

**CLONACIÓN, SECUENCIACION Y CARACTERIZACIÓN
BIOINFORMATICA DE CANDIDATOS A microRNAs DE YUCA
(*Manihot esculenta* Crantz Var. Tai 16)**

ANA CAROLINA BALLEEN TABORDA

**TRABAJO DE GRADO
Presentado como requisito parcial
Para optar al título de**

BIÓLOGA.

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE CIENCIAS
CARRERA DE BIOLOGÍA
Bogotá, D.C.
Julio de 2009**

NOTA DE ADVERTENCIA

Artículo 23 de la Resolución N° 13 de Julio de 1946

“La Universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de tesis. Solo velará por qué no se publique nada contrario al dogma y a la moral católica y por que las tesis no contengan ataques personales contra persona alguna, antes bien se vea en ellas el anhelo de buscar la verdad y la justicia”.

**CLONACIÓN, SECUENCIACION Y CARACTERIZACIÓN
BIOINFORMATICA DE CANDIDATOS A microRNAs DE YUCA
(*Manihot esculenta* Crantz Var. Tai 16)**

ANA CAROLINA BALLEEN TABORDA

APROBADO

Dr. Joe Tohme. Ph.D

Director

Dr. Wilson Terán. Ph.D

Codirector

Fausto Rodríguez, Biólogo

Codirector

Sonia Cuellar

Jurado

Silvia Restrepo

Jurado

**CLONACIÓN, SECUENCIACION Y CARACTERIZACIÓN
BIOINFORMATICA DE CANDIDATOS A microRNAs DE YUCA
(*Manihot esculenta* Crantz Var. Tai 16)**

ANA CAROLINA BALLEEN TABORDA

APROBADO

Dr. Ingrid Shuler, Ph.D
Decana Académica

Dr. Andrea Forero, Bióloga
Directora Carrera de Biología

DEDICATORIA

A Dios, a la Virgen y a la vida por darme la oportunidad de realizar mi trabajo de grado en el Centro Internacional de Agricultura Tropical, donde me he formado como persona y profesional de la ciencia. A mis padres por su amor y apoyo en este proceso de formación académica. A mis hermanas por su amistad incondicional. A mis amigos y compañeros de trabajo.

AGRADECIMIENTOS

A mi familia y amigos por su amistad, apoyo y amor incondicional.

A la Pontificia Universidad Javeriana por la formación académica y personal brindada durante mi carrera.

Al Dr. Joe Tohme. Líder del proyecto de agrobiodiversidad y biotecnología del Centro Internacional de Agricultura Tropical (CIAT) por la oportunidad, confianza y apoyo que me brindó durante el desarrollo de mi trabajo de grado y sobre todo en mi formación como profesional de la ciencia.

Al Dr. Wilson Terán. De la unidad de biotecnología vegetal de la Pontificia Universidad Javeriana. Por su labor en la codirección del Proyecto y asesoría en el desarrollo de mi trabajo de grado.

A la Dr. Diana Álvarez por la revisión del anteproyecto.

Germán Plata y Fausto Rodríguez. Asistentes de investigación del CIAT. Por su asistencia y compromiso y apoyo en el desarrollo de los procedimientos de laboratorio y conocimientos de bioinformática.

Al Dr. Ibrokhim Abdurakhmonov. Por su asesoría y dirección en la construcción de librerías de microRNAs.

A toda la unidad de biotecnología del CIAT. Adriana Bohórquez, Diana Bernal, Federico Roda, Jesús Beltrán, Marines Giraldo, Harold Suarez, Juliana Chacón, Sebastián Parra, Andrés Gutiérrez, Yamid Sanabria y Alicia V. por su amistad y colaboración en el desarrollo de los procedimientos de laboratorio.

TABLA DE CONTENIDOS

1.	INTRODUCCIÓN.....	1
2.	MARCO TEÓRICO Y REVISIÓN DE LITERATURA.....	2
2.1.	<i>Manihot esculenta</i> Crantz	2
2.2.	SmallRNAs y MicroRNAs	3
2.2.1.	Biogénesis	4
2.2.2.	MicroRNAs en plantas.....	6
2.2.3.	Promotores de microRNAs	8
2.2.4.	Antecedentes: Análisis experimentales y bioinformáticos.....	9
3.	FORMULACIÓN DEL PROBLEMA, PREGUNTA Y JUSTIFICACIÓN.....	12
3.1.	Formulación del problema	12
3.2.	Pregunta de investigación	13
3.3.	Justificación de la investigación.....	13
4.	OBJETIVOS	14
4.1.	General.....	14
4.2.	Específicos	14
5.	MATERIALES Y MÉTODOS	14
5.1.	Diseño de la Investigación	14
5.1.1.	Población de estudio y muestra de estudio.....	15
5.1.1.1.	Población de estudio.....	15
5.1.1.2.	Muestra de estudio	15
5.2.	Tratamientos	15
5.3.	Procedimiento de laboratorio	16
5.3.1.	Extracción de RNAs total y miRNAs con el kit <i>mirVANA</i>	16
5.3.2.	Cuantificación de las muestras	17
5.3.3.	Aislamiento de los microRNA a partir de gel de acrilamida.....	17
5.3.4.	Construcción, clonación y secuenciación de librerías de microRNAs mediante kit MiRCat siguiendo las pautas del protocolo de Bartel (Ver figura 2) (Lau <i>et al.</i> 2001, Lu <i>et al.</i> 2007).	18
5.3.4.1.	Ligación de adaptador en el extremo 3' de los microRNAs	19
5.3.4.2.	Ligación de adaptador en el extremo 5' de los microRNAs	19
5.3.4.3.	Síntesis de cDNA	20

5.3.4.4. Síntesis de segunda cadena (PCR) y cuantificación.....	20
5.3.4.5. Ligación de los amplicones al vector pGEMT-easy.....	20
5.3.4.6. Transformación de células bacterianas y PCR.....	21
5.3.5. Secuenciación del producto de PCR de los clones	21
5.4. Análisis bioinformático	22
5.4.1. Edición de secuencias en Sequencher 4.6.	22
5.4.2. Alineamiento de secuencias	23
5.4.3. Comparación con bases de datos.	23
5.4.3.1. BLAST vs. MIRBASE.	24
5.4.3.2. BLASTN vs. ESTs de <i>Manihot esculenta</i>	24
5.4.3.3. Recuperación de secuencias de ESTs.....	24
5.4.3.4. BLAST ESTs vs. GenBank de NCBI.....	24
5.4.3.5. BLAST vs. RNA ribosomal.	25
5.4.3.6. BLASTN vs. Genoma de <i>Ricinus communis</i>	25
5.4.3.7. Recuperación de las secuencias flanqueantes a partir de los hits con <i>Ricinus communis</i> y predicción de estructuras secundarias.....	25
5.4.4. Análisis de promotores tipo II.	26
5.4.5. Predicción de potenciales genes diana y sus funciones.....	26
6. RESULTADOS	27
6.1. Procedimiento de laboratorio	27
6.1.1. Extracción de RNAs total y miRNAs con el kit <i>mirVANA</i>	27
6.1.2. Cuantificación de las muestras	28
6.1.3. Aislamiento de los smallRNA a partir de gel de acrilamida	29
6.1.4. Construcción, clonación y secuenciación de librerías de smallRNAs.	30
6.1.4.1. Ligación de adaptador en el extremo 3' de los smallRNAs.....	30
6.1.4.2. Ligación de adaptador en el extremo 5' de los smallRNAs.....	31
6.1.4.3. Síntesis de segunda cadena (PCR) y cuantificación.....	31
6.2. Análisis bioinformático	33
6.2.1. Edición de secuencias en Sequencher 4.6.	33
6.2.2. Alineamiento de secuencias	35
6.2.3. Comparación con bases de datos.	36
6.2.3.1. BLAST vs. MIRBASE.	36

6.2.3.2. BLAST vs. ESTs <i>Manihot esculenta</i> y BLAST ESTs (secuencia completa) vs. NCB.....	36
6.2.3.3. BLAST potenciales microRNAs vs. RNA ribosomal.....	37
6.2.3.4. BLAST vs. Genoma de <i>Ricinus communis</i>	38
6.2.3.5. Recuperación de las secuencias flanqueantes a partir de los hits con <i>Ricinus communis</i> y predicción de estructuras secundarias.....	39
6.2.4. Análisis de promotores.....	40
6.2.5. Predicción de potenciales genes diana y sus funciones.....	41
7. DISCUSIÓN	42
7.1. RNA utilizado, ligaciones y PCR.....	42
7.2. Secuencias obtenidas	44
7.3. MicroRNAs y factores que afectan su clonación	45
7.4. E-values para resultados de BLAST.....	48
7.5. Comparaciones.....	49
7.6. Predicción de estructuras secundarias	52
7.7. Análisis de promotores.....	53
7.8. Potenciales genes diana	54
7.9. Funciones de los genes regulados (Ver tabla 9).....	55
8. CONCLUSIONES.....	56
9. RECOMENDACIONES Y TRABAJO FUTURO.....	57
10. REFERENCIAS BIBLIOGRAFICAS	58
11. ANEXOS	68

ÍNDICE DE TABLAS

Tabla 1: Lecturas de absorvancias a 260 nm y 280 nm, concentraciones y pureza del RNA total aislado a partir de plantas <i>in-vitro</i> y plantas de campo.....	28
Tabla 2: Concentración, volumen y cantidad del PCR cortado y purificado.	32
Tabla 3: Secuencias de potenciales microRNAs de yuca generadas con la librería.....	34
Tabla 4: Blast de los potenciales microRNAs de yuca con microRNAs de otras especies registrados en MIRBASE.....	36
Tabla 5: Blast de los potenciales microRNAs con ESTs de <i>Manihot esculenta</i> y Blast de la secuencia completa de los ESTs con la base de datos (nr) de NCBI.....	37
Tabla 6: Blast de los potenciales microRNAs de yuca con la base de datos de RNA ribosomal Universidad de Ghent	37
Tabla 7: Blast de los potenciales microRNAs de yuca con el genoma de <i>Ricinus communis</i>	38
Tabla 8: Elementos promotores de genes tipo II.....	40
Tabla 9: Proteínas posiblemente afectadas obtenidas a partir de la comparación de los ESTs en BLASTX de NCBI.....	42

INDICE DE FIGURAS

Fig. 1: Biogénesis de microRNAs a partir de genes y de intrones.....	6
Fig. 2: Procedimiento para generar y secuenciar una librería de microRNAs.....	18
Fig. 3: Procedimiento para editar y analizar las secuencias de microRNAs secuenciados y sus funciones en regulación.....	22
Fig. 4: RNA total de <i>Manihot esculenta</i> sembrado en gel de agarosa 1.5%.....	27
Fig. 5: RNA total de <i>Manihot esculenta</i> sembrado en gel de acrilamida 15%.....	29
Fig. 6: MicroRNAs ligados en su extremo 3' a un adaptador, sembrados en un gel de acrilamida 15%.....	30
Fig. 7: microRNAs ligados en su extremo 3' y 5' a adaptadores, sembrados en un gel de acrilamida 15%.....	31
Fig. 8: microRNAs amplificados con primers 454.	31
Fig. 9: Curva estándar de DNA (0.02 ng/ul a 1 ng/ul).	32
Fig. 10: Alineamiento o ensamblaje de los potenciales microRNAs utilizando MacClade.	35
Fig. 11: Estructuras secundarias obtenidas en mFold.	39
Fig. 12: Potenciales promotores de genes tipo II de la secuencia anterior.	40
Fig. 13: ESTs probablemente regulados por el único microRNA secuenciado.	41
Fig. 14: Distribución de longitudes de smallRNAs de <i>Populus trichocarpa</i>	45
Fig. 15: Distribución de frecuencias de smallRNAs de de <i>O. Sativa</i> y <i>Pinus contort</i>	45
Fig. 16: RNA total sembrado el gel de acrilamida de 24 especies de plantas diferente.....	46
Fig. 17: Distribución de frecuencias de smallRNAs de Yuca secuenciados.	47

ÍNDICE DE ANEXOS

Anexo 1: Extracción de RNA total con el kit mirVANA™ miRNA Isolation.....	68
Anexo 2: Procedimiento adicional para extracción de RNA total de raíces de plantas de campo.....	69
Anexo 3: Gel de acrilamida denaturante al 15% (urea 7M) y aislamiento de microRNAs.....	69
Anexo 4: Purificación de miRNAs, ligación 3', ligación 5' y PCR a partir de un método estándar de macerado y remojo "Crush and Soak" de gel de acrilamida.....	70
Anexo 5: Construcción y clonación de librería de microRNAs.....	70
Anexo 6: Cuantificación de la concentración.....	75
Anexo 7: Protocolo para secuenciación.....	75

RESUMEN

Los microRNAs (miRNAs) son smallRNAs endógenos que regulan la expresión de genes en plantas y animales. Estos RNAs de ~21 nucleótidos son sintetizados a partir de secuencias intrónicas o a partir de genes de microRNAs por la RNA pol II, cuyo precursor forma una estructura tallo y asa, el cual es a su vez procesado por la enzima Dicer 1 hasta secuencias más cortas que son reclutadas por el complejo RISC, el cual dirige la degradación de mRNAs blanco. Estos pequeños transcritos han sido descubiertos mediante predicciones bioinformáticas o mediante su clonación y secuenciación, y se ha visto que los miRNAs en plantas están involucrados en múltiples procesos biológicos, incluyendo diferenciación celular, desarrollo de órganos, cambios de fases, señalización, respuesta a estrés biótico y abiótico, entre otros. Por otro lado, la yuca (*Manihot esculenta* Crantz) es considerada como uno de los cultivos más importantes en las regiones tropicales del mundo. En este sentido, este estudio representa el primer esfuerzo por identificar miRNAs en yuca a través de una aproximación experimental (clonación y secuenciación) y posterior análisis bioinformático, en el cual se obtuvieron secuencias en 6 familias diferentes, una de las cuales, fue validada como candidato a microRNA basada en el genoma de *Ricinus communis*. Esta secuencia puede ser un transcrito cuyo origen es a partir de transposones, por lo que se explicaría el porqué está presente en la mayoría de los clones secuenciados y podría estar involucrada con procesos asociados a la fotosíntesis.

Palabras claves:

MicroRNAs, Desarrollo y crecimiento de plantas, Regulación de genes, Silenciamiento de RNA.

ABSTRACT

MicroRNAs (miRNAs) are endogenous smallRNAs that regulate gene expression in plants and animals. These RNAs of ~21 nucleotides are synthesized from intron sequences or microRNA's genes for the RNA pol II, which form a precursor stem-loop structure, which is in turn processed by the enzyme Dicer 1 to shorter sequences who are loaded into silencing complex RISC, where directs the degradation of target mRNAs. MiRNAs have been discovered using two basic approaches: bioinformatic prediction and direct cloning and sequencing, and lots of investigations indicate that are involved in multiple biological processes, including stem cell differentiation, organ development, phase change, signalling and response to biotic and abiotic environmental stresses. In other hand, cassava (*Manihot esculenta* Crantz) is considered one of the most important crops in tropical regions of the world. Thus, this study represents the first effort to identify miRNAs in cassava through an experimental approach (cloning and sequencing) and bioinformatic analysis, in which were obtained sequences in 6 different families, but only one, was validated as a candidate microRNA based on *Ricinus communis* genome. This sequence could be transcribed from transposons, which explain why it is present in the majority of clones sequenced and this sequence might be involved in processes associated with photosynthesis.

Key words:

MicroRNAs, plant development and growth, gene regulation, RNA silencing.

1. INTRODUCCIÓN

Desde el momento en que Francis Crick en 1958 formuló el dogma central de la biología molecular, que describe la transmisión y expresión de la información genética, es decir, que el DNA se transcribe a RNA mensajero y este a su vez es traducido a proteínas (Crick 1970) se observó que en cada nivel de expresión existían mecanismos de regulación génica: a nivel transcripcional, post-transcripcional y traduccional. A nivel posttranscripcional se creía que la regulación génica solo estaba asociada a las modificaciones en el RNA mensajero, pero a partir de 1993, se observó que cientos de transcritos pequeños podían ser factores muy importantes en la regulación génica a este nivel. Estas especies de transcritos denominados microRNAs, que fueron inicialmente descubiertos en *Caenorhabditis elegans*, miden alrededor de 20 nucleótidos (nt) de longitud y regulan importantes procesos en plantas y animales (Jones-Rhoades *et al.* 2006) debido a que producen el silenciamiento de la expresión génica mediante la interacción con secuencias complementarias en el RNA mensajero blanco (Zhang *et al.* 2006).

Recientes estudios han revelado el importante rol de los microRNAs en diversos procesos en las plantas, tales como crecimiento y desarrollo (Lee *et al.* 2004). Amén de lo anterior, el desconocimiento que aún tenemos sobre estos transcritos y sus funciones, en especial en plantas cultivadas (Lee *et al.* 2003, Bohnsack *et al.* 2004, Lund *et al.* 2004, Krol *et al.* 2006), exige que se desarrollen más trabajos como estos para identificar secuencias y en trabajos futuros caracterizarlas funcionalmente.

En este sentido, en Colombia este estudio representa un esfuerzo por identificar candidatos a miRNAs en yuca a través de una aproximación experimental (clonación y secuenciación de miRNAs). La aproximación experimental es necesaria debido a que el genoma de la yuca no se encuentra disponible aún, complicando la

identificación de microRNAs propios de la yuca mediante métodos computacionales. Debido a lo mencionado, se usará el genoma de *Ricinus communis* y los ESTs (expressed sequence tags) (Zhang *et al.* 2005) disponibles para esta variedad de yuca obtenidos por Sakurai *et al.* (2007).

Así pues, el objetivo del presente trabajo es realizar la extracción, clonación y secuenciación de los microRNAs expresados en plántulas de yuca cultivadas *in-vitro* y en campo. Las primeras, sometidas a tres condiciones diferentes: normal, calor y sequía. Con el fin de caracterizarlos bioinformáticamente, esto es, determinar su similitud con miRNAs reportados para otras especies de plantas, calcular sus estructuras secundarias, encontrar sus elementos promotores y predecir sus posibles roles en regulación de genes.

2. MARCO TEÓRICO Y REVISIÓN DE LITERATURA

2.1. *Manihot esculenta* Crantz

La yuca (*Manihot esculenta* Crantz) es un arbusto leñoso perenne que pertenece a la familia Euphorbiaceae. A esta familia pertenece *Ricinus communis*, la planta mas cercanamente emparentada a la yuca (Ceballos 2002) con el genoma secuenciado, en base a las cuales se realizó la validación de las secuencias obtenidas para yuca.

La yuca es originaria del continente americano y hay evidencias arqueológicas que indican que ha sido cultivada desde hace aproximadamente 5000 años en una extensa región geográfica comprendida entre México y el sur de Brasil. Esta especie es de gran importancia socioeconómica para los agricultores y consumidores en el mundo a nivel del trópico, ocupando el cuarto lugar entre los alimentos con más contenido de calorías en el mundo, después del arroz, la caña de azúcar y el maíz (Álvarez 1972, Domínguez *et al.* 1979).

Esta planta se caracteriza porque tiene una gran producción de carbohidratos y su valor nutricional es muy alto. Tanto sus raíces como sus hojas pueden ser utilizadas para el consumo humano, alimentación animal y adicionalmente para la industria. Además del valor económico y productivo, este cultivo se destaca por adaptarse a una gran diversidad de ambientes y su tolerancia a factores bióticos y abióticos adversos (Ceballos 2002).

2.2. SmallRNAs y MicroRNAs

Los smallRNAs son secuencias de RNA cortas de 15 a 35 nt de longitud, importantes en la regulación post-transcripcional. Hasta el momento se distinguen los microRNAs y tres clases de smallRNAs de interferencia tales como trans-acting siRNA (ta-siRNA), heterochromatic siRNA (hc-siRNA) y natural antisense siRNA (nat-siRNA) (Phillips *et al.* 2007).

