



UNIVERSIDAD NACIONAL DE COLOMBIA

Caracterización de miRNAs involucrados en respuesta a infección por bacterias del género *Xanthomonas* en arroz y yuca

Álvaro Luis Pérez Quintero

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Biología
Bogotá, Colombia

2012

Caracterización de miRNAs involucrados en respuesta a infección por bacterias del género *Xanthomonas* en arroz y yuca

Alvaro Luis Pérez Quintero

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título

de:

Magister en Ciencias Biología

Director:

(Ph.D.) Camilo Ernesto López Carrascal

Línea de Investigación:

Genética

Grupo de Investigación:

Manihot Biotec

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Biología

Bogotá, Colombia

2012

*"We alone on earth can rebel against the
tyranny of the selfish replicators"*

Richard Dawkins

*"Los computadores son inútiles. Solamente
pueden producir respuestas."*

Pablo Picasso

Agradecimientos

Especialmente a **Camilo López** por aventurarse en esta área de investigación, por depositar su confianza en mí y guiarme debidamente a lo largo de este trabajo.

A **Boris Szurek** por todo su apoyo, por invitarme a trabajar en el sistema arroz-Xoo, por guiarme en este proceso y por hacer de ésta parte de la investigación un proceso tan agradable.

A **Sebastien Cunnac** por su guía en el análisis de interacciones TAL-DNA, la colaboración en partes experimentales de arroz y por las provechosas discusiones generadas.

Al grupo **Manihot Biotec** por la compañía, la ayuda, las discusiones y los buenos momentos.

A **Ralf Koebnik** y al equipo **Effecteurs - Resistance des Plantes aux Bioagresseurs del Institute de Recherche pour le Development**, por la hospitalidad, la compañía y la colaboración en las partes experimentales.

A **Adriana Bernal** y el **grupo de investigación LAMFU de la Universidad de los Andes**, por los intercambios de información y la colaboración.

A **Andrés Quintero, Oscar Urrego y Pablo Vanegas** por su colaboración en los análisis bioinformáticos y experimentales en yuca.

A **Alejandra Muñoz Bodnar** por el acceso a la información de RNA-seq de yuca y la colaboración en estos análisis.

A **Luis Miguel Rodriguez** por el acceso al programa Storyteller y su colaboración en la predicción de sitios de unión de efectores TAL.

A **Andrés Zapata** por su colaboración con el análisis bioinformático y el desarrollo de algunas herramientas utilizadas en este trabajo.

A **Severine Lacombe** por su ayuda en los experimentos de northern blott en arroz.

A **Alexis Derepeer** por su ayuda en la creación de la plataforma web del programa Talvez.

A **Liliana López, Clara Bermudez y Luis Fernando Cadavid** por sus correcciones y observaciones en el planteamiento de este proyecto.

Al consejo de cede de la **Universidad Nacional de Colombia** por otorgarme la distinción Grado de Honor de Pregrado que me permitió cursar la maestría y al Sistema Nacional de Becas - Estudiantes Posgrado por su diligente administración.

Al programa **ECOS NORD** por la financiación de mi estadía en el Institute de Recherche pour le Development.

A **Colciencias** y la **DIB** por financiación parcial de este proyecto (Contrato 221-2008, quipu No Contract Number 221-2008).

A mis padres y amigos por el apoyo.

Resumen

Los microRNAs (miRNAs) son moléculas de RNAs pequeñas que juegan un papel importante en la regulación de la expresión génica en eucariotas mediante el silenciamiento de mRNAs complementarios (*targets*). En este trabajo se estudió su papel en interacciones con bacterias del género *Xanthomonas* en arroz y yuca. Mediante análisis bioinformático de librerías de RNAs pequeños se identificaron y cuantificaron miRNAs, y se estudió la expresión de los posibles *targets*. Se encontró que en yuca la respuesta mediada por miRNAs está caracterizada por la inducción de miRNAs que reprimen señalización por auxinas y por la represión de miRNAs involucrados en el control de genes de resistencia. Mientras que en arroz los miRNAs no parecen tener un papel crucial en defensa aunque la expresión de miRNAs varía en respuesta a distintas cepas bacterianas. Y se identificaron también posibles mecanismos de regulación transcripcional de miRNAs mediados por factores de transcripción vegetales y bacterianos.

Palabras clave: microRNAs, bacteria, *Xanthomonas*, arroz, yuca, efectores TAL.

Abstract

microRNAs (miRNAs) are small RNA molecules involved in the control of gene expression in eukaryotes through the silencing of complementary mRNAs (*targets*). This work addresses their role in interactions with bacteria from the genus *Xanthomonas* in rice and cassava. miRNAs were identified and quantified through the bioinformatic analyses of sRNA libraries and the expression of the possible *targets* was also studied. It was found that in cassava the response mediated by miRNAs is characterized mainly by

the induction of miRNAs that repress auxin silencing and by the repression of miRNAs involved in the control of resistance genes. While in rice miRNAs seemed not to have a crucial role in defense even though the expression of miRNAs was variable in response to different bacterial strains. This work also addresses the identification of possible mechanisms of transcriptional regulation of miRNAs mediated by vegetal and bacterial transcription factors.

Keywords: microRNAs, bacteria, *Xanthomonas*, rice, cassava, TAL effectors.

Contenido

	<u>Pág.</u>
Resumen	IX
Lista de tablas	XIV
Lista de Símbolos y abreviaturas	XV
Objetivos	XVII
Introducción	1
1. miRNAs en la interacción <i>Manihot esculenta</i>-<i>Xanthomonas axonopodis</i> pv. <i>manihotis</i>	7
1.1 Introducción: el sistema yuca-Xam	7
1.2 Resultados	10
1.2.1 Secuenciación de librerías de sRNAs en yuca	10
1.2.2 Identificación de miRNAs y pre-miRNAs	14
1.2.3 Expresión diferencial de miRNAs	15
1.2.4 Expresión diferencial de <i>targets</i> de miRNAs	18
1.3 Discusión	22
1.4 Materiales y métodos	24
1.4.1 Materiales vegetales e inoculación	24
1.4.2 Construcción y secuenciamiento de librerías de sRNAs	25
1.4.3 Análisis de datos de secuenciación	25
1.4.4 Identificación de miRNAs y premiRNAs	26
1.4.5 Expresión de <i>targets</i> de miRNAs en experimentos de RT-PCR semicuantitativo y RNA-seq	27
2. miRNAs en la interacción <i>Oryza sativa</i> -<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>	29
2.1 Introducción: el sistema arroz-Xoo.	29
2.2 Resultados	33
2.2.1 Secuenciación de librerías de sRNAs en arroz	33
2.2.2 Expresión diferencial de miRNAs	38
2.2.3 Expresión de <i>targets</i> de miRNAs en experimentos de RNA-seq y microarreglos	44
2.3 Discusión	49
2.4 Materiales y métodos	53
2.4.1 Materiales vegetales e inoculación	53
2.4.2 Construcción y secuenciamiento de librerías de sRNAs y RNA-seq.	53
2.4.3 Análisis de librerías de sRNAs	54

2.4.4	Análisis de librerías de RNA-seq	54
2.4.5	Northern blot.....	55
2.4.6	Q-PCR.....	55
3. Regulación transcripcional de miRNAs por factores de transcripción endógenos..... 57		
3.1	Introducción: Regiones promotoras de genes <i>MIRNA</i>	57
3.2	Resultados	59
3.2.1	Identificación de sitios de unión de factores de transcripción (TFBS) en promotores de genes <i>MIRNA</i> en plantas.....	59
3.2.2	Diferencias en frecuencias de TFBS entre tipos de genes y entre plantas.....	60
3.2.3	TFBS sobre-representados en familias involucradas en respuesta a bacterias en arroz y yuca.....	66
3.3	Discusión	68
3.4	Materiales y Métodos	71
3.4.1	Predicción de TFBS.....	71
3.4.2	Análisis estadístico	72
4. Regulación transcripcional de miRNAs por factores de transcripción bacterianos..... 73		
4.1	Introducción: Efectores TAL (transcription activator-like).....	73
4.2	Resultados	75
4.2.1	Desarrollo de un algoritmo de predicción de sitios de unión de efectores TAL.....	75
4.2.1	Predicción de genes <i>targets</i> de efectores TAL y expresión diferencial en respuesta a <i>Xanthomonas</i>	79
4.2.2	Interacción entre efectores TAL y genes de la vía de miRNAs.....	82
4.2.1	Interacción entre efectores TAL y genes <i>MIRNA</i>	84
4.3	Discusión	87
4.4	Materiales y métodos	89
4.4.1	Algoritmo de predicción de sitios de unión de efectores TAL: Tal vez.....	89
4.4.2	Sets de datos y predicciones	90
4.4.1	Análisis de microarreglos.....	90
5.	Conclusiones.....	92
6.	Recomendaciones y perspectivas	94
7.	Anexos	101
8.	Bibliografía	102

Lista de figuras

	<u>Pág.</u>
Figura 0-1: Modelo general de interacción planta-bacteria mediado por miRNAs.....	6
Figura 1-1: Distribución de tamaños de lecturas de librerías de sRNAs de yuca.	12
Figura 1-2: Expresión diferencial de miRNAs en respuesta a inoculación por Xam en yuca.	17
Figura 1-3: Expresión de <i>targets</i> de miRNAs a diferentes tiempos post-inoculación con Xam en experimentos de RT-PCR semicuantitativo (izquierda) y RNA-seq (derecha).	20
Figura 1-4: Boxplot mostrando los niveles de expresión de <i>targets</i> de miRNAs en datos de RNA-seq.....	21
Figura 2-1: Expresión del gen <i>Os11N3</i> evaluada por Q-PCR en muestras de RNA utilizadas para secuenciación.....	35
Figura 2-2: Distribución de tamaños de lecturas de librerías de sRNAs en arroz.....	36
Figura 2-3: Cluster jerárquico mostrando las relaciones entre tratamientos en librerías de sRNAs a partir de la correlación en perfiles de expresión de miRNAs.	41
Figura 2-4: Expresión diferencial de miRNAs en arroz en respuesta a Xoo BAI3.	42
Figura 2-5: Expresión diferencial de miRNAs en arroz en respuesta a Xoo PXO99A. ..	43
Figura 2-6: Librerías de RNA-seq de arroz inoculado con Xoo BAI3.	46
Figura 2-7: Expresión de <i>targets</i> de miRNAs diferencialmente expresados en respuesta a Xoo BAI3 en datos de RNA-seq.	47
Figura 2-8: Expresión de <i>targets</i> de miRNAs diferencialmente expresados en respuesta a Xoo PXO99 en datos de microarreglos.....	48
Figura 3-1: Frecuencia de nucleótidos en regiones de promotores de genes de plantas.	64
Figura 3-2: Cluster jerárquico basado en la correlación de distribuciones de frecuencias de TFBS en promotores de genes <i>MIRNA</i> en distintas plantas.	65
Figura 4-1: Desempeño de tres programas de predicción de sitios de unión de efectores TAL.	77
Figura 4-2: Red predicha de efectores TAL y sus genes <i>targets</i> en la interacción de arroz con distintas cepas de <i>Xanthomonas oryzae</i>	81
Figura 4-3: Interacción entre efectores TAL de cepas de <i>Xanthomonas oryzae</i> y el gen <i>OsHEN1</i>	83

Lista de tablas

	<u>Pág.</u>
Tabla 1-1: Estadísticas generales del procesamiento y mapeo de librerías de RNAs pequeños de yuca.	13
Tabla 2-1: Estadísticas generales del procesamiento y mapeo de librerías de RNAs pequeños de arroz.	37
Tabla 3-1: Diferencias en distribución de TFBS entre genes MIRNA y genes codificantes.	62
Tabla 3-2: TFBS sobrerrepresentados en genes MIRNA comparados con genes codificantes para proteínas en yuca, arroz y <i>Arabidopsis</i>	63
Tabla 3-3: TFBS sobre-representados en promotores de familias de miRNAs diferencialmente espresados en respuesta a <i>Xanthomonas</i> en yuca y arroz.....	67
Tabla 4-1. Predicción de <i>targets</i> naturales de efectores TALs en búsquedas de promotores utilizando tres programas de predicción.....	78

Lista de Símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término	Definición
Log ₂ FC	Cambio en expresión	log 2 (expresión 1/expresión 2)
cfu/ml	Concentración	Unidades formadoras de colonia por mililitro
OD600nm	Concentracion	Densidad óptica a 600 nanómetros

Símbolos con letras griegas

Símbolo	Definición
Δ	Delección de un gen

Abreviaturas

Abreviatura	Término	Abreviatura	Término
AGO	ARGONAUTE	Fx	Efector
aly	<i>Arabidopsis lyrata</i>	Genes R	Genes codificantes para proteínas R
ARF	<i>Auxin Response Factors</i>	gma	<i>Glycine max</i>
ath	<i>Arabidopsis thaliana</i>	GO	<i>Gene ontology</i>
ATP	<i>Adenosin tri-phosphate</i>	HEN1	HUA ENHANCER 1
bdi	<i>Brachypodium distachyon</i>	hpi	horas post-inoculación
CBB	<i>Cassava bacterial blight</i>	I	Inoculada
cis-nat-RNAs	<i>cis natural antisense RNAs</i>	ID	Identificación
cre	<i>Chlamydomonas reinhardtii</i>	LRR	<i>Leucin Rich Repeat</i>
csi	<i>Citrus sinensis</i>	mdo	<i>Malus domestica</i>
Ct	<i>Threshold cycle</i>	mes	<i>Manihot esculenta</i>
cv	cultivar	MFE	<i>Minimum folding energy</i>
DCL1	<i>Dicer-like 1</i>	MIRNA	Gen codificante para un miRNA
dpi	días post-inoculación	miRNA*	miRNA pasajero
EBEs	<i>Effector Binding Elements</i>	miRNAs	microRNAs
EMSA	<i>Electrophoretic mobility shift assay</i>	mtr	<i>Medicago truncatula</i>
ESTs	<i>Expressed Sequences Tags</i>	NBS	<i>Nucleotide binding domain</i>
ETI	<i>Effector-Triggered Immunity</i>	NI	No inoculada
ETS	<i>Effector-Triggered Susceptibility</i>	nr	No redundante

Abreviatura	Término	Abreviatura	Término
nt	Nucleótido	Xam	<i>Xanthomonas axonopodis</i> pv. manihotis
osa	<i>Oryza sativa</i>	Xoc	<i>Xanthomonas oryzae</i> pv. oryzicola
PAMPs	<i>Pathogen-Associated Molecular Patterns</i>	Xoo	<i>Xanthomonas oryzae</i> pv. oryzae
PCR	<i>Polymerase Chain Reaction</i>	zma	<i>Zea mays</i>
PMW	<i>Positional Weight Matrix</i>		
ppt	<i>Physcomitrella patens</i>		
pre-miRNA	Precursor de miRNA		
pri-miRNA	miRNA primario		
Proteínas R	Proteínas de resistencia		
PRRs	<i>Pathogen Recognition Receptors</i>		
ptc	<i>Populus trichocarpa</i>		
PTI	<i>PAMP Triggered Immunity</i>		
Pto	<i>Pseudomonas syringae</i> pv. tomato		
pv	Patovar		
Q-PCR	PCR cuantitativo		
rco	<i>Ricinus communis</i>		
RDR	<i>RNA-dependent RNA polymerase</i>		
RISC	<i>RNA Induced Silencing Complex</i>		
RLK	<i>Receptor-like kinase</i>		
rma	Robust Multichip Average		
RNA	Ácido ribonucleico		
ROC	<i>Receiver operating characteristics</i>		
RPKM	Reads per kilobase per million		
rRNAs	<i>Ribosomal RNA</i>		
RT-PCR	<i>Reverse transcription PCR</i>		
RVDs	<i>Repeat Variable Diresidues</i>		
sbi	<i>Sorghum bicolor</i>		
SGS	<i>Suppressor of gene silencing</i>		
siRNAs	<i>small interference RNAs</i>		
smo	<i>Selaginella moellendorffii</i>		
snoRNAs	<i>Small nucleolar RNA</i>		
sRNAs	RNAs pequeños (small RNAs)		
subsp	Subespecie		
T3SS	Sistema de Secreción tipo 3		
TAL	transcription-activator like		
ta-siRNAs	<i>trans-acting small interference RNAs</i>		
TF	Factor de transcripción		
TFBS	Sitios de unión de factores de transcripción		
tRNAs	<i>Transfer RNA</i>		
UV	Ultra violeta		
vvi	<i>Vitis vinifera</i>		

Objetivos

Objetivo general

Identificar los miRNAs y sus *targets* involucrados en respuesta a infección por bacterias del género *Xanthomonas* en arroz y yuca.

Objetivos específicos

- Identificar miRNAs diferencialmente expresados durante la infección por *Xanthomonas* en yuca y arroz mediante perfiles de expresión en librerías de RNAs pequeños.
- Validar experimentalmente la regulación de *targets* de miRNAs por sus respectivos miRNAs.
- Identificar elementos reguladores en las regiones promotoras de miRNAs de yuca y arroz.
- Relacionar los patrones expresión de miRNAs en respuesta a bacterias en arroz y yuca con el análisis de elementos reguladores.

Introducción

El arroz (*Oryza sativa*) y la yuca (*Manihot esculenta*) constituyen dos fuentes importantes de carbohidratos para más de 3 mil millones de seres humanos en el mundo, ocupando respectivamente el tercer y octavo lugar en producción mundial entre diferentes cultivos (<http://faostat3.fao.org/>, datos 2010). Ambos cultivos son afectados por bacterias del género *Xanthomonas*, particularmente las especies *Xanthomonas oryzae* pv. *oryzae* (Xoo) y *Xanthomonas axonopodis* pv. *manihotis* (Xam), las cuales pueden causar pérdidas de hasta el 50% de la producción en arroz [1, 2] y del 75% en yuca [3, 4], respectivamente. Xam y Xoo hacen parte de las 10 bacterias más importantes en fitopatología molecular [5]. El aumento en la incidencia de estos patógenos amenaza actualmente la seguridad alimentaria de miles de personas en el mundo lo cual se ve además exacerbado por los fenómenos de sobrepoblación y cambio climático [6-8]. Por esto, el estudio de los mecanismos moleculares de resistencia a bacterias fitopatógenas y el desarrollo de estrategias biotecnológicas para el mejoramiento de cultivos, se ha convertido en una prioridad para el sector agrícola.

Tradicionalmente, el estudio de la fitopatología molecular se ha basado en el análisis de tres tipos de proteínas que son las principales mediadoras durante las interacciones planta-patógeno: los PRRs (*Pathogen Recognition Receptors*), las proteínas de resistencia de plantas (R) y las proteínas efectoras de bacterias. Los PRRs son proteínas que reconocen estructuras y moléculas altamente conservadas en microorganismos conocidas como PAMPs (*Pathogen-Associated Molecular Patterns*). Las PRRs median un mecanismo basal de inmunidad conocido como PTI (*PAMP Triggered Immunity*). Para contrarrestar la PTI, patógenos adaptados para infectar plantas huésped particulares desarrollaron las proteínas efectoras, las cuales son inyectadas, en el caso de bacterias,

en la célula vegetal por el sistema de secreción tipo 3 (T3SS) para interferir o suprimir este tipo de inmunidad basal, generando susceptibilidad o ETS (*Effector-Triggered Susceptibility*). En respuesta, las plantas desarrollaron las proteínas R, las cuales reconocen proteínas efectoras particulares y activan una inmunidad conocida como ETI (*Effector Triggered Immunity*)[9].

Recientemente, se han identificado otros mecanismos importantes para regular las interacciones planta-patógeno, principalmente los mecanismos de silenciamiento génico mediados por RNAs pequeños (sRNAs). Estos mecanismos se encuentran en la mayoría de los eucariotas y se basan en la interacción de sRNAs (de 18 a 30 nucleótidos) con ácidos nucleicos complementarios induciendo su modificación. El silenciamiento génico mediado por sRNAs está involucrado en múltiples procesos, incluyendo el control de la expresión génica, modificaciones epigenéticas y defensa contra ácidos nucleicos invasores y transposones [10-12].

Entre los sRNAs más estudiados se encuentran los microRNAs (miRNAs) los cuales son moléculas de aproximadamente 21 nt que se originan a partir de genes nucleares. Un gen nuclear *MIRNA* (los genes de miRNAs se notan en cursiva y mayúsculas) es primero transcrito por la RNA polimerasa II en un miRNA primario (pri-miRNA). Éste es luego procesado a un RNA intermediario llamado precursor de miRNA (pre-miRNA) por la proteína *Dicer-like 1* (DCL1) en plantas [13-15]. Los pre-miRNAs tienen una estructura secundaria característica y estable, con una alta y negativa energía libre de plegamiento, siendo este un criterio fundamental para su identificación y anotación [16]. Luego, los pre-miRNAs son procesados a un dúplex miRNA:miRNA* por DCL1. Este dúplex miRNA/miRNA* contiene extremos 3' con 2 nucleótidos sobresalientes, los cuales son metilados en el núcleo de la célula por la proteína HEN1 (*HUA ENHANCER 1*), lo que le proporciona estabilidad al dúplex miRNA/miRNA* y lo protege de la degradación. Finalmente, el miRNA es exportado al citoplasma gracias a una proteína exportina, probablemente la proteína HASTY. Sólo la hebra activa del miRNA o miRNA maduro es incorporada al complejo RISC (*RNA Induced Silencing Complex*), mientras que la hebra pasajera (miRNA*) es degradada. El complejo RISC, al cual se asocia la proteína ARGONAUTE1 (AGO1), es luego el encargado de guiar la escisión de RNAs mensajeros complementarios al miRNA maduro conocidos como *targets* o blancos [13-15]. Los

miRNAs normalmente se encuentran representados por varios loci en el genoma y son funcionalmente redundantes, por esto, los miRNAs se agrupan en familias que incluyen moléculas con secuencias de más de un 90% de identidad [16].

Los miRNAs están también implicados en la biogénesis de otro tipo de RNAs pequeños, los *trans-acting small interference RNAs* (ta-siRNAs), estos RNAs se originan a partir de un gen *TAS*, cuyo transcrito es cortado gracias a la acción de un miRNA, el producto es luego convertido en RNA de doble cadena (dsRNA) gracias al complejo RDR6/SGS3, para posteriormente ser procesado por *Dicer-like 4* (DCL4) produciendo *small interference RNAs* (siRNAs) que se unirán a AGO1 o AGO7 para ejercer actividad de silenciamiento sobre otros RNAs mensajeros [17, 18].

Los miRNAs pueden cumplir una función en defensa contra bacterias mediante la reprogramación de la expresión génica. Tal mecanismo fue primero revelado en la interacción entre *Arabidopsis thaliana* y *Pseudomonas syringae* pv. tomato (Pto), donde se encontró que el miRNA miR393 es inducido en respuesta al péptido flg22 derivado de la flagelina, el PAMP más estudiado [19]. La inducción de miR393 interrumpe la señalización mediada por auxinas y en consecuencia disminuye el crecimiento bacteriano [19], probablemente debido a que la auxina induce la expresión de proteínas expansinas que debilitan las paredes celulares durante procesos de crecimiento [20]. Otros trabajos en el mismo patosistema, han encontrado que en respuesta a bacterias, es común la inducción de miRNAs involucrados en señalización de hormonas y que este tipo de regulación al parecer se encuentra asociada con respuestas de PTI [21, 22]. Por otro lado, se ha encontrado que la represión de miRNAs cuyos *targets* son genes R (asociados a ETI) juega un papel crucial en la defensa contra bacterias [23]. Además, la regulación de genes R por miRNAs y ta-siRNAs también es importante en interacciones de plantas con varios tipos de microorganismos incluyendo virus [24] hongos patógenos [25, 26], bacterias fijadoras de nitrógeno [27-29] y hongos formadores de micorrizas arbusculares [30].

Sin embargo, las bacterias pueden también defenderse de la maquinaria silenciamiento. Se ha encontrado que las bacterias pueden utilizar proteínas efectoras para suprimir la vía de silenciamiento por miRNAs y consecuentemente causar enfermedad [31, 32] de

forma análoga a los supresores de silenciamiento utilizados por virus [33]. De igual forma una posibilidad adicional, aún no explorada, es que las bacterias del género *Xanthomonas* puedan regular la transcripción de componentes de las vías de miRNAs para causar enfermedad utilizando un tipo de proteínas efectoras características del género (también presentes en *Ralstonia* spp.) conocidas como efectores TAL (transcription-activator like), que pueden actuar como factores de transcripción induciendo la expresión de genes en la planta [34] (Sección 4.1).

A partir de estas observaciones se puede proponer un modelo general de interacción en el que, en respuesta al reconocimiento de bacterias patógenas por la vía de PTI o ETI, se inducen miRNAs cuyos *targets* son reguladores negativos de respuesta de defensa (p. ej. algunas vías de hormonas) y se reprimen miRNAs cuyos *targets* son reguladores positivos de respuesta de defensa (p. ej. genes de resistencia) [35]. En respuesta las bacterias contra-atacan suprimiendo la vía de silenciamiento gracias a la translocación de proteínas efectoras, y posiblemente en el caso de *Xanthomonas* y *Ralstonia* alterando la transcripción de miRNAs o de componentes de la vía utilizando efectores TAL. Una representación de este modelo se puede ver en la **Figura 0-1**.

Este modelo sin embargo se basa principalmente en estudios en *Arabidopsis* [21, 22] y se desconoce si el mismo patrón de regulación se presenta en plantas filogenéticamente lejanas como arroz y yuca. Se desconoce además si existe variación en las respuestas con distintos tipos de bacterias, incluyendo bacterias del género *Xanthomonas*, uno de los grupos de bacterias fitopatógenas de mayor impacto a nivel mundial que infecta más de 400 especies de plantas [36].

Por otro lado, otro aspecto aún desconocido en la interacción planta-bacteria es el control de la transcripción de genes *MIRNA*. Varios trabajos han identificado algunos elementos claves implicados en la regulación de miRNAs y se sabe que las regiones promotoras de miRNAs contienen elementos comunes como cajas TATA, así como sitios de unión para factores de transcripción específicos, y algunos de estos elementos se han encontrado sobre-representados en familias involucradas en procesos específicos [37-39]. Sin embargo se desconoce por ejemplo si existen elementos sobre-representados en familias de miRNAs con papel en defensa contra bacterias y aun se desconoce en gran parte

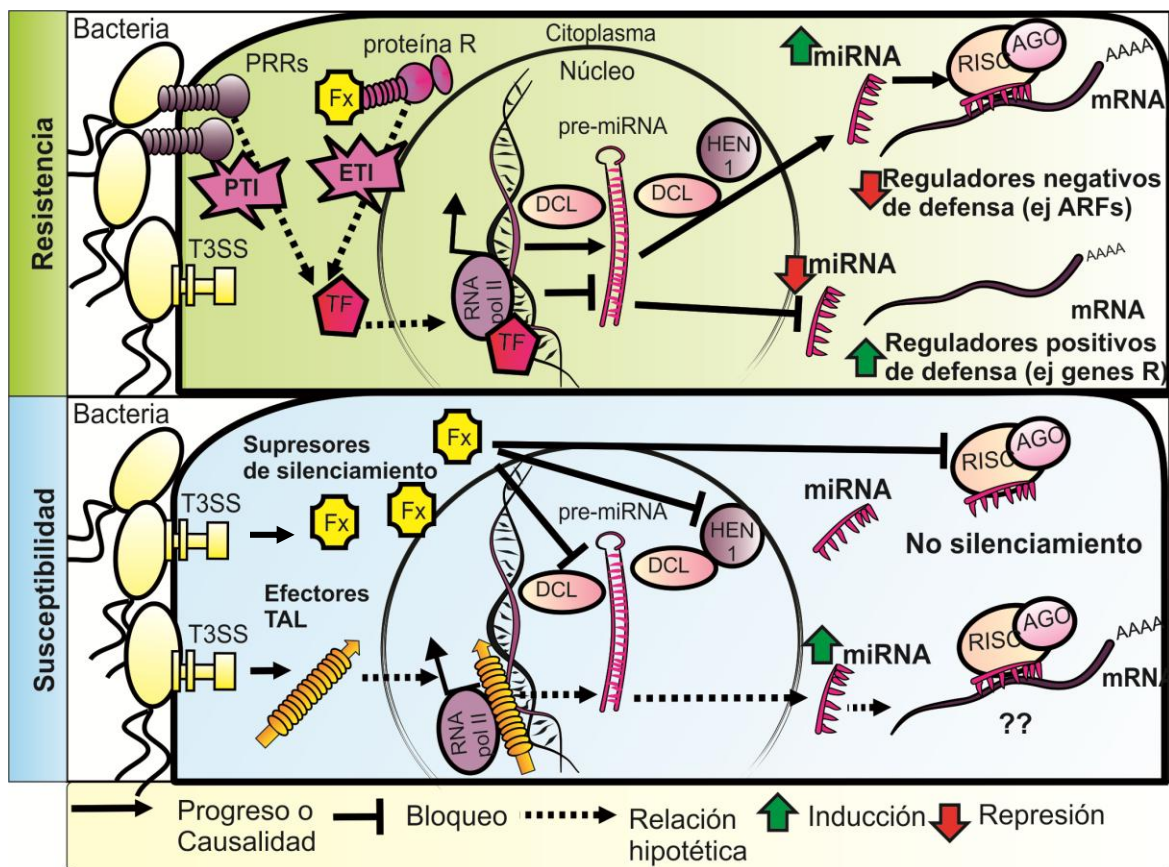
como cambia la regulación de genes *MIRNA* entre especies. De igual manera y como ya se ha mencionado, la regulación transcripcional de miRNAs mediada por factores de transcripción bacterianos (efectores TALs) es otro aspecto hasta ahora inexplorado .

En este trabajo se buscó analizar el papel de los miRNAs en interacciones planta-bacteria en los sistemas *M. esculenta*-Xam y *O. sativa*-Xoo, con base principalmente en el análisis bioinformático de librerías de RNAs pequeños secuenciadas por tecnología Solexa/illumina, una técnica que ha mostrado ser idónea para este tipo de estudios, dando resultados consistentes entre distintas réplicas técnicas y biológicas [21, 40]. El trabajo se encuentra organizado en cuatro capítulos donde se muestran los resultados de los temas principales así: Capítulo 1: miRNAs en la interacción *Manihot esculenta*-*Xanthomonas axonopodis* pv. manihotis. Capítulo 2: miRNAs en la interacción *Oryza sativa* -*Xanthomonas oryzae* pv. oryzae. Capítulo 3: Regulación transcripcional de miRNAs por factores de transcripción endógenos. Y Capítulo 4: Regulación transcripcional de miRNAs por factores de transcripción bacterianos.. Cada capítulo se presenta como un artículo independiente que consta de una breve introducción del tema específico, los resultados y discusión y luego, los materiales y métodos. Finalmente se presentan las conclusiones del trabajo en general y algunas recomendaciones y perspectivas.

Se espera que los resultados encontrados en este trabajo contribuyan no solamente al conocimiento general de los mecanismos moleculares de la interacción planta-patógeno sino que permitan además encaminar estrategias de mejoramiento basadas en la reprogramación génica mediada por miRNAs, el cual es un campo promisorio actual en la ingeniería de plantas [35].

Figura 0-1: Modelo general de interacción planta-bacteria mediado por miRNAs.

Se muestran dos posibles escenarios: en una interacción de resistencia se detecta la presencia del patógeno y se activan las vías de inmunidad (PTI o ETI) lo cual modificará la transcripción de algunos miRNAs. En una interacción susceptible es posible que la bacteria inyecte supresores de silenciamiento que interrumpirán distintos pasos de la vía de biosíntesis de miRNAs causando que no haya silenciamiento. Es posible también que la bacteria utilice efectores TAL para inducir la expresión de miRNAs con fines aun desconocidos. TF = factor de transcripción, Fx = efector.



1.miRNAs en la interacción *Manihot esculenta-Xanthomonas axonopodis* pv. manihotis.

1.1 Introducción: el sistema yuca-Xam

La yuca (*M. esculenta*, Euphorbiaceae) es un cultivo de vital importancia para la seguridad alimentaria siendo el componente básico de la dieta de más de 500 millones de personas en el mundo, es actualmente el quinto cultivo en producción mundial, teniendo particular importancia en las regiones sub-tropicales de África, Asia y América-Latina [41, 42]. En Colombia, la yuca ocupa el cuarto lugar en producción entre los diferentes cultivos alcanzando un total de 2.3 millones de toneladas anuales (<http://faostat3.fao.org>, datos 2010). Su valor nutricional proviene principalmente del alto contenido de almidón almacenado en sus raíces que puede alcanzar hasta el 80% de su peso, por lo que es además un cultivo promisorio para la producción de bioetanol y otros productos derivados del almidón [43, 44]. La yuca tiene además alta tolerancia a distintos tipos de estrés incluyendo alta humedad, sequía, altas temperaturas y suelos pobres, lo cual la hace ideal para cultivo en condiciones ambientales desfavorables para otros [41].

