



# UNIVERSIDAD NACIONAL DE ROSARIO

## FACULTAD DE CIENCIAS AGRARIAS

TESIS PARA OPTAR AL GRADO ACADÉMICO DE  
DOCTOR EN CIENCIAS AGRARIAS

*Caracterización de fragmentos génicos desconocidos que fueron  
asociados con la apomixis y/o la poliploidía en Paspalum notatum*

**Presentada por:**

**ANA CLAUDIA OCHOGAVÍA**

**Licenciada en Biotecnología**

**Directora:**

**Dra. SILVINA CLAUDIA PESSINO**

ZAVALLA- ARGENTINA

**2011**

## **AGRADECIMIENTOS**

Quisiera expresar mi agradecimiento a la Dra. Silvina Pessino por su predisposición, su apoyo y principalmente por confiar en mí para realizar este trabajo.

Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por haber posibilitado mi estudio de posgrado.

A la Facultad de Ciencias Agrarias de la Universidad Nacional de Rosario (FCA-UNR), lugar donde pude desarrollar este trabajo de investigación.

Al Dr. Juan Pablo Ortiz y al grupo de investigación del laboratorio de Biología Molecular de Ciencias Agrarias, por su contribución desinteresada. Y al Dr. Camilo Quarín del Instituto de Botánica del Nordeste (IBONE-CONICET) quien inició esta línea de trabajo.

A la Dra. Silvina Felitti, por su generosa colaboración y sus consejos. A mis compañeros de laboratorio, Micaela, Maricel, Lorena y Lucas por cada uno de sus aportes y por acompañarme día a día durante estos años.

A todas las personas con las que comparto mi vida: familiares, amigos, compañeros y conocidos.

Finalmente quiero agradecer a mis padres y mi abuelo por la incondicionalidad y el respaldo. A mi hermana por toda la bondad que me brindó y que siempre me acompañará. Y a mi compañero de vida, Martín, por todo el amor, la paciencia y por cada uno de los días juntos, pero especialmente por la hermosa familia que tenemos con esta personita maravillosa que viene en camino para llenar nuestras vidas de alegría.

Parte de los resultados informados en este trabajo de tesis fueron volcados en las siguientes publicaciones y reuniones científicas:

#### ***Publicaciones científicas***

1. **Ochogavía AC**, Seijo G, González AM, Cervigni G, Podio M, Duarte Silveira E, Machado Lacerda AL, Campos Carneiro V, Ortiz JP, Pessino SC. (2011) **Characterization of retrotransposon sequences expressed in inflorescences of apomictic and sexual *Paspalum notatum* plants.** Sexual Plant Reproduction 24 (3): 231- 246
2. **Ochogavía AC**, Seijo G, González AM, Pessino SC. **Expression of precursor of miRNAs sequences in apomictic and sexual *Paspalum notatum*** (Manuscrito en preparación)

#### ***Otras publicaciones relacionadas***

1. **Ochogavía AC**, Cervigni G, Selva JP, Echenique V, Pessino SC (2009) **Variation in cytosine methylation patterns during ploidy level conversions in *Eragrostis curvula*.** Plant Molecular Biology 70 (1-2):17-29
2. Mecchia MA, **Ochogavía AC**, Selva JP, Laspina N, Felitti S, Martelotto LG, Spangenberg G, Echenique V and Pessino SC (2007) **Genome polymorphisms and gene differential expression in a ‘back-and-forth’ ploidy-altered series of weeping lovegrass (*Eragrostis curvula*)** Journal of Plant Physiology 164(8):1051-61
3. Zappacosta D, **Ochogavía AC**, Meier M, Echenique V, Pessino S. **Epigenetic variation underlie changes in apomixis expressivity in *Eragrostis curvula*** (Manuscrito en preparación)

#### ***Presentaciones en congresos***

1. **Ochogavía AC**, Seijo G, González AM, Ortiz JP, Pessino SC. **Caracterización de transcritos homólogos a precursores de micro ARN en tejidos reproductivos de plantas sexuales y apomícticas de *Paspalum notatum*.** XL Congreso Argentino de Genética - III Simposio Latinoamericano de Citogenética y Evolución- I Jornadas Regionales SAG-NEA (18-21/09/2011)
2. **Ochogavía AC**, Seijo G, González AM, Cervigni G, Podio M, Duarte Silveira E, Machado Lacerda AL, Campos Carneiro V, Ortiz JP, Pessino SC. **Characterization of retrotransposon sequences expressed in inflorescences of apomictic and sexual**

***Paspalum notatum* plants.** XXVIII Reunión Argentina de Fisiología Vegetal (RAFV 2010) La Plata, Bs. As. (26-29/09/2010)

3. **Ochogavía A C**, Seijo G, González A M, Cevigni G, Podio M, Laspina N, Duarte Silveira, Machado Lacerda A L, Carneiro V T C, Pessino S C. **Caracterización de transcriptos no codificantes asociados a la expresión de la apomixis en *Paspalum notatum*.** II Congreso Internacional-REDBIO- Argentina, VII Simposio Nacional de Biotecnología REDBIO-Argentina (20-24/04/2009). **Trabajo premiado.**
4. **Ochogavía AC**; Seijo G; González AM; Laspina N; Tavares de Campos Carneiro V and Pessino SC. **Functional annotation and expression analysis of novel sequences associated to aposporous development.** XX International Congress on Sexual Plant Reproduction, Brasilia, Brasil. (04-08/08/2008)

#### ***Otras presentaciones relacionadas en congresos***

1. Rodrigo JM, **Ochogavía AC**, Meier M, Zappacosta D, Pessino SC, Echenique V. **Effects of stress and genomic plasticity on the reproductive mode of facultative apomictic plants of weeping lovegrass [*Eragrostis curvula* (Schrad.) Nees].** The 6th International Symposium on the Molecular Breeding of Forage and Turf - Buenos Aires-Argentina (15-19/03/2010)
2. Selva JP, Cervigni G, **Ochogavía AC**, Zappacosta D, Meier M, Pessino SC, Echenique V. **Global analysis of the genome, transcriptome and epigenome in diplosporous grass *Eragrostis curvula* (Schrad.) Nees.** XX International Congress on Sexual Plant Reproduction, Brasilia, Brasil. (04-08/08/2008).
3. Cervigni G, Selva JP, Mecchia M, **Ochogavía A**, Paniego, N, Cardone S, Polci P, Felitti S, Martelotto G, Zappacosta D, Meier M, Voigt P, Spangenberg G, Pessino S and Echenique V. **Genetic, epigenetic and transcriptional changes associated to ploidy and diplospory in *Eragrostis curvula* (Schrad.) Nees.** Plant and Animal XV Genome Conference, San Diego, E.E.U.U. (15-20/01/2008)

#### ***Premios***

1. **Mención especial de AAPRESID al trabajo “Caraterización de transcriptos no codificantes asociados a la expresión de la apomixis en *Paspalum notatum*”.** II Congreso Internacional- REDBIO- Argentina - VII Simposio Argentino de Biotecnología, REDBIO Argentina - Rosario. Fecha: 24 de Abril de 2009.

**INDICE**

Índice	i
Abreviaturas y anglicismos	iv
Resumen	vii
<b>1- INTRODUCCIÓN</b>	<b>1</b>
1.1- Apomixis: conceptos generales	1
1.1.1- Mecanismos citoembriológicos de la apomixis	2
1.1.2- Caracterización genética de la apomixis	6
1.1.3- Regulación de la expresión génica en la apomixis	7
1.1.4- Importancia agronómica de la apomixis	13
1.2- Poliploidía	14
1.2.1- Poliploidía: aspectos generales	14
1.2.2- Modificaciones genómicas que acompañan la poliploidización	16
1.2.3- Poliploidía y expresión génica	17
1.3- Relación entre la apomixis y la poliploidía	18
1.4- El género <i>Paspalum</i>	20
1.4.1- <i>Paspalum notatum</i> (Pasto horqueta, “Bahiagrass”)	20
1.4.2- La apomixis en <i>Paspalum notatum</i>	22
1.4.2.1- Caracterización genética de la apomixis en <i>P. notatum</i>	22
1.4.2.2- Estudios de la expresión génica en genotipos apomícticos y sexuales de <i>Paspalum notatum</i>	23
<b>2- OBJETIVOS</b>	<b>25</b>
2.1- Objetivos generales	25
2.2- Objetivos específicos	25
<b>3- MATERIALES Y MÉTODOS</b>	<b>27</b>
3.1- Material vegetal	27

---

3.2- Extracción de ADN	27
3.3- Extracción de ARN total	29
3.4- Amplificación de ADN genómico	29
3.5- Construcción de bibliotecas no clonadas de ADN	30
3.6- Amplificación de extremos 5' y 3' de ADNc	30
3.7- Clonado y transformación	33
3.8- Análisis bioinformático de secuencias	34
3.9- Transcripción reversa	36
3.10- PCR en tiempo real	37
3.11- Generación de marcadores moleculares de tipo SCAR (Sequence characterized amplified regions)	39
3.12- Mapeo genético	40
3.13- Hibridización <i>in situ</i> de tejidos reproductivos	41
3.13.1- Marcación de sondas con Digoxigenina	41
3.13.2- Hibridización de tejidos reproductivos	42
3.14- Display Diferencial Específico de Retrotransposones	44
3.15- Detección de secuencias blanco por 5' RACE	45
<b>4- RESULTADOS</b>	47
4.1- Búsquedas bioinformáticas iniciales	47
4.2- Amplificaciones genómicas de control	49
4.3- Extensión de los transcritos por experimentos 5' y 3' RACE	50
4.4- Análisis bioinformáticos de los fragmentos extendidos	53
4.5- Análisis casuístico de los transcritos extendidos	61
4.5.1- Estudio de transcritos homólogos a retrotransposones (N17 y N22)	61
4.5.1.1- Determinación del número de copias genómicas	61
4.5.1.2- Experimentos de mapeo de N17	63
4.5.1.3- Estudios de expresión por PCR en tiempo real	64
4.5.1.4- Estudios de expresión por Hibridización <i>in situ</i> de tejidos	66
4.5.1.5- Display Diferencial Específico de Retrotransposones	69
4.5.1.6- Caracterización de la familia de retrotransposones N22 mediante análisis filogenético	73

---

4.5.2- Estudio de transcritos similares a pre-miARN (N11 y N13)	74
4.5.2.1- Estudios de expresión por PCR en tiempo real	74
4.5.2.2- Estudios de expresión por Hibridización <i>in situ</i> de tejidos	78
4.5.2.3- Determinación del número de copias genómicas	81
4.5.2.4- Análisis 5' RACE de los genes blanco predichos	82
4.6- Análisis bioinformático del resto de los transcritos diferenciales	85
4.6.1- Estudio de los transcritos diferenciales por modo de reproducción	85
4.6.2- Estudio bioinformático de los transcritos diferenciales por nivel de ploidía	89
<b>6- DISCUSIÓN</b>	<b>97</b>
6.1- Análisis de los transcritos homólogos a retrotransposones	99
6.2- Análisis de transcritos homólogos a precursores de miARNs	106
6.3- Análisis <i>in silico</i> del resto de los transcritos bajo estudio	110
<b>7- CONCLUSIONES</b>	<b>115</b>
<b>8- REFERENCIAS</b>	<b>117</b>
ANEXOS	134
ANEXO I	134
ANEXO I- Sección A	134
ANEXO I- Sección B	145
ANEXO II	157
ANEXO II- Sección A	157
ANEXO II- Sección B	162

## **ABREVIATURAS Y ANGLICISMOS**

**A:** base nitrogenada púrica adenina

**AA:** resto aminoacídico

**ADN:** ácido desoxiribonucleico

**ADNc:** ácido desoxiribonucleico copia

**AFLP:** marcadores moleculares generados a partir de polimorfismos de fragmentos amplificados (del inglés, Amplified Fragment Length Polymorphism)

**Amplicón:** secuencia de ADN producto de la reacción en cadena de la polimerasa

**ARN:** ácido ribonucleico

**ARNasa:** enzima degradativa de ARN

**ARNi:** ARN de interferencia

**ARNm:** ácido ribonucleico de tipo mensajero

**Blast:** herramienta de búsqueda bioinformática por alineamientos locales con las bases de datos.

**BrEt:** bromuro de etidio

**BSA:** albúmina sérica bovina 1-acetilada

**C:** base nitrogenada pirimídica citocina

**cDNA-AFLP:** polimorfismos en el largo de los fragmentos amplificados a partir de ADN copia (del inglés, cDNA Amplified Fragment Length Polymorphism)

**cM:** centimorgan, unidad de medida de distancia genética

**Ct:** parámetro evaluado en los experimentos de PCR en tiempo real, indica el número de ciclos necesarios para que la señal de fluorescencia supere una línea de base umbral preestablecida

**CTAB:** bromuro de cetil - dimetilamonio

**DD:** Display Diferencial

**DDT:** oligonucleótido anclado (del inglés Differential Display PolyT)

**DEPC:** dietilpirocarbonato

**Display Diferencial:** técnica de biología molecular utilizada para identificar cambios en la representación de transcritos de ARN poliadenilados (del inglés, Differential Display)

**dNTP:** desoxinucleótidos trifosfato

**DTT:** ditioneitol

**EDTA:** ácido etilendiamino-tetracético

**EST:** etiquetas de secuencia expresadas. Secuencias de entre 150 y 300 nt que provienen de transcritos poliadenilados expresados (del inglés, Expressed Sequence Tags)

**e-value (eV):** valor esperado

**G:** base nitrogenada púrica guanina

**H<sub>2</sub>O<sub>d</sub>:** agua destilada

**Kinasa “check point”:** enzima de actividad kinasa involucrada en procesos clave del ciclo celular

**Kinasa:** enzima fosforilativa dependiente de ATP.

**Kit:** equipo comercial

**Knob DNA:** (en inglés) estructura compleja de ADN compuesta principalmente por un nudo de heterocromatina

**LTR:** repeticiones terminales largas encontradas en algunas familias de elementos repetitivos (del inglés, Long Terminal Repeats)

**Medio LB:** medio de cultivo bacteriano Luria-Bertani

**miARN:** micro ARN

**n:** número cromosómico

**NcoI:** enzima de restricción de corte frecuente

**nt:** nucleótidos

**p/v:** peso en gramos por cada 100 ml de volumen de solución

**pb:** pares de base

**PCR Real Time:** (en inglés) reacción de PCR analizada en tiempo real.

**PCR:** reacción en cadena de la polimerasa (del inglés, Polymerase Chain Reaction)

**Pre-miARN:** precursor de micro ARN maduro

**Pri-miARN:** precursor primario de micro ARN

**RACE:** amplificación rápida de extremos de ADNc, técnica de biología molecular (del inglés, Rapid Amplification of cDNA Ends)

**RAPD:** marcadores moleculares generados a partir de la amplificación de ADN al azar (del inglés, Random Amplification of Polymorphic DNA)

**RT-PCR:** reacción en cadena de la polimerasa de la transcripción reversa (del inglés, Reverse Transcription Polymerase Chain Reaction)

**SAGE:** técnica de biología molecular (del inglés: Serial Analysis of Gene Expression)

**Sall:** enzima de restricción de corte frecuente

**SCAR:** marcadores moleculares de regiones de secuencias caracterizadas (del inglés, Sequence Characterized Amplified Regions)

**SDS:** dodecil sulfato de sodio

**Buffer:** solución reguladora, solución tampón

**Stem-loop:** (en inglés) estructura secundaria de ARN tipo horquilla

**Super-SAGE:** técnica de biología molecular derivada de SAGE (del inglés: Super Serial Analysis of Gene Expression)

**T:** base nitrogenada pirimídica timina

**TAE:** Tris clorhídrico-ácido acético-EDTA (solución reguladora)

**Take off:** parámetro evaluado en los experimentos de PCR en tiempo real, indica el número de ciclos transcurridos cuando la derivada de la señal de fluorescencia deja de ser igual a cero.

**TBE:** Tris clorhídrico- ácido bórico – EDTA (solución reguladora)

**Tm:** temperatura de hibridización de ácidos nucleicos

**Tris:** N-Tris-hidroximetil-aminoetano

**v/v:** volumen en mililitros por cada 100 ml de volumen de solución

**x:** número cromosómico básico de una especie

**X-Gal:** 5-bromo-4cloro-3indolil-beta-D-galactopiranosido

**RESUMEN**

La apomixis es una forma de reproducción asexual vía semillas frecuentemente asociada a la poliploidía. El objetivo de este trabajo fue inferir la función de 58 transcritos desconocidos, que participan de los procesos de apomixis y/o poliploidización en *Paspalum notatum*. Utilizando amplificaciones rápidas de los extremos del cDNA se obtuvieron las secuencias completas de varios candidatos, que resultaron similares a elementos repetitivos portadores de segmentos génicos transduplicados o a precursores de miRNA de plantas. Para ambos tipos de secuencias se propuso un posible rol regulatorio, y se determinaron *in silico* los blancos putativos. Se eligieron secuencias representativas de cada uno de estos grupos, y se realizaron estudios experimentales para lograr una anotación funcional inequívoca. Para ello se analizó el número de copias genómicas, la posición en el genoma, la expresión cuantitativa durante el desarrollo reproductivo y la localización *in situ* de la actividad génica. También se realizaron análisis de display diferencial específicos para aislar retrotransposones con segmentos transduplicados relacionados y estudios de clivado específico de los blancos regulatorios de los miRNAs. Los resultados presentados en esta tesis aportan las primeras evidencias acerca del rol regulatorio que los retrotransposones y los miRNAs podrían estar desempeñando durante el desarrollo apomítico.

**ABSTRACT**

Apomixis is an asexual mode of reproduction via seeds commonly associated with polyploidy. The objective of this research work was to ascribe a functional identity to 58 unknown transcripts related to apomixis and/or polyploidization in *Paspalum notatum*. We used rapid amplification of cDNA ends to obtain full sequences for several candidates, which resulted homologous to retrotransposons carrying transduplicated gene segments or plant miRNA precursors. For both sequence types a regulatory role was proposed, and the identities of the putative targets were determined *in silico*. Representative sequences were selected to carry out experiments in order to achieve positive functional annotation. We analyzed gene copy numbers, positions in the genome, quantitative expressions at different developmental stages, and *in situ* localization of gene activity. Specific differential display analysis were conducted to isolate related retroelements containing gene transduplicated segments. Specific clivage analysis of the miRNA putative targets was also performed. Results presented here provide evidence on the regulatory role that both retrotransposons and miRNAs could be performing in apomixis development.

## 1- INTRODUCCIÓN

### 1.1- Apomixis: conceptos generales

El término apomixis deriva del griego, siendo *apo* un prefijo de significación negativa que podría traducirse como “falta o ausencia”; y *mixis* un vocablo que significa “unión o mezcla”. El término hace referencia a la ausencia de recombinación y de fecundación que caracteriza a este tipo de reproducción (Nogler, 1984).

La apomixis en las plantas angiospermas es comúnmente definida como “reproducción asexual vía semillas”. Existen dos tipos de reproducción asexual observadas en la clase *Angiospermae*: una es la propagación vegetativa, donde las plantas se reproducen vía estolones, rizomas, bulbos o enraizamiento de esquejes; y la otra es la apomixis, donde se producen semillas viables sin reducción meiótica ni fusión de gametas, que contienen embriones genéticamente idénticos a la planta madre (Cronquist, 1988).

La apomixis fue descrita por primera vez en 1841 en la especie dioica *Alchornea ilicifolia* por Smith (1841). Un único ejemplar femenino de esta especie había sido aislado al ser trasladado a los Kew Gardens de Londres desde Asia. Sin embargo, luego de florecer produjo semillas en abundancia en ausencia de un progenitor masculino, poniendo al carácter en evidencia (Smith, 1841). Unos pocos años más tarde, Gregor Mendel se convirtió en el autor de los primeros experimentos con plantas apomícticas, aunque éstos fueron realizados en forma involuntaria. Mendel utilizó cruza entre especies del género *Hieracium* para intentar confirmar los resultados obtenidos en sus estudios sobre la herencia de las arvejas de jardín. El investigador atribuyó erróneamente a una supuesta "autopolinización frecuente" la falta de segregación observada. Aunque nunca publicó sus resultados, los comentó en sus cartas a Carl Naegeli (Correns, 1905). Actualmente se conoce que muchas especies del género *Hieracium* son apomícticas.

La apomixis se encuentra ampliamente distribuida entre las angiospermas. Ocurre en más de 400 especies pertenecientes a 40 familias diferentes, siendo las más representadas las familias *Gramineae*, *Compositae*, *Rosaceae* y *Rutaceae* (Hanna and Bashaw, 1987; Ozias-Akins, 2006; Ozias-Akins and van Dijk, 2007). También se ha observado frecuentemente en las orquídeas (Naumova, 1993). Curiosamente, este carácter ocurre en algunas familias y especies aisladas dentro de cada taxón, lo que denota un claro origen polifilético (Ozias-Akins, 2006; Ozias-Akins and van Dijk, 2007). Además, una misma planta puede presentar ambos modos de reproducción, sexual y apomíctico, en

perfecto balance y con niveles variables de expresión, por lo que la capacidad de producir semillas clonales no excluye necesariamente la sexualidad. En ese caso se habla de *apomixis facultativa*. Las plantas que dan lugar a progenies idénticas a la planta madre y producen todos sus embriones por apomixis son denominadas *apomícticas obligadas* (Koltunow, 1993; Chaudhury *et al.*, 2001), aunque muchos autores coinciden en que siempre existe una tasa de sexualidad vestigial (revisado en Nogler 1984).

La apomixis es un carácter controlado genéticamente, comúnmente asociado a la ocurrencia de poliploidía. Sus bases moleculares no han sido aún completamente esclarecidas. Las especies apomícticas suelen formar complejos agámicos, en los cuales los individuos diploides son sexuales y los poliploides apomícticos (Nogler, 1984). Sin embargo, existen algunos casos poco frecuentes en los cuales se ha informado la existencia de apomixis a nivel diploide, por ejemplo *Boechera holboellii* (Schranz *et al.*, 2006) y *Paspalum rufum* (Siena *et al.*, 2008).

### 1.1.1- Mecanismos citoembriológicos de la apomixis

Resulta necesario comparar la reproducción la sexual con la apomíctica debido a que se considera que esta última ha evolucionado como un sistema de reproducción alternativo a la sexualidad a través de la reformulación genética y/o epigenética de sus programas de desarrollo (Savidan, 2000; Grimanelli *et al.*, 2001). La reproducción sexual de las plantas comprende la alternancia cíclica entre los estados de esporofito (la planta misma,  $2n$ ) y de gametofito (el grano de polen y el saco embrionario,  $n$ ). La meiosis posibilita la recombinación y reducción del número de cromosomas y da lugar a la formación de las esporas femeninas (megásporas) en el ovario y masculinas (micrósporas) en las anteras. Por medio de la megasporogénesis se generan cuatro células haploides. En el proceso sexual más común (formación de un saco de tipo *Polygonum* monospórico), tres de ellas degeneran y una se transforma en la megáspora funcional. Esta última sufre tres mitosis sucesivas que dan lugar a la formación de un saco embrionario octanucleado (gametofito femenino), un proceso conocido como megagametogénesis. Así, este saco de tipo *Polygonum* queda conformado por ocho núcleos haploides y siete células (la ovocélula, dos sinérgidas, una célula central binucleada, y tres antípodas). Por otra parte, las micrósporas desarrollan los granos de polen, compuestos por tres células haploides ( $n$ ), mediante un proceso de microgametogénesis. Dos de ellas constituyen los gametos

masculinos y contienen los núcleos generativos y la tercera es la célula vegetativa, responsable del crecimiento y avance del tubo polínico (Nogler, 1984).

La formación de la semilla en plantas sexuales involucra una doble fecundación, donde un primer gameto masculino ( $n$ ) se fusiona con la ovocélula ( $n$ ) para formar el cigoto ( $n + n$ ) y un segundo gameto masculino ( $n$ ) se fusiona con la célula central del saco embrionario ( $n + n$ ) para originar el endosperma ( $n + n + n$ ) (Crow and Kimura, 1965; Nogler, 1984).

Las plantas apomícticas son capaces de eludir o modificar algunos eventos de la secuencia del desarrollo reproductivo sexual, evitando la reducción meiótica y la fecundación, generando embriones perfectamente funcionales que se desarrollan en el óvulo a partir de una célula que contiene el número somático de cromosomas (Koltunow, 1993). Los embriones guardan una identidad genética completa con el organismo que le dio origen. Por ello, la apomixis también es denominada *agamospermia*.

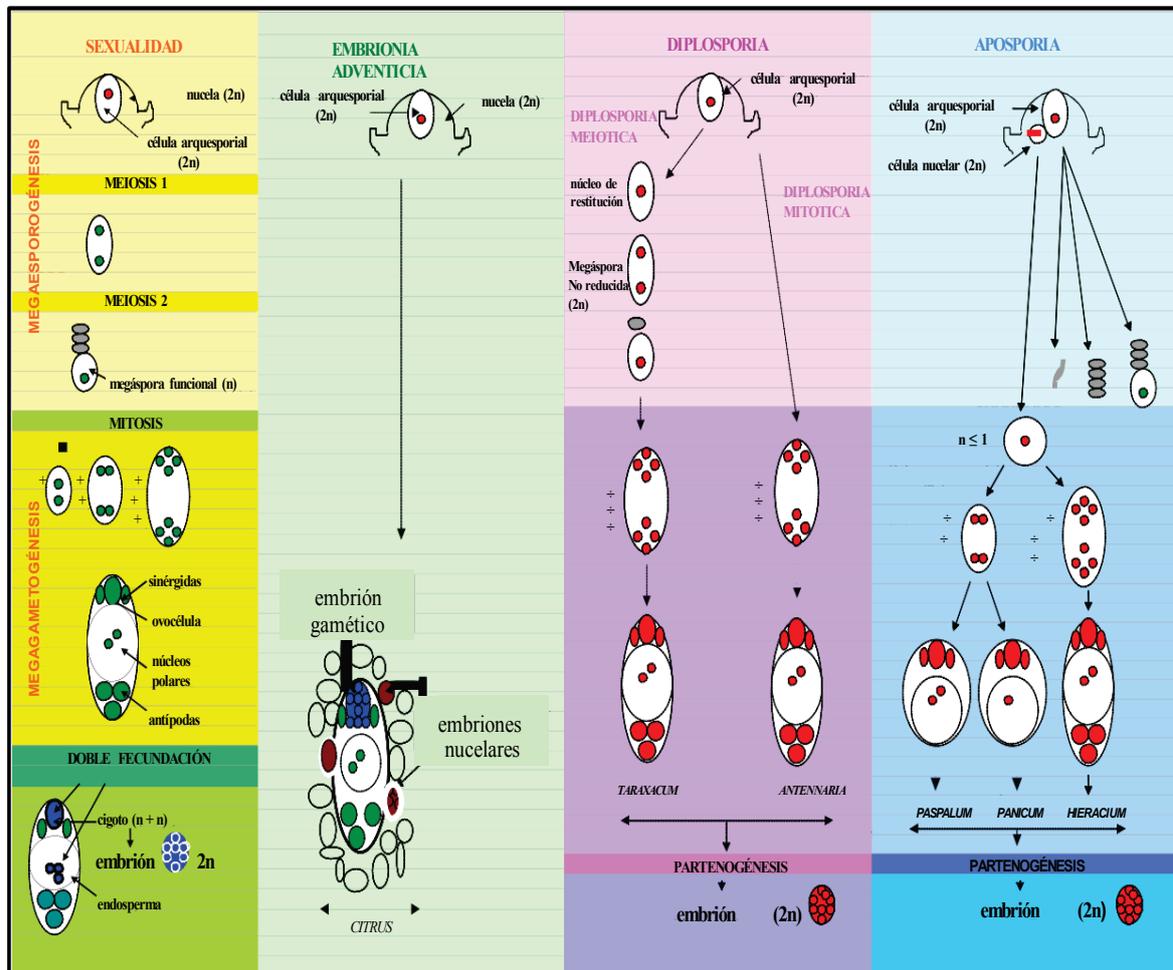
La apomixis se presenta en la naturaleza de dos formas generales: la *embriogénesis adventicia* y la *apomixis gametofítica*. En el primer caso, surgen embriones directamente por partenogénesis a partir de células de la nucela o de los integumentos del óvulo, por un proceso partenogenético denominado *embrionía adventicia* (Lakshmanan and Ambegaokar, 1984). De este modo, se forman embriones múltiples a partir de células nucelares (esporofíticas) que comparten el óvulo con el embrión de origen sexual y utilizan su endosperma para desarrollarse. Esta forma de apomixis aparece comúnmente en los cítricos. En la *apomixis gametofítica* se forman sacos embrionarios que contienen ovocélulas no reducidas, ya que la meiosis de la célula madre de la megáspora (CMM) es reemplazada por un proceso mitótico (Koltunow, 1993). De acuerdo al tipo celular involucrado en los procesos mitóticos durante la generación del megagametofito, la apomixis gametofítica puede clasificarse en: a) *diplosporia*: cuando el saco embrionario se origina a partir de la célula madre de la megáspora, directamente por mitosis o luego de una falla en la meiosis; o b) *aposporía*: cuando el saco embrionario se genera a partir de células somáticas del óvulo, habitualmente de la nucela. Tanto el proceso diplospórico como el apospórico derivan en la formación de un megagametofito que contiene todos sus núcleos no reducidos ( $2n$ ), inclusive el de la célula huevo, sin embargo son marcadas las diferencias morfológicas entre ambos tipos de sacos. Los sacos embrionarios diplospóricos suelen conservar la estructura típica de los sacos de origen meiótico de tipo *Polygonum* (7 células y 8 núcleos) (Savidan, 2000) con algunas excepciones como la de

*Eragrostis curvula* (Meier *et al.*, 2011). En las plantas diplospóricas el proceso sexual se ve comprometido ya que es la misma CMM la que sufre alteraciones en su desarrollo para generar los megagametofitos no reducidos, por lo que ambos procesos son excluyentes. Sin embargo existen individuos diplospóricos facultativos, con algunos óvulos que forman sacos embrionarios meióticos y otros sacos embrionarios no reducidos (Koltunow, 1993). Por otra parte, en los individuos apospóricos, los sacos no reducidos pueden coexistir en el mismo óvulo con los generados por sexualidad, ya que la CMM no se encuentra comprometida en la generación de los sacos no reducidos. Los sacos apospóricos presentan una constitución morfológica muy variable entre taxones o incluso dentro de un mismo taxón, entre las que se encuentran la morfología tipo *Hieracium* (Rosenberg, 1908; Rosenberg, 1930), tipo *Panicum* (Warmke, 1954), y las variantes del género *Paspalum* (Quarín and Burson, 1991). Se detalla en la Figura I un esquema comparativo de las vías de reproducción sexual y apomícticas en angiospermas.

A diferencia de las plantas sexuales, en las apomícticas (esporofíticas y gametofíticas), los embriones son generados por partenogénesis, aunque existe una tasa variable de formación de híbridos BIII (individuos generados por fecundación de la ovocélula no reducida). Generalmente, los núcleos polares no reducidos son fecundados para producir el endosperma (pseudogamia), aunque existen casos en los que éste se desarrolla en forma autónoma (Nogler, 1984). La mayoría de las especies apomícticas producen polen viable, aunque la tasa de funcionalidad del mismo puede verse disminuida. La presencia de polen viable permite la fecundación eventual de la ovocélula no reducida, conllevando a un aumento de la ploidía. Sin embargo, en varias especies apomícticas de los géneros *Taraxacum* (Cooper and Dick, 1949), *Poa*, *Pennisetum*, *Tripsacum* (Asker and Jerling, 1992) y *Hieracium* (Koltunow *et al.*, 1998), la formación del embrión es precoz, iniciando su desarrollo en el estadio previo a la antesis, limitando, de esta manera, la posibilidad de fecundación de la ovocélula no reducida.

Como se mencionó anteriormente, apomixis y sexualidad no son procesos mutuamente excluyentes, ya que una misma planta pueden aparecer simultáneamente sacos reducidos (meióticos) y no reducidos (provenientes de apomixis), incluso en una misma inflorescencia y aún en un mismo óvulo. Por lo tanto, en los genotipos apomícticos facultativos las progenes segregan como clases maternas ( $2n + 0$ ) y no-maternas o aberrantes. Existen tres tipos diferentes de individuos aberrantes que pueden encontrarse en la progenie de una planta apomíctica: 1) híbridos BIII ( $2n + n$ ) que resultan de la

fecundación de una ovocélula no reducida, 2) híbridos BII ( $n + n$ ) que resultan de la fecundación de una ovocélula reducida y 3) haploides ( $n + 0$ ) generados por partenogénesis a partir de una ovocélula reducida (Nogler, 1984).



**Figura I:** Comparación de mecanismos de reproducción de angiospermas. Se detalla un diagrama del mecanismo de desarrollo sexual, de apomixis esporofítica y de los dos tipos de apomixis gametofítica (diplosporía y aposporía). Durante la sexualidad, una célula de la nucela del óvulo se diferencia a CMM y genera cuatro megasporas haploides por meiosis. Una de ellas da origen a la formación de un saco embrionario(n). La ovocélula y los núcleos polares del saco son fecundados por los núcleos generativos del polen para generar el cigoto y el núcleo primario del endosperma, respectivamente. El cigoto da origen al embrión a través de sucesivas mitosis. En la apomixis esporofítica el embrión surge directamente de la célula somática del óvulo (usualmente de la nucela) y el embrión clonal utiliza el endosperma que se genera en un saco embrionario reducido que coexiste en el mismo óvulo. En la apomixis gametofítica se forma un megagametofito con células no reducidas, cuya ovocélula ( $2n$ ) genera un embrión por partenogénesis. Apomixis diplospórica: el saco embrionario no reducido es generado a partir de de la CMM, directamente por mitosis o luego de una falla en la meiosis. Apomixis apospórica: los sacos embrionarios no reducidos son generados a partir de células somáticas del óvulo, generalmente de la nucela. Ilustración tomada de Stein (2006).

