in roots (Sugaya *et al.*, 1989).

Small RNAs are systemic silencing signals able to move from cell to cell and over long distances. Systemic silencing occurs during PTGS of transgenes or viruses (Brosnan & Voinnet, 2011). Long distance movement of small RNAs has been observed predominantly to take place from the shoot-to-root direction. In this study, we detected ArT-sRNAs in untransformed leaves of composite plants. These ArT-sRNAs were quite scarce and poorly abundant; however, this result strongly suggests that small RNAs generated from hairy roots by expressing hairpin RNAs may trigger systemic silencing throughout the plant. In this sense, weak silencing of the β-glucuronidase (*GUS*) reporter gene was observed in leaves of *Lotus japonicus* transgenic lines that had hairy roots expressing a hairpin RNA complementary to the gene (Kumagai & Kouchi, 2003).

The detection of ArT-sRNAs mainly of 21- and 22-nt in length shows us that the T-DNA genes are silenced at the post-transcriptional level and that their transcripts are processed by DCL4 and DCL2 as occurs during regulation of RNA virus. Also, the identification of ArT-sRNAs as phasiRNAs points out an active role of one or more RDRs during the regulation of these genes. Besides, cleavage analysis of the T-DNA by ArT-sRNAs indicate that this small RNAs are loaded into AGO proteins. Most of them could be

loaded onto AGO1 which has an important role in antiviral silencing. The interactions detected for conserved microRNAs using the PAREsnip algorithm and the degradome data were mainly validated interactions. For this reason, the amount of interactions detected between ArT-sRNAs and host transcripts suggests that ArT-sRNAs may shape hairy roots phenotype through direct regulation of the host transcripts. Further experiments are required to truly validate this regulation.

We think that our results contribute to our understanding and usage of hairy roots. For example, we propose that constructs for overexpression used in hairy roots may consider the population of ArT-sRNAs to avoid silencing. Also, it would be interesting to determine if the defensive role of RNA silencing against the T-DNA genes has an impact on the efficiency of constructs designed to silence genes. In addition, we consider that studies to elucidate the function of several T-DNA genes, especially of *rolA*, have to take into account these observations. Our results also suggest that more genes of the T-DNA of *A. tumefaciens* could be silenced at the beginning of the infection (Dunoyer *et al.*, 2006). It would be interesting to know if silencing the T-DNA genes favours hairy root emergence or until differentiation takes place the RNA silencing mechanism regulates them. We do not discard that the T-DNA of *A. rhizogenes* may produce a protein like 6b (Wang *et al.*, 2011).

Acknowledgements

We thank the Unidad Universitaria de Secuenciación Masiva de DNA (UUSMD-UNAM) for sequencing services. We are grateful to the Instituto de Biotecnología-UNAM for giving us access to its computer cluster and to Dr. Oswaldo Valdés, Dra. Georgina Hernández and Lic. Olivia Santana for technical support. Financial support for this research was provided by Consejo Nacional de Ciencia y Tecnología (CONACYT). Pablo Peláez was supported by a fellowship from CONACYT and this work is part of his Ph.D. thesis at the PhD program in Biomedical Sciences (UNAM).

References

Axtell MJ. **2010**. A method to discover phased siRNA loci. In: Meyers BC, Green PJ, eds. *Plant MicroRNAs, Methods in molecular biology*. New York, NY, USA: Humana Press, 59-70.

Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, Braun L, Bergdoll M, Hakimi MA, Lagrange T, *et al.* 2010. Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. *Genes & development* **24**: 904–915.

Baulcombe D. **2004**. RNA silencing in plants. *Nature* **431**: 356–363.

Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, *et al.* 2004. Functional annotation of the Arabidopsis genome using controlled vocabularies. *Plant physiology* **135**: 745–755.

Blevins T, Rajeswaran R, Shivaprasad P V., Beknazariants D, Si-Ammour A, Park HS, Vazquez F, Robertson D, Meins F, Hohn T, *et al.* 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Research* **34**: 6233–6246.

Britton MT, Escobar MA, Dandekar AM. 2008. The Oncogenes of *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. In: Tzfira T, Citovsky V, eds. *Agrobacterium: from biology to biotechnology.* Springer, NY, 523-563.

Brosnan CA, Voinnet O. **2009**. The long and the short of noncoding RNAs. *Curr Opin Cell Biol* **21**: 416–425.

Carneiro M, Vilaine F. **1993**. Differential expression of the rolA plant oncogene and its effect on tobacco development. *The Plant journal : for cell and molecular biology* **3**: 785–792.

Chapman EJ, Carrington JC. **2007**. Specialization and evolution of endogenous small RNA pathways. *Nature reviews. Genetics* **8**: 884–896.

Chen H-M, Li Y-H, Wu S-H. **2007**. Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 3318–3323.

Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. **2006**. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science (New York, N.Y.)* **313**: 68–71.

Dunoyer P, Himber C, Voinnet O. **2005**. DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nature genetics* **37**: 1356–1360.

Dunoyer P, Himber C, Voinnet O. **2006**. Induction, suppression and requirement of RNA silencing pathways in virulent Agrobacterium tumefaciens infections. *Nature genetics* **38**: 258–263.

Estrada-Navarrete G, Alvarado-Affantranger X, Olivares J-E, Guillén G, Díaz-Camino C, Campos F, Quinto C, Gresshoff PM, Sanchez F. **2007**. Fast, efficient and reproducible genetic transformation of Phaseolus spp. by Agrobacterium rhizogenes. *Nature protocols* **2**: 1819–1824.

Fei Q, Xia R, Meyers BC. **2013**. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *The Plant cell* **25**: 2400–15.

Folkes L, Moxon S, Woolfenden HC, Stocks MB, Szittya G, Dalmay T, Moulton V. **2012**. PAREsnip: A tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. *Nucleic Acids Research* **40**.

Garcia D, Garcia S, Pontier D, Marchais A, Renou JP, Lagrange T, Voinnet O. **2012**. Ago Hook and RNA Helicase Motifs Underpin Dual Roles for SDE3 in Antiviral Defense and Silencing of Nonconserved Intergenic Regions. *Molecular Cell* **48**: 109–120.

Gasciolli V, Mallory AC, Bartel DP, Vaucheret H. **2005**. Partially redundant functions of arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-Acting siRNAs. *Current Biology* **15**: 1494–1500.

Georgiev MI, Agostini E, Ludwig-Müller J, Xu J. **2012**. Genetically transformed roots: From plant disease to biotechnological resource. *Trends in Biotechnology* **30**: 528–537.