Los microRNAs (miRNAs) son una clase de smallRNAs no codificantes de cadena sencilla, de aproximadamente 19 a 25 nt de longitud, generados a partir de un transcrito endógeno que incluye una estructura en tallo y asa. La mayoría de los miRNAs están conservados entre especies relacionadas y juegan importantes roles de regulación en animales y plantas, reprimiendo la expresión de genes que codifican a proteínas o a factores de transcripción (Bartel 2004, Jones-Rhoades *et al.* 2006). Muchos precursores de miRNAs de animales son aproximadamente de 70 nt de longitud mientras que en plantas las longitudes de los precursores varían ampliamente, algunos se extienden hasta los 300 nt (Wang *et al.* 2004b).

Existen dos tipos de RNAs que podrían confundirse los unos con los otros: los MicroRNAs y los siRNAs, los dos representan una familia de RNAs pequeños que son críticos en la regulación de genes, ambos tienen de 19 a 25 nt de longitud, están

sujetos a un procesamiento similar mediante DICER y se unen al mismo complejo de silenciamiento para regular los mRNAs. Estos transcritos difieren en la naturaleza de sus precursores: los siRNAs son procesados de RNAs de doble cadena y los microRNAs a partir de RNAs de cadena sencilla que forman una estructura secundaria en tallo y asa (Jones-Rhoades *et al.* 2006). MiRNAs y siRNAs también se diferencian por su origen genómico y la naturaleza de sus genes diana (Dugas *et al.* 2004); los siRNAs derivan de sus genes diana en sí mismos, si dicho gen diana es un transgen, un virus o un transposon, por el contrario, en el caso de los microRNAs los genes diana son aquellos genes propios que codifican a mRNAs o factores de transcripción (Dugas *et al.* 2004, Naqvi *et al.* 2009). Por último, tenemos la diferencia relacionada a su conservación evolutiva (Dugas *et al.* 2004), esto es, que muchos microRNAs están conservados entre especies relacionadas mientras que los siRNAs no lo están (Kidner *et al.* 2003, Jones-Rhoades *et al.* 2006).

Aunque, muchas familias de miRNAs conocidos, están conservados entre las especies de plantas (Bartel 2004), la gran mayoría de estos transcritos son encontrados como únicos para cada especie (Jones-Rhoades *et al.* 2006).

2.2.1. Biogénesis

La biogénesis de los miRNA presenta varios pasos (Ver figura 1). El primer paso ocurre en el núcleo, donde la RNA polimerasa II sintetiza a partir de genes de microRNAs una estructura secundaria de aproximadamente 60 - 80 nt de longitud, denominada pri-miRNA, precursor de microRNA con cap en 5' y una cola poli A en el extremo 3' (Bonnet *et al.* 2006).

Secuencialmente, la endonucleasa RNAsa de tipo III DICER (DCL1) remueve la región en tallo y asa de los pri-miRNAs para liberar transcritos más pequeños (pre-miRNA), que son procesados nuevamente por DCL1 para generar un dúplex de

microRNA-microRNA* que es exportado al citoplasma por la proteína HASTY (exportina) (Zhang *et al.* 2007).

Posteriormente, la secuencia complementaria al microRNA denominado microRNA* es rápidamente degradado y luego el miRNA maduro es incorporado a un complejo proteico RISC (RNA induced silencing complex), el cual se encarga de guiar a este complejo hacia mRNAs diana que son complementarios. Eventualmente, RISC+microRNA inhibe la traducción afectando directamente a los genes o a aquellos que codifican a factores de transcripción (Bartel 2004). En la actualidad se conoce que cerca del 50 % de los genes diana de los miRNAs son factores de transcripción (Bartel 2004, Jones-Rhoades *et al.* 2006, Zhang *et al.* 2006, Yang *et al.* 2007, Yin *et al.* 2008).

Según Ruby *et al.* (2007) y Naqvi *et al.* (2009) existe una ruta alternativa (Ver figura 1) para generar microRNAs en plantas, en el cual ciertos intrones que presentan las mismas características estructurales de los pre-microRNAs, son reconocidos por DICER 1 para generar pre-miRNAs, a los cuales se les denominan “mirtrons” (Ruby *et al.* 2007, Naqvi *et al.* 2009).

Existen al menos 4 tipos de enzimas DICER (DCL) en *Arabidopsis thaliana* funcionales que se expresan en respuesta a estrés ambiental y en condiciones normales en todos los tejidos pero diferencialmente entre estos durante el desarrollo. DCL 1 se encarga solamente de producir microRNAs (21 nt a 24 nt) a partir de estructuras en tallo y asa y a partir de secuencias repetidas o transposones (Liu *et al.* 2009). Por otro lado, DCL2, DCL3 y DCL4 generan diferentes tipos de siRNAs (RNAs de interferencia cortos) de 22, 24 y 21 nt respectivamente (Dolgosheina *et al.* 2008, Liu *et al.* 2009, Naqvi *et al.* 2009).

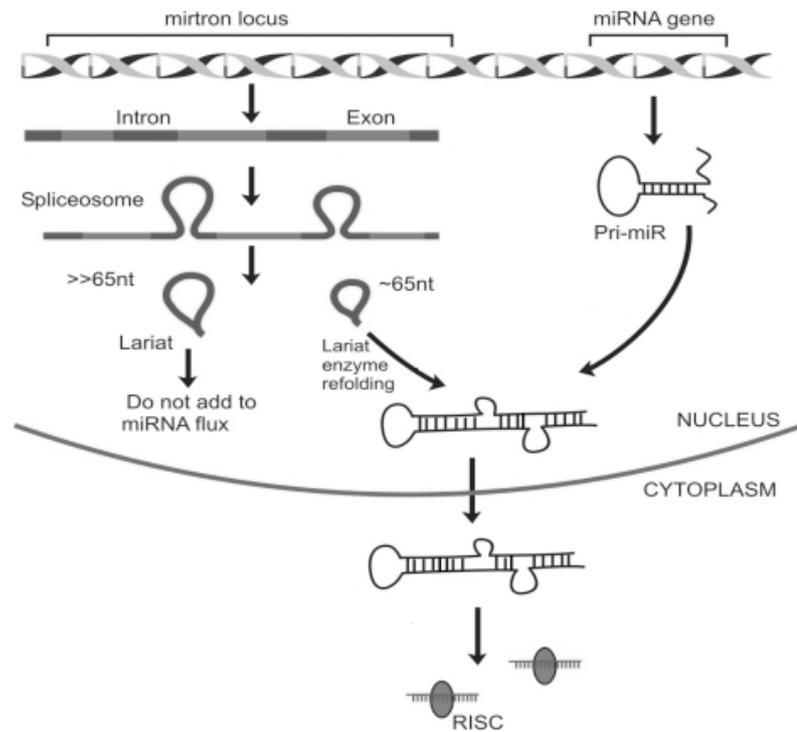


Fig. 1: Biogénesis de microRNAs a partir de genes y a partir de intrones (Naqvi *et al.* 2009).

2.2.2. MicroRNAs en plantas

Gracias a que miRNAs y mRNA en plantas, a diferencia de los animales, tienen una perfecta complementariedad, esta característica permite una fácil, rápida y confiable predicción de los genes mediante métodos computacionales (Reinhart *et al.* 2002, Jones-Rhoades *et al.* 2006). Si la complementariedad entre el miRNA y su mRNA blanco no es del 100 %, el mRNA es físicamente inalcanzable por el centro activo de la endonucleasa (slicer) de RISC para ser degradado, pero el silenciamiento es eficiente ya que se forma una horquilla entre el blanco y el miRNA, generando una represión traduccional en lugar de una digestión del mRNA diana (Jones-Rhoades *et al.* 2006).

Los microRNAs pueden tener diferentes patrones de expresión en diferentes tejidos, estados de desarrollo o bajo estímulos de estrés (abiótico, biótico, oxidativo, mecánico, de nutrientes, deshidratación), es decir, que se modula la expresión de genes de acuerdo a los cambios en el desarrollo o en el ambiente (Zhang *et al.* 2007). La máxima diversidad de microRNAs expresados, se da en órganos en desarrollo y en las flores (Rhoades *et al.* 2002, Sunkar *et al.* 2004, Axtell *et al.* 2005, Jones-Rhoades *et al.* 2006, Yao *et al.* 2007).

En general, los microRNAs participan en aspectos en el desarrollo, integridad del genoma y crecimiento de los órganos vegetales, incluyendo raíces, tallos, brotes y flores (Axtell *et al.* 2005, Jones-Rhoades *et al.* 2006, Zhang *et al.* 2007). Para dar un ejemplo, estos participan en procesos tales como: formación de los meristemos apicales y axilares (miR164), desarrollo de embriones, desarrollo vascular, identidad de los órganos florales (miR165), tiempo de floración (la transición de las plantas del crecimiento vegetativo a crecimiento reproductivo) (Zhang *et al.* 2007), organogénesis y morfogénesis de hojas (miR165), polaridad de los órganos, formación de raíces laterales y adventicias, fertilidad, entre otros (Rhoades *et al.* 2002, Sunkar *et al.* 2004, Jones-Rhoades *et al.* 2006, Yao *et al.* 2007, Zhang *et al.* 2007).

En el caso específico del desarrollo vegetativo de las plantas se ha observado que los miRNAs, como miR160 y miR167 (Rhoades *et al.* 2002), regulan componentes claves en la señalización de la hormona auxina. En estos casos, se ha visto que bajo deficiencia de miRNAs se presentan anomalías en el desarrollo de raíces y otros órganos (Zhang *et al.* 2006). Esto se explica por el hecho que la fitohormona auxina juega un rol crítico durante el crecimiento y desarrollo de las plantas, incluyendo la elongación, diferenciación, respuestas fototrópicas y gravitrópicas, dominancia apical y formación de raíces laterales y adventicias (Yang *et al.* 2007). Por otro lado hay evidencia que los miRNAs están involucrados en enfermedades de plantas causadas

por diferentes patógenos y pueden también estar involucrados en silenciamiento de genes inducido por virus (Zhang *et al.* 2006).

Los efectos de la regulación post-transcripcional mediada por microRNAs en el fenotipo vegetal se han demostrado mediante experimentos de sobreexpresión (Bonnet *et al.* 2006). Por ejemplo, altos niveles de miR172 en *Arabidopsis thaliana* puede causar floración temprana y defectos en la identidad floral, tal como la ausencia de pétalos y transformación de sépalos en carpelos. Así mismo, la baja expresión de los miRNAs causa algunas anormalidades en el desarrollo, como alteración en la forma de las hojas, fertilidad femenina, etc. (Zhang *et al.* 2006). En otro sentido, la sobreexpresión de aquellos miRNAs que juegan una función esencial en el desarrollo vascular de las plantas, disminuye la cantidad de mRNA blanco y causa una acelerada diferenciación de células vasculares y consecuentemente causan un sistema vascular alterado con xilema expandido y una región interfascicular (Kim *et al.* 2005).

2.2.3. Promotores de microRNAs

La mayoría de los miRNAs se presentan como loci de copia única en el genoma, pero algunos se encuentran formando clústers que son transcritos y expresados de manera coordinada, es decir, que pueden ser policistrónicos. Los genes de miRNAs son unidades transcripcionales delimitados comúnmente por un promotor que contiene una caja TATA corriente arriba del punto de inicio de la transcripción y una señal de poliadenilación en su extremo 3' (Kurihara *et al.* 2004, Jones-Rhoades *et al.* 2006).

Existe evidencia que demuestra que por lo menos algunos de los genes de microRNAs son genes de clase-II (genes transcritos por la RNA polimerasa II (Pol II)), esto determina el promotor y la presencia colas poli A y la estructura en Cap (propiedades únicas de genes tipo II) en los transcritos primarios (pri-microRNAs) de

los genes de microRNAs, como lo mostraron Houbaviy *et al.* (2005) encontrando una caja TATA localizada a -35 del inicio del pri-microRNA, el cual poseía una cap y estaba poliadenilado, y Xie *et al.* (2005) identificando los promotores de 52 genes de microRNA de *A. thaliana* que tenían una caja TATA en sus promotores centrales (Houbaviy *et al.* 2005, Xie *et al.* 2005, Zhou *et al.* 2007). Estos promotores son cruciales para la iniciación de la transcripción y determinan las condiciones específicas bajo las cuales la expresión génica se lleve a cabo.

Para entender el mecanismo y condiciones de la activación de los genes de microRNAs, se requiere conocer la localización de las regiones promotoras centrales. Con tal fin se pueden utilizar un programa de computador para la predicción de promotores teniendo en cuenta que los métodos actuales de predicción de promotores pueden no ser adecuados para genes de microRNAs (Zhou *et al.* 2007). Por eso es necesario confirmar tales predicciones en el laboratorio.

2.2.4. Antecedentes: Análisis experimentales y bioinformáticos

Los miRNAs han sido descubiertos usando tres aproximaciones: En primer lugar, clonación directa de librerías de RNAs pequeños de diferentes tejidos y condiciones. En segundo lugar, predicción bioinformática a partir de secuencias genómicas seguida por la validación experimental y por último un método tradicional, los screens genéticos gracias a los cuales se descubrieron los primeros miRNAs (Jones-Rhoades *et al.* 2006) (*lin-4* y *let-7*) en *Caenorhabditis elegans* en 1993 (Yang *et al.* 2007, Zhou *et al.* 2007). Pero hay que tener en cuenta que cuando los genomas están disponibles, los programas bioinformáticos son muy útiles para encontrar microRNAs verdaderos dadas las características de sus precursores, por el contrario, si el genoma no se ha secuenciado, una aproximación experimental es la mejor opción para identificar miRNAs (Yao *et al.* 2007).

El método más directo para descubrir los miRNAs es el aislamiento y clonación de los RNAs pequeños a partir de muestras biológicas. El procedimiento general que se utiliza es aislamiento de miRNAs, ligación a adaptadores, transcripción reversa, amplificación, clonación y por último la secuenciación (Jones-Rhoades *et al.* 2006).

Con respecto a lo anteriormente expuesto, el PCR se puede secuenciar sin necesidad de ser clonado, es decir mediante secuenciación a gran escala. Un tipo es la pirosecuenciación o secuenciación 454, la cual consiste en lo siguiente: Primero el DNA se mezcla con unas esferas que poseen la secuencia complementaria al primer B. Posteriormente estas esferas son emulsionadas con reactivos de PCR + luciferasa en una mezcla de aceite y agua. Después de realizar el PCR, éste se coloca en una placa (PicoTiterPlate) que posee multitud de pozos o pocillos, la cual es ubicada en el secuenciador FLX. La unión de los dNTPs a los nucleótidos complementarios provoca la liberación de una molécula de pirofosfato, y gracias a que la reacción tiene luciferasa, al ponerse en contacto con el pirofosfato, genera una luz (proporcional al número de nucleótidos incorporados) que es captada por el secuenciador (Margulies *et al.* 2005, Droege *et al.* 2008).

La identificación y cuantificación en plantas de pequeños RNAs mediante métodos de secuenciación fue primero implementado en *Arabidopsis*. Más de 2 millones de pequeños RNAs fueron secuenciados por MPSS (Massively Parallel Signature Sequencing) de flores y semillas de esta especie. Esto representa un avance significativo sobre los métodos tradicionales para la identificación de miRNAs. Una limitación de MPSS es que solamente es capaz de secuenciar en el extremo 5' 17 nucleótidos. Lu *et al.* (2007) fueron pioneros en el uso de un método alternativo, la pirosecuenciación, para identificar RNAs pequeños: esta es una técnica que permite analizar secuencias más largas (100-400 pb) a gran escala (Lu *et al.* 2007).

Substanciales progresos han sido desarrollados en las nuevas tecnologías de secuenciación. Solexa, Inc. ha desarrollado secuenciación de DNA por síntesis (SBS),

método que puede potencialmente generar > 10 millones de secuencias de 20-30 nt con alta precisión. Un método diferente de secuenciaciones se denomina SOLiD (Supported Oligo Ligation Detection) el cual es desarrollado por Biosystems, Inc (Lu *et al.* 2007).

Por otro lado los análisis bioinformáticos han sido usados como complementarios a la clonación para encontrar similitud entre los miRNAs y los genomas, buscar secuencias conservadas entre especies, predecir estructuras secundarias y encontrar los genes regulados (Yang *et al.* 2007).

Con el paso de los años, varios programas bioinformáticos han sido desarrollados para ayudar a identificar potenciales genes diana de los miRNAs conociendo la secuencia de mRNA blanco. Estos programas incluyen MIRANDA, MIRcheck, miRU y otros. Debido a que los miRNAs de plantas muestran una perfecta o casi perfecta complementariedad con sus mRNAs diana, es mucho más fácil predecir los genes diana empleando algoritmos de búsqueda (Yin *et al.* 2008).

Hasta la fecha, se han secuenciado alrededor de 4000 miRNAs (Zhang *et al.* 2007) los cuales están depositados en una actual edición del registro de miRNAs MIRBASE (<http://microrna.sanger.ac.uk/sequences/index.shtml>) (Griffiths-Jones 2004). La mayoría de los miRNAs de plantas han sido identificados para *Arabidopsis*, *O. sativa* y *Populus trichocarpa* cuyos genomas están secuenciados completamente (Yin *et al.* 2008). En plantas se han encontrado 863 miRNAs, 131 para *Arabidopsis*, 242 para arroz, 215 para *Populus*, 96 para maíz, 72 para *Sorghum*, 39 para *Physcomitrella*, 30 para *Medicago truncatula*, 22 para soya y 16 para caña de azúcar (Yao *et al.* 2007).

3. FORMULACIÓN DEL PROBLEMA, PREGUNTA Y JUSTIFICACIÓN

3.1. Formulación del problema

La yuca (*Manihot esculenta* Crantz) se caracteriza por su alta producción de carbohidratos y es considerada como uno de los cultivos más importantes en las regiones tropicales del mundo. Tanto sus raíces como sus hojas pueden ser utilizadas para consumo humano, alimentación animal, y adicionalmente para la industria (almidón y alcohol) (Ceballos 2002).

Además del valor económico, este cultivo ofrece otras ventajas como la tolerancia a la sequía, la capacidad de producir en suelos degradados y la tolerancia a los suelos ácidos y de baja fertilidad. A pesar de su enorme potencial productivo, su destacable adaptación a una gran diversidad de ambientes, su reconocida tolerancia a factores bióticos y abióticos, y su diversidad de usos, la yuca no ha logrado desarrollar todo su potencial. Esto se debe a que es muy vulnerable a virus, plagas y enfermedades y el proceso de mejoramiento genético es lento. Adicionalmente, la perecebilidad de las raíces para consumo es alta y debido a su volumen y contenido de agua genera altos costos para el transporte. A pesar de todas estas dificultades, la yuca permanece como cultivo de importancia a nivel global y se están dando pasos acelerados para resolver estos problemas (Ceballos 2002).

Debido a que existe evidencia que no todos los miRNAs están conservados entre plantas, si no que la gran mayoría de estos transcritos son propios y exclusivos de cada especie (Lindow *et al.* 2005, Jones-Rhoades *et al.* 2006, Moxon *et al.* 2008), a que solo se conoce un 1 % de estos transcritos y sus funciones y a que no hay estudios de RNAs pequeños en *Manihot esculenta*, el propósito de este estudio es secuenciar y encontrar la función de los miRNAs que se expresan en plántulas y plantas de yuca bajo diferentes condiciones, ya que se sabe que los miRNAs se

expresan en tejidos en desarrollo y hay una expresión diferencial bajo distintos estímulos ambientales (Rhoades *et al.* 2002, Bartel 2004).

3.2. Pregunta de investigación

¿Cuál es la secuencia de candidatos a miRNAs expresados en plantas de yuca sometidas a condiciones normales y de estrés por calor y sequía, y cuáles son sus potenciales genes diana?