### 1.1.2- Caracterización genética de la apomixis

La apomixis es un carácter controlado genéticamente, aunque sus bases moleculares no han sido esclarecidas completamente. Los estudios de la herencia del carácter se han llevado a cabo por medio de análisis de segregación en cruzas entre un progenitor femenino sexual y uno masculino apomítico. Al ser las plantas apomíticas altamente heterocigotas, se puede observar la segregación del modo de reproducción en la generación F<sub>1</sub> (Ozias-Akins and van Dijk, 2007). El modo de clasificación de las progenies en apomíticas y sexuales fue motivo de controversia durante varias décadas entre los investigadores dedicados al estudio del fenómeno. En los últimos años el criterio predominante es el modelo cualitativo de clasificación: se considera que una planta es apomítica cuando se observa en ella al menos un saco embrionario no reducido, independientemente del grado de expresividad del carácter (Pessino and Ortiz, 2010).

La tecnología de marcadores moleculares y los procedimientos de biología molecular han permitido la generación de conocimiento novedoso en relación a las bases moleculares de la apomixis (Pessino *et al.*, 2001). Se detectaron marcadores moleculares ligados a la aposporía en *Pennisetum* (Ozias-Akins *et al.*, 1998), *Cenchrus* (Roche *et al.*, 1999), *Brachiaria* (Pessino *et al.*, 1997; Pessino *et al.*, 1999), *Paspalum* (Pupilli *et al.*, 1997; Martínez *et al.*, 2001; Siena *et al.*, 2008), y *Poa pratensis* (Albertini *et al.*, 2001a; Albertini *et al.*, 2001b) y a la diplosporía en híbridos de maíz-*Tripsacum* (Leblanc *et al.*, 1995; Grimanelli *et al.*, 1998b), *Taraxacum* (Vijverberg *et al.*, 2004) y *Erigeron annuus* (Noyes and Rieseberg, 2000; Noyes, 2005; Noyes, 2006).

Los resultados de los estudios de control genético de la apomixis indicaron: 1) la aposporía y la diplosporía parecen estar controladas por *loci* diferentes (Leblanc *et al.*, 1995; Pessino *et al.*, 1997; Pessino *et al.*, 1999); 2) la apomixis apospórica estaría controlada por un único sector genómico en todas las especies estudiadas (Pessino *et al.*, 1997; Pupilli *et al.*, 1997; Ozias-Akins *et al.*, 1998; Pessino *et al.*, 1999; Roche *et al.*, 1999; Martínez *et al.*, 2001; Siena *et al.*, 2008), a excepción de *Poa pratensis* (Albertini *et al.*, 2001a; Albertini *et al.*, 2001b) donde la apomeiosis y la partenogénesis segregan separadamente; 3) la apomixis diplospórica estaría controlada por dos o más genes. En *Tripsacum* y *Erigerum annuus* se identificaron dos *loci* involucrados (para apomeiosis y partenogénesis, respectivamente) (Leblanc *et al.*, 1995; Grimanelli *et al.*, 1998b; Grimanelli *et al.*, 2001); (Noyes and Rieseberg, 2000; Noyes, 2005; Noyes, 2006). En *Taraxacum officinale* se hallaron tres *loci* asociados al carácter (para apomeiosis,

partenogénesis y desarrollo autónomo del endosperma, respectivamente) (Van Dijk and Bakx-Schotman, 2004; Vijverberg *et al.*, 2004). Una fuerte supresión de la recombinación se observó en las regiones que controlan la apomeiosis en *Tripsacum* y *Erigeron* (Grimanelli *et al.*, 1998b; Noyes and Rieseberg, 2000) pero no en *Taraxacum* (Vijverberg *et al.*, 2004).

En *Pennisetum squamulatum*, se determinó que el sector que controla la apomixis apospórica (ASGR: Aposporous Specific Genomic Region) presenta una longitud de casi medio brazo cromosomal y muestra una alta restricción de la recombinación (Ozias-Akins *et al.*, 1998; Akiyama *et al.*, 2005). Una situación similar se encontró en otras especies apospóricas, donde además esta región está asociada a un efecto pleiotrópico letal o ligada a un alelo con letalidad parcial que afecta a los gametos masculinos (Ozias-Akins *et al.*, 1998; Martinez *et al.*, 2001; Pupilli *et al.*, 2004; Stein *et al.*, 2004). Esta falta de recombinación alrededor de la región ASGR podría jugar un rol importante en el mantenimiento de la herencia del carácter al impedir la dispersión de un grupo de genes acoplados (revisado en Ozias-Akins *et al.* 2006).

Trabajos recientes en *Pennisetum squamulatum* y *Pennisetum ciliare*, han revelado que la región ASGR contendría tanto regiones ricas en genes como grandes sectores genómicos con abundancia de elementos repetitivos (Conner *et al.*, 2008). Pudo identificarse que varios de los genes que mapean en el cromosoma portador de la ASGR presentan expresión en óvulos (Zeng *et al.*, 2011).

### 1.1.3- Regulación de la expresión génica en la apomixis

Los estudios orientados a la identificación de genes diferencialmente expresados en los estadios tempranos del desarrollo apomíctico han permitido aislar transcritos de ARNm específicos del carácter en varias especies de gramíneas. Los primeros transcritos asociados a la apomixis fueron aislados en *Pennisetum ciliare* (Vielle-Calzada *et al.*, 1996). Estos estudios permitieron identificar un gen sin función caracterizada (*Pcs-2*) de expresión exclusiva en ovarios sexuales y otros dos (*Pca-2* y *Pca-3*) de expresión en ovarios apomícticos (Vielle-Calzada *et al.*, 1996). Asimismo, Chen *et al.* (1999) informaron la expresión de *asgl* (gen específico de la apomixis 1) en primordios florales de una accesión apomíctica de *Panicum maximum* asociada con la aparición de las células iniciales de la aposporia. La secuencia de *asgl* resultó similar a varios genes específicos de la semilla o el embrión de diferentes especies vegetales, entre ellos *rd22* (un gen expresado

en semillas e inducido por sequía en *A. thaliana*), *grp* (un gen que codifica una proteína rica en glicina de la pared celular de ovarios de *Phaseolus vulgaris*), *usp* (un gen que codifica a una proteína de semilla de *Vicia fava*), *plyg1* (un gen que codifica a un precursor de la cadena beta de poligalacturonasa de *Lycopersicon esculentum*) y *adr6p* (un gen regulado negativamente por auxina de *Glycine max*). La homología significativa con esta variedad de genes permitió asignar a *asg1* una nueva función dentro del complejo de formación del embrión y la semilla (Chen *et al.*, 1999).

En *Paspalum notatum*, la comparación de la expresión de genes en inflorescencias de plantas sexuales y apomícticas permitió la identificación de tres transcritos sobrexpresados en los genotipos apomícticos, relacionados con la cinesina KatD (proteína motora del citoesqueleto) de *A. thaliana* (Pessino *et al.*, 2001). Por otro lado, en *Brachiaria brizantha*, se encontraron genes diferencialmente expresados en plantas apospóricas similares a una proteína miosina, una exonucleasa de la superfamilia RecB, una MAP quinasa, una acuaporina y un factor translocador de proteínas (Rodrigues *et al.*, 2003).

Estudios de hibridización *in situ* de tejidos reproductivos desarrollados en *Panicum maximum* permitieron la evaluación de la expresión del gen *PmASG-1* durante la gametogénesis de un genotipo sexual obligado y uno apospórico (Chen *et al.*, 2005). Estos estudios permitieron comprobar que la expresión de *PmASG-1* se encuentra confinada en óvulos inmaduros exclusivamente del genotipo apomíctico y durante todo el desarrollo del saco apospórico, aunque también se detectó en el polen durante estadios tempranos del desarrollo (Chen *et al.*, 2005). Otros transcritos asociados al desarrollo apospórico pudieron ser identificados en esta especie a través de evaluación de ESTs (Expressed Sequence Tags), estudios de microarreglos y experimentos de RT-PCR y PCR en tiempo real (Yamada-Akiyama *et al.*, 2009). Aquí, se lograron identificar 181 secuencias con homologías significativas con otras especies, 10 de las cuales correspondieron a otras especies apomícticas. Los análisis de RT-PCR y PCR en tiempo real permitieron detectar que algunos de estos transcritos presentaban expresión específica en ovarios apospóricos, particularmente durante los estadios tempranos del desarrollo (Yamada-Akiyama *et al.*, 2009).

Del mismo modo, en *Pennisetum ciliare*, los experimentos de Display Diferencial y de Hibridización Sustractiva permitieron identificar nuevos genes de expresión diferencial en ovarios sexuales y apomícticos (Singh *et al.*, 2007). Así, pudieron ser identificados dos genes nuevos (*Pca21* y *Pca24*) sin caracterización previa en plantas, donde *Pca21* mostró

homologías con genes de trigo expresados durante el desarrollo reproductivo y *Pca24* resultó similar a proteínas de maíz. Los estudios de Northern Blot y de hibridación *in situ* de tejidos revelaron que estos genes se encontraron expresados en los sacos embrionarios inmaduros durante el desarrollo del gametofito femenino únicamente de ovarios apomícticos (Singh *et al.*, 2007).

Trabajos apuntados a lograr una caracterización de los cambios en el transcriptoma de genotipos sexuales y apomícticos fueron desarrollados para las especies apospóricas *Poa pratensis* (Albertini *et al.*, 2004) y *Paspalum notatum* (Laspina *et al.*, 2008). En *P. pratensis*, los análisis de cDNA-AFLP generaron 179 fragmentos derivados de transcritos de espiguillas (Albertini *et al.*, 2004) entre los que se encontró la cinesina, previamente aislada de *P. notatum* (Pessino *et al.*, 2001), y otros genes relacionados con la recepción y transducción de señales hormonales y el control de la estructura del citoesqueleto. Se aislaron luego las secuencias completas de dos transcritos de interés *PpSERK* y *APOSTART* cuyos patrones de expresión fueron corroborados por PCR en tiempo real y por hibridación *in situ* de tejidos. Los análisis en *P. notatum* identificaron secuencias génicas diferencialmente expresadas entre plantas sexuales y apomícticas pertenecientes a cascadas de señalización de tipo ERK y otros procesos celulares como control del ciclo celular, procesamiento de proteínas, señalización intercelular, actividad de transposones, regulación transcripcional y biosíntesis mediada por el retículo endoplasmático (Laspina *et al.*, 2008).

Estudios desarrollados en inflorescencias inmaduras de plantas diplospóricas y sexuales de *Eragrostis curvula* de diferentes ploidías, permitieron aislar 112 genes candidatos de expresión diferencial. Los perfiles de expresión de los genotipos tetraploide diplospórico y diploide sexual resultaron idénticos y diferentes al genotipo tetraploide sexual para un grupo de genes (Cervigni *et al.*, 2008a). La mayor parte de los genes que presentaron este patrón de expresión (44 de 50) estaban reprimidos en la tetraploide sexual. Para explicar estos inesperados resultados los autores formularon la hipótesis de que la falla parcial en el silenciamiento de la expresión de algunos genes durante la poliploidización en esta especie llevaría al desarrollo de la diplosporía (Cervigni *et al.*, 2008b).

Trabajos más recientes apuntados al estudio de genes involucrados en el desarrollo apomíctico fueron llevados a cabo por medio de experimentos SuperSAGE en varias especies del género *Boechea* (Sharbel *et al.*, 2009; Sharbel *et al.*, 2010). Estos estudios

permitieron identificar más de 2.000.000 de fragmentos de ARNm diferencialmente expresados entre óvulos microdisecionados de genotipos sexuales y apomícticos. Aproximadamente 577 transcritos se encontraron diferencialmente expresados durante un único estadio del desarrollo. Más de 590 transcritos se encontraron expresados heterocrónicamente entre óvulos sexuales y apomícticos en cada uno de los estadios del desarrollo apomíctico evaluados. La gran mayoría de ellos se encontraron subexpresados en el desarrollo temprano del óvulo apomíctico. Aproximadamente 1100 transcritos presentaron expresión diferencial significativa en cada uno de los estadios evaluados (Sharbel *et al.*, 2009; Sharbel *et al.*, 2010).

Del mismo modo, estudios recientes del transcriptoma de genotipos sexuales y apomícticos de *Paspalum simplex* por medio de cDNA-AFLP, permitieron la identificación de 202 amplicones de expresión diferencial (Polegri *et al.*, 2010). La mayor parte de estos transcritos se detectó en estadios específicos del desarrollo apomíctico. Las técnicas de hibridización permitieron demostrar que el 10% de los amplicones se encontraron dentro del locus ASGR de *P. simplex*. En cuanto a la identidad de los transcritos identificados, el 60% presentó homologías en las bases de datos, siendo proteínas involucradas en vías de transducción de señales y proteínas de unión a ADN las funciones biológicas las más representadas (Polegri *et al.*, 2010).

La comparación de transcriptomas de óvulos microdisecionados de un genotipo apomíctico dador de polen y una retrocruza 8 que contenía únicamente el cromosoma portador de la ASGR, originado en mismo dador de polen, permitió identificar numerosos transcritos que mapean en el cromosoma portador de la ASGR de *Penisetum squamulatum*. Para ello se realizaron estudios de secuenciación paralela masiva, que permitieron realizar un total de 343.637 lecturas de secuencias del óvulos del dador apomíctico y 332.567 de óvulos de la retrocruza (Zeng *et al.*, 2011). Sesenta y un transcritos fueron localizados *in silico* en el cromosoma portador de la ASGR. Sin embargo, sólo uno de ellos pudo ser localizado físicamente en esta región específica y además presentó expresión exclusiva en tejidos reproductivos. Este candidato mostró homologías con una proteína hipotética presente en sorgo y arroz portadora del gen de la transposasa (Zeng *et al.*, 2011). Es de destacar que diferentes tipos de elementos repetitivos fueron identificados en cada uno de los trabajos de estudios transcriptómicos asociados a la ocurrencia de la apomixis (Albertini *et al.*, 2004; Cervigni *et al.*, 2008b; Laspina *et al.*,

2008; Yamada-Akiyama *et al.*, 2009; Polegri *et al.*, 2010; Sharbel *et al.*, 2010; Zeng *et al.*, 2011), aunque no se profundizó en su estudio.

Por otra parte, varios estudios desarrollados en plantas sexuales modelo mutantes arrojaron concepciones innovadoras. Se identificó el gen receptor kinasa *SERK* (*SOMATIC EMBRYOGENESIS RECEPTOR KINASE*), el cual se encuentra expresado en células de raíces cultivadas por embriogénesis somática, y podría corresponder al iniciador de una cascada de señalización que inicia la embriogénesis (Schmidt *et al.*, 1997). Se demostró que la sobreexpresión del gen *SERK* en *Arabidopsis* (*AtSERK*) aumenta el potencial embriogénico de las células cultivadas. *AtSERK* se encuentra expresado en óvulos inmaduros, inclusive en la ovocélula (Hecht *et al.*, 2001). Entre otros genes asociados a la iniciación del embrión se describieron algunos factores de transcripción: *BBM* (*BABYBOOM*), *LEC1*, *LEC2* (*LEAFYCOTYLEDON 1 and 2*) o *WUS* (*WUSCHEL*) cuya sobreexpresión dirigía la formación de embriones en las plantas transgénicas (Lotan *et al.*, 1998; Stone *et al.*, 2001; Boutilier *et al.*, 2002). Aunque el desarrollo autónomo del endosperma no representa un mecanismo indispensable para el desarrollo de la apomixis, sin embargo este fenómeno ocurre en un número considerable de especies apomícticas. Los análisis de mutantes de *Arabidopsis* han permitido dilucidar algunos aspectos del inicio del desarrollo del endosperma. Las mutantes *fis* (*fertilization-independent seed, mea (medea)*), *fis2* y *fie* (*fertilization-independent endosperm*) presentaron desarrollo del endosperma sin fertilización (Chaudhury and Peacock, 1993; Chaudhury *et al.*, 1998). Por otra parte, el descubrimiento de mutantes de ARGONAUTA 9 (*AGO9*) de *Arabidopsis* permitió asociar por primera vez la regulación por RNA pequeños al desarrollo apomíctico, ya que las proteínas ARGONAUTA son los componentes clave de los complejos de silenciamiento por ARN (Mallory and Vaucheret, 2010). Los estudios demuestran que *ago9* controla la formación de gametas femeninas por medio de la restricción de la especificación del gametofito (Olmedo-Monfil *et al.*, 2010). Así, las mutaciones deletéreas en *ago9* permiten la diferenciación de múltiples células gaméticas capaces de iniciar la gametogénesis, generando un fenotipo que se asemeja a la aposporía. Además, las mutaciones producidas en SUPPRESSOR OF GENE SILENCING 3 (Gen supresor del silenciamiento 3) y RNA\_DEPENDENT RNA POLYMERASE 6 (ARN polimerasa 6 dependiente de ARN) generaron fenotipos análogos al de las mutantes *AGO9*, indicando la necesidad del movimiento de pequeños ARN (ARNsi) hacia el exterior de las células acompañantes para controlar la especificación de las células

gaméticas (Peragine *et al.*, 2004; Yoshikawa *et al.*, 2005). AGO9 interactúa principalmente con ARNs de 24 nucleótidos de longitud derivados de elementos transponibles, y su función sería la de regular la expresión principalmente de retroelementos del tipo *Gypsy*, *Athila* y *CACTA* en las gametas femeninas y sus células accesorias (Olmedo-Monfil *et al.*, 2010). Estos resultados fueron los primeros en sugerir un importante rol de los retrotransposones y del mecanismo de silenciamiento por ARN pequeños en la regulación de la expresión génica del desarrollo apomítico. Recientemente ha sido reportada la caracterización de nuevas mutantes en maíz con un desarrollo reproductivo similar al apomítico y con la generación de gametas viables no reducidas (Singh *et al.*, 2011). El locus mutado correspondió a la proteína *AGO104* (un subtipo de ARGONAUTA) la expresión de los transcritos del gen resultaron confinados principalmente en ovarios y anteras en el estadio cercano a la meiosis (Singh *et al.*, 2011). Según los autores, *ago104*, es similar a su homólogo funcional *ago9* de *Arabidopsis*, aunque AGO9 actuaría en tejidos somáticos para reprimir el desarrollo de células germinales y AGO104 actuaría en células germinales para reprimir el desarrollo de células somáticas (Eckardt, 2011). La comparación de los perfiles de expresión génica generales en los ovarios diseccionados de las plantas mutantes y las silvestres arrojaron diferencias muy significativas, indicando una disrupción masiva de la expresión génica tanto de regiones repetitivas como de genes codificantes para proteínas en los ovarios mutantes (Singh *et al.*, 2011). Por otra parte, las mutantes *ago104* presentaron hipometilación en regiones centroméricas y repeticiones de ADN knob, indicando el requerimiento de *AGO104* para la metilación de estas regiones y el rol regulatorio específico de la metilación del ADN dependiente de *AGO104* en el control de la expresión de genes relacionados a la apomixis en el ovario (Eckardt, 2011; Singh *et al.*, 2011). En consistencia con estos hallazgos, han sido detectados altos niveles de *Dmc102* (proteína homóloga a la metiltransferasa *CMT3* de *Arabidopsis*) en las células de la nucela de maíz, (García-Aguilar *et al.*, 2010). Se conoce (en *Arabidopsis* y en *Oryza*) que estas proteínas participan en la regulación de la expresión génica, particularmente de elementos repetitivos, por medio de mecanismos de metilación del ADN en la vía regulada por ARNs (Chan, 2008; Wu *et al.*, 2010). Además las mutantes simultáneas *Dmc2* y *Dmc3* (una proteína homóloga muy cercana a *Dmc2*) de maíz producen la formación de sacos embrionarios ectópicos, algunos de ellos provenientes de CMM múltiples (García-Aguilar *et al.*, 2010). Algunos trabajos desarrollados en *Triticum* postularon que la demetilación de secuencias repetitivas como los LTRs afectan significativamente la expresión de los genes

adyacentes (Kashkush *et al.*, 2003; Kashkush and Khasdan, 2007; Yaakov and Kashkush, 2011a). Los esfuerzos actuales se encuentran enfocados en la identificación de los diferentes *loci* afectados por la demetilación mediada por ARNsi en las células blanco durante la gametogénesis (Armenta-Medina *et al.*, 2011; Baroux *et al.*, 2011).

#### **1.1.4- Importancia agronómica de la apomixis**

El gran desafío que enfrenta la agricultura en los próximos cincuenta años consiste en duplicar o triplicar la producción actual de alimentos para atender adecuadamente las necesidades de una población mundial creciente. Las investigaciones realizadas por biólogos y botánicos han permitido identificar varios caracteres estructurales o fisiológicos cuyo uso controlado podría aumentar el rendimiento y la calidad de los cultivos, contribuyendo a reducir la incidencia de la pobreza y el hambre a nivel mundial. Entre ellos, la apomixis es posiblemente la de mayor impacto (Koltunow *et al.*, 1995; Vielle-Calzada *et al.*, 1996).

A pesar de su amplia distribución, la apomixis no se encuentra presente en especies vegetales de producción agronómica masiva, aunque sí aparece en las gramíneas forrajeras, la mandioca, la manzana, los citrus, el mango y las frutillas (Savidan, 2000). Se ha especulado que su transferencia a especies mayores como los cereales (maíz, sorgo, arroz y otros) produciría un enorme impacto en la agricultura y la producción de alimentos, posibilitando la fijación de combinaciones híbridas, su mantenimiento y su multiplicación por semillas (Vielle-Calzada *et al.*, 1996).

El uso controlado de esta reproducción clonal vía semillas representaría un impacto potencial enorme para la agricultura, sobre todo en los países con una alta incidencia de pobreza extrema. La introducción de la apomixis en los cereales permitiría la fijación indefinida de la heterosis, habilitando a los campesinos pobres a resembrar parte de su cosecha sin la consecuente pérdida del vigor híbrido. Eso reduciría notablemente los costos de producción y aumentaría los rindes, posibilitando una mayor producción de alimentos en áreas marginales. Los planes de mejoramiento se facilitarían. Ya no sería necesario mantener líneas parentales homocigotas en condiciones de aislamiento para producir los híbridos. Cualquier genotipo (homocigota o heterocigota) podría utilizarse como parental para producir cruza. Las progenies F<sub>1</sub> segregantes podrían ser evaluada en numerosos ambientes particulares de interés, y los individuos con comportamiento superior podrían ser reproducidos clonalmente a través de la introducción de la apomixis. Así, los proyectos

de mejoramiento se simplificarían enormemente, reduciendo tanto el costo como el tiempo empleado en los mismos. Nuevas variedades mejoradas podrían ser producidas muy fácilmente, aún aquellas destinadas a nichos ambientales de superficie reducida. Eso favorecería a los productores pequeños de Latinoamérica, África Subsahariana y Asia Central, que actualmente utilizan variedades mejoradas destinadas a grandes áreas de producción masiva, pobremente adaptadas a las condiciones muy diversas y particulares de sus pequeñas parcelas. Así, la apomixis permitiría maximizar los rendimientos en amplias regiones donde los hábitats son muy diversos, en lugar de restringir la producción a zonas homogéneas de gran extensión, como sucede en la actualidad (Toenniessen, 2001).

Por otra parte, la apomixis facilitaría la obtención de nuevas variedades híbridas, que no sufrirían las típicas restricciones en la fertilidad impuestas por la meiosis. Esto permitiría aprovechar de manera óptima la variabilidad natural encerrada en el germoplasma salvaje. Además, muchas plantas que actualmente se reproducen vegetativamente podrían ser propagadas a partir de semillas. Se facilitaría el uso de transformantes, ya que no sería necesario lograr la homocigosis para evitar la segregación de un transgen. El impacto potencial del uso controlado de la apomixis únicamente para el aprovechamiento máximo de la heterosis en el arroz se calcula en 2.5 a 4 billones de dólares anuales (Pessino and Ortiz, 2010).

## **1.2- Poliploidía**

### **1.2.1- Poliploidía: aspectos generales**

La poliploidización es el fenómeno por el cual el número de cromosomas de un organismo es aumentado a través de la adición de juegos cromosomales completos, presentes en uno o más organismos ancestrales (Grant, 1981). Las primeras investigaciones registradas sobre el tema fueron publicadas en los comienzos del siglo XX (Winkler, 1916; Winge, 1917) siendo un tema de sustancial vigencia en la actualidad (Wendel, 2000; Jeffrey Chen *et al.*, 2004; Comai, 2005; Ma and Gustafson, 2005).

Si bien la poliploidización ha sido tradicionalmente reconocida como un proceso prominente de especiación y adaptación en plantas (Soltis and Soltis, 1995; Ramsey and Schemske, 1998; Comai, 2005), se ha demostrado que las duplicaciones genómicas también ocurrieron frecuentemente en levaduras, insectos, anfibios, reptiles y peces (Lewis, 1980b), y se han encontrado evidencias sobre el origen poliploide del genoma de

mamíferos (Miklos and Rubin, 1996). Se considera que aproximadamente el 70% de las angiospermas ha experimentado uno o más eventos de duplicación genómica en algún punto de su historia evolutiva (Goldblatt, 1980; Lewis, 1980a; Materson, 1994). Se ha propuesto que la ocurrencia frecuente de la poliploidía en las angiospermas está ligada a un rol evolutivo, ya que la condición poliploide confiere ventajas adaptativas respecto a los progenitores diploides. Dichas ventajas abarcan desde un vigor superior, cambios en el tamaño de las plantas y tiempo de floración, hasta una mayor viabilidad, mayor rendimiento reproductivo y homeostasis fisiológica exacerbada (Levin, 1983; Soltis *et al.*, 2009). Estos caracteres presentes en los poliploides, tales como tolerancia a estrés, apomixis, aumento de tamaño de órganos y biomasa, podrían permitirles colonizar nuevos nichos ecológicos o bien representar características agronómicas deseables (Grant, 2002; Blanc and Wolfe, 2004; Comai, 2005). Además los autopoliploides particularmente, como resultado de su herencia polisómica, poseen una heterocigosis más elevada respecto de sus padres diploides y pueden mantener varios alelos para un mismo *locus*. Esto último aporta razones genéticas contundentes para poder comprender el éxito de los autopoliploides en las poblaciones naturales (Osborn *et al.*, 2003; Soltis and Soltis, 2003; Comai, 2005).

Los individuos poliploides pueden generarse a partir de la reproducción sexual al involucrarse gametas no reducidas ( $2n$ ), o alternativamente, mediante la duplicación somática de cromosomas. La poliploidía es clasificada tradicionalmente en función a la procedencia de los complementos genómicos que compondrán el poliploide. Según este principio, el término autopoliploide se utiliza para denotar a los poliploides que se generan a partir de cruces entre individuos de una misma especie y para aquellos que provienen de una duplicación del genoma, que puede ser espontánea o derivar de tratamientos con agentes químicos como la colchicina. Por su parte, el término alopoliploide se refiere a poliploides derivados de híbridos entre individuos de distintas especies (Stebbins, 1947).

Los alo y autopoliploides han sido identificados habitualmente por el modo de herencia de sus cromosomas. Los alopoliploides presentan un apareamiento bivalente y herencia disómica, mientras que los autopoliploides exhiben apareamiento multivalente y herencia polisómica (Stupar *et al.*, 2007). Esta clasificación de los poliploides resulta insuficiente para agrupar los numerosos casos intermedios comúnmente hallados en la naturaleza, donde los individuos presentan algunos sectores del genoma que divergen más que otros, y por lo tanto surgen comportamientos mixtos entre herencia polisómica y

disómica. Se ha identificado a éstos últimos con el nombre de *alopoliploides segmentales* y *aloautoploiploides* (Ramsey and Schemske, 2002).

### 1.2.2- Modificaciones genómicas que acompañan la poliploidización

Los estudios más recientes de secuenciación y mapeo comparativo han evidenciado que las plantas con flores poseen genomas con una considerable redundancia génica y la presencia de una alta proporción de pseudogenes (secuencias genómicas generadas a partir de secuencias modificadas de genes funcionales, generalmente luego de duplicaciones génicas/genómicas) (Zou *et al.*, 2009). Se conoce actualmente que, por ejemplo, muchos de los genes de *Arabidopsis* se encuentran duplicados, lo que demuestra la posible ocurrencia de dos o más rondas de duplicación genómica a lo largo del desarrollo evolutivo de esta especie (Bennett, 2003; Blanc and Wolfe, 2004; Yu *et al.*, 2010). Se cree que estas rondas de doblado cromosómico frecuentemente son seguidas por una serie de modificaciones genómicas que conducen a una restauración del comportamiento meiótico diploide. Además se establecen nuevos patrones de expresión génica orientados a mantener la homeostasis. Se conoce que la síntesis de aloploiploides de especies de *Brassica* y trigo provoca grandes cambios genéticos rápidos, que suceden inmediatamente luego de la poliploidización (Song *et al.*, 1995; Ozkan *et al.*, 2001). Además, ha sido demostrado que en trigo la aloploiploidización genera hipo o hipermetilación en el 43% de los casos, siendo la hipometilación predominante durante las primeras tres generaciones subsiguientes a la síntesis (Kraitshtein *et al.*, 2010). La evidencia general apunta a la ocurrencia de procesos genéticos y epigenéticos que incluirían rearrreglos del genoma, modulación de la actividad de retrotransposones y modificación del patrón de metilación de citosinas del genoma, aunque el grado de importancia de estos fenómenos es altamente dependiente de la especie (Kashkush *et al.*, 2003; Comai, 2005; Pessino and Martelotto, 2006; Yaakov and Kashkush, 2011b).

Asimismo, se ha demostrado la ocurrencia de cambios genómicos rápidos en autoploiploides. Fueron reportados polimorfismos genéticos para una serie euploide (4x-2x-4x) en *Eragrostis curvula*, donde frecuentemente se observó que las formas alélicas del progenitor tetraploide natural desaparecían durante la dihaploidización y eran luego recobradas al restaurar el nivel de ploidía original a través del tratamiento con colchicina (Mecchia *et al.*, 2007), un comportamiento que se calificó como “revertante”. Del mismo modo, los estudios epigenéticos desarrollados en esta serie permitieron identificar el

mismo tipo de restauración comparando los pasajes 4x a 2x y 2x a 4x (Ochogavía *et al.*, 2009). Asimismo, los análisis de expresión en estos genotipos permitieron detectar alteraciones reversibles asociadas a los cambios de ploidía (Mecchia *et al.*, 2007; Cervigni *et al.*, 2008a). Un análisis de las correlaciones de cada uno de estos factores permitió inferir que la expresión génica se encuentra fuertemente correlacionada con las modificaciones epigenéticas y un poco menos marcadamente con las modificaciones genéticas (Ochogavía *et al.*, 2009).

Por otra parte, los análisis desarrollados en un sistema euploide de *P. notatum* conformado por un genotipo diploide y su autotetraploide derivado, generado por tratamiento con colchicina, permitieron determinar que las modificaciones a nivel genómico involucraron principalmente la pérdida o modificación de las formas alélicas que se encontraban presentes en el genotipo diploide original (Martelotto *et al.*, 2007). Estas modificaciones genéticas eran acompañadas por un rearrreglo parcial de transcriptoma, involucrando alteraciones en la expresión génica del nuevo tetraploide (Martelotto *et al.*, 2005). Estudios en los patrones de metilación de citosinas globales desarrollados con esta serie permitieron identificar que, si bien el nivel de metilación entre plantas sexuales y apomícticas permaneció relativamente constante, los genotipos tetraploides presentaron mayor variabilidad que los diploides y nuevos epialelos surgieron luego de la tetraploidización (Rodríguez *et al.*, 2011). A partir de un dendrograma de distancias epigenéticas, se identificaron agrupamientos entre los genotipos tetraploides apomícticos y por otro lado los genotipos diploides y un genotipo tetraploide de reproducción sexual derivado. El análisis y secuenciación de estas bandas diferenciales permitió aislar genes de bajo número de copia, elementos repetitivos y secuencias genómicas (Rodríguez *et al.*, 2011).

### **1.2.3- Poliploidía y expresión génica**

La duplicación de los genomas, tanto en alo- como en auto-poliploides, involucra un aumento en el número de copia de los genes (Wendel, 2000). Ha sido demostrado en una serie euploide de maíz, compuesta por genotipos monoploides, diploides, triploides y tetraploides, con un fondo genético común, que la poliploidía tiene un efecto general de incrementar los niveles de expresión de los genes en proporción a la dosis génica conferida por el nivel de ploidía (Guo *et al.*, 1996). Sin embargo, algunos genes presentan efectos particulares, como marcada sobreexpresión o inactivación, y algunos respondieron al tipo

de poliploidía (par-impar) (Guo et al., 1996). Estos efectos de dosaje alélico han sido observados en varias especies modelo, para muchos genes regulatorios del desarrollo (Guo et al., 1996; Michaels and Amasino, 1999).

La cantidad de producto expresado resulta crítica para la correcta funcionalidad celular por lo que un aumento global del número de copia de cada uno de los genes duplicados alteraría negativamente las complejas redes regulatorias. Varias investigaciones han estudiado la respuesta del transcriptoma a los cambios del nivel de ploidía. Se han reportado cambios epigenéticos y de silenciamiento génico en *A. thaliana* luego del proceso de aloploidización (Comai et al., 2000; Lee and Chen, 2001). En trigo, han sido observadas rápidas pérdidas génicas, silenciamiento génico y activación transcripcional de retroelementos en el período inmediatamente posterior a la aloploidización (Kashkush et al., 2003).

En cuanto a la identidad de los transcriptos alterados durante las variaciones en el nivel de ploidía, una comparación de las modificaciones en el transcriptoma de inflorescencias de un citotipo diploide y su tetraploide derivado desarrollada en la especie *P. notatum*, permitió detectar que muchos de los genes expresados diferencialmente se encontraron relacionados a la modificación de la estructura de la cromatina, regulación de la transcripción, proteólisis, plegamiento de proteínas y metabolismo de lípidos (Martelotto et al., 2005; Laspina et al., 2008). Estudios genéticos y de metilación de ADN realizados para algunos de estos genes no detectaron diferencias genéticas ni epigenéticas entre las líneas, por lo que la expresión diferencial de estos transcriptos parecería estar relacionada con el efecto de otros elementos regulatorios, como la actividad de factores de transcripción (Martelotto et al., 2005; Martelotto et al., 2007). Del mismo modo, se identificaron numerosos genes controlados por los niveles de ploidía en genotipos de *E. curvula* pertenecientes a pocas familias funcionales (Cervigni et al., 2008a).