Guivarc'h A, Carneiro M, Vilaine F, Pautot V, Chriqui D. **1996**. Tissue-specific expression of the rolA gene mediates morphological changes in transgenic tobacco. *Plant molecular biology* **30**: 125–134.

Hamilton AJ, Baulcombe DC. **1999**. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science (New York, N.Y.)* **286**: 950–952.

Hofacker IL. **2003**. Vienna RNA secondary structure server. *Nucleic Acids Research* **31**: 3429–3431.

Kumagai H, Kouchi H. **2003**. Gene silencing by expression of hairpin RNA in Lotus japonicus roots and root nodules. *Molecular plant-microbe interactions : MPMI* **16**: 663–668.

Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.

Ma Z, Coruh C, Axtell MJ. **2010**. Arabidopsis lyrata small RNAs: transient MIRNA and small interfering RNA loci within the Arabidopsis genus. *The Plant cell* **22**: 1090–1103.

Martínez De Alba AE, Jauvion V, Mallory AC, Bouteiller N, Vaucheret H. **2011**. The miRNA pathway limits AGO1 availability during siRNA-mediated PTGS defense against exogenous RNA. *Nucleic Acids Research* **39**: 9339–9344.

Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ. **2009**. RNA-mediated chromatin-based silencing in plants. *Current Opinion in Cell Biology* **21**: 367–376.

Moissiard G, Voinnet O. **2006**. RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 19593–19598.

Morel J-B, Godon C, Mourrain P, Béclin C, Boutet S, Feuerbach F, Proux F, Vaucheret H. **2002**. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *The Plant cell* **14**: 629–639.

Mourrain P, Béclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, *et al.* 2000. Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101**: 533–542.

De Paoli E, Dorantes-Acosta A, Zhai J, Accerbi M, Jeong D-H, Park S, Meyers BC, Jorgensen RA, Green PJ. **2009**. Distinct extremely abundant siRNAs associated with cosuppression in petunia. *RNA (New York, N.Y.)* **15**: 1965–1970.

Park W, Li J, Song R, Messing J, Chen X. **2002**. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Current Biology* **12**: 1484–1495.

Peláez P, Sanchez F. **2013**. Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences. *Frontiers in plant science* **4**: 343.

Peláez P, Trejo MS, Iniguez LP, Estrada-Navarrete G, Covarrubias AA, Reyes JL, Sanchez F. **2012**. Identification and characterization of microRNAs in Phaseolus vulgaris by high-throughput sequencing. *BMC Genomics* **13**: 83.

Prüfer K, Stenzel U, Dannemann M, Green RE, Lachmann M, Kelso J. **2008**. PatMaN: Rapid alignment of short sequences to large databases. *Bioinformatics* **24**: 1530–1531.

Qi X, Bao FS, Xie Z. **2009**. Small RNA deep sequencing reveals role for Arabidopsis thaliana RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS ONE* **4**.

Schommer C, Bresso EG, Spinelli SV, Palatnik JF. 2012. Role of MicroRNA miR319 in Plant Develompent. In: Sunkar R, ed. *MicroRNAs in Plant development and Stress Responses, Signaling and Communication in Plants.* Springer Berlin Heidelberg, 29-47.

Shimura H, Pantaleo V, Ishihara T, Myojo N, Inaba J ichi, Sueda K, Burgyán J, Masuta C. **2011**. A Viral Satellite RNA Induces Yellow Symptoms on Tobacco By Targeting a Gene Involved in Chlorophyll Biosynthesis Using the RNA Silencing Machinery. *PLoS Pathogens* **7**.

Smith NA, Eamens AL, Wang MB. **2011**. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathogens* **7**.

Stocks MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T, Moulton V. **2012**. The UEA sRNA workbench: A suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. *Bioinformatics* **28**: 2059–2061.

Sugaya S, Hayakawa K, Handa T, Uchimiya H. 1989. Cell-specific expression of the rolC gene of the TL-DNA of Ri plasmid in transgenic tobacco plants. *Plant cell physiology* **305:** 649-653

Tepfer D. **1984**. Transformation of several species of higher plants by Agrobacterium rhizogenes: sexual transmission of the transformed genotype and phenotype. *Cell* **37**: 959–967.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. **2013**. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**: 178–192.

Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gasciolli V, Mallory AC, Hilbert JL, Bartel DP, Crété P. **2004**. Endogenous trans-acting siRNAs regulate the accumulation of arabidopsis mRNAs. *Molecular Cell* **16**: 69–79.

Voinnet O, Rivas S, Mestre P, Baulcombe D. **2003**. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant Journal* **33**: 949–956.

Wang M, Soyano T, Machida S, Yang J-Y, Jung C, Chua N-H, Yuan YA. **2011**. Molecular insights into plant cell proliferation disturbance by Agrobacterium protein 6b. *Genes & development* **25**: 64–76.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. **2009**. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189–1191.

Weiberg A, Wang M, Lin F-M, Zhao H, Zhang Z, Kaloshian I, Huang H-D, Jin H. **2013**. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science (New York, N.Y.)* **342**: 118–23.

Xie Z, Fan B, Chen C, Chen Z. **2001**. An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 6516–6521.

Figure legends

Fig. 1 Small RNAs in hairy roots and callus. (a) Sequence length distribution of small RNAs. Average percentage (Y-axis) of unique (blue bars) and redundant (green bars) sequences of 18-26 nt length (X-axis) for the hairy roots and the callus libraries. (b) Abundance and distribution of ArT-sRNAs. The small RNAs that aligned to the plus strand (red bars) and to the minus strand (black bars) of the T-DNA of *A. rhizogenes* (greenish blue long bar) are indicated for the hairy roots (upper part) and callus (lower part) libraries. The abundance scale (right) differs between the libraries. In the bottom, the coding regions and the borders from the T-DNA are represented (greenish blue bars). (c) Comparison of the number of ArT-sRNAs detected in hairy roots and callus. Abundance levels were graphed as transcripts per million.

Fig. 2 Accumulation of microRNAs in hairy roots and callus. (a) Northern blot analysis of microRNAs in hairy roots and wild type roots of common bean. The signal of the U6 small nuclear RNA was used as loading control. The accumulation ratios between wild type roots and hairy roots were calculated (upper numbers in gels). A different member of 22-nt in length (*) was detected also for the miR167 family. (b) Differential accumulation analysis of microRNAs from hairy roots and callus libraries. Frequencies of reads were normalized (TPM) and represented in logarithmic scale (Y-axis).

Fig. 3 Movement of ArT-sRNAs. Schematic representation of the movement of ArT-sRNAs (black lines) from hairy roots (gray) to non-transformed leaves (green) in composite plants.