Sin embargo, los cultivos de yuca pueden verse fuertemente afectados por la bacteriosis vascular o añublo bacteriano (*Cassava bacterial blight*, CBB). Esta enfermedad es causada por la bacteria Gram negativa *Xanthomonas axonopodis* pv. manihotis (Xam) la cual tiene distribución global, puede ocasionar pérdidas de más del 75% dependiendo de condiciones ambientales y en casos severos ha contribuido a situaciones de hambruna en países africanos [3, 4, 45]. Los síntomas de CBB son variados, en plantas adultas con la aparición de manchas angulares en el lado abaxial de la hojas y exudados en el tallo

de color naranja, a medida que la enfermedad avanza se puede dar marchitamiento y eventualmente la muerte total de la planta [4, 8]. Comúnmente la bacteria infecta la planta por heridas, hidátodos o estomas, una vez dentro se multiplica en los espacios intercelulares del mesófilo u en el tejido vascular [46].

Varios estudios han revelado gran variabilidad genética en distintas poblaciones de Xam en distintos rangos geográficos [8, 47-49], la cual se ve también reflejada en los distintos grados de virulencia de diferentes cepas de Xam sobre diferentes variedades de yuca [8]. Recientemente se reveló mediante la secuenciación del genoma de 65 cepas de Xam que existe gran variabilidad en repertorios de proteínas efectoras, existiendo únicamente 9 efectores (de 25 analizados, sin incluir efectores tipo TAL) compartido por todas las cepas analizadas [50]. Las distintas cepas analizadas tienen también repertorios variables de efectores tipo TAL, aunque sólo se dispone de la secuencia de 7 de estos efectores: 5 para la cepa Xam668 y 2 para la cepa XamCIO151 [50]. Uno de estos efectores tipo TAL, llamado TALE1_{xam} se sabe tiene un papel crucial en la patogenicidad. TALE1_{xam} actúa como factor de transcripción, induciendo la expresión de uno o varios genes determinantes para la enfermedad [51], dentro de los cuales un gen que codifica para un factor B3 de choque térmico sobresale como un fuerte candidato [Muñoz-Bodnar *et al* in prep].

Al igual que existe variable virulencia entre cepas de Xam, distintas variedades de yuca muestran distintos grados de resistencia a la bacteria [52]. La mayoría de las variedades cultivadas son sin embargo susceptibles a las cepas mas prevalentes de Xam, y las variedades resistentes naturales no están bien adaptadas a regiones agroecológicas particulares o no exhiben características agronómicas deseables [8, 45]. En los últimos años se ha intentado descifrar las bases moleculares de la resistencia a Xam, la cual se ha considerado poligénica y aditiva (cuantitativa) [53]. Se ha logrado identificar dos genes de yuca posiblemente implicados en resistencia contra Xam [54-56] y se ha analizado la regulación transcripcional de algunos genes en respuesta a Xam, encontrando principalmente inducción de genes involucrados estrés oxidativo y degradación de proteínas, y la represión de genes relacionados con fotosíntesis y metabolismo [57]. Actualmente existen varias iniciativas para aumentar la resistencia en yuca mediante la

inserción de genes de resistencia (a partir de variedades resistentes o de otras plantas) en variedades susceptibles [8].

Una posibilidad relativamente poco explorada es la modulación de la regulación transcripcional dada por miRNAs por distintos métodos como sobre-expresión inducible de miRNAs, diseño de miRNAs artificiales o *target mimics* [58-61]. Algunas ventajas de estas estrategias incluirían, por ejemplo, la regulación simultánea de varios genes o vías de interés [35]. Sin embargo, para esto es necesario conocer cómo se da naturalmente la regulación por miRNAs en este sistema.

Hasta la fecha son pocos los estudios de miRNAs en yuca. Los primeros estudios estuvieron basados en secuencias de ESTs en donde se lograron identificar hasta 20 secuencias correspondientes a genes *MIRNAs* pertenecientes a familias conservadas en otras plantas [61-65] y sólo 10 se encuentran depositados en la base de datos de referencia de miRNAs: miRBase [66]. El reciente secuenciamiento del genoma de yuca [67] (actualmente en su segunda versión borrador) ha permitido identificar un amplio repertorio de genes *MIRNA* de familias conservadas, llegando a una estimación de 169 loci diferentes [65, 68]. Adicionalmente, para algunos de estos miRNAs se ha inferido un papel en estrés abiótico, principalmente en resistencia a sequía [63, 68]

En este capítulo se presentan los resultados del análisis de librerías de RNAs pequeños de yuca para encontrar miRNAs involucrados en defensa contra Xam. Se caracteriza el repertorio de miRNAs de yuca (conservados y especie-específicos). Se identifican miRNAs diferencialmente expresados en respuesta a la bacteria, se predicen los *targets* para estos miRNAs y en algunos casos se hace la validación experimental de algunos *targets* mediante RT-PCR. Los resultados acá presentados han sido publicados en Pérez-Quintero *et al.* [69].

1.2 Resultados

1.2.1 Secuenciación de librerías de sRNAs en yuca

Se construyeron dos librerías de RNAs pequeños utilizando RNA extraído de hojas y tallos de yuca de la variedad resistente a Xam MBRA685. Una librería fue no inoculada (**NI**) y la otra inoculada (**I**) con la cepa virulenta Xam CIO15. La librería inoculada se construyó a partir de una mezcla de RNA extraído a varios tiempos post-inoculación (0 hpi –horas post-inoculación-, 6hpi, 24hpi, 2 dpi –días post-inoculación-, 5dpi, 7dpi y 15dpi). Se hizo de esta forma principalmente porque se desconoce a qué tiempos se activan las defensas en este sistema y también esperando detectar solamente las respuestas que fuesen robustas y consistentes en el tiempo. A partir de este tratamiento se pueden encontrar miRNAs involucrados en defensa tipo PTI y ETI, sin poder diferenciar entre las dos respuestas.

Tras remover adaptadores y secuencias de baja calidad se obtuvieron 15 y 11 millones de lecturas para las librerías NI e I respectivamente (~1.5 y 1.2X cobertura) (**Tabla 1-1**). Para ambas librerías se encontró la distribución de tamaños esperada en lecturas no-redundantes (lecturas únicas) (**Figura 1-1a**): picos en 21 y 24 nt, 24nt>21nt [70, 71]. Sin embargo la distribución de lecturas totales (**Figura 1-1b**) mostró diferencias entre las dos librerías y se desvía de la distribución esperada (picos de igual tamaño en 21 y 24nt) [70, 71], esto sugiere la presencia de abundantes lecturas correspondientes a productos de degradación, otros tipos de RNAs reguladores, e incluso fragmentos de adaptadores persistentes aun luego del proceso de *trimming*. Efectivamente un alto porcentaje de secuencias correspondían a varias familias de RNAs pequeños nucleares (snoRNAs), tRNAs y rRNAs (36% NI y 28% I). Sin embargo, la presencia de estas secuencias no afectó la distribución esperada para miRNAs (**Figura 1- 1c y d**, sección 0) y, aunque los porcentajes encontrados de miRNAs fueron relativamente bajos (**Tabla 1-1**), esto se ve compensado por la profundidad de las librerías.

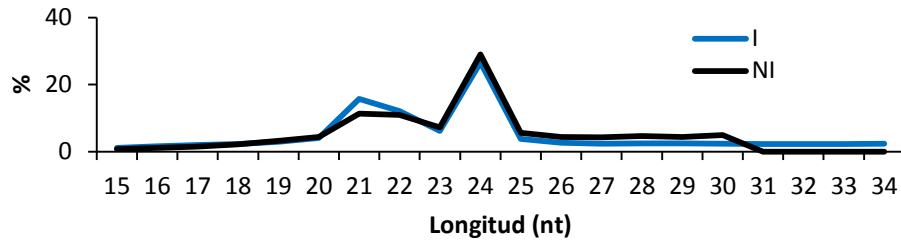
Las lecturas fueron también mapeadas contra el genoma de yuca pero se conservaron las lecturas no mapeadas teniendo en cuenta que el genoma está aun incompleto y que

muchos miRNAs pueden aparecer expresados en las librerías pero sin tener un locus identificado en el genoma. Se filtraron las lecturas correspondientes a otras familias de RNAs y se obtuvieron respectivamente 5,655,383 y 5,187,027 de lecturas de las librerías NI e I, las cuales se emplearon para todos los análisis subsecuentes. Ambas librerías fueron también mapeadas contra el genoma de Xam para verificar la presencia de ácidos nucleicos provenientes de la bacteria únicamente en la librería inoculada. De acuerdo a lo esperado, menos secuencias de la librería NI (0.8%) comparado con la librería I (5.2%) mostraron similitud con secuencias de Xam. La mayoría de las lecturas de NI que mapeaban contra el genoma de Xam también mapeaban contra el genoma de yuca (91%) y corresponderían por tanto a regiones compartidas entre ambos genomas. Mientras que en las lecturas de la librería inoculada solo el 33% correspondía a estas regiones compartidas, las lecturas restantes pueden representar RNA de distintos tipos proveniente de Xam incluyendo RNA no-codificante.

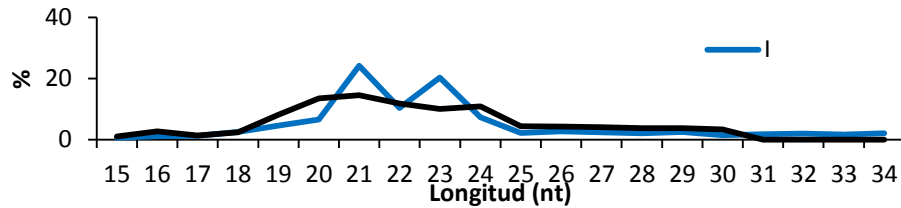
Figura 1-1: Distribución de tamaños de lecturas de librerías de sRNAs de yuca.

Se muestran los porcentajes para cada tamaño entre **a)** lecturas no-redundantes (únicas), **b)** lecturas redundantes o repetitivas, y lecturas totales mapeadas contra **c)** miRNAs conservados y **d)** miRNAs nuevos específicos de yuca.

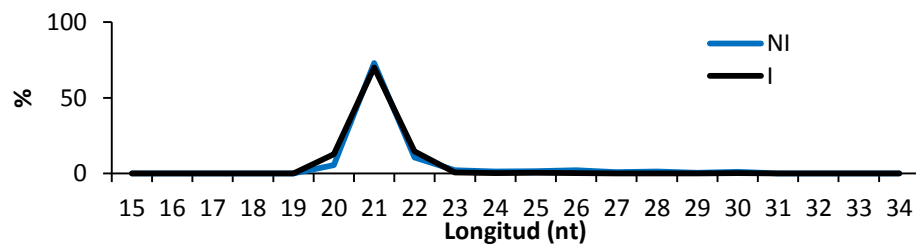
a) Lecturas no-redundantes



b) Lecturas totales



c) miRNAs conservados



d) miRNAs nuevos

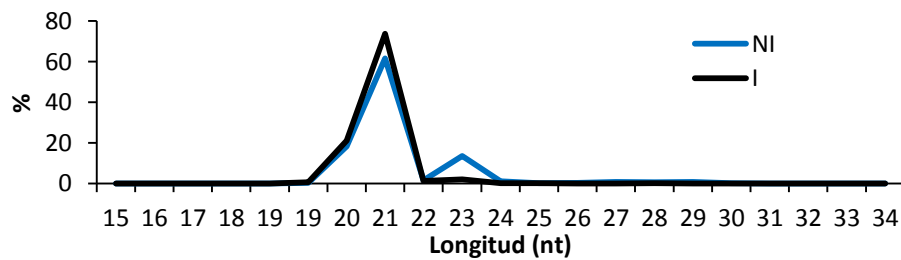


Tabla 1-1: Estadísticas generales del procesamiento y mapeo de librerías de RNAs pequeños de yuca.

Características	No inoculada (lecturas)	Inoculada (lecturas)
Total	31,434,907	15,968,516
Alta calidad, sin adaptador	15,084,481	11,871,330
Mapea contra el genoma de yuca	8,844,221	7,226,346
Mapea contra el genoma de yuca (nr)	1,183,868	653,480
No mapea rRNAs, snoRNAs, tRNAs*	5,655,383	5,187,027
No mapea rRNAs, snoRNAs, tRNAs (nr)ψ	2,906,014	1,664,361
Mapea genoma Xam	181,808 (3.2%)	374,034 (7.2%)
Mapea genoma Xam (nr)	6,952 (0.2%)	80,850 (4.8%)
Mapea regiones intergénicas de yuca	522,080 (9.2%)	599,178 (11.5%)
Mapea regiones intergénicas de yuca (nr)	87,590 (3%)	64,536 (3.8%)
Mapea miRNAs conocidos	42,356 (0.7%)	122,273 (2.3%)
Mapea miRNAs conocidos (nr)	4,921 (0.1%)	3,398 (0.2%)
Familias de miRNAs conocidos	53	48

nr = no redundante, *Los porcentajes y valores normalizados en este trabajo fueron calculados con base en este valor. ψ Los porcentajes y valores normalizados en este trabajo para lecturas no redundantes fueron calculados con base en este valor

1.2.2 Identificación de miRNAs y pre-miRNAs

Para identificar miRNAs conservados, lecturas de ambas librerías fueron mapeadas contra un set de miRNAs conocidos de miRBase [66]. 56 familias conservadas se encontraron expresadas en yuca, las lecturas que mapeaban contra estos miRNAs eran en su mayoría de una longitud de 21 bases (la distribución de lecturas únicas fue similar a la de lecturas totales). Los niveles de expresión de estas familias fueron confirmados utilizando 3 programas para el mapeo (blastn, bwa y miRprof) [72-74], encontrándose poca variación. Se presentan los resultados obtenidos con miRprof que obtuvo los mayores conteos. Adicionalmente, se mapearon miRNAs provenientes de la base de datos PMRD [62], la cual incluye miRNAs sin confirmación experimental, únicamente 7 familias exclusivas de PMRD se encontraban representadas en las lecturas de yuca, estas familias fueron tenidas en cuenta para la identificación de pre-miRNAs pero dado que no se encontró ningún precursor en el genoma, fueron excluidas de los análisis posteriores.

Para la identificación de pre-miRNAs las lecturas correspondientes a las 56 familias conservadas fueron mapeadas contra el genoma y se analizó la región adyacente (485 regiones en total). Como resultado, 116 regiones cumplían con los criterios principales para ser considerados pre-miRNAs reales (estructura de *hairpin* y alta estabilidad). Estos pre-miRNAs representaban 24 familias (**Anexo A**). El que pre-miRNAs no pudieran ser identificados para ciertas familias se puede explicar considerando que el genoma de yuca está incompleto o por la presencia de pre-miRNAs no-canónicos que no cumplen algunos de los criterios aceptados para miRNAs.

Para identificar miRNAs nuevos, lecturas de tamaños entre 20 y 24 nt provenientes de ambas librerías fueron mapeadas contra regiones intergénicas de yuca (**Tabla 1-1**) y se analizó la región adyacente ± 150 nt (en total 64,876 regiones). Se utilizaron los programas miRcheck[75], miPred [76] y miReap [77] para filtrar posibles pre-miRNAs. Los candidatos que tuviesen soporte de al menos un programa y que además cumplieran con los criterios mencionados anteriormente (estructura de *hairpin* y alta estabilidad) fueron considerados pre-miRNAs candidatos. Se identificaron en total doce nuevos pre-

miRNAs provenientes de 12 familias distintas, y se nombraron Cass_miRA a Cass_miRL. **(Anexo A).**

Las lecturas que mapearon contra estos miRNAs nuevos fueron en su mayoría de longitud de 21 nt, como se observó para miRNAs conservados (**Figura 1-1**) Una búsqueda utilizando blastn contra la base de datos de nucleótidos del Genbank reveló ausencia de similitud (e-value <0.001, >80% cobertura alineamiento) con secuencias de plantas diferentes a yuca para 20 de las 12 familias nuevas identificadas, mientras que dos familias (Cass_miRK y Cass_miRL) presentaban similitud con secuencias no-codificantes de *Populus thricocarpa* y *Helianthus petiolaris* respectivamente.

Se predijeron *targets* para todos los miRNAs identificados contra los genes anotados en yuca utilizando una versión modificada del software miRanda [78, 79] y el programa psRNA*target* [80] **(Anexo B)**. Un alto porcentaje de *targets* identificados fueron factores de transcripción (24%), lo cual es común para muchos miRNAs conocidos en plantas [10]. Otros *targets* comúnmente predichos fueron kinasas (7%), proteínas de unión a DNA (3%), y proteínas de resistencia (3%). Cuatro familias (miR167, miR397, miR894 y Cass_miRJ) tenían entre sus posibles *targets* genes involucrados en biosíntesis o metabolismo de almidón.

1.2.3 Expresión diferencial de miRNAs

Los niveles de expresión de miRNAs fueron cuantificados como el número de lecturas normalizadas (lecturas por millón) mapeadas contra cada miRNA maduro en las librerías. 10 familias de miRNAs conservadas tenían altos niveles de inducción ($\log_2FC > 2$) en respuesta a la infección con Xam (**Figura 1-2**) incluyendo las familias miR160 y miR167 que son conocidas por tener como *targets* factores de respuesta a auxinas (ARFs), al igual que las familias miR393 y miR390, también involucradas en regulación de respuesta a auxinas [21]. En el caso de miR390 esta regulación se hace por medio de la producción de ta-siRNAs [81]. Los *targets* predichos para estas familias en el genoma de yuca correspondían igualmente a los *targets* conocidos. Para algunas de las familias inducidas fue posible confirmar la represión del *target* respectivo, incluyendo miR394,

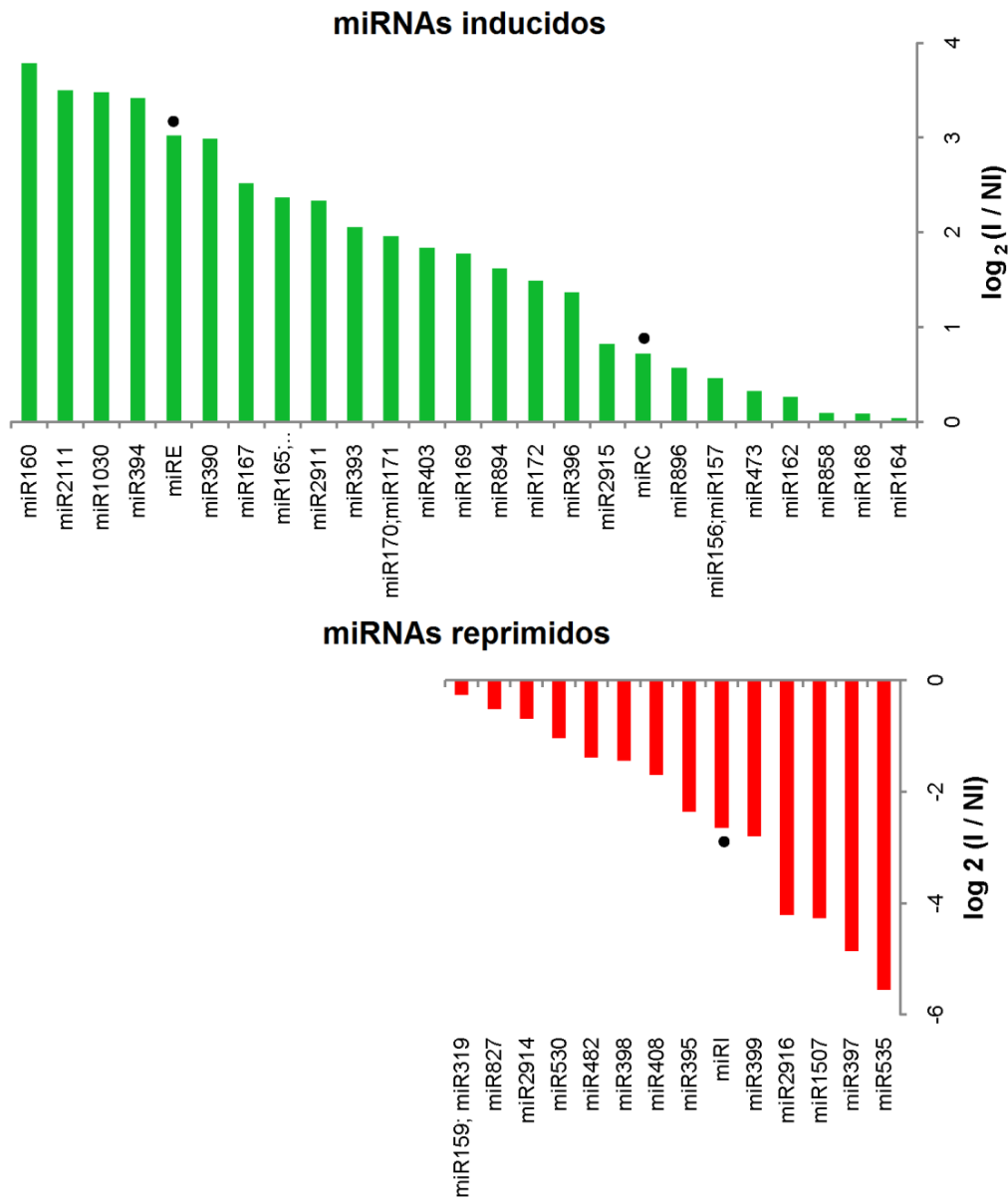
miR165 y miR171 las cuales tienen como blancos respectivos una proteína de la familia F-box, un factor de transcripción ATHB y un factor de transcripción *Scarecrow-like* (Sección 0, **Figura 1-3**).

Por otro lado, la expresión de siete familias se vio reducida en respuesta a la infección por Xam ($\log_2FC < -2$) (**Figura 1-2**). Estas familias reprimidas incluían miR535, miR395 y miR482 las cuales tienen entre sus *targets* predichos varias proteínas con dominios de resistencia tipo NBS-LRR o LRR. Otras familias reprimidas fueron miR397, miR398 y miR408 ($\log_2FC < -1.4$), estas familias están involucradas en la regulación de las vías de cobre mediante el silenciamiento de lacasas, superóxido dismutasas de cobres y plantacyaninas, respectivamente [82]. Fue posible confirmar la inducción de una proteína con dominio de resistencia LRR predicha como *target* de la familia miR535 y de una plantacyanina predicha como *target* de la familia miR408 (Sección 0, **Figura 1-3**).

Los nuevos miRNAs identificados en yuca tenían en general niveles de expresión muy bajos. Únicamente tres nuevos miRNAs (Cass_miRC, Cass_mirE and Cass_miR) tenían valores de expresión mayores a 10 lecturas por millón en al menos una de las librerías. Por tanto, aunque algunas de estas mostraron niveles diferenciales de expresión en respuesta a la bacteria, no fueron consideradas dado el bajo nivel de expresión. Es conocido, en otros sistemas que miRNAs especie-específicos tienden a tener valores de expresión bajos comparados con miRNAs conservados [83]. Cass_miRE fue altamente inducido en respuesta a la bacteria ($\log_2FC = 3$) y la represión de su *target* predicho, una kinasa similar a *FERONIA* de *A. thaliana*, fue observada por RT-PCR (Sección 0, **Figura 1-3**). Se sabe que este gen está involucrado en interacciones con patógenos y en desarrollo de polen [84, 85]

Figura 1-2: Expresión diferencial de miRNAs en respuesta a inoculación por Xam en yuca.

Se calculó el logaritmo en base 2 de la razón entre el número de lecturas normalizadas en la librería inoculada (I) vs la librería no inoculada (NI). Solo se presentan los resultados de familias con al menos 10 lecturas normalizadas en al menos una de las librerías. Los puntos negros indican miRNAs nuevos específicos de yuca.



1.2.4 Expresión diferencial de *targets* de miRNAs

Para validar la regulación de los *targets* predichos para los miRNAs se evaluó la expresión de estos *targets* mediante RT-PCR semicuantitativo. Se espera que la expresión de los *targets* de miRNAs inducidos disminuya en respuesta a la bacteria y lo contrario para los *targets* de miRNAs reprimidos. La expresión de estos *targets* se estudió en tejidos de hoja de la variedad resistente MBRA685 (utilizada para la construcción de las librerías) a 0, 5 y 7 dpi y en la variedad susceptible MCOL1522 a 0, 4 y 8 dpi. Por disponibilidad de material la expresión de algunos *targets* no fue estudiada en MBRA685.

Se encontró que en su mayoría los *targets* respondían al comportamiento esperado (**Figura 1-3**). Para los *targets* de miRNAs inducidos se observó una disminución de la expresión en al menos uno de los tiempos post inoculación, y en algunos casos la expresión parece regresar a niveles normales en el último tiempo observado. De igual manera, para los *targets* de miRNAs reprimidos se pudo ver aumento en los niveles de expresión en al menos uno de los tiempos de inoculación. Notablemente, las diferencias en expresión de *targets* en algunos casos fueron más marcados en la variedad susceptible que en la variedad resistente (p.ej. miR394 y miR167a). Es posible que en la variedad resistente cambios drásticos de expresión se den en tiempos más tempranos no analizados en estos ensayos. Es difícil descifrar la relevancia biológica de este tipo de comportamientos en parte porque las librerías fueron construidas a partir de RNA de varios tiempos. Por tanto, una cuantificación específica de miRNA en ciertos tiempos es aun necesaria para tener una correlación precisa con la expresión de *targets*.

La expresión de *targets* se analizó también a partir de datos de un experimento de RNA-seq orientado a determinar el efecto del efector TALE1_{xam} en yuca [51]. En este experimento se inocularon tallos de yuca de la variedad susceptible MCOL1522 con la cepa de Xam no patógena Orst4 la cual no posee el efector TALE1_{xam} y con la cepa Orst4 complementada con el efector TALE1_{xam}. Se secuenció el transcriptoma de estas plantas a 0 5 y 7 días post inoculación con dos réplicas técnicas y se calcularon los valores de expresión normalizados (RPKM = reads per kilobase per million) para cada

tratamiento [Muñoz-Bodnar *et al.*, in prep]. Se analizó la expresión de los *targets* de miRNAs evaluados por RT-PCR en los datos de RNA-seq, encontrándose que se mantiene la tendencia general de represión de *targets* de miRNAs inducidos, con algunas excepciones (miR167b, miR171), los cambios sin embargo no fueron dramáticos (**Figura 1-3**). Para algunos *targets* no se encontraron lecturas suficientes en los datos de RNA-seq para determinar su expresión (miR394 y miR535). Para el único *target* de miRNA reprimido evaluado (miR408) no se observó claramente inducción de la expresión. No se observaron tampoco diferencias notorias entre los tratamientos con y sin el efector TALE1_{Xam}.

Sin embargo los datos de RNA-seq pudieron ser utilizados para analizar el comportamiento de un mayor número de *targets* de miRNAs para observar mejor una tendencia (**Figura 1-4**). Se calculó el cambio en expresión (Log2FC) de los *targets* reprimidos e inducidos a cada tiempo post inoculación respecto al tiempo 0 para cada tratamiento. Se utilizaron únicamente los *targets* predichos con ambos algoritmos de predicción utilizando miRanda [78, 79] y psRNA*target* [80]. Aunque se observó gran variación en los niveles de expresión se encontró que los *targets* de miRNAs reprimidos en respuesta a bacteria eran significativamente más inducidos que los *targets* de miRNAs inducidos (Wilcoxon p-value <0.1). Curiosamente, a los 7 días post-inoculación los rangos de Log2FC fueron más estrechos en las plantas inoculadas con Orst4+ TALE1_{Xam} que en plantas inoculadas con Orst4, esto sucedió tanto para *targets* de miRNAs inducidos y reprimidos (**Figura 1-4**). Esto sugiere que la presencia del efector restringe la expresión de los *targets* de miRNAs al menos a los 7dpi, esto puede estar relacionado con una actividad de supresor de silenciamiento del efector [31], la cual podría estar dada por la inducción de genes supresores de miRNAs gracias a la actividad de factor de transcripción de este efector. El hecho de que el o los probables genes inducidos por TALE1_{Xam} sean igualmente factores de transcripción dificulta dilucidar la vía por la cual se puede dar esta actividad de supresión [51]

Figura 1-3: Expresión de *targets* de miRNAs a diferentes tiempos post-inoculación con Xam en experimentos de RT-PCR semicuantitativo (izquierda) y RNA-seq (derecha).

Se muestra el tipo de interacción evaluada (resistencia de la variedad / agresividad de la cepa). Los números entre paréntesis corresponden a los identificadores de los genes estudiados. En experimentos con MBRA685 se eliminaron tiempos tempranos ya que los niveles del gen control (tubulina) no eran iguales a los otros tratamientos. Los datos en los experimentos de RNA-seq corresponden a valores promedio de RPKM entre 2 réplicas técnicas, las barras de error muestran la desviación estándar.

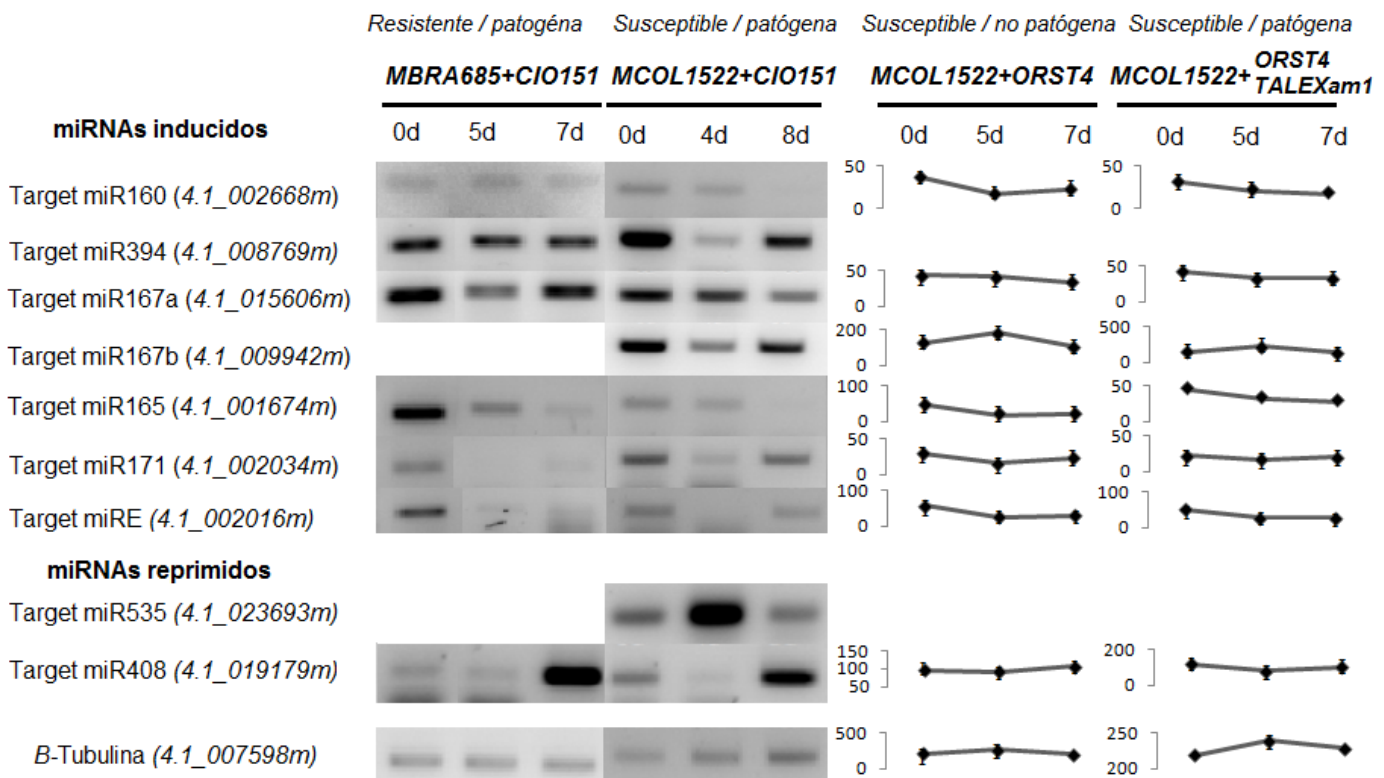
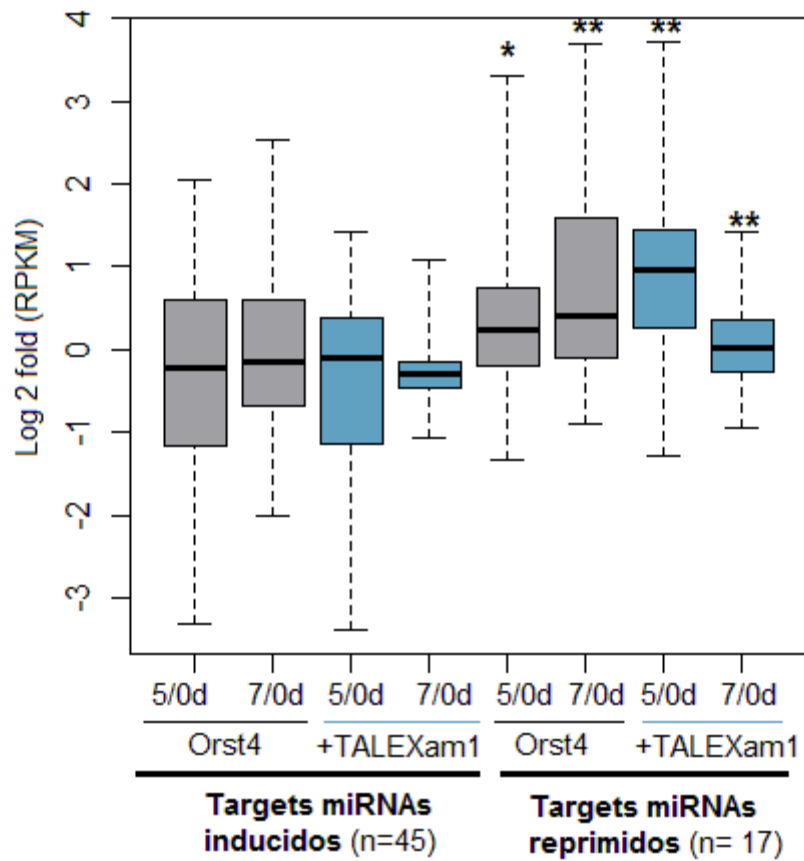


Figura 1-4: Boxplot mostrando los niveles de expresión de *targets* de miRNAs en datos de RNA-seq.