### **1.3- Relación entre la apomixis y la poliploidía**

El hecho de que los niveles de ploidía condicionan el desarrollo de la apomixis dificulta la caracterización del control genético de la misma. Los individuos apomícticos exhiben el carácter generalmente a partir del nivel tetraploide o mayor (Asker and Jerling, 1992; Savidan, 2000). Sin embargo, se ha identificado ocasionalmente la presencia de sacos apomeióticos, junto a los sexuales, en algunos genotipos diploides naturales de los géneros *Paspalum* (Normann et al., 1989; Quarín et al., 1989; Quarín and Normann,

1990), *Potentilla* (Asker and Jerling, 1992), *Boechera* (Kantama *et al.*, 2007) y *Brachiaria* (Naumova *et al.*, 1999).

Un estudio reciente desarrollado en *Paspalum rufum* ha demostrado la existencia de sacos apospóricos funcionales en óvulos de la especie a nivel diploide (Siena *et al.*, 2008). En este trabajo se identificaron cinco individuos clonales idénticos a la planta madre, obtenidos por autopolinización de un individuo diploide. Además, no se obtuvo progenie materna durante los cruzamientos del mismo diploide con otros genotipos (Siena *et al.*, 2008). Esto sugiere que el genotipo materno, en caso de existir disponibilidad de polen heterólogo, se comportaría de manera sexual, pero permitiría la ocurrencia ocasional de agamosperma en las autopolinizaciones. Estas observaciones indican que la apomeiosis sería funcional a nivel diploide, al menos en algunos individuos de estas gramíneas.

En general las especies apomícticas presentan razas de bajos niveles de ploidía (usualmente diploides) que se reproducen por sexualidad y otras de mayor nivel de ploidía (como tri, tetra o pentaploides) que se reproducen por apomixis. Estos sistemas integrados por individuos sexuales y apomícticos de distinto nivel de ploidía se conocen como *complejos agámicos* y se consideran estructuras reproductivas evolucionadas (Daurelio *et al.*, 2004). La variabilidad está asegurada en el nivel diploide, donde ocurre la sexualidad. Ocasionalmente, podrían generarse nuevos poliploides a partir de los diploides, lo que trasladaría la variabilidad a niveles de ploidía superiores, que se reproducen clonalmente de manera muy eficiente (Daurelio *et al.*, 2004).

Existen evidencias experimentales en el género *Paspalum* de que la duplicación del nivel de ploidía, desde el estado diploide al tetraploide inducido por tratamiento con colchicina provocaría un cambio en el modo de reproducción de sexual a apomíctico en algunos genotipos (Quarín *et al.*, 2001). Estos hallazgos permitieron hipotetizar que los factores determinantes de la apomixis podrían encontrarse presentes a nivel diploide, al menos en algunos casos, pero que su expresión se encontraría reprimida (Quarín *et al.*, 2001). En el trabajo de Quarín *et al.* (2001) se postuló la existencia de un determinante genético de la apomixis presente a nivel diploide, pero con una expresión condicionada por la actividad de uno o varios genes adicionales sujetos a efectos de dosaje.

Otra hipótesis apunta a categorizar a la apomixis como consecuencia de esta desestabilización del genoma (Quarín *et al.*, 2001; Spillane *et al.*, 2004; Cervigni *et al.*, 2008a), ya que la redundancia génica generada durante la poliploidización representa una pérdida de balance en la expresión de genes, con la consecuente desregulación de la vía

sexual. Muchos otros autores plantearon posibles explicaciones respecto a la causa de la estrecha relación entre la apomixis y la poliploidía. Nogler (1984) propuso, a partir de sus observaciones en *Ranunculus auricomus*, la existencia de un alelo dominante determinante de la apomixis que sólo podía ser transferido por gametas diploides. Así, las gametas haploides generaban únicamente genotipos sexuales. Observaciones similares del fenómeno fueron realizadas en *Tripsacum* (Grimanelli *et al.*, 1998a) y *Penisetum* (Roche *et al.*, 1999). Una hipótesis alternativa es que una condición epigenética asociada con la poliploidía posibilitara la expresión del carácter. Así, a nivel diploide el/los factor/es determinante/s de la apomixis se encontraría/n presente/s, pero su expresión sería posible únicamente en entornos poliploides (Rodríguez *et al.*, 2011).

#### **1.4- El género *Paspalum***

El género *Paspalum* pertenece a la familia de las gramíneas (subfamilia *Panicoideas*, tribu *Paniceas*), y comprende más de 400 especies (Chase, 1929). Muchas de estas especies constituyen importantes recursos forrajeros para las regiones tropicales y subtropicales de América, donde se encuentran ampliamente distribuidos (Burton, 1974). La poliploidía está presente en aproximadamente en el 80 % de las especies del género, existiendo una gran variación en el nivel de ploidía que va desde 3x hasta excepcionalmente 16x (Burton, 1948). Generalmente, las especies de *Paspalum*, se encuentran conformadas por citotipos diploides de reproducción sexual y poliploides apomícticos o sexuales (Quarín and Norman, 1987; Quarín, 1992). Además, muchas especies tetraploides apomícticas incluyen razas coespecíficas diploides, autoincompatibles y de reproducción sexual (Quarín, 1992).

El gran número de especies de *Paspalum* existentes en la naturaleza, la amplia distribución geográfica que presentan y sus excepcionales características forrajeras, hacen del género uno de los más importantes y promisorios para el desarrollo de cultivares forrajeros adaptados a las regiones cálidas y templadas de América.

##### **1.4.1- *Paspalum notatum* (Pasto horqueta, “Bahigrass”)**

*Paspalum notatum* Flüggé, de nombre vulgar pasto horqueta o “bahigrass”, es una especie perenne distribuida desde México a Argentina. Es formadora de estolones y de gruesos rizomas, que la afirman fuertemente al terreno, y despliega densas matas radiculares (Soreng *et al.*, 2009). Presenta hojas chatas y glabras que pueden medir desde

20 cm hasta 1 m de longitud. Cuenta con una lígula anidada con pelos cortos. Las bases foliares al término de cada rizoma usualmente tienen un matiz púrpúreo. Las inflorescencias se conforman de dos racimos terminales, de 5 a 10 cm de longitud, con espiguillas superpuestas en dos filas insertadas a lo largo de la cara inferior (Figura II). Dentro de cada espiguilla se encuentra la flor con estigma y estambres de color negro. El fruto es de tipo cariopse oval, de dimensiones que abarcan los 2 a 3,5 mm de longitud (Soreng *et al.*, 2009).



**Figura II:** Fotografía de individuos de la especie *Paspalum notatum* tomada en el predio del Instituto de Botánica del Nordeste (IBONNE-CONICET). Se observa su fuerte desarrollo vegetativo y la morfología de la inflorescencia al que hace referencia su nombre vulgar, pasto “horqueta”. La especie presenta una amplia adaptabilidad pudiendo alcanzar un alto desarrollo vegetativo en suelos de fertilidad variable.

Las razas tetraploides se distribuyen en tierras abiertas, sabanas y pasturas hasta 2.000 metros sobre el nivel del mar desde México central hasta el centro de Argentina, siendo el principal constituyente de los campos naturales del sur de Brasil, Paraguay, noreste de Argentina y Uruguay (Chase, 1929). La forma diploide es nativa de las cuencas de los ríos Paraná y Uruguay en el noreste de Argentina y se propuso que su centro de origen sería la isla Berduc, en la provincia de Santa Fe (Burton, 1967).

Fiel a su género, esta especie presenta varios niveles de ploidía y un sistema reproductivo complejo con citotipos diploides autoincompatibles de reproducción sexual y poliploides (3x, 4x, 5x) con reproducción apomíctica, pseudógamos y autocompatibles.

Aproximadamente el 98% de los poliploides son tetraploides y la forma natural de reproducción más frecuente en la especie es la apomixis apospórica (Quarín, 1992).

*Paspalum notatum* Flüge cuenta con un número básico de cromosomas  $x = 10$ . Los citotipos diploides presentan un número  $2n = 2x = 20$  y los tetraploides  $2n = 4x = 40$ . Los citotipos triploides  $2n = 3x = 30$  y pentaploides  $2n = 5x = 50$  son muy poco frecuentes en la naturaleza (Burton, 1948).

#### 1.4.2- La apomixis en *Paspalum notatum*

La apomixis en *Paspalum notatum* es del tipo gametofítica apospórica pseudógama. Como se mencionó anteriormente, la gran mayoría de los individuos naturales de esta especie son tetraploides apospóricos, y una minoría, diploides y de reproducción sexual. No se ha descrito la presencia de plantas tetraploides completamente sexuales de *P. notatum* en la naturaleza, aunque sí se han obtenido experimentalmente mediante tratamientos con colchicina de diploides o por cruzamientos entre tetraploides apomícticos facultativos generados artificialmente y tetraploides apomícticos naturales (Forbes and Bourton, 1961; Quarín *et al.*, 2001).

Los citotipos diploides presentan sacos embrionarios meióticos similares a los de tipo *Polygonum*, aunque las antípodas proliferan. Se observa únicamente un saco embrionario por óvulo. Por otro lado, los tetraploides presentan sacos embrionarios no reducidos, donde se distingue la ovocélula acompañada por una o dos sinérgidas y una célula central binucleada (Quarín and Burson, 1991). Los sacos no reducidos no presentan antípodas y suelen ser múltiples (Quarín and Burson, 1991).

##### 1.4.2.1- Caracterización genética de la apomixis en *Paspalum notatum*

En *P. notatum* la región ASGR (Apospory Specific Genomic Region) abarca aproximadamente 36 Mpb (Stein *et al.* 2007). Se considera que varios genes incluidos en la misma podrían operar juntos en el control de la apomixis y serían transferidos como una unidad funcional (Stein *et al.*, 2004). Además, las regiones cercanas a este locus presentan apareamiento preferencial de cromosomas, a diferencia del resto del genoma donde predomina una clara herencia polisómica (Stein *et al.*, 2004).

En *P. notatum* se encontraron evidencias citológicas y moleculares de la existencia de una posible inversión en la región que gobierna la apomixis (Stein *et al.*, 2004). Las razas apomícticas presentan meiosis masculinas irregulares. También se observó en

estudios de sintenia *in silico* que sectores genómicos ubicados en ambos extremos distales del cromosoma 12 de arroz mapean ligados en *P. notatum* (Pupilli *et al.*, 2004). De haberse producido una inversión alrededor de la región ASGR de *P. notatum*, podría explicarse en parte la alta tasa de inviabilidad de gametos masculinos y la fuerte distorsión de la segregación del carácter en esta especie (Stein *et al.*, 2004; Martinez *et al.*, 2007; Stein *et al.*, 2007).

#### **1.4.2.2- Estudios de la expresión génica en genotipos apomícticos y sexuales de *Paspalum notatum***

Los primeros trabajos de aislamiento de ARNms relacionados a la ocurrencia de la apomixis en *P. notatum* fueron reportados por Pessino *et al.* (2001). Como se mencionó anteriormente, en estos estudios se logró identificar una secuencia sobreexpresada en plantas apomícticas homóloga a la cinesina KatD de maíz, que posteriormente también fue aislada de inflorescencias apomícticas y sexuales de *Poa pratensis* (Albertini *et al.*, 2004).

Posteriormente, nuestro grupo produjo un segundo trabajo, con una caracterización más detallada del transcriptoma de inflorescencias de plantas apomícticas y sexuales de esta especie (Laspina *et al.*, 2008). Estos estudios permitieron identificar 65 transcriptos únicos controlados diferencialmente por el modo de reproducción (Laspina *et al.*, 2008). Cuarenta y cinco de ellos correspondieron a secuencias génicas mayoritariamente agrupadas en unas pocas clases ontológicas. Entre estas se encontraron proteínas involucradas en cascadas de transducción de señales: kinasas, receptores kinasas, fosfatasas y proteínas ancladas a membrana plasmática. Algunos de estos genes también habían sido aislados en *Poa pratensis*, asociados a la ocurrencia de la apomixis (Albertini *et al.*, 2004). Muchos de los genes diferencialmente expresados entre plantas sexuales y apomícticas, tanto en *Poa* como en *Paspalum*, parecerían pertenecer a una vía de señalización de tipo ERK (extracelular receptor kinase) mediada por ortólogos de la proteína Ras y Fosfolipasa C (Mishra *et al.*, 2006). Otros grupos de transcriptos aislados en *P. notatum* correspondieron a genes de procesamiento de proteínas, como proteínas ribosomales, proteasas, y poliubiquitinas, de las cuales, algunas de ellas también fueron identificadas en *Poa* (Albertini *et al.*, 2004). Y también se aislaron secuencias cuya función se asocia a la biosíntesis de proteínas y la funcionalidad del retículo endoplasmático (Laspina *et al.*, 2008). Asimismo, se identificaron numerosas secuencias relacionadas a la regulación de la actividad de transposones, como ARN polimerasas dirigidas por ADN,

transposasas, y proteínas no caracterizadas de transposones. Muchos de los transcriptos aislados pudieron ser clasificados como genes de regulación transcripcional y genes de control del ciclo celular, entre estos últimos se encontraron kinasas “check-point” y proteínas supresoras de apoptosis celular. También, se identificaron secuencias relacionadas a las actividades bioenergéticas celulares, como el *cytP450*, la proteína NADHdh y Adh (Laspina *et al.*, 2008). Estas secuencias también habían sido identificadas en *E. curvula* como asociadas a la apomixis/poliploidía (Cervigni *et al.*, 2008b). Del total de 65 transcriptos controlados diferencialmente por el modo de reproducción aislados por Laspina *et al.* (2008), 22 no pudieron ser anotados funcionalmente por no mostrar homologías significativas en las bases de datos.

Un tipo análogo de análisis comparativo había sido realizado con la finalidad de indagar la naturaleza de los genes diferencialmente expresados según el nivel de ploidía en *P. notatum* (Martelotto *et al.*, 2005). Este estudio realizado a partir de inflorescencias de plantas diploides y tetraploides (ambas sexuales) de *P. notatum*, permitió identificar un grupo de 48 transcriptos de expresión diferencial que podrían ser los responsables de determinar el transcriptoma poliploide particular necesario para la manifestación de la apomixis (Martelotto *et al.* 2005). De este grupo, 20 correspondieron a secuencias únicas sin homologías en las bases de datos.

En esta tesis doctoral se presenta una caracterización estructural y funcional del grupo de transcriptos desconocidos de *P. notatum*, compuesto por los 22 diferencialmente expresados en plantas tetraploides sexuales y apomícticas (Laspina *et al.*, 2008), los 20 diferencialmente expresados en plantas sexuales diploides y tetraploides (Martelotto *et al.*, 2005) y 16 transcriptos aislados posteriormente en experimentos de Display Diferencial en este último sistema (Silvina Pessino, datos no publicados). Estos fragmentos de secuencia conocida y función desconocida representan una importante fuente potencial de información sobre las bases moleculares de la apomixis y de la influencia de la poliploidía en el moldeado de los genomas vegetales.

## 2- OBJETIVOS

### 2.1- Objetivo general

El objetivo general de este proyecto fue contribuir a la caracterización de las bases moleculares de la aposporia en *Paspalum notatum*, mediante el estudio de un grupo de 58 secuencias génicas desconocidas que fueron asociadas al carácter en estudios previos.

### 2.2- Objetivos específicos

*Objetivo específico 1:* Detectar homologías crípticas o presencia de dominios estructurales definidos en los 58 fragmentos génicos incógnita. *Hipótesis asociada 1:* Las actualizaciones de las bases de datos públicas y la implementación de nuevas herramientas bioinformáticas permitirán detectar los genes ortólogos putativos de las secuencias incógnitas en las especies modelo.

*Objetivo específico 2:* Detectar y descartar posibles contaminaciones. *Hipótesis asociada 2:* algunas de las secuencias aisladas de los experimentos de Display Diferencial podrían provenir de contaminaciones en las muestras.

*Objetivo específico 3:* Extender las secuencias de ADNc completas de los genes que permanezcan sin anotación funcional. *Hipótesis asociada 3:* Los fragmentos de ADNc pueden ser ampliados por amplificación de las secuencias flanqueantes.

*Objetivo específico 4:* Una vez obtenidas las secuencias de ADNc expandidas, repetir los análisis bioinformáticos. *Hipótesis asociada 4:* Los fragmentos de mayor tamaño de ADNc tienen mayor posibilidad de revelar homologías significativas en las búsquedas bioinformáticas.

*Objetivo específico 5:* Realizar un análisis casuístico funcional de candidatos selectos modelo para cuantificar su perfil de expresión en el desarrollo, localizar a nivel tisular la expresión (hibridización *in situ*), determinar el número de copias genómicas, ubicar la posición de los genes en el genoma y aislar otras secuencias relacionadas.

*Hipótesis asociada 5: Algunos de los transcritos aislados son elementos repetitivos que contienen segmentos génicos transduplicados y otros son precursores de microARN.*

### 3- MATERIALES Y MÉTODOS

#### 3.1- Material vegetal

El material vegetal utilizado en este trabajo consistió en dos individuos tetraploides sexuales de origen experimental del género *Paspalum notatum*, Q4188 ( $2n=4x=40$ ) y C4-4X ( $2n=4x=40$ ), cuyos modos de reproducción fueron comprobados por citoembriología y marcadores moleculares (Quarín *et al.*, 2001; Quarín *et al.*, 2003); y el genotipo apomítico obligado Q4117 ( $2n=4x=40$ ), colectado de una población natural (Ortiz *et al.*, 1997). Se trabajó también con un total de 62 individuos de una población de mapeo de tipo “pseudo-testcross”, segregante para la aposporia, proveniente del cruzamiento de los genotipos Q4188 x Q4117. Esta población fue generada por el Ing. Camilo Quarín y colaboradores en el Instituto de Botánica del Nordeste (IBONE-CONICET) y había sido utilizada para la construcción de un mapa genético completo de la especie a nivel tetraploide, donde se identificó y caracterizó la región de la apomixis (Stein *et al.*, 2007).

Las espiguillas de inflorescencias inmaduras del genotipo Q4117 y Q4188 fueron utilizadas para las extracciones de ARN total en diferentes estadios del desarrollo. Dichos estadios fueron evaluados en función de la macromorfología de la inflorescencia y el desarrollo del polen, utilizando la metodología y calendario reproductivo publicado por Laspina *et al.* (2008). Las flores centrales de una de las hileras de las inflorescencias fueron seleccionadas y analizadas para determinar su estadio de desarrollo. Tanto las condiciones de crecimiento de las plantas como el momento de recolección fueron cuidadosamente verificados y estandarizados para evitar cualquier efecto de factores ambientales o ritmos circadianos endógenos sobre la expresión génica. Se seleccionaron espiguillas en estadio pre-meiótico (0), pre-meiosis tardía / meiosis (I/II), post-meiosis (IV/V) y anthesis (VI), según el calendario reproductivo publicado por Laspina *et al.* (2008). Por otro lado se utilizaron cortes de inflorescencias del genotipo Q4117 y C4-4x en estadio pre-meiosis tardía / meiosis (I/II) para realizar hibridaciones *in situ* de tejidos. Este estadio coincide con el momento inmediato previo al surgimiento de las células iniciales de la aposporia.

#### 3.2- Extracción de ADN

Para la extracción de ADN se utilizó el método CTAB (Sahgai-Marroof *et al.*, 1984) a partir de 6 g de hojas jóvenes, sanas y limpias de *P. notatum*. Las hojas fueron molidas

finamente en presencia de N<sub>2</sub> líquido y luego introducidas en un tubo Falcon de 50 ml. Se adicionaron 10 ml de solución reguladora de extracción CTAB (Tris HCl 0,1M pH=7,5; NaCl 0,7 M; EDTA 0,05 M pH=8; CTAB 2%; β-Mercaptoetanol 0,14M) previamente llevado a una temperatura de 65°C. Se incubó 60 min con agitación a 65 °C y luego las muestras fueron enfriadas 15 min a temperatura ambiente. A continuación se agregaron 5 ml de cloroformo manteniendo una agitación suave durante 10 min para facilitar el contacto con el material vegetal. La mezcla se centrifugó a 3500 rpm durante 10 min y el sobrenadante fue transvasado a un nuevo tubo Falcon de 50 ml. Se agregaron 5 ml de cloroformo, se mantuvo en agitación durante 10 min y se realizó una nueva centrifugación de 10 min a 3500 rpm. Posteriormente se extrajo la solución acuosa del sistema bifásico formado y se le agregaron 5 ml de 2-propanol a 4°C. Se colectó el ADN con ansas de vidrio esterilizadas, se lo sumergió en 1 ml de solución de lavado durante 10 min (EtOH 75%; NaOAc 0,2M) y posteriormente se lo dejó secar 30 min a 37°C. Una vez seco, se colocó el ADN en un tubo Eppendorf de 1,5 con 1 ml de solución reguladora TE (Tris-HCl 10 mM pH8, EDTA 0,1 mM). Se incubó durante 2 hs a temperatura ambiente en agitación suave para lograr la disolución del ADN y se realizó un tratamiento con ARNasa agregando 5% v/v de la enzima (4 mg/ml) e incubando 30 min a 37°C. Luego, se precipitó el ADN agregando 5% v/v de NaCl 5 M y 2,5 vol de EtOH absoluto. Se incubó durante toda la noche a -20°C y al día siguiente se centrifugó a 13000 rpm durante 20 min. Se descartó el sobrenadante, se agregó 1 ml de EtOH 70% para disolver los restos de sal y se centrifugó a 13000 rpm durante 20 min. Se descartó nuevamente el sobrenadante y se incubó a 37°C durante 1 h para eliminar todo resto de solución de lavado. Finalmente se agregaron 250 µl de agua ultrapura estéril y se incubó toda la noche a temperatura ambiente con agitación para lograr la resuspensión total. El ADN extraído en solución se cuantificó y analizó en un espectrofotómetro Perkin-Elmer Lambda Bio+, midiendo el índice de absorbancia a 260 nm, y considerando la relación 1 OD = 50 µg/µl. La pureza de las muestras se controló midiendo la relación de Abs260nm/Abs280nm, se consideraron aceptables relaciones entre 1,7 y 1,9. La integridad del ADN se estimó realizando electroforesis de 300 ng de ADN en geles de agarosa al 0.8 % p/p, corridos a 60mA. Las muestras de ADN fueron llevadas a una concentración de 300 ng/ µl, alicuotadas y almacenadas a -80°C. Las soluciones de trabajo fueron diluidas a una concentración de 30 ng/µl y almacenadas a -20°C.

### 3.3- Extracción de ARN total

La extracción de ARN total se realizó a partir de inflorescencias de los genotipos Q4117 y Q4188 en cuatro estadios del desarrollo reproductivo. Las inflorescencias fueron seleccionadas y recolectadas durante las primeras horas de la mañana. Se realizó un análisis macromorfológico y del desarrollo del polen para determinar el estadio reproductivo de las espiguillas y así poder clasificarlas según el calendario reproductivo de *P. notatum* (Laspina *et al.*, 2008). Una vez asignado el estadio, se partió de 30 mg de tejido reproductivo y se utilizó el equipo de extracción de ARN SV Total RNA Isolation Kit (Promega), siguiendo rigurosamente las instrucciones del fabricante. Tanto la calidad como la cantidad de ARN extraído fueron determinadas espectrofotométricamente en un equipo Perkin-Elmer Lambda Bio+, midiendo el índice de absorbancia a 260 nm, y considerando la relación  $1 \text{ OD} = 40 \mu\text{g}/\mu\text{l}$ . La pureza de las muestras se controló midiendo la relación de  $\text{Abs}_{260\text{nm}}/\text{Abs}_{280\text{nm}}$  (una muestra pura muestra una relación de 2, se consideraron aceptables relaciones entre 1,8 y 2,2). La integridad del ADN se estimó realizando electroforesis de 300 ng de ARN en geles de agarosa al 1 % corridos a 60 mA, considerando que una muestra íntegra presenta dos bandas bien definidas que corresponden a los ARN ribosomales 18S y 5S y una banda difusa de bajo peso molecular que corresponde a los ARN de transferencia. Las muestras fueron diluidas hasta alcanzar una concentración de  $300 \text{ ng}/\mu\text{l}$ , alicuotadas y almacenadas a  $-80^\circ\text{C}$ .

### 3.4- Amplificación de ADN genómico

Se utilizó el ADN genómico extraído de hojas jóvenes de los genotipos Q4188 y Q4117 para realizar aplicaciones de PCR con la finalidad de localizar la presencia de los fragmentos candidatos en el genoma de *P. notatum*. Las reacciones se llevaron a cabo en un volumen final de 25  $\mu\text{l}$  incluyendo 60 ng de ADN total, 1X de solución reguladora GoTaq (Promega), 0,2 mM de dNTPs, 0,2  $\mu\text{M}$  de cebadores específicos diseñados para cada gen (ver Tabla 1) y 1,5 U de polimerasa Go Taq (Promega). El termociclado se realizó en el equipo My Cycler Thermal (Biorad) y consistió en: 5 min a  $94^\circ\text{C}$  seguido de 40 ciclos de 30 s a  $94^\circ\text{C}$ , 1 min a  $55-69^\circ\text{C}$  y 1 min a  $72^\circ\text{C}$ . La fase de elongación fue de 5 min a  $72^\circ\text{C}$ . Los fragmentos amplificados fueron visualizados por medio de electroforesis en geles de agarosa 1,5 % p/v en cubas horizontales LabNet (E-0350; 1007-10) a 60 mA con solución reguladora TAE (Tris-acetato 0,04M y EDTA 1mM) y posteriormente teñidos con bromuro de etidio (EtBr) 1% v/v. Los productos de amplificación que no pudieron ser

detectados en agarosa se analizaron por electroforesis en geles de poliacrilamida 5% p/v, en solución reguladora TBE (Tris 0,108 % p/v; ácido bórico 0,55 % p/v y EDTA 2 mM) y fueron revelados utilizando tinción con plata, utilizando el protocolo de BioRad (BioRad Silver Stain, BioRad).

### **3.5- Construcción de bibliotecas no clonadas de ADN**

Se construyeron bibliotecas no clonadas a partir de ARN total de inflorescencias inmaduras en estadio pre-meiosis tardía / meiosis (I/II) de los genotipos Q4117 y Q4188. Estas construcciones fueron realizadas en colaboración con el grupo de trabajo de la Dra. Vera Campos Carneiro del Instituto EMBRAPA-CENARGEN (Brasil). Se utilizó el equipo de extracción de ARN SV Total RNA Isolation Kit (Promega). Los ARNm fueron purificados utilizando Dynabeads (Dyna) y las bibliotecas se construyeron empleando el Marathon cDNA Amplification kit (BD Biosciences Clontech). Una vez aislados los ARNm se llevó a cabo la retrotranscripción para generar el ADNc doble hebra y luego se ligaron los adaptadores Marathon (Clontech) siguiendo las instrucciones del fabricante. Las bibliotecas fueron diluidas en una proporción 1/250 de solución reguladora Tricina (tricina 0,01M y EDTA 0,1 mM) y almacenadas a -20°C.

### **3.6- Amplificación de extremos 5' y 3' de ADNc**

La técnica de RACE (Rapid Amplification of cDNA Ends) se llevó a cabo según las indicaciones del Marathon cDNA Amplification kit (Clontech)(Chenchik *et al.*, 1996). Se diseñaron dos pares de cebadores anidados en base a las secuencias a extender, utilizando el programa Primer3 (<http://biotools.umassmed.edu/bioapps/primer3/>), y se estudiaron las posibilidades de generar dímeros de cebadores con el programa Oligo Analyzer 1.0. Los parámetros establecidos para el diseño incluyeron una longitud de los cebadores de entre 23 y 28 nt, un contenido de GC de 50 a 70 % y una temperatura de hibridización (T<sub>m</sub>) superior o igual a 67°C. En la Tabla 1 se detalla la secuencia de cada uno de los cebadores diseñados para los experimentos de RACE.

Las reacciones de PCR se llevaron a cabo en un volumen final de 50 µl con 2 µl de la dilución de la biblioteca Marathon, 1X de solución reguladora GoTaq (Promega), 0,2 mM de dNTPs, 0,2 µM de cebadores específicos del gen, 0,2 µM de cebadores específicos del adaptador de la biblioteca (AP1 o AP2) y 1,5 U de polimerasa Go Taq (Promega). Las condiciones de ciclado de PCR fueron las siguientes: 1 min a 94°C seguido de 30 ciclos de

30 s a 94°C y 4 min a 68°C (tanto la temperatura de hibridización del cebador como la de polimerización fueron de 68°C). Se realizaron dos rondas de amplificación con el cebador específico externo seguidas de al menos dos rondas de amplificación del cebador anidado. Además se incluyeron controles negativos los cuales contenían el cebador específico y AP1 pero no incluían el ADN molde; y controles positivos, donde se utilizaron dos cebadores específicos diseñados en sentido opuesto, de modo de corroborar la presencia en la biblioteca del fragmento a extender. Con la finalidad de obtener fragmentos de extensión 5' y 3' únicos, se optimizó la temperatura de hibridización de cada reacción variando la T<sub>m</sub> en el rango 68-72°C. Los productos de amplificación fueron separados por medio de electroforesis en geles de agarosa 1,5 % p/v en cubas horizontales LabNet (E-0350; 1007-10) a 60 mA con solución reguladora TAE (Tris-acetato 0,04M y EDTA 1 mM) y posteriormente teñidos con bromuro de etidio (EtBr) 1% v/v para permitir su visualización. Una vez obtenidas bandas discretas, se procedió a recuperar y aislar cada una de ellas, cortando con un bisturí el sector del gel que las contenía y utilizando el SV Wizard Gel and PCR Clean up System (Promega) para purificar el ADN. El producto purificado fue reamplificado en las mismas condiciones utilizadas para las reacciones originales de RACE. Estos productos de amplificación fueron clonados en vectores p-GEM T Easy (Promega), y las construcciones fueron electrotransferidas a cepas DH5α de *Escherichia coli* (ver debajo la técnica de clonado). Las bacterias transformadas fueron cultivadas en medio líquido toda la noche, los plásmidos fueron purificados utilizando SV Wizard Plus Minipreps (Promega) y secuenciados por MacroGen Inc. (Korea).

**Tabla 1: Secuencias de cebadores diseñados para 3' y 5' RACE**

Nombre del cebador*	Secuencia
A-2-upper1	5'-TGGAGCGGGCAAAGTAAGAAAGGAACGA-3'
A-2-upper2	5'-GCCTGTCGGCAAACGAAGTGGCATAC-3'
A-2-lower1	5'-CATTGATGGCGTGGTGGGCATGAGTTAT-3'
A-2-lower2	5'-TAGGCCGCTTACCGAACGCAACTCG-3'
A-7-upper1	5'-TCTATCCCAGACGCCAACCTGA-3'
A-7-upper2	5'-CTACCCCTCAGCCTGCCACAATG-3'
A-7-lower1	5'-GGTGCAATAGACAAGGGCGGTTACA-3'
A-7-lower2	5'-GATGGCTCCCACTACGCCACAGT-3'
A-11-upper1	5'-GAAAATGCTGCCCGGGTGGTTGTGTT-3'
A-11-upper2	5'-CCACTCGCATCGGCACCATCCCAACTTA-3'
A-11-lower1	5'-CCTTACGCCGGGCCGCATACGA-3'
A-11-lower2	5'-AACCAACCCGGGGCAGCATTTTCAGTGA-3'
A-13-upper1	5'-GCCCTTAGCCATTTACACTTATTTAG-3'
A-13-upper2	5'-GGAGTTCCACCACCTCTCTTTGTTG-3'
A-13-lower1	5'-TATCAGGGGTTTATGCATGTGGGGTTTC-3'

A-13-lower2	5'-TATCTTCACTCGACTTGTACCGTTTAGG-3'
A-17-upper1	5'-TTGTGAGTATAGAGTGCGGGAGTCG-3'
A-17-upper2	5'-CAGTTCTAGGATTGAGTAAAGTTC-3'
A-17-lower1	5'-AAGTAAGGATGAAATAACAAAGGTAAGA-3'
A-17-lower2	5'-GCATGCACAAGGTTCCCGTAAA-3'
A-22-upper1	5'-AGGGTTCAGCCGCCTCGAATCAGC-3'
A-22-upper2	5'-ATCAGCCGTCGCCAGAAATCATCAACAG-3'
A-22-lower1	5'-CCAACAGAAAGAGGGGCACTGGACACAA-3'
A-22-lower2	5'-CTGCTGGAGGGAGCTGGTAACTTCATCC-3'
A-38-upper1	5'-TGGACCTGTCGCGTGGTGATGATAA-3'
A-38-upper2	5'-CAGTGTATCGCGGACACGCAATC-3'
A-38-lower1	5'-CAGAGCTTACCAACAGCGTCACA-3'
A-38-lower2	5'-GACGATGTAACGCACGCGAAGAA-3'
A-43-upper1	5'-CGAGGCAATGCTCGACTCAACTCT-3'
A-43-upper2	5'-AACGCGTTGGATGCATAGCTTGAG-3'
A-43-lower1	5'-TCAGGTTTGGTATGGTGGGGGAAA-3'
A-43-lower2	5'-TACAGGCGGCCGGGAATTCATTAG-3'
A-49-upper1	5'-GGGCCTTATGGGAGATTTAGAA-3'
A-49-upper2	5'-CCACCCCATGAGCAATGTATGTGA-3'
A-49-lower1	5'-TGGCTTGGAATCCTCCATATCAA-3'
A-49-lower2	5'-ACATACATTGCTCATGGGGTGGT-3'
A-56-upper1	5'-GCAAACATGAACCAAAGGCAGCCACAC-3'
A-56-upper2	5'-CCAACTAGGGCAACGTGGCTGTCCAT-3'
A-56-lower1	5'-TCCCATCAAGTTCCGGCAAGGTTTACC-3'
A-56-lower2	5'-TCGGAGCCATGGAGACCAGCTTTTCATT-3'
A-58-upper1	5'-TCGCCCATCTATCTGGCCCCACGAC-3'
A-58-upper2	5'-ACGCTGATCCATCGTCCTACAAAAGAGA-3'
A-58-lower1	5'-TCGTCGCCCATGTTTAGAGAATGATTG-3'
A-58-lower2	5'-TCGCCCATGTTTAGAGAATGATTGGAGT-3'
A-70-upper1	5'-TTGCAGCGCAATGATCAGGCTAT-3'
A-70-upper2	5'-CGAAGTGGACAGCGGAATCAGAA-3'
A-70-lower1	5'-CGCCTTCTTAGCCCCTTGCTGT-3'
A-70-lower2	5'-CTGGCGCTGCCTTGAATTATCC-3'
A-77-upper1	5'-AGTGCAGGAGTTATCCCAGCATCAA-3'
A-77-upper2	5'-TACTGCTGGCACATTACCGCTACCT-3'
A-77-lower1	5'-CCAGCTCCAGCTGCTCAGTTAATTC-3'
A-77-lower2	5'-GCTGCCCAGCAGACAGATCAATAAA-3'
A-79-upper1	5'-CTCGATTCCGTTGGCGCCACTGAAGG-3'
A-79-upper2	5'-GAAAGGCCGCAGCAGCGTGTGAGAGC-3'
A-79-lower1	5'-TAGGGCCTTCCCCGTGCTTCCACTCCTC-3'
A-79-lower2	5'-GCCACCCTCCCGCACAGCTCTCACA-3'
A-103-upper1	5'-GGTCAGGCCCGATGGGGTATCCA-3'
A-103-upper2	5'-GTCGAATACAGCGGTCAGCGTGAAGTAA-3'
A-103-lower1	5'-TCAGGCCCGGTGCAACAAAGAG-3'
A-103-lower2	5'-AAACAGGCATATCTTCCGTCCCGCTAAT-3'
A-104-upper1	5'-CTGCGGTTGCTGTGGAAAAGTAGTGTGG-3'
A-104-upper2	5'-GTTCCCTCTTCTCCCCGCCAGCATCA-3'
A-104-lower1	5'-GGCAGACGCGGGTACGGTAGGAAGAGA-3'
A-104-lower2	5'-TTGATGCTGGCGGGGAGAAGAGG-3'
AP1	5'-CCATCCTAATACGACTCACTATAGGGC-3'
AP2	5'-ACTCACTATAGGGCTCGAGCGGC-3'

\* la nomenclatura de los cebadores incluye el nombre del gen, seguido del tipo de cebador (upper o lower) y el número (1 o 2). AP1 y AP2 son los cebadores comerciales que se unen a los adaptadores provistos por el kit de construcción de la biblioteca no clonada Marathon (Clontech).