Fig. 4 Size and strand origin of ArT-sRNAs. (a) Sequence length distribution of ArT-sRNAs found in the hairy roots and callus libraries. (b) Average percentage of unique and redundant sequences that aligned with the plus strand (gray) and the minus strand (blue) of the T-DNA for each of the small RNA libraries.

672 **Fig. 5** ArT-sRNAs mapped to T-DNA genes. Length distribution (column plots)
673 and strand origin (circular plots) of unique and redundant reads of ArT-sRNAs for
674 particular T-DNA genes (region according to base-pair positions in gb: EF433766).
675 **Fig. 6** PhasiRNA-producing loci from the T-DNA. (a) Radial graphs representing
676 the phasing register of small RNAs as performed by Axtell (2010). The 21 spokes of each
677 graph correspond to the possible phasing registers. (b) Analysis of phasiRNA-producing
678 loci represented in radial graphs. The genes (˜), the region considered in the analysis (*)
679 and the distinct small RNAs that mapped the region are indicated (ǂ). The region
680 considered for the *rolB* and *CUS* genes is greater (˘). Loci with phase ratios of less than
681 0.25 were discarded. The *TAS3* transcript was used as control. (c) Identification of
682 phasiRNAs as described by Chen *et al.* (2007) taking into account ArT-sRNAs with one
683 (upper part) or more than two (lower part) absolute reads (Ⱨ). The strand of the loci (Σ), the
684 number of phased ArT-sRNAs (∧) and the P-value (£) are indicated.
685 **Figure 7** Quantitative PCR analysis of transcripts in hairy roots and callus. The
686 expression of transcripts was normalized with the reference gene *EF1*. The expression
687 values of the transcripts from the callus sample were used as background to show
688 differential expression (n=3; *P<0.05; **P<0.005).
689 **Figure 8** Cleavage analysis of *rolA* by ArT-sRNAs. (a) T-plot showing different
690 interactions between the *rolA* transcript and ArT-sRNAs. The interactions between *rolA*
691 and the most abundant ArT-sRNA of 22-nt and its variants are shown (gray boxes). In the
692 T-DNA (X-axis), the coding region for *rolA* is indicated (bottom). (b) Multiple alignment
693 of *rolA* sequences from different Ri plasmids. The ArT-sRNA of 22-nt is indicated (box
694 and red arrow).
695 **Figure 9** Host candidate targets regulated by ArT-sRNAs. (a) GO categorization of
696 98 predicted common bean targets of ArT-sRNAs. Percentages correspond to the total
697 number of targets within the category (Y-axis). (b) T-plots related to six different
698 interactions identified between abundant host degradome fragments and ArT-sRNAs. Base
699 pair interactions are indicated (gray boxes).
700
701
702

Fig. 1

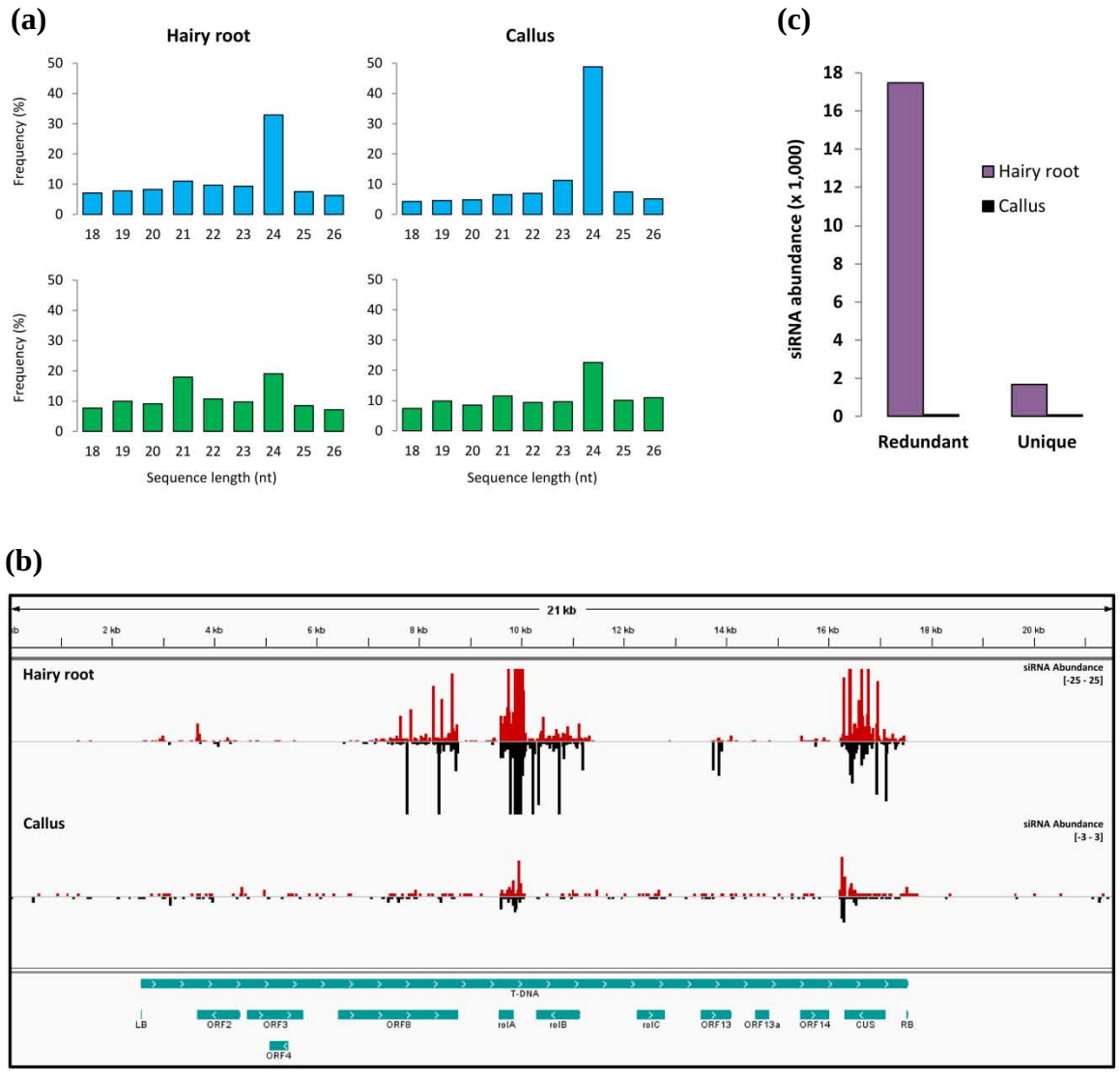

(a)
Hairy root
Callus
Frequency (%)
Sequence length (nt)
(c)
siRNA abundance (x 1,000)
Hairy root
Callus
Redundant
Unique
(b)
21 kb
Hairy root
siRNA Abundance
[-25 - 25]
Callus
siRNA Abundance
[-3 - 3]
T-DNA
LB
ORF2
ORF3
ORF4
ORF8
rolA
rolB
rolC
ORF13
ORF13a
ORF14
CUS
RB