3.3. Justificación de la investigación

Actualmente, la yuca es un cultivo muy importante en las regiones tropicales; más de 163 millones de toneladas son producidas en el mundo cada año y sin duda su valor económico y alimenticio es muy alto. La yuca es el cuarto producto básico más importante después del arroz, el trigo y maíz, y es un componente en la dieta de más de 1000 millones de personas en la región del trópico (Ceballos 2002).

En la actualidad el desconocimiento que aún tenemos sobre estos, su procesamiento post-transcripcional y sus funciones (Lee *et al.* 2003, Bohnsack *et al.* 2004, Lund *et al.* 2004, Krol *et al.* 2006), hace que sea necesario desarrollar trabajos de este tipo, ya que el conocimiento acerca de los procesos de regulación génica en donde están involucrados los miRNAs se está convirtiendo en una herramienta para estudiar los procesos de desarrollo, crecimiento y fisiología de las plantas (Ceballos 2002, Zhang *et al.* 2006).

Dentro de las nuevas líneas de investigación, la información acerca de miRNAs se podría utilizar con varios propósitos: 1). Diseñar miRNAs artificiales para reprimir la expresión de genes diana con el objetivo de estudiar la función de genes de interés

(Jones-Rhoades *et al.* 2006), 2). Identificar métodos de defensa o resistencia de las plantas a diversos estreses ambientales, ataque de insectos e infección por patógenos.

4. OBJETIVOS

4.1. General

Clonación, secuenciación y caracterización bioinformática de un conjunto de microRNAs expresados en plantas bajo condiciones normales y de estrés por calor y sequía, y predicción de sus posibles roles en regulación.

4.2. Específicos

1. Clonar y Secuenciar los miRNAs de *Manihot esculenta* var. Tai 16 expresados bajo condiciones normales y de estrés
2. Identificar microRNAs de *Manihot esculenta* Var. Tai 16 conservados con otras especies de plantas.
3. Caracterizar bioinformáticamente la población de RNAs pequeños obtenidos de plantas de yuca bajo condiciones normales y de estrés por calor y sequía
4. Predecir los genes diana de los miRNAs secuenciados

5. MATERIALES Y MÉTODOS

5.1. Diseño de la Investigación

Es una investigación de tipo descriptiva, cuyo propósito es caracterizar los potenciales microRNA presentes en yuca que se expresan en plántulas cultivadas *in-vitro* bajo condiciones normales de crecimiento, bajo estrés de sequía y calor, y plantas de campo bajo condiciones normales. Dicha investigación incluye el

aislamiento de los miRNAs, su clonación y secuenciación, comparación con otros miRNAs registrados en MIRBASE, comparación con ESTs propios de yuca, análisis de su estructura secundaria, análisis de sus promotores y genes diana.

5.1.1. Población de estudio y muestra de estudio

5.1.1.1. Población de estudio

Plantas de *Manihot esculenta* Crantz Var. Tai 16 en cultivo.

5.1.1.2. Muestra de estudio

Muestras de hojas y raíces de *Manihot esculenta* variedad Tai 16 cultivadas *in-vitro* y en campo en el CIAT (Centro Internacional de Agricultura Tropical. Palmira, Valle del Cauca), se empaquetaron en sobres de aluminio de 0.2 g de tejido, se congelaron inmediatamente en nitrógeno líquido y se almacenaron a -80° C hasta el momento de la extracción del RNA (Abdurakhmonov *et al.* 2008).

5.2. Tratamientos

Para maximizar la diversidad de miRNAs, se realizó la extracción de RNA total a partir de varios grupos de plantas sometidas a diferentes condiciones (Rhoades *et al.* 2002; Axtell *et al.* 2005), dejando en claro que el objetivo no es el de comparar las secuencias para cada tratamiento.

Los siguientes tratamientos fueron utilizados, ya que estos mismos fueron aplicados por Sakurai *et al.* (2007) para obtener los ESTs para *Manihot esculenta* variedad Tai 16.

1. Plantas *in-vitro* de dos meses de edad bajo condiciones normales de crecimiento o control.
2. Plantas *in-vitro* de dos meses de edad sometidas a estrés por sequía. Las plantas se retiran del medio y se dejan a temperatura ambiente por 3 y 6 horas, al cabo de las cuales se congelan las plantas completas.
3. Plantas *in-vitro* de dos meses de edad sometidas a altas temperaturas (incubadas en un horno a 40°C). A las 6 y 24 horas de sometidas al tratamiento son congeladas.
4. Plantas *in-vitro* de dos meses de edad sometidas a estrés de calor por 24 h y posteriormente 4 horas a temperatura ambiente. A las 28 horas son congeladas.
5. Plantas cultivadas en campo bajo condiciones normales de crecimiento o control. Se toman por separado las hojas y raíces jóvenes, se lavan con agua tratada con DEPC y se congelan inmediatamente. En el caso de las raíces estas fueron peladas y ralladas en nitrógeno líquido.

5.3. Procedimiento de laboratorio

5.3.1. Extracción de RNAs total y miRNAs con el kit *mirVANA*

La extracción o aislamiento de los RNAs totales de las plantas cultivadas *in-vitro* y de campo, se realizó en tres grupos (*in-vitro*, raíz-campo y hoja-campo) con el kit *mirVANA*TM miRNA Isolation (AMBION, USA, Abdurakhmonov *et al.* 2008), el cual consta de dos pasos principales, extracción orgánica con fenol-cloroformo y extracción final de RNA total, incluyendo los RNAs de menor peso molecular, en columnas de fibra de vidrio, usando las soluciones proporcionadas en el kit (Ver Anexo 1). Para las raíces de campo, debido a la presencia de almidón en este tejido (Ceballos 2002), se realizó la extracción de RNA total con el kit *mirVANA*TM, pero con una extracción orgánica con fenol ácido (pH 4.3): cloroformo: alcohol isoamílico

(24:24:1) previa al procedimiento del kit (Ver Anexo 2). Para confirmar la calidad y pureza de la extracción fueron analizados por electroforésis 5 μ l de los RNA totales en un gel de agarosa 1.5 % en TBE 1X, esperando observar 3 bandas definidas de RNA: 28S (4.6 a 5.2 kb), 18S (2 kb) y RNAs de bajo peso molecular (*mirVana*TM miRNA Isolation Kit, Bryant 1998, Farrell 2005).

5.3.2. Cuantificación de las muestras

Para la cuantificación de cada una de las muestras (*in-vitro*, raíz-campo y hoja-campo) se hizo la dilución de 3 μ l del RNA en 297 μ l de agua libre de RNAsas y se realizó la lectura de absorbancias a 260 nm (absorción máxima del RNA de luz ultravioleta), 280 nm y 230 nm en el espectrofotómetro BECKMAN DU-240 (*mirVana*TM miRNA Isolation Kit, Rapley 1998); se calculó la concentración en ng/ μ l con la siguiente fórmula: concentración de RNA total (ng/ μ l) = 40*A₂₆₀*factor de dilución; (*mirVANA*TM miRNA Isolation kit, Rapley 1998, Farrell 2005), y se calculó la cantidad (ug) multiplicando la concentración por el volumen. Los mejores resultados se obtienen si 50-100 ug de RNA total son usados (*miRCat*TM microRNA cloning kit) o 50-500 ug según Lau *et al* (2001) y Lu *et al* (2007).

5.3.3. Aislamiento de los microRNA a partir de gel de acrilamida

Se mezclaron todos los RNAs totales extraídos, para tener un 'pool' de todos los tratamientos. El cual se corrió en un gel de acrilamida denaturante al 15% (urea 7M) (dPAGE) (Ver Anexo 3) en una cámara de electroforesis vertical PROTEAN (BIORAD), por alrededor de 15 minutos a 270 voltios. El gel se tiñó con SYBR-Gold (Invitrogen), se observó en un trans-iluminador de luz azul para cortar la fracción de RNA deseado (entre 15 y 30 nt). Para ello se uso un marcador de RNA de 21 nt (*miSPIKE*, Integrated DNA Technologies, USA) y también el marcador 10 pb

sembrado en un pozo contiguo al RNA para confirmar el tamaño (miRCat™ microRNA cloning kit, Lau *et al.* 2001, Lu *et al.* 2007, Abdurakhmonov *et al.* 2008). Los miRNAs que fueron cortados del gel de acrilamida, se purificaron usando el método estándar de macerado y remojo “Crush and Soak” del gel y a través de las columnas NAP-5 (GE healthcare, USA). Los RNA filtrados se precipitaron con etanol absoluto frío, acetato de sodio 3M (pH 5.3) y glicógeno (10 mg/ml) y luego se resuspendieron en 10 ul de H2O-DEPC (Ver Anexo 4) (miRCat™ microRNA cloning kit, Lu *et al.* 2007).

5.3.4. Construcción, clonación y secuenciación de librerías de microRNAs mediante kit MiRCat siguiendo las pautas del protocolo de Bartel (Ver figura 2) (Lau *et al.* 2001, Lu *et al.* 2007).

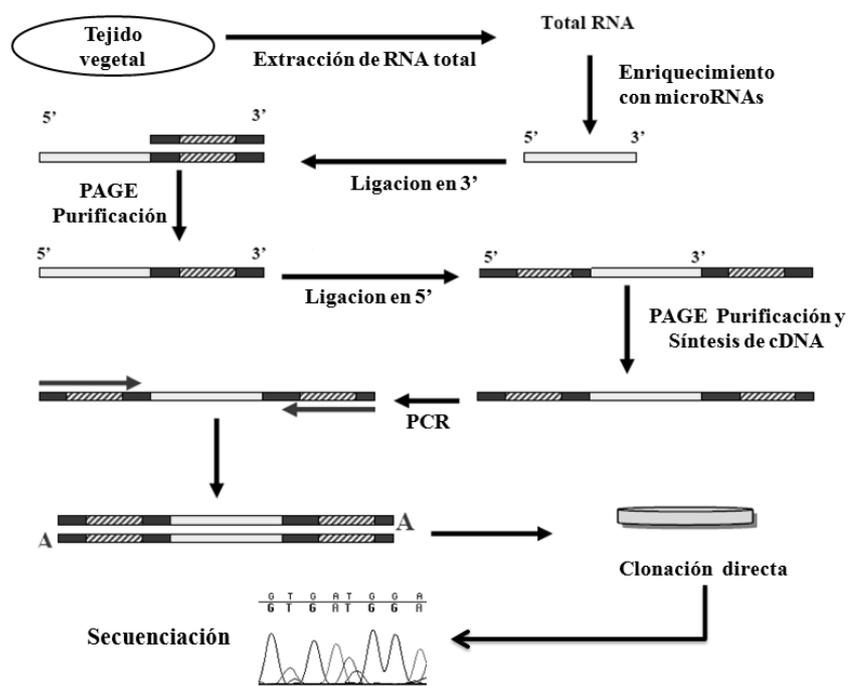


Fig. 2: Procedimiento para generar y secuenciar una librería de microRNAs. (Tomado y modificado de MiRCat smallRNAs cloning kit Technical manual)

La construcción de la librería de miRNAs se realizó de acuerdo con los protocolos de de los kits MiRCat™ microRNA cloning (Integrated DNA Technologies, USA), SuperScript™ III Reverse Transcriptase (Invitrogen), Advantage™ 2 Polymerase Mix (CLONTECH) y pGEM®-T Easy Vector Systems (Lagos-Quintana *et al.* 2001).

5.3.4.1. Ligación de adaptador en el extremo 3' de los microRNAs

Una vez la fracción de smallRNAs fue aislada y purificada, éstos se ligaron en su extremo 3' a un adaptador adenilado (5'-rAppCTGTAGGCACCATCAATddC-3') usando RNA ligasa T4 en la ausencia de ATP. Esta ligación se incubó a 22 °C por dos horas y posteriormente se limpió (Ver Anexo 5a.). Luego en un gel de poliacrilamida 15 % denaturante (urea 7M) se corrió la ligación en 3', por medio de electroforesis horizontal, y se cortó la fracción de gel entre 37 y 43 nt (como control de tamaño se uso miSPIKE (21 nt) quién también se ligó, piSPIKE (31 nt) y el marcador 10 pb), la cual se maceró y filtró en columnas NAP-5 (GE healthcare, USA), y por último se precipitó el filtrado (Ver Anexo 4) (Abdurakhmonov *et al.* 2008).

5.3.4.2. Ligación de adaptador en el extremo 5' de los microRNAs

Se ligaron en sus extremos 5' con un adaptador quimérico o híbrido (DNA/RNA) (5'-TGGAATrUrCrUrCrGrGrGrCrArCrCrArArGrGrU-3') usando RNA ligasa T4 en presencia de 10 mg/ml de ATP. Esta ligación se incubó a 22 °C por dos horas y posteriormente se limpió (Ver Anexo 5b.). Luego en un gel de poliacrilamida 15 % denaturante (urea 7M) se corrió la ligación en 5' y se cortó la fracción de gel entre 55 y 65 nt (teniendo como control de tamaño piSPIKE y el marcador 10 pb), la cual se maceró y filtró en columnas NAP-5 (GE healthcare, USA), y por último se precipitó el filtrado (Ver Anexo 4) (Abdurakhmonov *et al.* 2008).

5.3.4.3. Síntesis de cDNA

A partir de los smallRNAs ligados a adaptadores en sus extremos 3' y 5', se procedió con la síntesis de cDNA con la enzima transcriptasa reversa SuperScript™ III (INVITROGEN) usando el primer reverso complementario al adaptador 3' (5'-GATTGATGGTGCCTACAG 3', Tm = 50.2°C) (Ver Anexo 5c.) (Lu *et al.* 2007, Shao-Yao 2006, Abdurakhmonov *et al.* 2008)

5.3.4.4. Síntesis de segunda cadena (PCR) y cuantificación

Directamente se hizo la síntesis de segunda cadena por PCR usando los primers modificados de IDT para pirosecuenciación (For: 5'-GCCTCCCTCGCGCCATCAGTGGGAATTCTCGGGCACC -3'; Rev: 5'-GCCTTGCCAGCCCGCTCAGGATTGATGGTGCCTACAG -3') (Ver Anexo 5d.). Los amplicones en el rango esperado de 95 a 105 pb fueron purificados de un gel de poliacrilamida 15% con las columnas NAP-5 (Ver Anexo 4) (Lu *et al.* 2007, Abdurakhmonov *et al.* 2008) y su concentración fue medida con Quant-iT™ picoGreen (<http://probes.invitrogen.com/media/pis/mp07581.pdf>) en el espectrofluorometro TECAN (Ver Anexo 6).

5.3.4.5. Ligación de los amplicones al vector pGEMT-easy

A continuación, se procedió a realizar la ligación del PCR al vector pGEM®-T easy (Ver Anexo 5e.); la cantidad de PCR a usar en la ligación dependió de la proporción molar inserto: vector que se quiso utilizar, en este caso la proporción fue 7.6:1. Por lo tanto, la cantidad de PCR que se utilizó fue de 12.7 ng, según la siguiente fórmula:

$$\text{ng inserto} = \frac{50 \text{ ng vector} * 0.1 \text{ kb inserto}}{3.0 \text{ kb vector}} * \frac{7.6}{1} \quad (\text{pGEM®-T Easy Vector Systems kit}).$$

5.3.4.6. Transformación de células bacterianas y PCR

Después de limpiar el producto de ligación con Sec-Butanol y etanol (Ver Anexo 5f.), se electroporaron las células bacterianas competentes TOPO junto con el vector que contenía el PCR y se dejaron recuperar en medio SOC por una hora a 37°C. Acto seguido, se plaquearon cajas de petri (Carbenicilina 100 ug/ml, IPTG 0,5 mM y x-GAL 80 ug/ml) con diferentes volúmenes de transformación y se dejaron crecer por una noche a 37°C. El siguiente paso, fue seleccionar 158 colonias blancas, para picarlas y sembrarlas en dos placas de 96 con LB líquido y carbenicilina 100 ug/ml, las cuales se dejaron crecer a 37 °C toda la noche. Al siguiente día se alicuotaron 150 µl de cada uno de los clones para guardar como librería de smallRNAs (Ver Anexo 5g.). Se realizó un PCR (Ver Anexo 5h.) sobre las 158 colonias seleccionadas y dos blancos con los primer SP6 y T7. Para confirmar la presencia del inserto se visualizaron los PCRs en un gel de agarosa 4 % en TBE 1X, junto con un marcador de peso molecular (10 pb) para confirmar el tamaño, esperando obtener tamaños de aproximadamente de 240 pb (Abdurakhmonov *et al.* 2008).

5.3.5. Secuenciación del producto de PCR de los clones

De los 158 PCR (sobre los clones) se seleccionaron 96 buscando la más alta variabilidad (en tamaño) posible y los PCRs se pasaron a una nueva placa para su limpieza y precipitación con PEG 20%. Se visualizaron 2 ul del producto de precipitación en un gel de agarosa 4% de los 10 ul disponibles a partir de la limpieza y se procedió a preparar la reacción de secuencia con el PCR, mix BidDye (dNTPs marcados), buffer y primer SP6. El exceso de oligonucleótidos marcados fue removido por el método estándar de precipitación con etanol. Una vez precipitados, los amplicones fueron secados y resuspendidos en 3 ul de buffer carga (formamida, blue dextran) y se sembraron en un gel de acrilamida 4%, usando el ABI PRISM™ 377 DNA Sequencher de PERKIN ELMER para secuenciar por el método de

Sanger(Ver Anexo 7) (ABI PRISM® Big Dye terminator V1.1. cycle sequencing Ready Reaction kit, Abdurakhmonov *et al.* 2008)

5.4. Análisis bioinformático

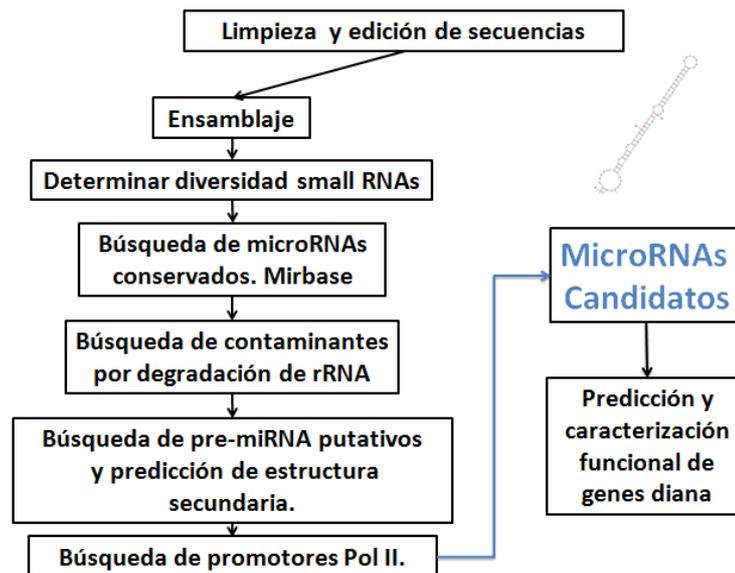


Fig. 3: Procedimiento para editar y analizar las secuencias de microRNAs secuenciados y sus funciones en regulación

5.4.1. Edición de secuencias en Sequencher 4.6.

Para analizar la diversidad de miRNAs, las secuencias obtenidas a partir de la secuenciación fueron editadas con SEQUENCHER 4.6. Este paso consistió en quitar las secuencias ambiguas, secuencias de vector pGEMT-easy y de primers 454. A partir de esto se pudo saber cuántos de los clones tenían buenos insertos dentro del rango de tamaño esperado, y de estos obtener el número de reads y la longitud de cada secuencia y cada contig (Abdurakhmonov *et al.* 2008).

Debido a que el PCR se liga al vector pGEMT-easy de dos maneras diferentes (5'-3' y 3'-5'), se verificó en SEQUENCHER 4.6 que la orientación de la secuencias correspondiera a la dirección correcta 5'-3' de los transcritos corroborando en los cromatogramas la secuencia de los primers a cada lado, es decir, que se tuviera lo siguiente:

5' -

GCCTCCCTCGCGCCATCAGTGGAATTCTCGGGCACCAAGGU -transcrito-
CTGTAGGCACCATCAATCCTGAGCGGGCTGGCAAGGC-3' (Abdurakhmonov
et al. 2008).