### 3.7- Clonado y transformación

Los fragmentos obtenidos de amplificaciones por PCR se integraron a un vector apto para transformación de bacterias. Se utilizó en todos los casos el plásmido pGEM-T Easy (Promega). La reacción de ligación se llevó a cabo siguiendo las instrucciones del fabricante: en un volumen final de 5  $\mu$ l, se incorporaron 1,5  $\mu$ l de producto de PCR, 1X de solución reguladora Rapid Ligation (Promega), 25 ng de vector pGEM-T Easy (Promega) y 1,5 U de ADN ligasa T4. La incubación se realizó a 4 °C durante toda una noche.

La transformación de cepas DH5 $\alpha$  de *E. coli* con los plásmidos recombinantes se llevó a cabo mediante electroporación de alta eficiencia. Se utilizó el equipo BioRad Micropulse, respetando las recomendaciones descritas en las guías de aplicaciones del equipo. La preparación de células electrocompetentes comenzó con la inoculación de 500 ml de medio de cultivo líquido LB (peptona de carne 1% p/v; extracto de levadura 0,5% p/v y NaCl 1% p/v pH 7) con 5 ml de cultivo fresco de *E. coli*. Se incubó a 37°C con agitación durante aproximadamente 3 hs hasta alcanzar un índice OD<sub>600</sub> de 0,5-0,7. Luego, el cultivo fue transferido a tubos Falcon de 50 ml y centrifugado a 4°C a 3500 rpm durante 15 min. Se descartó el sobrenadante y las células fueron resuspendidas cuidadosamente en 500 ml de glicerol 10% v/v frío. Se procedió a centrifugar nuevamente a 3500 rpm a 4°C durante 15 min y se descartó el sobrenadante. Las células fueron resuspendidas esta vez en 250 ml de glicerol 10% v/v y luego centrifugadas en las mismas condiciones. Se descartó el sobrenadante y las células fueron disueltas en 20 ml de glicerol 10% v/v y nuevamente centrifugadas en las condiciones detalladas anteriormente. Finalmente el sobrenadante fue descartado y las células fueron resuspendidas en 2 ml de glicerol 10% v/v. La suspensión fue fraccionada en tubos de 1,5 ml y éstos conservados a -80°C.

Para la electroporación se incubaron en hielo 40  $\mu$ l de suspensión celular electrocompetente con 5  $\mu$ l de plásmido recombinante durante 1 min. Se transfirió la mezcla a la cubeta y se aplicó un micropulso. Se agregó 1 ml de medio SOC (extracto de levadura 0,5 % p/v; triptona de carne 2% p/v; NaCl 10 mM; KCl 2,5 mM; MgCl<sub>2</sub> 10 mM; MgSO<sub>4</sub> 10 mM; glucosa 20 mM) a la cubeta, se mezcló cuidadosamente con pipeta para integrar las células al medio y se transfirió a tubos Eppendorf de 1,5 ml. Se incubó a 37°C con agitación durante 1 h y se procedió a depositar 200  $\mu$ l del cultivo en medio selectivo: medio LB (peptona de carne 1% p/v; extracto de levadura 0,5 % p/v y NaCl 1 % p/v, pH 7), ampicilina 100  $\mu$ g/ml, X-Gal 60  $\mu$ g/ml y agar 1,5 % p/v.

Las placas se incubaron toda la noche a 37 °C. Las colonias transformadas se observaron de color blanco y las no transformadas de color azul. Para la confirmación de la identificación de colonias portadoras de los plásmidos recombinantes con el fragmento de interés se realizó una PCR de colonia (Colony PCR). Se seleccionaron colonias blancas y se utilizó un palillo estéril para inocular 20 µl de H<sub>2</sub>O<sub>d</sub> estéril en tubos de 0,2 ml. Esta suspensión se incubó en termociclador 10 min a 100°C para lograr una lisis celular. A continuación se llevó a cabo una reacción de PCR en un volumen final de 25 µl, con 2 µl de lisado, 0,5 µM de cebadores M13 (M13F: 5'-GTAAAACGACGGCCAGT-3' y M13R: 5'-GCGGATAACAATTCACACAGG-3'), 0,2 mM de dNTPs, 1X de solución reguladora GoTaq (Promega) y 1,5 U de polimerasa Go Taq (Promega). El programa de ciclado consistió en 3 min iniciales a 94°C seguido de 30 ciclos de 30 s a 94°C; 30 s a 55°C y 1 min a 72°C, con una fase de elongación final de 5 min a 72°C. Los productos de amplificación fueron visualizados por medio de electroforesis en geles de agarosa 1,5 % p/v en cubas horizontales a 60 mA con solución reguladora TAE (Tris-acetato 0,04M y EDTA 1mM) y posteriormente teñidos con BrEt 1% v/v.

Para el aislamiento de plásmidos recombinantes se utilizó el kit Wizard Plus SV Minipreps DNA Purification System (Promega), siguiendo las indicaciones establecidas por los fabricantes. Se seleccionaron las 4 colonias blancas por placa de estriado para ser cultivadas en 5 ml de medio selectivo LB líquido con ampicilina 100 µg/ml durante toda la noche a 37°C con agitación. Para la extracción de los plásmidos se trabajó con los reactivos, concentraciones y volúmenes indicados en el equipo comercial. Finalmente, los plásmidos fueron eluidos en 100 µl de agua estéril libre de nucleasa (Promega) y almacenados a -20°C. La calidad y cantidad de ADN plasmídico fueron evaluadas en un espectrofotómetro Perkin-Elmer Lambda Bio+. Los fragmentos clonados fueron secuenciados en MacroGen Inc. (Korea).

### **3.8- Análisis bioinformático de secuencias**

Las secuencias y cromatogramas provistos por MacroGen Inc. (Korea) fueron analizados con el programa Chromas 2.0. Se identificaron y eliminaron los sectores provenientes a ADN plasmídico utilizando el programa VecScreen del National Centre for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>). Las secuencias correspondientes a los cebadores específicos fueron detectadas utilizando el programa Sequencer 4.10.1. Se realizaron alineamientos entre los

productos de amplificación obtenidos por medios de 5' y 3' RACE implementando el programa Clustal W2 del European International Institute (EBI-EMBL) (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) y el programa BL2Seq del NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) para detectar las regiones solapadas y obtener consensos. Los análisis de similitud de secuencia con las bases de datos públicas se realizaron utilizando el programa BLAST 2.2.25 del sitio del NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Las búsquedas fueron confinadas a las especies de plantas verdes (green plants taxid: 33090) indagando en las siguientes bases de datos: colección de nucleótidos no redundantes de plantas (nr/nt); secuencias referencia de ARNm (refseq\_rna); y secuencias proteicas no redundantes (nr). Además se estudió la similitud en el genoma de *Arabidopsis thaliana* mediante la herramienta TAIR BLAST 2.2.8 del sitio Arabidopsis Information Resource (<http://www.arabidopsis.org/Blast/>). Asimismo se analizó la similitud con secuencias genómicas y génicas de las especies *Oryza sativa*, *Sorghum bicolor* y *Zea mays* por medio del programa BLASTn y tBLASTx del sitio Gramene (<http://www.gramene.org/multi/blastview>). Se investigó la presencia de fragmentos abiertos de lectura (ORF) utilizando el programa ORFinder del servidor NCBI ([www.ncbi.nlm.nih.gov/gorf/gorf.html](http://www.ncbi.nlm.nih.gov/gorf/gorf.html)). También, se indagaron las bases de datos de secuencias repetitivas de gramíneas (TIGR Gramineae Repeats Database) utilizando las herramientas de búsqueda disponibles en el sitio de J. Craig Venter Institute (<http://www.tigr.org/db.shtml>) (Ouyang and Buell, 2004) Los parámetros de búsqueda utilizados en cada uno de los análisis fueron los preestablecidos para cada programa y se siguieron los criterios de búsqueda y análisis publicados por Larrinua y Belmar (2008).

Por otro lado se analizaron las similitudes con microARN maduros y con precursores de miARN en las bases de datos del Instituto Sanger (<http://www.mirbase.org/search.shtml>). Se analizó la estabilidad de estructuras secundarias de tipo horquilla para los fragmentos que presentaron similitudes con precursores de miARN de plantas utilizando el programa mFold 2.3 disponible en el sitio de University at Albany, State University of New York (<http://mfold.rna.albany.edu/?q=mfold>). Para esto se estableció una temperatura de plegamiento 25°C, la cual es adecuada para las especies de plantas (Zuker, 2003). Se utilizó el programa miRU (Plant miRNA Potencial Target Finder, <http://bioinfo3.noble.org/miRNA/miRU.htm>) para identificar los posibles genes blanco de los precursores de miARN en el transcriptoma de *Oryza sativa* (Zhang *et al.*, 2005). La estabilidad de la hibridación entre el posible precursor de miARN y su target

putativo fue determinada en función a su valor de  $\Delta G$  utilizando el programa RNAhybrid disponible en el sitio de la Universidad de Bielefeld (<http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/submission.html>). Se consideraron las indicaciones publicadas por Schwab et al (2005) para realizar las predicciones de los genes blanco putativos. Estas incluyen: i) el apareamiento perfecto en el extremo 5' del miARN, con una sola base desapareada entre los nucleótidos 2-12 pero no en el sitio de corte (entre los nucleótidos 10 y 11) y ii) hasta 3 bases desapareadas en el extremo 3' del miARN pero no más de dos contiguas. Se considera además, como control de estabilidad de hibridización, que la energía libre calculada entre el miARN y su blanco debe ser igual o superior al 72% de la energía libre de complementariedad perfecta.

La presencia de sectores similares a secuencias proteicas o familias de dominios proteicos fue analizada utilizando los programas InterProScan Sequence Search del sitio EBI-EMBL (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>), Conserved Domains de NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) y ScanProsite del sitio EXPASy y SIB (Swiss Institute of Bioinformatics) (<http://ca.expasy.org/tools/scanprosite/>).

Los sectores que presentaron similitud con elementos repetitivos aislados por medio del experimento de Display Diferencial Específico de Transposones a partir de N22, fueron analizados en la plataforma Phylogeny (<http://www.phylogeny.fr/>). Se seleccionó el programa de alineamiento T-Coffe y el programa de construcción de árboles filogenéticos PHYLYP 3.66 (Fast/Dist + BioNJ/Neighbor) estableciendo un número de 1000 iteraciones (bootstraps). Para la visualización del árbol construido se seleccionó el programa TreeDyn de la plataforma Phylogeny.

### 3.9- Transcripción reversa

Se realizó la transcripción reversa del ARNm de *P. notatum* utilizando la enzima Superscript II Reverse Transcriptase (Invitrogen) y siguiendo las instrucciones del fabricante. La reacción se llevó a cabo en un volumen final de 20  $\mu\text{l}$ . Se incorporaron, 10 ng de oligo(dT)<sub>12-18</sub>, 1  $\mu\text{g}$  de ARN total, 1  $\mu\text{l}$  de dNTP (10 mM) y H<sub>2</sub>O<sub>d</sub> estéril hasta llegar a un volumen de 12  $\mu\text{l}$ . Se incubó a 65°C por 5 min y la mezcla se mantuvo en hielo. A continuación, se agregaron 4  $\mu\text{l}$  de solución reguladora 5X First-Strand (Invitrogen) y 2  $\mu\text{l}$  de DTT (0,1 M) y luego de homogeneizar la mezcla se incubó a 42°C durante 2 min. Se adicionaron 200 U de Superscript II RT (Invitrogen), mezclando suavemente con pipeta, y

se incubó 60 min a 42°C. La reacción de retrotranscripción se inactivó incubando durante 15 min a 70°C. El ADNc sintetizado fue conservado a -20°C.

### 3.10- PCR en tiempo real

Los experimentos de PCR en tiempo real se llevaron a cabo, tanto para precisar los valores de expresión relativa de los transcritos de interés en cuatro estadios del desarrollo reproductivo de *P. notatum* entre plantas sexuales y apomícticas, como también para determinar el número de copia de dichas secuencias en el genoma. En ambos casos los genotipos utilizados fueron Q4117 (apomíctico) y Q4188 (sexual).

Para el análisis de la expresión génica se utilizaron muestras de ARN extraídas de espiguillas colectadas en estadio pre-meiosis (0), pre-meiosis tardía/meiosis (I/II), post-meiosis (IV/V) y anthesis (VI) según el calendario reproductivo (Laspina *et al.*, 2008). Para la extracción del ARN se utilizó el equipo comercial SV Total RNA Isolation Kit (Promega) y se llevó a cabo la retrotranscripción a partir de 1 µg de ARN total con la enzima Superscript II Reverse Transcriptase (Invitrogen), como se indica en el apartado anterior. El ADNc sintetizado fue diluido en 4 partes de H<sub>2</sub>O<sub>d</sub> estéril para conformar la solución de trabajo. La PCR en tiempo real se llevó a cabo en un volumen de 15 µl incluyendo los siguientes reactivos: 0,1 µM de cebadores específicos, 1 X de Mezcla Real (Biodynamics) y 20 ng de ADNc de las soluciones de trabajo. Se utilizaron dos controles de igual expresión: el gen de la Glucosa 6-Fosfato Deshidrogenasa (G6PdH) y el de la subunidad β de la Tubulina (Tub). Estos genes constitutivos fueron reportados por otros grupos de trabajo como genes de expresión estable entre plantas sexuales y apomícticas en especies de gramíneas (Albertini *et al.* 2004, Duarte-Silveira *et al.* 2009).

Los cebadores específicos fueron diseñados utilizando el programa Primer3 (<http://biotools.umassmed.edu/bioapps/primer3/>) y sintetizados por IDT (<http://www.idtdna.com/home/home.aspx>). Las secuencias de los cebadores utilizados en estos experimentos se detallan en la Tabla 2. Las reacciones de PCR en tiempo real se realizaron por triplicado en todos los casos y utilizando dos repeticiones biológicas para cada una de las determinaciones. Además se incorporaron controles negativos a los que no se les agregó el ADNc. Las amplificaciones se llevaron a cabo en el equipo Rotor-Gene Q Thermocycler (Quiagen) utilizando la siguiente programación: 2 min a 94°C, 45 ciclos de 15 s a 94°C, 30 s a 57-65°C y 40 s a 72°C. La curva de hibridización del producto de amplificación se construyó al final del ciclado de la siguiente manera: ciclos de 10 s desde

72°C a 95°C aumentando la temperatura 0,2°C luego del ciclo 2. Los datos y reportes de cada amplificación fueron extraídos utilizando el programa Rotor-Gene Q Series Software 1.7 y la cuantificación relativa se realizó considerando los valores de “take off” (despegue) y de eficiencia para cada reacción individualmente. El análisis estadístico se desarrolló con el programa REST-RG (Relative Expression Software Tool V 2.0.7 para Rotor Gene, Corvett Life Sciences) utilizando simultáneamente ambos controles para normalizar cada valor de expresión relativa.

Para la determinación del número de copias genómico, las reacciones de PCR en tiempo real se llevaron a cabo en un volumen final de 25 µl. Se utilizaron 0,2 µM de cebadores específicos (Tabla 2), 1X de Mezcla Real (Biodynamics) y la masa correspondiente calculada de plásmido recombinante o de ADN genómico. Las curvas de calibración fueron construidas incluyendo 5, 50, 500 y 5000 copias de plásmido. Se realizaron cuatro diluciones de ADN genómico de los genotipos Q4117 y Q4188:  $30 \times 10^{-9}$  g/µl,  $30 \times 10^{-11}$  g/µl,  $30 \times 10^{-12}$  g/µl y  $30 \times 10^{-13}$  g/µl. Tanto las reacciones de amplificación de la curva de calibración como las de las diluciones del ADN genómico se llevaron a cabo por triplicado, en este caso también se incluyeron controles negativos sin templado. Las amplificaciones se realizaron en el equipo Rotor-Gene Q Thermocycler (Quiagen) utilizando la siguiente programación: 2 min a 94°C, 45 ciclos de 15 s a 94°C, 30 s a 62-65°C y 40 s a 72°C. La curva de hibridización del producto de amplificación se construyó al final del ciclado de la siguiente manera: ciclos de 10 s desde 72°C a 95°C aumentando la temperatura 0,2°C luego del ciclo 2. Los datos de cada amplificación fueron extraídos utilizando el programa Rotor-Gene Q Series Software 1.7 y la cuantificación absoluta se realizó considerando los valores de “take off” y de eficiencia para cada reacción individualmente. Sólo fueron considerados los puntos abarcados en el rango lineal de las curvas de calibración. Se incluyó un gen de referencia de igual número de copias entre el genotipo sexual y el apomítico, N20, el cual codifica al un ortólogo del gen LORELEI de *A. thaliana* (Felitti *et al.*, 2011). La familia LORELEI posee 4 miembros en arábido. Finalmente se construyeron gráficos comparativos con los valores de cuantificación absoluta para los dos genotipos con sus respectivos desvíos estándar.

**Tabla 2: Secuencias de cebadores utilizados en experimentos PCR en tiempo real**

Nombre del cebador*	Secuencia
A-11-upper1	5'-GAAAATGCTGCCCCGGGTGGTTGTGTT-3'
A-11-lower1	5'-CCTTACGCCGGGCCGCCATACGA-3'
A-13-upper2	5'-GGAGTTCCACCACCCTCTCTTTGTTT-3'
A-13-lower2	5'-TATCTTCACTCGACTTGTACCGTTTAGG-3'
A-17-upper1	5'-TTGTGAGTATAGAGTGCGGGAGTCG-3'
A-17-lower2	5'-GCATGCACAAGGTTCCCGTAAA-3'
A-22-upper1	5'-AGGGTTCAGCCGCCTCGAATCAGC-3'
A-22-lower1	5'-CCAACAGAAAGAGGGGCACTGGACACAA-3'
Tar-N11 upper	5'-GAAAGGGTATCAGGACGCTCTCC-3'
Tar-N11 lower	5'-GCACCCTGATACTGTCCACCTCC-3'
Tar-N13 upper	5'-CAAGCCAAAGCACACACACCTC-3'
Tar-N13 low3	5'-GATCAGGTTACAGCCACATCCT-3'
N20 upper	5'-GTGTGGCAGTAGCGTTGTTGTTCC-3'
N20 lower	5'-TACGGCCGACCCACCATTAGCA-3'
G6PdH upper	5'-TGAATCTAGTCCATCCGCTTG-3'
G6PdH lower	5'-TCATCAGGCAGGGAAGCTA-3'
Tub upper	5'-GTGGAGTGGATCCCCAACAA-3'
Tub lower	5'-AAAGCCTTCCTCCTGAACATGG-3'

\* la nomenclatura de los cebadores incluye el nombre del gen (11, 13, 17, 22, TarN11, TarN13, N20, G6PdH y Tub) seguido del tipo de cebador (upper o lower) y el número (1 o 2).

### 3.11- Generación de marcadores moleculares de tipo SCAR (Sequence Characterized Amplified Regions)

Para la generación de marcadores moleculares SCAR se realizaron amplificaciones sobre el ADN de *P. notatum*. Inicialmente se utilizaron los padres de la población de mapeo (Q4117 y Q4188) para detectar bandas polimórficas útiles con cada una de las combinaciones posibles de cebadores específicos detallados en la Tabla 1. Las amplificaciones se llevaron a cabo en un volumen final de 12,5 µl con 30 ng de ADN genómico, 1X de solución reguladora Go Taq (Promega), 0,1 mM de dNTPs, 0,1 µM de cebadores específicos y 0,75 U de polimerasa Go Taq (Promega). El ciclado consistió en una desnaturalización inicial de 5 min a 94°C, seguido de 40 ciclos de 30 s a 94°C, 1 min a 57-65°C y 1:30 min a 72°C, con un último paso de elongación a 72°C durante 5 min.

Los productos de las amplificaciones se analizaron en geles de secuencia de poliacrilamida 5 % p/v en condiciones desnaturalizantes (acrilamida/bisacrilamida 19:1, urea 7,5 M). Se indujo la polimerización utilizando TEMED (6 x 10<sup>-5</sup> % v/v) y persulfato de amonio (0,06 % p/v). La solución catalizada se vertió entre vidrios tratados con Bind Silano (solución de trabajo: Bind Silano 0,3% v/v, EtOH 95% y ácido acético 0,5%) y Repel Silano, respectivamente. La electroforesis se realizó en una cuba vertical Hoefer SQ3 Sequencer utilizando TBE 1X (Tris 0,108 % p/v; ácido bórico 0,55 % p/v y EDTA

2mM) como solución reguladora. Se realizó un precalentamiento del gel durante 1 h a 60W de potencia y se sembraron las muestras previamente desnaturalizadas por incubación de 5 min a 94°C en solución STR 1X (NaOH 10 mM; formamida 95 % v/v; azul de bromofenol 0,05 % p/v y xilencianol 0,05 % p/v). El tiempo de electroforesis, a 60W de potencia, dependió del tamaño de los fragmentos, variando entre 45 a 90 min y luego se procedió a la detección de las bandas siguiendo el protocolo de tinción con plata de BioRad. Brevemente, en la primera fase de la detección, la fijación, la placa de vidrio con el gel adherido se sumergió en ácido acético glacial 10 % v/v. Luego de realizar tres lavados de 3 min con H<sub>2</sub>O<sub>d</sub> se llevó a cabo la tinción sumergiendo la placa en nitrato de plata 0,1 % p/v y formaldehído 0,055 % v/v durante 30 min. Para el revelado se realizó un lavado de 10 s en H<sub>2</sub>O<sub>d</sub> y luego se incubó con la solución de revelado, previamente refrigerada (carbonato de sodio 3 % p/v; formaldehído 0,055 % v/v y tiosulfato de sodio 0,002 % p/v) hasta la aparición de bandas (aproximadamente 10 min, dependiendo de la temperatura de la solución de revelado). La reacción de revelado fue detenida por la adición de un (1) volumen de solución de fijación y luego se transfirió la placa con el gel a una cubeta con H<sub>2</sub>O<sub>d</sub> durante al menos 20 min para eliminar los restos de las soluciones utilizadas en la detección. Los patrones de bandas observados en los geles fueron digitalizados empleando el equipo multifunción EPSON Stylus SX218.

Una vez detectados patrones polimórficos entre los padres de la población, se analizaron 10 individuos F<sub>1</sub>, 5 apomícticos y 5 sexuales con la finalidad de determinar los posibles patrones de segregación de las bandas. Los análisis se llevaron a cabo de la manera descrita en el párrafo anterior. Finalmente se amplió el análisis de la segregación de las bandas polimórficas a 62 individuos de la población de mapeo. La presencia / ausencia de bandas fue determinada visualmente y digitalizada en archivos específicos de los programas de mapeo.

### 3.12- Mapeo genético

Los datos colectados a partir de los marcadores moleculares de tipo SCAR fueron analizados por medio de un test  $\chi^2$  para determinar bondad de ajuste ( $p \leq 0,05$ ) entre el número esperado de genotipos para cada proporción de segregación y el valor observado de los mismos (Stein *et al.*, 2007). Los datos provenientes de los fragmentos polimórficos entre los padres amplificadas en dosis simple (SDAF), que mostraron una relación de segregación de 1:1, fueron incorporados a las matrices del mapa genético de *P. notatum*

(Stein *et al.*, 2007). Se calculó la distancia entre marcadores utilizando las unidades de mapeo del tipo centimorgan (cM) derivadas de la función de mapeo publicada por Kosambi (Kosambi, 1944). El análisis de ligamiento de los SDAFs se llevó a cabo utilizando el programa MAPMAKER/EXP 3.0 (Lander *et al.*, 1987) considerando un valor de mínimo de LOD de 6.0-2.0 y una frecuencia de recombinación de  $\theta = 0.30$  (Wu *et al.*, 1992).

### 3.13- Hibridización *in situ* de tejidos reproductivos

#### 3.13.1- Marcación de sondas con digoxigenina

Para la marcación de las sondas de ARN se utilizó el kit DIG RNA Labelling Kit Sp6/T7 (Roche), siguiendo las especificaciones del fabricante. Inicialmente se digirió el plásmido recombinante que contenía el inserto a marcar con dos enzimas de restricción, en forma alternativa. Se incubaron, por un lado, 10  $\mu\text{g}$  de plásmido con 1 U de la enzima de restricción *Sall*, solución reguladora de actividad 1X (Roche) y 10  $\mu\text{g}/\mu\text{l}$  de BSA en un volumen final de 20  $\mu\text{l}$ ; y, en otra fracción, la misma masa de plásmido con 1U de *NcoI*, solución reguladora de actividad 1X (Roche) y 10  $\mu\text{g}/\mu\text{l}$  de BSA en un volumen final de 20  $\mu\text{l}$ . Se realizó una incubación de 2 h y 30 min a 37°C y luego se agregaron 180  $\mu\text{l}$  de H<sub>2</sub>O<sub>d</sub> tratada con DEPC. Se llevó a cabo la purificación del plásmido mediante una extracción con fenol/cloroformo (1:1), seguida de una con cloroformo únicamente. Luego se precipitó el ADN plasmídico en 5% v/v de NaCl 5 M y 2,5 volúmenes de etanol absoluto y se incubó a -20 °C durante toda una noche. Seguidamente, se centrifugó a 12000 rpm durante 20 min y se descartó el sobrenadante. Al precipitado se le agregó 0,5 ml de solución de lavado (etanol 70% v/v) y se centrifugó a 12000 rpm por 20 min. El precipitado fue deshidratado a 37 °C en estufa durante al menos 1 hora y se resuspendió en 20  $\mu\text{l}$  de H<sub>2</sub>O<sub>d</sub> estéril tratada con DEPC.

Una vez obtenidas las digestiones con cada una de las enzimas, se prepararon dos mezclas de marcación: la primera de ellas se conformó en un volumen final de 24  $\mu\text{l}$  con 1  $\mu\text{g}$  de plásmido digerido con *NcoI*, 1 X de Mezcla de dNTP Marcado (Roche), 1 X solución reguladora de transcripción (Roche), 20 U de inhibidor de RNasa y 40 U de ARN polimerasa SP6; la segunda se conformó en un volumen final de 24  $\mu\text{l}$  con 1  $\mu\text{g}$  de plásmido digerido con *Sall*, 1 X de mezcla de dNTP marcado (Roche), 1 X solución

reguladora de transcripción (Roche), 20 U de inhibidor de RNasa y 40 U de ARN polimerasa T7. Ambas mezclas fueron incubadas a 37°C durante 2 h y luego tratadas con 20 U de ADNasa a 37°C durante 15 min. Finalmente la reacción de degradación de ADN se detuvo adicionando EDTA hasta alcanzar una concentración de 15 mM. Debido a que es necesario que las sondas tengan una longitud de 150-200 pb para aumentar la eficiencia de la hibridización, se realizó la fragmentación de las mismas por medio de una reacción de hidrólisis alcalina. En un volumen final de 100 µl se incorporaron a la mezcla de transcripción, 60 mM de Na<sub>2</sub>CO<sub>3</sub> y 40 mM de NaHCO<sub>3</sub> y se incubó a 60°C durante el tiempo determinado según la siguiente ecuación:  $T = (L_o - L_f) / (k \times L_o \times L_f)$ , donde T es el tiempo de hidrólisis a 60°C, L<sub>o</sub> la longitud inicial del fragmento (en Kpb), L<sub>f</sub> la longitud deseada del fragmento (en Kpb), y k la constante de hidrólisis (0,11 Kb/min). Finalmente se procedió a purificar las sondas marcadas mediante precipitación del ARN con 5% v/v de NaCl 5 M y 2,5 volúmenes de etanol absoluto. La sonda hidrolizada se controló por electroforesis en geles de agarosa 1,5 % p/v.

### 3.13.2- Hibridización de tejidos reproductivos

Los experimentos de hibridización fueron realizados en colaboración con el Dr. Guillermo Seijo y la Dra. Ana María Gonzalez del Instituto de Botánica del Nordeste (IBONNE-CONICET) utilizando el protocolo descrito por Dusi (2001) e incluyendo algunas modificaciones menores. Para la preparación de los portaobjetos se trataron los vidrios con solución de lavado (Extrán 5% p/v) durante 4 h y luego se enjuagaron con agua corriente. Se incubaron durante 4 h en HCl 10% v/v y se los lavó en H<sub>2</sub>O<sub>d</sub>. Los vidrios tratados se dejaron secar a temperatura ambiente y luego fueron autoclavados. Una vez esterilizados se los sumergió en una solución de fijación compuesta por 100 µg/ml de poli-L- lisina y 10 mM de Tris HCl a pH=8 durante 10 min y se los secó a temperatura ambiente durante 24 h.

Se colectaron inflorescencias inmaduras de los genotipos Q4117 y C4-4x de *P. notatum* en estadio premeiosis tardía/meiosis (I/II). El material fue disectado en una placa de Petri en H<sub>2</sub>O<sub>d</sub> tratada con DEPC y se sumergió en la solución de fijación (paraformamida 4 % v/v; glutaraldehído 0,25 % v/v; solución reguladora fosfato 0,01 M a pH= 7) durante 3-6 hs. Posteriormente el tejido fue deshidratado removiendo el fijador mediante incubaciones de 30 min en soluciones de etanol de concentración creciente (10%, 30%, 50%, 70%, 90% y 98%). La última incubación, en etanol absoluto, se realizó por un

período de 1 h. El material deshidratado fue sumergido en una serie de diluciones de xilol / etanol (xilol: etanol 1:3, 1:1 y 3:1) incubándose 30 min con cada solución, seguida de una incubación final de 60 min en xilol 100%. A continuación el material fue tratado en una mezcla de parafina (58-60°C) y xilol en una proporción 2:1, durante 4-16 h a temperatura ambiente. Este paso fue reiterado variando la temperatura de incubación a 42°C y luego a 60°C. La solución parafina: xilol fue suplantada por parafina y se incubó durante 2 días sustituyendo la parafina periódicamente. Para el montaje, se utilizaron moldes con parafina derretida donde se incluyeron las muestras y se dejaron reposar 30 min a temperaturas descendientes (60°C, 58°C, temperatura ambiente y 4°C). Los bloques fueron seccionados cuidadosamente, en ángulo de 6° con 7 µm de espesor y fueron depositados en papel absorbente. Luego los cortes fueron trasladados en los portaobjetos tratados con poli-L-lisina, agregando 200 µl de H<sub>2</sub>O<sub>d</sub> tratada con DEPC, y secados en una placa precalentada a 42°C. Se procedió a retirar la parafina del tejido fijado mediante la incubación con xilol durante 10 min, xilol: etanol (1:1) durante 10 min, etanol 100% 10 min, etanol:agua (2:1) 10 min y finalmente dos incubaciones en H<sub>2</sub>O<sub>d</sub> tratada con DEPC durante 5 min.

Se realizó la prehibridización agregando a cada uno de los cortes 100 µl de solución proteasa (proteínasa K 1 µg/ml y TrisHCl 0,05M a pH=7,5) e incubando durante 10 min en cámara húmeda. Se realizaron dos lavados con H<sub>2</sub>O<sub>d</sub> tratada con DEPC durante 10 min. Para conformar la solución de hibridación se mezclaron 6 µl de ARN total (600 ng/µl) y 6 µl de sonda marcada en un volumen final de 15 µl y la solución fue desnaturalizada a 85°C durante 5 min. Se adicionaron 85 µl de solución reguladora de hibridización (TrisHCl 10 mM a pH=7,5; NaCl 300 mM; formamida deionizada 50% v/v; EDTA 10 mM; Denhart 1X; sulfato de dextrán 10 % p/v). Los 100 µl de mezcla de hibridización se depositaron sobre el tejido fijado y fueron incubados a 42°C en cámara húmeda y oscuridad durante toda la noche. Posteriormente se realizaron lavados en serie de 30 min a 37°C, utilizando SSC 4X, SSC 2X, SSC 1X y SSC 0,5X (composición SSC 20X: NaCl 3M y citrato de sodio 0,3M en agua bidestilada, pH= 7,2).