Fig. 2

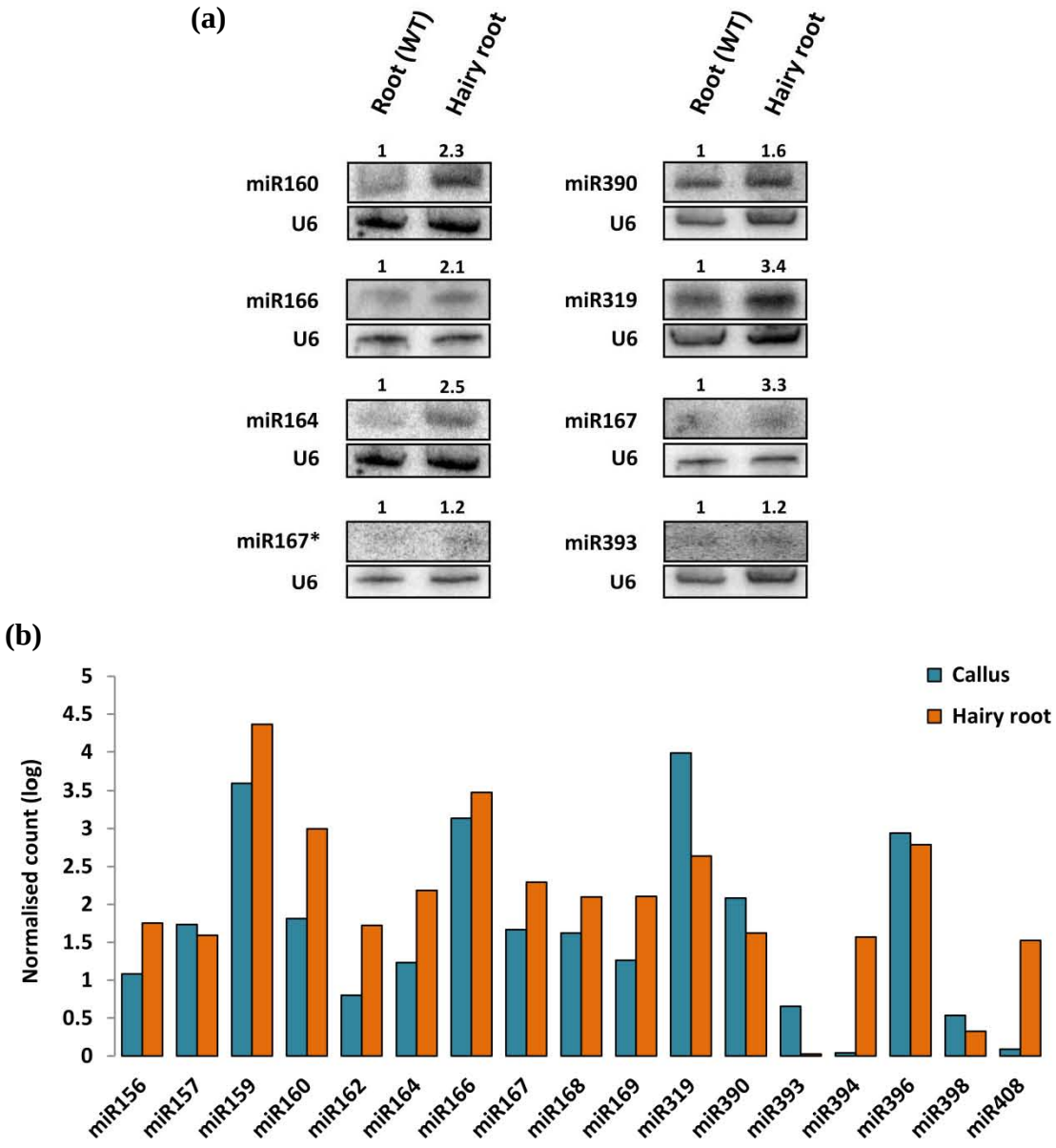
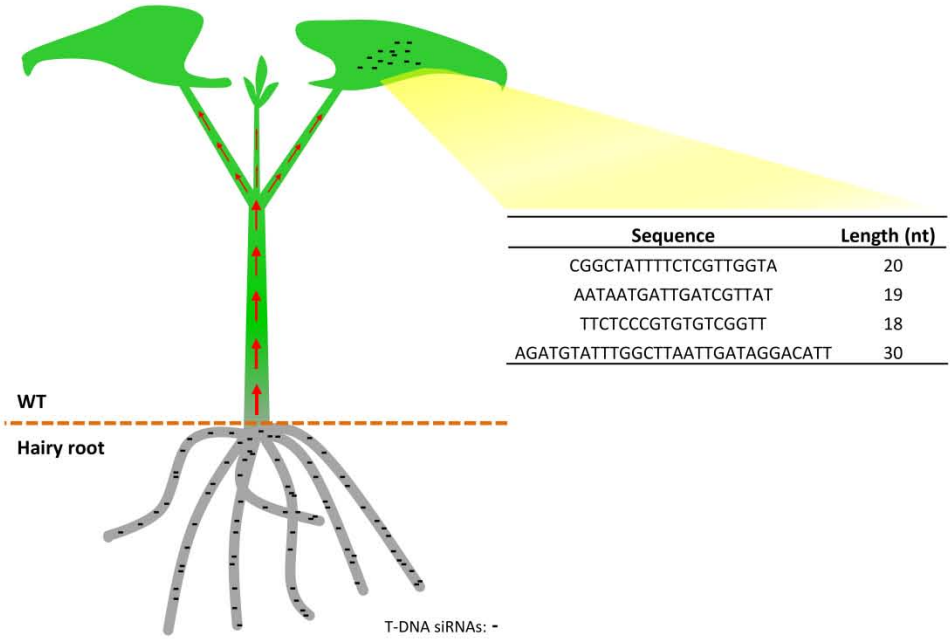