Se muestra el cambio en la expresión entre plantas inoculadas a 5 y 7 días respecto a plantas no inoculadas, con la cepa de Xam no patógena Orst4 y con la misma cepa complementada con el efector TALE1Xam. Asteriscos indican significancia de pruebas Wilcoxon a dos colas entre miRNAs reprimidos y miRNAs inducidos para cada tratamiento (* = p-value<0.1, ** = p-value<0.05)



1.3 Discusión

En este trabajo se identificaron familias de miRNAs nuevas y conservadas en el genoma de yuca y se analizó su expresión en librerías de sRNAs. Cabe notar que el número de loci de pre-miRNAs conservados identificados en este trabajo (114) es inferior a aquellos reportados por otros trabajos de hasta 163 loci [65, 68]. Estas diferencias se deben principalmente a que en este trabajo se buscaron únicamente los pre-miRNAs para los que existiese evidencia de expresión en librerías de sRNAs y además los criterios para identificación de estos loci fueron mucho más estrictos.

En el modelo general del papel de miRNAs en la interacción con bacterias (**Figura 0-1**), se propone que miRNAs cuyos *targets* son reguladores negativos de defensa son inducidos en respuesta a bacterias y que de igual manera se reprimen miRNAs cuyos *targets* son reguladores positivos de defensa [86, 87]. En este trabajo la expresión diferencial de miRNAs de yuca encontrada en respuesta a Xam concuerda en gran parte con este modelo. Por un lado, se encontró inducción de miRNAs involucrados en regulación de señalización de auxinas: miR160, miR167, miR390 y miR393, al igual que se ha encontrado en *Arabidopsis* [19, 21]. Recientemente además estos resultados fueron ampliados al estudiar la regulación de TAS3 en yuca, un locus de ta-siRNAs que es procesado por la familia miR390, encontrándose que en respuesta a la bacteria existe también aumento en la producción de estos ta-siRNAs cuyos *targets* son principalmente factores de respuesta a auxinas [Quintero *et al.*, sometido].

La conservación de los patrones de expresión en yuca y uca y *Arabidopsis* apunta a un mecanismo de defensa conservado, probablemente relacionado con PTI, dado que algunas de estas familias se han visto inducidas en respuesta al PAMP flg22 [22]. La interrupción de señalización de auxinas parece ser una estrategia importante para impedir el crecimiento bacteriano en plantas. Trabajos recientes han mostrado que la interrupción de estas vías por mecanismos enzimáticos es importante para conferir resistencia de amplio espectro, y que algunos patógenos secretan auxinas durante la infección para hacer a la planta vulnerable mediante el debilitamiento de la pared celular [20, 88].

De igual manera la represión de miRNAs mostrada en este trabajo concuerda con el modelo general, pues por un lado tres familias reprimidas (miR535, miR395 y miR482) tenían entre sus *targets* predichos principalmente genes codificantes para proteínas con dominios presentes en las proteínas R. Recientemente se ha encontrado que la regulación de genes *R* por miRNAs y otros sRNAs es un mecanismo prevalente en plantas y parece tener un papel predominante en respuestas de defensa [23, 24, 29, 89]. En solanáceas por ejemplo, se propone que la regulación de genes tipo NBS-LRR por la familia miR482 disminuye el costo fisiológico que implica para la planta la regulación de este tipo de genes, este costo está asociado con la necesidad de impedir la activación de estas proteínas en ausencia del patógeno. Se propone que estos genes se mantienen a un bajo nivel de expresión controlados por miRNAs y que son todos activados luego de la detección de PAMPs y efectores [23], prescindiendo así de los mecanismos usuales de regulación de este tipo de proteínas que normalmente implican cambios conformacionales (en algunos casos dimerización) e hidrólisis de ATP en presencia del efector del patógeno [90].

Otras familias reprimidas fueron miR397, miR398 y miR408, que se sabe están involucradas en regulación de cobre y se han encontrado diferencialmente expresadas en varios tipos de estrés biótico y abiótico [82, 91, 92]. El cobre es ampliamente utilizado como pesticida en agricultura por su efecto antibacterial [93], y en el caso de la interacción de arroz con *Xanthomonas oryzae* se ha mostrado que un efector tipo TAL modifica la distribución de cobre en la planta para hacerla susceptible [94]. Sin embargo artículos más recientes señalan que el proceso involucrado en este mecanismo es en realidad la distribución de azúcares en el sistema vascular [95, 96].

Los patrones de regulación acá encontrados parecen darse de igual manera en variedades susceptibles y resistentes, aunque, la caracterización de estas variedades no es aún precisa, dada la naturaleza cuantitativa de la resistencia en yuca. Es posible que las diferencias entre resistencia y susceptibilidad, si están dadas por miRNAs, estén relacionadas con la activación temprana o tardía de estas vías de regulación.

Se encontró además curiosamente que en presencia al efector TALE1_{Xam} parece haber supresión del silenciamiento, evidenciado en menores niveles de inducción o represión de *targets* de miRNAs. Esto es similar a lo que se ha encontrado para algunos efectores de Pto [31]. De confirmarse estos resultados, sería la primera vez que se encuentra este tipo de actividad para un efector tipo TAL, aunque existen ya indicios de que la disrupción de la vía de miRNAs por TALs es relevante para inmunidad, particularmente por la ubiquidad de efectores TALs en cepas de *Xanthomonas oryzae* que se predice inducen la expresión de HEN1 [97], aunque el efecto que esto pueda tener en la vía de miRNAs es aún desconocido (Sección 4.2.2).

Este trabajo muestra la caracterización de miRNAs en respuesta a infección por Xam y reafirma la importancia de estas moléculas en la interacción planta-patógeno. Éstos datos son promisorios pero aun preliminares y se necesita más validación experimental para entender completamente el impacto de la vía de miRNAs en esta interacción, particularmente dada la heterogeneidad de los experimentos y datos acá. utilizados

Resultados adicionales de los análisis de estas librerías (p.ej. expresión de hebra pasajera miRNA* y transitividad de *targets*) son detallados en Pérez-Quintero *et al.* [69].

1.4 Materiales y métodos

1.4.1 Materiales vegetales e inoculación

Plantas de yuca fueron plantadas a partir de corte de estacas y mantenidas en invernadero a temperatura de 26°-30°C, con un fotoperiodo de 12 horas de luz y humedad relativa de 80%. La variedad de yuca MBRA685 (resistente a Xam-CIO151) fue utilizada para la construcción de librerías de RNAs pequeños y las variedades MBRA685 y MCOL1522 (susceptible a Xam-CIO151) fueron utilizadas para los experimentos de RT-PCR. Por disponibilidad de material la expresión de algunos genes sólo pudo ser estudiada en la variedad susceptible.

Plantas de seis semanas fueron inoculadas con cultivos de 36 horas de la cepa Xam CIO151 en hojas y tallos (sólo hojas fueron utilizadas para RT-PCR). Las hojas fueron inoculadas mediante perforación de seis agujeros en el mesófilo y colocando una gota de 5uL de cultivo Xam-MgCl₂ a una concentración OD_{600nm} = 0.002 (1 x10⁸cfu/ml). Dos foliolos por hoja y tres hojas por planta fueron inoculadas. Los tallos fueron inoculados por punción. Al menos tres plantas fueron inoculadas por tiempo de colección para cada experimento.

1.4.2 Construcción y secuenciamiento de librerías de sRNAs

Para la librería inoculada, hojas y tallos fueron colectados de plantas inoculadas (0, 6 y 24hpi, y 2,5,7 y 15dpi). Se hicieron extracciones de RNA utilizando un método basado en LiCl - fenol ácido:cloroformo. La extracciones de distintos tiempos (al menos 6 plantas por tiempo) fueron mezcladas en igual cantidad tras medición de la concentración con nanómetro. Para la librería no-inoculada, una sola extracción a partir de plantas no tratadas fue utilizado. Las librerías fueron construidas en el BC Cancer Agency's Michael Smith Genome Sciences Centre (<http://www.bcgsc.ca/>) y secuenciadas utilizando el Illumina Genome Analyzer IIx (<http://www.illumina.com/>). Para esto se ligaron adaptadores a la muestra de RNA, se sintetizó cDNA y se secuenció la fracción de 10 a 50nt separadas en un gel de policarilamida denaturante.

1.4.3 Análisis de datos de secuenciación

Se aplicaron varios filtros de calidad a los datos brutos. Se eliminaron secuencias con un valor de castidad menor a 0.6 en las primeras 25 bases (parámetro estándar del secuenciador) y secuencias con más del 50% de la longitud con bases con un puntaje *Phred* menor a 30. Se eliminó la secuencia de los adaptadores a través del desarrollo de un programa en C++; este programa eliminaba cualquier secuencia con fragmentos mayores a 10nt de la secuencia original que tuvieran matches perfectos con la secuencia de adaptadores utilizados en las librerías. Sólo se removían secuencias menores a 10nt si se encontraban a mas de 15nt en la secuencia original. El programa utilizaba la

herramienta wordfinder de la suite EMBOSS [98] para encontrar fragmentos de adaptador (parámetros: *minimun match score =70, alignment width = 6, gap opening 0.0 + gap extension 10 or gap opening 500 + gap extension 10*). Secuencias menores a 15nt así como secuencias de baja complejidad (más de 15 nucleótidos idénticos consecutivos) fueron removidas utilizando comandos unix.

Las lecturas fueron mapeadas contra snoRNAs, tRNAs y rRNAs conocidos obtenidos de la base de datos Rfam [99] y luego removidas. El mapeo fue hecho utilizando blastn (parámetros: *e-value<0.0001, ungapped, word size = 4, -max_target_seqs 1 -best_hit_score_edge 0.05 -best_hit_overhang 0.25*). A menos que se indique, estos fueron los parámetros utilizados para todos los análisis de mapeo.

Las lecturas también fueron mapeadas contra el genoma de yuca v.4. [67] (*parámetro: e-value<1e-5, 100% identidad*) y contra el genoma de la cepa Xam CIO151 [50][Arrieta et al. in prep] (*parámetros; e-value<1e-5, 100% identidad*).

1.4.4 Identificación de miRNAs y premiRNAs

Para identificar miRNAs conservados filogenéticamente, lecturas de ambas librerías de sRNAs fueron mapeadas contra el set de miRNAs maduros del grupo *Viridiplantae* obtenido de miRBase version 16 [66] y el set completo de miRNAs de la abse de datos PMRD [62]. Lecturas con menos de dos mismatches con un miRNA conocido fueron considerados miRNAs conservados [16]. El perfil de expresión de miRNAs conservados fue obtenido con la herramienta miRProf del kit UEA sRNA [72], y fueron confirmados utilizando mapeo con blastn [74] y bwa [73]. Para la identificación de miRNAs nuevos, las lecturas de ambas librerías fueron mapeadas contra el genoma de yuca.

Para los análisis de pre-miRNAs se utilizó la región adyacente (-150, +150 nt) a las posiciones mapeadas de una lectura de interés. Se eliminaron las regiones que tuvieran solapamiento de más del 25% de la longitud con regiones codificantes.

Se tomaron dos criterios principales para considerar una región como un pre-miRNA real: características estructurales identificadas a partir de la predicción de estructuras secundarias y una prueba estadística de la estabilidad de la estructura. Las estructuras secundarias fueron obtenidas con RNAFold del paquete Vienna RNA [100]. Estas estructuras fueron analizadas con un script en perl, se consideraron pre-miRNAs reales si tenían menos de 6 *mismatches* entre el miRNA maduro predicho y la hebra pasajera miRNA*, igualmente debían tener máximo 3 desacoplamientos (bultos) asimétricos en la estructura [16, 101]. Por otro lado se calculó la significancia de la mínima energía de doblamiento (MFE) utilizando una prueba estadística: se generaron 1000 secuencias aleatorias para cada precursor manteniendo la misma composición de bases y la frecuencia de dinucleótidos utilizando ushuffle [102]. Se calculó la MFE para las estructuras de las secuencias aleatorias y se calculó el p-value como el porcentaje de estructuras aleatorias con una MFE igual o menor al precursor original [103]. Los pre-miRNAs reales debían tener un p-value menor a 0.05.

Los candidatos para pre-miRNAs nuevos fueron analizados y filtrados utilizando miRcheck [75] (parámetros por defecto) MIREAP [77] (*parámetros* $B= 55$, $a = 19$, $b = 24$, $u = 1,000$, $e = -10kcal/mol$, $d= 200$, $p = 7$, $v = 10$, $s = 100$, $f = 10$) y miPred [76] (parámetros por defecto).

Para la búsqueda de *targets* se utilizó una versión modificada del programa miRanda version v. 2.0 [104] optimizada para criterios de pares miRNA:*target* en plantas [78]. Se buscaron también *targets* utilizando psRNA*target* v [80]. La predicciones se hicieron sobre las regiones codificantes anotadas en el genoma de yuca obtenidas de <http://www.phytozome.net/cassava>.

1.4.5 Expresión de *targets* de miRNAs en experimentos de RT-PCR semicuantitativo y RNA-seq

Para los experimentos de RT-PCR semicuantitativo se extrajo RNA de hojas de las variedades MBRA685 y MCOL1522 de plantas no-inoculadas e inoculadas (4 o 5 dpi y 7 u 8 dpi). Se aplicó tratamiento con DNAsa I a las muestras (Fermentas), posteriormente

se sintetizó cDNA utilizando First Strand cDNA Synthesis con oligo dTs (Fermentas). Se normalizó la cantidad de cDNA luego del PCR utilizando como referencia el gen de *tubulina* (F= 5'-GATCCTACTGGGAAGTACATTGG-3', R= 5'-GATCATTCTCCACCAACTGA-3'). Se diseñaron primers específicos para *targets* predichos de miRNAs que cubrieran la región conteniendo el sitio de clivaje predicho para el miRNA. Los PCRs se realizaron con 32 ciclos = 94°C por 30s, 54-60°C por 30s y 72°C for 30s.

Los datos de RNA-seq fueron obtenidos de un experimento en que se obtuvo el RNA de plantas inoculadas con la cepa de Xam no patógena Orst4, y con la cepa Orst4 complementada con el efector TALE1_{Xam}. Se secuenciaron muestras obtenidas a tiempos de inoculación 0dpi, 4dpi y 8dpi para cada tratamiento, y se realizaron 2 réplicas técnicas para cada tratamiento. Las librerías fueron mapeadas sobre el set de genes anotados de yuca utilizando seqmap [105] y se obtuvieron valores de expresión normalizados (RPKM) utilizando Rseq [106]. Para los análisis presentados se utilizó el promedio de las dos réplicas técnicas para cada tratamiento tras comprobar que no existía alta variación entre ambas réplicas. Se identificaron genes con expresión diferencial significativa en respuesta al efector utilizando un procedimiento similar al descrito en la sección 2.4.4. Los detalles de este trabajo serán presentados en (Muñoz-Bodnar *et al.* 2012).

2. miRNAs en la interacción *Oryza sativa* -*Xanthomonas oryzae* pv. *oryzae*

2.1 Introducción: el sistema arroz-Xoo.

El arroz (*O. sativa*, Poaceae) es sin duda un cultivo de gran importancia a nivel mundial, ocupa el primer lugar en consumo y tercer lugar en producción total a nivel mundial con 690 millones de toneladas anuales, superado únicamente por maíz y caña de azúcar (<http://faostat3.fao.org>, datos 2010). Más del 80% de la producción está destinada al consumo directo por humanos, alcanzando a representar hasta el 65% del consumo calórico en algunos países [107]. En Colombia, este cultivo ocupa el tercer lugar en producción con un total de 2.4 millones de toneladas anuales (<http://faostat3.fao.org>, datos 2010). Es por otro lado una especie importante en investigación agrícola. Los genomas de las dos subespecies principales de arroz: *indica* y *japonica*, fueron secuenciados en el 2002, convirtiendolo en la segunda especie de plantas con genoma completamente secuenciado y en el primer modelo genético de cereales y en general de monocotiledóneas [108, 109].

De igual forma el arroz ha sido también una planta modelo en el estudio de interacciones planta-bacteria, particularmente en su interacción con la especie *Xanthomonas oryzae*, la

cual tiene además gran impacto económico [1]. Existen principalmente dos patovares de esta especie que atacan arroz: *Xanthomonas oryzae* pv. *oryzae* (Xoo) y *Xanthomonas oryzae* pv. *oryzicola* (Xoc), las cuales a pesar de estar estrechamente relacionadas, invaden la planta de distintas formas causando diferentes enfermedades [1].

Xoo es el agente causal de la bacteriosis vascular de arroz (Rice Bacterial Blight), en esta enfermedad la bacteria invade principalmente el sistema vascular tras entrar a la planta mediante heridas, tricomas, e hidátodos en las puntas o márgenes foliares. La enfermedad se caracteriza inicialmente por la presencia de manchas acuosas que luego se expanden a lo largo de los vasos vasculares tornándose cloróticas y eventualmente necróticas, típicamente estas lesiones se extienden de la punta de la hoja hacia abajo [5, 110, 111]. Al cabo de varios días de infección la bacteria exuda a través de los hidátodos convirtiéndose en fuente de infección secundaria [111]. Por el contrario, Xoc es el agente causal de la raya foliar bacteriana del arroz (Rice Bacterial Leaf Streak), enfermedad en la cual la bacteria entra también por heridas pero invade únicamente el apoplasto del tejido del mesófilo [1, 110]. Los síntomas inician en cualquier parte de la hoja, típicamente como manchas amarillas y se expanden longitudinalmente delimitadas por los vasos principales [1, 110].

La bacteriosis vascular se encuentra en regiones tropicales y templadas y ha sido detectada en cultivos en Australia, América Latina, Asia, África y los Estados Unidos [1, 96, 112]. Las pérdidas reportadas para esta enfermedad oscilan entre el 10% al 50% de la producción [2]. Mientras que la raya foliar bacteriana es una enfermedad esporádica, restringida principalmente a Asia sub-tropical, con pérdidas reportadas de hasta el 20% [1, 2]. Este trabajo tratará principalmente la interacción entre arroz y Xoo siendo ésta la bacteria de mayor importancia económica. Algunos aspectos de la interacción con Xoc son tratados brevemente en la sección 4.2.2.

Existe gran diversidad genética en poblaciones de Xoo, con más de 30 razas distintas identificadas en diferentes partes del mundo. Actualmente se cuenta con la secuencia completa del genoma de tres cepas: MAFF311018, KACC10331 y PXO99A, provenientes de Japón, Korea y Filipinas, respectivamente [113-115]. Entre estas, la cepa PXO99A es altamente virulenta sobre múltiples variedades de arroz [114]. Se han caracterizado hasta

la fecha distintos factores implicados en la virulencia de estas cepas, particularmente se ha encontrado que estas bacterias necesitan efectores tipo TAL para una colonización exitosa y todas las cepas secuenciadas tienen efectivamente amplios repertorios de estos efectores (más de 15 en cada cepa) [114, 116, 117].

En este sistema precisamente, se ha caracterizado uno de los mejores ejemplos de cómo la bacteria utiliza estos efectores para modificar procesos fisiológicos en la planta y desarrollar la infección. Tres de los efectores TALs para los que se ha identificado un papel determinante en virulencia se unen a las regiones promotoras de genes de la familia de Nodulinas Mtn3 (también conocidos como *SWEET*) induciendo su expresión [118, 119]. Entre estos, el efector PthXo1, el principal determinante de virulencia de la cepa PXO99A, induce la expresión del gen *Os8N3* (*Xa13/ OsSWEET11/ LOC_Os08g42350*) [94, 119, 120]. Esta familia de proteínas está involucrada en el transporte de azúcares hacia los vasos vasculares y se cree que la inducción de esta familia es estimulada por la bacteria para satisfacer sus requerimientos nutricionales [95, 96]. Estas proteínas pueden además funcionar eliminando cobre de los vasos vasculares, lo que representaría también una ventaja para la bacteria dado el papel bactericida del cobre [94].

Recientemente se inició la caracterización de cepas de Xoo provenientes de África, donde la enfermedad se registró por primera vez en 1979 y se ha expandido en relación directa con la expansión e intensificación de cultivos de arroz en este continente [121]. Análisis genéticos mostraron que las cepas africanas forman un clado distinto de las cepas asiáticas y parecen estar más relacionadas con cepas de Xoc. Sin embargo los síntomas causados por las cepas de Xoo africanas son idénticos a los de la bacteriosis vascular causada por cepas asiáticas [122]. Notablemente, las cepas africanas tienen números reducidos de efectores TAL en el genoma, usualmente menos de 5, sin embargo parecen igualmente ser determinantes para la virulencia [122]. En una de las cepas Xoo africanas estudiadas XooBA13, se logró identificar un efector TAL determinante para la virulencia, llamado TALC [123]. Curiosamente este efector también induce la expresión de un gen de la familia *Mtn3*, el gen *Os11N3* (*OsSWEET14, LOC_Os11g31190*), confirmando nuevamente que la inducción de genes de esta familia es importante para la colonización del sistema vascular [123].

Hasta la fecha se han identificado más de 20 genes de resistencia a distintas cepas de Xoo y el uso de variedades resistentes sigue siendo una de las mejores estrategias para combatir la enfermedad [1], aunque en muchos casos la bacteria puede eventualmente escapar de la acción de estos genes [124]. Los genes de resistencia identificados en arroz son de distintos tipos incluyendo genes que codifican proteínas de tipo NBS-LRR y RLK (Receptor-like kinase). Existe también la presencia de genes de resistencia recesivos [116], varios de los genes de resistencia recesivos identificados en arroz corresponden de hecho a variantes de genes que han evitado la inducción por efectores TAL mediante mutación en sus regiones promotoras [34, 119, 125]. Precisamente este mecanismo fue recientemente aprovechado para generar plantas de arroz resistentes a algunas cepas de Xoo mediante la edición artificial de la región promotora de uno de los genes de la familia *MtN3* (*Os11N3*) normalmente inducido por efectores TAL [126]. De manera similar, se ha utilizado tecnología de miRNAs artificiales para lograr resistencia mediante el silenciamiento tejido-específico del alelo dominante (susceptible) *Xa13*, normalmente inducido por PthXo1 [127].

A pesar de la importancia de la enfermedad y de que se conocen muchos de los mecanismos moleculares involucrados en la resistencia y la susceptibilidad de la planta, aún no se ha estudiado el papel que los miRNAs pueden jugar en esta interacción. Se han realizado sin embargo muchos estudios de miRNAs en arroz, y se han identificado en total 591 loci de miRNAs agrupados en 333 familias [66]. Algunos se han identificado mediante análisis masivos de librerías de sRNAs con la correspondiente validación experimental [70, 128-132]. Es probable que la identificación de los loci de miRNAs ya haya llegado a su punto de saturación. Se ha identificado expresión tejido-específica de miRNAs [70, 133], así como miRNAs diferencialmente expresados en distintos tipos de estrés incluyendo respuestas a frío [134], metales pesados [135], ausencia de fósforo [70], ausencia de azufre [70], sequía [136], salinidad [137], virus [132, 138], hongos [128], entre otros. Notablemente, a partir de trabajos en arroz se descubrió que miRNAs de plantas podían tener actividad en células de mamíferos al ser ingeridas en la dieta [139], lo que sin duda representa un hito en la investigación de miRNAs y probablemente tendrá grandes repercusiones en investigación biomédica.

En este capítulo se presentan los resultados del análisis de librerías de RNAs pequeños de arroz obtenidas a partir de plantas infectadas con la cepa Xoo BAI3 y Xoo BAI3ΔHRP (mutante para el T3SS), y se comparan con datos disponibles de un experimento similar con las cepas PXO99A y PXO99AΔHRC (también mutante para el T3SS). A partir de estos análisis se identifican miRNAs posiblemente involucrados en PTI y ETS y se analizan los perfiles de expresión de los posibles *targets* en datos de RNA-seq y microarreglos.

2.2 Resultados

2.2.1 Secuenciación de librerías de sRNAs en arroz

Para este trabajo se construyeron librerías de RNAs pequeños a partir de RNA extraído de plantas de arroz (*O. sativa* subsp. japonica cv. Nipponbare) sometidas a tres tratamientos: plantas control inoculadas con agua, plantas inoculadas con la cepa africana Xoo BAI3 y plantas inoculadas con la bacteria mutante Xoo BAI3ΔhrcC [123]. En esta última cepa se presenta una mutación en el gen *hrcC*, el cual codifica para una proteína constitutiva y fundamental para el ensamblaje y funcionamiento del T3SS, por tanto la bacteria no puede inyectar efectores en la planta [140]. El RNA fue extraído a las 24 hpi ya que en este tiempo se presentan eventos relevantes para la infección en este sistema, tales como la inducción del gen *Os11N3* por el efector TALC en la cepa silvestre [123]. Para cada tratamiento se disponen de librerías construidas a partir de dos réplicas biológicas (distintos grupos de plantas inoculadas). En este trabajo estas librerías serán designadas B-Control (plantas tratadas con MgCl₂), B3 (plantas inoculadas con la cepa africana BAI3) y B3ΔHRC (plantas inoculadas con la cepa africana BAI3 mutada en el T3SS), acompañada del número de cada réplica (1 y 2). Previo a la secuenciación se confirmó mediante Q-PCR la inducción del gen *Os11N3* en las muestras de plantas inoculadas con XooBAI3 mientras que en las plantas inoculadas con Xoo BAI3ΔhrcC o en las plantas control no se observó expresión de este gen (**Figura 2-1**). A partir de estos tratamientos se podrán observar diferencias en miRNAs involucrados en PTI

(diferencialmente expresados en B3 Δ HRC) y dependientes de efectores o involucrados en ETS (diferencialmente expresados en B3).

Adicionalmente, existe disponible un set librerías de RNAs pequeños de un experimento similar presente en el repositorio de secuencias de sRNAs de arroz MPSS [141]. Estas librerías fueron construidas a partir de tejidos de arroz (*Oryza sativa* subsp. japonica cv. Nipponbare) inoculadas con la cepa asiática Xoo PXO99A, la cepa Xoo PXO99A Δ hrcC, mutante para el T3SS, y plantas inoculadas con MgCl₂ 10mM como control. En estos datos no existen réplicas. Estas librerías son designadas en este trabajo como P-control, PXO y PXO Δ HRC. Este experimento permitirá de igual forma identificar miRNAs involucrados en PTI y ETS.

Para el experimento empleando la cepa BAI3 se obtuvieron librerías sin adaptadores y con alta calidad, por lo que no se aplicó ningún filtro adicional, con la excepción de la eliminación de lecturas menores a 15 nt. De igual manera en el experimento empleando la cepa PXO99A, las librerías se encuentran disponibles sin adaptadores y no se encuentran valores de calidad asociados. La distribución de tamaños de las lecturas totales fue similar entre las librerías (**Figura 2-2**) y corresponde a las distribuciones observadas en otros experimentos en arroz: gran abundancia de lecturas de 21 y 24 nt y gran diversidad de lecturas de 24nt [70].

Se excluyeron lecturas que no mapearan contra el genoma de arroz, así como lecturas que mapearan contra tRNAs, rRNAs y snoRNAs, obteniendo en promedio 10 millones de lecturas por librería en el experimento BAI3 y 3.5 millones por librería en el experimento PXO99A (**Tabla 2-1**). Adicionalmente las lecturas totales de cada librería fueron mapeadas contra el genoma de Xoo PXO99A (el genoma de Xoo BAI3 no está disponible), se verificó una mayor cantidad de lecturas provenientes de la bacteria en las librerías de plantas inoculadas (tanto silvestre como mutante) comparada con las librerías control (**Tabla 2-1**). Para el experimento BAI3, en las librerías inoculadas se presentaron diferencias en el porcentaje de lecturas de la bacteria entre réplicas biológicas, esto puede indicar variación en la población de bacterias entre las réplicas (**Tabla 2-1**).

Figura 2-1: Expresión del gen *Os11N3* evaluada por Q-PCR en muestras de RNA utilizadas para secuenciación.

Se muestra la expresión relativa (Ct) del gen *Os11N3* comparada con un gen constitutivo (*actina*) en muestras de RNA obtenido de 3 réplicas biológicas en 3 tratamientos. Se muestra la desviación estándar de 2 réplicas técnicas.

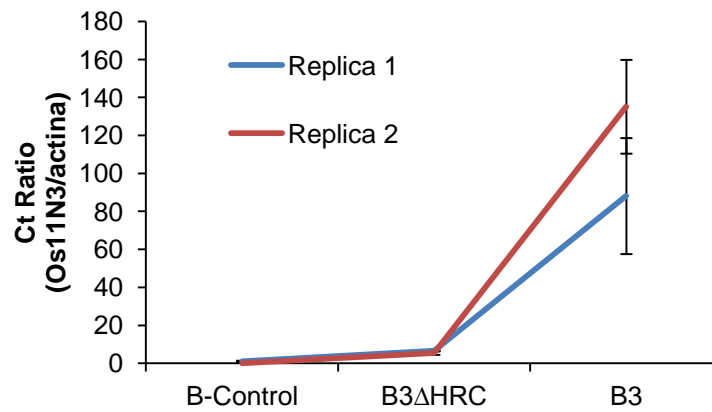


Figura 2-2: Distribución de tamaños de lecturas de librerías de sRNAs en arroz.

Se muestran los porcentajes para cada tamaño en los experimentos con Xoo BAI3 (izq.) y Xoo PXO99 (der.) entre a) lecturas totales, b) lecturas no-redundantes o repetitivas c) lecturas totales mapeadas contra miRNAs de arroz y d) lecturas no-redundantes mapeadas contra miRNAs de arroz. Se muestra también e) la distribución de tamaños de los miRNAs descritos en arroz y en todas las plantas, depositados en miRBase.