5.4.2. Alineamiento de secuencias

En este mismo software (SEQUENCHER) se agruparon todas las secuencias obtenidas en contigs con el 100% de porcentaje de alineamiento. Tanto para los contigs formados como para las secuencias que no formaron ninguno (de tamaños entre 12 y 18 nt), fueron exportadas las secuencias consenso en formato fasta a un formato de texto plano. Este archivo se guardó en la base de datos interna del CIAT y se alinearon las secuencias utilizando PRANK (desde el terminal) para generar un ensamblaje, el cual fue editado en MacClade 4.06 OS X, para agruparlos según sus semejanzas y observar cuantas familias se habían secuenciado (Golubchik *et al.* 2007, Abdurakhmonov *et al.* 2008, <http://www.macclade.org/>), y no para plantear una hipótesis de homología.

5.4.3. Comparación con bases de datos.

Para todas las búsquedas de similitud mediante el algoritmo de BLAST, se tomaron solo aquellas secuencias de las cuales se obtuvieran hits con un E-value menor a 2 y con un porcentaje de identidad o similaridad en el alineamiento de 100 % (Yin *et al.* 2008).

5.4.3.1. BLAST vs. MIRBASE.

Para encontrar microRNAs conservados se hizo una búsqueda de similitud (usando el algoritmo de BLASTN) de los potenciales miRNAs de yuca con secuencias de miRNAs maduros de otras especies, depositados en MIRBASE. Lo anterior se realizó desde la página de MIRBASE (<http://microrna.sanger.ac.uk/sequences/search.shtml>), y desde el terminal utilizando la base de datos de miRNAs depositada en el servidor de CIAT (Altschul *et al.* 1997, Korf *et al.* 2003). Los miRNAs maduros de otras especies deben tener solo de 0-3 nucleótidos no apareados con la secuencia denominada “query” o miRNAs secuenciados (Abdurakhmonov *et al.* 2008, Yin *et al.* 2008)

5.4.3.2. BLASTN vs. ESTs de *Manihot esculenta*.

Se realizó una búsqueda de similitud usando el algoritmo de BLASTN, de los potenciales miRNAs en contra de los ESTs de *Manihot esculenta* (Altschul *et al.* 1997; Abdurakhmonov *et al.* 2008) con el objetivo de encontrar similaridad entre los potenciales miRNAs secuenciados y las secuencias de ESTs de yuca (Korf *et al.* 2003, Zhang *et al.* 2005).

5.4.3.3. Recuperación de secuencias de ESTs.

Las secuencias completas de los ESTs, contra los cuáles se produjo un hit de similaridad al compararlos con los potenciales microRNAs, se recuperaron a partir de la base de datos de ESTs de yuca de CIAT (Korf *et al.* 2003).

5.4.3.4. BLAST ESTs vs. GenBank de NCBI

Las secuencias completas de los ESTs recuperados, se compararon en NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) con las secuencias depositadas en la base de datos de “Nucleotide collection (nr/nt)”, para descartar que correspondieran a proteínas ó a productos de degradación de RNA ribosomal (National Center for Biotechnology Information, Abdurakhmonov *et al.* 2008).

5.4.3.5. BLAST vs. RNA ribosomal.

Para confirmar lo que se obtuviera en el paso anterior, es decir descartar que los potenciales microRNAs correspondieran a proteínas o productos de degradación RNAs ribosomales de mayor peso molecular, se realizó una búsqueda de similitud usando el algoritmo de BLASTN, contra las secuencias de la base de datos de RNA ribosomal de la Universidad de Ghent (<http://bioinformatics.psb.ugent.be/webtools/rRNA/blastrrna.html>).

5.4.3.6. BLASTN vs. Genoma de *Ricinus communis*.

Debido a que el genoma de *Manihot esculenta* no está disponible y dado que *Ricinus communis* es la euforbiácea más cercana a Yuca, se realizó una búsqueda de similitud (BLASTN 2.2.16) solamente con los potenciales miRNAs que no correspondían a RNA ribosomal o a proteínas (desde el terminal utilizando la base de datos interna de CIAT) en contra del genoma *Ricinus communis*, con el objetivo de encontrar secuencias genómicas similares y a partir de las secuencias flanqueantes a las regiones de similitud, predecir las estructuras secundarias.

5.4.3.7. Recuperación de las secuencias flanqueantes a partir de los hits con *Ricinus communis* y predicción de estructuras secundarias.

Solo a partir de aquellas secuencias de potenciales microRNAs que no correspondían a RNA ribosomal, se recuperó las secuencias flanqueantes 80 pb a cada lado de la región de similitud (Lu *et al.* 2008) en el genoma de *Ricinus communis*, con las cuáles se predijo las estructuras secundarias de los potenciales pre-microRNAs usando MFOLD 3.2 (Wang *et al.* 2004b, Lu *et al.* 2008, <http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi>, Korf *et al.* 2003).

5.4.4. Análisis de promotores tipo II.

Para analizar los potenciales microRNAs de yuca, se buscaron dos promotores de genes tipo II, la caja TATA y la caja CAAT en la secuencia 30165 de *Ricinus communis*. Para esto se recuperó la secuencia flanqueante 2000 pb corriente arriba del inicio del pre-miRNA (es decir 2080 nt corriente arriba del microRNA), la cual se analizó mediante GENOMATIX matinspector (http://www.genomatix.de/cgi-bin/matinspector_prof/mat_fam.pl?s=ffb0277337c7815e59140a74e40061fe). Por lo tanto, se esperaba obtener cualquiera de los promotores antes mencionados en las posiciones 1 a 500 pb, 500 a 1000 pb, 1000 a 1500 pb y a 1500 según Zhou *et al.* (2007) y Higo *et al.* (1999).

5.4.5. Predicción de potenciales genes diana y sus funciones

Después de confirmar que las secuencias tuvieran los elementos de regulación característicos de genes tipo II, se procedió a buscar los potenciales mRNAs regulados por los potenciales miRNAs de yuca detectados.

Los potenciales genes diana se analizaron utilizando miRanda V1.9 (microRNA Target scanning Algorithm, http://cbio.mskcc.org/research/sander/data/miRNA2003/miranda_new.html), desde el terminal, comparando la secuencia de microRNA validado

(GCTGACCCTGAAGTTCAT) con la base de datos de ESTs de yuca (Enright *et al.* 2003, Korf *et al.* 2003), con el objetivo de encontrar secuencias complementarias de ESTs que podrían estar siendo regulados por el potencial miRNA secuenciado.

Con el propósito de encontrar la función de los potenciales genes blanco, es decir a que proteínas son codificantes, se recuperó la secuencia completa de los ESTs de yuca (Korf *et al.* 2003) para compararlos con la base de datos swissprot de NCBI mediante BLASTX (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Lu *et al.* 2008), teniendo en cuenta que los hits estuvieran en frame (complemetarios).

6. RESULTADOS

6.1. Procedimiento de laboratorio

6.1.1. Extracción de RNAs total y miRNAs con el kit *mirVANA*

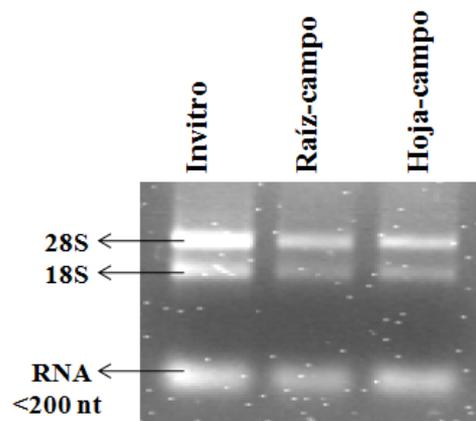


Fig. 4: RNA total de *Manihot esculenta* sembrado en gel de agarosa 1.5%.

En la figura 4 se muestra el RNA total, aislado de plántulas sembradas *in-vitro* y de plantas de campo, sembrado en un gel de agarosa al 1.5 %. En esta foto se observan

tres bandas, dos superiores la cuales corresponden a RNA ribosomal 28S y 18S y una banda inferior que corresponde a RNAs de tamaños menores a 200 nt, por debajo de la cual se espera que estén presentes los microRNAs (Bryant 1998).

6.1.2. Cuantificación de las muestras

Tabla 1: Lecturas de absorbancias a 260 nm y 280 nm, concentraciones y pureza del RNA total aislado a partir de plantas *in-vitro* y plantas de campo.

	<i>A260</i>	<i>A280</i>	<i>Pureza</i> <i>260/280;260/230</i>	<i>Concentración</i> <i>(ng/ul)</i>	<i>Cantidad</i> <i>(ug)</i>
Plantas In-vitro	0.08	0.052	1.53; 1.9	320	28.8
Raíces de plantas de campo	0.05	0.038	1.31; 1.8	200	18
Hojas de plantas de campo	0.039	0.026	1.5; 1.9	156	14
Total				676	60.8

El RNA total inicial para la construcción de la librería de microRNAs fueron 60.8 ug y la pureza dada por las proporciones de las absorbancias a 260 nm con respecto a 280 nm para cada muestra se muestran en la tabla 1.

6.1.3. Aislamiento de los smallRNA a partir de gel de acrilamida

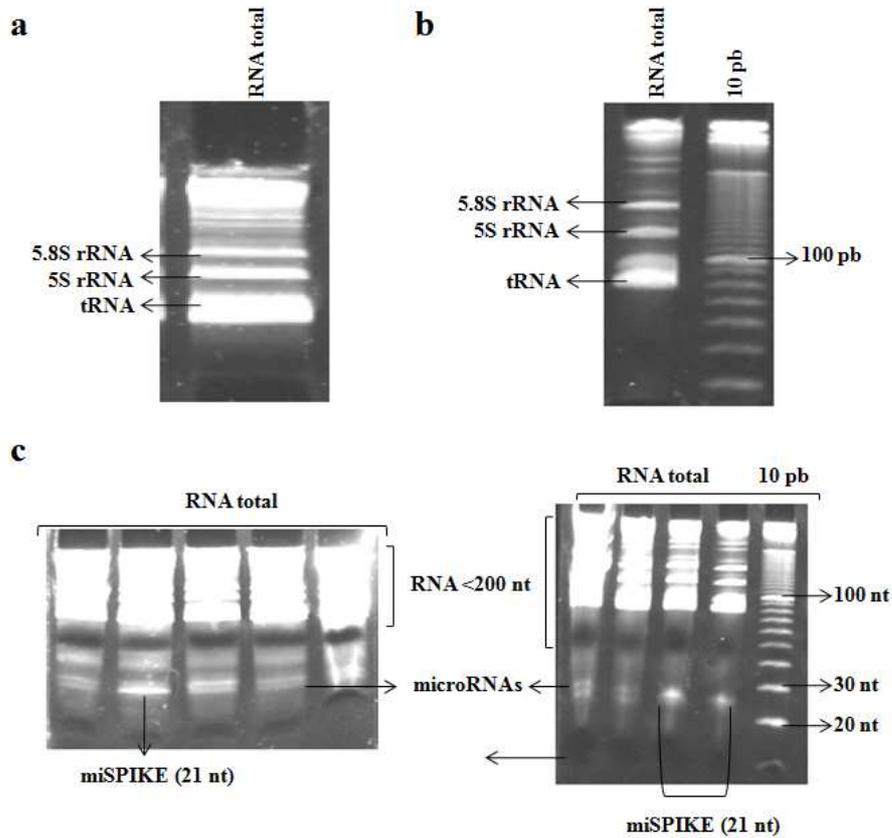


Fig. 5: RNA total de *Manihot esculenta* sembrado en gel de acrilamida 15%. **a** y **b** Bandas de RNA ribosomal y de transferencia y marcador de peso molecular que incluye tamaños cada 10 pb. En **c** dos bandas de smallRNAs detectadas.

En la figura 5 se observa el RNA total aislado a partir de yuca. En **a** y **b** están los RNAs de mayor tamaño no degradados, los cuales corresponden a bandas de rRNA 5.8S, 5S y RNA de transferencia, y en la figura **c** se pueden observar los smallRNAs dentro del rango esperado entre 20 y 30 nt (Bryant 1998).

6.1.4. Construcción, clonación y secuenciación de librerías de smallRNAs.

6.1.4.1. Ligación de adaptador en el extremo 3' de los smallRNAs

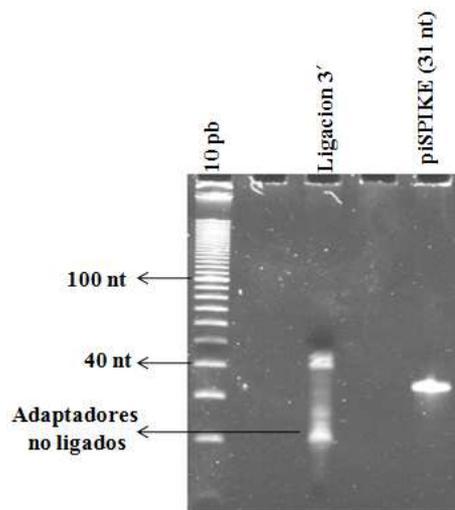


Fig. 6: smallRNAs ligados en su extremo 3' a un adaptador, sembrados en un gel de acrilamida 15%.

En la figura 6 se observan los smallRNAs, ligados en su extremo 3' a un adaptador, de un tamaño de 40 a 45 nt aproximadamente (MiRCatTM smallRNA cloning kit Technical manual).

6.1.4.2. Ligación de adaptador en el extremo 5' de los smallRNAs

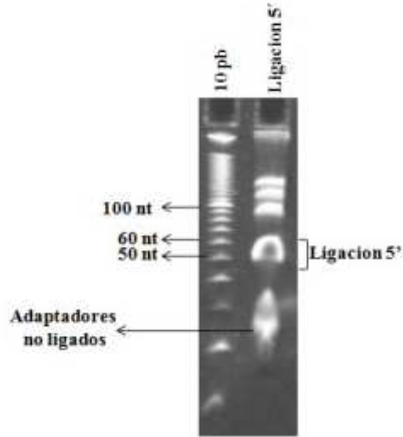


Fig. 7: smallRNAs ligados en su extremo 3' y 5' a adaptadores, sembrados en un gel de acrilamida 15%.

En la figura 7 se observan los smallRNAs ligados en ambos extremos a un adaptador de un tamaño de 50 a 65 nt aproximadamente (MiRCat™ smallRNA cloning kit Technical manual). Se observan tres bandas inespecíficas superiores, es por ello se lleva a cabo a purificación del producto de ligación de tamaño deseado.

6.1.4.3. Síntesis de segunda cadena (PCR) y cuantificación

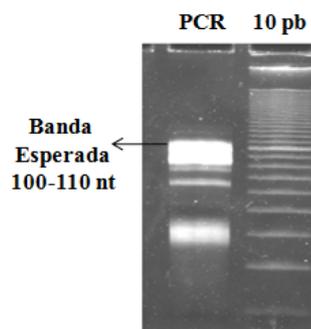


Fig. 8: smallRNAs amplificados con primers 454.

En la figura 8 se observa el producto de amplificación con el tamaño correcto de 100-110 pb. Se pueden detectar otras bandas inespecíficas de menor peso molecular, es por ello que se cortó y purificó solo la fracción de gel correspondiente al PCR de tamaño deseado.

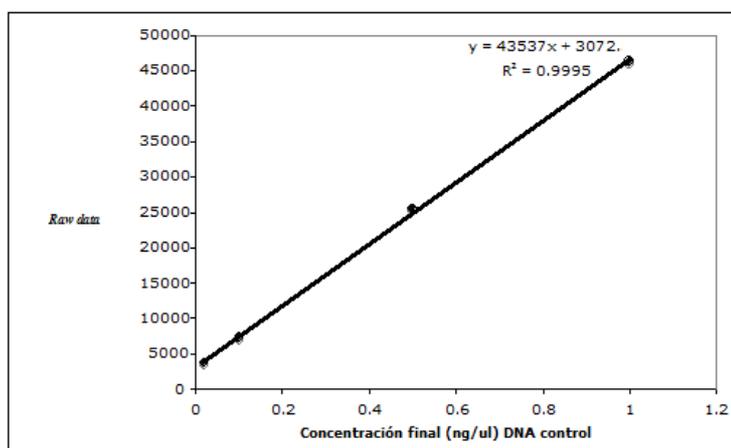


Fig. 9: Curva estándar de DNA (0.02 ng/ul a 1 ng/ul).

En la figura 9 se observa la curva estándar de DNA de concentración conocida en base a la cual se cuantificó la concentración del PCR.

Tabla 2: Concentración, volumen y cantidad del PCR cortado y purificado.

<i>Concentración final esperada (ng/ul) DNA control</i>	<i>Raw data (dato crudo)</i>	<i>Concentración final obtenida (ng/ul) DNA control y blanco</i>	<i>Concentración (ng/ul) obtenida para el PCR. ∇</i>	<i>Factor de dilución</i>	<i>Volumen (ul)</i>	<i>Cantidad (ng)</i>
1	46292	0.992707	-	100	-	-
0.5	25516	0.515504	-	100	-	-
0.1	7267	0.096343	-	100	-	-
0.02	3745	0.015447	-	100	-	-
0	2854	-0.00502	-	100	-	-
-	8601	-	0.12698	100	20	253.96

∇: A partir de la formula de la regresión lineal ($y = 43537x + 3072.5$), dada por la curva estándar, se despejó x y se reemplazó el valor de y con el Raw data obtenido para el PCR cuantificado así: $x = (8601-3072.5)/43537$. El valor de x obtenido, 0.127 corresponde a la concentración la cual es multiplicada por el factor de dilución y luego por el volumen para obtener la cantidad del PCR, 253.96 ng.

El Valor de R2 para la curva estándar fue de 0.9995.

Conociendo la cantidad de PCR que se tiene, para llevar a cabo la ligación al vector, 12.7 ng de amplificado corresponden a 1 ul de la muestra.

6.2. Análisis bioinformático

6.2.1. Edición de secuencias en Sequencher 4.6.

De los 96 clones individuales secuenciados, 67 de ellos tenían un inserto de longitudes entre 12-18 nt (Ver tabla 3), el resto de secuencias correspondían a clones con insertos de tamaños menores a 9 nucleótidos y secuencias difíciles de editar (varios picos para una sola base nucleotídica o picos muy altos) probablemente a la cantidad de PCR secuenciado o calidad de la secuencia.

Tabla 3: Secuencias de potenciales microRNAs de yuca generadas con la librería.