(a)
Root (WT)
Hairy root
Root (WT)
Hairy root
1
2.3
miR160
U6
1
1.6
miR390
U6
1
2.1
miR166
U6
1
3.4
miR319
U6
1
2.5
miR164
U6
1
3.3
miR167
U6
1
1.2
miR167*
U6
1
1.2
miR393
U6
(b)
Callus
Hairy root
Normalised count (log)
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
miR156
miR157
miR159
miR160
miR162
miR164
miR166
miR167
miR168
miR169
miR319
miR390
miR393
miR394
miR396
miR398
miR408

Fig. 3



Sequence	Length (nt)
CGGCTATTTTCTCGTTGGTA	20
AATAATGATTGATCGTTAT	19
TTCTCCCGTGTGTCGGTT	18
AGATGTATTTGGCTTAATTGATAGGACATT	30

Fig. 4

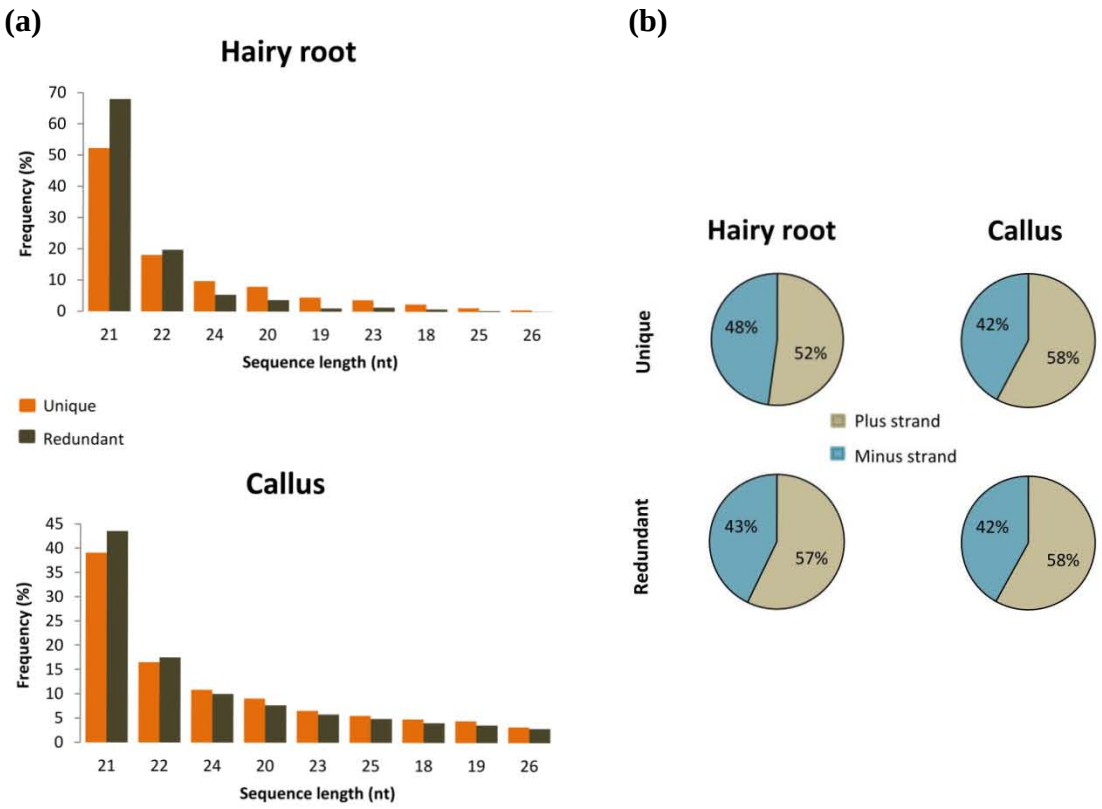

(a)
Hairy root
Frequency (%)
70
60
50
40
30
20
10
0
21 22 24 20 19 23 18 25 26
Sequence length (nt)
Unique
Redundant
Callus
Frequency (%)
45
40
35
30
25
20
15
10
5
0
21 22 24 20 23 25 18 19 26
Sequence length (nt)
(b)
Hairy root
Callus
Unique
48%
52%
42%
58%
Plus strand
Minus strand
Redundant
43%
57%
42%
58%

Fig. 5

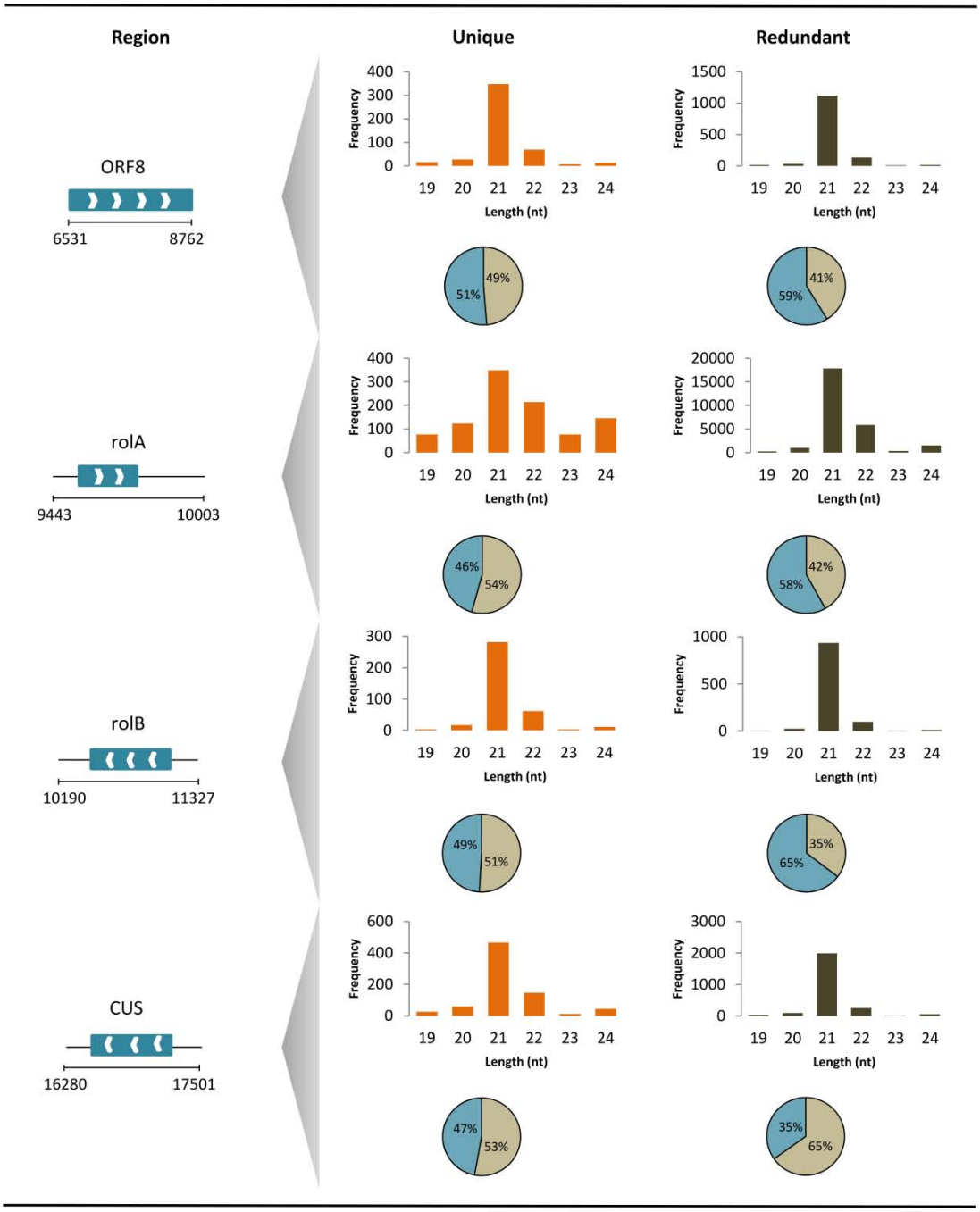
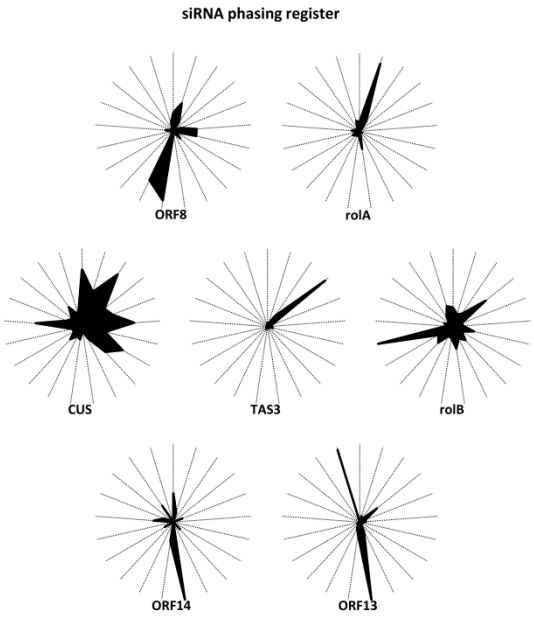