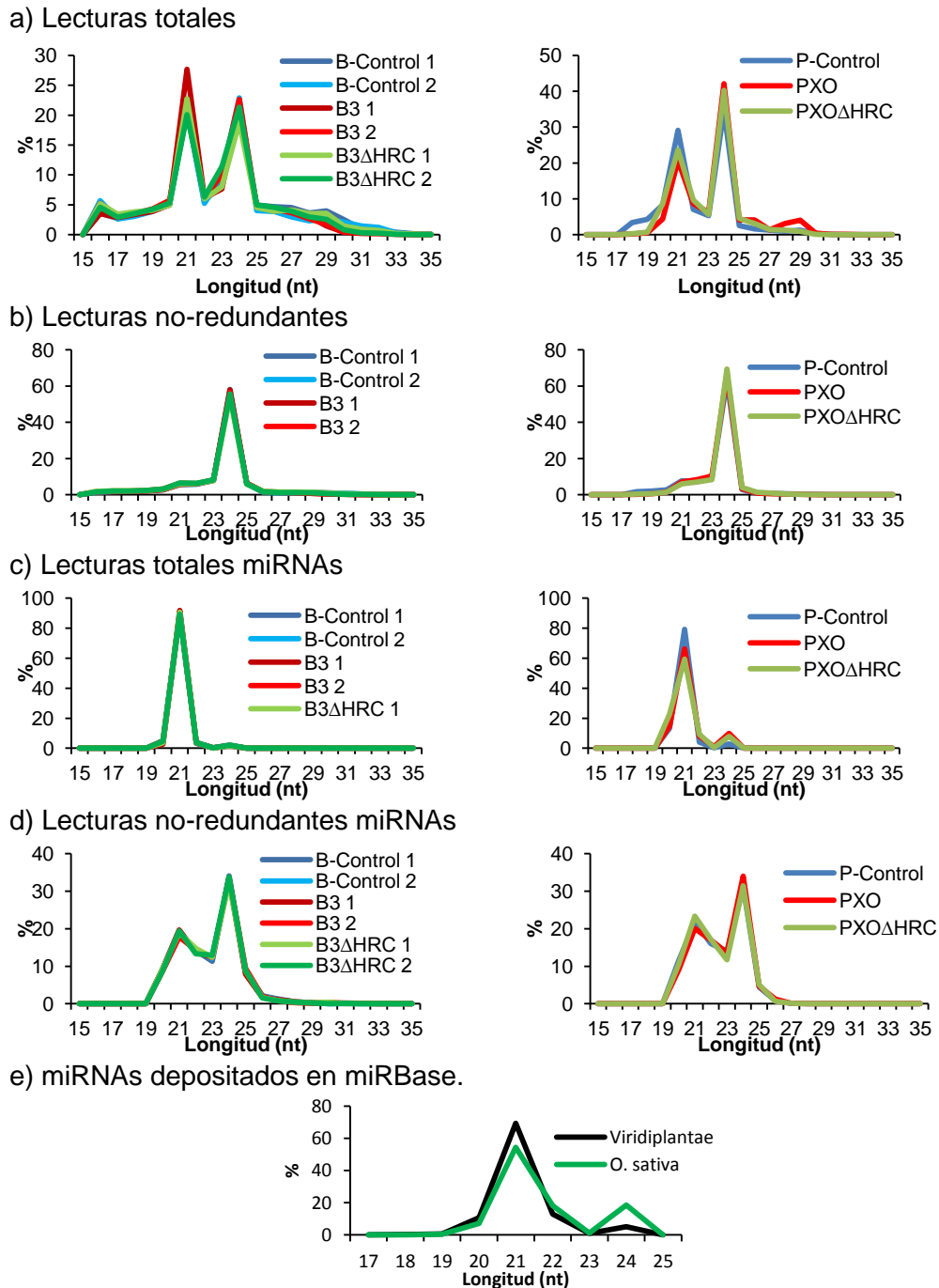


Tabla 2-1: Estadísticas generales del procesamiento y mapeo de librerías de RNAs pequeños de arroz.

nr = no redundante, *Los porcentajes y valores normalizados en este trabajo fueron calculados con base en este grupo de lecturas. ψ Los porcentajes y valores normalizados en este trabajo para lecturas no redundantes fueron calculados con base en este grupo de lecturas. Los

	Xoo BAI3						Xoo PXO99A		
	B-Control		B3		B3 Δ HRC 1		P-Control	PXO	PXO Δ HRC
	1	2	1	2	1	2			
Lecturas totales	1.4E+07	1.6E+07	1.6E+07	1.6E+07	1.5E+07	1.5E+07	4.7E+06	7.4E+06	4.6E+06
Mapea genoma arroz	1.2E+07	1.3E+07	1.1E+07	1.2E+07	1.1E+07	1.2E+07	3.9E+06	4.0E+06	3.5E+06
Mapea genoma arroz (nr)	1.6E+06	2.0E+06	1.6E+06	1.9E+06	1.5E+06	1.8E+06	1.0E+06	7.4E+05	3.5E+05
No tRNA, rRNA (%)*	84.2	88.4	85.6	87.5	83.5	85.9	92.2	95.8	95.1
No tRNA, rRNA, (% nr) ψ	96.4	97.2	97.1	97.4	96.6	97.0	97.7	98.2	97.4
Genoma XooPXO99 (%)	0.044	0.041	8.478	3.050	6.085	1.910	0.011	2.844	3.612
Genoma XooPXO99 (%nr)	0.136	0.122	6.691	4.011	4.277	2.281	0.032	4.644	4.472
miRNA (%)	17.4	17.9	25.8	18.0	20.6	15.9	27.3	21.9	32.3
miRNA (%nr)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.3
Familias miRNA	191	192	189	189	181	187	167	171	143

porcentajes de lecturas mapeadas contra el genoma de Xoo PXO99 son calculados sobre lecturas totales.

2.2.2 Expresión diferencial de miRNAs

Dada la abundancia de miRNAs ya identificados en arroz y la cantidad de estudios que se han realizado en esta planta [70, 128-132], se decidió para esta parte del trabajo analizar únicamente la expresión de miRNAs ya identificados y no focalizarse en la búsqueda de miRNAs nuevos. Se mapearon las lecturas de las librerías de sRNAs de arroz contra el set de miRNAs de arroz disponibles en miRBase, identificándose la expresión de al menos 140 familias distintas en cada librería (**Tabla 2-1**). Se realizó igualmente la predicción de *targets* de para las familias de miRNA encontradas y en algunos casos se contrastaron estas predicciones con validaciones experimentales encontradas en la literatura. (**Anexos C y D**)

En las librerías de arroz las distribuciones de tamaños de lecturas mapeadas contra miRNAs fueron diferentes cuando se consideraban lecturas totales y lecturas únicas. Se encontró básicamente gran abundancia de lecturas de miRNAs de 21 nt y gran diversidad de lecturas de miRNAs de 24nt (**Figura 2-2**). Esto puede estar relacionado con el mayor número de miRNAs de 24nt que han sido identificados en arroz en comparación con otras plantas (**Figura 2-2**).

Con el fin de analizar similitudes entre las distintas librerías, se calculó la correlación entre niveles de expresión de los miRNAs y se calcularon distancias euclidianas a partir de las correlaciones. Esto permitió identificar que de modo general, en cada experimento, los perfiles de expresión de las librerías inoculadas (con bacteria silvestre o mutante) eran más similares entre sí que con las librerías control. Se observó además que se presentaba una gran diferencia entre las librerías del experimento con Xoo PXO99 y del experimento con Xoo BAI3. También se puede observar que los tratamientos del experimento Xoo PXO99 son más disímiles entre sí que los tratamientos del experimento Xoo BAI3 (**Figura 2-3**).

Notablemente, se observó que las librerías correspondientes al mismo tratamiento en distintas réplicas biológicas (p.ej. B3_1 y B3_2) en el experimento Xoo BAI3 no se agrupaban, indicando que los perfiles de expresión de miRNAs en plantas infectadas

fueron diferentes entre réplicas biológicas. Esto puede estar relacionado con los distintos niveles de infección en las dos réplicas inferidos a partir de la cantidad de lecturas correspondientes al genoma de la bacteria obtenido en las librerías (**Tabla 2-1**).

Posteriormente se realizó un análisis del cambio de expresión de cada familia de miRNAs en los distintos tratamientos. Para ello se relacionó en cada experimento la expresión de cada familia de miRNA en las plantas inoculadas con la bacteria silvestre (expresión dependiente de efectores) y la expresión en las plantas inoculadas con la bacteria mutante (expresión dependiente de PTI). En los experimentos con ambas cepas se observa una tendencia general en la que los niveles de inducción de la expresión de los miRNAs con la bacteria silvestre son similares a los niveles de inducción con la bacteria mutante, indicando que la presencia de efectores causa pocas modificaciones en la respuesta mediada por miRNAs en este sistema.

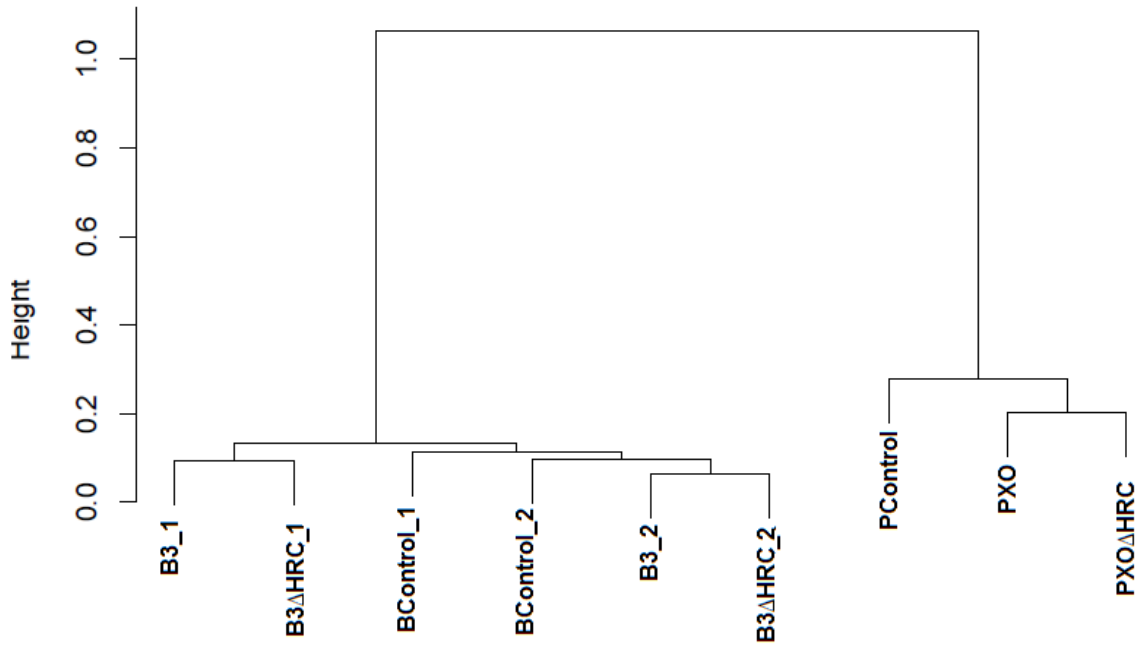
En el experimento con Xoo BAI3, pocas familias de miRNAs presentaron niveles altos de inducción o represión ($|\text{Log}_2\text{FC}| > 2$), y ninguna familia mostró niveles significativos de inducción en las dos réplicas biológicas (**Figura 2-4**). En general, se observaron mayores niveles de inducción o represión en las librerías de la réplica biológica 1, las cuales posiblemente presentaban una mayor población de bacterias (**Figura 2-4**). En consecuencia, se consideraron como diferencialmente expresadas las familias que tuviesen altos niveles de inducción o represión en al menos una de las dos librerías.

Bajo estos criterios se identificaron seis familias diferencialmente expresadas en respuesta a Xoo BAI3. Tres familias fueron inducidas en respuesta a B3 y a B3 Δ HRC1, una familia fue reprimida en respuesta a las dos bacterias y dos familias fueron inducidas en respuesta únicamente a B3 (expresión dependiente de efectores) (**Figura 2-4**). Se intentó validar mediante northern blot la expresión diferencial de estas familias pero no se obtuvieron resultados satisfactorios en ninguno de los casos. Estos resultados se pueden explicar considerando la baja expresión de estas familias de miRNAs (menos de 500 lecturas por millón en todas las librerías). Solo se pudo detectar la expresión de una familia no diferencialmente expresada: miR159a la cual presentó más de 10000 lecturas por millón en todas las librerías (**Figura 2-4**).

Por otro lado, en respuesta a Xoo PXO99A se encontraron 47 familias de miRNAs diferencialmente expresadas en respuesta a la bacteria ($|\text{Log}_2\text{FC}| > 2$). De estas, algunas familias fueron diferencialmente expresadas específicamente en respuesta a efectores o en respuesta a PTI, sin embargo los mayores niveles de inducción o represión se observaron en familias expresadas diferencialmente en ambos tratamientos (22 familias) (**Figura 2-5**). Estos resultados parecen confirmar la idea de que los efectores en la respuesta mediada por miRNAs no tienen un gran efecto.

Figura 2-3: Cluster jerárquico mostrando las relaciones entre tratamientos en librerías de sRNAs a partir de la correlación en perfiles de expresión de miRNAs.

Se calculó la correlación entre perfiles de expresión por el método Spearman. Se empleó la matriz de correlación para calcular distancias euclidianas para luego realizar un cluster jerárquico. La altura (*height*) equivale a las distancias entre tratamientos.



2.2.3 Expresión de *targets* de miRNAs en experimentos de RNA-seq y microarreglos

Para validar la expresión de los *targets* de miRNAs diferencialmente expresados en respuesta a Xoo BAI3, se generaron librerías de RNA-seq construidas a partir de las mismas muestras utilizadas para las librerías de sRNAs. Estas librerías fueron mapeadas contra los genes de arroz y se calcularon valores de expresión normalizados (RPKM) para cada gen. Se utilizaron los perfiles de expresión de RPKM para calcular distancias entre las librerías de la misma forma que se llevo a cabo para las librerías de sRNAs y se encontró un comportamiento similar: diferencias entre ambas réplicas biológicas y cercanía entre muestras inoculadas. Sin embargo en la réplica 1 el perfil de expresión de B3ΔHRC fue más cercano al de las muestras control que a B3. (**Figura 2-6**).

En estas librerías se verificó nuevamente la inducción de *Os11N3*, encontrándose mayor inducción en la réplica 1, lo cual concuerda con los mayores niveles de bacteria inferidos a partir de las librerías de sRNAs y con los mayores niveles de expresión diferencial de miRNAs. Sin embargo en el experimento de Q-PCR se observó el comportamiento contrario para las dos réplicas biológicas (**Figura 2-6**).. Esta discordancia se puede deber al mayor nivel de variación en este experimento (**Figura 2-1**). pero no afecta los resultados generales obtenidos y analizados.

Se estudió la expresión de los *targets* de familias de miRNAs diferencialmente expresadas en cada réplica , esperando encontrar una relación inversa entre los niveles de expresión de miRNAs y los *targets*. Se utilizaron para estoprincipalmente *targets* predichos con altos puntajes por los programas de predicción así como aquellos que tuviesen validación experimental reportada. Para ninguno de los *targets* se encontró expresión diferencial significativa en respuesta a la bacteria (p-value prueba EdgeR >0.1) (**Figura 2-7**). Tampoco se presentaron diferencias significativas entre la expresión de *targets* de miRNAs inducidos y *targets* de miRNA reprimidos (Wilcoxon p-value >0.1).

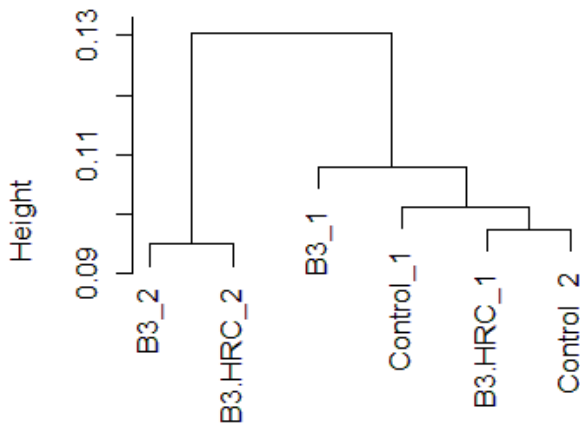
Para determinar la expresión de los *targets* predichos de miRNAs diferencialmente expresados en respuesta a Xoo PXO99A se dispone de datos de microarreglos obtenidos a partir de plantas sometidas a los mismos tratamientos (GSE36272). Se obtuvieron los valores de expresión normalizados para los genes de arroz en este experimento en la base de datos PlexDB [142]. Se verificó en este experimento la inducción de *Os8N3*, blanco del efector TAL PthXo1, en plantas inoculadas con PXO99A enocntrándose que efectivamente era inducido. Posteriormente se analizó el cambio en expresión de los *targets* de los distintos grupos de miRNAs diferencialmente expresados (por ejemplo inducidos por efectores). En general, no se encontró expresión diferencial significativa para ninguno de los *targets* de miRNAs diferencialmente expresados, incluyendo los *targets* de las familias de interés mencionadas en la sección anterior **(Figura 2-8)**..

Estos resultados parecen indicar que los cambios en expresión de miRNAs observados en esta interacción pueden no ser lo suficientemente pronunciados como para causar cambios detectables en los niveles de expresión de genes.

Figura 2-6: Librerías de RNA-seq de arroz inoculado con Xoo BAI3.

a) Cluster jerárquico mostrando las relaciones entre tratamientos en librerías de RNA-seq a partir de la correlación en perfiles de expresión de genes. Se calculó la correlación entre perfiles de expresión por el método Spearman. Se empleó la matriz de correlación para calcular distancias euclidianas para luego realizar un cluster jerárquico. La altura (*height*) equivale a las distancias entre tratamientos. b) Valores de expresión normalizados (RPKM) del *gen Os11N3* en librerías de RNA-seq.

a) Distancias entre librerías



b) Os11N3

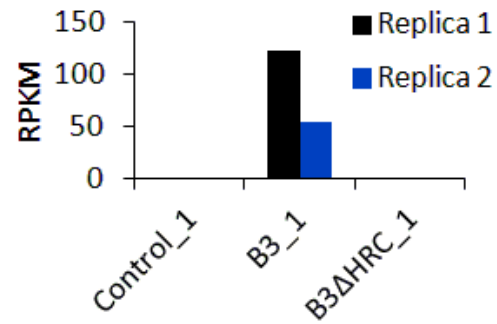
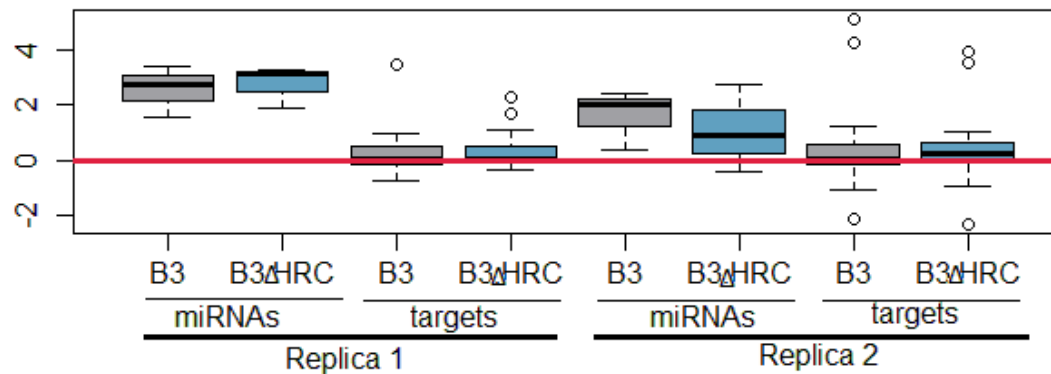


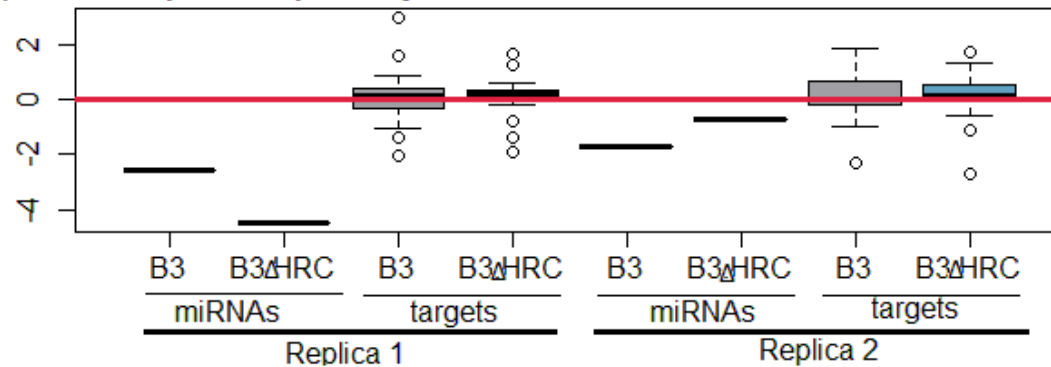
Figura 2-7: Expresión de *targets* de miRNAs diferencialmente expresados en respuesta a Xoo BAI3 en datos de RNA-seq.

a-c) Se muestra la expresión relativa (\log_2FC) respecto al control de los distintos grupos de miRNAs diferencialmente expresados en librerías de sRNAs y de sus *targets* predichos en datos de RNA-seq en cada réplica. La línea roja indica ausencia de expresión diferencial ($\log_2FC = 0$)

a) miRNAs inducidos por PTI y efectores



b) miRNAs reprimidos por PTI y efectores



c) miRNAs inducidos por efectores

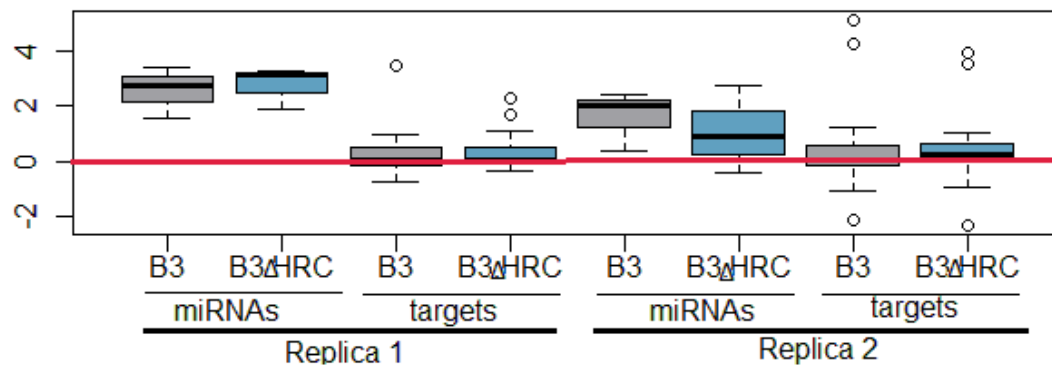
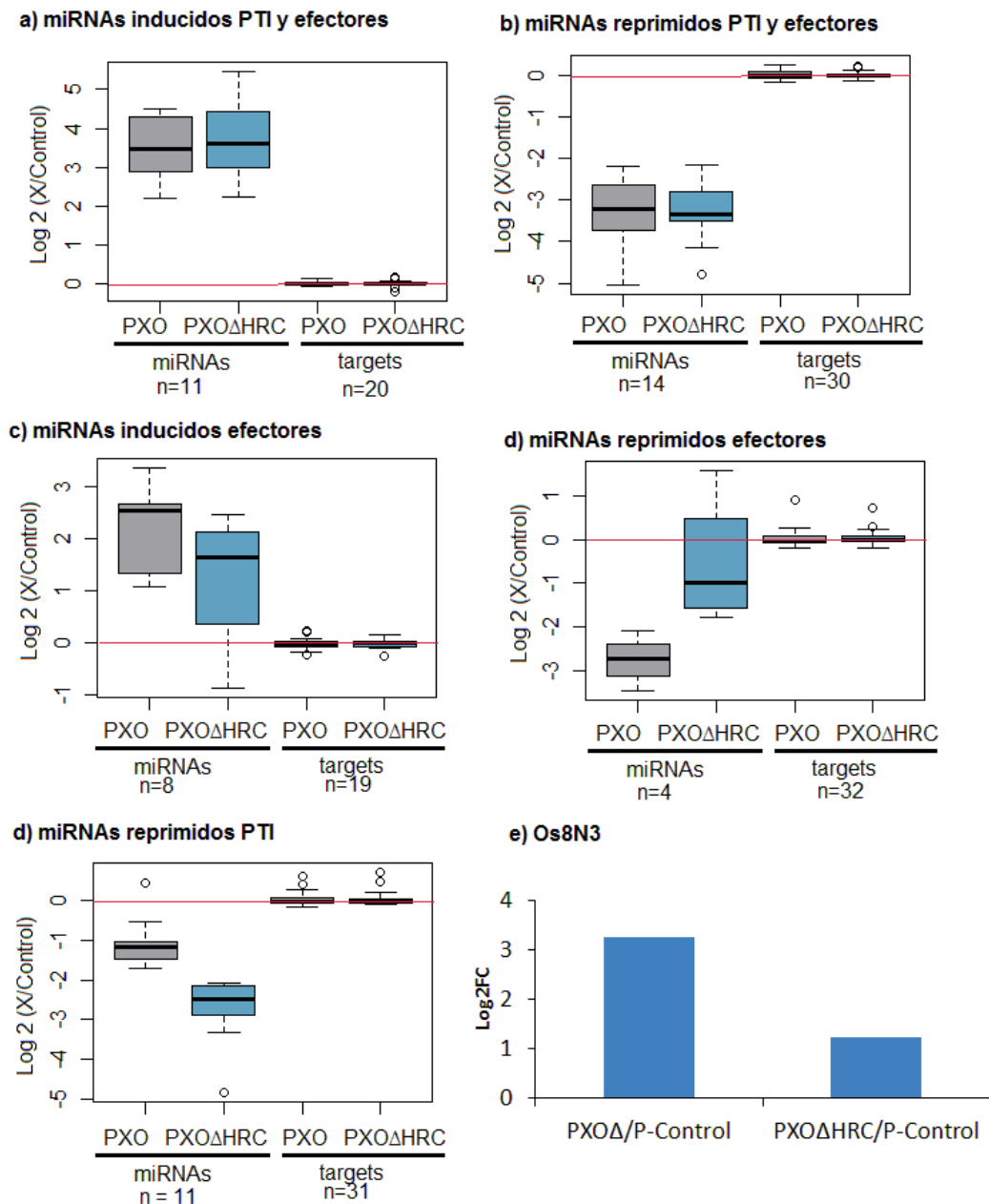


Figura 2-8: Expresión de *targets* de miRNAs diferencialmente expresados en respuesta a Xoo PXO99 en datos de microarreglos.

a-d) Se muestra la expresión relativa (\log_2FC) respecto al control de los distintos grupos de miRNAs diferencialmente expresados en librerías de sRNAs y de sus *targets* predichos en datos de microarreglos. e) Expresión del gen *Os8N3* en datos de microarreglos. La línea roja indica ausencia de expresión diferencial ($\log_2FC = 0$)



2.3 Discusión

En este trabajo se identificaron familias de miRNAs diferencialmente expresadas en arroz en respuesta a dos cepas de *Xanthomonas oryzae* pv. *oryzae*, la cepa Africana Xoo BAI3 y la cepa asiática Xoo PXO99. Se encontraron pocas familias diferencialmente expresadas, especialmente en respuesta a Xoo BAI3 en comparación a lo encontrado en otros trabajos [21, 22]. Y notablemente, no se pudo relacionar la expresión diferencial de los miRNAs con cambios en la expresión de genes *targets* en datos de microarreglos o RNA-seq. Esto se puede deber a que los niveles de expresión diferencial de los miRNAs no sean lo suficientemente pronunciados para causar cambios considerables en la expresión. Es posible que los niveles de bacteria existentes en el tiempo de inoculación observado no sean suficientes para elicitar una respuesta. A partir de los resultados observados en Xoo BAI3 en las dos réplicas técnicas se tienen indicios que, efectivamente los niveles de inducción de miRNAs dependen de la cantidad de bacteria encontrada. En otros trabajos se ha reportado la ausencia de patrones robustos de expresión de genes en interacciones compatibles o de susceptibilidad [143].

A pesar de que no se relacionó la expresión de miRNAs y *targets*, es posible que las familias de miRNAs acá encontradas sean indicadores de mecanismos más robustos que se presenten en otros tiempos tras la infección o con mayores poblaciones de la bacteria, por tanto es importante analizar posibles efectos de las familias de miRNAs encontradas como diferencialmente expresadas. Por un lado, las familias de miRNAs diferencialmente expresadas en arroz fueron diferentes a las reportadas en *Arabidopsis* [86, 87] y a las encontradas en yuca (este trabajo, Capítulo 1). Incluso algunas familias que se sabe son inducidas en respuesta a bacterias en *Arabidopsis* y yuca, tales como miR393, miR394 y miR160 [13, 69], fueron reprimidas en respuesta a Xoo PXO99A (silvestre y mutante).

La mayoría de las familias encontradas diferencialmente expresadas en estos tratamientos corresponden a familias exclusivas de arroz o restringidas a la familia Poaceae, lo cual sugiere un mecanismo especie-específico de regulación de la vía de miRNAs en respuesta a bacterias. Ya se ha señalado la posibilidad de que algunas de

las respuestas de defensa mediadas por miRNAs sean diferentes en ciertos clados [23, 144]. Particularmente, el papel crucial de la familia miR482 en la regulación de genes tipo NBS-LRR en Solanáceas se ha relacionado con la expansión de este tipo de genes y de ésta familia en ese grupo de plantas [23]. Otra evidencia adicional de regulación especie-específica viene del hecho de que la mayoría de las familias diferencialmente expresadas en este sistema correspondían a familias exclusivas de arroz o restringidas al clado de las gramíneas.

Por otro lado, resulta curioso que las respuestas encontradas fueron marcadamente diferentes entre las cepas estudiadas. Muchas más familias se encontraron diferencialmente expresadas en la interacción con la cepa asiática PXO99A que con la cepa africana BAI3, aunque es posible que estas diferencias se deban a factores experimentales puesto que las librerías estudiadas para ambas cepas fueron procesadas de manera distinta, y no se conocen por completo los detalles del manejo experimental de las librerías correspondientes al tratamiento con PXO99A.

Es posible sin embargo, que estas diferencias estén también asociadas a factores biológicos. Particularmente la aparición más reciente de las cepas africanas de Xoo, de forma que es posible que en arroz aún no se hayan desarrollado mecanismos de reconocimiento y defensa para esta cepa y en consecuencia el papel de los miRNAs no sea tan importante. Se sabe que las cepas africanas parecen ser más cercanas genéticamente a las cepas del patovar Xoc [122] y se sabe las respuestas de defensa en arroz son muy diferentes en respuesta a estos dos patovars al punto que, aunque se han encontrado más de 20 genes de resistencia a cepas de Xoo en distintas variedades, no se ha encontrado aún un gen de resistencia a Xoc [1]. Sería interesante realizar un estudio similar empleando la bacteria Xoc, lo cual daría mayor sustento o refutaría esta hipótesis.

Únicamente en dos casos se observó la misma respuesta con las dos cepas: las familias miR1427 y miR1320 fueron inducidas en respuesta a bacterias silvestres y mutantes en ambos experimentos. La familia miR1320 presenta entre sus *targets* predichos proteínas tipo *remorines* y *clathrines*, las cuales están asociadas a membranas e implicadas en tráfico vesicular, respectivamente [134]. Esta familia de miRNAs se ha encontrado que es

reprimida en respuesta a estrés por frío, sin embargo se desconoce qué impacto tiene su regulación sobre procesos fisiológicos [134]. Por otro lado, la familia miR1427 tiene como único *target* predicho un gen codificante para una proteína similar a harpin-induced protein 1 (*HIN1*, *LOC_Os02g16030*). Las proteínas de tipo HIN1 se encuentran directamente implicadas en defensa contra patógenos, se sabe que son inducidas por las proteínas Harpin de algunas bacterias y que en respuesta desencadenan una respuesta hipersensible asociada a ETI [145]. Genes de este tipo se han visto inducidos en arroz en respuesta a Xoo [146, 147], lo cual contradice el comportamiento esperado dada la inducción de este miRNA. Sin embargo la expresión de este locus particular aun no ha sido estudiada y no se encontró expresado en este trabajo ni en experimentos de microarreglos ni RNA-seq.

Otras familias de interés diferencialmente expresadas en respuesta a XooPXO99 incluyen miR168 (inducida en PXO y PXO Δ HRC) la cual se sabe está implicada en la regulación misma de la vía miRNAs mediante su acción sobre proteínas de la familia AGO y se sabe está involucrada en respuesta de defensas contra virus [148, 149]. También la familia miR1432 se vio inducida en respuesta a la bacteria silvestre y la mutante. Esta familia tiene entre sus *targets* principalmente genes codificantes para proteínas con dominios EF-hand que comúnmente actúan como sensores de Ca⁺ y se sabe están implicadas en interacción con microorganismos benéficos [150]. Además algunos factores de transcripción dependientes de estas proteínas actúan como reguladores negativos de defensa contra bacterias [151, 152]. Curiosamente, es posible que este tipo de proteínas, específicamente proteínas tipo Calmodulina, sean blancos preferidos de efectores TAL de *Xanthomonas* en arroz [Pérez-Quintero *et al*, sometido] (Sección 0).

Finalmente, otro resultado sorprendente en este trabajo fue la similitud entre las respuestas asociadas a PTI y las respuestas asociadas a ETI en terminos de expresión diferencial de miRNAs. Para ninguna de las cepas se pudo establecer un efecto claro de los efectores sobre la expresión de miRNAs. Esto sugiere nuevamente que las respuestas de miRNAs pueden no ser importantes durante la inmunidad en arroz. Particularmente, no se encontraron evidencias que sugieran un papel de supresión de

silenciamiento de miRNAs para los efectores de estas bacterias como se ha encontrado en otros sistemas [31]. Se encontraron sin embargo, algunas familias que parecen ser inducidas específicamente en respuesta a efectores; la posibilidad de que estas sean reguladas por efectores TAL es explorada en la sección 4.2.3.

En conclusión la respuesta a bacterias en arroz mediada por miRNAs parece no atenerse al modelo general propuesto a partir de otras plantas (**Figura 0-1**). Esto puede significar que en arroz la regulación mediada por miRNAs no es crucial para la defensa y es posible que ésta esté mediada principalmente por las interacciones tradicionalmente estudiadas como interacción gen-por-gen entre efectores y genes de resistencia [116]. Sin embargo es importante ampliar estos resultados estudiando otros tiempo de inoculación o estudiando respuestas que puedan estar mediadas por otros tipos de RNAs pequeños.