La detección inmunológica se llevó a cabo siguiendo las instrucciones del kit DIG RNA Labelling Kit Sp6/T7 (Roche). Los portaobjetos fueron incubados en solución reguladora 1 (0,1M TrisHCl pH=7,5 y 0,15M NaCl) durante 5 min y luego en solución reguladora de bloqueo (BSA 2% p/v en solución reguladora 1) durante 30 min. Se mantuvo 5 min en solución reguladora 1 antes de agregar los anticuerpos anti-dig AP (fragmentos

Fab) diluidos en solución reguladora 1 en una proporción 1:1000. La incubación con anticuerpos se realizó a temperatura ambiente durante 1 h. Se llevaron a cabo dos lavados subsiguientes de 15 min en solución reguladora 1 y luego se incubó 5 min en solución reguladora 3 (0,1M TrisHCl pH=7,5; 0,1 M NaCl; 0,05M MgCl<sub>2</sub> pH=9,5). Los portaobjetos fueron sumergidos en la solución de revelado (NBT 4,5 µl; BCIP 4,5 µl y 1 ml de solución reguladora 3) en oscuridad durante 30 min. Por último, la reacción de revelado fue detenida adicionando solución de terminación (TrisHCl 0,01M pH=8; EDTA 1 mM), se incubó 10 min y se realizaron dos lavados finales de 5 min con H<sub>2</sub>O<sub>d</sub> tratada con DEPC. Las muestras hibridizadas fueron evaluadas en microscopio óptico LEICA M205.

### 3.14- Display Diferencial Específico de Retrotransposones

Los experimentos de Display Diferencial fueron basados en el protocolo general informado por Liang y Pardee (1992), aplicando algunas modificaciones, como se indica en Laspina *et al.* (2008). Se utilizaron oligonucleótidos anclados de estructura general 3'X(AgC)T<sub>12</sub>5', donde X representa a los nucleótidos A, C, G o T (DDT1, DDT2, DDT3 y DDT4 respectivamente) para realizar las retrotranscripciones de ARN total extraído de inflorescencias inmaduras en estadio pre meiótico tardío / meiótico (I/II) (Laspina *et al.*, 2008). Para las reacciones de amplificación se utilizaron cada uno de los cebadores poliT anclados y cebadores específicos diseñados sobre N22 (A-22up2short: 5'-CTGGTAACTTCATCC-3'; A-22low2short 5'-CCAGAAATCATCAAC-3') en lugar de los decámeros al azar del protocolo original. De este modo, cada uno de los cebadores específicos fue combinado con cada cebador poliT anclado. La reacción de amplificación se llevó a cabo en un volumen final de 25 µl agregando 2,5 µl de una dilución 1/20 de ADNc generado por retrotranscripción, 2,5 µM de cebadores poliT anclados, 0,7 µM de cebadores específicos, 0,05 mM de dNTPs, 1,5 mM de MgCl<sub>2</sub>, 1X de solución reguladora GoTaq (Promega) y 2 U de polimerasa Go Taq (Promega). Todas las muestras, inclusive los controles negativos, fueron amplificadas por duplicado. El programa de ciclado consistió en una desnaturalización inicial de 3 min a 94°C seguido de 40 ciclos de 20 s a 94°C, 20 s a 38°C y 30 s a 72°C y finalmente una fase de elongación de 5 min a 72°C. Las muestras fueron incubadas durante 5 min a 95°C con 1X de solución de siembra desnaturalizante (formaldehído 95% v/v; azul de bromofenol) y analizadas en geles de poliacrilamida (5% p/v) revelados por tinción con plata. Se contabilizaron sólo las bandas diferenciales duplicadas localizadas en la región media del gel.

Las bandas seleccionadas para ser secuenciadas se cortaron directamente del gel de poliacrilamida previamente humedecido con H<sub>2</sub>O<sub>d</sub>. El fragmento de gel se procesó en porciones de mínimo tamaño con un bisturí y luego se incubó en 50 µl de solución reguladora de elución (acetato de amonio 0,5 M y EDTA 1 mM, pH=8) a 37°C durante toda una noche. A continuación, las muestras fueron centrifugadas a 12000 rpm durante 10 min y se recuperó cuidadosamente el sobrenadante. El ADN se precipitó agregando NaCl 5 M 5% v/v y 2,5 volúmenes de etanol absoluto. Los tubos se incubaron a -20 °C durante toda una noche. Seguidamente, se centrifugaron a 12000 rpm durante 20 min y se descartó el sobrenadante. Al precipitado se le agregó 0,5 ml de solución de lavado (etanol 70% v/v) y se centrifugó a 12000 rpm por 20 min. El precipitado fue deshidratado en estufa a 40 °C durante al menos 1 hora y luego se resuspendió en 15 µl de a estéril H<sub>2</sub>O<sub>d</sub>.

Las bandas recuperadas fueron reamplificadas y clonadas en el vector pGEM-T Easy y transferidas por electrotransformación a *E. coli* DH5α. Las células transformadas fueron seleccionadas en medio selectivo sólido (medio LB; ampicilina 100 µg/ml; X-Gal 60 µg/ml; y agar 1,5 % p/v). La presencia de los fragmentos de interés se confirmó mediante PCR directa de colonia y los plásmidos recombinantes fueron posteriormente aislados utilizando el equipo comercial Wizard Plus SV Minipreps DNA Purification System (Promega).

Los fragmentos clonados fueron secuenciados en Macrogen Inc. (Korea) y sus secuencias comparadas con las bases de datos públicas de NCBI (green plants non-redundant database) y TIGR (Gramineae plant repeat database).

### 3.15- Detección de secuencias blanco por 5' RACE

Con la finalidad de poder aislar las secuencias de los posibles genes blanco clivadas en el sitio correspondiente al miRNA maduro putativo, y así validar su naturaleza como blanco regulatorio del miARN, se llevó a cabo un experimento de Amplificación Rápida del Extremo 5' (5' RACE) de estas secuencias. Se diseñaron tres cebadores anidados sobre la secuencia blanco utilizando el programa Primer3 (<http://biotools.umassmed.edu/bioapps/primer3/>). Los parámetros establecidos para el diseño incluyeron una longitud de los cebadores de entre 23 y 28 nt, contenido de GC de 50 a 70 % y la temperatura de hibridización (T<sub>m</sub>) superior o igual a 67°C. Las posibilidades de formación de dímeros de cebadores con AP1 y AP2 fueron analizadas con el programa Oligo Analyzer 1.0. En la

Tabla 3 se detalla la secuencia de cada uno de los cebadores diseñados para amplificar las secuencias blanco.

Las reacciones de PCR se llevaron a cabo en un volumen final de 50  $\mu$ l con 2  $\mu$ l de la dilución de la biblioteca Marathon, 1X de solución reguladora GoTaq (Promega), 0,2 mM de dNTPs, 0,2  $\mu$ M de cebadores específicos del blanco, 0,2  $\mu$ M de cebadores específicos del adaptador de la biblioteca (AP1 y AP2, ver Tabla 1) y 1,5 U de polimerasa Go Taq (Promega). Las condiciones de ciclado de la PCR fueron las siguientes: 1 min a 94°C seguido de 30 ciclos de 30 s a 94°C y 4 min a 68°C. Se realizaron dos rondas de amplificación con el cebador específico externo, seguidas de al menos dos rondas de amplificación del segundo cebador anidado y dos rondas de amplificación con el tercer cebador anidado. Además se incluyeron controles negativos (sin ADNc molde). Los productos de amplificación fueron visualizados por medio de electroforesis en geles de poliacrilamida (5% p/v) revelados por tinción con plata. Las bandas del tamaño esperado fueron cortadas directamente del gel de poliacrilamida y extraídas en solución reguladora de elución (acetato de amonio 0,5 M y EDTA 1 mM, pH=8). Posteriormente fueron precipitadas en 5% v/v de NaCl 5 M y 2,5 volúmenes de etanol absoluto. El producto purificado fue reamplificado en las mismas condiciones que las reacciones de amplificación de 5'RACE. Estos productos fueron clonados en vectores p-GEM T Easy (Promega) y electrotransferidos a cepas DH5 $\alpha$  de *E. coli*. Los plásmidos recombinantes fueron purificados utilizando el kit SV Wizard Plus Minipreps (Promega) y secuenciados en Macrogen Inc. (Korea).

**Tabla 3: Secuencias de cebadores utilizados en experimentos de detección de secuencias blanco 5' RACE**

<b>Nombre del cebador</b>	<b>Secuencia</b>
Tar-N11 low1	5'-TCCCGGCTACACCGCTCTGCCTGCTC-3'
Tar-N11 low2	5'-ACTTGACCGCCTCCTCATCTCGCTCAGC-3'
Tar-N11 low3	5'-GATGCCGGATGATTCTCTGTAGCACGAC-3'
Tar-N13 low1	5'-GCCGCACAGGTTTTAGGAAGAAC-3'
Tar-N13 low2	5'-CCTTCCTCCAGGCACCTATTGAT-3'
Tar-N13 low3	5'-GATCAGGTTACAGCCACATCCT-3'

## 4- RESULTADOS

### 4.1- Búsquedas bioinformáticas iniciales

Como un paso preliminar a los intentos de caracterización de función de las secuencias objeto de estudio en esta tesis, se ampliaron las búsquedas bioinformáticas que habían sido realizadas por Laspina *et al.* (2008) y Martelotto *et al.* (2005). Las 20 secuencias diferenciales entre plantas sexuales y apomícticas que no presentaron anotación funcional y 2 secuencias que habían presentado homologías con genes bacterianos publicadas por Laspina *et al.* 2008 se detallan en el Anexo I (Sección A). Las secuencias originales diferenciales entre plantas tetraploides y diploides de *P. notatum*, publicadas por Martelotto *et al.* 2005 y 16 aisladas posteriormente (Silvina Pessino, comunicación personal) se muestran en el mismo Anexo I (Sección A).

Los veintidos (22) transcritos diferenciales por modo de reproducción (N2, N7, N11, N13, N17, N22, N38, N43, N49, N56, N58, N70, N77, N82, N86, N87, N99, N103, N104, N106, N116, N119) y los treinta y seis (36) transcritos diferenciales por nivel de ploidía (L14, L20, L24, L29, L45, L47, L48, L49, L53, L55, L61, L66, L72, L73, L75, L76, L85, L86, L87, L90, L94, L95, L774, L784, L794, L802, L812, L832, L834-1, L834-3, L844-2, L852-1, L852-2A, L852-2B, L854, L884) fueron analizados con herramientas bioinformáticas para detectar similitud con otras secuencias de las bases de datos públicas GenBank, EMBL y DDBJ. Todas las secuencias fueron depuradas y luego comparadas contra las bases de datos. Se estudió la presencia de restos de secuencias de vectores de transformación que pudiesen estar contaminando las búsquedas con el programa VecScreen del NCBI (<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>). Se trabajó con los mecanismos de búsqueda de los servidores NCBI del National Centre for Biotechnology Information (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), TAIR del Arabidopsis Information Resource (<http://www.arabidopsis.org/Blast/>) y del sitio específico de gramíneas: Gramene (<http://www.gramene.org/multi/blastview>).

A través de estos estudios bioinformáticos iniciales se pudieron detectar nuevas similitudes con proteínas de especies relacionadas de plantas. En el caso de los transcritos diferenciales por el modo de reproducción, se detectó un alto nivel de similitud para la secuencia N116 con una proteína hipotética de maíz (gb|EU975538.1|,  $eV= 3e-10$ ) (Tabla 4). Por otro lado, trabajando con el grupo de fragmentos diferenciales por nivel de ploidía, se detectaron similitudes significativas con secuencias proteicas de especies relacionadas

para 15 fragmentos. Este importante aumento de la capacidad de detección de homologías con genes codificantes para proteínas respecto al trabajo previo de Martelotto *et al.* (2005) se debe posiblemente a la mayor cantidad de información contenida en las bases de datos de secuencias incorporada en el periodo 2005-2010. Las homologías correspondieron principalmente a proteínas hipotéticas sin función caracterizada de maíz y de arroz y se encuentran detalladas en la Tabla 4. Además, una secuencia fue descartada como candidata a sufrir regulación por ploidía, y fue clasificada como posible contaminación, por presentar homología completa con vectores de clonado comerciales (Tabla 4).

Si bien estos estudios previos permitieron inferir una posible anotación funcional de 15 transcritos diferenciales por nivel de ploidía y un transcrito diferencial por modo de reproducción, 39 transcritos permanecieron sin anotación funcional (19 diferenciales por modo de reproducción y 20 diferenciales por nivel de ploidía).

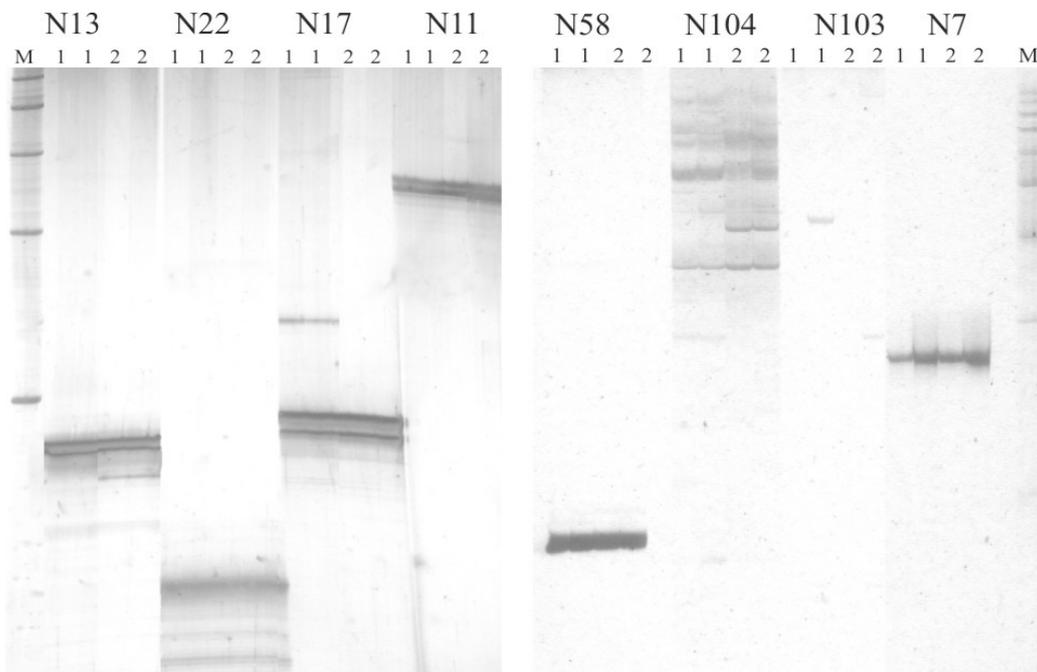
**Tabla 4: Transcritos diferenciales por nivel de ploidía que presentaron homologías con proteínas en las búsquedas bioinformáticas iniciales**

Nombre	ID	Detalle	eV	Regulado por
N116*	gb EU975538.1	Proteína hipotética de maíz	3e-10	Modo de reproducción
L20*	gb EU957444.1	Proteína hipotética de maíz	8e-20	Nivel de ploidía
L24*	gb EU966467.1	Proteína hipotética de maíz	3e-55	Nivel de ploidía
L47*	dbj AK224581.1	cDNA expresado en panoja de <i>Oryza officinalis</i>	4e-10	Nivel de ploidía
L53	gb EU964722.1	Proteína hipotética de maíz	1e-5	Nivel de ploidía
L49*	gb EU951333.1	mRNA de proteína de maíz	1e-21	Nivel de ploidía
L66	LOC100194403	mRNA de proteína de maíz	1e-7	Nivel de ploidía
L774*	LOC100192571	mRNA de proteína de maíz	2e-10	Nivel de ploidía
L802	gb EU975659.1	mRNA de ankirina kinasa de maíz	9e-6	Nivel de ploidía
L844*	gb EU960848.1	mRNA de proteína tipo josephin-2 de maíz	1e-15	Nivel de ploidía
L854*	Os02g0267000	Proteína de <i>Oryza sativa</i> tipo peptidasa I	2e-23	Nivel de ploidía
L75*	Os03g0826600	Proteína de <i>Oryza sativa</i> tipo fosfoesterasa	2e-26	Nivel de ploidía
L76*	gb EU970386.1	Proteína hipotética de maíz	1e-18	Nivel de ploidía
L87*	gb EU365401.1	NADHdh mitocondrial de <i>Bambusa oldhamii</i>	1e-54	Nivel de ploidía
L94*	gb BT018512.1	mRNA de proteína de maíz	2e-24	Nivel de ploidía
L95*	LOC10021731	Proteína hipotética de maíz	7e-49	Nivel de ploidía

\* Se señalan las homologías asignadas a los transcritos diferencialmente expresados (eV del orden de  $e^{-10}$  o menores)

#### 4.2- Amplificaciones genómicas de control

Debido a que muchas de las secuencias permanecían sin posible anotación funcional, se realizaron amplificaciones a partir de ADN genómico de *P. notatum* para descartar contaminaciones. Se diseñaron oligonucleótidos que permitieron amplificar pequeños segmentos internos de cada secuencia, y se los utilizó para generar amplicones para dos genotipos, uno de ellos apomítico (Q4117) y el otro sexual (Q4188). Esto permitió, en los casos de amplificación positiva, descartar la posibilidad de contaminaciones. Estos experimentos fueron realizados en colaboración con la Lic. Maricel Podio. La Figura III muestra dos fracciones de geles de poliacrilamida donde se observan los patrones de los fragmentos genómicos correspondientes a las secuencias N7, N11, N13, N17, N22, N58, N103 y N104.



**Figura III:** Sección de geles de poliacrilamida: amplificaciones de ADN genómico con cebadores específicos diseñados sobre los transcritos donde 1 corresponde al genotipo apomítico Q4117 y 2 al genotipo sexual Q4188. M: marcador de peso molecular.

Se analizaron 16 fragmentos, todos ellos relacionados con el modo de reproducción. Solo 12 fueron amplificados a partir de ADN genómico de *P. notatum*, de los cuales 9 presentaron patrones monomórficos de bandas del tamaño esperado (N7, N11, N13, N22, N38, N56, N58, N70 y N77) y 3 mostraron polimorfismos entre el genotipo sexual y el apomítico (N17, N43 y N104). Los polimorfismos detectados en este

experimento serán posteriormente utilizados para localizar el fragmento en el genoma. Un total de 4 fragmentos no pudieron ser amplificados a partir del genoma de *P. notatum* (N2, N49, N86 y N103) con 4 combinaciones de cebadores. Estos clones fueron clasificados como posibles contaminaciones.

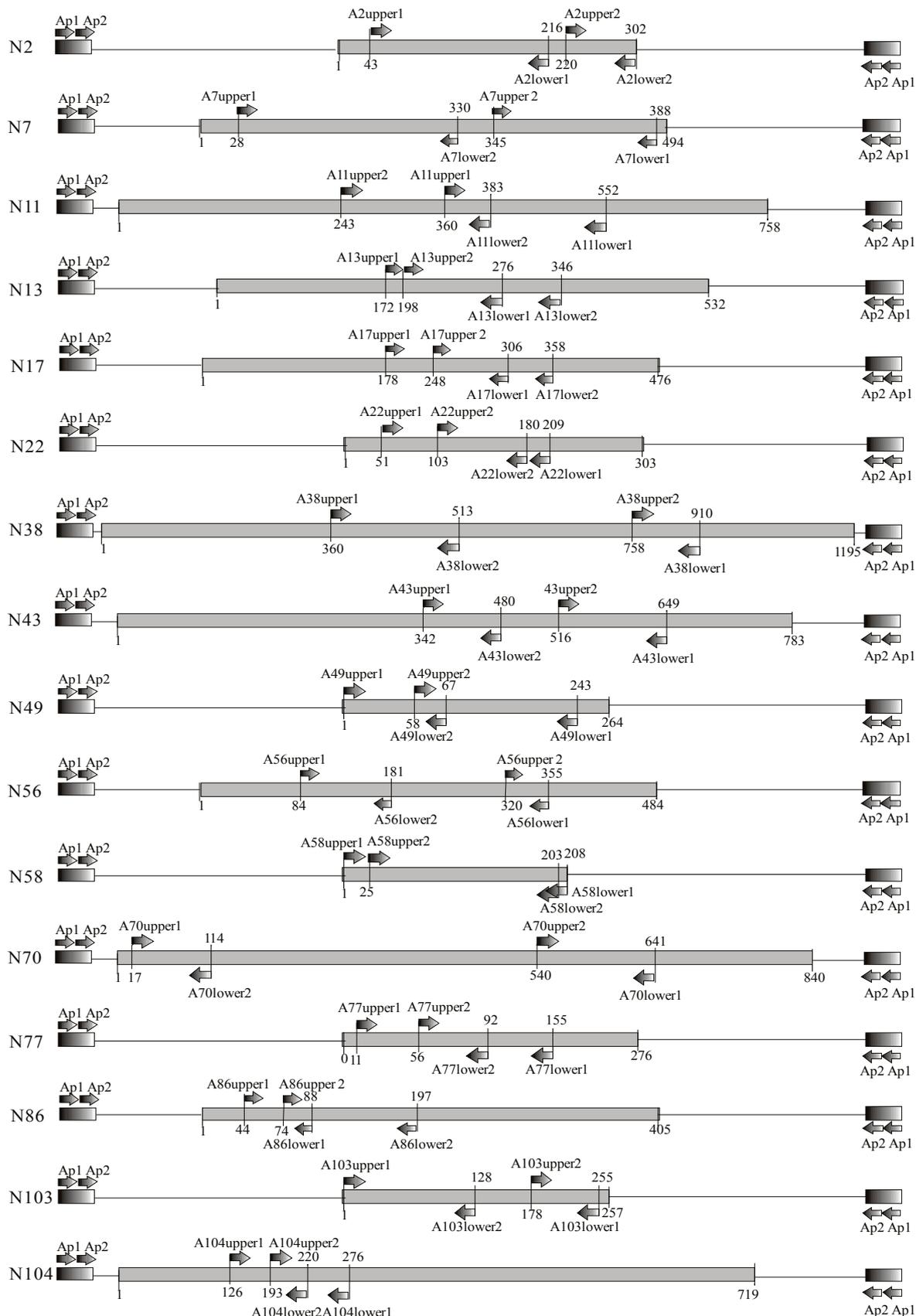
#### **4.3- Extensión de los transcritos por experimentos 5' y 3' RACE**

Para intentar asignar un posible rol funcional a algunas de estas secuencias se decidió extenderlas utilizando la técnica de RACE (Rapid amplification of cDNA ends), de manera de contar con fragmentos de secuencia más extensos que facilitarían la detección de homologías en las bases de datos.

Como estrategia de trabajo, se decidió iniciar el trabajo de extensión de las regiones flanqueantes enfocándonos en un subgrupo de transcritos diferenciales por el modo de reproducción seleccionados al azar.

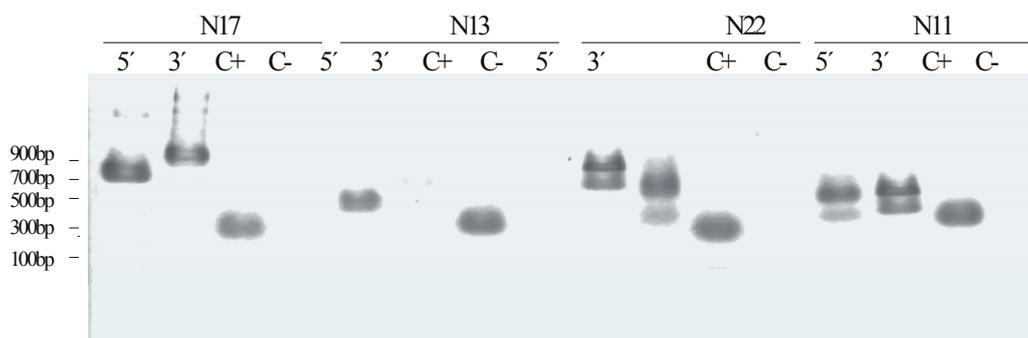
Con la finalidad de aislar las regiones flanqueantes y obtener las secuencias extendidas se utilizó la técnica de RACE (Chenchik *et al.*, 1996). Se construyeron dos bibliotecas no-clonadas Marathon (Clontech), en colaboración con el grupo de trabajo de la Dra. Vera Tavares de Campos Carneiro (Embrapa CENARGEN, Brasilia, Brasil). Las bibliotecas fueron desarrolladas a partir de ADNc de inflorescencias inmaduras de los genotipos Q4117 (apomítico) y Q4188 (sexual) de *P. notatum* en estadio de desarrollo I (Laspina *et al.*, 2008).

Se diseñaron dos pares de cebadores para cada transcrito, respetando los parámetros establecidos en el Marathon cDNA Amplification kit (Clontech). En la Figura IV se detallan las posiciones de los cebadores sobre los diferentes transcritos, en una representación esquemática del diseño de amplificación. Los transcritos aislados como sobrexpresados en el genotipo apomítico fueron investigados dentro de la biblioteca Q4117 y los sobrexpresados en el genotipo sexual fueron amplificados desde la biblioteca Q4188.



**Figura IV:** Estrategia de amplificación RACE. Representación esquemática de los fragmentos iniciales y los oligonucleótidos diseñados para extenderlos hacia los extremos. AP1 y AP2 indican la posición de los cebadores que hibridizan sobre los adaptadores terminales de la biblioteca Marathon. Se señala la posición específica de los cebadores sobre la secuencia a amplificar.

Se obtuvieron amplicones para 8 de los 16 transcritos estudiados (N7, N11, N13, N17, N22, N43, N56 y N104), sin embargo no fue posible aislar bandas puras para el resto de los clones (N2, N38, N49, N58, N86, N70, N77, N103) analizadas por 5' y 3' RACE. La Figura V muestra una sección de un gel de agarosa, donde se muestran ampliificaciones de los sectores 5' y 3' para cada gen y los controles utilizados.



**Figura V:** Sección de geles de agarosa: ampliificaciones 5' y 3' RACE con cebadores específicos diseñados sobre los transcritos N11, N13, N17 y N22.

Los amplicones fueron aislados, clonados y secuenciados. Se detectaron y eliminaron los restos de vectores de transformación utilizando el programa VecScreen (<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>). Las secuencias fueron ensambladas en contiguos con los fragmentos originales, y comparadas con las bases de datos mediante las herramientas bioinformáticas disponibles. Los alineamientos entre secuencias se llevaron a cabo utilizando el programa ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) y BL2seq (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

El diseño de cebadores permitió extender la secuencia de N7 hacia el extremo 3' (de 494 a 540 nt), sin embargo, la extensión 5' no generó ningún amplicón. En este caso se pudo obtener la secuencia completa ya que este fragmento aislado contenía la cola de poliA del transcripto. En el caso de N11, los cebadores permitieron amplificar varias bandas que se solaparon con el transcripto original (758 nt), sin embargo no se logró extenderlo hacia ninguno de los extremos. Para el clon N13 se obtuvo un fragmento elongado hacia el sector 5' ampliando la secuencia de 532 a 718 nt. En el caso de N17 se obtuvieron productos de amplificación con los cebadores diseñados para extender hacia el sector 3'. La extensión de N17 permitió ampliar la secuencia desde 476 a 907 nt. El caso de N22 resultó excepcional ya que se obtuvieron varios amplicones tanto hacia el sector 3' como hacia el 5', sin embargo los cebadores no permitieron extender el transcripto

original. En su lugar se lograron aislar varias secuencias relacionadas que no pudieron ser analizadas como un consenso, ya que no se solapaban con la secuencia obtenida originalmente por Laspina *et al.* (2008). Si bien se detectaron los cebadores en los extremos de las secuencias amplificadas, las regiones solapadas internas no mostraban homología perfecta con N22. Las secuencias fueron analizadas independientemente y fueron consideradas como diferentes miembros de una familia de transcritos relacionados. Se aislaron 5 secuencias relacionadas: A43, A44, A45, A46 y A47 (de: 849, 827, 784, 643 y 802 nt de longitud, respectivamente). La amplificación RACE de N43 arrojó resultados similares que los obtenidos para N22 ya que si bien se obtuvieron varios amplicones tanto para el sector 5' como para el 3' del transcripto, y se detectaron los cebadores en los extremos de las secuencias, tampoco se logró obtener un consenso. Aquí se aislaron 5 transcritos: A11, A12, A13, A15 y A16 de 629, 676, 849, 775 y 169 nt de longitud, respectivamente. Estos fueron estudiados como secuencias relacionadas pertenecientes a una familia de transcritos. En el caso de N56 se logró amplificar un segmento con el cebador específico directo, sin embargo no se alcanzó la extensión del transcripto ya que el sector solapado quedó comprendido dentro de la secuencia original (464 nt). En el caso de N104, si bien se logró aislar un producto de (325 nt), el fragmento permitió extender hacia el 5' sólo en 69 nt la secuencia original (Consenso N104: 786 nt). Cada una de las secuencias obtenidas por medio de los experimentos de 5' y 3' RACE se presentan en el Anexo I Sección B.

#### **4.4- Análisis bioinformáticos de los fragmentos extendidos**

Los estudios de similitud fueron realizados con cada una de las secuencias extendidas por RACE (N7, N11, N13, N17, N22, N43, N56 y N104). Las búsquedas generales se llevaron a cabo en la plataforma de búsqueda del sitio NCBI comparando con las bases de datos de secuencias nucleotídicas y proteicas no redundantes (nr/nt, refseq\_rna y nr) en las especies de plantas verdes (green plants taxid: 33090). Además se estudió la similitud con las secuencias nucleotídicas de *Arabidopsis thaliana*, *Oryza sativa*, *Sorghum bicolor* y *Zea mays*, utilizando las herramientas de búsqueda de los sitios TAIR y Gramene. Se estudió la similitud en las bases de datos de elementos repetitivos de gramíneas (TIGR) y en las bases de datos de microARN (Sanger). Todos los resultados obtenidos durante las búsquedas bioinformáticas se encuentran resumidos en la sección A del Anexo II.

Las búsquedas realizadas con la secuencia consenso de N7 (540 nt) no permitieron identificar homologías significativas con genes codificantes en los sitios NCBI, TAIR y Gramene. No se identificaron homologías significativas en las bases de datos de miARN de Sanger, ni en las bases de datos de elementos repetitivos TIGR. Cuando la secuencia fue traducida en los 6 marcos de lectura posibles utilizando la herramienta Transeq y sus homologías con proteínas conocidas o dominios proteicos fueron indagadas con el programa InterProScan, sólo se detectó un sector de 21 AA (377 a 540 de la secuencia nucleotídica) con el patrón de péptido señal (Signal-P-NN euk) en el marco -2. Se completó el estudio realizando la predicción de péptido señal y la detección del sitio de corte con el programa SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP/>) donde se obtuvo un resultado altamente positivo para 163 nt del sector 3', con una probabilidad de predicción de 0,99, siendo el valor máximo igual a 1. El sitio de corte predicho quedó localizado entre el AA 20 y 21 (nt 376-377). Se realizaron nuevas búsquedas tBlastx en NCBI con el sector 5' de la secuencia nucleotídica de N7 (1-376) y se detectó un sector similar a lo largo de toda la secuencia correspondiente a una proteína hipotética de *Glycine max* (FJ014716.1), que contaba con un 98% de identidad con Glucosa-1-fosfato adeniltransferasa de *A. thaliana*. La búsqueda de dominios conservados en las bases de datos Uniprot (Universal Protein Resource) por medio de la herramienta MotifScan ([http://myhits.isb-sib.ch/cgi-bin/motif\\_scan](http://myhits.isb-sib.ch/cgi-bin/motif_scan)) permitió identificar un dominio conservado correspondiente al sitio de fosforilación Protein Kinasa C (AA 88 a 93 en el marco -3). Estos resultados sugieren que el transcripto podría corresponder a una secuencia proteica del tipo kinasa muy poco conservada en especies vegetales, sin embargo, las dimensiones de los valores esperados de las búsquedas no representan suficiente evidencia para poder afirmar esta funcionalidad.