Region
Unique
Redundant
ORF8
6531
8762
rolA
9443
10003
rolB
10190
11327
CUS
16280
17501
Frequency
Length (nt)
19 20 21 22 23 24
51%
49%
59%
41%
46%
54%
58%
42%
49%
51%
65%
35%
47%
53%
35%
65%

Fig. 6

(a)



(b)

Phased sRNA loci from the T-DNA

Gene~	Region*	Distinct Seqs.ǂ	Phase ratio
ORF8	8649 - 8942	37	0.25
rolA	9541 - 10020	350	0.4
rolBˇ	10231 - 11360	269	0.3
ORF13	13681 - 13980	26	0.33
ORF14	15721 - 16010	16	0.36
CUSˇ	16381 - 17340	386	0.47
TAS3	398 - 712	81	0.6

(c)

PhasiRNA-generating loci from T-DNA

Min. Abundance (1)ǀ

Region*	Gene	StrandΣ	Distinct Seqs.ǂ	Distinct Phased Seqs.^	P-value£
10251-10502	rolB	W	61	8	3.01E-03
11074-11325	rolB	W	38	6	4.55E-03
13711-13962	ORF13	C	25	5	3.50E-03
15731-15982	ORF14	C	13	5	1.22E-04
16404-16655	CUS	W	139	14	4.45E-04
17053-17304	CUS	W	50	7	4.11E-03
8660-8911	ORF8	C	30	7	1.49E-04
9731-9982	rolA	C	226	17	2.27E-03
	TAS3	C	58	9	3.79E-04

Min. Abundance (2)ǀ

Region*	Gene	StrandΣ	Distinct Seqs.ǂ	Distinct Phased Seqs.^	P-value£
10737-10988	rolB	W	25	5	3.50E-03
11074-11325	rolB	W	15	4	3.15E-03
16404-16655	CUS	W	80	11	1.82E-04
17293-17544	CUS	W	14	4	2.38E-03
8723-8974	ORF8	C	3	3	8.20E-05
9582-9833	rolA	C	91	12	1.12E-04
	TAS3	C	30	9	1.05E-06