2.4 Materiales y métodos

2.4.1 Materiales vegetales e inoculación

Para el experimento con la cepa Xoo BAI3, plantas de arroz de la variedad Nipponbare fueron cultivadas a partir de semillas y mantenidas en invernadero a 26°C y con humedad relativa del 80%. Plantas de tres semanas fueron inoculadas con suspensiones bacterianas de cultivos de 48 horas de las cepas Xoo BAI3 y Xoo BAI3 Δ HRC, descritas en [123], a una OD_{600nm} = 0.5. La inoculación se realizó por infiltración utilizando una jeringa sin punta, a ~ 8 cm del ápice de la hoja más joven. Para cada réplica biológica se inocularon al menos 8 plantas. Se comprobó el desarrollo de los síntomas esperados en plantas inoculadas pero no recolectadas en cada réplica 7dpi.

2.4.2 Construcción y secuenciamiento de librerías de sRNAs y RNA-seq.

La extracción de RNA se llevó a cabo empleando el método de Trizol (Invitrogen) a partir de tejidos de arroz infiltrados con bacteria o con agua a 24hpi. Muestras de RNA correspondientes a dos réplicas biológicas fueron utilizadas para secuenciación de librerías de RNAs pequeños, secuenciación de RNA mensajero y experimentos de Q-PCR para verificar la expresión del gen control *Os11N3*. Previo a la secuenciación se confirmó la calidad del RNA extraído utilizando un 2100 Bioanalyzer (Agilent technologies). Las librerías de sRNAs y RNA-seq fueron construidas y secuenciadas por la empresa Fasteris (<http://www.fasteris.com/>).

Los datos correspondientes al experimento Xoo PXO99 no fueron desarrollados en este trabajo, se obtuvieron a partir de bases de datos disponibles. Librerías de sRNAs: (http://mpss.udel.edu/rice_sRNA/) y datos de microarreglos: (<http://www.plexdb.org/>, Exp OS66).

2.4.3 Análisis de librerías de sRNAs.

Se verificó la calidad de las librerías (Experimento Xoo BAI) utilizando el paquete FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), encontrándose la mayoría de secuencias con puntajes Phred promedio mayores a 30. Las secuencias se encontraban ya sin adaptadores.

Las secuencias de ambos experimentos fueron mapeadas contra el genoma de arroz (*O. sativa* cv Nipponbare MSU7) obtenido de phytozome (<http://www.phytozome.net/>). El genoma de Xoo PXO99A (NC_010717.1) y el set de snoRNAs, tRNAs y rRNAs conocidos fueron obtenidos de la base de datos Rfam [99]. Estos mapeos fueron realizado con el programa bowtie (*parámetros l= 15, n= 2*) [153]. Se eligió bowtie por mayor rapidez, sin embargo los alineamientos mostraron ser similares a los de blast en alienamientos de pruebas hechos con subsets de las librerías.

Para el cálculo de la expresión de miRNAs conservados y la predicción de *targets* se utilizaron procedimientos similares a los descritos en la sección 1.2.2. Principalmente se obtuvo el perfil de expresión de miRNAs de arroz al mapearlos contra el set de miRNAs descritos para esta especie en miRBase version 17 [66]. Se obtuvieron estos perfiles con la herramienta miRprof del kit UEA sRNA [72] y con blastn [74] (*parámetros e-value<0.0001, ungapped, word size = 4, -max_target_seqs 1 -best_hit_score_edge 0.05 -best_hit_overhang 0.25*) (Nota= bowtie no fue adecuado para mapeos de secuencias cortas contra secuencias cortas por eso se utilizó blast en este caso). Se reportan los resultados obtenidos con miRprof como total de lecturas mapeadas miRNAs/total lecturas de librerías (mapeadas contra genoma de arroz y libres de tRNAs y rRNAs). La predicción de *targets* se llevo a cabo utilizando psRNA*target* (parámetros por defecto), contra los genes anotados en la versión MSU7 del genoma de arroz [80].

2.4.4 Análisis de librerías de RNA-seq

Tras comprobar la calidad de las librerías estas fueron mapeadas contra los genes anotados en el genoma de arroz utilizando seqmap [105] y se calcularon valores de

expresión normalizados (RPKM) para cada gen utilizando Rseq, con parámetros para lecturas pareadas [106]. Los RPKM fueron utilizados para analizar el cambio en expresión de los *targets* de miRNAs diferencialmente expresados. Adicionalmente para identificar los genes con valores de expresión diferencial más significativos en ambas réplicas biológicas se utilizó el paquete de R EdgeR [154]. Este programa utiliza un modelo negativo binomial y una prueba similar al test exacto de Fisher para normalizar los valores de expresión e identificar los genes diferencialmente expresados.

2.4.5 Northern blot

Se estudió la expresión de miRNAs utilizando un protocolo de northern blot adecuado para la expresión de RNAs pequeños similar al descrito en [155]. Brevemente, se extrajo RNA a partir de hojas inoculadas utilizando el kit de purificación de RNAs pequeños miRNeasy mini kit (Qiagen). Para la hibridación se utilizaron 20ug de RNA total separados en un gel de poliacrilamida 15% y transferidos a una membrana Zeta-Probe GT (Biorad) por capilaridad. Las membranas fueron fijadas bajo luz UV y se realizó la hibridación durante una noche a 40°C con sondas de oligonucleótidos marcadas con ³²P complementarias a las secuencias de miRNA. Se utilizó como control una sonda marcada de 31nt complementaria a U6 snRNA.

2.4.6 Q-PCR

El PCR cuantitativo se realizó con un equipo Mx3005P (Stratagene) y el kit QPCR MESA Green (EUROGENETEC) de acuerdo a las especificaciones del fabricante. Las condiciones de qPCR luego de estandarización fueron 15 min 95°C, 15 sec 95°C, 30 sec 61°C 40x y 30 sec 72°C. Se verificó la ausencia de DNA genómico en las muestras comparando con controles de RNA y se verificó la especificidad de los primers utilizados analizando la curva de disociación. Se utilizó de referencia el gen correspondiente a la *actina* en arroz. Los primers utilizados para *Os11N3* fueron F= CTACCTGGCCCCACTGC y R= GTGCGCACCACCAGCC, y para *actina* F = GCGTGGACAAAGTTTTCAACCG y R = TCTGGTACCCTCATCAGGCAT.

3.Regulación transcripcional de miRNAs por factores de transcripción endógenos

3.1 Introducción: Regiones promotoras de genes *MIRNA*.

En eucariotas la regulación transcripcional está regulada fundamentalmente por el reclutamiento de factores de transcripción (TFs) a elementos reguladores *cis*. Los TFs interactúan con elementos específicos del DNA, con otros TFs, y con la maquinaria de transcripción para regular la expresión de los genes a los que se unen. En plantas existen más de 1500 TFs, cada uno controlando la expresión de decenas o miles de genes en complejas redes de señalización. Los elementos *cis* o sitios de unión de factores de transcripción (TFBS) son elementos de DNA funcional que influyen la actividad transcripcional [156, 157]. En algunas ocasiones múltiples TFBS constituyen un modulo de regulación *cis* que integra las señales de múltiples TFs resultando en control combinatorial y en patrones altamente específicos de expresión de genes. La conservación de estos elementos de DNA en distintas plantas hace relativamente fácil su identificación y la inferencia de patrones de regulación de genes a partir de la presencia o combinación de estos [156, 157]

Los genes *MIRNA* son transcritos por la RNA polimerasa II al igual que la mayoría de los genes codificantes para proteínas, y por tanto están sometidos al mismo control por factores de transcripción. En la mayoría de los genes de *MIRNA* analizados se han encontrado elementos comunes de promotores tipo Pol-II como la caja TATA en proporciones similares a genes codificantes [38]. Se sabe además que la región promotora de los genes *MIRNA* se encuentra comúnmente a menos de 1000pb curso arriba del inicio del pre-miRNA, especialmente en el caso de miRNAs de regiones intergénicas, no policistrónicas [37, 38, 158]. Esto es de vital importancia para la predicción de TFBS puesto que en muchos casos el sitio de inicio de la transcripción (TSS) no está caracterizado para miRNAs.

Varios trabajos han caracterizado TFBS en genes *MIRNA* en varias plantas. Se ha encontrado que algunos TFBS parecen estar sobre-representados en genes *MIRNA* con respecto a genes codificantes para proteínas, por ejemplo los motivos AtMYC2, ARF, SORLREP3, y LFY así como la caja TATA fueron encontrados sobre-representados en genes *MIRNA* en *Arabidopsis* [37]. Se han encontrado también diferencias en frecuencias de TFBS entre distintas especies de plantas [38], y entre microRNAs conservados y especie-específicos [158]. De igual manera se han encontrado TFBS sobre-representados en grupos de miRNAs involucrados en varios procesos incluyendo respuestas a estrés oxidativo [159], UV-B [160, 161], salinidad [54, 162], sequía [54, 163], baja temperatura [164], así como en procesos de organogénesis y desarrollo [165].

Esto resalta la importancia del estudio de los procesos de regulación transcripcional de genes *MIRNA* para entender su papel en procesos fisiológicos. Además el que gran parte de miRNAs tengan preferencialmente entre sus *targets* factores de transcripción [166] señala un mecanismo de regulación recíproca entre factores de transcripción y miRNAs que se considera parte crucial de las redes de regulación génica en plantas [167].

En este capítulo se buscó entender las diferencias de regulación de genes *MIRNA* observadas entre yuca y arroz, haciendo un análisis extensivo de las frecuencias de TFBS en distintas plantas, y por otro lado se buscó identificar TFBS que fuesen determinantes para la regulación transcripcional de miRNAs en respuesta a bacterias.

3.2 Resultados

3.2.1 Identificación de sitios de unión de factores de transcripción (TFBS) en promotores de genes *MIRNA* en plantas

Para tener una perspectiva amplia de los posibles mecanismos de regulación asociados a miRNAs, se analizaron las regiones promotoras de los pre-miRNAs (1000 nucleótidos *upstream* del extremo 5' del pre-miRNA) identificados en yuca y arroz (abreviadas respectivamente *mes* y *osa* en estos análisis), y en otras 15 especies de plantas para las que se cuenta con un genoma secuenciado y al menos 50 pre-miRNAs depositados en miRBase [66]: *Arabidopsis lyrata* (*aly*), *Arabidopsis thaliana* (*ath*), *Brachypodium distachyon* (*bdi*), *Chlamydomonas reinhardtii* (*cre*), *Citrus sinensis* (*csi*), *Glycine max* (*gma*), *Malus domestica* (*mdm*), *Medicago truncatula* (*mtr*), *Physcomitrella patens* (*ppt*), *Populus trichocarpa* (*ptc*), *Ricinus communis* (*rco*), *Sorghum bicolor* (*sbi*), *Selaginella moellendorffii* (*smo*) *Vitis vinifera* (*vvi*) y *Zea mays* (*zma*). Adicionalmente se analizaron también las regiones promotoras de genes codificantes para proteínas en estas 17 especies.

Se utilizó el programa PlantTFBS [104] para identificar sitios de unión de factores de transcripción en las regiones promotoras de miRNAs y genes en estas especies. En total se escanearon 3812 regiones promotoras correspondientes a miRNAs, y 596701 correspondientes a genes. Se escanearon 99 sitios de unión para factores de transcripción (TFBS), utilizando las matrices de peso posicional descritas en [104], las cuales están basadas principalmente en datos de *Arabidopsis* obtenidos de las bases de datos AGRIS [168], PlantCare [169] y AtProbe (<http://exon.cshl.org/cgi-bin/atprobe/atprobe.pl>). Y para verificar que los sitios de unión identificados por el programa no se deban simplemente al azar se generaron sets de secuencias de miRNAs y genes aleatorias manteniendo la distribución de nucleótidos de las secuencias originales.

Se calculó la frecuencia relativa de promotores que contenían al menos un sitio de unión para cada factor de transcripción en cada grupo de genes y se evaluó la distribución de estas frecuencias. Se encontró que la distribución era significativamente diferente entre promotores de miRNAs aleatorizados y no aleatorizados (Kolmogorov-Smirnov p-value = 6.44E-05) y entre promotores de genes aleatorizados y no aleatorizados (Kolmogorov-Smirnov p-value = 1.56E-10), indicando que efectivamente los TFBS identificados no se dan únicamente por azar. Sin embargo, algunos TFBS se encontraron en frecuencias muy altas (>50%) en las secuencias aleatorizadas, estos correspondían principalmente a sitios de unión cortos (5-8 nt) o de baja complejidad como las cajas TATA, GATA y CAAT **(Anexo E)**.

3.2.2 Diferencias en frecuencias de TFBS entre tipos de genes y entre plantas

Se procedió luego a comparar los resultados obtenidos para TFBS en promotores de miRNAs con los promotores de genes. Primero, para verificar que las diferencias encontradas no estuvieran determinadas por la distribución de nucleótidos en los promotores, se analizó esta distribución para cada especie en los promotores de miRNAs, en genes, y en todo el genoma. Para ninguna de las especies se encontraron diferencias significativas entre la distribución de nucleótidos entre los promotores de miRNAs y los promotores de genes (Kolmogorov-Smirnov p-value > 0.05), y en muchos casos la distribución de nucleótidos en promotores reflejaba la distribución de nucleótidos de los genomas **(Figura 3-1)**. Además la distribución de nucleótidos se encontró que variaba mucho entre plantas siendo similar entre plantas cercanas filogenéticamente. De estos resultados se puede esperar que, de existir diferencias en TFBS de miRNAs y genes estas se deberán a posiblemente procesos de selección y no a diferencias en distribución de nucleótidos, y que se espera encontrar similitudes en TFBS en plantas filogenéticamente cercanas.

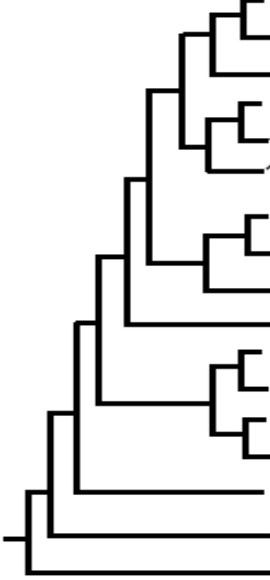
Se evaluó la distribución de frecuencias TFBS entre promotores de miRNAs y de genes, para 9 de 17 plantas se encontraron diferencias significativas en esta comparación **(Tabla 3-1)**. Notablemente no se encontraron diferencias en la distribución de frecuencias

de TFBS entre promotores de miRNAs y genes para dos grupos de plantas: Gramíneas (*B. dystachyon*, *O. sativa*, *S. bicolor*, *Z. mays*) y plantas fijadoras de nitrógeno (*G. max*, *M. domestica*, *M. truncatula*). Es posible que estos resultados estén influenciados por el número de genes *MIRNA* descritos en las plantas pues entre las plantas en las que no se encontraron diferencias significativas se encontraban cinco plantas con más de 200 genes *MIRNA* descritos (**Tabla 3-1**).

Se hicieron pruebas de Fisher comparando la frecuencia de cada TFBS en promotores de miRNA contra promotores de genes codificantes para buscar TFBS sobre-representados en alguno de los dos grupos, en las plantas de interés en este trabajo (yuca y arroz) y en la planta modelo *A. thaliana*. En yuca se encontraron 5 TFBS sobre-representados en promotores de miRNAs, en Arabidopsis 2, y en arroz uno (**Tabla 3-2**). Ningún factor fue común en las tres plantas.

Finalmente, para saber si las diferencias entre TFBS en promotores de miRNAs de distintas plantas estaban relacionadas con la filogenia, se utilizaron las frecuencias de TFBS de cada planta para calcular una matriz de correlación, los valores de correlación se convirtieron en distancias entre las distintas plantas y se realizó un clúster jerárquico con estas distancias (**Figura 3-2**). Se puede ver que las frecuencias de TFBS reflejan parcialmente relaciones filogenéticas. algunas especies perteneciente a las mismas familias (Gramineas y Fabaceas) se agrupan en clústers, lo cual sugiere que nuevamente existen patrones de regulación transcripcional similares entre plantas filogenéticamente cercanas.

Tabla 3-1: Diferencias en distribución de TFBS entre genes MIRNA y genes codificantes.

Arbol filogenetico	Especie	P-value	Genes <i>MIRNA</i>
	<i>M. esculenta (mes)</i>	0.00*	127
	<i>R. comunis (rco)</i>	0.00*	63
	<i>P. trichocarpa (ptc)</i>	0.00*	318
	<i>M. truncatula (mtr)</i>	0.69	669
	<i>G. max (gma)</i>	0.46	503
	<i>M. domestica (mdo)</i>	0.11	178
	<i>A. thaliana (ath)</i>	0.00*	299
	<i>A. lyrata (aly)</i>	0.02*	198
	<i>C. sinensis (csi)</i>	0.00*	53
	<i>V. vinifera (vvi)</i>	0.01*	163
	<i>S. bicolor (sbi)</i>	0.21	206
	<i>Z. mays (zma)</i>	0.21	170
	<i>O. sativa (osa)</i>	0.69	589
	<i>B. dystachion (bdi)</i>	0.21	135
	<i>S. moellendorffii (smo)</i>	0.00*	57
	<i>P. patens (ppt)</i>	0.15	228
	<i>C. reinhardtii (cre)</i>	0.00*	49

El p-value corresponde al resultado de una prueba Kolmogorov-Smirnov entre la distribución de TFBS de genes *MIRNA* y genes codificantes de proteínas. Se muestra el número total de genes *MIRNA* analizados en cada especie. El árbol filogenético se obtuvo de www.phytozome.net. * p-value <0.05.

Tabla 3-2: TFBS sobrerrepresentados en genes MIRNA comparados con genes codificantes para proteínas en yuca, arroz y *Arabidopsis*.

TFBS	p-value	% MIRNA	% Genes codificantes
<i>M. esculenta</i>			
ATHB2 binding site motif	0.003	8.73	0.00
AtMYC2 BS in RD22	0.000	0.00	16.17
Ibox promoter motif	0.006	34.92	16.82
RAV1-B binding site motif	0.007	0.00	8.36
SORLIP1	0.001	0.00	11.15
SORLREP3	0.007	7.94	0.00
T-box promoter motif	0.000	30.95	0.00
W-box promoter motif	0.001	32.54	13.38
<i>O. sativa</i>			
LFY consensus binding site motif	0.001	42.69	20.24
<i>A. thaliana</i>			
ATB2/AtbZIP53/AtbZIP44/GBF5 BS in ProDH	0.000	22.48	0.00
AtMYB2 BS in RD22	0.007	7.72	0.00
BoxII promoter motif	0.000	0.00	24.25

El p-value corresponde al resultado de una prueba exacta de Fisher entre la frecuencia de cada TFBS en genes *MIRNA* comparado con genes codificantes de proteínas para cada planta.

Figura 3-1: Frecuencia de nucleótidos en regiones de promotores de genes de plantas. Se muestra la frecuencia de nucleótidos en las regiones promotoras (1kb *upstream*) de **a)** genes *MIRNA* y **b)** genes codificantes para proteínas. Se muestra también la frecuencia de nucleótidos en **c)** todo el genoma y se señalan las **d)** relaciones filogenéticas entre las plantas analizadas.

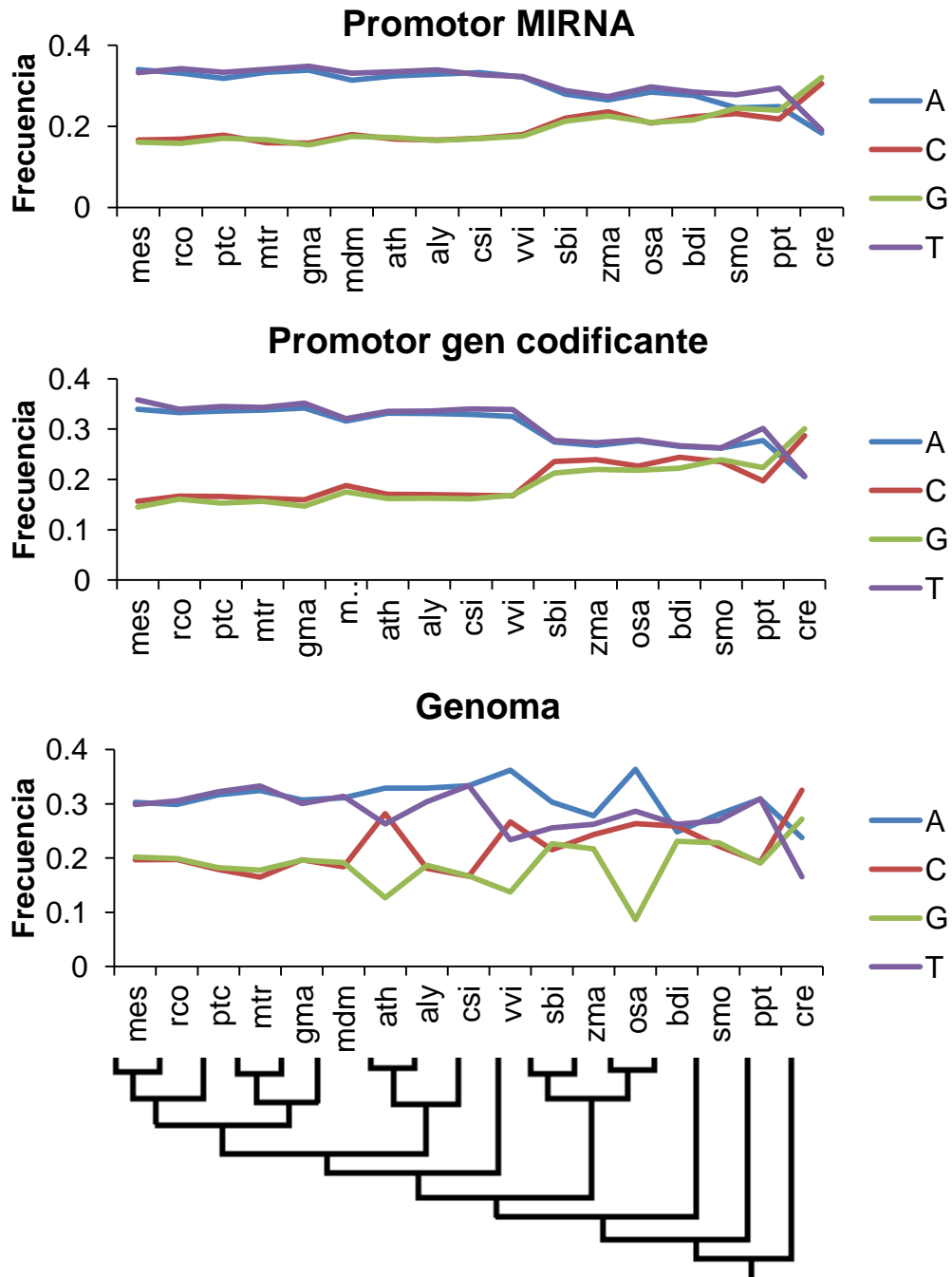
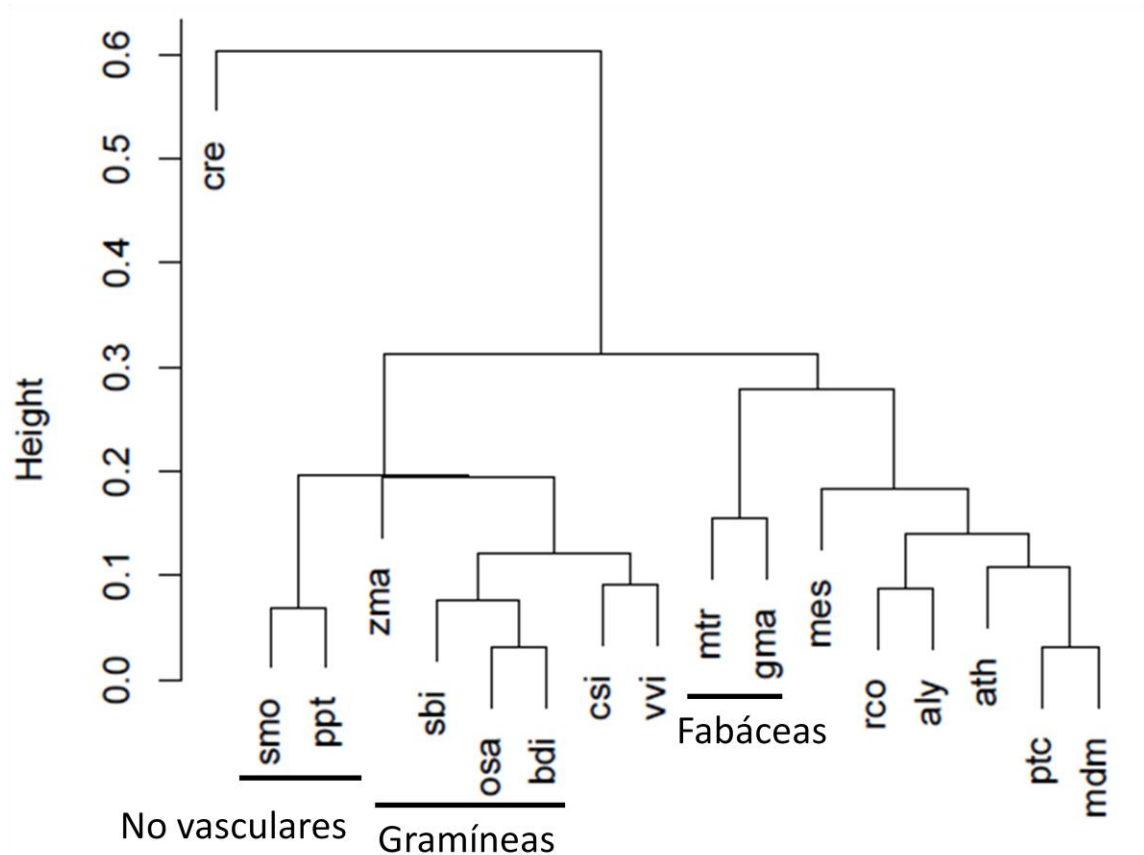


Figura 3-2: Cluster jerárquico basado en la correlación de distribuciones de frecuencias de TFBS en promotores de genes *MIRNA* en distintas plantas.

Se calculó la correlación entre distribuciones utilizando el método pearson. Se empleó la matriz de correlación para calcular distancias euclidianas para luego realizar un cluster jerárquico. La altura (*height*) equivale a las distancias entre plantas.



3.2.3 TFBS sobre-representados en familias involucradas en respuesta a bacterias en arroz y yuca.

Para buscar TFBS relacionados con las respuestas a bacterias se hicieron pruebas de Fisher entre las frecuencias de los distintos TFBS en promotores de genes *MIRNA* de familias inducidas en respuesta a la bacteria (para arroz se diferenciaron los grupos que respondían a PTI, Efectores o ambos) y se comparó con las frecuencias en familias no inducidas (incluyendo reprimidas). El proceso se realizó también para las familias reprimidas.

En yuca se encontraron seis TFBS sobre-representados en familias inducidas y ninguno en familias reprimidas (**Tabla 3-3**). En arroz, en el experimento con Xoo BAI3, no se encontraron elementos sobre-representados en ningún grupo, probablemente porque el número de familias en cualquier grupo (p.ej. inducidas por efectores) nunca fue mayor a 3. Por otro lado en el experimento con Xoo PXO99A se lograron identificar elementos sobre-representados en familias que eran inducidas o reprimidas de igual manera en PTI y con efectores pero no en familias inducidas o reprimidas en un tratamiento específico (**Tabla 3-3**).

Tabla 3-3: TFBS sobre-representados en promotores de familias de miRNAs diferencialmente expresados en respuesta a *Xanthomonas* en yuca y arroz.

Yuca-Xam			
miRNAs inducidos			
TFBS	p-value	% Inducidos	% No-inducidos
LFY consensus binding site motif	1.10E-17	51	44.7
T-box promoter motif	5.90E-11	33.3	28.9
Box II promoter motif	1.10E-06	27.8	19.7
CCA1 binding site motif	3.20E-03	13.7	6.6
SORLREP3	6.80E-03	11.8	5.3
SORLIP5	2.90E-02	9.8	2.6
Arroz-XooPXO99A			
miRNAs inducidos			
TFBS	p-value	% Inducidos	% No-inducidos
CArG promoter motif	1.64E-08	23.1	8.9
DRE-like promoter motif	2.11E-12	23.1	6.2
SORLREP1	2.54E-04	7.7	0.6
miRNAs reprimidos			
TFBS	p-value	% Reprimidos	% No-reprimidos
BoxII promoter motif	1.11E-34	34.3	17.6
MYB1 binding site motif	2.69E-06	8.6	2.9

3.3 Discusión

En este trabajo se analizó la presencia de TFBS en regiones promotoras de genes *MIRNA*. Se encontró que en muchas plantas no existen diferencias significativas en la distribución de TFBS en genes *MIRNA* comparado con genes codificantes para proteínas, especialmente en plantas para las que un gran número de miRNAs ha sido identificado (p.ej. arroz y *Medicago*). Curiosamente, en este trabajo no se identificaron los mismos TFBS de *Arabidopsis* que se encontraron sobre-representados en genes *MIRNA* respecto a genes codificantes obtenido en un trabajo previo utilizando una estrategia similar [37]. Una posible razón para explicar estas diferencias es considerando el hecho de que actualmente existen muchos más genes *MIRNA* identificados en *Arabidopsis* (299 genes) respecto a los utilizados en Megraw *et al.* (59 genes) [37]. De esta manera se resalta la importancia del número de loci estudiados para estos estudios. Desafortunadamente se desconoce si la identificación de estos genes en algunas plantas ha alcanzado el punto de saturación.

Por otro lado se encontró que tanto la distribución de nucleótidos y la distribución de TFBS en regiones promotoras de genes *MIRNA* es muy variable, siendo más similar entre plantas filogenéticamente relacionadas. Esto concuerda con patrones de regulación de miRNAs especie-específicos comúnmente encontrados al comparar por ejemplo arroz y *Arabidopsis* [38] y con la multiplicidad de funciones observadas para algunas familias de miRNAs en distintas plantas [35, 39]. Así mismo se resalta la importancia de estudiar plantas de distintos grupos taxonómicos para trabajos como estos en que se busca identificar la función de miRNAs en distintos procesos. No es sorprendente por tanto, que los miRNAs acá identificados como importantes para la interacción con *Xanthomonas* sean diferentes en arroz y yuca. De igual forma estos resultados explican también que las respuestas a la bacteria en yuca sean similares a las reportadas en *Arabidopsis* [21, 22] estando estas plantas más cercanas filogenéticamente.

Entre los TFBS sobre-representados en yuca se encontró el de unión para factores de transcripción LFY, el cual está involucrado principalmente en procesos de desarrollo floral

mediante la activación de genes tipo *AGAMOUS* (*AG*), *APETALA* (1, 2 y 3) [170]. Este motivo fue abundante en miembros de la familia 172, la cual es inducida en respuesta a Xam y tienen entre sus *targets* principalmente genes tipo *AP2* involucrados en desarrollo floral [171-173]. Esto sugiere la presencia de un bucle de regulación entre factores de transcripción y miRNAs en desarrollo floral, y es posible que en respuesta a bacterias se favorezca la acción de los miRNAs para frenar el desarrollo floral. La inducción de esta familia se discute también en la sección 4.2.3.

Se encontraron también sobre-representados elementos relacionados con la supresión del ciclo circadiano y las respuestas a luz: CCA1, T-box y SORLREP3 [174, 175]. Se sabe que existe una estrecha conexión entre procesos de defensa, ritmos circadianos y respuesta a luz [175]: Por ejemplo procesos normalmente controlados por el reloj circadiano como la apertura de estomas o por la luz como la síntesis de ácido salicílico son modificados durante los procesos de defensa [175]. No es sorprendente por tanto que la inducción de miRNAs pueda estar regulada por factores similares [176]. Por otro lado, se esperaba encontrar sobre-representación de motivos de unión de factores de respuesta auxinas (ARF), sin embargo estos fueron abundantes en la familia miR160 pero no en otras familias inducidas. Se sabe que las familias involucradas en señalización de auxinas normalmente participan en bucles regulatorios en los que los miRNAs son a la vez regulados por sus *targets*. En yuca esto parece ser cierto al menos para la familia miR160 [37, 177].

En arroz se encontraron sobre-representados en las familias de miRNAs en general motivos diferentes a los que se encontraron en yuca, lo cual afianza las diferencias en los mecanismos de transcripción que se infieren a partir del análisis de los promotores y de las librerías de miRNAs. Sin embargo, los elementos que se encontraron sobre-representados parecen estar involucrados en los mismos procesos. El motivo CarG que se encontró sobre-representado en miRNAs inducidos en respuesta a PXO99A se sabe está implicado en el control de la floración al igual que LFY [178] y el motivo SORLREP1, también encontrado en familias inducidas, es otro motivo implicado en respuestas a luz [179].