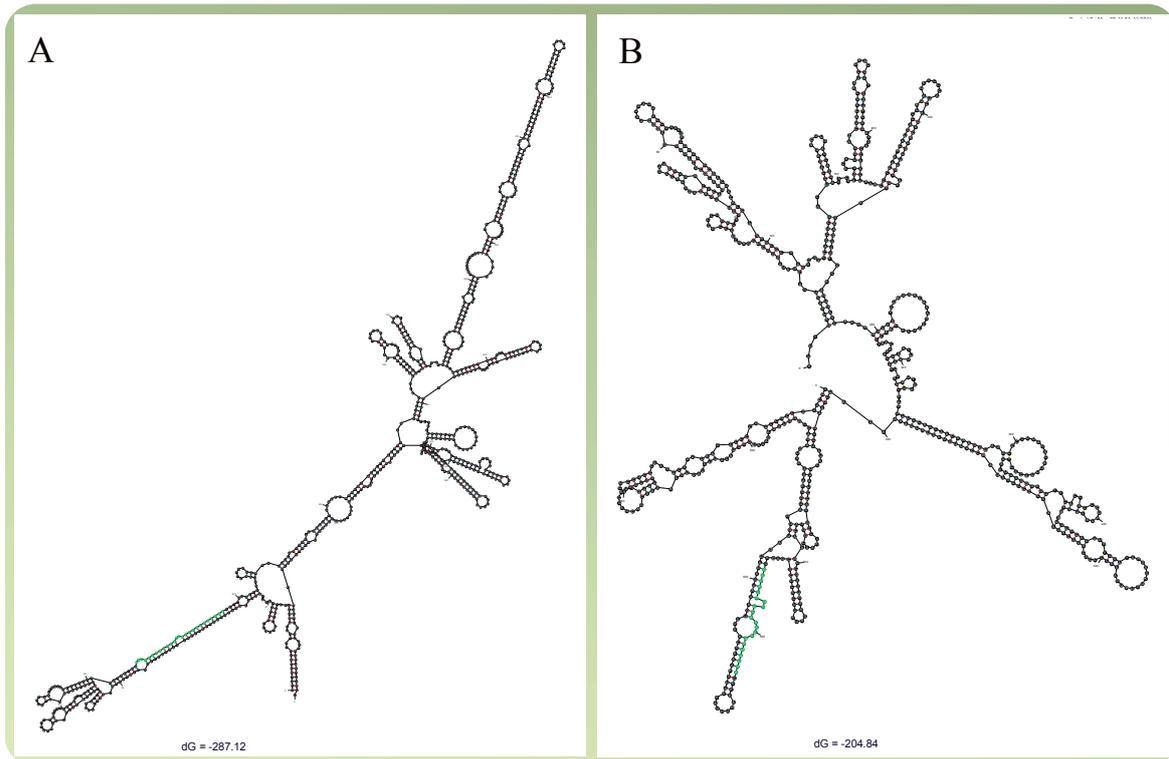
Si bien la secuencia N11 original era de una longitud considerable (758nt), sólo se identificó similitud en un pequeño sector de 26 nt (nt 170 a 195) con una proteína hipotética Os03g0199100 de *O. sativa* (eV: 0,065, correspondencia nucleotídica 24/26, identidad 92%). Por otra parte, cuando se realizaron las búsquedas de rutina en las bases de datos de miARN del Instituto Sanger, se detectó un precursor de miARN de *A. thaliana* ath-MIR165b con una importante similitud de secuencia con N11 (score: 103). Sin embargo, la homología se dispersaba en el sector correspondiente al miARN maduro de ath\_MIR 165b, y en su lugar se localizaba el fragmento de 26 nt Os03g0199100 (170-195 nt). Se estudió la estabilidad de la estructura secundaria *in silico* y la posición de la región

conservada dentro de la misma. Se desarrollaron los plegamientos de las hebras (+) y (-) del consenso N11 a 25°C utilizando el programa Mfold 2.3. Sólo la hebra (-) presentó una estructura secundaria estable ( $\Delta G = -274$ ) y el sector conservado de *O. sativa* se localizó precisamente en la región característica del tallo. Se detectaron en este sector las tres burbujas distintivas presentes en los pre-miARN de plantas: en el sector inferior del tallo (loop distal), en el centro (miARN:ARNm) y por encima del tallo (loop proximal) (Schwab *et al.*, 2005). En la Figura VI se muestra el plegamiento consenso N11 (-), donde se señala en color verde el sector del miARN maduro predicho. Se analizó el sector conservado de la secuencia N11 como posible miARN maduro realizando la predicción de las secuencias blanco putativas por medio del programa miRU: Plant miRNA Potential Target Finder. Este programa analiza la perfecta o casi perfecta complementariedad entre los miARN de plantas y sus ARNm blanco, para predecir los potenciales transcriptos regulados (Schwab *et al.*, 2005; Zhang, 2005). El programa arrojó que dos genes de arroz cumplían con las características necesarias del blanco regulatorio: Os03g0199100 (Match: 1,5) y Os03g10240.1 (Match: 1,5) (Rhoades *et al.*, 2002). Los alineamientos entre estas dos secuencias de ARNm permitieron inferir que ambas serían alelos del mismo gen ya que su identidad fue del 99% en 2129 nt y el valor esperado de 0,0. Posteriormente se analizó la estabilidad de la hibridización entre el miARN y su blanco putativo con el programa RNAhybrid (Rehmsmeier *et al.*, 2004). El programa calcula la energía libre de la complementariedad perfecta y la energía libre calculada, siendo el valor suficiente para la predicción que la energía libre calculada represente al menos el 72% del  $\Delta G$  de complementariedad perfecta (Schwab *et al.*, 2005). El resultado de la energía libre calculada entre el miARN y su ARNm complementario predicho fue del 87 % de la energía libre obtenida por complementariedad perfecta. A partir de estos resultados se planteó la hipótesis de que el clon N11 podría corresponder a un precursor primario de miARN relacionado al pre-miARN de *A. thaliana* ath-MIR165b, que tendría como blanco regulatorio al gen ortólogo de Os03g0199100. Con la finalidad de corroborar esta hipótesis se desarrollaron experimentos de detección del gen blanco predicho por 5'RACE (Ver sección 4.5.2.4).

Por otro lado, la secuencia consenso N13 (718nt) tampoco mostró homologías en las búsquedas ordinarias en NCBI, TAIR y Gramene. Al igual que en el caso de N11, se identificaron similitudes de fragmentos cortos muy conservados en especies de plantas. Estos fragmentos correspondieron mayoritariamente a regiones intergénicas y una

secuencia codificante de *Zea mays* (gb|BT068773.1) a lo largo de 22 nucleótidos, desde 547 a 568 (eV: 0,12, correspondencia nucleotídica 22/22, identidad 100%). Este gen de maíz codifica para una proteína no caracterizada que presenta un dominio ETR, el cual integra proteínas involucradas en los sistemas de señalización de dos componentes que son fosforiladas por histidin-quinasas y luego forman homodímeros funcionales (Entrez /BT068773.1). No se detectaron marcos de lectura abiertos de tamaño proteico en el transcripto, ni homologías con elementos repetitivos de gramíneas. La comparación con las bases de datos de miARN reveló sectores similares a varios pre-miARNs de plantas. El mejor resultado fue obtenido para el precursor osaMIR1877 (score 135) en un sector de 105 nt. Las predicciones de estructura secundaria con Mfold 2.3 desarrolladas para ambas hebras (+ y -) permitieron detectar que sólo la hebra (-) presentaba una estructura secundaria estable ( $\Delta G = -204,64$ ). El sector conservado de 22 nt se localizó en plegamiento típico de horquilla y en este caso también se observaron todas las burbujas características de los pre-miARN de plantas descritas por Song (2010), sin embargo la estructura central (miARN:ARNm) no se ajustó a las dimensiones halladas comúnmente. La Figura VI muestra el plegamiento predicho para N13 (-), y se señala en verde el sector conservado que podría corresponder al miRNA maduro. El análisis con el programa miRU permitió identificar un transcripto que cumplía con las características previstas por Rhoades *et al.* (2002). Estos resultados sugieren que el clon N13 podría corresponder a un precursor primario relacionado con el pre-miARN de arroz osaMIR1877, el cual tendría como blanco regulatorio al ARNm similar a BT068773 de maíz.

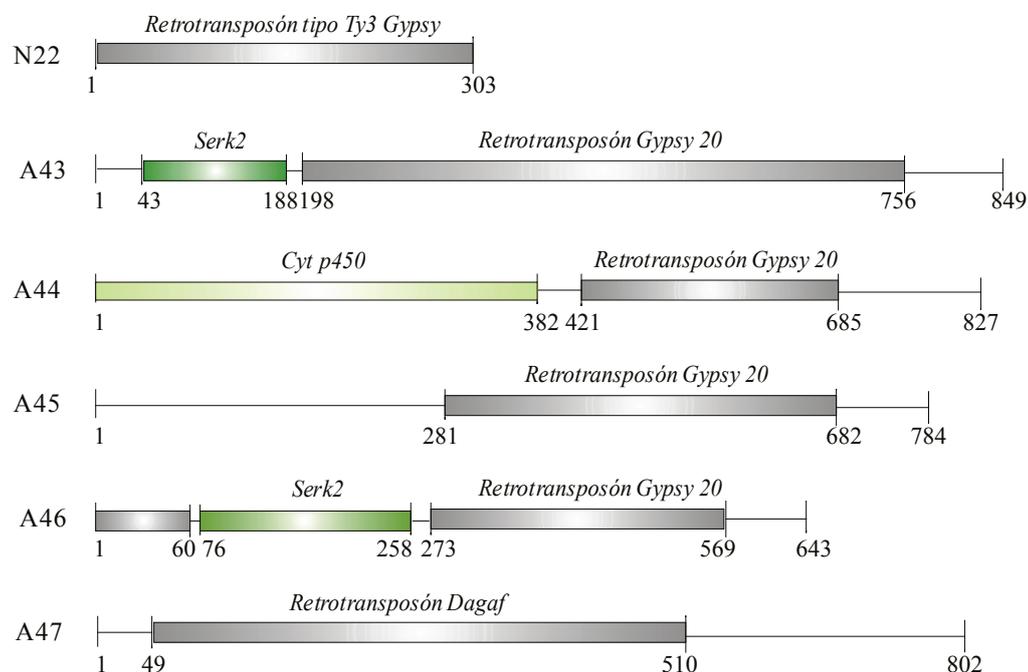
Las búsquedas en NCBI, TAIR y Gramene para la secuencia extendida de N17 revelaron una homología parcial con retrotransposones. Al ser comparada con las bases de datos de elementos repetitivos en el sitio TIGR se detectó una importante similitud con un sector de al menos 200 nt con la región LTR de un retrotransposón Aboov de posición centromérica de *Oryza sativa* (ORSiCMCM00100010; eV:  $9,2e^{-06}$ ). Asimismo, se detectó un pequeño sector conservado dentro de la secuencia extendida N17 (de 816 a 878 nt) similar a la proteína hipotética no caracterizada Os11g0269700 de *O. sativa* (eV:  $2e^{-11}$ , 98% de identidad). Estos resultados indican que N17 podría corresponder a un retrotransposón de posición centromérica portador de un pequeño fragmento génico transducido de 62 nt, similar a Os11g0269700 de *O. sativa*.



**Figura VI:** Predicciones de las estructuras secundarias espontáneas adoptadas por los transcritos N11 y consenso N13 a 25°C halladas con el programa Mfold 2.3. Al pie de los diagramas se muestra la Energía libre ( $\Delta G$ ) calculada para cada estructura. El panel A corresponde al plegamiento de la hebra (-) de N11, se señala en color verde el sector del miARN maduro predicho. El panel B presenta el plegamiento de la hebra (-) del consenso N13, se señala en color verde el sector del miARN maduro predicho.

Por otra parte, las nuevas búsquedas de baja astringencia realizadas con la secuencia N22 en las bases de datos de elementos repetitivos en el sitio TIGR permitieron detectar incipientes homologías con el sector codificante para la proteína Gag del retrotransposón Ty3-*Gypsy* ORSiTERT00200152 (eV: 0,97). Como se comentó en la sección anterior, los experimentos de RACE realizados permitieron el aislamiento de cinco fragmentos relacionados pero no idénticos (A43, A44, A45, A46 y A47) que no pudieron ser considerados como una secuencia consenso sino que fueron analizados independientemente como miembros de una familia de transcritos. El clon A43 mostró homología con un retrotransposón *Gypsy* 20 de *Arabidopsis lyrata* (gb|EU558521.1|, eV:  $3e^{-65}$ ) entre los nucleótidos posicionados 198-756, mientras que entre los nucleótidos 43-188 mostró homología con el gen *serk2*: gen de la embriogénesis somática tipo receptor kinasa 2 de *Poa pratensis* (emb|AJ841697.1| eV:  $3e^{-28}$ ). El clon A44 mostró homología con el mismo retrotransposón *Gypsy* 20 (gb|EU558521.1|, eV:  $8e^{-35}$ ) entre los nucleótidos 421

y 685, mientras que en la posición 1-382 se detectó homología con el gen del citocromo p450 de *Zea mays* (gb|EU967060.1|, eV:  $6e^{-100}$ ). El clon A45 reveló homología con el mismo retrotransposón *Gypsy 20* (gb|EU558521.1|, eV:  $1e^{-16}$ ) pero no se detectaron homologías con secuencias génicas de plantas. El clon N46 mostró homologías con el mismo retrotransposón *Gypsy 20* (gb|EU558521.1|, eV:  $2e^{-60}$ ) en los intervalos 4-60 y 273-569, mientras que en el sector interno (76-255 nt) presentó homologías con el gen *serk2* de *Poa pratensis* (emb|AJ841697.1| eV:  $2e^{-54}$ ). El clon A47 resultó similar a otro retrotransposón ZRSiTERTO00146, dagaf\_576C20-1, eV:  $3e^{-8}$ ). Estos resultados indican que todos los fragmentos derivados de N22 podrían corresponder a miembros de una familia de retrotransposones, algunos de los cuales portan segmentos de genes codificantes de plantas. Sorprendentemente, estos fragmentos codificantes corresponden a genes que habían sido asociados al desarrollo de la apomixis. Tanto el gen *serk2* como el *cytP450* fueron reportados como diferencialmente regulados durante el desarrollo apomítico en *Poa pratensis*, *Paspalum notatum* y *Eragrostis curvula* (Albertini *et al.*, 2005; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008). En la figura VII se muestra un diagrama de las características y disposiciones de las secuencias derivadas de N22.



**Figura VII:** Representación esquemática de las secuencias relacionadas a N22 amplificadas por RACE. Las secuencias similares a retrotransposones se representan en rectángulos grises y los que corresponden a genes codificantes de plantas se encuentran coloreados. Las posiciones exactas se demarcan con la numeración de nucleótidos que se encuentra debajo de cada rectángulo.

El estudio bioinformático de la secuencia N43 reveló un sector de 111 nt correspondiente a contaminaciones con vectores de clonado. La contaminación fue eliminada y las búsquedas realizadas en el sitio NCBI con la nueva secuencia N43 (672 nt) permitieron hallar homologías significativas (nt 456-639) con una inserción de varios tipos de transposones de *A. thaliana*: CACTA2, SINE9, MITE7, *Gypsy* 20 (eV:  $2e^{-11}$ ). Los experimentos de RACE permitieron aislar 5 transcritos a partir de N43 (A11, A12, A13, A15 y A16). Si bien todas las secuencias contuvieron los cebadores específicos diseñados sobre N43, no fue posible construir una secuencia consenso entre ellas. Al igual que en el caso de N22, cada uno de los fragmentos aislados fueron analizados independientemente. A11 mostró muy altas homologías con el mismo tipo de inserción de transposones tipo SINE9, MITE7 y *Gypsy* 20 de *A. thaliana* que N43 (gb|EU558521.1|, eV:  $5e^{-113}$ ) a lo largo de casi toda la secuencia (desde 132 a 672). Se realizaron búsquedas restringidas con los primeros 132 nucleótidos de A11 y se detectó una secuencia similar a una proteína hipotética de *Z. mays* (gb|BT035386.1|), la cual se conoce que contiene dos dominios kinasa y un dominio de unión a ATP (NCBI Entrez/ GENE ID: 100193188 LOC100193188, eV:  $2e^{-8}$ ). La secuencia A12 presentó homologías significativas con una proteína hipotética de *Sorghum bicolor* (ref|XM\_002441009.1| eV:  $6e^{-156}$ ) a lo largo de casi toda la secuencia (170-652). Dicha proteína estaba conformada por un dominio PNGaseA: N-acetyl-beta-D-glucosaminyl asparaginasa amidasa A (pFam: 12222) cuya función es la de degradar cadenas polipeptídicas. Se detectó además que los primeros 170 nt de A12 resultaron similares a las mismas inserciones de transposones de A11 (gb|EU558521.1|, eV: 0.056). En cuanto a la secuencia A13, se detectaron similitudes con secuencias conocidas repetitivas sólo para el fragmento correspondiente a los cebadores específicos en las búsquedas en NCBI. Sin embargo en la base de datos de elementos repetitivos se halló un sector interno de 179 nt similar a un retrotransposon denominado *angela\_SK14\_4* (TRSiTERT00100121, eV: 0,63). La secuencia A15 no mostró homologías en las búsquedas ordinarias de NCBI, aunque en las bases de datos TIGR se identificó similitud con la región LTR del retrotransposón *Sukkula* de *Hordeum vulgare* (HRSiTERTO00038, eV: 0,47) desde 586 a 659 nt. Las búsquedas en Gramene permitieron detectar similitudes en sectores pequeños correspondientes a secuencias codificantes no caracterizadas localizadas en el cromosoma 3 y 4 de arroz. Algunas de estos genes contienen dominios de retrotranscriptasas e integrasas por lo que se podría deducir que corresponderían a retrotransposones no caracterizados. Finalmente, la

secuencia A16 presentó una homología a lo largo de toda su secuencia con una proteína hipotética de *S. bicolor* (ref|XM\_002488912.1|, eV:  $5e^{-55}$ ), sin embargo no se pudieron identificar sus características funcionales ni dominios proteicos de esta última. De manera similar a N22, estos resultados podrían indicar que los fragmentos derivados de N43 corresponderían a transcritos relacionados a retrotransposones, algunos de los cuales portarían segmentos transduplicados de genes de plantas.

Como se había mencionado en el apartado anterior, los experimentos RACE no permitieron elongar la secuencia N56. Se obtuvo un amplicón a partir de los cebadores orientados hacia el sector 3', sin embargo el fragmento (164 nt) se encontró abarcado dentro de la secuencia original. Se realizaron nuevas búsquedas informáticas exhaustivas con la secuencia N56 original. No se identificaron similitudes con proteínas conocidas en las bases de datos de NCBI, tampoco con secuencias repetitivas de gramíneas en TIGR. Las búsquedas en el sitio Gramene y en las bases de datos TIGR arrojaron similitudes muy bajas, que no permitían arribar a conclusiones. Tampoco se detectaron homologías significativas con precursores de miARN de plantas. Sin embargo, se identificó homología con un miARN maduro de *A. thaliana*: ath-miR861-3p (eV: 0.67, Score: 82), localizado en la posición 197-216. Se analizó la estructura secundaria del transcripto para establecer si existía la posibilidad de que se tratase de un precursor de miARN con secuencia poco conservada en otras especies. Si bien ambas hebras presentaron plegamientos estables ( $\Delta G = -186$  y  $-155,58$ , respectivamente), sólo la hebra (-) localizó el sector homólogo al miARN maduro de *A. thaliana* en una estructura tipo horquilla. Se realizaron predicciones en el sitio miRU y se analizaron cada una de las consideraciones planteadas por Schwab et al (2005) para la detección del tipo y número de bases incorrectamente apareadas entre miARN y blanco regulatorio de un mismo genotipo. Si bien se identificó un único gen (At1g53010.1, Score: 2,5) que cumplió con las características de apareamiento de bases, su energía de hibridización fue del 55,3% de la energía libre de complementariedad perfecta. Es por esto que no se pudo predecir el blanco regulatorio en *A. thaliana*. Estos resultados sugieren que el transcripto podría corresponder a un precursor de miARN, sin embargo, al no poder predecir su gen blanco no sería posible confirmar esta funcionalidad putativa.

Finalmente los estudios realizados con la secuencia consenso N104 (786 nt) mostraron similitudes en un sector de 60 nt (2-62) con una proteína no caracterizada de *Sorghum bicolor* (XM\_002445379.1, eV:  $6e^{-12}$ ). Se identificó además similitud del fragmento inmediatamente consecutivo (83-162) con una región promotora de un gen

rd29A de *Zoysia japonica* (gb|EU346948.1|, eV:  $5e^{-6}$ ) del que se conoce únicamente su capacidad de ser inducida en condiciones de estrés. No se identificaron homologías significativas con las bases de datos de elementos repetitivos, ni de precursores de miARN.

#### 4.5- Análisis casuístico de los transcritos extendidos

Con la finalidad de obtener una caracterización más completa de los transcritos en estudio se seleccionaron al azar cuatro de ellos: dos que resultaron similares a elementos repetitivos (N17 y N22) y dos a precursores de miARN (N11 y N13). Se desarrollaron experimentos con el objetivo de poner a prueba esta asignación funcional hipotética y confirmar la expresión diferencial en plantas apomícticas y sexuales.

##### 4.5.1- Estudio de transcritos homólogos a retrotransposones (N17 y N22)

###### 4.5.1.1- Determinación del número de copias genómicas

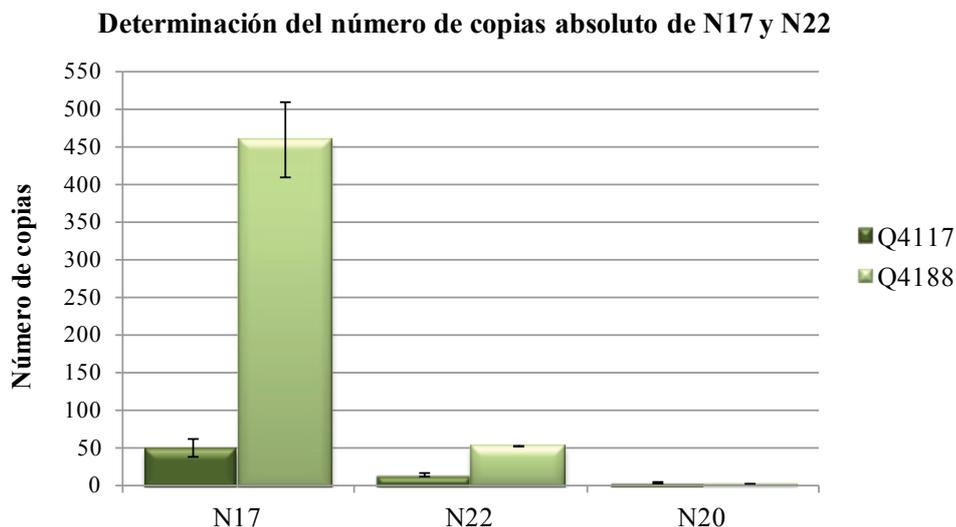
Se estudió la representación de los dos fragmentos (N17 y N22) en el genoma de *P. notatum* por medio de la estimación cuantitativa del número de copias genómicas en los genotipos Q4117 (tetraploide apomíctico) y Q4188 (tetraploide sexual). Para esto se desarrollaron amplificaciones de PCR cuantitativa en tiempo real. Se incluyó el gen N20 como control de bajo número de copias e igual representación en el genotipo sexual y apomíctico (Felitti *et al.*, 2011). N20 codifica para un ortólogo putativo del gen LORELEI de *A. thaliana* (Felitti *et al.*, 2011). La familia LORELEI posee sólo cuatro miembros en *arabidopsis* (Capron *et al.*, 2008).

Para el cálculo del número de copias génicas por PCR en tiempo real se consideró el valor C informado para el genoma diploide de *P. notatum*, que es de 0,6 pg (Jarret *et al.*, 1995). Se construyeron curvas de calibración para N17 y N22 incluyendo 5, 50, 500, 5000 y 50000 copias de los respectivos insertos contenidos en plásmidos recombinantes. Los valores de despegue (“take off”) en los genotipos Q4117 y Q4188 fueron determinados por triplicado. Se analizaron 4 diluciones del ADN genómico ( $30 \times 10^{-9}$  g/ $\mu$ l,  $30 \times 10^{-11}$  g/ $\mu$ l,  $30 \times 10^{-12}$  g/ $\mu$ l y  $30 \times 10^{-13}$  g/ $\mu$ l) y sólo fueron considerados para su posterior extrapolación los puntos abarcados en el rango lineal de las curvas. Los datos y reportes de cada amplificación fueron analizados con el programa Rotor-Gene Q Series Software 1.7.

Los experimentos permitieron detectar un promedio de número de copia para N17 de 50,5 por genoma monoploide en el genotipo apomíctico y 460,4 en el sexual. Se

seleccionó un test no paramétrico (Test de Fiedman) de análisis de la varianza para evaluar la significancia entre ambas variables, y para evitar el viciado de las curvas de distribución, ya que el número de datos resultó entre 3 y 5 para cada punto. La diferencia en el número de copias entre los genotipos resultó altamente significativa ( $p \leq 0,0001$ ,  $\alpha = 0,05$ ). De la misma manera, el número de copias por genoma monoploide de N22 resultó significativamente diferencial, siendo los valores promedio de 16,3 para el genotipo apomítico y 53,2 para el sexual ( $p \leq 0,0001$ ,  $\alpha = 0,05$ ). Por otro lado, para el gen de referencia N20 (proteína anclada a GPI-LORELEI), se obtuvieron los valores promedio de 4,33 en Q4117 y 2,83 en Q4188 por genoma monoploide. En este caso la diferencia no resultó estadísticamente significativa ( $p = 0,18 > \alpha = 0,05$ ).

A partir de estos resultados fue posible confirmar que N17 y N22 serían miembros de familias de elementos repetitivos que se encuentran moderadamente representados en el genoma de *P. notatum*. Se detectó además que el número de copia resultó diferencial entre el genotipo apomítico y el sexual, con una mayor representación en el genotipo sexual para ambos transcritos. Como era de esperar, el gen N20 presentó un bajo número de copia (3 a 4 copias por genoma monoploide) pero no se detectó un patrón diferencial entre el genotipo sexual y el apomítico (Figura VIII).

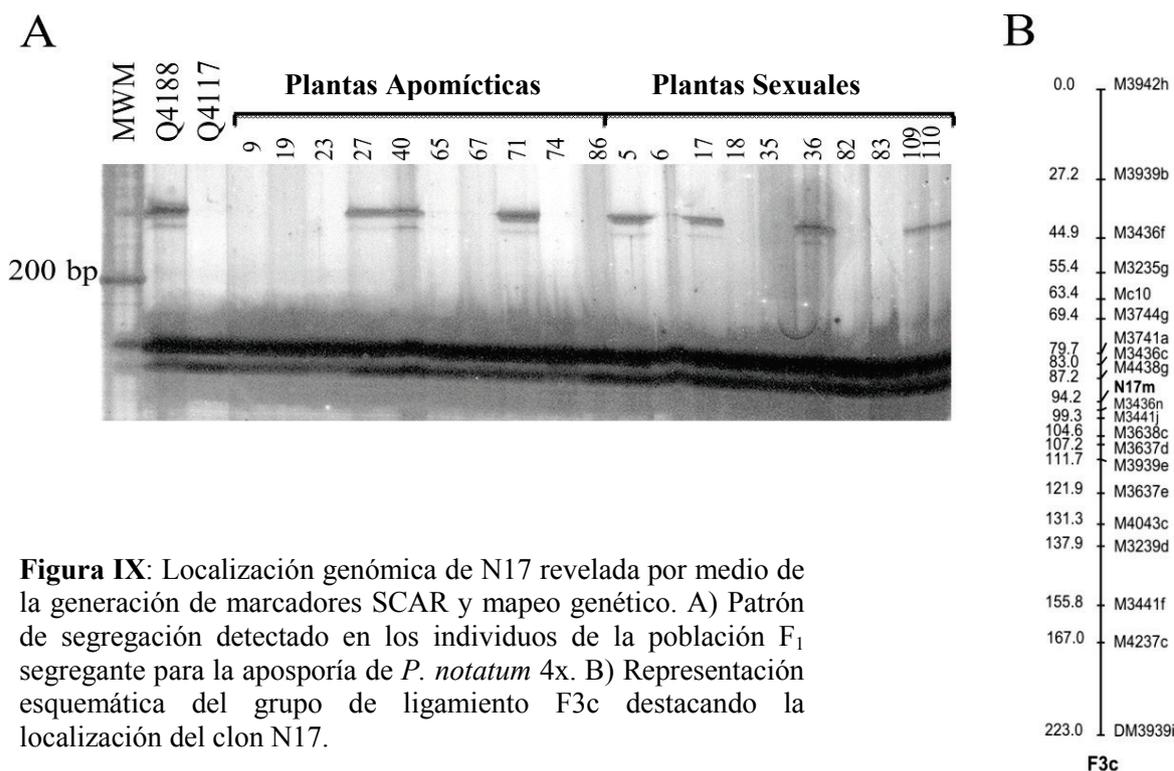


**Figura VIII:** Estimación del número de copia por genoma monoploide en el genotipo apomítico (Q4117) y el genotipo sexual (Q4188) de *P. notatum*. Tanto N17 como N22 mostraron un patrón típico de secuencias repetitivas y se observó un valor significativamente mayor en el genotipo sexual. El gen de referencia N20 se encontró igualmente representado con bajo número de copias en ambos genotipos.

#### 4.5.1.2- Experimentos de mapeo de N17

Los resultados obtenidos en las búsquedas bioinformáticas de los transcritos extendidos por RACE, indicaron homología entre N17 y un retrotransposón de posición centromérica. Se llevaron a cabo experimentos de mapeo genético con la finalidad de determinar la localización de N17 en el genoma de *P. notatum*. Se generaron marcadores de tipo SCAR a partir de ADN genómico de los genotipos parentales Q4117 y Q4188 y 62 individuos F<sub>1</sub> de la población segregante para la aposporía de tipo pseudo-test-cross F<sub>1</sub>, la que había sido anteriormente utilizada para la construcción de un mapa genético de *P. notatum* 4x (Stein *et al.*, 2007). Durante las amplificaciones de ADN genómico parental (Figura III) se identificaron polimorfismos útiles para mapeo. Se controló la segregación de la banda polimórfica en los 62 individuos F<sub>1</sub> (Figura IX A) y se utilizó el programa MAPMAKER 3.0 para realizar los análisis de ligamiento (Lander *et al.*, 1987).

Los resultados arrojados por el programa de mapeo indicaron que N17 resultó ligado en fase de acoplamiento al grupo F3c localizado en una región cromosómica central de baja recombinación (Figura IX B). Por otra parte, se detectó que N17 se encontraba ligado en fase de acoplamiento a una distancia de aproximadamente 30 cM a otro gen expresado diferencialmente entre inflorescencias de plantas sexuales y apomícticas de *P. notatum*: la kinesina Mc10 (Stein *et al.*, 2007). Esta localización cromosómica de N17 en un sector centromérico de baja recombinación coincide con las homologías detectadas en las búsquedas bioinformáticas previas. Debido a que N17 no resultó ligado genéticamente a la región que gobierna la aposporía en *P. notatum*, no hay evidencias que indiquen que este transcrito constituiría un determinante genético primario de este sistema reproductivo. Sin embargo, debe considerarse que se trata de un elemento repetido en el genoma, y que solamente se mapeó una copia del mismo. La localización de esa copia en una región no recombinante central está de acuerdo con los análisis de homología, que indican una anotación como retrotransposón centromérico.



**Figura IX:** Localización genómica de N17 revelada por medio de la generación de marcadores SCAR y mapeo genético. A) Patrón de segregación detectado en los individuos de la población F<sub>1</sub> segregante para la aposporía de *P. notatum* 4x. B) Representación esquemática del grupo de ligamiento F3c destacando la localización del clon N17.

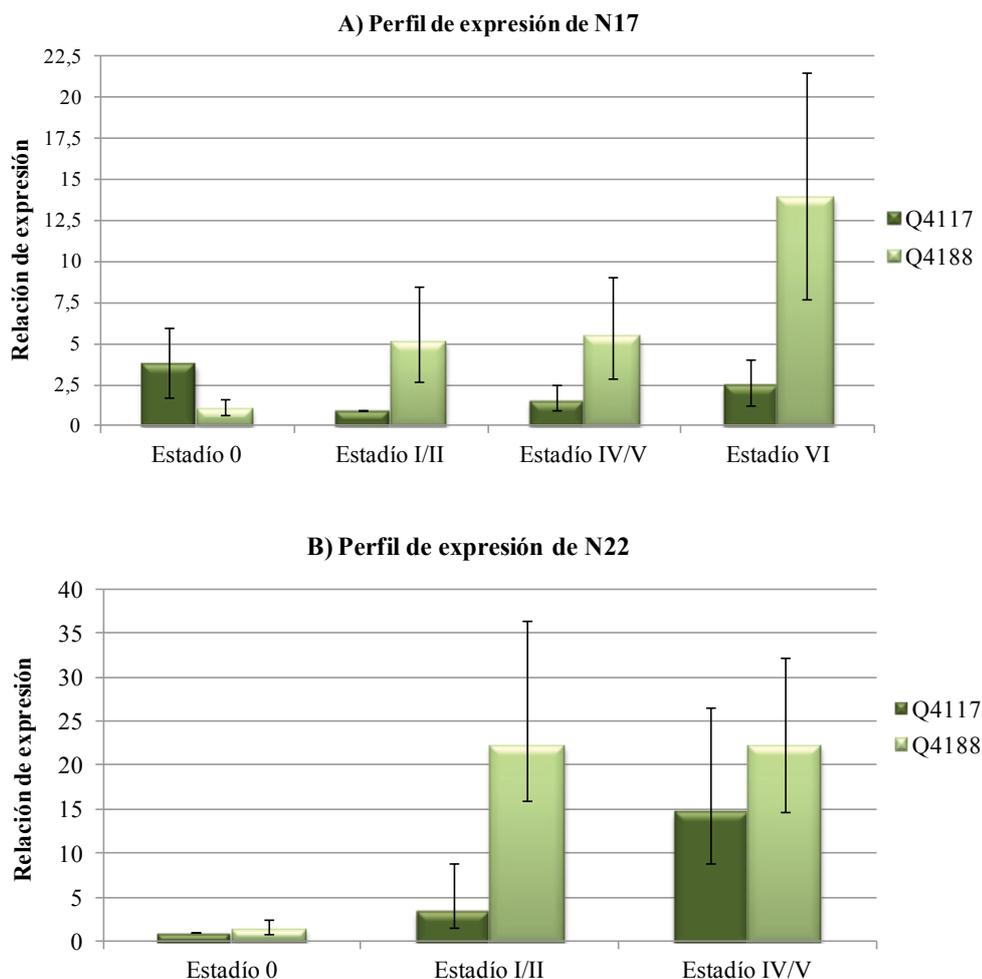
#### 4.5.1.3- Estudios de expresión por PCR en tiempo real

Con la finalidad de caracterizar y comparar el perfil de expresión de los transcritos N17 y N22 en las diferentes fases del desarrollo reproductivo entre plantas sexuales y apomícticas, se llevó a cabo una cuantificación relativa de la expresión por medio de PCR en tiempo real. Para ello se utilizaron muestras de ARN extraídas de espiguillas colectadas en estadio de pre-meiosis (0), pre-meiosis tardía/meiosis (I/II), post-meiosis (IV/V) y anthesis (VI), según el calendario reproductivo reportado por Laspina *et al.* (2008). Se utilizaron cebadores diseñados sobre los fragmentos originales y se incorporaron Glucosa-6- fosfato deshidrogenasa y  $\beta$  Tubulina como genes de referencia. Este gen fue reportado como uno de los genes constitutivos más estables en tejidos reproductivos de plantas sexuales y apomícticas en la especie apospórica *Brachiaria brizantha* (Duarte Silveira *et al.*, 2009). Las amplificaciones fueron realizadas por triplicado utilizando dos repeticiones biológicas para cada una de las determinaciones. Además se incorporaron controles negativos a los que no se les agregó el ADNc. El análisis estadístico se desarrolló con el programa REST-RG (Relative Expression Software Tool V 2.0.7 para Rotor Gene, Corvett Life Sciences).