Fig. 7

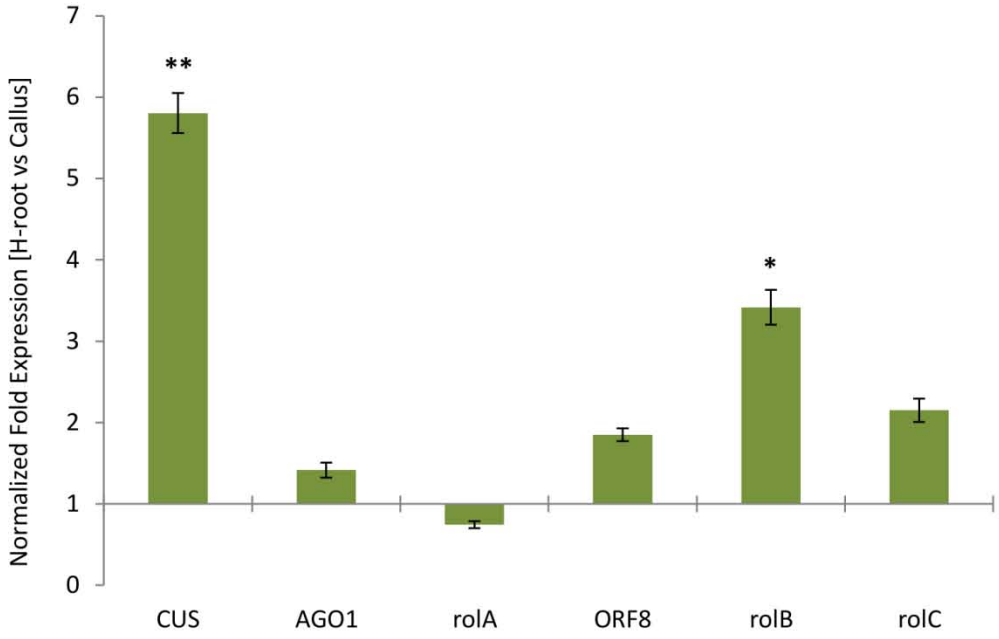
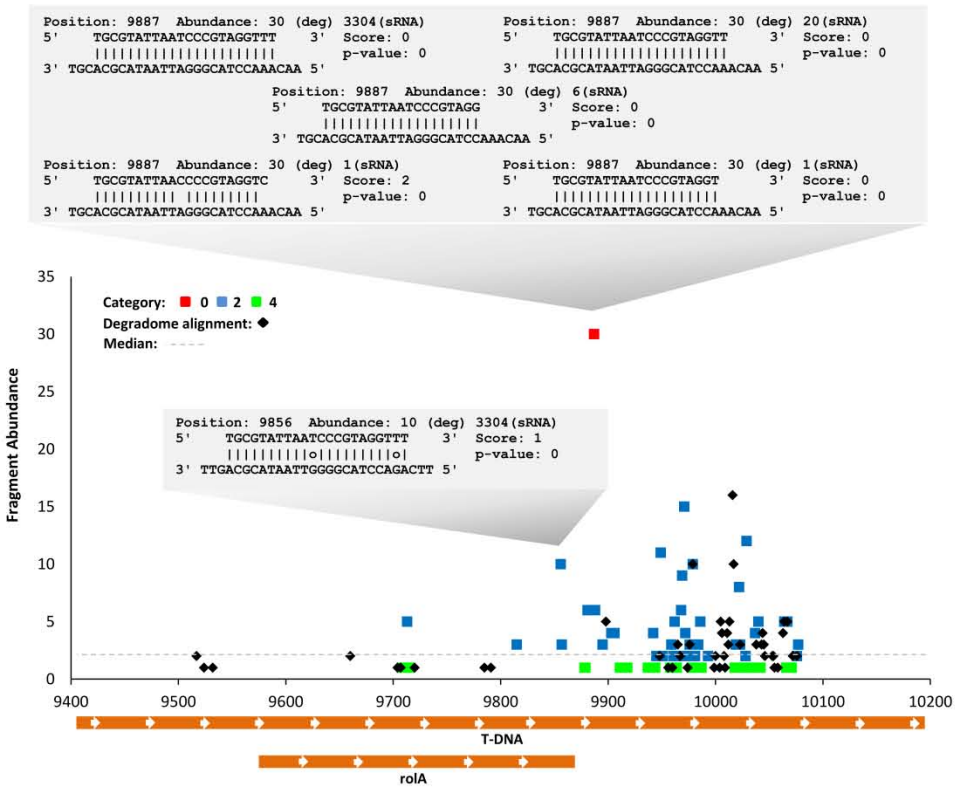

Normalized Fold Expression [H-root vs Callus]
7
6
5
4
3
2
1
0
**
*
CUS
AGO1
rolA
ORF8
rolB
rolC

Fig. 8

(a)


Position: 9887 Abundance: 30 (deg) 3304(sRNA)
5' TGCGTATTAATCCCGTAGGTTT 3' Score: 0
p-value: 0
3' TGCACGCATAATTAGGGCATCCAAACAA 5'
Position: 9887 Abundance: 30 (deg) 20(sRNA)
5' TGCGTATTAATCCCGTAGGTT 3' Score: 0
p-value: 0
3' TGCACGCATAATTAGGGCATCCAAACAA 5'
Position: 9887 Abundance: 30 (deg) 6(sRNA)
5' TGCGTATTAATCCCGTAGG 3' Score: 0
p-value: 0
3' TGCACGCATAATTAGGGCATCCAAACAA 5'
Position: 9887 Abundance: 30 (deg) 1(sRNA)
5' TGCGTATTAACCCCGTAGGTC 3' Score: 2
p-value: 0
3' TGCACGCATAATTAGGGCATCCAAACAA 5'
Position: 9887 Abundance: 30 (deg) 1(sRNA)
5' TGCGTATTAATCCCGTAGGT 3' Score: 0
p-value: 0
3' TGCACGCATAATTAGGGCATCCAAACAA 5'
Category: 0 2 4
Degradome alignment:
Median:
Fragment Abundance
Position: 9856 Abundance: 10 (deg) 3304(sRNA)
5' TGCGTATTAATCCCGTAGGTTT 3' Score: 1
p-value: 0
3' TTGACGCATAATTGGGGCATCCAGACTT 5'
T-DNA
rolA

(b)