Curiosamente, el motivo Box II que fue encontrado sobre-representado en familias de miRNAs inducidas en yuca, aparece sobre-representado en familias reprimidas en arroz. Es posible que este fenómeno esté relacionado con el comportamiento inverso que se observó para algunas familias entre las dos plantas, por ejemplo miR160 y miR393 (inducidas en yuca, reprimidas en arroz). Este motivo también se ha encontrado que está involucrado en respuestas a luz [180].

En conclusión las diferentes respuestas a bacterias encontradas en yuca y arroz pueden estar relacionadas con diferencias en distribuciones de TFBS, sin embargo existen los elementos sobre-representados en familias diferencialmente expresadas en respuesta a bacterias en ambas plantas que parecen estar implicados en procesos similares.

Aunque estos resultados dan muchas luces sobre los procesos de transcripción de miRNAs, la aproximación utilizada es relativamente simplista pues no se considera patrones complejos de interacción entre distintos factores de transcripción o la posibilidad de transcripción no-independiente en miRNAs policistronicos. De igual manera, se está asumiendo en este trabajo que los sitios de unión y los factores de transcripción analizados son conservados entre plantas (los modelos acá utilizados provienen principalmente de *Arabidopsis*), lo cual puede no ser cierto. Estos aspectos aun no se pueden corregir en un análisis masivo de regulación entre distintas especies principalmente por ausencia de suficiente información para modelarlos adecuadamente.

Finalmente, una dificultad añadida en este análisis, particularmente para la identificación de elementos sobre-representados en grupos funcionales es la imposibilidad de diferenciar la expresión de loci de *MIRNA* que son regulados independientemente pero que tienen la misma secuencia en los miRNAs maduros. Desafortunadamente, la expresión de los pre-miRNAs, los cuales son diferentes para todos los loci, es transitoria y difícil de cuantificar en este tipo de experimentos [181].

3.4 Materiales y Métodos

3.4.1 Predicción de TFBS

Para hacer la identificación de TFBS primero se extrajeron las regiones promotoras de genes *MIRNA* y genes codificantes para proteínas en distintas plantas. Para esto se extrajeron las secuencias de los pre-miRNAs de las plantas *aly*, *at*, *bdi*, *cre*, *csi*, *gma*, *mdm*, *mtr*, *ppt*, *ptc*, *rco*, *sbi*, *smo*, *vvi*, *gma*, *mdm*, *mtr*, *osa*, *ppt*, *ptc*, *rco*, *sbi*, *smo*, *vvi* y *zma* de miRBase [66], para *mes* (yuca) se utilizaron los pre-miRNAs identificados en este trabajo. Los pre-miRNAs se mapearon contra los genomas respectivos utilizando Blastn, obtenidos de <http://www.phytozome.net/> v. 8.0, (*parametros :e-value<0.0001, ungapped, word size = 4, -max target seqs 1 -best hit score edge 0.05 -best hit overhang 0.25*). Se extrajo la región 1000 bases curso arriba del inicio (5') del pre-miRNA utilizando fastacmd. Se conservaron solo las secuencias de promotores de más de 500nt de longitud puesto que en algunos genomas esta region se encontraba incompleta.

Las regiones promotoras de los genes anotados se extrajeron utilizando la herramienta biomart de <http://www.phytozome.net/>. Para el control negativo las secuencias de promotores de genes *MIRNA* y de genes codificantes fueron aleatorizadas utilizando shuffleseq de la suite EMBOSS [98].

Para la identificación de los TFBS se utilizaron las PMWs y los scripts del programa PlantTFBS detallado en [104]. Las secuencias correspondientes a las PMWs de los 99 factores de transcripción fueron confirmadas en la última versión de Agris [168]. PlantTFBS utiliza una función de *log-likelihood* para asignar puntajes a los posibles sitios de unión. Para esto primero se calcula la composición de nucleótidos del set de secuencias a escanear, se genera un set de secuencias posibles de unión para cada TFBS y se calcula el menor puntaje posible para un sitio de unión aceptado. La función del puntaje calcula la tasa de probabilidad de que una base se encuentre en una posición de un sitio de unión de acuerdo a la PMW sobre la probabilidad de esta base dada la distribución de nucleótidos y se hace la sumatoria (normalizada por longitud) de estas probabilidades y se reportan los sitios de unión con un puntaje mayor al mínimo puntaje

posible para cada PMW [104]. En la búsqueda se asignó un número de *pseudocounts* (probabilidad base para cualquier base en cualquier posición) de 0.02.

3.4.2 Análisis estadístico

Las frecuencias de TFBS para cada planta y para cada grupo de genes se calculó como el porcentaje de genes que contenían al menos un sitio de unión identificado para el factor de transcripción respectivo. Las pruebas de diferencias en distribución de frecuencias se hicieron con la prueba Kolmogorov-Smirnov. Esta prueba es ideal para detectar diferencias entre distribuciones empíricas y ha sido utilizada en análisis similares de regiones promotoras [182].

Las diferencias en frecuencias de TFBS particulares entre grupos se analizaron mediante pruebas de Fisher, las cuales han sido extensivamente utilizadas para la identificación de elementos sobre-representados en regiones promotoras [183-185].

Para calcular distancias entre plantas basadas en la distribución de frecuencias se calculó una matriz de correlación entre las distribuciones utilizando el método de Pearson, esta matriz se utilizó para calcular distancias euclidianas y realizar un cluster jerárquico con *hclust*. Todos los análisis se hicieron con paquetes de Bioconductor en R [186].

4.Regulación transcripcional de miRNAs por factores de transcripción bacterianos.

4.1 Introducción: Efectores TAL (transcription activator-like)

Los efectores TAL pertenecen a una familia de proteínas bacterianas identificadas en los géneros *Xanthomonas* y *Ralstonia*. Estos efectores son translocados a la célula por el sistema de secreción tipo 3 (T3SS). Actúan imitando la actividad de factores de transcripción endógenos y se unen a la región promotora de genes de la planta. Estos efectores pueden contribuir a la virulencia de la bacteria mediante la inducción de genes de susceptibilidad [34, 121, 187].

Los TALs tienen secuencias de aminoácidos altamente conservadas que contienen una señal de secreción y translocación T3SS en la región N-terminal, una señal de localización nuclear y un dominio de activación transcripcional en la región C-terminal. La región central del efector está compuesta por una serie de repeticiones en tandem, cada una de 34 aminoácidos y en algunos casos excepcionales de 35. Los aminoácidos de estas repeticiones son idénticos, salvo los de las posiciones 12 y 13 los cuales son hipervariables y se denominan RVDs (Repeat Variable Diresidues) [97, 188]. La unión de los TALs al DNA es altamente específico y se debe a la combinación de RVDs dado que estos pueden tener preferencias de unión por ciertos nucleótidos. Por ejemplo una repetición de un TAL con el RVD "HD" (His y Asp en los aminoácidos 12 y 13) se une preferencialmente a citosina [97, 188]. De modo que la secuencia de RVDs de un TAL determina la secuencia preferida de unión. Recientemente, se pudo descifrar el código de especificidades RVD-DNA lo cual ayuda a la predicción e identificación de sitios de unión

de TALs en genomas de plantas [97, 188]. Además, recientemente las estructuras cristalográficas de TALs mostraron la manera en que interactúan con la doble hélice de DNA en forma helicoidal manteniendo los RVDs hacia el centro de la hélice, en contacto directo con el DNA [189-192]. Se descifró también que los dos aminoácidos del RVD tienen diferentes roles en la interacción, el primer aminoácido de los RVDs (el aminoácido 12 de cada repetición) no está en contacto directo con el DNA sino que tiene un papel estabilizador de la conformación local, mientras que el segundo aminoácido determina la especificidad mediante interacción directa con los nucleótidos mediante puentes de hidrógeno o interacciones de van der Waals [190, 192].

El descubrimiento del código TAL-DNA es un acontecimiento importante no solo en fitopatología sino en biotecnología. Utilizando las especificidades de los RVDs mejor caracterizados (HD -> C, NI -> A, NG -> T, NN -> G, NS -> NN) y aprovechando la naturaleza modular de los TALs, se han podido crear TALs artificiales con cualquier especificidad deseada, no solo en plantas sino también en otros eucariotas [193]. Se ha podido también fusionar las repeticiones de los TALs a otros dominios de proteínas de interés, notablemente nucleasas, para permitir la edición específica de regiones deseadas del genoma [193]. También se han creado fusiones con dominios de represión para impedir la transcripción de genes blanco [194]. De manera importante algunas de estas construcciones ya están siendo utilizadas para producir fenotipos deseados, como resistencia a enfermedades, en plantas y animales [126, 195, 196].

A pesar de que existen actualmente más de 100 genomas secuenciados de especies del género *Xanthomonas*, que infectan plantas de interés agronómico como banana, tomate, pimiento fríjol, yuca y arroz [36, 50, 197], de que existen más de 100 efectores TAL secuenciados (<http://www.xanthomonas.org/>) y de los avances en biotecnología de TALs, es un número reducido (<30) de *targets* naturales de estos efectores en plantas hospederas los que se han estudiado y validado [97, 118, 119, 123, 188, 198-200]. De manera importante, se desconoce si existen genes de *MIRNA* que sean inducidos por TALs durante la infección y el papel que esto pueda tener en la interacción.

Actualmente se encuentra disponible únicamente un programa de predicción de sitios de unión de TALs (llamados effector binding elements o EBEs) el cual está disponible en la

suite TALE-NT [201]. A pesar de su éxito en la identificación de algunos *targets* [97], el programa falla en detectar algunos pares TAL-DNA conocidos [97, 118].

En esta parte del trabajo se buscó identificar sitios de unión de TALs (EBEs) en las regiones promotoras de genes *MIRNA* y comparar esta predicción con perfiles de expresión de miRNAs y sus *targets* en arroz y yuca. Para esto fue necesario el desarrollo de un nuevo algoritmo para la identificación de estas regiones (llamado Talvez, disponible en <http://bioinfo.mpl.ird.fr/cgi-bin/talvez/talvez.cgi>) que mostró tener mejor desempeño que otras estrategias en un set de datos control.

4.2 Resultados

4.2.1 Desarrollo de un algoritmo de predicción de sitios de unión de efectores TAL

Se diseñó un nuevo programa para la predicción de sitios de unión de efectores TAL, el cual se denominó Talvez. Este programa está basado en gran parte en la estrategia utilizada en el programa plantTFBS[104], utilizado para la búsqueda de sitios de unión de factores de transcripción, presentado en la sección 3.4.1. Básicamente el programa toma una secuencia de RVDs y la transforma en una matriz de peso posicional asignando las probabilidades de unión de cada RVD a los posibles nucleótidos. Posteriormente se utiliza esta matriz para escanear las regiones promotoras y asignar un puntaje a cada posible sitio de unión de acuerdo a la función de *log-likelihood* empleada en [104]. El programa utiliza un código de especificidades RVD-DNA mejorado respecto al publicado en [97, 188], teniendo en cuenta las propiedades estructurales recientemente descritas [190, 192]. Adicionalmente, Talvez emplea una corrección por posición, que permite tolerar más desviaciones de las especificidades RVD-DNA hacia el extremo C-terminal del TAL (luego de 15 repeticiones), puesto que existen evidencias que sugieren que esta región no es determinante para la actividad del TAL [Pérez-Quintero *et al.* sometido] [188, 199].

Se comparó el desempeño de este programa con el algoritmo previamente disponible TALE-NT [201], y con el algoritmo Storyteller, desarrollado por el grupo Effecteurs - Resistance des Plantes aux Bioagresseurs del Institute de Recherche pour le Development (IRD), disponible en <http://bioinfo-prod.mpl.ird.fr/xantho/tales/>. TALE-NT utiliza una estrategia similar a Talvez, basada en PMWs, pero con una función diferente para asignar puntajes que no tiene en cuenta la composición de bases de la secuencia escaneada [201]. Storyteller por el contrario, utiliza una estrategia basada en modelos ocultos de Markov: brevemente, se toma la secuencia de RVDs de un TAL, se traduce a un set de 10000 secuencias de unión posibles y a partir de ellas se construye un modelo oculto, el cual se utiliza para escanear las secuencias promotoras deseadas utilizando Hmmer [202].

Los tres programas fueron probados contra un set de controles que comprendían 74 interacciones TAL-DNA confirmadas experimentalmente obtenidas de la literatura [97, 118-120, 123, 126, 188, 198-200, 203, 204]. De estas, 36 eran interacciones positivas donde se mostró la inducción de un gen curso abajo del sitio de unión del TAL y 38 eran interacciones negativas, las cuales eran usualmente similar en secuencia a las regiones positivas pero que no mostraban actividad de inducción. Se escaneó el set de entrenamiento utilizando los tres programas y se encontró que los tres programas asignaban puntajes significativamente mayores a las interacciones positivas que a las negativas (Wilcoxon one-tail , $p\text{-value} < 0.001$) (**Figura 4-1**). Los tres programas también mostraron altas tasas de verdaderos positivos y bajas tasas de falsos positivos, sin embargo el desempeño de Talvez y Storyteller fue mejor que el de TALE-NT) (**Figura 4-1**).

Una prueba adicional del desempeño de los programas se realizó buscando los *targets* naturales de un grupo de efectores TAL conocidos en los genomas de las plantas hospedero respectivas (*Arabidopsis* y arroz) (**Tabla 4-1**). La búsqueda se hizo sobre todas las regiones promotoras (1000 bases antes del sitio de inicio de la traducción) de los genes anotados en los genomas. Los resultados de la búsqueda fueron ordenados de acuerdo al puntaje y se anotó la posición del *target* conocido en las predicciones de cada programa. Se puede ver que con mayor frecuencia Talvez le asignó un mejor puntaje en

las búsquedas a los *targets* naturales conocidos (**Tabla 4-1**). Con base en estos resultados se decidió utilizar Talvez para todos los análisis posteriores.

Figura 4-1: Desempeño de tres programas de predicción de sitios de unión de efectores TAL.

a) Boxplot mostrando los puntajes de predicción de tres programas sobre un set control de interacciones TAL-DNA positivas (+) y negativas (-). Los puntajes fueron escalados respecto al puntaje máximo obtenido con cada programa para facilitar la comparación y adicionalmente los puntajes de TALENT se escalaron x^{-1} por ser inversos a los de los demás programas, esta transformación no altera la estructura de los datos. ** Indica diferencias significativas entre positivos vs negativos (Wilcoxon one-tailed p-value <0.001). **b)** Gráficas ROC (Receiver operating characteristics) mostrando las tasas de verdaderos y falsos positivos para cada programa calculadas a partir de las predicciones del set control. La línea punteada indica el desempeño teórico de un programa de clasificación aleatorio en el que la tasa de falsos positivos es igual a la de verdaderos positivos.

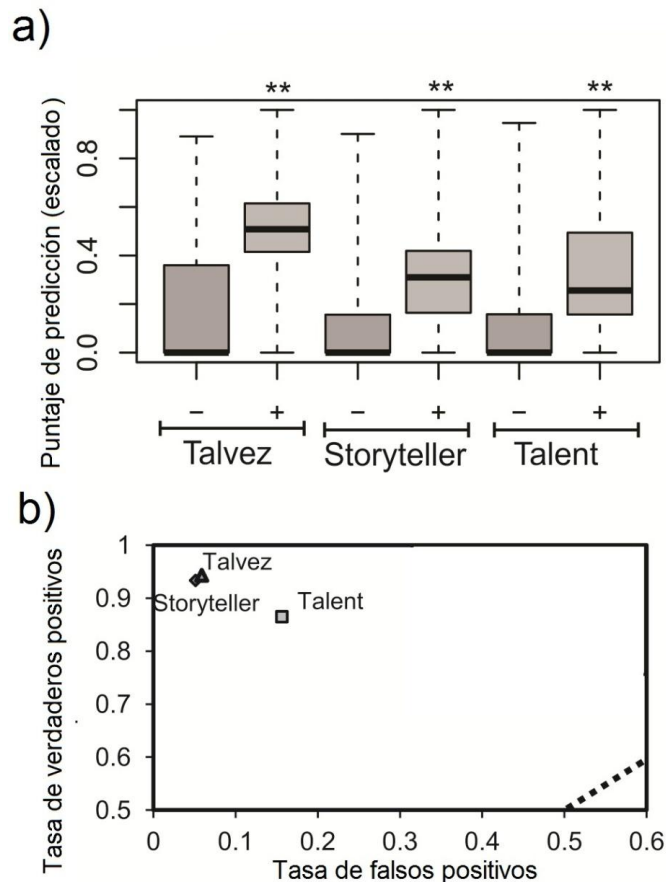


Tabla 4-1. Predicción de *targets* naturales de efectores TALs en búsquedas de promotores utilizando tres programas de predicción.

TAL	Cepa	Target Gene	Locus ID	EBE	Ref.	Talvez posición	Storyteller posición	Talent posición
AvrXa27	Xoo PXO99A	<i>Xa27</i>	LOC Os06g39810	-84	[200, 205]	674	∅	3186
AvrXa7	Xoo PXO86	<i>Os11N3</i>	LOC Os11g31190	-259	[118]	191	344	443
AvrXa7	Xoo PXO86		LOC Os04g19960	-60	[126]	172	∅	∅
Hax2	Xca 5	<i>PAP1</i>	AT1G56650	-130	[188]	180	438	∅
PthXo1	Xoo PXO99A	<i>Os8N3</i>	LOC Os08g42350	-251	[119, 188]	1	1	1
PthXo3	Xoo PXO61	<i>Os11N3</i>	LOC Os11g31190	-261	[118]	1	∅	∅
PthXo6	Xoo PXO99A	<i>OsTFX1</i>	LOC Os09g29820	-136	[188, 198]	7	2	2
PthXo7	Xoo PXO99A	<i>OsTFIIAγ1</i>	LOC Os01g73890	-469	[188, 198]	1	6	2
Tal1c	Xoo PXO99A	<i>OsHen1</i>	LOC Os07g06970	-217	[97, 206]	1	1	1
Tal9A	Xoc BLS256	<i>OsHen1</i>	LOC Os07g06970	-206	[97, 206]	1	1	1
TalC	Xoo BAI3	<i>Os11N3</i>	LOC Os11g31190	-319	[123]	1	2	1

Se muestra la posición (*rank*) de cada *target* en las predicciones hechas sobre el conjunto de promotores del genoma respectivo, con cada programa. Se señala en amarillo la mejor predicción para cada *target*, ∅ indica que el *target* no fue predicho. EBE indica la distancia en nucleótidos desde la primera base (5') del sitio de unión hasta el sitio de inicio de la traducción. La referencia en la literatura para la interacción se indica en Ref.

4.2.1 Predicción de genes *targets* de efectores TAL y expresión diferencial en respuesta a *Xanthomonas*.

Para continuar evaluando el desempeño del nuevo programa se decidió analizar qué tanto se relacionaban las predicciones hechas por Talvez con datos de expresión de genes. Se escogió el modelo arroz-*Xanthomonas oryzae* dada la abundancia de datos de microarreglos y de TALs secuenciados. Para esto se empleó el programa Talvez para identificar genes *target* para todos los TALs de cepas de *Xanthomonas oryzae* para las cuales existe datos de microarreglos de genes de arroz en respuesta a la infección (Xoo BAI3, Xoo PXO86, Xoo PXO99A, Xoo MAFF311018, Xoc BLS256, 71 TALs).

Se identificó la intersección entre los 200 *targets* predichos con mayor puntuación para cada TAL y los 100 genes con mayor inducción ($\log_2FC > 2$, $pvalue < 0.001$) en respuesta a la bacteria en comparaciones de microarreglos. Estos criterios fueron escogidos basados en la expresión de *targets* conocidos en experimentos relevantes.

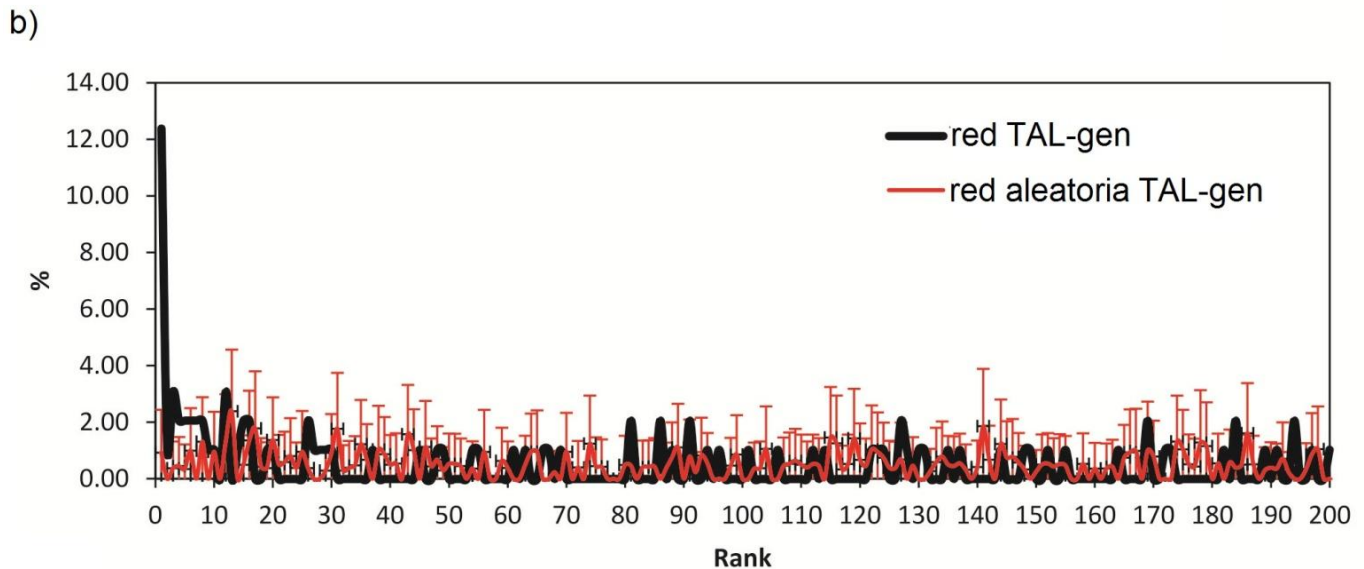
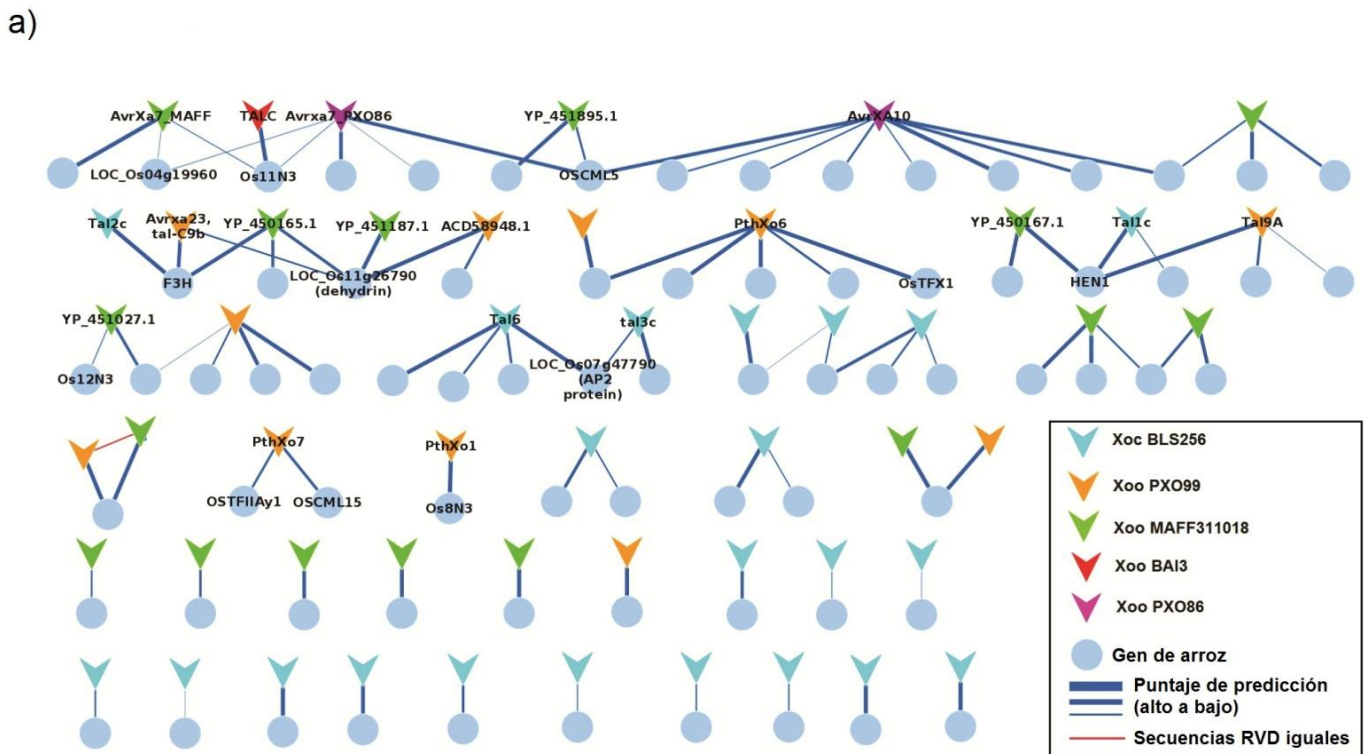
Al realizar esta comparación se encontró un set de 77 genes de arroz potencialmente regulados por 51 efectores TAL. Los genes y los respectivos TAL se organizaron en una red de 96 interacciones, de las cuales 19 fueron interacciones uno-a-uno (un TAL y un gen). Las interacciones restantes se organizaron en módulos discretos que correspondían a múltiples TALs regulando un mismo gen, o viceversa, varios genes regulados por un mismo TAL. Los genes de la red fueron muy heterogéneos en anotación y no mostraron enriquecimiento de términos GO utilizando Análisis de Enriquecimiento Singular (SEA) [207]. Sin embargo un posible elemento en común de estos genes es que al menos el 57% eran genes altamente conectados (grado de conectividad > 5) en una red de interacciones funcionales de arroz [208]. Esto sugiere que los TALs podrían converger hacia hubs en redes génicas, como se describió para otros efectores tipo III en la interacción *Arabidopsis-P. syringae* [209]. No sería sorprendente por tanto que los TALs regularan también genes *MIRNA* dado su papel central en este tipo de redes [167].

Adicionalmente, esta red se comparó con 100 redes obtenida a partir de aleatorizar la secuencia de RVDs de los TALs utilizados y repetir el proceso. De manera importante, la red original, comparada con las redes aleatorias estaba compuesta de un mayor número de elementos y mostraba además prevalencia de genes predichos con altos puntajes (**Figura 4-2**). indicando que las relaciones encontradas son relevantes biológicamente. 12% de las interacciones en la red TAL-gen correspondían a asociaciones entre un TAL y su *target* predicho con mayor puntaje (**Figura 4-2**), esto puede traducirse en una frecuencia del 12% en que el *target* predicho con mejor puntaje para un TAL es diferencialmente expresado en respuesta a la bacteria.

La red además está compuesta de interacciones previamente desconocidas como las inducción de nuevos miembros de la familia Mtn3 y la convergencia de varios TALs sobre genes de la familia de Calmodulinas o genes involucrados en regulación de Calcio. Estas interacciones serán discutidas en detalle en otros trabajos [Pérez-Quintero *et al.* sometido]

Figura 4-2: Red predicha de efectores TAL y sus genes *targets* en la interacción de arroz con distintas cepas de *Xanthomonas oryzae*.

a) Representación jerárquica de la red TAL-gen. Los genes están representados por círculos y los TALs por polígonos (los colores corresponden a distintas cepas). Se muestran los nombres para algunas interacciones conocidas y de interés. b) Distribución de porcentajes de los puntajes de los pares TAL-gen en redes predichas. Se presenta el promedio de 100 redes aleatorias obtenidas a partir de aleatorizar la secuencia de RVDs de los TALs utilizados en la red original.



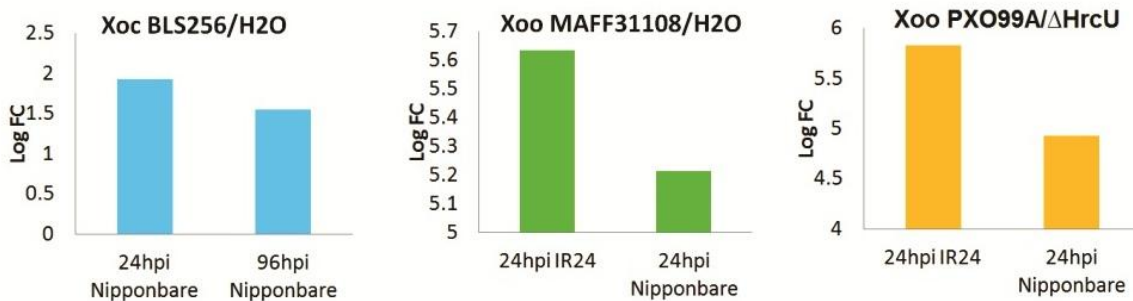
4.2.2 Interacción entre efectores TAL y genes de la vía de miRNAs.

Al analizar la red de interacciones de TAL y genes para arroz y *Xanthomonas*, se encontró que un gen implicado en la vía de miRNAs era un *target* predicho para varios TALs de diferentes cepas, el gen *OsHEN1* (LOC Os07g06970), el cual codifica para una metiltransferasa encargada de la metilación de los miRNAs antes de su exportación al citoplasma [210]. Se ha identificado previamente que este gen es un posible *target* de Tal1c (XocBLS256) y Tal9A (XooPXO99A) y que es inducido en respuesta a bacterias conteniendo estos TALs [97]. La inducción de este gen en respuesta a XooPXO99A se pudo ver en los datos de microarreglos utilizados en la sección 2.2.3 En este trabajo se encontró que también *OsHEN1* es un posible *target* de un TAL de la cepa XooMAFF311018 y es inducido en respuesta a esta bacteria. Además también se predijo como *target* para dos TALs diferentes de las cepas XooJXOIII y XooKACC1033 (para las cuales no existen datos de expresión de genes) (**Figura 4-3**).

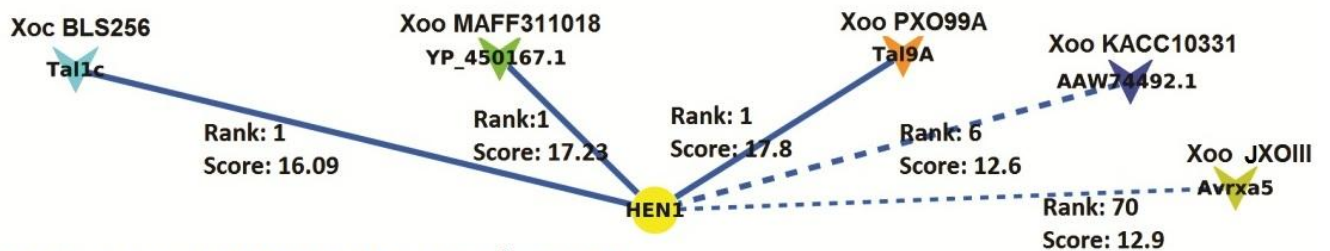
Figura 4-3: Interacción entre efectores TAL de cepas de *Xanthomonas oryzae* y el gen *OsHEN1*.

a) Nivel de inducción del gen *OsHEN1* en respuesta a la cepa indicada en distintos tratamientos (diferentes variedades o tiempos de inoculación). b) Se muestra el modulo de la red de TALs y genes compuesta por *OsHEN1*, se señalan los puntajes de predicción y la posición para cada TAL utilizando talvez. c) Secuencia de RVDs de los TALs que inducen *OsHEN1* y el sitio de unión respectivo para cada TAL en la región promotora de *OsHEN1*.

a) Expresión en microarreglos



b) Red TAL-gen



c) Secuencias de RVDs y sitio de unión (EBE)

Tal1c-> HD-HD-HD-HD-HD-NG-HD-NN-HD-NG-HG-NN-HD-N*-NG-NG
 YP_450167.1-> HD-HD-HD-NG-N*-NG-HD-S*-HG-NI-NI-NN-HD-NN-ND-HD-NI-HD-HG-NG
 Tal9A-> HD-HD-HD-NG-N*-NN-HD-HD-N*-NI-NI-NN-HD-HI-ND-HD-NI-HD-NG-NG
 Avrxa5-> HD-HD-HD-NG-N*-NN-HD-HD-N*-NI-NN-HD-HI-ND-HD-NI-HD-HD-NG-NG
 AAW74492.1-> HD-NG-N*-N*-NI-NI-NN-HD-HI-ND-HD-NI-HD-NG-NG

EBE -> -217 TCCCCCTCGCTTCCTTCCTAAACCCCACTT.... 0 ATG - 3'

----- Sin datos de microarreglos

4.2.3 Interacción entre efectores TAL y genes *MIRNA*.

Se desconoce hasta ahora algún caso en el que un efector TAL induzca la expresión de un gen *MIRNA*. Para estudiar esta posibilidad se realizaron predicciones de sitios de unión de efectores TAL sobre regiones promotoras de genes *MIRNA* y se estudió luego la expresión del miRNA y sus posibles *targets* en librerías de sRNAs y datos de RNA-seq o microarreglos, esperando que existiera inducción del miRNA en respuesta del TAL y consecuentemente represión de los *targets* del miRNA. Se realizó este estudio con las tres cepas *Xanthomonas* tratadas en este trabajo para las que se tienen datos de expresión de miRNAs y genes: Xam CIO151, Xoo BAI3 y Xoo PXO99A.