Los datos obtenidos a partir de los experimentos de cuantificación relativa de N17 se detallan en el gráfico superior presentado en la Figura X. Las relaciones de expresión

fueron calculadas con respecto a la planta apomítica en estadio I/II (estadio de premeiosis tardía/ meiosis, el punto de menor expresión del experimento). Se observó en el genotipo sexual, que los niveles de expresión aumentaron desde 1,1 (error estándar: 0,428 - 1.318) a 5,097 hacia el estadio meiótico (error estándar: 2,442 - 9,496), a 5,498 en post-meiosis (error estándar: 2,835 - 9,040) y ascendieron a 13,88 durante la antesis (error estándar: 7,607 - 22,457). Mientras que para el genotipo apomítico, en estadio pre-meiótico, se detectó una relación de expresión levemente superior a la sexual (3,73, error estándar: 1,984 - 5,972), que luego descendió a 1,539 (error estándar: 0,812 - 2,542) y se mantuvo en estos valores hasta el final del desarrollo apomítico (2,28 en postmeiosis, error estándar: 1,278 - 3,663; y 2,42 en antesis, error estándar: 1,273 - 4,097). Estos resultados sugieren que N17 se encontraría activado en el genotipo sexual a partir del estadio meiótico y permanecería de esta manera hasta alcanzada la antesis, a diferencia del genotipo apomítico que mantendría un nivel bajo de expresión durante todo el desarrollo.

Los datos obtenidos a partir de los experimentos de cuantificación relativa de N22 se detallan en el gráfico inferior de la Figura X. Las relaciones de expresión fueron calculadas con respecto a la planta apomítica en estadio pre-meiótico (el punto de menor expresión del experimento). Se observó que en el genotipo apomítico, los niveles de expresión aumentaron a 3,476 en meiosis (Error estándar: 1,62-8,91) y a 14,64 en post-meiosis (Error estándar: 6,92-33,00). Mientras que en el genotipo sexual, en estadio pre-meiótico, la relación de expresión se mantuvo baja (1,44, Error estándar: 0,41-1,13), aumentó drásticamente a 22,18 en meiosis (Error estándar: 16,05-36,40) y conservó esta tendencia en estadio post-meiótico (22,15, Error estándar: 14,62-32,26). Estos resultados sugieren que N22 se encontraría activado en el genotipo sexual durante el estadio meiótico y permanecería así en estadio postmeiótico, a diferencia del genotipo apomítico que se induciría hacia el final del desarrollo reproductivo. Cabe destacar que si bien las predicciones bioinformáticas de N22 y las amplificaciones RACE permitieron identificar que pertenecería a una familia de elementos repetitivos relacionados, este experimento midió únicamente los perfiles de expresión del clon original N22 en el calendario reproductivo.



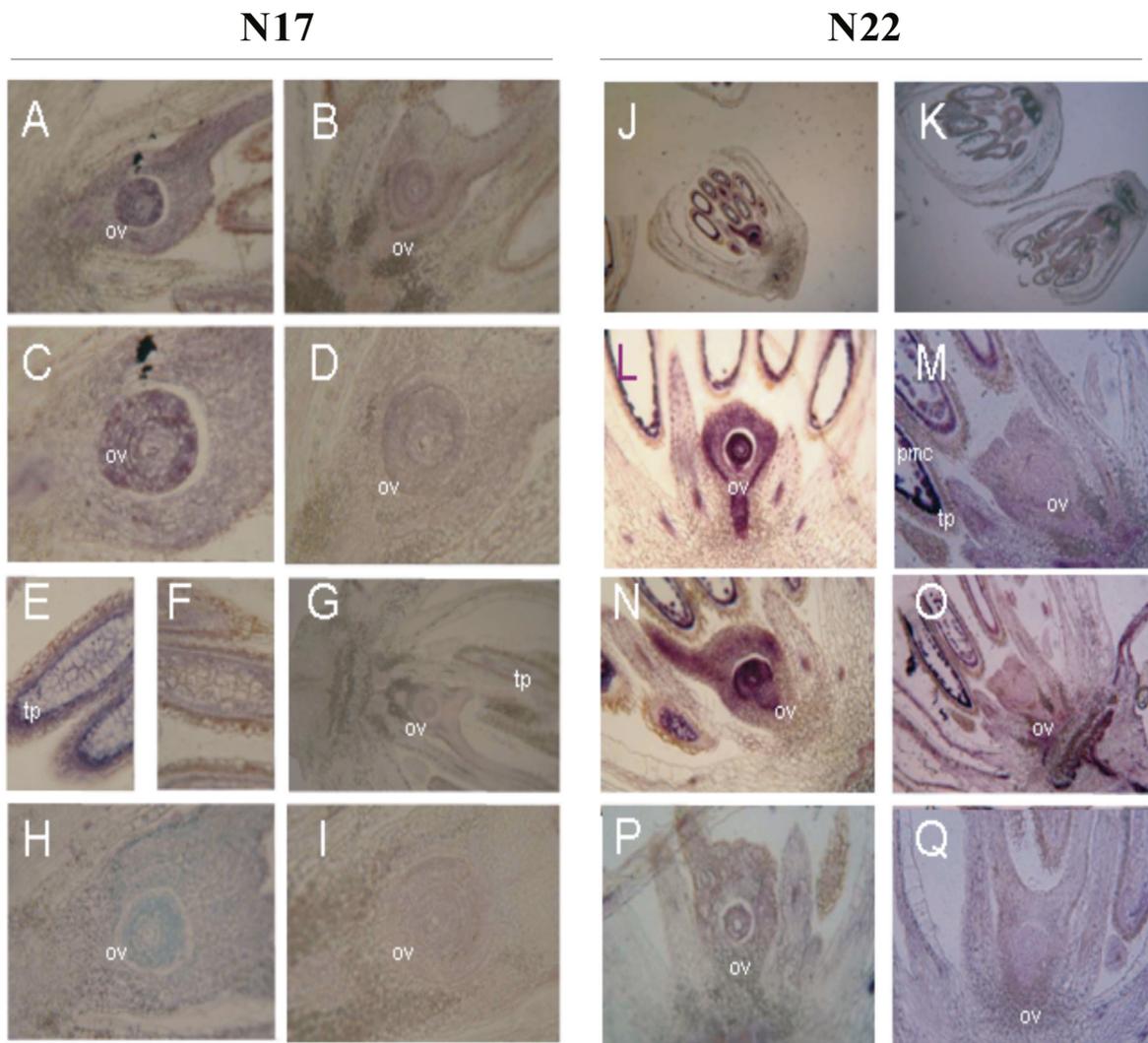
**Figura X:** Relaciones de expresión relativa de N17 y N22 analizadas por PCR en tiempo real en los genotipos Q4117 (apomítico) y Q4188 (sexual) de *P. notatum*. En el panel A se observan los patrones de expresión del transcrito N17 en diferentes estadios del desarrollo del ovario: premeiótico (0), pre-meiosis tardía/ meiosis (I/II), post-meiosis (IV/V) y antesis (IV). En el panel B se observan los perfiles de expresión del transcrito N22. Se estudiaron los estadios premeiótico (0), pre-meiosis tardía/meiosis (I/II) y post-meiosis (IV/V).

#### 4.5.1.4- Estudios de expresión por Hibridación *in situ* de tejidos

Se llevaron a cabo experimentos de expresión *in situ* de tejidos reproductivos con la finalidad de determinar la localización tisular de los transcritos N17 y N22, expresados en el genotipo apomítico (Q4117) y sexual (C4-4x). Estos análisis fueron realizados en colaboración con los Dres. Ana María González y Guillermo Seijo (IBONE, CONICET). Se utilizaron las sondas de ARNm sintetizadas a partir de ambas hebras de los fragmentos N17 y N22. Se examinaron cortes de inflorescencias inmaduras en estadio pre-meiótico tardío (I). Este estadio corresponde al período que precede inmediatamente la emergencia de la célula madre de la megáspora (MMC) en la nucela.

Los cortes del material incluido fueron hibridizados con las sonda antisentido (NcoI/SP6) y sentido (SalI/T7) las que permitieron la identificación de las secuencias complementarias en el transcriptoma de *P. notatum*. La sonda antisentido N17 (NcoI/SP6) originó una señal moderada en las células del tapetum de anteras y señal intensa en óvulos del genotipo sexual, mientras que no se detectó señal significativa en los órganos estudiados del genotipo apomíctico. La sonda sentido N17 (SalI/T7) generó una señal muy tenue y uniforme en los órganos de ambos genotipos (Figura XI, panel izquierdo). Estos resultados permitieron confirmar la expresión de N17 en tejidos reproductivos de *P. notatum*, indicando que la hebra sentido del retrotransposón putativo se encontraría expresada en anteras y óvulo del genotipo sexual.

De manera similar, la sonda antisentido de N22 (NcoI/SP6) originó una fuerte señal en anteras y en óvulos del genotipo sexual, mientras que se detectó señal significativa únicamente en anteras del genotipo apomíctico. La sonda sentido N22 (SalI/T7) generó una señal muy tenue y uniforme en los órganos de ambos genotipos (Figura XI, panel derecho). Estos resultados permitieron confirmar la expresión de la hebra sentido N22 en órganos reproductivos de ambos genotipos. Pero se identificó una importante diferencia entre ambos, observándose una marcada señal en el óvulo del genotipo sexual, la que no se detectó en el genotipo apomíctico.



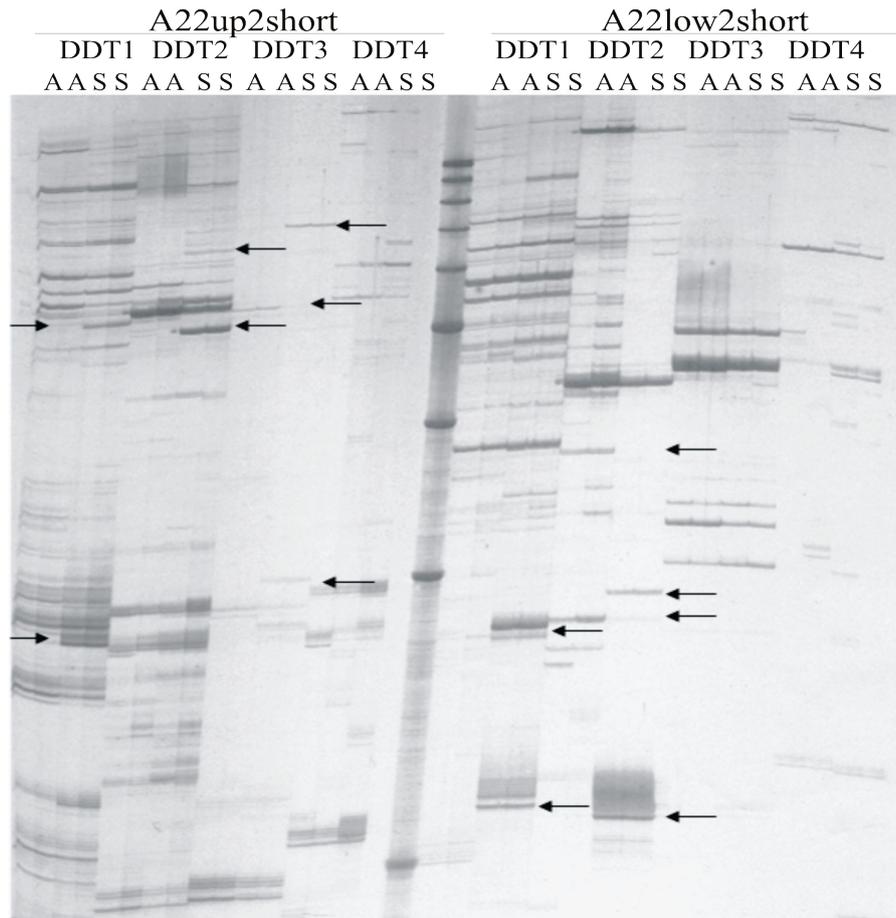
**Figura XI:** Fotografías de hibridización *in situ* de tejidos reproductivos de los genotipos sexual (C4=4x) y apomíctico (Q4117). Panel Izquierdo: hibridizaciones de las sondas originadas en N17. Panel derecho: hibridizaciones de las sondas N22. **A)** Genotipo sexual hibridizado con la sonda antisentido N17 NcoI, donde se detecta intensa señal en el óvulo. **B)** Genotipo apomíctico hibridizado con la sonda antisentido N17 NcoI, donde no se observa señal en el óvulo. **C)** Imagen aumentada del panel A. **D)** Imagen aumentada del panel B. **E)** Anteras del genotipo sexual hibridizado con la sonda antisentido N17 NcoI, donde se detecta señal moderada en células del tapetum. **F)** Anteras del genotipo apomíctico hibridizado con la sonda antisentido N17 NcoI, donde no se detecta señal. **G e I)** Genotipo apomíctico hibridizado con la sonda sentido N17 Sall, donde no se observa señal. **H)** Genotipo sexual hibridizado con la sonda sentido N17 Sall, donde no se detecta señal. **J)** Fotografía panorámica de tejido reproductivo del genotipo sexual hibridizado con la sonda antisentido N22 NcoI. **K)** Fotografía panorámica de tejido reproductivo del genotipo apomíctico hibridizado con la sonda antisentido N22 NcoI. **L y N)** Genotipo sexual hibridizado con la sonda antisentido N22 NcoI, donde se detecta intensa señal en el óvulo y anteras. **M y O)** Genotipo apomíctico hibridizado con la sonda antisentido N22 NcoI, donde se detecta señal intensa en anteras y no en el óvulo. **P)** Genotipo sexual hibridizado con la sonda sentido N22 Sall, donde no se detecta hibridización. **Q)** Genotipo apomíctico hibridizado con la sonda sentido N22 Sall, donde no se detecta señal. *Ov* óvulo, *tp* tapetum, *Pmc* células madres del polen.

#### 4.5.1.5- Display Diferencial Específico de Retrotransposones

En base a los resultados obtenidos en los experimentos de RACE para N22, donde se logró aislar un grupo de cinco transcritos relacionados (A43, A44, A45 y A46 y A47) que presentaron homologías significativas con transposones y tres de éstos contuvieron pequeños fragmentos de genes previamente asociados a la apomixis (Albertini *et al.*, 2004; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008), se construyó la hipótesis de que estos fragmentos serían miembros de una familia compleja de transcritos que se encontrarían sujetos a regulación diferencial durante el desarrollo de la apomixis. Con la finalidad de aislar y caracterizar miembros adicionales de la familia de transcritos de expresión diferencial entre plantas sexuales y apomícticas, se diseñó el experimento de Display Diferencial específico para retrotransposones homólogos de esta familia. El ensayo se llevó a cabo evaluando los genotipos Q4117 (apomíctico) y Q4188 (sexual) en base a la técnica original de Display Diferencial informada por Liang y Pardee (1992) pero incorporando cebadores específicos pequeños (15 nt) diseñados a partir de la secuencia original de N22 para amplificar secuencias relacionadas con N22.

Las amplificaciones se realizaron por duplicado y se visualizaron en geles de poliacrilamida 5% p/v (Figura XII). Se contabilizaron sólo las bandas duplicadas de la sección media del gel (mayor intensidad). De un total de 107 bandas, 48 resultaron polimórficas (44,85%) y la mayoría de estas bandas se originaron a partir de las amplificaciones del genotipo sexual (34 bandas, 77% de las bandas polimórficas).

Se logró aislar, clonar y secuenciar un total de 13 fragmentos polimórficos, cuyas secuencias fueron comparadas con las bases de datos públicas de los sitios de J. Craig Venter Institute (<http://www.tigr.org/db.shtml>), NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) y Gramene (<http://www.gramene.org/multi/blastview>). Todos los fragmentos presentaron sectores homólogos a elementos repetitivos. La mayoría de las bandas clonadas correspondieron a retrotransposones de tipo *Gypsy*. Diez de ellas incluyeron segmentos génicos transduplicados de longitud variable. La información relativa al tipo de elementos identificados y sus segmentos transduplicados contenido se presenta en la Tabla 6. Las secuencias se encuentran enumeradas en el Anexo I- Sección B.



**Figura XII:** Segmento de gel de poliacrilamida del experimento de Display Diferencial Específico de Retrotransposones. En la parte superior se detalla el nombre de los cebadores específicos utilizados (N22 upper short y N22 lower short), el nombre de cada cebador DDT (1, 2,3 y 4) y la nomenclatura de las plantas (A: Q4117 apomítica; S: Q4188: sexual).

Las secuencias ADD10-1, ADD12-2 y ADD5-2 presentaron sectores similares al retrotransposón Ty3 *Gypsy*, mientras que ADD10-1 además presentó un pequeño sector similar a una proteína tipo PDI (Proteína Disulfuro Oxidoreductasa) de maíz. Se conoce que estas proteínas pertenecen a la superfamilia de las tiorredoxinas (PF00085) y se encuentran asociadas al metabolismo de oxidoreducción celular.

Por otro lado, ADD7-1 presentó un sector similar a un retrotransposón tipo Copia RIRE7 y otro sector similar a un ARNm de *A. thaliana* con que compartió un 98% de identidad con una proteína POT (proteína Transportadora de Oligopéptidos dependiente de Protones) de la misma especie. Estas proteínas pertenecen a la superfamilia de transportadores generales de sustratos (InterPro:IPR016196) y se encuentran localizadas en la membrana plasmática.

Tres de las secuencias aisladas: ADD9-1, ADD16-1 y ADD17-2, presentaron altas homologías con proteínas del tipo LTP (Proteína Transportadora de Lípidos) específicas de anteras y las tres secuencias contuvieron además sectores similares a diferentes retrotransposones no caracterizados. Las proteínas LTP pertenecen a la familia AAI\_LTSS (PF00234) que involucra 188 miembros presentes en 266 especies de plantas superiores. Se conoce que esta familia proteica participa de las reacciones de defensa frente a patógenos, transporte lipídico a través de la membrana plasmática y almacenamiento de nutrientes.

Cinco de las secuencias aisladas: ADD12-1, ADD13-2, ADD2-2, ADD6-2 y ADD8-2 presentaron sectores similares a retrotransposones no caracterizados y cuatro de ellas contuvieron además sectores similares a secuencias proteicas. ADD13-2 presentó muy alta homología con una proteína hipotética de *O. sativa* y la información disponible en las bases de datos públicas permitió asociarla con un dominio de tipo syntaxina. Este tipo de proteínas participan de la conformación de receptores de membrana relacionados con el transporte vesicular intracelular. ADD2-2 presentó similitudes con un clon de ADNc de *Phyllostachys edulis* que resultó idéntico a una proteína de *O. sativa* que pertenece a la familia del complejo adaptador de clatrina. Estas proteínas cumplen un importante rol en la biosíntesis proteica ya que median el transporte desde el retículo endoplasmático al aparato de Golgi. La secuencia ADD6-2 presentó un sector similar a una proteína hipotética de *O. sativa*. Las indagaciones de bases de datos de dominios proteicos permitieron identificar un dominio de función desconocida tipo DUF1645 localizado en muchas proteínas de plantas superiores y que fue informada su funcionalidad en levaduras como parte de receptores de señales hormonales (AR781). La secuencia ADD8-2 presentó similitud con una proteína hipotética de *O. sativa* cuya función no se encuentra caracterizada.

Finalmente la secuencia ADD7-2 mostró un sector similar a secuencias tipo LTR (secuencias Repetitivas Terminales Largas) de un retrotransposón de localización centromérica del cromosoma 8 de arroz y en otro sector con una proteína hipotética de maíz que aumenta su expresión en condiciones de estrés por frío pero aún no se encuentra caracterizada estructuralmente.

**Tabla 6: Lista de elementos repetitivos (ER) relacionados con N22, identificados por Display Diferencial o por RACE**

Clon	ID del ER	Anotación	eV	ID de gen transduplicado <sup>a</sup>	Descríp. del gen transdup.	eV
N22	ORSiTERT00200152	Retrotransp. rir8 Ty3- <i>Gypsy</i>	0.97	-	-	-
A43	gb EU558521.1	Retrotransp. <i>Gypsy20</i>	1e <sup>-65</sup>	CAH56436.1	<i>serk2</i>	3e <sup>-28</sup>
A44	gb EU558521.1	Retrotransp. <i>Gypsy20</i>	8e <sup>-35</sup>	LOC100191539	<i>cyt P450</i>	8e <sup>-100</sup>
A45	gb EU558521.1	Retrotransp. <i>Gypsy20</i>	1e <sup>-16</sup>	-	-	-
A46	gb EU558521.1	Retrotransp. <i>Gypsy20</i>	2e <sup>-60</sup>	CAH56436.1	<i>serk2</i>	2e <sup>-54</sup>
A47	ZRSiTERTO00146	Retrotransp. dagaf_576C20-1	3e <sup>-8</sup>	-	-	-
ADD12-1	ORSiTERTO00318	Retrotransp. rn_128-73	7e <sup>-4</sup>	-	-	-
ADD13-2	ZRSiTERTO00278	Retrotransp. opie_af090447-2	0.9	NM_001054388	Proteína Hipotética	2e <sup>-61</sup>
ADD16-1	ZRSiTERTO00248	Retrotransp. grande_af546188-1	0.4	NM_001112535	Proteína LTP	2e <sup>-11</sup>
ADD17-2	gb AF187823	transposasa DOPD y DOPA	0.02	BT086830.1	Proteína LTP	2e <sup>-11</sup>
ADD2-2	gi 68445404	Transposón nDart1-201	0.75	emb FP097948	ARNm	8e <sup>-9</sup>
ADD5-2	ORSiTERT00200157	Retrotransp. Ty3- <i>Gypsy</i>	0.72	-	-	-
ADD6-2	ORSiTERTO00412	Retrotransposon	0.37	NM_001149478	Proteína Hipotética	0.003
ADD7-1	ORSiTERT00200004	Retrotransp. Copia RIRE 7	0.99	AK319168.1	ARNm	0.32
ADD7-2	ORSiCMCM00100019	LTR-Retrotransp. centromérico	0.55	AY108819.2	Proteína Hipotética	0.026
ADD8-2	ZRSiTERTO00235	Retrotransp. Ji5bl2	0.99	EU976081.1	Proteína Hipotética	2e <sup>-4</sup>
ADD9-1	ORSiTETNO00120	Transposón AnacB1	0.86	NM_001112535	Prot. LTP de anteras	2e <sup>-17</sup>
ADD10-1	ORSiTERT00200080	Retrotransp. Ty3 <i>Gypsy</i>	0.99	U90944.1	Proteína PDI	0.99
ADD12-2	ORSiTERT00200080	Retrotransp. Ty3 <i>Gypsy</i>	0.99	-	-	-

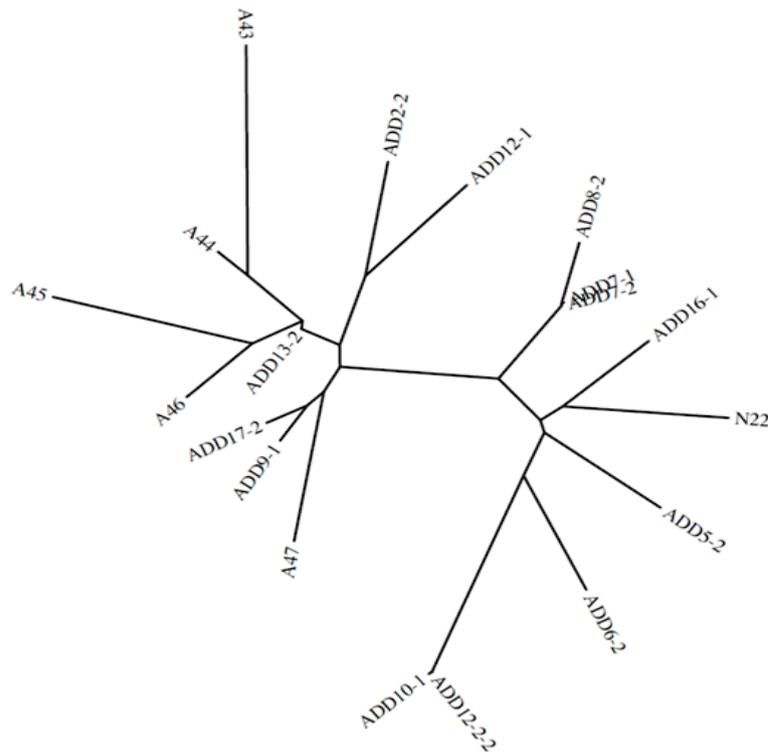
<sup>a</sup> Algunos de los elementos repetitivos contuvieron fragmentos génicos internos.

Estas observaciones sugieren que N22 es un miembro de una familia de retrotransposones relacionados, que tienen expresión diferencial en el genotipo apomítico y sexual. Si bien no todas las anotaciones funcionales pudieron ser asignadas, las proteínas codificadas por los segmentos transduplicados presentaron las siguientes funciones celulares: biosíntesis proteica, transporte proteico y lipídico, metabolismo de oxidoreducción y vías de transducción de señales hormonales, particularmente recepción de la señal.

#### **4.5.1.6- Caracterización de la familia de retrotransposones N22 mediante análisis filogenético**

Con la finalidad de profundizar el estudio de las relaciones filogenéticas en la familia N22, se analizaron los fragmentos correspondientes a elementos repetitivos de cada una de las secuencias aisladas por Display Diferencial Específico de Transposones. Para tal fin, se implementaron los programas de la plataforma Phylogeny (<http://www.phylogeny.fr/>). Se seleccionaron los programas T-Coffe y PHYLYP 3.66 estableciendo un número de 1000 repeticiones (bootstraps) (Figura XIII). Las secuencias empleadas para la construcción del árbol filogenético se encuentran detalladas en el Anexo I- Sección B.

Se pudieron identificar dos grandes grupos. El primero de ellos estaba compuesto por los elementos repetitivos del tipo Ty3-*Gypsy* (ADD5-2, ADD12-2 y ADD10.1), donde se incluyó la secuencia N22 y los siguientes retrotransposones no caracterizados: Ji5bl2(ADD8-2), Copia RIRE 7 (ADD7-1), LTR centromérico (ADD7-2), grande\_af546188-1 (ADD16-1) y el sector correspondiente a ADD6-2. En el segundo grupo se incluyeron las secuencias homólogas a retrotransposones de tipo *Gypsy*-20 (A43, A44, A45 y A46), secuencias correspondientes a transposones: transposasa DOPD y DOPA (ADD17-2), AnacB1 (ADD9-1), ADD2-2 (nDart1-201); y los siguientes retrotransposones no caracterizados: dagaf\_576C20-1 (A47), opie\_af090447-2 (ADD13-2), rn\_128-73 (ADD12-1). Sin embargo, se desconoce el significado biológica de este patrón de agrupamiento.



**Figura XIII:** Árbol filogenético donde se muestra el agrupamiento de fragmentos similares a elementos repetitivos asociados al modo de reproducción. Se diferencian dos grupos principales: uno relacionado con retroelementos Ty3 Gypsy, donde se encuentra N22, y otro con Gypsy 20.

#### 4.5.2- Estudio de transcritos similares a pre-miARN (N11 y N13)

##### 4.5.2.1- Estudios de expresión por PCR en tiempo real

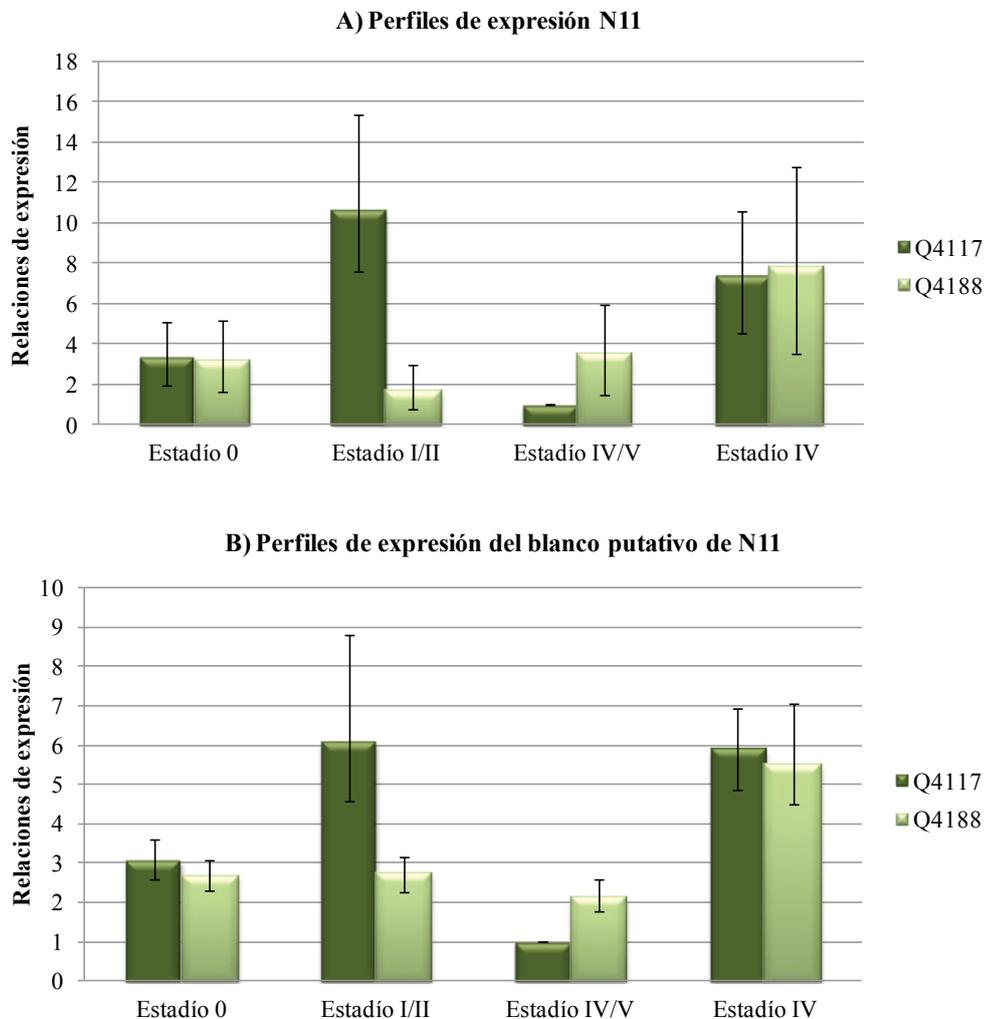
A partir de los resultados obtenidos en las búsquedas bioinformáticas de los transcritos extendidos de N11 y N13 se construyó la hipótesis de que estos fragmentos podrían corresponder a precursores de microARN. Se realizaron las predicciones bioinformáticas de sus respectivos ARNm blancos regulatorios, donde se asignó el gen Os03g0199100 para N11 y a BT068773.1 para N13 (Ver el punto 4 de esta sección)

Se llevaron a cabo experimentos de PCR en tiempo real con la finalidad de determinar los perfiles de expresión de los transcritos y sus blancos predichos durante el desarrollo reproductivo de *P. notatum* tanto en plantas sexuales como apomícticas. Para ello se utilizaron muestras de ARN extraídas de espiguillas colectadas en estadio pre-meiótico (0), pre-meiosis tardía/meiosis (I/II), post-meiosis (IV/V) y antesis (VI) según el calendario reproductivo reportado por Laspina *et al.* (2008). Además se diseñaron cebadores sobre las secuencias de ambos transcritos y sobre la secuencia de sus blancos

putativos. Estos últimos cebadores fueron diseñados de manera de abarcar el sitio de clivaje del posible blanco y así circunscribir los valores de cuantificación relativa únicamente a la secuencia blanco intacta. Se utilizaron dos controles de igual expresión: el gen de la Glucosa 6-Fosfato Deshidrogenasa (G6PdH) y el de la subunidad  $\beta$  de la Tubulina (Tub). Estos genes fueron reportados por otros grupos de trabajo como genes de expresión estable entre plantas sexuales y apomícticas en especies de gramíneas. G6PdH resultó ser uno de los más estables entre plantas de ambos modos de reproducción en la especie apospórica *Brachiaria brizantha* (Duarte Silveira *et al.*, 2009). Por otra parte, la subunidad  $\beta$  de la tubulina representa un sector muy conservado entre especies y fue utilizado como control de igual expresión entre plantas sexuales y apomícticas en *Poa pratensis* (Albertini *et al.* 2004). Se analizaron dos repeticiones biológicas de los genotipos Q4117 y Q4188 y todas las amplificaciones fueron realizadas por triplicado. Además se incorporaron controles negativos a los que no se les agregó el ADNc. El análisis estadístico se desarrolló con el programa REST-RG (Relative Expression Software Tool V 2.0.7 para Rotor Gene, Corvett Life Sciences).