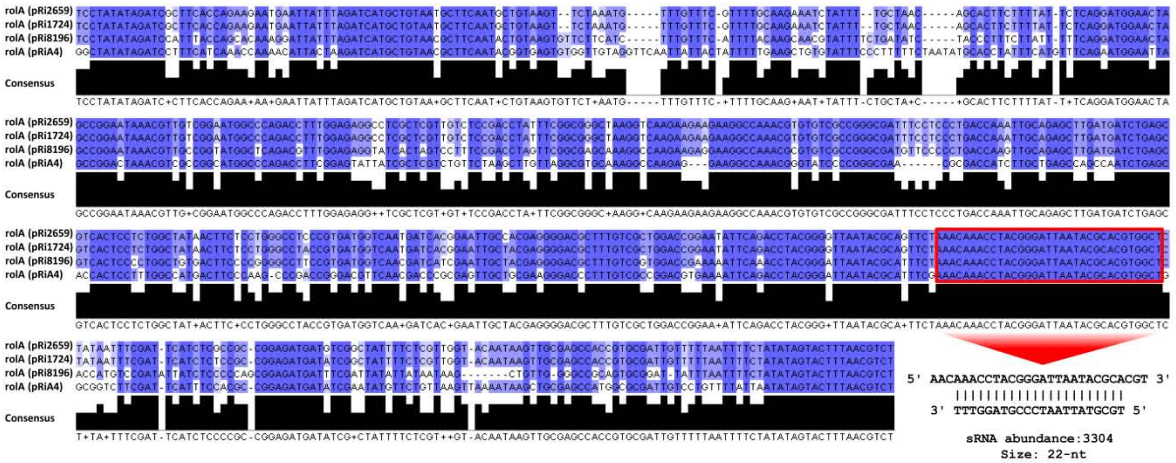
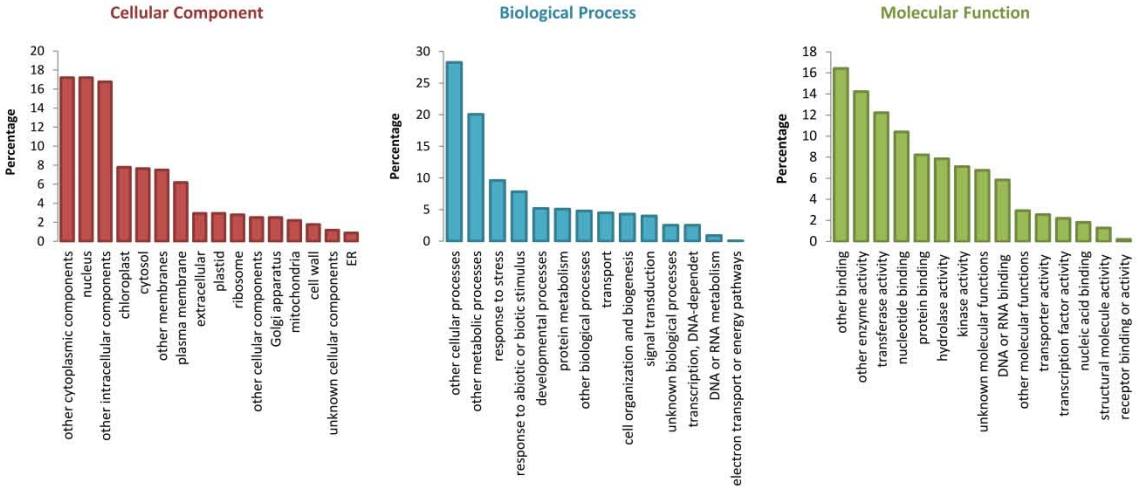
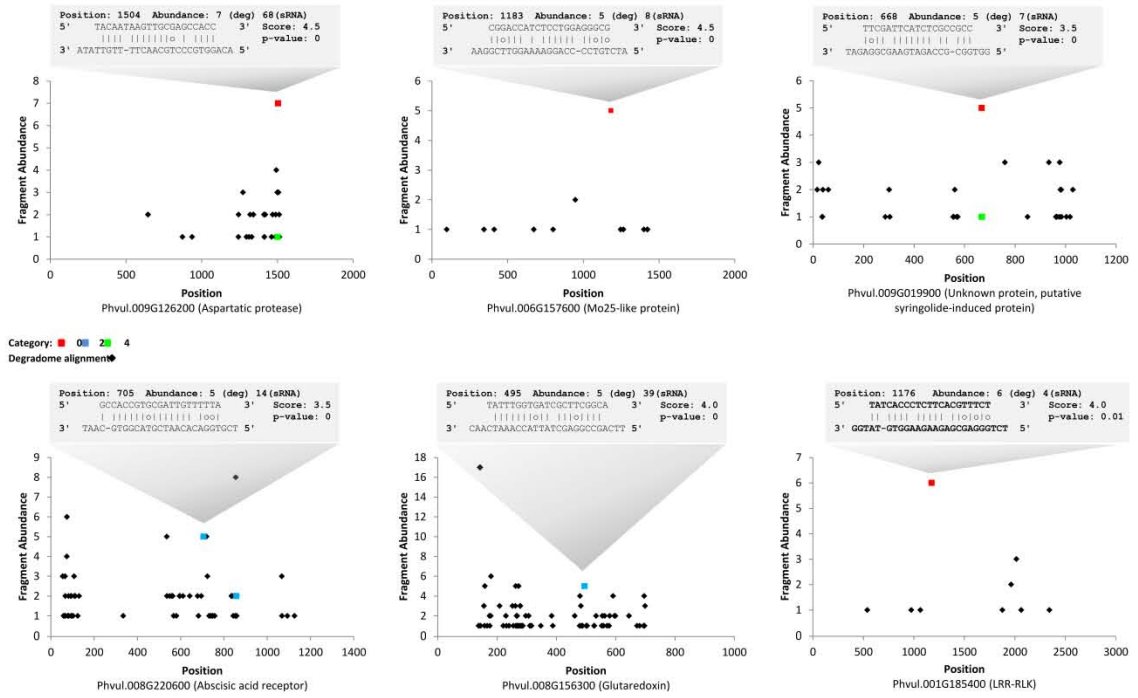

rolA (pRi2659)
rolA (pRi1724)
rolA (pRi8196)
rolA (pRiA4)
Consensus
5' AACAAACCTACGGGATTAATACGCACGT 3'
3' TTTGGATGCCCTAATTATGCGT 5'
sRNA abundance:3304
Size: 22-nt

Fig. 9

(a)


Cellular Component
Biological Process
Molecular Function
Percentage
other cytoplasmic components
nucleus
other intracellular components
chloroplast
cytosol
other membranes
plasma membrane
extracellular
plastid
ribosome
other cellular components
Golgi apparatus
mitochondria
cell wall
unknown cellular components
ER
other cellular processes
other metabolic processes
response to stress
response to abiotic or biotic stimulus
developmental processes
protein metabolism
other biological processes
transport
cell organization and biogenesis
signal transduction
unknown biological processes
transcription, DNA-dependet
DNA or RNA metabolism
electron transport or energy pathways
other binding
other enzyme activity
transferase activity
nucleotide binding
protein binding
hydrolase activity
kinase activity
unknown molecular functions
DNA or RNA binding
other molecular functions
transporter activity
transcription factor activity
nucleic acid binding
structural molecule activity
receptor binding or activity

(b)


Fragment Abundance
Position
Phvul.009G126200 (Aspartatic protease)
Phvul.006G157600 (Mo25-like protein)
Phvul.009G019900 (Unknown protein, putative syringolide-induced protein)
Category: 0 2 4
Degradome alignment
Phvul.008G220600 (Abscisic acid receptor)
Phvul.008G156300 (Glutaredoxin)
Phvul.001G185400 (LRR-RLK)

Supporting Information

Table S1 Phasing-generating loci from *Phaseolus vulgaris.*

Table S2 Oligonucleotide probes used for miRNA detection by northern blot.

Table S3 MicroRNAs used for the identification of validated targets.

Table S4 Analysis of conserved miRNA targets.

Table S5 Primers used for qRT-PCR.

Table S6 Small RNAs (18-26nt) derived from the Ri plasmid T-DNA region**.**

Table S7 PAREsnip analysis of the T-DNA derived sRNAs against the degradome fragments aligning the T-DNA.

Table S8 Predicted targets for T-DNA derived sRNAs in *P. vulgaris.*

Table S9 PAREsnip analysis of the T-DNA derived sRNAs against the transcripts of *P. vulgaris*.

Table S10 Identification of *P. vulgaris* targets from PAREsnip analysis.

Fig. S1 Number of raw small RNA reads derived from the T-DNA.

Fig. S2 RNA secondary structure of T-DNA loci

Fig. S3 5´ terminal nucleotide of T-DNA-derived sRNAs

Fig. S4 Alignment of degradome sequences to the T-DNA

Fig. S5 Number of raw degradome reads derived from the T-DNA