- **TALE1_{Xam} y miR172e**

En el caso de Xam CIO151, actualmente sólo se han secuenciado dos efectores TAL [50]. Uno de estos efectores: TALE1_{Xam}, se sabe que es importante para la virulencia de la bacteria y se dispone además de datos de RNA-seq de plantas inoculadas con bacterias con y sin el efector (Sección 1.4.5.). Se realizó la predicción de sitios de unión de TALE1_{Xam} sobre los promotores de los genes anotados en yuca incluyendo los promotores de genes *MIRNA*. Únicamente un locus de miRNA tuvo un puntaje alto en las predicciones: miR172e (posición 82 entre los genes con mayor puntaje de predicción). La familia 172 se identificó entre las familias inducidas en respuesta a Xam (**Figura 1-2**) y entre sus *targets* predichos se encontraban principalmente factores de transcripción tipo AP2 y RAP2.

La secuencia del miRNA maduro correspondiente a miR172e es única y diferente de la de los demás miembros de la familia, lo cual permitió mapear específicamente lecturas de las librerías de sRNAs provenientes de este locus. Se encontró que miR172e (siguiendo el comportamiento general de la familia miR172) es inducido en respuesta a Xam. Luego, se quiso ver si la expresión de los *targets* de miR172e disminuía en presencia del efector TALE1_{Xam} en datos de RNA-seq. Para esto, se estudió la expresión de 8 *targets* de miR172e para los cuales se tenían datos en el experimento de RNA-seq

(algunos de estos *targets* eran compartidos con otros loci de la familia). Únicamente uno de los posibles *targets* mostró el comportamiento deseado (represión en presencia del efector TALE1Xam), que corresponde al gen cassava4.1 001818, codificante para una phosphatidylinositol kinasa, sin embargo este gen no se encuentra entre los genes significativamente diferencialmente expresados en presencia del efector (Muñoz-Bodnar *et al.* prep). Es difícil saber si el comportamiento de este gen es consecuencia o no de la inducción del miRNA.

Los demás *targets* estudiados de esta familia (principalmente factores de transcripción) mostraron ser reprimidos en respuesta a la bacteria con o sin la presencia del efector, y de igual manera los demás miembros de la familia también son inducidos en respuesta a Xam en las librerías de sRNAs. Esto sugiere que la inducción de esta familia está relacionada con defensa tipo PTI, y es posible que la inducción de miR172e sea parte del mecanismo general de defensa de la planta y no esté relacionado con la actividad de TALE1_{Xam}. Esta familia, se sabe está implicada en la represión transcripcional del proceso de floración y es inducida en respuesta a múltiples tipos de estrés, probablemente para reducir costos fisiológicos a la planta (como la reproducción) en condiciones no óptimas [171-173].

▪ **TALC y miR399c**

Hasta la fecha se han identificado sólo tres efectores TAL en la cepa Xoo BAI3. Uno de estos efectores, TALC, tiene un papel en la virulencia de esta cepa mediante la inducción de un gen de la familia Mtn3: *Os11N3* [123]. Se llevó a cabo la predicción de sitios de unión para los tres efectores TALs sobre las regiones de miRNAs encontrándose 11 pares TAL/MIRNA con altos puntaje de predicción. Sin embargo solo uno de estos pares correspondía a una familia de miRNA inducida en respuesta a efectores en las librerías de sRNAs: TALC-miR399c.

En arroz existen tres loci de la familia mir399 que tienen la misma secuencia de miRNA maduro que miR399c, por eso no es posible identificar correctamente la expresión de este miRNA particular en las librerías de sRNAs. Sin embargo los niveles de expresión

para el miRNA maduro de estos tres loci muestra inducción en plantas inoculadas con Xoo BAI3 y no con Xoo BAI3ΔHRC, en una de las réplicas analizadas. Entre los *targets* predichos para estos miRNAs se ha confirmado al menos su acción sobre un gen codificante para una Ubiquitin conjugating enzyme (*LOC_Os05g48390*) [131]. Se ha observado en otros sistemas, que las bacterias se aprovechan de la actividad de este tipo de enzimas para suprimir defensas [211, 212].

Sin embargo, la expresión de este gen y la de los demás *targets* predichos para esta familia en datos de RNA-seq no obedece al comportamiento esperado dada la inducción del miRNA, la ausencia de asociación entre expresión de miRNAs y *targets* en estos análisis es discutida en la sección 2.3. De igual manera se estudió la expresión de los posibles *targets* de miR399c en datos de microarreglos (GSE19844, [123]) en los que se compara la expresión en plantas de arroz inoculadas con Xoo BAI3 y con una cepa mutante para el efector TALC (XooBAI3ΔTALC) pero nuevamente no se observó ninguna expresión diferencial.

- **TAL-3Ca y miR827.**

Por último se hizo el mismo análisis para el sistema arroz-Xoo PXO99A. En esta bacteria se han identificado 18 efectores TAL y para al menos tres se han identificado *targets*, aunque se le ha dado un papel predominante en virulencia únicamente a PthXo1 [119, 188, 198]. Tras predecir los *targets* de estos TALs se encontraron 40 posibles asociaciones TAL-MIRNA en este sistema. Únicamente una de estas asociaciones correspondía a una familia inducida en respuesta a efectores: la familia miR827. Los loci miR827a y miR827b los cuales conforman un pre-miRNA policistrónico y producen el mismo miRNA maduro, fueron ambos predichos como inducidos por el TAL-3Ca el cual no tiene ningún blanco conocido en arroz y no resultó tampoco asociado a ningún gen en la red construida en este trabajo. Únicamente uno de tres *targets* predichos de esta familia, una proteína de membrana (*LOC_Os04g48390*) tenía expresión registrada en los microarreglos pero no se encontró como diferencialmente expresado en respuesta a la bacteria.

4.3 Discusión

Uno de los logros de este trabajo es el diseño de una herramienta más precisa para la predicción de sitios de unión de efectores TAL en secuencias de DNA. Esto, junto a datos de expresión permitió la predicción de genes potencialmente inducidos por efectores de *Xanthomonas* durante la infección en arroz y yuca incluyendo genes *MIRNA*.

El nuevo programa, Talvez, fue comparado con otros dos algoritmos existentes utilizando un set control y mostró mejor desempeño en asignar altos puntajes a *targets* conocidos. Las diferencias entre Talvez y los demás programas probablemente se deban al uso de un código mejorado de especificidades TAL-DNA, y a la incorporación de una corrección por posición que le da mayor peso a las primeras repeticiones para asignar los puntajes.

Existen sin embargo aun aspectos pocos caracterizados de la unión TAL-DNA que podrían tenerse en cuenta para futuras versiones del programa. Por ejemplo, recientemente se mostró que diferentes RVDs tenían diferentes contribuciones a la unión al DNA, teniendo los RVDs que forman puentes de hidrógeno con el DNA como NN o HD mayor contribución que RVDs como NI o NG, que se unen por fuerzas de van der Waals y pueden en algunos casos incluso ser insuficientes para tener actividad inductora [213]. Un conocimiento más detallado de las contribuciones de distintos tipos de RVDs sobre *targets* naturales podrá eventualmente ser útil para mejorar las predicciones de sitios de unión. De igual manera otros aspectos como el contexto genómico del sitio de unión (p.e. distancia al sitio de unión de la traducción) podrán eventualmente incorporarse a la predicción una vez se conozca mas sobre su efecto.

El programa diseñado en este trabajo ha sido utilizado para identificar posibles *targets* de efectores TAL y contrastarlos con datos de microarreglos, encontrándose que comúnmente un alto valor de predicción en el programa esta asociado con inducción en respuesta a bacterias. De esta forma se encontró que un gen de la vía de miRNAs: *OsHEN1*, parece ser un *target* clave de cepas de *Xanthomonas oryzae* (al menos para 5 cepas). Esta situación ha sido previamente reportada para dos cepas [97]. Este es uno de sólo dos genes que son predichos como *target* de cepas de Xoc y de Xoo. Esto hace

pensar que la inducción de este gen por parte de cepas de *X. oryzae* puede ser un requerimiento universal para virulencia, aunque se desconoce el efecto que la inducción de este gen pueda tener en la vía de miRNAs y en la fisiología de la planta. Para la cepa Xoo BAI3, estudiada en este trabajo, aún no se ha encontrado un TAL que tenga entre sus posibles *targets* este gen. Igualmente, no se ha encontrado este gen entre los posibles *targets* de los TALs secuenciados de cepas de Xam [Osorio *et al.* in prep] [50].

Por otro lado, en este trabajo se estudió la posibilidad de que la expresión de genes MIRNA fuese modificada por efectores TAL de dos cepas de Xoo y una de Xam. Aunque se predijeron varias interacciones TAL-MIRNA, únicamente una interacción para cada cepa presentó una relativamente solida evidencia experimental (inducción del respectivo miRNA en librerías de RNAs pequeños). En todos los casos, sin embargo, estas asociaciones correspondían a puntajes relativamente bajos de predicción para el sitio de unión y además no se pudo establecer una clara asociación entre los niveles de inducción de los miRNAs y la expresión sus posibles *targets*. Estos datos sugiere que los *MIRNAs* no son en verdad inducidos por estos TALs o alternativamente que su inducción es muy baja para causar un efecto en los *targets*. No se descarta, sin embargo, la posibilidad de que otros efectores no estudiados en este trabajo tengan la capacidad de inducir genes *MIRNA* y que esto tenga un efecto fisiológico importante en la planta. El que la transcripción de la mayoría de los genes *MIRNA* este regulada por los mismos factores que la mayoría de los genes indicaría que pueden igualmente ser regulados por factores de transcripción bacterianos. La herramienta acá desarrollada para la predicción de sitios de unión de efectores TAL será de gran ayuda para estudiar esta hipótesis una vez se tengan datos de expresión de miRNAs y genes en respuesta a otras cepas que cuenten con otros repertorios de efectores TAL.

4.4 Materiales y métodos

4.4.1 Algoritmo de predicción de sitios de unión de efectores TAL: Talvez

Una interface web y el código fuente de Talvez se encuentra disponible en <http://bioinfo-prod.mpl.ird.fr/xantho/tales> (login = effectors, password = l00p). La estrategia de Talvez está basada en gran parte en el algoritmo de búsqueda de sitios de unión de factores de transcripción descrita en [104]. El programa está basado en los lenguajes de programación perl y java.

Talvez busca sitios de unión de efectores TAL a través de varios pasos: (I) Para cada secuencia de RVDs se crea una matriz (PMW) de acuerdo a las especificidades de unión RVD-DNA. Si se especifica la corrección por posición, se utiliza una versión escalada del código que tolera más coincidencias imperfectas entre RVDs y DNA. Se añade además una probabilidad de base (pseudocounts) a todos los elementos de la PMW de forma que cualquier base es permitida en cualquier posición al menos con la probabilidad de base (por defecto = 0.001). (II) Se calcula la composición de nucleótidos de las secuencias de DNA a escanear. (III) Se asignan puntajes a todos los posibles sitios de unión en la secuencia de DNA utilizando la función de probabilidad descrita en [104]. (IV) Se reportan los resultados que se encuentren por encima de un puntaje definido por el usuario y son ordenados de acuerdo al puntaje. El programa escanea únicamente la hebra positiva de las secuencias y no permite la inserción de gaps. La versión web permite escanear las regiones promotoras de los genes identificados en los genomas de 25 plantas, obtenidos de Phytozome (<http://www.phytozome.net/>).

El código RVD-DNA utilizado en el programa está basado en especificidades obtenidas de la literatura [97] y fue modificado para ajustar nuevos aspectos de la interacción (**Anexo F**). Se asignaron iguales probabilidades a RVDs que terminasen en el mismo aminoácido. Se modificaron las especificidades para la posición 0, permitiendo la unión de citosina y timina, teniendo la citosina menor probabilidad de unión.

4.4.2 Sets de datos y predicciones

Se recopiló un set de datos control para probar el desempeño de programas de predicción de datos basado en interacciones TAL-DNA publicadas que hubiesen sido validadas experimentalmente mediante ensayos GUS, EMSA, y/o QRT-PCR [97, 118-120, 123, 126, 188, 198-200, 203, 204, 206] (**Anexo G**). Un subset de interacciones correspondientes a *targets* naturales de TAL en los genomas de arroz y Arabidopsis fue utilizado para ensayos control de búsquedas en todo el genoma.

Para las predicciones con Talvez se utilizaron los siguientes parámetros: pseudocounts '1e-05', minimum score '9', position correction '19'. Las predicciones con storyteller se hicieron con los siguientes parámetros: rounds '1e5', noise '0.5', noise-shape 'hvaa-dependent', max e-value '700', minscore '2', gap probability '1e-3'. Las predicciones con Talent se hicieron con parámetros por defecto. Las regiones promotoras de genes fueron obtenidas de www.phytozome.com y las regiones promotoras de miRNAs fueron obtenidas como fue descrito en la sección 3.4.1.

Los resultados de escanear el set de interacciones control con cada programa fueron analizados así: para cada TAL y cada programa se estableció un puntaje mínimo definido como el mínimo puntaje para una interacción positiva conocida, todos los sitios de unión predichos sobre este puntaje fueron considerados predicciones positivas, y debajo de este puntaje predicciones negativas. De esta forma un verdadero positivo corresponde a una interacción positiva conocida con un puntaje mayor a este puntaje mínimo.

4.4.3 Análisis de microarreglos

Datos de expresión de genes de plantas inoculadas en respuesta a *Xanthomonas oryzae*, con sus respectivos controles fueron obtenidos de los siguientes sets de datos de la base de datos GEO datasets (<http://www.ncbi.nlm.nih.gov/gds>): Comparative transcriptional profiling of rice undergoing infection by *X. oryzae* pv. *oryzae* or by *X. oryzae* pv. *oryzicola* (GSE16793), “affy_Xoo_rice-Transcriptomics-based identification of Xoo strain BAI3 Talc *targets* in rice” (GSE19844) y “Comparison of transcriptional

responses of two susceptible rice cultivars to strains of *Xanthomonas oryzae*" (GSE36272), también utilizado en la sección 2.2.3. Se obtuvieron datos de expresión normalizados para estos experimentos en la base de datos PlexDB [142] y fueron comprobados utilizando la función `rma` (Robust Multichip Average) del paquete `affy` de Bioconductor en R [214].

Se consideraron como genes de interés aquellos que resultaron diferencialmente expresados en alguna de las siguientes comparaciones: (I) Plantas inoculadas con cepas silvestres comparadas con plantas no tratadas. (II) Plantas inoculadas con cepas silvestres comparadas con plantas mutantes para el T3SS. (III) Plantas inoculadas con cepas silvestres comparadas con plantas mutantes para TALs específicos. La identificación de genes diferencialmente expresados se realizó con el paquete `limma` de bioconductor [215]. Se consideraron diferencialmente expresados los genes que tuviesen \log_2FC mayor o igual a 1, con un p -value menor a 0.05. Para cada experimento se tomaron únicamente los 100 genes más inducidos en respuesta a la bacteria ordenados de acuerdo al \log_2FC .

Los sets de genes diferencialmente expresados fueron contrastados con los sets de *targets* predichos y las interacciones fueron representadas en una red utilizando `cytoscape` [216].

5. Conclusiones

En este trabajo se estudio la expresión de miRNAs y sus targets en respuesta a bacterias del género *Xanthomonas* en yuca y arroz y los posibles mecanismos de regulación transcripcional de estos miRNAs por factores de transcripción vegetales y bacterianos. Se basó este trabajo en un modelo general del papel de los miRNAs en interacciones planta-bacteria relacionado con las respuestas de inmunidad conocidas tradicionalmente: ETI y PTI.

En yuca se encontró que la expresión diferencial de miRNAs correspondía a este modelo general y que la respuesta era similar a la reportada en *Arabidopsis*. Se inducen miRNAs que reprimen reguladores negativos de defensa como son genes involucrados en señalización de auxinas y se reprimen miRNAs que normalmente regulan genes de resistencia y genes involucrados en distribución de cobre. Se pudo relacionar la expresión diferencial de los miRNAs con sus targets, y se encontró un posible efecto de represión de silenciamiento para un efector tipo TAL secretado por la bacteria.

En contraste en arroz no se encontró una respuesta clara mediada por miRNAs y no se pudo asociar la expresión diferencial de éstos con la expresión de targets. De modo que aunque se encontraron familias diferencialmente expresadas estas pueden no tener un efecto fisiológico importante o pueden tenerlo bajo condiciones diferentes a las evaluadas. Las familias que se encontraron fueron diferentes a las que se esperaban de acuerdo al modelo general y variaron de acuerdo a la cepa de *X. oryzae* estudiada. Por otro lado no mostraron mucha variación en presencia o ausencia de efectores secretados por la bacteria. Y notablemente, la respuesta parece ser cuantitativa, dependiendo de la población bacteriana que se encuentre en la planta, por lo que es posible que se necesite

un umbral de bacteria para tener una respuesta robusta que se evidencie en la expresión diferencial de genes.

Se estudiaron las regiones promotoras de los miRNAs en estas plantas para entender la posibles diferencias en regulación y se encontraron diferencias en la distribución de sitios de unión de factores de transcripción en estas regiones para la las plantas estudiadas, siendo esta distribución similar entre yuca y *Arabidopsis* lo cual puede estar relacionado con las respuestas similares encontradas. Se identificaron además elementos sobre-representados en las familias diferencialmente expresadas y estos elementos parecen estar relacionados con procesos de respuesta a luz, regulación del ritmo circadiano y regulación de floración.

Finalmente se estudió la posible regulación de miRNAs por efectores bacterianos tipo TAL los cuales actúan como factores de transcripción. Aunque se encontraron posibles asociaciones entre estos efectores y algunos miRNAs para cepas de Xam y Xoo éstas no se pudieron relacionar con expresión diferencial de genes. Sin embargo, se encontró una fuerte asociación entre los efectores TAL de varias cepas de Xoo y la vía de miRNAs mediante la inducción del gen HEN1 involucrado en la metilación de miRNAs, aunque el efecto fisiológico que esto pueda tener es aún desconocido.

Los resultados de este trabajo recalcan la importancia de estudiar proceso de inmunidad mediados por miRNAs en plantas de distintos grupos filogenéticos y muestran que puede haber gran variación entre estas respuestas inclusive en respuesta a cepas similares genéticamente. Es además el primer estudio que analiza la expresión por miRNAs en respuesta a bacterias del género *Xanthomonas*, uno de los géneros de bacterias más importantes en fitopatología por su impacto en un gran número de cultivos y por su habilidad de causar directamente cambios transcripcionales en las plantas.

6.Recomendaciones y perspectivas

Aunque estos resultados son promisorios se necesita aún ampliar el conociendo sobre la regulación de miRNAs en estas plantas y mejorar algunos aspectos de esta investigación. Particularmente los sets de datos utilizados para yuca fueron bastantes heterogéneos y pueden tener mucho ruido, y se recomienda hacer el estudio a tiempos específicos e incluir réplicas técnicas y biológicas. De igual manera en arroz se deberían estudiar otros tiempos para determinar con certeza si existe o no una respuesta mediada por miRNAs. Por otro lado se debe también mejorar la estrategia para hacer asociaciones más acertadas entre la expresión de miRNAs y los targets, se recomendaría estudiar librerías de RNAs degradados (*degradome sequencing*) para verificar el clivaje de los targets predichos.

Se debe demás tener en cuenta que gran parte de los resultados aquí generados son producto de análisis *in silico* y es aún necesaria la validación experimental de varios aspectos como la expresión de los miRNAs mediante northern blot o Q-PCR de RNAs pequeños, el clivaje de los targets mediante RACE-PCR y la inducción de genes por efectores TAL mediante ensayos EMSA o GUS.

Se espera estos resultados puedan guiar estrategias de mejoramiento en estas plantas, particularmente en yuca en la que parece que los miRNAs sí juegan un papel en defensa. Algunas de estas estrategias pueden incluir:

- El diseño de miRNAs artificiales dirigidos contra factores de respuesta auxinas y su transformación en plantas bajo promotores inducibles en respuesta bacterias para fortalecer la respuesta que se da naturalmente o regular ARFs que no se encuentren bajo control de miRNAs.
- La transformación de plantas con target mimics (con promotores inducibles en respuesta bacterias) específicos que bloqueen la actividad de miRNAs que regulen normalmente genes de resistencia.
- El diseño de TALs artificiales que induzcan la expresión de miRNAs con un papel en defensa contra bacterias. En Arroz podría por ejemplo inducirse artificialmente la expresión de miRNAs conservados que controlen los ARFs y ver si esto se traduce en una respuesta de resistencia.

A pesar de las dificultades inherentes a los ensayos bioinformáticos los resultados de este trabajo fueron en gran parte satisfactorios y se cumplió con los objetivos propuestos. Parte de los resultados acá presentados fueron publicados en un artículo [69] ,la exploración bibliográfica necesaria para esta investigación fue aprovechada para la elaboración de artículos de revisión [58, 59] y un capítulo de libro [35]. Partes de estos resultados fueron también expuestos en varios congresos, incluyendo:

- 4th Xanthomonas genomics conference 2012, Anger. France
- GCP21-II, Second Scientific Conference of the Global Cassava Partnership for the 21st Century. 2012. Kampala, Uganda.
- Effectome 2011, Lauret, France.
- 1er Congreso Colombiano de Biología Computacional 2011 (trabajo ganador). Bogotá, Colombia.
- XXX Congreso Colombiano y XVI Latinoamericano de Fitopatología. 2011. Bogotá Colombia

Y permitieron además que se le otorgara al grupo de investigación Manihot Biotec el Premio semilleros ADN 2011 de investigación otorgado por la la Asociación de Biotecnología Vegetal Agrícola (<http://www.agrobio.org.co/news/view.php?id=MTg3>).

Además, los métodos y los sets de datos acá generados se esperan sean valiosos para futuras investigaciones y en algunos casos ya están siendo utilizados. Por ejemplo, las técnicas empleadas para la identificación de miRNAs sirvieron para un análisis en la

planta con potencial anti-malaria *Artemisia annua* [217]. Las librerías de RNAs pequeños de yuca fueron aprovechadas para la identificación de otros RNAs reguladores: nat-siRNAs y cis-nat-RNAs [Quintero *et al.*, sometido] así como ncRNAs provenientes de la bacteria [Bernal *et al.*, sometido], también se utilizaron en un proyecto de análisis masivo de RNAs provenientes de genomas de cloroplastos en varias plantas [Urrego *et al.*, in prep] e hicieron parte de un set de datos utilizado para crear una red de interacciones de genes con papel en inmunidad en yuca [Leal *et al.*, sometido]. En arroz se planea utilizar estos datos para estudiar no solo miRNAs sino otros tipos de RNAs pequeños involucrados en defensa y además los datos de RNA-seq permitirán estudiar nuevos transcritos que puedan ser inducidos por efectores TAL en la interacción. Y finalmente, el programa Talvez diseñado para la predicción de *targets* de efectores TAL ha permitido la identificación de funciones previamente desconocidas entre TALs de *Xanthomonas oryzae* [Pérez-Quintero *et al.*, sometido] y fue utilizado para la predicción de varios sitios de unión de efectores de distintas cepas de Xam [51] [Osorio *et al.*, in prep], de la cepa de Xoo africana MAI1 [Pesce *et al.*, in prep] y de la especie patógena de cítricos *X. citri* [Escalon *et al.*, in prep].

Se espera estos resultados sigan siendo útiles y permitan avances en el entendimiento de las bases moleculares de la interacción planta-patógeno.

7. Anexos

Los anexos se encuentran como archivos separados en el CD que acompaña este trabajo.

Anexo A. Ubicación genómica y características de miRNAs identificados en yuca.

Anexo B. Targets de miRNAs predichos en yuca.

Anexo C. Expresión de miRNAs en arroz en respuesta a Xoo BAI3 y Xoo PXO99A

Anexo D. Targets de miRNAs predichos en arroz.

Anexo E. Frecuencias de TFBs en regiones promotoras de MIRNAs y genes codificantes.

Anexo F. Especificidades TAL-DNA utilizadas por el programa Talvez

Anexo G. Set de interacciones TAL-DNA utilizadas como control para evaluar la predicción de sitios de unión de efectores.

8. Bibliografía

1. Niño-Liu DO, Ronald PC, Bogdanove AJ: **Xanthomonas oryzae pathovars: model pathogens of a model crop**. *Molecular plant pathology* 2006, **7**(5):303-324.
2. Ou SH, Commonwealth Mycological Institute (Great Britain): **Rice diseases**, 2nd edn. Kew, Surrey, UK: Commonwealth Mycological Institute; 1985.
3. Verdier V: **Bacteriosis vascular (o añublo bacteriano) de la yuca causada por Xanthomonas axonopodis pv. manihotis**. En CIAT eds *La yuca en el Tercer Milenio Sistemas modernos de producción procesamiento utilización y comercialización* 2002:148-159.
4. Lozano JC: **Cassava bacterial blight: a manageable disease**. *Plant Disease* 1986, **70**:1089-1093.
5. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Beer SV, Machado MA *et al*: **Top 10 plant pathogenic bacteria in molecular plant pathology**. *Molecular plant pathology* 2012, **13**(6):614-629.
6. Prudencio YC: **The food security stabilization roles of cassava in Africa**. *Food Policy* 1994, **19**(1):57-64.
7. Nguyen NV: **Challenges to ensuring food security through rice**. In., vol. 1: CABI LA - English; 2006: 9 pp.-9 pp.
8. López CE, Bernal AJ: **Cassava Bacterial Blight: Using Genomics for the Elucidation and Management of an Old Problem**. *Tropical Plant Biology* 2012.
9. Chisholm ST, Coaker G, Day B, Staskawicz BJ: **Host-microbe interactions: shaping the evolution of the plant immune response**. *Cell* 2006, **124**(4):803-814.
10. Kidner CA, Martienssen RA: **The developmental role of microRNA in plants**. *Current opinion in plant biology* 2005, **8**(1):38-44.
11. Obbard DJ, Gordon KHJ, Buck AH, Jiggins FM: **The evolution of RNAi as a defence against viruses and transposable elements**. *Review Literature And Arts Of The Americas* 2009(October 2008):99-115.
12. Holeski LM, Jander G, Agrawal AA: **Transgenerational defense induction and epigenetic inheritance in plants**. *Trends in ecology & evolution* 2012, **27**(11):618-626.
13. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function**. *Cell* 2004, **116**(2):281-297.
14. Tang G, Reinhart BJ, Bartel DP, Zamore PD: **A biochemical framework for RNA silencing in plants**. *Genes & development* 2003, **2**:49-63.
15. Zhu JK: **Reconstituting plant miRNA biogenesis**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(29):9851-9852.
16. Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ *et al*: **Criteria for annotation of plant MicroRNAs**. *The Plant cell* 2008, **20**(12):3186-3190.
17. Allen E, Xie Z, Gustafson AM, Carrington JC: **microRNA-directed phasing during trans-acting siRNA biogenesis in plants**. *Cell* 2005, **121**(2):207-221.

18. Yoshikawa M, Peragine A, Park MY, Poethig RS: **A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis**. *Genes & development* 2005, **19**(18):2164-2175.
19. Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD: **A plant miRNA contributes to antibacterial resistance by repressing auxin signaling**. *Science* 2006, **312**(5772):436-439.
20. Fu J, Liu H, Li Y, Yu H, Li X, Xiao J, Wang S: **Manipulating broad-spectrum disease resistance by suppressing pathogen-induced auxin accumulation in rice**. *Plant physiology* 2011, **155**(1):589-602.
21. Zhang W, Gao S, Zhou X, Chellappan P, Chen Z, Zhang X, Fromuth N, Coutino G, Coffey M, Jin H: **Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks**. *Plant molecular biology* 2011, **75**(1-2):93-105.
22. Li Y, Zhang Q, Zhang J, Wu L, Qi Y, Zhou JM: **Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity**. *Plant physiology* 2010, **152**(4):2222-2231.
23. Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BaCM, Baulcombe DC: **A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs**. *The Plant cell* 2012, **24**(3):859-874.
24. Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B: **MicroRNA regulation of plant innate immune receptors**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**(5):1790-1795.
25. Lu S, Sun YH, Amerson H, Chiang VL: **MicroRNAs in loblolly pine (Pinus taeda L.) and their association with fusiform rust gall development**. *The Plant journal : for cell and molecular biology* 2007, **51**(6):1077-1098.
26. Chen L, Ren Y, Zhang Y, Xu J, Zhang Z, Wang Y: **Genome-wide profiling of novel and conserved Populus microRNAs involved in pathogen stress response by deep sequencing**. *Planta* 2012, **235**(5):873-883.
27. Lelandais-Briere C, Naya L, Sallet E, Calenge F, Frugier F, Hartmann C, Gouzy J, Crespi M: **Genome-wide Medicago truncatula small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules**. *The Plant cell* 2009, **21**(9):2780-2796.
28. Li H, Deng Y, Wu T, Subramanian S, Yu O: **Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation**. *Plant physiology* 2010, **153**(4):1759-1770.
29. Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, González AJ, Yan Z, Kitto SL, Grusak Ma *et al*: **MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs**. *Genes & development* 2011, **25**(23):2540-2553.
30. Devers EA, Branscheid A, May P, Krajinski F: **Stars and symbiosis: microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis**. *Plant physiology* 2011, **156**(4):1990-2010.
31. Navarro L, Jay F, Nomura K, He SY, Voinnet O: **Suppression of the microRNA pathway by bacterial effector proteins**. *Science* 2008, **321**(5891):964-967.
32. Mosher RA, Baulcombe DC: **Bacterial pathogens encode suppressors of RNA-mediated silencing**. *Genome biology* 2008, **9**(10):237.
33. Dunoyer P, Lecellier CH, Parizotto EA, Himber C, Voinnet O: **Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing**. *The Plant cell* 2004, **16**(5):1235-1250.
34. Bogdanove AJ, Schornack S, Lahaye T: **TAL effectors: finding plant genes for disease and defense**. *Current opinion in plant biology* 2010, **13**(4):394-401.
35. López CP-QAL: **The Micromics Revolution : MicroRNA-Mediated Approaches to Develop Stress-Resistant Crops**. *Improving Crop Resistance to Abiotic Stress* 2012.
36. Ryan RP, Vorholter FJ, Potnis N, Jones JB, Van Sluys MA, Bogdanove AJ, Dow JM: **Pathogenomics of Xanthomonas: understanding bacterium-plant interactions**. *Nature reviews Microbiology* 2011, **9**(5):344-355.
37. Megraw M, Baev V, Rusinov V, Jensen ST, Kalantidis K, Hatzigeorgiou AG: **MicroRNA promoter element discovery in Arabidopsis**. *RNA* 2006, **12**(9):1612-1619.