<i>Nombre de secuencia</i>	<i>Secuencia SANGER</i>	<i>Complemento reverso</i>	<i>Secuencia del transcrito</i>	<i>Long. (nt)</i>	<i># Reads</i>	<i>Familia</i>
53-miRNA02H07.SP6	GGGTCCGGGCACG	CGTGCCCGGACCC	CGUGCCCGGACCC	13	1	1
me-smRNA05	GGTCCGGGCACGACG	CGTCGTGCCCCGGACC	CGUCGUGCCCCGGACC	15	2	1
94-miRNA01F12.SP6	GGTCCGGGCACGACGG	CCGTCGTGCCCCGGACC	CCGUCGUGCCCCGGACC	16	1	1
03-miRNA01A03.SP6	CGTGCCCGGACC	GGTCCGGGCACG	CGUGCCCGGACC	12	1	1
34-miRNA01G04.SP6	GTCGGCACGACGCTC	GAGCGTCGTGCCGAC	GAGCGUCGUGCCGAC	15	1	2
me-smRNA04	GCTGACCCTGAAGTTCAT	ATGAACTTCAGGGTCAGC	GCUGACCCUGAAGUUCAU	18	2	3
me-smRNA02	CTGACCCTGAAGTTCAT	ATGAACTTCAGGGTCAG	CUGACCCUGAAGUUCAU	17	17	3
07miRNA01B02.SP6	CTGACCTGAAGTTCAT	ATGAACTTCAGGTCAG	CUGACCCUGAAGUUCAU	16	1	3
me-smRNA03	TGACCCTGAAGTTCAT	ATGAACTTCAGGGTCA	UGACCCUGAAGUUCAU	16	13	3
70-miRNA01F10.SP6	CCTGAAGTTCAT	ATGAACTTCAGG	CCUGAAGUUCAU	12	1	3
me-smRNA01	GACCCTGAAGTTCAT	ATGAACTTCAGGGTC	GACCCUGAAGUUCAU	15	22	3
41-miRNA01A06.SP6	GACCCTGAAATTCAT	ATGAATTCAGGGTC	GACCCUGAAAUUCAU	15	1	3
96-miRNA01H12.SP6	GACCCTGAGATTCAT	ATGAATCTCAGGGTC	GACCCUGAGAUUCAU	15	1	3
20-miRNA01D05.SP6	CTATTTACTCGTACAT	ATGTACGAGTAAATAG	AUGUACGAGUAAAUAG	16	1	4
21-miRNA01E01.SP6	CTGGCCGAGTTA	TAACTCGGCCAG	CUGGCCGAGUUA	12	1	5
75-miRNA01G10.SP6	GGCGTCCGAATTGTAGT	ACTACAATTCGGACGCC	GGCGUCCGAAUUGUAGU	17	1	6
Total					67	

En la tabla 3 están listados los smallRNAs secuenciados con su complemento reverso, la secuencia del transcrito, longitud, número de clones por cada secuencia (reads) y por último la familia. En esta tabla se observa que no todas las secuencias obtenidas a partir de la secuenciación corresponden a la orientación correcta de los transcritos, es decir que también se secuenciaron los complementos reversos. La secuencias obtenidas tienen longitudes entre 12 y 18 nucleótidos.

6.2.3. Comparación con bases de datos.

6.2.3.1. BLAST vs. MIRBASE.

Tabla 4: Blast de los potenciales microRNAs de yuca con microRNAs de otras especies registrados en MIRBASE.

<i>Nombre de secuencia</i>	<i>Hit resultante</i>	<i>Long. (nt)</i>	<i>E- value</i>
53-miRNA02H07.SP6	Human cytomegalovirus miR-US33-5p	11	0.082
me-smRNA05	Human cytomegalovirus miR-US33-5p	11	0.12
94-miRNA01F12.SP6	Human cytomegalovirus miR-US33-5p	11	0.14
03-miRNA01A03.SP6	Human cytomegalovirus miR-US33-5p	11	0.062
34-miRNA01G04.SP6	No Hit		
me-smRNA04	No Hit		
me-smRNA02	No Hit		
07miRNA01B02.SP6	No Hit		
me-smRNA03	No Hit		
70-miRNA01F10.SP6	No Hit		
me-smRNA01	No Hit		
41-miRNA01A06.SP6	No Hit		
96-miRNA01H12.SP6	No Hit		
20-miRNA01D05.SP6	<i>Drosophila melanogaster</i> miR-987	9	2.2
21-miRNA01E01.SP6	No Hit		
75-miRNA01G10.SP6	No Hit		

Cuando se compararon los potenciales microRNAs con otros de otras especies desde la página de MIRBASE no se obtuvo ningún resultado, pero cuando se hizo el BLAST desde el terminal con los microRNAs de otras especies depositados en la base de datos de CIAT se obtuvieron los resultados de la tabla 4. Se muestra la longitud de los hits resultantes y los E-values, para evaluar su significancia.

6.2.3.2. BLAST vs. ESTs *Manihot esculenta* y BLAST ESTs (secuencia completa) vs. NCBI.

Las secuencias completas de los ESTs que fueron similares con los potenciales microRNAs de yuca se muestran en la tabla 5. También en la tabla 5 se presenta el resultado de la comparación de los ESTs con genbank de NCBI, en donde solo los ESTs similares a los potenciales microRNAs 53-miRNA02H07.SP6, me-smRNA05,

94-miRNA01F12.SP6, 03-miRNA01A03.SP6 y 75-miRNA01G10.SP6 corresponden a RNA ribosomal 26S. Mientras que con los otros ESTs que no correspondían a RNA ribosomal, ni RNA mensajero (proteínas) se continuó el análisis.

Tabla 5: Blast de los potenciales microRNAs con ESTs de *Manihot esculenta* y Blast de la secuencia completa de los ESTs con la base de datos (nr) de NCBI.

<i>Nombre de secuencia</i>	<i>Hit resultante potenciales microRNAs vs. ESTs de yuca</i>	<i>BLAST de los ESTs contra NCBI (nr)</i>
53-miRNA02H07.SP6 me-smRNA05 94-miRNA01F12.SP6 03-miRNA01A03.SP6 34-miRNA01G04.SP6	scaff_16862	Gen de RNA ribosomal 26S
me-smRNA04 me-smRNA02 me-smRNA03 me-smRNA01 07miRNA01B02.SP6	scaff_12250 scaff_11891 gb CK643622.1 CK643622 CAS044_J01 scaff_11312	No hit Otro No hit
70-miRNA01F10.SP6 41-miRNA01A06.SP6	CAS042_C08.f CAS010_C13.f CAS045_L14.f CAS037_G18.r scaff_14159 scaff_20057 scaff_17393	No hit No hit
96-miRNA01H12.SP6 20-miRNA01D05.SP6 21-miRNA01E01.SP6 75-miRNA01G10.SP6	CAS009_K07.f No Hit No Hit scaff_16862	No hit Gen de RNA ribosomal 26S

6.2.3.3. BLAST potenciales microRNAs vs. RNA ribosomal.

Tabla 6: Blast de los potenciales microRNAs de yuca con la base de datos de RNA ribosomal Universidad de Ghent.

<i>Nombre de secuencia</i>	<i>Secuencia potencial microRNAs</i>	<i>RNA ribosomal de:</i>
53-miRNA02H07.SP6 me-smRNA05 94-miRNA01F12.SP6 03-miRNA01A03.SP6 34-miRNA01G04.SP6 me-smRNA04 me-smRNA02	CGTGCCCGGACCC CGTCGTGCCCCGACC CCGTCGTGCCCCGACC CGTGCCCGGACC GAGCGTCGTGCCGAC GCTGACCCTGAAGTTCAT CTGACCCTGAAGTTCAT	Planta Planta Planta Planta No Hit No Planta No planta

07miRNA01B02.SP6	CTGACCTGAAGTTCAT	No planta
me-smRNA03	TGACCCTGAAGTTCAT	No planta
70-miRNA01F10.SP6	CCTGAAGTTCAT	No Planta
me-smRNA01	GACCCTGAAGTTCAT	No planta
41-miRNA01A06.SP6	GACCCTGAAATTCAT	No Planta
96-miRNA01H12.SP6	GACCCTGAGATTCAT	No hit
20-miRNA01D05.SP6	ATGTACGAGTAAATAG	No planta
21-miRNA01E01.SP6	CTGGCCGAGTTA	No hit
75-miRNA01G10.SP6	GGCGTCCGAATTGTAGT	Planta

En la tabla 6 se presentan los resultados que se obtuvieron al comparar los potenciales microRNAs con la base de datos de RNA ribosomal, lo cual confirma los resultados de la tabla 5 en donde 5 de las secuencias obtenidas corresponden a RNA ribosomal de plantas.

6.2.3.4. BLAST vs. Genoma de *Ricinus communis*.

Tabla 7: Blast de los potenciales microRNAs de yuca con el genoma de *Ricinus communis*.

Nombre de secuencia	Hit resultante potenciales microRNAs vs. genoma <i>Ricinus communis</i>	Long. (nt)	E-value
34-miRNA01G04.SP6	No Hit		
me-smRNA04	gnl ricinusWGS contig30165	16	0.22
	gnl ricinusWGS contig28320	15	0.86
me-smRNA02	gnl ricinusWGS contig30165	16	0.15
	gnl ricinusWGS contig28320	15	0.58
07miRNA01B02.SP6	gnl ricinusWGS contig30184	15	0.58
me-smRNA03	gnl ricinusWGS contig30165	16	0.15
	gnl ricinusWGS contig28320	15	0.58
70-miRNA01F10.SP6	No Hit		
me-smRNA01	gnl ricinusWGS contig30165	15	0.58
	gnl ricinusWGS contig28320	15	0.58
41-miRNA01A06.SP6	No Hit		
96-miRNA01H12.SP6	No Hit		
20-miRNA01D05.SP6	No Hit		
21-miRNA01E01.SP6	No Hit		

En la tabla 7 se muestra el resultado obtenido al comparar los potenciales microRNAs (que no son productos de degradación de RNA ribosomal o mensajero), con el genoma de *Ricinus communis*. Se puede ver que los hits son casi tan largos como los potenciales microRNAs y también se muestran los E-values de las comparaciones.

6.2.3.5. Recuperación de las secuencias flanqueantes a partir de los hits con *Ricinus communis* y predicción de estructuras secundarias.

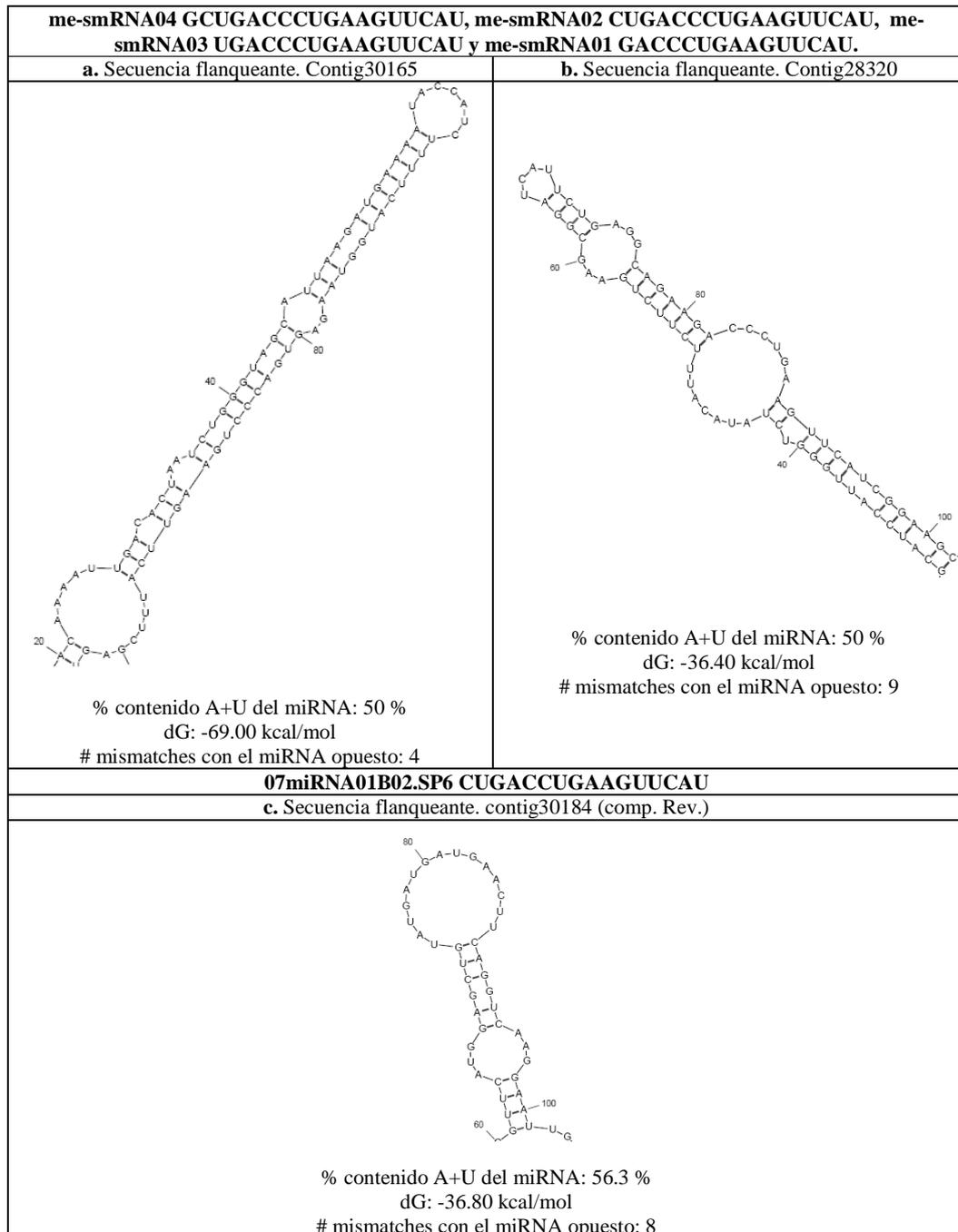


Fig. 11: Estructuras secundarias obtenidas en mFold.

En la figura 11 se presentan las estructuras secundarias de las tres secuencias genómicas de *Ricinus communis*, con la información sobre contenido de Adenina+Guanina, valor de la energía libre de doblamiento y el número de nucleótidos no apareados entre el microRNA y el microRNA* (opuesto).

6.2.4. Análisis de promotores.

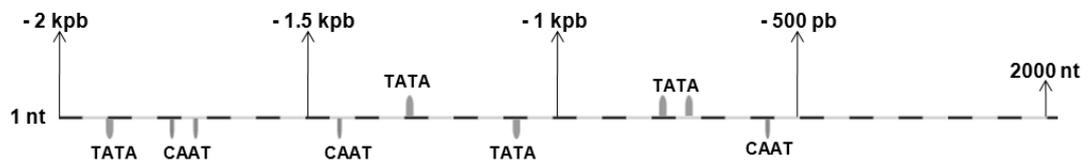


Fig. 12: Potenciales promotores de genes tipo II de pre-miRNA candidato (Contig30165 de *Ricinus communis*).

En la figura 12 se observan elementos reguladores de genes tipo II, tales como la caja TATA y la CAAT, en varias posiciones. El valor de similaridad y secuencia para cada elemento promotor de la secuencia flanqueante del potencial pre-miRNA están en la tabla 8 (Higo *et al.* 1999).

Tabla 8: Elementos promotores de genes tipo II.

<i>Promotor de genes tipo II</i>	<i>posición</i>	<i>similaridad (1.0 perfecta complementariedad)</i>	<i>Secuencia</i>
Caja TATA (1, 2, 3, 4 y 5) para unión de factores de transcripción	99-113	0.953	aagtTATAaatcaaa
	706-720	0.992	actcTATAaatagag
	922-936	0.960	tagtTATAaatagat
	1218-1232	0.958	tcatTATAaataaaa
	1271-1285	0.958	tcatTATAaataaaa
Caja CAAT para unión de factores de transcripción	228-236	0.990	ctCCAAtca
	276-284	0.976	ccCCAAtaa
	567-575	0.989	ctCCAAtgc
	1433-1441	1	ctcCAATta

6.2.5. Predicción de potenciales genes diana y sus funciones

Cuando se comparó el microRNA, validado con el genoma de *Ricinus communis*, con ESTs de yuca mediante miRanda, se obtuvieron los resultados de la figura 13:

scaff_14545		scaff_18940	
Query:	3' TACTTGAAGTCC-CAGTCG 5'	Query:	3' TACTTGAAGTCCCAGTCg 5'
			:
Ref:	5' ATGAACCTCAGGTGTCTGC 3'	Ref:	5' ATGAACTTGAGGGTAGGg 3'
scaff_19125		scaff_15138	
Query:	3' tAC-TTGAAGTCCCAGTCg 5'	Query:	3' TACTTGAAGTCCCAGTCg 5'
			:
Ref:	5' qTGCAACTTCAGGGACAGa 3'	Ref:	5' AAGAACTTCAGGGATAGa 3'
scaff_19798		scaff_12532	
Query:	3' tACTTGAAGTCCCAGTCG 5'	Query:	3' tACTTGAAGTCCCAGTCG 5'
			:
Ref:	5' gTGAAC-TCAGGG-CAGC 3'	Ref:	5' cTGAACTTCAGAGTCGGC 3'
scaff_19657		scaff_13242	
Query:	3' TACTTGAAGTCCCAGTcg 5'	Query:	3' TACTTGAAGTCCCAGTCg 5'
	:		:
Ref:	5' ATGGACTTCAGGGTCata 3'	Ref:	5' ATGAACTTGAGGGTAGGg 3'
scaff_12596		scaff_16213	
Query:	3' TACTTGAAGTCCCAGTCG 5'	Query:	3' TACTTGAAGTCCCAGTCG 5'
Ref:	5' ATGAACTTGAGGTTCCAGC 3'	Ref:	5' AGGAACTTCAGGG-CTGC 3'
CAS008_A19.f		CAS008_L20	
Query:	3' tACTTGAAGTCCCAGTCg 5'	Query:	3' TACTTGAAGTCCCAGTCG 5'
			:
Ref:	5' gTG-ACTTCAGGCTCAGt 3'	Ref:	5' ATGAAC-TCAGGTTCCGGC 3'
CAS010_M12.r		CAS037_G18.r	
Query:	3' TACTTGAAGTCCCAGTCG 5'	Query:	3' tACTTGAAGTCCCAGTCg 5'
			:
Ref:	5' AGGAACTTCAGGGGAAGC 3'	Ref:	5' tTGAATTTTCAGGGTGAGt 3'
CAS045_L14.f		CAS049_I05	
Query:	3' TACTTGAAGTCCCAGTCg 5'	Query:	3' tACTTGAAGTCCCAGTCg 5'
Ref:	5' ATGAACTTCAGGAGCTGt 3'	Ref:	5' gTGAACTTCAGGG-CAct 3'
gi 119018556 gb DV455675.1 DV455675		gi 119019110 gb DV456666.1 DV456666	
Query:	3' TACTTGAAGTCCCAGTCg 5'	Query:	3' TACTTGAAGTCCCAGTCg 5'
	:		:
Ref:	5' ATGGACTTCAGGATCatt 3'	Ref:	5' AAGAACTTCAGGGATAGa 3'

Fig. 13: ESTs probablemente regulados por el único microRNA validado.

En la figura 13 se muestran los ESTs que serían potenciales genes regulados la secuencia de candidato a microRNA. Se observa que los “query” (secuencia problema) son altamente complementarios a los ESTs, de 2 a 4 nucleótidos no apareados (Yin *et al.* 2008).

Tabla 9: Proteínas posiblemente afectadas obtenidas a partir de la comparación de los ESTs con la base de datos swissprot (BLASTX) de NCBI.

<i>Nombre de secuencia</i>	<i>Secuencia potencial microRNAs</i>	<i>Proteína</i>
me-smRNA04	GCTGACCCTGAAGTTCAT	Proteína de formación de tilacoide 1
me-smRNA02	CTGACCCTGAAGTTCAT	Proteína reguladora de luz en el cloroplasto
me-smRNA03	TGACCCTGAAGTTCAT	
me-smRNA01	GACCCTGAAGTTCAT	

En la tabla 9 se muestran las proteínas codificadas por los ESTs probablemente regulados por los microRNAs, al compararlos con la base de datos de proteínas swissprot de NCBI. Estas proteínas fueron encontradas en *Ricinus communis* y están asociadas a procesos relacionados con la fotosíntesis.

7. DISCUSIÓN

7.1. RNA utilizado, ligaciones y PCR

Todos los RNAs totales se mezclaron en un ‘pool’, porque el objetivo del trabajo no fue comparar las secuencias obtenidas en cada tratamiento, y por otro lado, ya que el método de purificación no permite un tratamiento cuantitativo del número de secuencias de cada smallRNA, no se obtiene mucha información haciendo una librería de cada tratamiento por separado.