38. Zhou X, Ruan J, Wang G, Zhang W: **Characterization and identification of microRNA core promoters in four model species**. *PLoS computational biology* 2007, **3**(3):e37.
39. Kruszká K, Pieczynski M, Windels D, Bielewicz D, Jarmolowski A, Szweykowska-Kulinska Z, Vazquez F: **Role of microRNAs and other sRNAs of plants in their changing environments**. *Journal of plant physiology* 2012.
40. Willenbrock H, Salomon J, Sokilde R, Barken KB, Hansen TN, Nielsen FC, Moller S, Litman T: **Quantitative miRNA expression analysis: comparing microarrays with next-generation sequencing**. *RNA* 2009, **15**(11):2028-2034.
41. El-Sharkawy MA: **Cassava biology and physiology**. *Plant molecular biology* 2004, **56**(4):481-501.
42. Fao: **OECD-FAO AGRICULTURAL OUTLOOK 2010-2019 HIGHLIGHTS**. In.; 2010: 88-88.
43. Blagbrough IS, Bayoumi SAL, Rowan MG, Beeching JR: **Cassava: An appraisal of its phytochemistry and its biotechnological prospects**. *Phytochemistry* 2010, **71**(17-18):1940-1951.
44. Ziska LH, Runion GB, Tomecek M, Prior SA, Torbet HA, Sicher R: **An evaluation of cassava, sweet potato and field corn as potential carbohydrate sources for bioethanol production in Alabama and Maryland**. *Biomass and Bioenergy* 2009, **33**(11):1503-1508.
45. Wydra K, Verdier V: **Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics**. *Agriculture, Ecosystems and Environment* 2002, **93**(211-226).
46. Verdier V, Restrepo S, Mosquera G, Jorge V, Lopez C: **Recent progress in the characterization of molecular determinants in the Xanthomonas axonopodis pv. manihotis-cassava interaction**. *Plant molecular biology* 2004, **56**(4):573-584.
47. Restrepo S, Velez CM, Verdier V: **Measuring the Genetic Diversity of Xanthomonas axonopodis pv. manihotis Within Different Fields in Colombia**. *Phytopathology* 2000, **90**(7):683-690.
48. Wydra K, Zinsou V, Jorge V, Verdier V: **Identification of Pathotypes of Xanthomonas axonopodis pv. manihotis in Africa and Detection of Quantitative Trait Loci and Markers for Resistance to Bacterial Blight of Cassava**. *Phytopathology* 2004, **94**(10):1084-1093.
49. Restrepo S, Velez CM, Duque MC, Verdier V: **Genetic structure and population dynamics of Xanthomonas axonopodis pv. manihotis in Colombia from 1995 to 1999**. *Applied and environmental microbiology* 2004, **70**(1):255-261.
50. Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T *et al*: **High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**(28):E1972-1979.
51. Castiblanco LF, Gil J, Rojas A, Osorio D, Gutierrez S, Munoz-Bodnar A, Perez-Quintero AL, Koebnik R, Szurek B, Lopez C *et al*: **TALE1 from Xanthomonas axonopodis pv. manihotis acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants**. *Molecular plant pathology* 2012.
52. Ogunjobi AA, Fagade OE, Dixon AGO, Bandyopadhyay R: **Assessment of large population of cassava accessions for resistant to cassava bacterial blight infection in the screen house environment**. *Journal of Agricultural Biotechnology and Sustainable Development* 2010, **2**(6):87-91.
53. Jorge V, Fregene M, Vélez CM, Duque MC, Tohme J, Verdier V: **QTL analysis of field resistance to Xanthomonas axonopodis pv. manihotis in cassava**. *TAG Theoretical and Applied Genetics* 2001, **102**(4):564-571.
54. Contreras Nieto E, Lopez CE: **EXPRESIÓN DE DOS GENES CANDIDATOS A RESISTENCIA CONTRA LA BACTERIOSIS VASCULAR EN YUCA Expression Of Two Resistance Gene Candidates Against Cassava Bacterial Blight In Cassava**. *Acta biol Colomb* 2008, **13**(2):175-188.

55. Lopez CE, Quesada-Ocampo LM, Bohorquez A, Duque MC, Vargas J, Tohme J, Verdier V: **Mapping EST-derived SSRs and ESTs involved in resistance to bacterial blight in *Manihot esculenta***. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* 2007, **50**(12):1078-1088.
56. Jorge V, Fregene MA, Duque MC, Bonierbale MW, Tohme J, Verdier V: **Genetic mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz)**. *TAG Theoretical and Applied Genetics* 2000, **101**(5-6):865-872.
57. Lopez C, Soto M, Restrepo S, Piégu B, Cooke R, Delseny M, Tohme J, Verdier V: **Gene expression profile in response to *Xanthomonas axonopodis* pv. *manihotis* infection in cassava using a cDNA microarray**. *Plant molecular biology* 2005, **57**(3):393-410.
58. Sablok G, Perez-Quintero AL, Hassan M, Tatarinova TV, Lopez C: **Artificial microRNAs (amiRNAs) engineering - On how microRNA-based silencing methods have affected current plant silencing research**. *Biochemical and biophysical research communications* 2011, **406**(3):315-319.
59. Pérez-quintero ÁL, López C: **Artificial microRNAs and their applications in plant molecular biology** *MicroARNs artificiales y sus aplicaciones en biología molecular de plantas*. *Network* 2010.
60. Liu Q, Chen YQ: **A new mechanism in plant engineering: the potential roles of microRNAs in molecular breeding for crop improvement**. *Biotechnology advances* 2010, **28**(3):301-307.
61. Amiteye S, Corral JM, Sharbel TF: **Overview of the potential of microRNAs and their target gene detection for cassava (*Manihot esculenta*) improvement**. *Afr J Biotechnol* 2011, **10**(14):2562-2573.
62. Zhang ZZ, Yu J, Li D, Liu F, Zhou X, Wang T, Ling Y, Su Z: **PMRD: plant microRNA database**. *Nucleic acids research* 2010, **38**(Database issue):D806-D813.
63. Zeng C, Wang W, Zheng Y, Chen X, Bo W, Song S, Zhang W, Peng M: **Conservation and divergence of microRNAs and their functions in Euphorbiaceae plants**. *Nucleic acids research* 2010, **38**(3):981-995.
64. Pérez-Quintero ÁL, Zapata A, López C: **Identification of Conserved miRNAs in Cassava (*Manihot esculenta*)**. *Acta Biológica Colombiana* 2010, **15**(2):149-168.
65. Lertpanyasampatha M, Pramoolkit P, Kongsawadworakul P, Viboonjun U, Chrestin H, Narangajavana J: **Conserved microRNAs and their targets in rubber tree and the related Euphorbiaceae species**. *Journal of Plant Biochemistry and Biotechnology* 2012.
66. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ: **miRBase: tools for microRNA genomics**. *Nucleic acids research* 2008, **36**(Database issue):D154-158.
67. Prochnik S, Marri PR, Desany B, Rabinowicz PD, Kodira C, Mohiuddin M, Rodriguez F, Fauquet C, Tohme J, Harkins T *et al*: **The Cassava Genome: Current Progress, Future Directions**. *Trop Plant Biol* 2012, **5**(1):88-94.
68. Patanun O, Lertpanyasampatha M, Sojikul P, Viboonjun U, Narangajavana J: **Computational Identification of MicroRNAs and Their Targets in Cassava (*Manihot esculenta* Crantz.)**. *Molecular biotechnology* 2012.
69. Pérez-Quintero ÁL, Quintero A, Urrego O, Vanegas P, López C: **Bioinformatic identification of cassava miRNAs differentially expressed in response to infection by *Xanthomonas axonopodis* pv. *manihotis***. *BMC plant biology* 2012, **12**(1):29-29.
70. Jeong DH, Park S, Zhai J, Gurazada SG, De Paoli E, Meyers BC, Green PJ: **Massive Analysis of Rice Small RNAs: Mechanistic Implications of Regulated MicroRNAs and Variants for Differential Target RNA Cleavage**. *The Plant cell* 2011, **23**(12):4185-4207.
71. Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC: **Genome-wide profiling and analysis of *Arabidopsis* siRNAs**. *PLoS biology* 2007, **5**(3):e57.
72. Moxon S, Schwach F, Dalmay T, Maclean D, Studholme DJ, Moulton V: **A toolkit for analysing large-scale plant small RNA datasets**. *Bioinformatics* 2008, **24**(19):2252-2253.
73. Li H, Durbin R: **Fast and accurate long-read alignment with Burrows-Wheeler transform**. *Bioinformatics* 2010, **26**(5):589-595.

74. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *Journal of molecular biology* 1990, **215**(3):403-410.
75. Jones-Rhoades MW, Bartel DP: **Computational identification of plant microRNAs and their targets, including a stress-induced miRNA.** *Molecular cell* 2004, **14**(6):787-799.
76. Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z: **MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features.** *Nucleic acids research* 2007, **35**(Web Server issue):W339-344.
77. Chen X, Li Q, Wang J, Guo X, Jiang X, Ren Z, Weng C, Sun G, Wang X, Liu Y *et al*: **Identification and characterization of novel amphioxus microRNAs by Solexa sequencing.** *Genome biology* 2009, **10**(7):R78.
78. Perez-Quintero AL, Neme R, Zapata A, Lopez C: **Plant microRNAs and their role in defense against viruses: a bioinformatics approach.** *BMC plant biology* 2010, **10**:138.
79. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS: **Human MicroRNA targets.** *PLoS biology* 2004, **2**(11):e363.
80. Dai X, Zhao PX: **psRNATarget: a plant small RNA target analysis server.** *Nucleic acids research* 2011, **39**(Web Server issue):W155-159.
81. Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, Nussaume L, Crespi MD, Maizel A: **miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth.** *The Plant cell* 2010, **22**(4):1104-1117.
82. Abdel-Ghany SE, Pilon M: **MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis.** *The Journal of biological chemistry* 2008, **283**(23):15932-15945.
83. Cuperus JT, Fahlgren N, Carrington JC: **Evolution and functional diversification of MIRNA genes.** *The Plant cell* 2011, **23**(2):431-442.
84. Kessler SA, Shimosato-Asano H, Keinath NF, Wuest SE, Ingram G, Panstruga R, Grossniklaus U: **Conserved Molecular Components for Pollen Tube Reception and Fungal Invasion.** *Science* 2010, **330**(6006):968-971.
85. Escobar-Restrepo J-M, Huck N, Kessler S, Gagliardini V, Gheyselinck J, Yang W-C, Grossniklaus U: **The FERONIA receptor-like kinase mediates male-female interactions during pollen tube reception.** *Science* 2007, **317**(5838):656-660.
86. Ruiz-Ferrer V, Voinnet O: **Roles of plant small RNAs in biotic stress responses.** *Annual review of plant biology* 2009, **60**:485-510.
87. Katiyar-Agarwal S, Jin H: **Role of small RNAs in host-microbe interactions.** *Annual review of phytopathology* 2010, **48**:225-246.
88. Ding X, Cao Y, Huang L, Zhao J, Xu C, Li X, Wang S: **Activation of the indole-3-acetic acid-amido synthetase GH3-8 suppresses expansin expression and promotes salicylate- and jasmonate-independent basal immunity in rice.** *The Plant cell* 2008, **20**(1):228-240.
89. Yi H, Richards EJ: **A cluster of disease resistance genes in Arabidopsis is coordinately regulated by transcriptional activation and RNA silencing.** *The Plant cell* 2007, **19**(9):2929-2939.
90. Takken FL, Albrecht M, Tameling WI: **Resistance proteins: molecular switches of plant defence.** *Current opinion in plant biology* 2006, **9**(4):383-390.
91. Liu HH, Tian X, Li YJ, Wu CA: **Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana.** *RNA* 2008:836-843.
92. Sunkar R, Zhu JK: **Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis.** *The Plant cell* 2004, **16**(8):2001-2019.
93. Borkow G, Gabbay J: **Putting copper into action: copper-impregnated products with potent biocidal activities.** *The FASEB journal official publication of the Federation of American Societies for Experimental Biology* 2004, **18**(14):1728-1730.
94. Yuan M, Chu Z, Li X, Xu C, Wang S: **The bacterial pathogen Xanthomonas oryzae overcomes rice defenses by regulating host copper redistribution.** *The Plant cell* 2010, **22**(9):3164-3176.

95. Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, Frommer WB: **Sucrose efflux mediated by SWEET proteins as a key step for phloem transport.** *Science* 2012, **335**(6065):207-211.
96. Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, Guo WJ, Kim JG, Underwood W, Chaudhuri B *et al*: **Sugar transporters for intercellular exchange and nutrition of pathogens.** *Nature* 2010, **468**(7323):527-532.
97. Moscou MJ, Bogdanove AJ: **A Simple Cipher Governs DNA Recognition by TAL Effectors.** *Science* 2009, **326**(5959):1501-1501.
98. Rice P, Longden I, Bleasby A: **EMBOSS: the European Molecular Biology Open Software Suite.** *Trends in genetics : TIG* 2000, **16**(6):276-277.
99. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, Eddy SR, Gardner PP, Bateman A: **Rfam 11.0: 10 years of RNA families.** *Nucleic acids research* 2012.
100. Lorenz R, Bernhart SH, Honer Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, Hofacker IL: **ViennaRNA Package 2.0.** *Algorithms for molecular biology : AMB* 2011, **6**:26.
101. Ambros V: **A uniform system for microRNA annotation.** *RNA* 2003, **9**(3):277-279.
102. Jiang M, Anderson J, Gillespie J, Mayne M: **uShuffle: a useful tool for shuffling biological sequences while preserving the k-let counts.** *BMC bioinformatics* 2008, **9**:192.
103. Freyhult E, Gardner PP, Moulton V: **A comparison of RNA folding measures.** *BMC bioinformatics* 2005, **6**:241.
104. Megraw M, Hatzigeorgiou AG: **MicroRNA promoter analysis.** *Methods Mol Biol* 2010, **592**:149-161.
105. Jiang H, Wong WH: **SeqMap: mapping massive amount of oligonucleotides to the genome.** *Bioinformatics* 2008, **24**(20):2395-2396.
106. Salzman J, Jiang H, Wong WH: **Statistical Modeling of RNA-Seq Data.** *Statistical Science* 2011, **26**(1):62-83.
107. Fairhurst T, Dobermann A: **Rice in the global food supply.** *World* 2002, **5**(7,502):454,349-511,675.
108. Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X *et al*: **A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica).** *Science* 2002, **296**(5565):79-92.
109. Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H *et al*: **A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica).** *Science* 2002, **296**(5565):92-100.
110. EPPO: **Xanthomonas oryzae.** *EPPO Bulletin* 2007, **37**(3):543-553.
111. Mew T, Alvarez A, Leach J, Swings J: **Focus on bacterial blight of rice.** *Plant Disease* 1993, **77**(1):5-12.
112. Verdier V, Vera Cruz C, Leach JE: **Controlling rice bacterial blight in Africa: needs and prospects.** *Journal of biotechnology* 2012, **159**(4):320-328.
113. Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS *et al*: **The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331, the bacterial blight pathogen of rice.** *Nucleic acids research* 2005, **33**(2):577-586.
114. Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, Tsuge S, Furutani A, Ochiai H, Delcher AL, Kelley D *et al*: **Genome sequence and rapid evolution of the rice pathogen Xanthomonas oryzae pv. oryzae PXO99A.** *BMC genomics* 2008, **9**:204.
115. Ochiai H, Inoue Y, Takeya M, Sasaki A, Kaku H: **Genome sequence of Xanthomonas oryzae pv. oryzae suggests contribution of large numbers of effector genes and insertion sequences to its race diversity.** *Japan Agricultural Research Quarterly* 2005, **39**(4):275.
116. White FF, Yang B: **Host and pathogen factors controlling the rice-Xanthomonas oryzae interaction.** *Plant physiology* 2009, **150**(4):1677-1686.
117. White FF, Potnis N, Jones JB, Koebnik R: **The type III effectors of Xanthomonas.** *Molecular plant pathology* 2009, **10**(6):749-766.

118. Antony G, Zhou J, Huang S, Li T, Liu B, White F, Yang B: **Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-11N3.** *The Plant cell* 2010, **22**(11):3864-3876.
119. Yang B, Sugio A, White FF: **Os8N3 is a host disease-susceptibility gene for bacterial blight of rice.** *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(27):10503-10508.
120. Yuan T, Li X, Xiao J, Wang S: **Characterization of Xanthomonas oryzae-responsive cis-acting element in the promoter of rice race-specific susceptibility gene Xa13.** *Molecular plant* 2011, **4**(2):300-309.
121. Kay S, Hahn S, Marois E, Hause G, Bonas U: **A bacterial effector acts as a plant transcription factor and induces a cell size regulator.** *Science* 2007, **318**(5850):648-651.
122. Gonzalez C, Szurek B, Manceau C, Mathieu T, Sere Y, Verdier V: **Molecular and pathotypic characterization of new Xanthomonas oryzae strains from West Africa.** *Molecular plant-microbe interactions : MPMI* 2007, **20**(5):534-546.
123. Yu YH, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng JX, Verdier V, Szurek B: **Colonization of Rice Leaf Blades by an African Strain of Xanthomonas oryzae pv. oryzae Depends on a New TAL Effector That Induces the Rice Nodulin-3 Os11N3 Gene.** *Mol Plant Microbe In* 2011, **24**(9):1102-1113.
124. Vera Cruz CM, Bai J, Ona I, Leung H, Nelson RJ, Mew TW, Leach JE: **Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation.** *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**(25):13500-13505.
125. Iyer-Pascuzzi AS, McCouch SR: **Recessive resistance genes and the Oryza sativa-Xanthomonas oryzae pv. oryzae pathosystem.** *Molecular plant-microbe interactions : MPMI* 2007, **20**(7):731-739.
126. Li T, Liu B, Spalding MH, Weeks DP, Yang B: **High-efficiency TALEN-based gene editing produces disease-resistant rice.** *Nature biotechnology* 2012, **30**(5):390-392.
127. Li C, Wei J, Lin Y, Chen H: **Gene silencing using the recessive rice bacterial blight resistance gene xa13 as a new paradigm in plant breeding.** *Plant cell reports* 2012, **31**(5):851-862.
128. Lu C, Jeong D-H, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D *et al*: **Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs).** *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(12):4951-4956.
129. Sunkar R, Girke T, Jain PK, Zhu JK: **Cloning and characterization of microRNAs from rice.** *The Plant cell* 2005, **17**(5):1397-1411.
130. Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu JK: **Identification of novel and candidate miRNAs in rice by high throughput sequencing.** *BMC plant biology* 2008, **8**:25.
131. Li YF, Zheng Y, Addo-Quaye C, Zhang L, Saini A, Jagadeeswaran G, Axtell MJ, Zhang W, Sunkar R: **Transcriptome-wide identification of microRNA targets in rice.** *The Plant journal : for cell and molecular biology* 2010, **62**(5):742-759.
132. Guo W, Wu G, Yan F, Lu Y, Zheng H, Lin L, Chen H, Chen J: **Identification of Novel Oryza sativa miRNAs in Deep Sequencing-Based Small RNA Libraries of Rice Infected with Rice Stripe Virus.** *PloS one* 2012, **7**(10):e46443.
133. Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C: **A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains.** *Genome research* 2008, **18**(9):1456-1465.
134. Lv DK, Bai X, Li Y, Ding XD, Ge Y, Cai H, Ji W, Wu N, Zhu YM: **Profiling of cold-stress-responsive miRNAs in rice by microarrays.** *Gene* 2010, **459**(1-2):39-47.
135. Huang SQ, Peng J, Qiu CX, Yang ZM: **Heavy metal-regulated new microRNAs from rice.** *Journal of inorganic biochemistry* 2009, **103**(2):282-287.
136. Zhao B, Liang R, Ge L, Li W, Xiao H, Lin H, Ruan K, Jin Y: **Identification of drought-induced microRNAs in rice.** *Biochemical and biophysical research communications* 2007, **354**(2):585-590.

137. Jian X, Zhang L, Li G, Wang X, Cao X, Fang X, Chen F: **Identification of novel stress-regulated microRNAs from *Oryza sativa* L.** *Genomics* 2010, **95**(1):47-55.
138. Du P, Wu J, Zhang J, Zhao S, Zheng H, Gao G, Wei L, Li Y: **Viral infection induces expression of novel phased microRNAs from conserved cellular microRNA precursors.** *PLoS pathogens* 2011, **7**(8):e1002176.
139. Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X, Cai X *et al*: **Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA.** *Cell research* 2012, **22**(1):107-126.
140. Yuan J, He SY: **The *Pseudomonas syringae* Hrp regulation and secretion system controls the production and secretion of multiple extracellular proteins.** *Journal of bacteriology* 1996, **178**(21):6399-6402.
141. Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC: **Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA.** *Nucleic acids research* 2006, **34**(Database issue):D731-D735.
142. Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA: **PLEXdb: gene expression resources for plants and plant pathogens.** *Nucleic acids research* 2012, **40**(Database issue):D1194-1201.
143. Tao Y, Xie Z, Chen W, Glazebrook J, Chang HS, Han B, Zhu T, Zou G, Katagiri F: **Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*.** *The Plant cell* 2003, **15**(2):317-330.
144. Khraiweh B, Zhu JK, Zhu J: **Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants.** *Biochimica et biophysica acta* 2012, **1819**(2):137-148.
145. Rakwal R, Agrawal GK, Tamogami S, Iwahashi H: **Transcriptional profiling of OsHin1 in rice plants: a potential role in defense/stress and development.** *Plant Science* 2004, **166**(4):997-1005.
146. Kottapalli KR, Rakwal R, Satoh K, Shibato J, Kottapalli P, Iwahashi H, Kikuchi S: **Transcriptional profiling of indica rice cultivar IET8585 (Ajaya) infected with bacterial leaf blight pathogen *Xanthomonas oryzae* pv *oryzae*.** *Plant physiology and biochemistry : PPB / Societe francaise de physiologie vegetale* 2007, **45**(10-11):834-850.
147. Grewal RK, Gupta S, Das S: ***Xanthomonas oryzae* pv *oryzae* triggers immediate transcriptomic modulations in rice.** *BMC genomics* 2012, **13**:49.
148. Vaucheret H, Vazquez F, Cr  t   P, Bartel DP: **The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development.** *Genes & development* 2004, **18**(10):1187-1197.
149. V  rallyay E, V  l  czi A, Agyi A, Burgy  n J, Havelda Z: **Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation.** *The EMBO journal* 2010, **29**(20):3507-3519.
150. Levy J, Bres C, Geurts R, Chalhoub B, Kulikova O, Duc G, Journet EP, Ane JM, Lauber E, Bisseling T *et al*: **A putative Ca²⁺ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses.** *Science* 2004, **303**(5662):1361-1364.
151. Kim KC, Fan B, Chen Z: **Pathogen-induced *Arabidopsis* WRKY7 is a transcriptional repressor and enhances plant susceptibility to *Pseudomonas syringae*.** *Plant physiology* 2006, **142**(3):1180-1192.
152. Kim MC, Chung WS, Yun DJ, Cho MJ: **Calcium and calmodulin-mediated regulation of gene expression in plants.** *Molecular plant* 2009, **2**(1):13-21.
153. Langmead B: **Aligning short sequencing reads with Bowtie.** *Current protocols in bioinformatics / editorial board, Andreas D Baxevanis [et al]* 2010, **Chapter 11**:Unit 11 17.
154. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**(1):139-140.
155. Chappell L, Baulcombe D, Molnar A: **Isolation and cloning of small RNAs from virus-infected plants.** *Current protocols in microbiology* 2006, **Chapter 16**:Unit 16H 12.
156. Priest HD, Filichkin SA, Mockler TC: **Cis-regulatory elements in plant cell signaling.** *Current opinion in plant biology* 2009, **12**(5):643-649.

157. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR *et al*: **Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes**. *Science* 2000, **290**(5499):2105-2110.
158. Cui X, Xu SM, Mu DS, Yang ZM: **Genomic analysis of rice microRNA promoters and clusters**. *Gene* 2009, **431**(1-2):61-66.
159. Petrov V, Vermeirssen V, De Clercq I, Van Breusegem F, Minkov I, Vandepoele K, Gechev TS: **Identification of cis-regulatory elements specific for different types of reactive oxygen species in Arabidopsis thaliana**. *Gene* 2012, **499**(1):52-60.
160. Zhou X, Wang G, Zhang W: **UV-B responsive microRNA genes in Arabidopsis thaliana**. *Molecular systems biology* 2007, **3**:103.
161. Jia X, Ren L, Chen QJ, Li R, Tang G: **UV-B-responsive microRNAs in Populus tremula**. *Journal of plant physiology* 2009, **166**(18):2046-2057.
162. Ding D, Zhang L, Wang H, Liu Z, Zhang Z, Zheng Y: **Differential expression of miRNAs in response to salt stress in maize roots**. *Annals of botany* 2009, **103**(1):29-38.
163. Zhou L, Liu Y, Liu Z, Kong D, Duan M, Luo L: **Genome-wide identification and analysis of drought-responsive microRNAs in Oryza sativa**. *Journal of experimental botany* 2010, **61**(15):4157-4168.
164. Zhou X, Wang G, Sutoh K, Zhu JK, Zhang W: **Identification of cold-inducible microRNAs in plants by transcriptome analysis**. *Biochimica et biophysica acta* 2008, **1779**(11):780-788.
165. Joung JG, Fei Z: **Identification of microRNA regulatory modules in Arabidopsis via a probabilistic graphical model**. *Bioinformatics* 2009, **25**(3):387-393.
166. Croft L, Szklarczyk D, Jensen LJ, Gorodkin J: **Multiple independent analyses reveal only transcription factors as an enriched functional class associated with microRNAs**. *BMC systems biology* 2012, **6**:90.
167. Mejia-Guerra MK, Pomeranz M, Morohashi K, Grotewold E: **From plant gene regulatory grids to network dynamics**. *Biochimica et biophysica acta* 2012, **1819**(5):454-465.
168. Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E: **AGRIS: the Arabidopsis Gene Regulatory Information Server, an update**. *Nucleic acids research* 2011, **39**(Database issue):D1118-1122.
169. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S: **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences**. *Nucleic acids research* 2002, **30**(1):325-327.
170. Moyroud E, Minguet EG, Ott F, Yant L, Pose D, Monniaux M, Blanchet S, Bastien O, Thevenon E, Weigel D *et al*: **Prediction of regulatory interactions from genome sequences using a biophysical model for the Arabidopsis LEAFY transcription factor**. *The Plant cell* 2011, **23**(4):1293-1306.
171. Frazier TP, Sun G, Burklew CE, Zhang B: **Salt and drought stresses induce the aberrant expression of microRNA genes in tobacco**. *Molecular biotechnology* 2011, **49**(2):159-165.
172. Jung JH, Seo PJ, Kang SK, Park CM: **miR172 signals are incorporated into the miR156 signaling pathway at the SPL3/4/5 genes in Arabidopsis developmental transitions**. *Plant molecular biology* 2011, **76**(1-2):35-45.
173. Zhu QH, Helliwell CA: **Regulation of flowering time and floral patterning by miR172**. *Journal of experimental botany* 2011, **62**(2):487-495.
174. Wang ZY, Tobin EM: **Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression**. *Cell* 1998, **93**(7):1207-1217.
175. Hudson ME, Quail PH: **Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data**. *Plant physiology* 2003, **133**(4):1605-1616.
176. Roden LC, Ingle RA: **Lights, rhythms, infection: the role of light and the circadian clock in determining the outcome of plant-pathogen interactions**. *The Plant cell* 2009, **21**(9):2546-2552.

177. Meng Y, Huang F, Shi Q, Cao J, Chen D, Zhang J, Ni J, Wu P, Chen M: **Genome-wide survey of rice microRNAs and microRNA-target pairs in the root of a novel auxin-resistant mutant.** *Planta* 2009, **230**(5):883-898.
178. Tilly JJ, Allen DW, Jack T: **The CArG boxes in the promoter of the Arabidopsis floral organ identity gene APETALA3 mediate diverse regulatory effects.** *Development* 1998, **125**(9):1647-1657.
179. Li Y, Swaminathan K, Hudson ME: **Rapid, organ-specific transcriptional responses to light regulate photomorphogenic development in dicot seedlings.** *Plant physiology* 2011, **156**(4):2124-2140.
180. Perisic O, Lam E: **A tobacco DNA binding protein that interacts with a light-responsive box II element.** *The Plant cell* 1992, **4**(7):831-838.
181. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA: **Conservation and divergence of plant microRNA genes.** *The Plant journal : for cell and molecular biology* 2006, **46**(2):243-259.
182. Zhang C, Xuan Z, Otto S, Hover JR, McCorkle SR, Mandel G, Zhang MQ: **A clustering property of highly-degenerate transcription factor binding sites in the mammalian genome.** *Nucleic acids research* 2006, **34**(8):2238-2246.
183. Spangler JB, Ficklin SP, Luo F, Freeling M, Feltus FA: **Conserved non-coding regulatory signatures in Arabidopsis co-expressed gene modules.** *PloS one* 2012, **7**(9):e45041.
184. Ho Sui SJ, Mortimer JR, Arenillas DJ, Brumm J, Walsh CJ, Kennedy BP, Wasserman WW: **oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes.** *Nucleic acids research* 2005, **33**(10):3154-3164.
185. Frith MC, Fu Y, Yu L, Chen JF, Hansen U, Weng Z: **Detection of functional DNA motifs via statistical over-representation.** *Nucleic acids research* 2004, **32**(4):1372-1381.
186. Reimers M, Carey VJ: **Bioconductor: an open source framework for bioinformatics and computational biology.** *Methods in enzymology* 2006, **411**:119-134.
187. Scholze H, Boch J: **TAL effectors are remote controls for gene activation.** *Current opinion in microbiology* 2011, **14**(1):47-53.
188. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U: **Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors.** *Science* 2009, **326**(5959):1509-1512.
189. Murakami MT, Sforca ML, Neves JL, Paiva JH, Domingues MN, Pereira AL, Zeri AC, Benedetti CE: **The repeat domain of the type III effector protein PthA shows a TPR-like structure and undergoes conformational changes upon DNA interaction.** *Proteins* 2010, **78**(16):3386-3395.
190. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N: **Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors.** *Science* 2012, **335**(6069):720-723.
191. Bradley P: **Structural modeling of TAL effector-DNA interactions.** *Protein Science* 2012, **21**(4):471-474.
192. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL: **The crystal structure of TAL effector PthXo1 bound to its DNA target.** *Science* 2012, **335**(6069):716-719.
193. Bogdanove AJ, Voytas DF: **TAL effectors: customizable proteins for DNA targeting.** *Science* 2011, **333**(6051):1843-1846.
194. Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, Zhu JK: **Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein.** *Plant molecular biology* 2012, **78**(3):311-321.
195. Mahfouz MM, Li L: **TALE nucleases and next generation GM crops.** *GM crops* 2011, **2**(2):99-103.
196. Sun N, Liang J, Abil Z, Zhao H: **Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease.** *Molecular bioSystems* 2012, **8**(4):1255-1263.
197. Rodriguez RL, Grajales A, Arrieta-Ortiz M, Salazar C, Restrepo S, Bernal A: **Genomes-based phylogeny of the genus Xanthomonas.** *BMC microbiology* 2012, **12**(1):43.

198. Sugio A, Yang B, Zhu T, White FF: **Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OstFIIAgamma1* and *OstFX1* during bacterial blight of rice.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(25):10720-10725.
199. Kay S, Hahn S, Marois E, Wieduwild R, Bonas U: **Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors *AvrBs3* and *AvrBs3Δrep16*.** *The Plant Journal* 2009, **59**(6):859-871.
200. Romer P, Recht S, Lahaye T: **A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens.** *Proceedings of the National Academy of Sciences* 2009, **106**(48):20526-20531.
201. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ: **TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction.** *Nucleic acids research* 2012.
202. Finn RD, Clements J, Eddy SR: **HMMER web server: interactive sequence similarity searching.** *Nucleic acids research* 2011, **39**(Web Server issue):W29-37.
203. Romer P, Recht S, Strauss T, Elsaesser J, Schornack S, Boch J, Wang S, Lahaye T: **Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*.** *The New phytologist* 2010, **187**(4):1048-1057.
204. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P: **Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription.** *Nature biotechnology* 2011, **29**(2):149-153.
205. Tian D, Yin Z: **Constitutive heterologous expression of *avrXa27* in rice containing the R gene *Xa27* confers enhanced resistance to compatible *Xanthomonas oryzae* strains.** *Molecular plant pathology* 2009, **10**(1):29-39.
206. Hummel AW, Doyle EL, Bogdanove AJ: **Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak.** *The New phytologist* 2012, **195**(4):883-893.
207. Du Z, Zhou X, Ling Y, Zhang Z, Su Z: **agriGO: a GO analysis toolkit for the agricultural community.** *Nucleic acids research* 2010, **38**(Web Server issue):W64-70.
208. Lee I, Seo YS, Coltrane D, Hwang S, Oh T, Marcotte EM, Ronald PC: **Genetic dissection of the biotic stress response using a genome-scale gene network for rice.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(45):18548-18553.
209. Mukhtar MS, Carvunis AR, Dreze M, Epple P, Steinbrenner J, Moore J, Tasan M, Galli M, Hao T, Nishimura MT *et al*: **Independently evolved virulence effectors converge onto hubs in a plant immune system network.** *Science* 2011, **333**(6042):596-601.
210. Yang Z, Ebright YW, Yu B, Chen X: **HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide.** *Nucleic acids research* 2006, **34**(2):667-675.
211. Abramovitch RB, Janjusevic R, Stebbins CE, Martin GB: **Type III effector *AvrPtoB* requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity.** *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(8):2851-2856.
212. Craig A, Ewan R, Mesmar J, Gudipati V, Sadanandom A: **E3 ubiquitin ligases and plant innate immunity.** *Journal of experimental botany* 2009, **60**(4):1123-1132.
213. Streubel J, Blucher C, Landgraf A, Boch J: **TAL effector RVD specificities and efficiencies.** *Nature biotechnology* 2012, **30**(7):593-595.
214. Gautier L, Cope L, Bolstad BM, Irizarry RA: **affy--analysis of Affymetrix GeneChip data at the probe level.** *Bioinformatics* 2004, **20**(3):307-315.
215. Smyth G: **limma: Linear Models for Microarray Data** *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. In. Edited by Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S: Springer New York; 2005: 397-420.

-
216. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: **Cytoscape 2.8: new features for data integration and network visualization.** *Bioinformatics* 2011, **27**(3):431-432.
 217. Perez-Quintero AL, Sablok G, Tatarinova TV, Conesa A, Kuo J, Lopez C: **Mining of miRNAs and potential targets from gene oriented clusters of transcripts sequences of the anti-malarial plant, *Artemisia annua*.** *Biotechnology letters* 2011.