A pesar de haber obtenido exitosamente una librería de smallRNAs (pero de secuencias muy cortas con respecto a los picos esperados en 21 y 24 nt), se quiso hacer un análisis sobre la calidad, pureza y cantidad del RNA total inicial utilizado,

tamaño de los microRNAs aislados del gel de acrilamida, tamaño de las ligaciones y PCR, para corroborar que todo el procedimiento de laboratorio se realizó correctamente y los tamaños cortados fueron los correctos.

El RNA total inicial utilizado, como se observa en la figura 4, es de buena calidad y no está degradado, ya que se diferencian claramente la bandas de RNA ribosomal característicos del RNA total cuando se visualizan en geles de agarosa. Estas son dos bandas superiores la cuales corresponden a RNA ribosomal 28S y 18S, siendo la primera dos veces más intensa que la segunda (Bryant 1998, Lu *et al.* 2007). También se observa una banda inferior que corresponde a RNAs de tamaños menores a 200 nt, por debajo de esta última banda están ubicados los smallRNAs, esto es cierto ya que cuando se hace el análisis electroforético en geles de acrilamida del RNA total, en el rango de 20 a 30 nt se detectan claramente 2 bandas de smallRNAs (figura 5c) (Lu *et al.* 2007). La calidad del RNA también se puede ver en la figura 5a y 5b, donde las bandas de rRNA 5.8S, 5S y RNA de transferencia se ven claramente definidas (Bryant 1998).

El RNA total inicial para la construcción de la librería de microRNAs fue de 60.8 ug, lo que está de acuerdo a la cantidad inicial necesaria para obtener los mejores resultados según el kit miRCat™ microRNA clonning (50-100 ug) o según Lau *et al.* (2001) y Lu *et. al* (2007) 50-500 ug (Ver Tabla 1). Por otro lado, el RNA total podría ser usado para clonar miRNAs, pero la masa absoluta de RNAs de bajo peso molecular es muy baja en comparación con los RNA más largos, por esta razón el paso de enriquecimiento (separación de RNAs de alto y bajo peso molecular) con smallRNAs es recomendado para la construcción de las librerías (Lu *et al.* 2007).

Por otro lado, las ligaciones de los smallRNAs a un adaptador en el extremo 3', en el extremo 5' y el PCR tienen los tamaños esperados de 40-45 nt (Ver figura 6), 60 nt (Ver figura 7) y 100-110 nt respectivamente (Ver figura 8) (Abdurakhmonov *et al.*

2008), por lo tanto, se puede decir que las ligaciones y el PCR se realizaron exitosamente.

7.2. Secuencias obtenidas

Inicialmente se quería hacer la clonación y secuenciación directa de los smallRNAs solo para mirar la diversidad de estos en el PCR inicial. Si se obtenía una gran diversidad de estos transcritos (similar a la de Abdurakhmonov *et al.* (2008), como se mencionara mas adelante), el PCR o los amplicones se utilizarían para ser secuenciados por el método de pirosecuenciación (secuenciación basada en pirofosfato que genera 400.000 reads de 400 a 500 pb) (Fahlgren *et al.* 2007) en un GS FLX de 454 (Genome Sequencer FLX System, 454 sequencing ROCHE, <http://www.454.com/applications/small-rna.asp>, Soares *et al.* 2009). Es por ello que se utilizaron en el momento de hacer el PCR primers modificados o “fusion primers”, es decir las secuencias de los primers del kit MiRCat (complementarios a los linkers), ligados a un primer A o un primer B de 19 pb en su extremo 3’ (miRCat™ microRNA cloning kit technical manual).

Para este tipo de secuenciación tan solo son necesarios 100 ng de producto de amplificación, para lo cual habrían disponibles 253.96 ng de PCR (esta lectura es confiable ya que el valor de R2 para la curva estándar fue de 0.9995, este valor se acerca a 1 (Ver figura 9) (Ver tabla 2)). En este sentido, si se hubiese obtenido secuencias entre tamaños de 20 a 25 nts, se hubiera secuenciado el PCR por 454. Se optó entonces por no secuenciar por el método de pirosecuenciación, y el análisis se hizo con las secuencias obtenidas inicialmente por el método convencional de Sanger (Hafner *et al.* 2008).

En resumen, se encontraron 6 familias diferentes de secuencias de smallRNAs, una de ellas conservada con *Ricinus communis* y con genes regulados asociados con la

fotosíntesis. No hay que descartar que las restantes 5 familias de secuencias sean transcritos propios de yuca, porque este trabajo es el primero en esta especie y la validación de los transcritos se hizo con base al genoma de otra especie.

7.3. MicroRNAs y factores que afectan su clonación

En general, la mayor frecuencia de microRNAs en especies de plantas está en los 21 nt y 24 nt (Lu *et al.* 2007), lo cual se puede corroborar en los histogramas de frecuencias para las especies *Populus trichocarpa* (Comparative Sequencing of plants small RNAs Project), *O. Sativa* y *Pinus contort* de las figuras 14 y 15 (Dolgosheina *et al.* 2008).

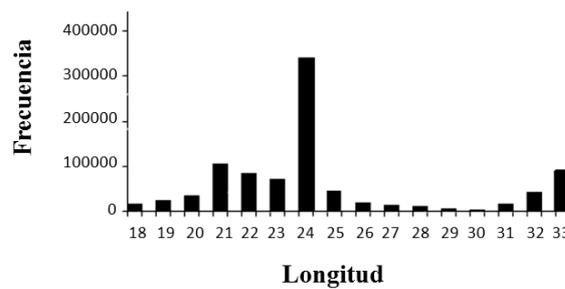


Fig. 14: Distribución de longitudes de smallRNAs de *Populus trichocarpa* (Comparative Sequencing of plants smallRNAs Project).

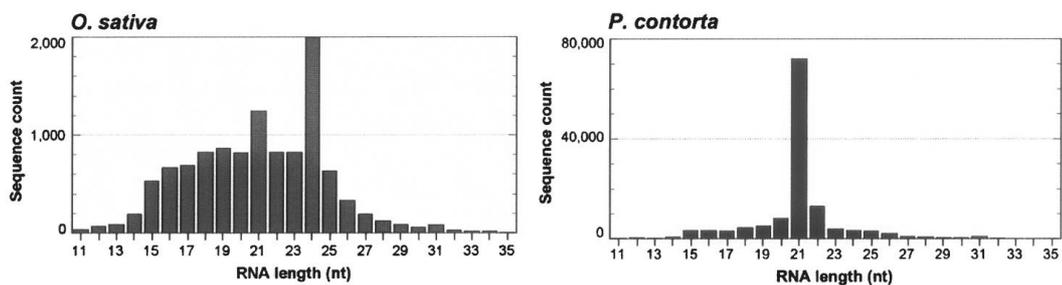


Fig. 15: Distribución de frecuencias de smallRNAs de de *O. Sativa* y *Pinus contort* (Dolgosheina *et al.* 2008).

Los dos picos, en 21 y 24 nt se pueden observar no solo en los resultados de la secuenciación de estos transcritos, si no desde el momento en que el RNA total se siembra en gel de acrilamida. En la siguiente figura (figura 16) se pueden detectar dos bandas de smallRNAs en 24 especies diferentes de plantas (sin ningún tratamiento). En algunas especies se expresan primordialmente transcritos de 24nt, en otros de 21 nt y en otras ambos tamaños en igual proporción se pueden presentar (Dolgosheina *et al.* 2008). Para *Manihot esculenta* (angiosperma eudicotiledónea (Ver figura 16, línea 20 -24)) se esperarían resultados similares a los de las líneas 20 a 24, esto si se observó, ya que esta especie bajo los tratamientos utilizados expresa en mayor proporción transcritos de 24 nt (ver figura 5) y niveles más bajos de RNAs de 21 nt (Lu *et al.* 2007); esto también se apoya en el hecho que la expresión de las enzimas DICER cambia dependiendo del estrés abiótico bajo el cual están sometidas y se ha confirmado que la expresión de DCL 1 (genera los microRNAs) decrece cuando las plantas están sometidas a sequía (Liu *et al.* 2009), lo anterior no quiere decir que no se obtengan smallRNAs bajo este tipo de estrés, si no que se expresan en menor proporción y por lo tanto son más difíciles de detectar.

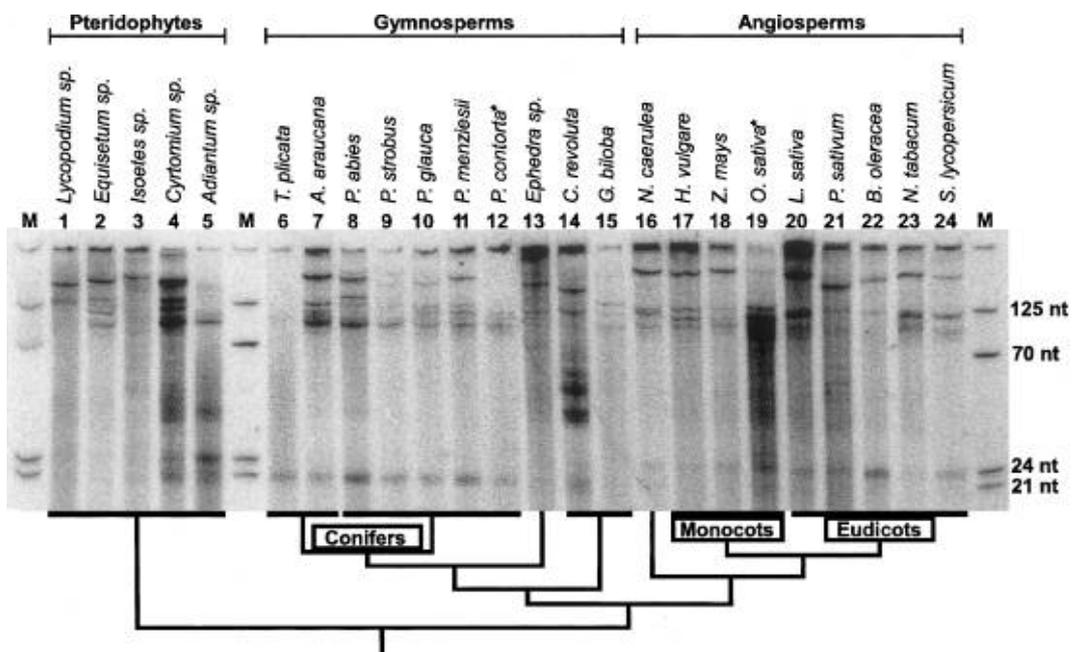


Fig. 16: RNA total sembrado el gel de acrilamida de 24 especies de plantas diferentes (Dolgosheina *et al.* 2008).

A partir del análisis de las secuencias, se puede afirmar que no hubo consistencia entre los tamaños de los smallRNAs (21 y 24 nt) aislados del gel de acrilamida (Ver figura 5) y los tamaños de las secuencias clonadas y secuenciadas. En la tabla 3 se observa que hubo una tendencia clara a secuenciar transcritos de bajo peso molecular (12 a 18), de menor tamaño a lo que se esperaba obtener (ver Figura 17), y en su mayoría agrupados en 4 secuencias diferentes (básicamente la misma secuencia pero con 1 a 3 nucleótidos adicionales) (Ver figura 10), esto puede deberse a que estas secuencias estaban más representadas en la muestra con respecto a otras secuencias (debido a que se derivan de transposones (Liu *et al.* 2009)).

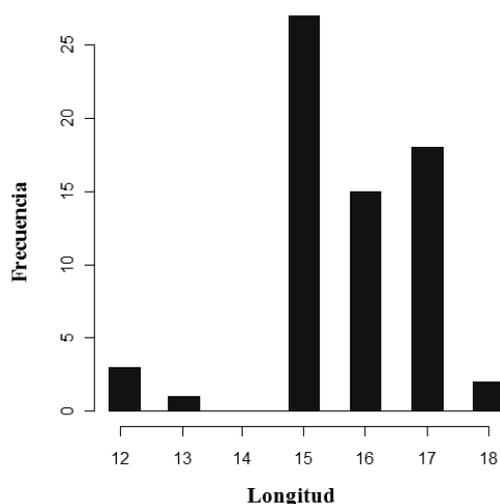


Fig. 17: Distribución de frecuencias de smallRNAs de Yuca secuenciados.

En las figuras 14 y 15 se puede observar que los smallRNAs pueden tener tamaños entre 11 y 34 nucleótidos con picos en 21 y 24 nt. Las secuencias obtenidas de yuca, al tener tamaño entre 15 a 18 nt (histograma de frecuencias figura 17), están dentro del rango de RNAs pequeños en plantas (aunque hay ausencia de otros tamaños), por

lo tanto, aunque una secuencia se pudo validar como candidato a microRNA en base al genoma de *Ricinus communis*, no podemos descartar que los demás sean transcritos únicos para yuca.

Abdurakhmonov *et al.* (2008) secuenció 6691 clones de 11 librerías diferentes comenzando con un RNA total de buena calidad, de los cuales solo encontró 583 microRNA diferentes; en este sentido si en 6691 clones se obtienen 583 secuencias diferentes, entonces se espera encontrar para yuca $(583/6691)*96 = 8.36$ secuencias diferentes por cada 96 clones secuenciados, en este trabajo se encontraron 6 familias diferentes, lo que no se aleja mucho del valor esperado. La librería secuenciada para yuca es casi tan diversa en secuencias como la de Abdurakhmonov *et al.* (2008). Sin embargo, y como se dijo anteriormente, dado que el tamaño de los smallRNAs obtenidos es menor del esperado para las dos fracciones más comunes de microRNAs de plantas (21 y 24 nt) se consideró que no tenía sentido someter ésta librería a secuenciación masiva.

7.4. E-values para resultados de BLAST

El valor esperado o E-value es un parámetro que describe el número de “hits” esperados por azar cuando se está buscando similitud de una secuencia problema en una base de datos de un tamaño particular, y es usado para cuantificar la significancia de los resultados. Este decrece exponencialmente cuando el puntaje del alineamiento (S) aumenta. En otras palabras el E-value es el número de alineamientos diferentes que son esperados al azar para determinado valor de S (score: es la suma del valor de los scores individuales de cada nucleótido) para una base de datos de tamaño determinado (NCBI, <http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html>, Korf *et al.* 2003).

Por lo anteriormente expuesto, los E-values obtenidos las búsquedas de similitud mediante BLAST (con MIRBASE y genoma de *Ricinus communis*) son valores

relativamente altos y aparentemente no significativos (desde 0.082 hasta 2.2) (Ver tabla 7). Entonces si se sabe que existe una relación inversa entre los E-values y los scores, y que cuando se busca homología de secuencias tan cortas en una base de datos tan grande, los scores siempre van a ser bajos, y al obtener valores bajos para S, se tendrán inevitablemente E-values con puntajes altos. (NCBI, <http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html>, Korf *et al.* 2003).

7.5. Comparaciones.

Inicialmente se hizo la comparación de las secuencias obtenidas con la base de datos MIRBASE con el objetivo de encontrar secuencias conservadas con otras especies de plantas. Con respecto a los resultados de esta comparación (hecha desde el terminal) se puede decir que los hits de secuencias homólogas a los potenciales microRNAs de yuca secuenciados, al tener una longitud tan corta (9 y 11 nt), y al no haber obtenido hits con microRNAs de especies de plantas (solo con virus de humano y mosca) (Ver Tabla 4), no se tomarán en cuenta estos resultados y se da por hecho que las secuencias de miRNAs de yuca pueden ser únicos de esta especie y por lo tanto no se generaron resultados significativos al hacer estas comparaciones, porque se sabe que la mayoría de los microRNAs no están conservados entre las especies (Moxon *et al.* 2008). Lo cual se apoya con los resultados obtenidos al realizar el blast desde la página de MIRBASE, en donde no se obtuvo ningún resultado.

Como se planteara más adelante, es probable que las secuencias obtenidas tengan su origen a partir de secuencias repetidas, en este sentido, otro motivo por el cual se cree que no se encontró ningún hit al hacer la comparación con la base de datos MIRBASE, es el hecho que transcritos originados de secuencias genómicas repetidas han sido excluidos de esta base de datos (Piriyaongsa *et al.* 2007).

Posteriormente, se quiso verificar que las secuencias obtenidas no correspondían a productos de degradación de RNA ribosomales o a genes codificantes (proteínas) (Ver tabla 5 y 6). Los resultados mostraron que 5 de las secuencias y sus ESTs de yuca similares, eran secuencias de RNA ribosomal 26S, por esta razón, se excluyeron de los análisis posteriores. Por el contrario, con las 11 secuencias restantes que no correspondían a RNA ribosomal ni proteínas, se continuó con el análisis bioinformático.

El siguiente paso fue hacer una búsqueda de similitud mediante BLAST de los potenciales microRNAs de yuca vs. genoma de *Ricinus communis* (Ver tabla 7), para poder extraer las secuencias adyacentes (80 pb a cada lado) y hacer el análisis de su estructura secundaria y de sus promotores.

Cuando se hizo la búsqueda de similitud en el genoma de *Ricinus communis* y se consideraron e-values > 2 , se obtuvieron más hits de lo esperado al azar, lo que concuerda con que el smallRNA corresponde a un transposon, lo que es característico en plantas (Moxon *et al.* 2008) y sobre todo cuando están bajo algún estrés (Wessler 1996), lo cual también se corroboró a su vez cuando se aplicó RepeatMasker con la base de datos de secuencias repetidas de *Arabidopsis*, obteniéndose como origen de este potencial microRNA un tipo de retrotransposon denominado ATGP5I LTR de la familia Gypsy.

En este sentido, los microRNAs pueden derivarse secuencias repetidas tales como MITEs (miniature inverted-repeat transposable elements), long interspersed elements (LINEs), short interspersed elements (SINEs) y long terminal repeat (LTR) retrotransposons (Piriyapongsa *et al.* 2008). Piriyapongsa *et al.* (2008) propusieron que algunos microRNAs de plantas son codificados por un tipo de DNAs cortos TEs (transposable elements) conocidos como MITEs. Estos TEs pueden ser expresados como transcritos de regiones intrónicas (Kurihara *et al.* 2004, Piriyapongsa *et al.* 2007) de longitudes de 80 a 500 pb, que pueden doblarse en forma de hebilla

(estructuras características de los pre-miRNA), son reconocidos por la maquinaria enzimática de DICER y son procesados hasta secuencias maduras de microRNAs (Piriyaongsa *et al.* 2007, Piriyaongsa *et al.* 2008).

En plantas varios precursores de microRNAs se originan de secuencias genómicas repetidas. Tal como se obtuvo con la clonación de smallRNAs de yuca, secuencias repetidas son rutinariamente observadas durante la clonación de microRNAs, pero muchas veces al ser interpretados como rasiRNAs (repeat-associated small interfering RNAs) derivados de dsRNA (Double-stranded RNA) (Aravin *et al.* 2003, Chen *et al.* 2005, Moxon *et al.* 2008), son excluidos de MIRBASE u otras bases de datos, por lo que se está subestimando los efectos en regulación de muchos microRNAs (Aravin *et al.* 2003), y por este motivo, en el caso que las secuencias obtenidas para yuca tengan su origen a partir de secuencias repetidas, se puede explicar por qué no se obtuvieron resultados al compararlos con la base de datos de microRNAs (MIRBASE).

Por otro lado, se ha evidenciado que en plantas los transposones se activan cuando éstas están bajo estrés biótico o abiótico (Wessler 1996). Se ha observado que la temperatura tiene efectos sobre varios elementos transponibles, es posible que en muchos casos estos efectos no sean directos; es decir, que un organismo que está bajo condiciones de estrés puede inducir una o varias respuestas a nivel fisiológico o genético, y estos a su vez actuarán activando maquinarias transcripcionales (Capy *et al.* 2000), con respecto a esto, como las plantas de yuca crecidas *in-vitro* estuvieron sometidas a estrés abiótico (sequía y calor) es probable que en las plantas se hayan activado transposones, y los microRNAs derivados de estos se hayan clonado y secuenciado más fácilmente.

En adición, estas secuencias repetidas de TEs se pueden definir como un mecanismo para generar múltiples microRNAs y así regular más fácilmente los genes diana a lo largo del genoma o compensar la expresión de genes de microRNAs que presenten mutaciones (Piriyaongsa *et al.* 2007), por otro lado, la inserción de transposones en

nuevos lugares del genoma puede dirigir la creación de nuevos microRNAs durante la evolución de los mamíferos y posiblemente en otros organismos tales como plantas. Los microRNAs nuevos, derivados de elementos transponibles, pueden ser particularmente relevantes, ya que pueden regular genes que no son conservados entre especies (Smalheiser *et al.* 2005).

7.6. Predicción de estructuras secundarias

Luego de obtener las secuencias flanqueantes, se realizó un análisis de las estructuras secundarias que forman sus pre-microRNAs a una temperatura de plegamiento de 37°C (Yin *et al.* 2008).

La predicción de las estructuras secundarias de estas secuencias genómicas, se obtuvo teniendo en cuenta los siguientes parámetros: gran parte de la secuencia del miRNAs debe estar incluido en el loop terminal de la estructura en hebilla; los potenciales precursores deben tener un valor bajo (negativo) de energía libre, lo que nos indica que son hélices en óptimo doblamiento; el contenido A+U debe estar entre 30 y 70 % y por último el miRNA debe tener máximo 6 nucleótidos no apareados con la secuencia opuesta (miRNA*) del otro brazo del loop (Lu *et al.* 2008, Yin *et al.* 2008).

Según lo anterior, la secuencia de microRNA incluida en el Contig30165 de *Ricinus communis* (Ver figura 11) es un buen candidato a microRNA ya que está inmersa en una estructura en tallo y asa, la energía libre de doblamiento es muy baja (-69.00 kcal/mol), tiene 50 % de contenido A+U y tiene 4 nucleótidos no apareados con el microRNA opuesto, mientras que en las otras dos secuencias no hay evidencia que sean microRNAs ya que tienen más de 6 nucleótidos no apareados con el microRNA*, por lo cual se excluyen de los siguientes análisis (Lu *et al.* 2008).

7.7. Análisis de promotores.

Para el análisis de promotores de genes tipo II solo se analizó el Contig30165, ya que cumplió con los parámetros característicos de las estructuras secundarias de los pre-microRNAs (Lu *et al.* 2008).

Para entender el mecanismo y funciones de la activación de los genes de microRNAs, se requiere localizar sus regiones promotoras centrales. Houbaviy *et al.* (2005) encontraron una caja TATA localizada a -35 del inicio del pri-microRNA. Xie *et al.* (2005) identificaron los promotores de 52 genes de microRNA de *A. thaliana*, y mostraron que estos tenían una caja TATA en sus promotores centrales. Ohler *et al.* (2004) por su lado escaneó secuencias de 1000 pb corriente arriba de los genes de microRNAs de *Drosophila* con el objetivo de encontrar motivos conocidos correspondientes a promotores, pero no detectó una preferencia consistente para ningún motivo reportado para genes que codifican a microRNAs.

Zhou *et al.* (2007) mediante un nuevo método de predicción de promotores para genes de microRNAs CoVote (common query voting), mostraron que la mayoría de los promotores tienen motivos de entre 5-10 pb de longitud. La mayoría de los motivos de los promotores se encontró que están localizados a -500 pb corriente arriba de la región de pre-microRNA, otros motivos significativos fueron identificados de -1k a -500 pb, otros de -1.5k a -1kpb y de -2k a -1.5kpb corriente arriba del pre-miRNA. En este mismo trabajo se presentan las secuencias de los motivos correspondientes a promotores de genes que codifican a miRNAs, también depositados en la página <http://cic.cs.wustl.edu/microrna/promoters.html>. En *C. elegans*, uno de los motivos más significativos tiene una secuencia consenso TTTCAATTTTTC, el cual es un elemento Inr (iniciador). La caja TATA es un motivo muy bien conocido en los promotores centrales para genes eucariotas de clase II, y descubierto para genes de microRNAs por Zhou *et al.* (2007) en *A. thaliana* y *O. sativa* (TATATATA).

De acuerdo con lo anterior, la secuencia de 2000 pb corriente arriba del inicio del pre-microRNAs contiene elementos promotores tales como caja TATA y caja CAAT, características de genes tipo II (genes de microRNAs) con valores de similitud cercanos a 1 (perfecta complementariedad). Las secuencias de los motivos varían en algunos nucleótidos pero básicamente corresponden a los mismos promotores (ver fig. 12 y tabla 8).

Con lo mencionado y a los análisis realizados previamente, se puede concluir que en la librería construída se clonó eficientemente (86%) una secuencia de potencial microRNA (familia 3, Ver tabla 3) que se validó con base al genoma de *Ricinus communis* y posiblemente derivada de transposones.

7.8. Potenciales genes diana

Encontrar los mRNAs diana es esencial para entender las funciones biológicas de los microRNAs. Diferentes métodos se han utilizado para predecir genes diana de miRNAs de plantas y animales. En plantas los mRNAs regulados pueden ser predichos por similitud entre las secuencias ya que se ha visto que los miRNAs de plantas tienen una perfecta complementariedad con los RNAs mensajeros (Lai 2004, Moxon *et al.* 2008).

Como resultado de la búsqueda de mRNAs diana mediante MIRANDA, se obtuvieron múltiples hits, pero solamente 18 ESTs son regulados por la secuencia candidato a microRNA (Ver figura 13), cuya secuencia es GCTGACCCTGAAGTTCAT. Los criterios fueron los siguientes: 1) 4 o menos nucleótidos no apareados en el sitio complementario entre las secuencias de los miRNAs y los potenciales mRNAs diana, 2) Un solo nucleótido no apareado de la región complementaria al microRNA entre las posiciones 2 y 12, pero no en las

posiciones 10 o 11, y 3). Hasta 3 nucleótidos no apareados adicionales entre 12 nt y 23 nt pero no más de dos nucleótidos no apareados continuos en esta región (Yin *et al.* 2008).

7.9. Funciones de los genes regulados (Ver tabla 9).

En la literatura no se ha reportado que los microRNAs estén asociados directamente a la fotosíntesis, y aunque se ha identificado que están involucrados en crecimiento y desarrollo (Zhang *et al.* 2007), no todos necesariamente participan en estos procesos. A continuación se mostrarán algunos procesos regulados por la secuencia de microRNA sugerido para de yuca, los cuales están relacionados con la fotosíntesis.

• Proteína de formación de tilacoide 1

El desarrollo de cloroplastos normales es un paso crucial para la sobrevivencia de eucariotas fotosintéticos. La regulación de genes que codifican a proteínas de los cloroplastos puede tener varios efectos sobre algunos procesos metabólicos que ocurren en este organelo, y por ende en el crecimiento y desarrollo normal de las plantas. El gen *Thf1* juega un papel crucial en el proceso dinámico de la biogénesis de la membrana tilacoidal mediada por vesículas (Wang *et al.* 2004a). Es decir, que si los tilacoides no están bien formados, la fotosíntesis no se puede llevar a cabo.

El desarrollo de las membranas de los tilacoides depende del transporte de vesículas de membrana en el interior del cloroplasto y la posterior fusión de vesículas en el interior de este. El gen de formación 1 del tilacoide (*Thf1*) controla este importante paso requerido para la organización normal de estas vesículas dentro de las grana maduras (grupo de tilacoides) y en última instancia en el desarrollo de las hojas. Altos niveles de *Thf1* se inducen gracias a la luz y decrecen en condiciones de oscuridad. Interrupciones en la expresión de este gen, resulta en plantas muy delgadas (ya que se ve afectada la fotosíntesis), con variedad de patrones de hojas, algunas de estas

presentando decoloración (blanco o amarillo), pero lo que se ve más afectado es la organización de las estructuras tilacoidales (reducidas cantidades de tilacoides y la distancia entre las granas más amplias a diferencia de un fenotipo normal) debido a la acumulación de vesículas en el interior de los cloroplastos (Wang *et al.* 2004a).

- **Proteína reguladora de luz en el cloroplasto**

La luz sirve como fuente de energía y regulador de señal para la fotosíntesis. La irradianza (incidencia de luz por unidad de superficie) y la calidad del espectro de luz a la que son expuestas las plantas controla la capacidad de las hojas para conducir la fotosíntesis y procesos asociados tal como transporte de electrones y fotofosforilación. La luz regula estas actividades fotosintéticas por el control de los niveles de proteínas cloroplásticas asociadas con cada función. Burkey (1992) identificó un polipéptido de membrana tilacoidal que regula la cantidad de luz en respuesta a la irradianza. Las membranas de las plantas que crecen bajo alta irradianza presentan mayores cantidades de proteínas en las membranas a diferencia de aquellas que estaban expuestas a menor cantidad de luz (Burkey 1992).

De acuerdo a lo anterior, la secuencia de potencial microRNA de yuca está involucrada en un importante proceso de formación de los cloroplastos y regulación de la cantidad de luz. No se debería asignar funciones específicas a la secuencia analizada, sin realizar estudios de expresión, por lo tanto, este trabajo crea las bases para hacer los estudios necesarios que puedan indicar nuevas formas de regulación de procesos, como la fotosíntesis.

8. CONCLUSIONES

La metodología utilizada permitió la caracterización bioinformática y funcional de un conjunto de smallRNAs entre 12 y 18 nt. Pero no se caracterizó la población de microRNAs de los tamaños usuales (21, 24 nt).

Los smallRNAs no validados, pueden ser sin embargo, microRNAs propios de yuca, que han divergido significativamente de *Ricinus*, y por tanto no se prestan al análisis comparativo que se planteó. Para descartarlos como microRNAs necesitaríamos tener disponible el genoma de la yuca.

9. RECOMENDACIONES Y TRABAJO FUTURO

Hacer ensayos de northern blot del miRNA candidato en las diferentes condiciones de estrés.

Dar explicación al porque no se obtuvieron secuencias de tamaño entre 19 y 30 nt en la librería secuenciada, evaluando los tamaños obtenidos en cada uno de los pasos desde la obtención del PCR hasta la secuenciación.

En este sentido, se observó que el tamaño de los smallRNAs del PCR fue de 20 a 30 nucleótidos. Para corroborar que se clonaron secuencias dentro de este rango (20 - 30 nt), se aisló el plásmido de la librería total, obteniéndose un tamaño correcto entre 20 y 30 nt.

Por otro lado, los productos de secuenciación tenían tamaños entre 12 y 18 nt, lo que indica que hubo un sesgo al momento de elegir las colonias. En este paso solo se eligieron colonias blancas aquellas en donde el inserto inactiva la producción de B-galactosidasa. No se eligieron colonias azules, porque se consideraron como negativas, pero estas últimas también pueden contener los insertos siempre y cuando el tamaño de estos sea múltiplo de 3 y no contenga codones de parada para la producción de la proteína (pGEMT-easy technical manual). En este sentido, el paso siguiente será secuenciar las colonias azules.

10. REFERENCIAS BIBLIOGRAFICAS

454 life sciences. Sequencing services. Protocols. 2008. Cloning small RNAs for sequencing with 454 Technology.

http://www.454.com/downloads/hannon_smallRNA-cloning_protocol2.pdf.

Abdurakhmonov, I.Y., Devor, E.J., Buriev, Z.T., Huang, L., Makamov, A., Shermatov, S.E., Bozorov, T., Kushanov, F.N., Mavlonov, G.T., Abdugarimov, A., 2008. Small RNA regulation of ovule development in the cotton plant, *G. hirsutum* L. BMC Plant Biol 8, 93.

ABI PRISM® Big Dye terminator V1.1. cycle sequencing Ready Reaction kit. Technical manual.

http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_042775.pdf. <http://www.cogentech.it/Sequencing/pdf/bdterm.pdf>.

Advantage® 2 Polymerase Mix (CLONTECH) Technical manual. <http://www.clontech.com/images/pt/PT3281-1.pdf>.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25, 3389-3402.

Alvarez-luna, E. 1972. Cassava (*Manihot esculenta* Crantz) a tropical root crop. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 3 p.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., Tuschl, T., 2003. The small RNA profile during *Drosophila melanogaster* development. *Dev Cell* 5, 337-350.

Axtell, M.J., Bartel, D.P., 2005. Antiquity of microRNAs and their targets in land plants. *Plant Cell* 17, 1658-1673.

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-297.

BigDye® Terminator v1.1 Cycle Sequencing Kit. Technical manual. <http://www.cvm.tamu.edu/dnacore/Documents/bigdyev1.1.pdf>.

Bohnsack M, Czaplinski K, Gorlich D. 2004. Exportin 5 is a RanGTPdependent dsRNA-binding protein that mediates nuclear export of premiRNAs. *RNA* 10: 185–191.

Bonnet, E., Y. Van de Peer & P. Rouzé. 2006. The small RNA world of plants. *New Phytologist*. 171: 451–468.

Bryant, S and Manning, D. L. 1998. Formaldehyde Gel Electrophoresis of Total RNA. Vol. 86. RNA isolation and characterization protocols. Rapley, R. and Manning, D. L., eds. Humana Press , Totowa, NJ. Páginas 69-72. Capítulo 13.

Burkey, K.O., 1992. Novel Light-Regulated Chloroplast Thylakoid Membrane Protein. *Plant Physiol* 98, 1211-1213.

Capy, P., Gasperi, G., Biemont, C., Bazin, C., 2000. Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85 (Pt 2), 101-106.

Ceballos, H. 2002. La yuca en Colombia y el mundo: nuevas perspectivas para un cultivo milenario. P. 1-13. En: La yuca en el tercer milenio: sistemas modernos de producción y procesamiento, utilización y comercialización. Cali. Colombia. CIAT.

Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D., Sander, C., Zavolan, M., Tuschl, T., 2005. The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev* 19, 1288-1293.

Crick, F. 1970. Central dogma of molecular biology. *Nature*. 227: 561-563.

Comparative sequencing of plants small RNAs project.
http://smallrna.udel.edu/project_data.php?download=1&src=http%3A%2F%2Fsmallrna.udel.edu%2F

Dolgosheina, E.V., Morin, R.D., Aksay, G., Sahinalp, S.C., Magrini, V., Mardis, E.R., Mattsson, J., Unrau, P.J., 2008. Conifers have a unique small RNA silencing signature. *RNA* 14, 1508-1515.

Domínguez, O., & L. F. Ceballos. 1979. Clasificación taxonómica y morfología de la planta de yuca (*Manihot esculenta* Crantz). Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 10 p.

Droege, M., Hill, B., 2008. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. *J Biotechnol* 136, 3-10.

Dugas, D.V., Bartel, B., 2004. MicroRNA regulation of gene expression in plants. *Curr Opin Plant Biol* 7, 512-520.

Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., Marks, D.S., 2003. MicroRNA targets in *Drosophila*. *Genome Biol* 5, R1.

Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Law, T.F., Grant, S.R., Dangl, J.L., Carrington, J.C., 2007. High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS ONE* 2, e219.

Farrell, R. E. 2005. RNA methodologies. A laboratory guide for isolation and characterization. Third edition. Academic Press. Boston. Pages 163-178.

Golubchik, T., Wise, M.J., Easteal, S., Jermiin, L.S., 2007. Mind the gaps: evidence of bias in estimates of multiple sequence alignments. *Mol Biol Evol* 24, 2433-2442.

Griffiths-Jones, S. 2004. The microRNA registry. *Nucl. Acids Res.* 32: D109–D111.

Hafner, M., Landgraf, P., Ludwig, J., Rice, A., Ojo, T., Lin, C., Holoch, D., Lim, C., Tuschl, T., 2008. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 44, 3-12.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* 27, 297-300.

Houbaviy, H.B., Dennis, L., Jaenisch, R., Sharp, P.A., 2005. Characterization of a highly variable eutherian microRNA gene. *Rna* 11, 1245-1257.

ImProm-II™ Reverse Transcription System (PROMEGA) Technical manual. <http://www.promega.com/tbs/tm236/tm236.pdf>.

Integrated DNA Technology. <http://www.idtdna.com>.

Jones-Rhoades, M.W., Bartel, D.P., Bartel, B., 2006. MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57, 19-53.

Kidner, C.A., Martienssen, R.A., 2003. Macro effects of microRNAs in plants. *Trends Genet* 19, 13-16.

Kim, J., Jung, J.H., Reyes, J.L., Kim, Y.S., Kim, S.Y., Chung, K.S., Kim, J.A., Lee, M., Lee, Y., Kim, V.N., Chua, N.H., Park, C.M., 2005. MicroRNA directed cleavage of ATHB15 mRNA regulates vascular development in Arabidopsis inflorescence stems. *Plant J.* 42, 84–94.

Krol J, Krzyzosiak W. 2006. Structure analysis of microRNA precursors. *Methods Mol Biol* 342: 19–32.

Korf, I., M. Yandell & J. Nedell. 2003. BLAST. Primera edición. Editorial O'REILLY.USA.

Kurihara, Y., Watanabe, Y., 2004. Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci U S A* 101, 12753-12758.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T., 2001. Identification of novel genes coding for small expressed RNAs. *Science* 294, 853-858.

Lai, E.C., 2004. Predicting and validating microRNA targets. *Genome Biol* 5, 115.

Lau, N., Lee P. Lim, Earl G. Weinstein, and D. P. Bartel. 2001. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in *Caenorhabditis elegans*. *Science*. 294 (5543): 858-862.

Lee Y, Ahn C, Han J, Choi H, Kim J, et al. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425: 415–419.

Lee Y, Kim M, Han J, Yeom K, Lee S, Baek, s., Kim, N. 2004. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23: 4051–4060.

Lindow, M. & A. Krogh. 2005. Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics*.6:119

Liu, Q., Feng, Y., Zhu, Z., 2009. Dicer-like (DCL) proteins in plants. *Funct Integr Genomics*.

Lu, C., Meyers, B.C., Green, P.J., 2007. Construction of small RNA cDNA libraries for deep sequencing. *Methods* 43, 110-117.

Lu, Y., c, Gan, Q., Chi, X. Qin, S., 2008. Identification and Characterization of MicroRNAs and Their Targets in Grapevine (*Vitis vinifera*). *Agricultural Sciences in China*. 7 (8), 929-943.

Lund E, Guttinger S, Calado A, Dahlberg J, Kutay U. 2004. Nuclear export of microRNA precursors. *Science* 303: 95–98.

Manning, V.A., Andrie, R.M., Trippe, A.F., Ciuffetti, L.M., 2004. Ptr ToxA requires multiple motifs for complete activity. *Mol Plant Microbe Interact* 17, 491-501.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani,

V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.

MiRBase. <http://microRNAs.sanger.ac.uk>.

MirCat, MicroRNA cloning kit technical manual.
http://www.idtdna.com/Support/Technical/TechnicalBulletinPDF/miRCat_User_Guide.pdf.

MiRU. Target Finder. Plant microRNA Potential Target Finder,
<http://bioinfo3.noble.org/miRNA/miRU.htm>.

mirVANA miRNA Isolation kit Instruction manual.
http://www.ambion.com/techlib/prot/fm_1560.pdf.

Moxon, S., Jing, R., Szittyá, G., Schwach, F., Rusholme Pilcher, R.L., Moulton, V., Dalmay, T., 2008. Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res* 18, 1602-1609.

National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/>.

Naqvi, A.R., Islam, M.N., Choudhury, N.R., Haq, Q.M., 2009. The fascinating world of RNA interference. *Int J Biol Sci* 5, 97-117.

pGEM®-T and pGEM®-T Easy Vector Systems instruction manual.
<http://www.promega.com/tbs/tm042/tm042.pdf>.

Phillips, J.R., Dalmay, T., Bartels, D., 2007. The role of small RNAs in abiotic stress. *FEBS Lett* 581, 3592-3597.

Piriyapongsa, J., Jordan, I.K., 2008. Dual coding of siRNAs and miRNAs by plant transposable elements. *Rna* 14, 814-821.

Piriyapongsa, J., Marino-Ramirez, L., Jordan, I.K., 2007. Origin and evolution of human microRNAs from transposable elements. *Genetics* 176, 1323-1337.

Rapley, R and Heptinstall, J. 1998. UV Spectrophotometric Analysis of Ribonucleic Acids. Vol. 86. RNA isolation and characterization protocols. Rapley, R. and Manning, D. L., eds. Humana Press , Totowa, NJ. Paginas 65-68. Capítulo 12.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., Bartel, D.P., 2002. MicroRNAs in plants. *Genes Dev* 16, 1616-1626.

Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., Bartel, D.P., 2002. Prediction of plant microRNA targets. *Cell* 110, 513-520.

Ruby, J.G., Jan, C.H., Bartel, D.P., 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83-86.

Sakurai, T., Plata, G., Rodríguez-Zapata, F., Seki, M., Salcedo, A., Toyoda, A., Ishiwata, A., Tohme, J., Sakaki, J., Shinozaki, K. and Ishitani, M. 2007. Sequencing analysis of 20,000 full-length cDNA clones from cassava reveals lineage specific expansions in gene families related to stress response. *BMC Plant Biology*. 7:66.

Shao-Yao Ying. 2006. *MicroRNA Protocols*. Segunda edición. Humana Press, Japón. 189 p.

Smalheiser, N.R., Torvik, V.I., 2005. Mammalian microRNAs derived from genomic repeats. *Trends Genet* 21, 322-326.

Soares, A.R., Pereira, P.M., Santos, B., Egas, C., Gomes, A.C., Arrais, J., Oliveira, J.L., Moura, G.R., Santos, M.A., 2009. Parallel DNA pyrosequencing unveils new zebrafish microRNAs. *BMC Genomics* 10, 195.

Sunkar, R. & J.K. Zhu. 2004. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*, *Plant Cell*. 16: 2001–2019.

SuperScript® III Reverse Transcriptase kit. Instruction manual (INVITROGEN).
<http://www.vdrc.at/fileadmin/conferences/VDRC/protocols/1stStrandcDNA.pdf>.

Wang, Q., Sullivan, R.W., Kight, A., Henry, R.L., Huang, J., Jones, A.M., Korth, K.L., 2004a. Deletion of the chloroplast-localized Thylakoid formation1 gene product in *Arabidopsis* leads to deficient thylakoid formation and variegated leaves. *Plant Physiol* 136, 3594-3604.

Wang, X.J., Reyes, J.L., Chua, N.H., Gaasterland, T., 2004b. Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol* 5, R65.

Wessler, S.R., 1996. Turned on by stress. Plant retrotransposons. *Curr Biol* 6, 959-961.

Yan, S., Yan, C.J., Gu, M.H., 2008. [Molecular mechanism of leaf development]. *Yi Chuan* 30, 1127-1135.

Yang, T., L. Xue & L. An. 2007. Functional diversity of miRNA in plants. *Plant Science* 172: 423–432.

Yao Y., G. Guo, Z. Ni, R. Sunkar, J. Du, J. Zhu & Q. Sun. 2007. Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). *Genome Biology*. 8: R96.1 –R96.13.

Yin, Z., C. Li, X. Han & F. Shen. 2008. Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). *Gene* 414 : 60 – 66.

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., Carrington, J.C., 2005. Expression of Arabidopsis MIRNA genes. *Plant Physiol* 138, 2145-2154.

Zhang, B., Pan, X., Cobb, G.P., Anderson, T.A., 2006. Plant microRNA: a small regulatory molecule with big impact. *Dev Biol* 289, 3-16.

Zhang, B.H., Pan, X.P., Wang, Q.L., Cobb, G.P., Anderson, T.A., 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Res* 15, 336-360.

Zhang, B., Wang, Q., Pan, X., 2007. MicroRNAs and their regulatory roles in animals and plants. *J Cell Physiol* 210, 279-289.

Zhou, X., Ruan, J., Wang, G., Zhang, W., 2007. Characterization and identification of microRNA core promoters in four model species. *PLoS Comput Biol* 3, e37.

11. ANEXOS

Anexo 1: Extracción de RNA total con el kit mirVANA™ miRNA Isolation (AMBION, USA).

Los 0.2 g de cada muestra se colocó en un mortero previamente enfriado y se maceró varias veces con nitrógeno líquido hasta obtener un polvo fino. Antes de que el tejido se descongelara, se agregaron 2 ml buffer de lisis (1 ml por cada 0.1 g de tejido).

Extracción Orgánica

1. De la mezcla tomar alícuotas de 500 µl para transferirlas a tubos 1.5 nuevos. Luego agregar a cada tubo 50 µl (es decir 1/10 del volumen) de miRNA homogenate additive, mezclar por inversión o en vortex y llevar a hielo por 10 min.
2. Agregar un volumen igual de ácido fenol-cloroformo al que había antes de agregarle el miRNA homogenate additive (es decir si el volumen original era de 500 µl, se le adicionan 500 µl de fenol cloroformo).
3. Dar vortex por 30 a 60 segundos. Centrifugar por 5 minutos a 10.000 rpm para separar la fase acuosa de la fase orgánica. Transferir el sobrenadante o fase acuosa a un nuevo tubo, teniendo en cuenta el volumen que se obtiene.

Extracción de RNA total

4. Adicionar 1.25 volúmenes de etanol 100% (a temperatura ambiente) a la fase acuosa obtenida de la extracción orgánica, mezclar por inversión o en vortex.
5. Pipetear 700 µl y pasarlo por un filtro a otro tubo, después centrifugar alrededor de 15 seg a 10000 rpm. Descartar el filtrado.
6. Agregar 700 µl de la solución 1 de lavado al filtro y centrifugar por 10 seg a 10.000. Descartar la fase acuosa.
7. Agregar 500 µl de la solución de lavado 2/3. Centrifugar 10 seg a 10.000. Descartar la fase acuosa. Repetir paso.

8. Transferir el filtro a un tubo Nuevo y aplicar al centro de este 100 µl de agua libre de RNAsas, calentada previamente a 95°C. Centrifugar por 20 a 30 seg a máxima velocidad. Luego recuperar la fase acuosa que queda en el tubo rica en RNA. Descartar filtro, y almacenar a -80 °C o -20 °C.

Anexo 2: Procedimiento adicional para extracción de RNA total de raíces de plantas de campo.

Agregar a 200 ug de tejido macerado, 500 ul de buffer de lisis (acetato de sodio 100 mM pH 6.1, EDTA 1 mM y SDS 4%). Dar vortex por 1 a 2 minutos. Luego agregar 500 ul de fenol acido (pH 4.3):cloroformo: alcohol isoamílico (24:24:1), dar vortex por 1 a 2 minutos y centrifugar a 10500 rpm por 5 minutos a 4°C. Rescatar el sobrenadante y colocarlo en un tubo nuevo de 1,5 ml. Proceder con el paso 1 del Anexo 1 (CIAT).

Anexo 3: Gel de acrilamida denaturante al 15% (urea 7M) y aislamiento de microRNAs (miRCat™ microRNA cloning kit, (Lu *et al.* 2007).

1. En una cámara de electroforesis horizontal (PROTEAN de BIORAD), servir un gel de acrilamida 15%, urea 7M (para un volumen de 15 ml, mezclar 7.2 g de urea, 5.6 ml de acrilamida 40 %, 1.5 de TBE 10 X, 0.15 de amberlitas, completar con agua libre de RNAsas hasta 15 ml: Para polimerizar el gel agregar 75 ul de persulfato de amonio y 15 ul de TEMED), poner peine y dejar polimerizar por 30 minutos a temperatura ambiente. Retirar el peine y poner el gel en una cámara con TBE 1X a 50°C. Lavar los pozos y correr 5 ul de buffer carga verificar que haya quedado bien preparado. Agregar 1 volumen de volumen de buffer carga al RNA y denaturar a 80°C. Luego correr el RNA total, si es posible en un mismo pozo (Lu *et al.* 2007).

Anexo 4: Purificación de miRNAs, ligación 3', ligación 5' y PCR a partir de un método estándar de macerado y remojo "Crush and Soak" de gel de acrilamida (miRCat™ microRNA cloning kit).

Una vez el fragmento de gel esta en los tubos 1.5 ml:

1. Usando maceradores plásticos para microtubos 1.5 ml, macerar bien el gel.
2. Agregar un volumen igual de agua libre de nucleasas (pesar el trozo de gel para llevarlo a un volumen de 1 µl/mg).
3. Dar vortex por 15-30 segundos.
4. Calentar la mezcla a 70°C por 10 minutos.
5. Dar vortex por 15-30 segundos.
6. Llevar a -80°C por 30 minutos.
7. Repetir pasos 3 a 6 dos veces más.
8. Después del paso final a -80°C, dejar que el macerado se descongele a temperatura ambiente y centrifugar a 1200 rpm por 5 minutos.
9. Transferir el sobrenadante a un Nuevo tubo y completar el volumen a 500 ul con agua libre de RNAsas.
10. Antes de filtrar el RNA, lavar las columnas NAP-5 3 veces con agua libre de RNAsas.
11. Servir la muestra sobre las columnas y esperar hasta que la muestra haya entrado completamente.
12. Agregar 1ml de agua pura y en un tubo nuevo coleccionar el filtrado.
13. Precipitar con 3 ul de glicógeno (10 mg/ml), 1/10 de acetato de sodio 3M (pH 5.3) y 2.5 volúmenes etanol absoluto (-20°C). Centrifugar, lavar pellet con etanol 100 % y resuspender en 10 ul de H₂O-DEPC .

Anexo 5: Construcción y clonación de librería de microRNAs (miRCat™ microRNA cloning kit) (Lu *et al.* 2007).

a. Ligación de adaptador en 3’: En un tubo 1.2 ml, agregar lo siguiente para un volumen final de 20 μ l (miRCat™ microRNA cloning kit):

<i>Reactivo y concentración</i>	<i>Cantidad (μl)</i>
microRNAs purificados	10 μ l
Linker 3’ (50 μ M)	1 μ l
Buffer de ligación 10X	2 μ l
Potenciador de ligación	6 μ l
RNA ligasa T4 (1 U/ μ l)	1 μ l

Incubar esta reacción a 22°C por 2 horas. Y luego,

1. Agregar 80 μ l de IDTE (pH 7.5)
2. Transferir todo el volumen a un tubo nuevo 1.5 ml
3. Agregar 3 μ l de glicógeno (10 mg/ml)
4. Agregar 1/10 del volumen (10 μ l) de 3.0 M NaOAc
5. Agregar 2.5 volúmenes de (250 μ l) -20°C 100% EtOH
6. Mezclar por inversión o vortex
7. Incubar a -80°C por 30 min.
8. Centrifugar a máxima velocidad (~16000 x g) por 10 min.
9. Descartar el sobrenadante
10. Secar el pellet completamente y lavar con etanol 100 %
11. Resuspender en 10 μ l de agua.

b. Ligación de adaptador en 5’: En un tubo 1.2 ml, agregar lo siguiente para un volumen final de 20 μ l (miRCat™ microRNA cloning kit):

<i>Reactivo y concentración</i>	<i>Cantidad (μl)</i>
Ligación 3’ purificada	8 μ l
Linker 5’ (50 μ M)	1 μ l
Buffer de ligación 10X	2 μ l
Potenciador de ligación	6 μ l
ATP (10 mM)	2 μ l
RNA ligasa T4 (5 U/ μ l)	1 μ l

Incubar esta reacción a 22°C por 2 horas. Y luego,

1. Agregar 80 µl de IDTE (pH 7.5)
2. Transferir todo el volumen a un tubo nuevo 1.5 ml
3. Agregar 3 µl de glicógeno (10 mg/ml)
4. Agregar 1/10 del volumen (10 µl) de 3.0 M NaOAc
5. Agregar 2.5 volúmenes de (250 µl) -20°C 100% EtOH
6. Mezclar por inversión o vortex
7. Incubar a -80°C por 30 min.
8. Centrifugar a máxima velocidad (~16000 x g) por 10 min.
9. Descartar el sobrenadante
10. Secar el pellet completamente y lavar con etanol 100 %
11. Resuspender en 10 µl de agua.

c. Síntesis de cDNA o transcripción reversa. Agregar los siguientes componentes un tubo de 0.2 µl para una reacción total de 20.0 µl (miRCat™ microRNA cloning kit):

<i>Reactivo y concentración</i>	<i>Cantidad (µl)</i>
microRNAs ligados a adaptadores	11 µl
dNTPs (10 mM)	1 µl
RT primer (10 µM)	2 µl
Incubar a 65 °C por 5 minutos, poner en hielo y agregar:	
5X First Strand Buffer	4 µl
0.1 M DTT	1 µl
RNAsa-OUT™ (40 U/µl)	1 µl
SuperScript™ III RT (200 U/µl)	1 µl
Incubar a 50°C por una hora, seguido por 15 minutos a 70 °C. Almacenar a -20°C	

d. Amplificación PCR: En tubos 0.2 ml preparar la siguiente reacción de volumen final 50 µl (miRCat™ microRNA cloning kit, Advantage™ 2 Polymerase Mix kit):

<i>Reactivo y concentración</i>	<i>Cantidad (ul)</i>
cDNA (o Agua para control negativo o blanco)	3 µl
Agua pura	35 µl
10X PCR Buffer Advantage	5 µl
dNTPs (40 mM)	4 µl
Forward Primer (10 µM)	1 µl
Reverse Primer (10 µM)	1 µl
Taq polymerase Advantage	1 µl

Condiciones de PCR:

95°C por 10 min.
 35 ciclos { 95°C por 30 seg.
 52°C por 30 seg.
 72°C por 30 min.
 72°C por 5 min.
 4°C sostenido.

e. Reacción de clonación en vector pGEM®-T Easy: Para un volumen final de reacción de 10 µl, mezclar (pGEM®-T Easy Vector Systems kit):

<i>Reactivo y concentración</i>	<i>Cantidad (ul)</i>
buffer de DNA ligasa 2X	5 µl
Producto de PCR	X µl *
Vector (50 ng)	1 µl
DNA ligasa	1 µl
Mezclar suavemente e incubar la reacción por 1 hora a temperatura ambiente.	

* La cantidad (ul) de PCR a ligar dependerá de la proporción molar inserto: vector que se quiera manejar, la proporción utilizada es de 7.6:1. Por lo tanto, la cantidad de PCR utilizada es de 12.7 ng, según la siguiente fórmula: En donde la fracción 7.6/1 corresponde a la proporción 7.6:1 inserto: vector, si se desea cambiar esta proporción se puede hacer.

$$\text{ng inserto} = \frac{50 \text{ ng vector} * 0.1 \text{ ng inserto}}{3.0 \text{ kb vector}} * \frac{7.6}{1}$$

f. Limpieza de producto de ligación al vector: A los 10 μ l de ligación agregar 40 μ l de agua y 500 μ l de sec-butanol, dar vortex suavemente y centrifugar a 14000 r.p.m por 10 minutos. Descartar con una micropipeta el butanol (fase superior) y agregar 1 ml de etanol absoluto frío. Centrifugar a 14000 r.p.m por 25 minutos, posteriormente, descartar etanol, dejar secar y resuspender en 10 μ l de agua (Shao-Yao 2006).

g. Transformación de células bacterianas: A 2 μ l de reacción de clonación agregar 22 μ l de células bacterianas y mezclar suavemente con la pipeta. Cuidadosamente transferir la mezcla a una celda evitando la formación de burbujas. Electroporar las muestras a un voltaje mayor a 2.1 kV y a una capacitancia de 400 μ F. Inmediatamente transferir las células transformantes a un tubo con 1 ml de medio SOC e incubar en un shaker a 37°C por 1 hora. Por otro lado, se debe preparar LB sólido con carbenicilina (50 μ g/ml), IPTG (0.5 mM) y Xgal (80 μ g/ml) en varias cajas de petri de plástico. Platear varias de ellas, cuando las células hayan crecido por 1 hora, con diferentes volúmenes de transformantes (50 μ l, 100 μ l y 150 μ l) para que al menos una placa tenga colonias separadas. Incubar a 37°C por toda la noche. Al siguiente día se debe elegir las placas con colonias separadas, picar con un palillo las colonias de color blanco y sembrar cada una en un pozo de una placa 96, en cada uno de los cuales hay 200 μ l de LB líquido con carbenicilina. Incubar a 37 °C por toda la noche (Shao-Yao 2006).

h. PCR con clones: Preparar un coctel de PCR dependiendo de la cantidad de reacciones con los siguientes reactivos, para un volumen final de cada reacción de 50 μ l (Shao-Yao 2006).

<i>Reactivo y concentración</i>	<i>Cantidad (μl)</i>
Colonia de bacteria con inserto	5 μ l
Agua pura	32.2 μ l
10X PCR Buffer	5 μ l

Mg 25 mM	3 µl
dNTPs (20 mM)	1.3 µl
Primer T7	1.5 µl
Primer SP6	1.5 µl
Taq polymerase	1

Condiciones de PCR:

95°C por 10 min.
95°C por 30 seg.
57°C por 30 seg.
72°C por 1 min.
35 ciclos { 95°C por 30 seg.
52°C por 30 seg.
72°C por 1 min.
72°C por 10 min.
4°C sostenido.

Anexo 6: Cuantificación de la concentración (http://www.roswellpark.org/files/1_2_1/Genomics/MA%20protocols/PicoGreen%20dsDNA%20quantitation.pdf).

En primer lugar se debe preparar la curva estándar de DNA, en base a la cual se realizará la lectura del DNA de concentración desconocida. Esto consiste en preparar en TE 1X 4 concentraciones diferentes del DNA control (1 ng/ul, 0.5 ng/ul, 0.1 ng/ul y 0.02 ng/ul) y un blanco. Nuestra muestra problema debe prepararse con un factor de dilución de 50. Todo esto debe prepararse en una placa, para agregar 100 ul de picogreen y ser leído por el fluorometro.

Anexo 7: Protocolo para secuenciación (ABI PRISM® Big Dye terminator V1.1. cycle sequencing Ready Reaction kit).

a. Limpieza de producto de PCR de clones:

1. A un PCR de 50 ul adicionar 1 vol de PEG 20% NaCl 2.5 M (10 g de Polietilen glycol (8.000) + 7.3 g de NaCl) y tapar con cinta la placa.
2. Mezclar por vortex, dar spin para bajar la muestra de la tapa e incubar a temperatura ambiente por 15 minutos.
3. Centrifugar a 2800 RPM por 30 minutos a 4°C y descartar el sobrenadante por inversión.
4. Adicionar 100 ul de Etanol 70%, centrifugar a 2800 RPM por 10 minutos a 4°C y descartar el sobrenadante por inversión.
5. Invertir placas de 96 sobre una toalla de papel y dar un spin (a 500 RPM) para quitar el exceso de etanol.
6. Resuspender en 10 ul de Agua y revisar 2 ul en gel de azarosa 4%.

b. Reacción de secuencia: En una nueva placa de 96 preparar:

<i>Reactivo y concentración</i>	<i>Cantidad (ul)</i>
PCR	3 µl
Terminator Ready reaction Mix BigDye	2 µl
Primer SP6	2µl
Buffer 5X	1 µl

Condiciones de PCR:

30 ciclos {
95°C por 5 min.
95°C por 30 seg.
50°C por 10 seg.
60°C por 4 min.
4°C sostenido.

c. Limpieza de secuencia:

1. Agregar 25 ul de Acetato de Na 120 mM + Etanol 96% (1 ul de Acetato de Na 3M y 24 ul de etanol 100%) a la reacción de secuencia de 10 ul.
2. Incubar por 20 min. A -20°C, centrifugar a 14000 RPM por 30 min. A 4°C y descartar el sobrenadante.
3. Agregar 250 ul de Etanol 70%, centrifugar a 14000 RPM por 15 min. A 4°C y descartar el sobrenadante
4. Dejar secar, resuspender en 3 ul de buffer de carga y servir en el secuenciador ABI en un gel de acrilamida 4%.