Las relaciones de expresión fueron calculadas con respecto a la planta apomíctica en estadio post-miótico del genotipo apomíctico (el punto de menor expresión del experimento tanto para N11 como para N13). Se observó una expresión diferencial entre plantas apomícticas y sexuales y una notable coordinación de las relaciones de expresión para N11 y su gen blanco predicho. En el caso de N11 se observó que en el genotipo apomíctico, los niveles de expresión en estadio pre-meiótico se detectaron en el valor de 3,34 (error estándar: 5,087 - 1,95), aumentaron a 10,54 en meiosis (error estándar: 15,53 - 7,62), descendieron al valor de referencia 1 en post-meiosis y presentaron un nuevo pico de 7,37 (error estándar: 10,55 - 4,56) en estadio de antesis. Mientras que en el genotipo sexual, en estadio pre-meiótico, meiótico y postmeiótico, la relación de expresión se mantuvo en valores moderados/bajos y no mostraron diferencias significativas entre ellos (3,33 error estándar: 5,14 - 1,64; 1,7 error estándar 2,78-069; y 3,56 error estándar: 5,56 - 2,49 respectivamente), observándose un aumento a 7,84 en antesis (error estándar: 12,76 - 3,35). En el caso del gen blanco de N11 predicho, también se detectaron dos aumentos significativos de expresión, uno en el estadio meiótico de 6,07 (error estándar: 8,8 - 4,56) y uno en antesis de 5,88 (error estándar: 6,95 - 4,95), mientras que en el genotipo sexual, en estadio pre-meiótico, meiótico y postmeiótico, la relación de expresión se mantuvo en valores bajos y no mostraron diferencias significativas entre ellos (2,67 con un error

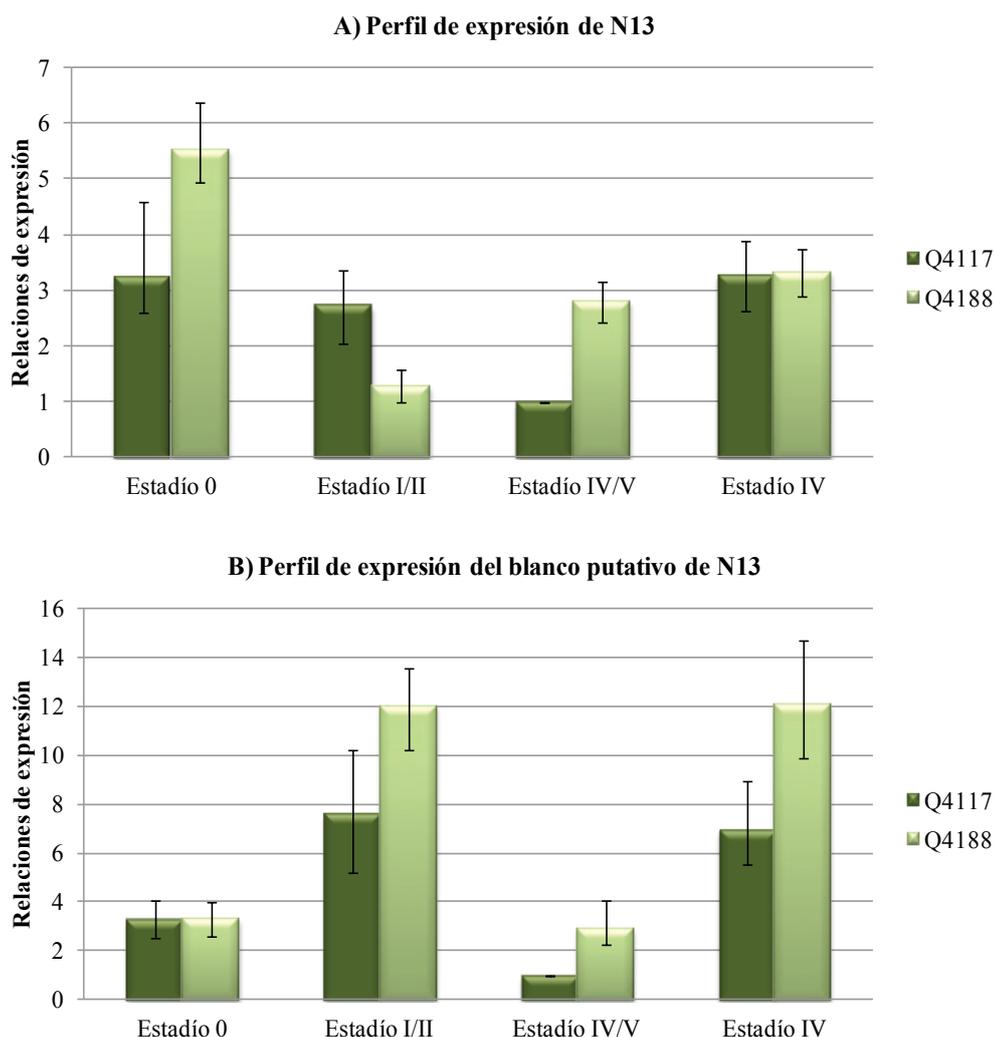
estándar de 3,09 - 2,29; 2,75 con un error estándar de 3,13 - 2,27; y 2,14 con un error estándar de 2,6 - 1,7, respectivamente), observándose un aumento significativo hasta 5,49 en antesis (error estándar: 7,04 - 4,48) (Figura XIV). Estos resultados sugieren que N11 se encontraría activado en el genotipo apomíctico durante el estadio meiótico y el de antesis, a diferencia de lo que ocurre en el genotipo sexual, donde la inducción se produce hacia el final del desarrollo reproductivo. Del mismo modo, el gen blanco predicho aumentaría su expresión en meiosis y antesis en el genotipo apomíctico y sólo en antesis en el sexual, mostrando una correlación completa con la expresión del precursor. Estos resultados podrían indicar que una disminución en el procesamiento del precursor se correlaciona con un aumento de la concentración del transcrito blanco intacto, sugiriendo una relación funcional entre ambos transcritos.



**Figura XIV:** Relaciones de expresión relativa analizadas por PCR en tiempo real en diferentes estadios del desarrollo del ovario: prememiótico (0), pre-meiosis tardía/ meiosis (I/II), post-meiosis (IV/V) y antesis (VI) en los genotipos Q4117 (apomíctico) y Q4188 (sexual) de *P. notatum*. En el panel superior se observan los perfiles de N11(A) y el panel inferior (B) el de su blanco predicho.

Por otra parte, en el caso de N13, los perfiles de expresión del precursor y el blanco putativo también resultaron coordinados. En el caso del genotipo apomítico, las relaciones de expresión de los estadios pre-meiosis, meiosis y antesis resultaron significativamente superiores al del estadio postmeiótico (3,23 con un error estándar de 4,46 - 2,6; 2,72 con un error estándar de 3,36 - 2,05; y 3,26 con un error estándar de 3,89 - 2,615, respectivamente). En el genotipo sexual se detectó un nivel de expresión de 5,5 en pre-meiosis (Error estándar: 6,36 - 4,94), un drástico descenso en meiosis a 1,28 (Error estándar: 1,58 - 1) y luego un aumento en post-meiosis a 2,79 (Error estándar: 3,16 - 2,61) cuya tendencia se mantuvo en estadio de antesis (3,32 con un error estándar de 3,74 - 2,89). En el caso del blanco predicho para N13, en el genotipo apomítico, se detectaron relaciones de expresión de los estadios pre-meiosis, meiosis y antesis que resultaron significativamente superiores a la del estadio postmeiótico (3,2 con un error estándar de 4,09 - 2,55; 7,74 con un error estándar de 10,24 - 5,17; y 6,87 con un error estándar de 8,92 - 5,55, respectivamente). Mientras que para el genotipo sexual, se detectaron aumentos significativos de expresión en meiosis y en antesis (12,0 y 12,1 con errores estándar de 13,53 - 10,22 y 14,66 - 9,92 respectivamente). Estos resultados sugieren que en el genotipo apomítico tanto N13 como su blanco predicho se encontrarían activados en pre-meiosis, meiosis y antesis disminuyendo su expresión en post-meiosis (Figura XV). En el genotipo sexual, N13 mantuvo una expresión elevada en pre-meiosis, post-meiosis y antesis, mientras el patrón de su blanco predicho presentó su mayor expresión en meiosis y en antesis.

Para el caso en N13, de manera análoga al de N11, también se observó que un aumento en la expresión del precursor se asocia a un aumento en la expresión del blanco en el estadio inmediato posterior y sugiere una relación funcional entre ambos transcritos.



**Figura XV:** Relaciones de expresión relativa analizadas por PCR en tiempo real en diferentes estadios del desarrollo del ovario: premeyótico (0), pre-meiosis tardía/ meiosis (I/II), post-meiosis (IV/V) y antrio (VI) en los genotipos Q4117 (apomítico) y Q4188 (sexual) de *P. notatum*. En el panel superior se observan los perfiles de N13(A) y en el panel inferior (B) de su blanco predicho.

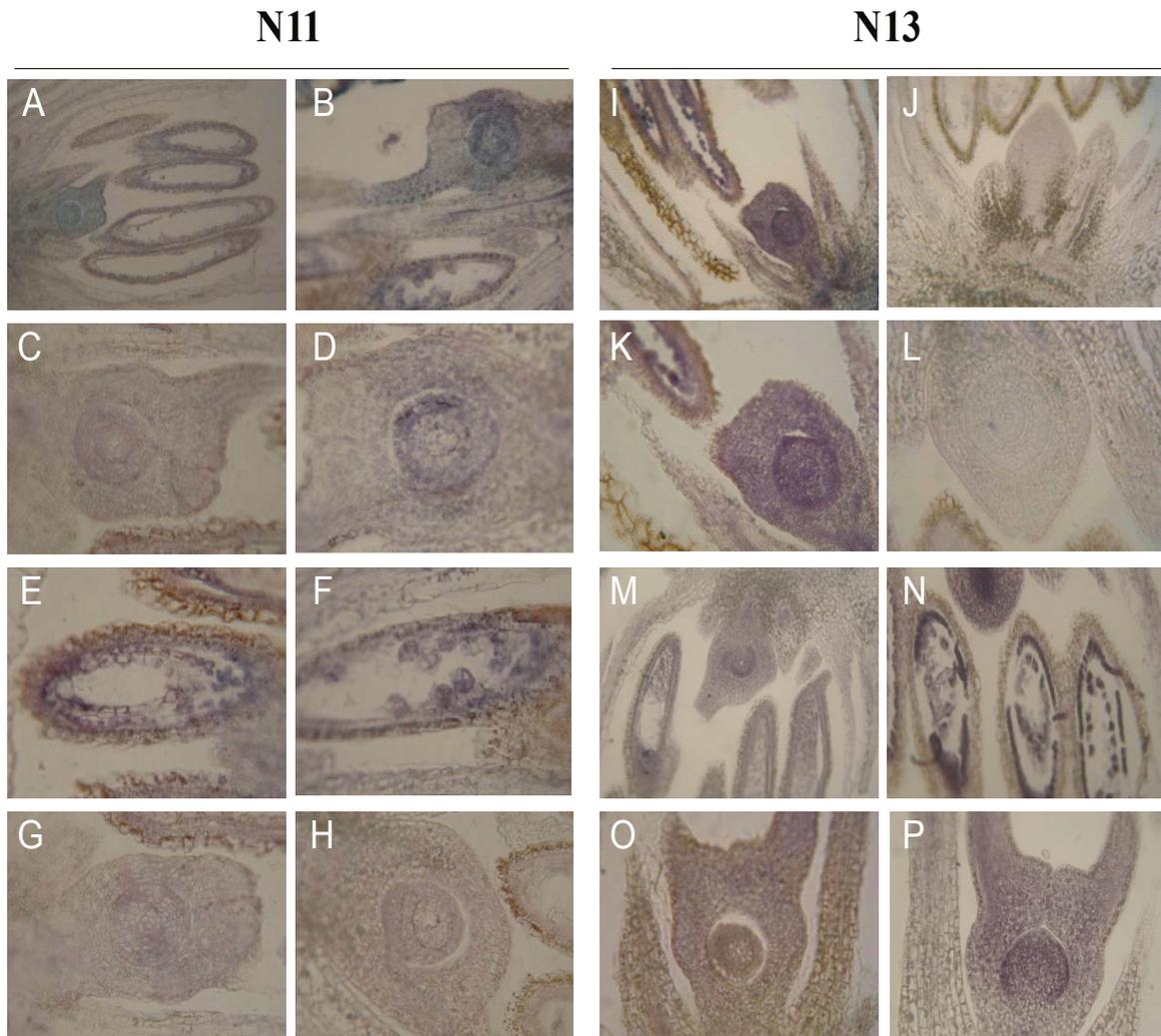
#### 4.5.2.2- Estudios de expresión por Hibridización *in situ* de tejidos

Se llevaron a cabo experimentos de expresión *in situ* de tejidos reproductivos con la finalidad de determinar la localización tisular de los transcritos N11 y N13 en el genotipo apomítico (Q4117) y sexual (C4-4x). Se utilizaron sondas de ARNm sintetizadas a partir de ambas hebras de los fragmentos. Se examinaron cortes de inflorescencias inmaduras en estadio pre-meiótico tardío (I), en el cual se había detectado la mayor diferencia de expresión entre ambos genotipos por PCR en tiempo real para N11 y N13.

Los cortes del material incluido fueron hibridizados con sondas (+) y (-). En este caso no utilizaremos los términos “sonda antisentido” y “sonda sentido”, ya que no se trata

de transcritos donde se puedan identificar regiones codificantes, como en el caso de los ARNm codificantes para proteínas y los retroelementos. Asignaremos arbitrariamente la denominación (+) a una de las sondas, y por consiguiente la denominación (-) a la complementaria. La sonda N11 (-) (NcoI/SP6) originó una señal muy pobre en los tejidos reproductivos de ambos genotipos. En cambio, la sonda N11 (+) (Sall/T7) generó una señal de intensidad media en el tapetum, células madres del polen (PMC) y en el ovario de ambos genotipos. Se detectó una señal levemente más intensa en el óvulo del genotipo apomítico en comparación con el genotipo sexual (Figura XVI, panel izquierdo). Estos resultados permitieron confirmar la expresión de N11 en el tejido reproductivo de *P. notatum*, siendo la hebra (-) del pre-miARN putativo la que se expresa en el óvulo del genotipo apomítico, en coincidencia con los resultados de plegamiento, que indican que son las hebras (-) las que presentan estructuras de tallo y burbujas.

De manera similar, la sonda (-) de N13 (NcoI/SP6) originó una señal uniforme y de intensidad media en anteras y en óvulos de ambos genotipos, mientras que la sonda (+) N13 (Sall/T7) generó una señal muy intensa en óvulos y anteras sólo en el genotipo sexual (Figura 4.13, panel derecho). Nuevamente, estos resultados permitieron confirmar la expresión de la hebra (-) N13 en órganos reproductivos de ambos genotipos, en coincidencia con las predicciones de plegamiento.



**Figura XVI:** Fotografías de hibridación *in situ* de tejidos reproductivos del genotipo sexual (C4-4x) y apomítico (Q4117). Panel Izquierdo: hibridaciones de las sondas N11. Panel derecho: hibridaciones de las sondas N13. **A)** Fotografía panorámica de tejido reproductivo del genotipo sexual hibridizado con la sonda N11 Sall, donde se detecta una señal media en el tapétum. **B)** Genotipo apomítico hibridizado con la sonda N11 Sall, se observa una señal moderada en tapetum e intensa en óvulo. **C)** Imagen del ovario de C4-4x hibridizado con N11 Sall, la señal es muy baja. **D)** Imagen aumentada del panel B donde se observa expresión en el óvulo del genotipo apomítico. **E)** Anteras de C4-4x hibridizado con la sonda N11 Sall, donde se detecta señal moderada en células del tapetum. **F)** Anteras de Q4117 hibridizado con la sonda N11 Sall, se detecta señal moderada en células del tapetum. **G)** Genotipo sexual hibridizado con la sonda N11 NcoI donde no se detecta señal. **H)** Genotipo apomítico hibridizado con la sonda N11 NcoI donde no se detecta señal. **I)** Genotipo sexual hibridizado con la sonda N13 Sall, se observa una señal muy intensa en óvulos y anteras, **J)** Genotipo apomítico hibridizado con la sonda N13 Sall, donde no se detecta señal. **K)** Corte de ovario del genotipo sexual hibridizado con la sonda N13 Sall, se observa una intensa señal en el óvulo. **L)** Genotipo apomítico hibridizado con la sonda N13 Sall, no se detecta señal. **N)** Anteras del genotipo sexual hibridizado con la sonda N13 Sall, se observa una intensa señal. **M, O y P)** Cortes de anteras y ovario del genotipo sexual hibridizado con la sonda N13 NcoI, donde se observa señal moderada en ambos tejidos.

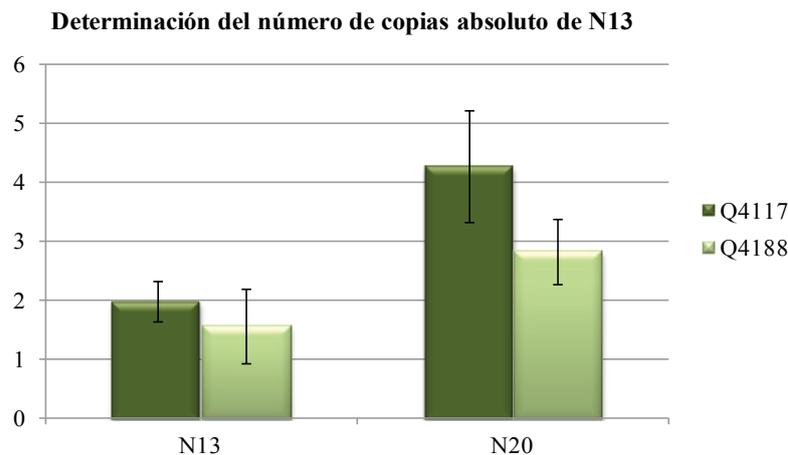
#### 4.5.2.3- Determinación del número de copias genómicas

Se llevó a cabo una estimación cuantitativa del número de copias genómicas de N13 en los genotipos Q4117 y Q4188 por medio de amplificaciones de PCR en tiempo real. Se utilizó el gen N20, que codifica para la proteína anclada a GPI LORELEI, como control de igual representación (Felitti *et al.*, 2011). Para el cálculo del número de copias genómicas por PCR en tiempo real se consideró el valor C informado para el genoma diploide, siendo 0,6 pg la masa por genoma monoploide de *P. notatum* (Jarret *et al.*, 1995). Se construyó la curva de calibración para N13 incluyendo 5, 50, 500 y 5000 copias de plásmidos recombinantes. Los valores de “take off” (despegue) de ambos genotipos fueron determinados por triplicado. Se analizaron 4 diluciones del ADN genómico y sólo fueron considerados los puntos abarcados en el rango lineal de las curvas para su posterior extrapolación. Los datos y reportes de cada amplificación fueron analizados con el programa Rotor-Gene Q Series Software 1.7.

Los experimentos permitieron detectar un promedio de número de copia por genoma monoploide de 1,98 y de 1,56 para los genotipos apomítico y sexual, respectivamente. Se estudió la significancia de la diferencia por medio de un test de Fiedman y no se obtuvieron diferencias significativas entre modos de reproducción ( $p=0,667 > \alpha=0,05$ ). Asimismo, no se detectaron diferencias significativas entre los genotipos para el gen de referencia N20 (Proteína anclada a GPI-LORELEI) ( $p=0,18 > \alpha=0,05$ ), donde los valores promedio por genoma monoploide fueron de 4,33 para Q4117 y de 2,83 para Q4188.

Estos resultados indican que el gen N13 se encuentra representado entre 1 y 2 veces en el genoma monoploide de *P. notatum*, y no se observan diferencias de representación entre el genotipo apomítico y el sexual (Figura XVII).

La estimación cuantitativa del número de copia genómico de N11 en los genotipos Q4117 y Q4188 por medio de amplificaciones de PCR en tiempo real no pudo ser concluida, ya que los valores estadísticos de ajuste a la recta de concentraciones conocidas del plásmido se encontraron fuera de los parámetros establecidos ( $R^2=0,43$ ). Esto pudo deberse a una falta de integridad del plásmido N11 o a su baja concentración.



**Figura XVII:** Estimación del número de copia por genoma monoploide en el genotipo apomítico (Q4117) y el genotipo sexual (Q4188) de *P. notatum*. N13 mostró una baja representación en el genoma. El gen de referencia N20 se encontró igualmente representado, con bajo número de copias, en ambos genotipos.

#### 4.5.2.4- Análisis 5' RACE de los genes blanco predichos

Con la finalidad de confirmar la naturaleza de los pre-miARN predichos (N11 y N13) y detectar la degradación específica del blanco regulatorio inducida por el miARN maduro correspondiente, se llevaron a cabo experimentos de 5' RACE a partir de dos bibliotecas no clonadas Marathon (Clontech) provenientes de los genotipos Q4117 (apomítico) y Q4188 (sexual). La validación completa de la naturaleza del miARN requiere la detección del blanco regulatorio predicho escindido entre el nucleótido 10 y 11 de la secuencia homóloga al respectivo miARN maduro, que se produce por la actividad de la proteína PAZ-PIWI ARGONAUTA durante el primer paso de la degradación (Rhoades *et al.*, 2002).

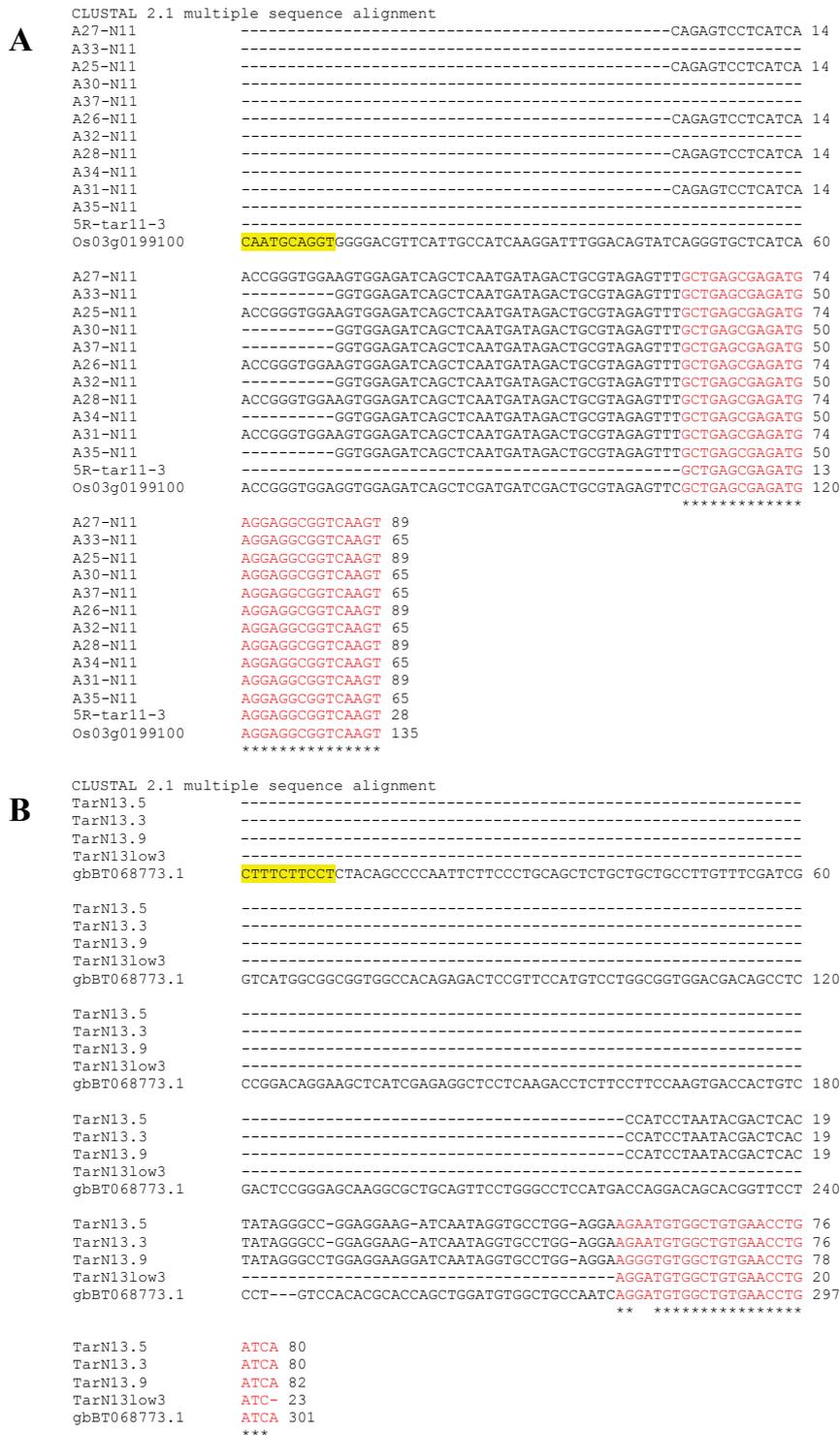
Para desarrollar el experimento se diseñaron tres cebadores anidados sobre las secuencias blanco putativas, de manera de aumentar la especificidad de la detección. Se llevaron a cabo amplificaciones por duplicado para cada ARNm en ambas bibliotecas, utilizando el cebador más externo y el cebador específico de adaptador Marathon: AP2. Posteriormente se realizó una segunda reacción de PCR a partir del producto de amplificación anterior con el cebador de posición intermedia y AP2, seguida de una tercera amplificación con el cebador interno y el cebador de adaptador AP2. Los productos de amplificación obtenidos a partir de esta última reacción fueron separados en geles de acrilamida 5% y revelados por tinción con plata. Se identificaron varias bandas para cada

transcripto, las que fueron clonadas y secuenciadas. Las secuencias fueron analizadas y comparadas mediante diferentes programas de alineamiento.

Para el ARNm blanco putativo de N11 (Os03g0199100) únicamente se aislaron bandas a partir de la biblioteca de ADNc proveniente del genotipo apomítico (Q4117). Se secuenciaron múltiples bandas de menor tamaño que el esperado y se identificaron dos productos de degradación que difirieron entre ellos en su longitud (66 y 89 nt). Específicamente estos fragmentos se encontraron a 69 nt y 46 nt del sitio de escisión primario de ARGONAUTA. Es interesante destacar que ambas longitudes difieren en 23 nt, fragmento compatible con un procesamiento por DICER y que se encuentran distanciados del sitio predicho de corte en un múltiplo de esa cantidad (Figura XVIII). Si bien no fueron detectados los productos primarios de la degradación, estos fragmentos podrían corresponder a productos secundarios procesados por enzimas tipo DICER (Berstein *et al.*, 2001; Lee *et al.*, 2003a). Una hipótesis que podría plantearse a partir de lo observado es que al alcanzarse el equilibrio cinético de la reacción de degradación, el primer producto se encontraría menos representado que los productos secundarios y que el blanco regulatorio estaría siendo procesado por una enzima tipo DICER.

En el caso de N13, únicamente pudieron ser aislados fragmentos provenientes de la biblioteca de ADNc del genotipo sexual (Q4188). Aquí, tres secuencias correspondieron a secciones del ARNm blanco esperado (gb|BT068773.1) (tar13-9, tar13-5 y tar13-3). Si bien se detectó el sector correspondiente al último cebador anidado, los fragmentos de menor tamaño no alcanzaron el sitio de corte predicho (Figura XVIII). Sin embargo, pudo observarse que todos los fragmentos aislados presentaron la misma longitud y difirieron en 221 nucleótidos del sitio de corte predicho. Comprendiendo que 220 es múltiplo de 22 y haciendo analogía con lo observado para N11, estos fragmentos podrían ser productos de degradación procesados por enzimas análogas a DICER.

Nuestros resultados sugieren que los transcritos N11 y N13 estarían siendo clivados en sitio correspondiente dentro del sector complementario al miARN maduro y luego estarían siendo procesados en fragmentos regulares de 23 o 22 nt por una enzima análoga a DICER.



**Figura XVIII:** Alineamiento de las secuencias aisladas de los ARNm correspondientes a los genes blanco predichos. En el panel superior (A) se observa el alineamiento de las 11 secuencias clonadas a partir de la biblioteca del genotipo apomítico, homólogas al gen blanco predicho de N11 (Os03g0199100). En el panel inferior (B) se observa el alineamiento de las 3 secuencias correspondientes al ARNm blanco predicho de N13 ([gb|BT068773.1](#)). Los asteriscos debajo de las secuencias señalan el sector del cebador. Las bases subrayadas indican las mutaciones puntuales entre el gen de *O. sativa* y de *Z. mays* y el de *P. notatum*. El sector del miARN se señala en color amarillo sobre la secuencia de los genes blanco predichos.

## 4.6- Análisis bioinformático del resto de los transcriptos diferenciales

### 4.6.1- Estudio de los transcriptos diferenciales por modo de reproducción

Del total de 22 transcriptos diferenciales por modo de reproducción aislados inicialmente por Display Diferencial (Laspina *et al.*, 2008), uno de ellos presentó similitud con secuencias proteicas (N116), 8 pudieron ser extendidos mediante 5' y 3' RACE (N7, N11, N13, N22, N43, N56 y N104) y 4 fragmentos no pudieron ser amplificados en el genoma de *P. notatum* (N2, N49, N86 y N103) (Ver puntos 2 y 3 de esta sección). Estos últimos fueron clasificados como posibles contaminaciones. El resto de los transcriptos (N38, N58, N70, N77, N82, N87, N99, N106 y N119) fueron analizados mediante búsquedas bioinformáticas exhaustivas en las bases de datos de elementos repetitivos de gramíneas (TIGR) y de precursores de miARN (Sanger). Además se indagó la presencia de dominios conservados en las bases de datos ProSite y Pfam (Sanger) y se realizaron las predicciones de blancos regulatorios en los casos de las secuencias similares a precursores de miARN.

Los estudios realizados con las 9 secuencias, permitieron detectar que cinco de ellas (N58, N70, N82, N99 y N119) resultaron similares a elementos repetitivos. Dos incluían fragmentos de genes transduplicados, según se deducía de las homologías con secuencias de especies relacionadas (Tabla 7). N58 (208 nt) presentó sectores similares a múltiples regiones genómicas de especies relacionadas de plantas en las búsquedas generales en NCBI y TAIR (eV~0,3). Las comparaciones con el genoma de arroz permitieron identificar 68 secuencias similares esparcidas en todos los cromosomas, muchas de las cuales correspondieron a retrotransposones y otras a regiones intergénicas. Las búsquedas en las bases de datos de elementos repetitivos permitieron identificar similitud de secuencia a lo largo de 110 nt (53% de la secuencia) con el retrotransposón *opie\_092E12-1* de maíz. Asimismo, durante las búsquedas realizadas en el sitio Gramene, para el clon N70 (840 nt), se identificaron 15 pequeños sectores similares en los cromosomas 7 y 8 de arroz los que correspondieron mayoritariamente a sectores intergénicos. Las comparaciones con las bases de datos de elementos repetitivos de gramíneas permitieron identificar similitudes a lo largo de 143 nt con un retrotransposón Ty3 *Gypsy* de arroz. No se detectaron similitudes de relevancia en el sitio NCBI, TAIR y Sanger. La secuencia N82 (552 nt) presentó homologías en las bases de datos de elementos repetitivos de gramíneas con el retrotransposón *milt\_573L14-1* a lo largo de 332 nt (desde 169 a 491) y el sector inicial de

la secuencia (posición 1-169) con un ARNm no caracterizado de maíz. El clon N99 (740 nt) mostró similitudes en NCBI (posición 2-246) con una secuencia no caracterizada de un retrotransposón de maíz y en 159 nt (posición 308-467) con una secuencia proteica de *O. sativa*. Esta proteína corresponde a un factor de transcripción de tipo GRAS (PF03514) que se encuentra relacionado con procesos del desarrollo. Por otro lado, para N119 (223 nt), los resultados más significativos fueron obtenidos mediante las búsquedas en las bases de datos TIGR donde se identificó una secuencia similar al gen de la transposasa de un transposón de tipo Rim 2-M47 de *O. sativa* a lo largo de toda su secuencia. A su vez se identificaron dos secuencias similares del cromosoma 5 de arroz durante las búsquedas en el sitio Gramene, ambas localizadas en regiones intergénicas.

Por otra parte, una (1) de las 9 secuencias estudiadas resultó similar a un pre-miARN (Tabla 7). Las búsquedas en las bases de datos de NCBI para N87 (970 nt) permitieron identificar un pequeño sector conservado (desde 1 a 38) similar a ARN mensajeros no caracterizados de varias especies de plantas. Las comparaciones con pre-miARN permitieron detectar similitud con una secuencia de *A. thaliana*. Por el contrario, las búsquedas en las bases de datos de elementos repetitivos de gramíneas y en los sitios TAIR y Gramene no arrojaron similitudes significativas. Se estudiaron las posibles estructuras secundarias con el programa mFold 2.3 y se implementaron las estrategias de predicción de ARNm blanco regulatorio. Si bien se obtuvieron plegamientos espontáneos para ambas hebras, (+) y (-) ( $dG = -276$ ), el sector conservado no se localizó en una estructura típica de horquilla, en cambio, en todas las estructuras estudiadas se halló en medio de una burbuja. Se analizó la existencia de posibles blancos regulatorios del sector conservado por medio del programa miRU, sin embargo todos los resultados obtenidos correspondieron a ARNm que contaban con al menos 8 desapareamientos de bases entre el miARN maduro y su blanco predicho. Por este motivo, las energías de hibridación calculadas con el programa RNAhybrid resultaron inferiores al 50% de la energía de complementariedad perfecta. En este caso la predicción del pre-miARN resultó negativa considerando las especificaciones planteadas por Schwab *et al.* (2005) para tal fin.

Finalmente, 3 de las 9 secuencias estudiadas no mostraron similitudes relevantes en las bases de datos de plantas (Tabla 7). Las búsquedas realizadas con N38 (1211 nt), N77 (276 nt) y N106 (461 nt) no permitieron identificar homologías significativas con genes codificantes en los sitios NCBI, TAIR y Gramene. Las comparaciones con las bases de elementos repetitivos de gramíneas y de precursores de miARN no arrojaron resultados

significativos manteniendo los parámetros preestablecidos en los programas de las plataformas utilizadas. En búsquedas extensivas realizadas contra las bases de datos de todas las especies disponibles en NCBI se detectó muy alta homología para N106, a lo largo de toda la secuencia, con una proteína de *Pseudomonas syringae* (YP\_235263.1, eV:  $1e^{-43}$ ). Esto indicaría que el clon N106 podría corresponder a una contaminación. Para confirmar este resultado, correspondería indagar la existencia del fragmento en el genoma de *P. notatum* por medio de amplificaciones de PCR.

**Tabla 7: Estudios bioinformáticos de las secuencias expresadas diferencialmente entre plantas sexuales y apomíticas aisladas por Display Diferencial**

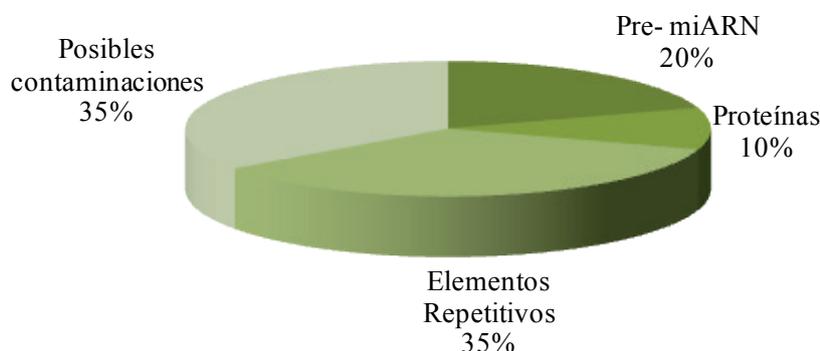
Clon	ID principal	Anotación principal	eV	ID de gen	Anotación del gen <sup>(1)</sup>	eV	Conclusión
N38	-	-	-	-	-	-	Posible contaminación
N58	ZRSiTERTOOT00251	Retrotransposón opie_092E12-1	0,5	-	-	-	Posible retrotransposón de gramíneas
N70	ORSiTERT00200156	Retrotransposón Ty3-Gypsy	0,7	-	-	-	Posible Retrotransposón
N82	ZRSiTERTOOT00248	Retrotransposón milt_573L14-1	$6,3e^{-5}$	gb EU940865.1	ARNm de <i>Zea mays</i>	$5e^{-4}$	Posible Retrotransposón con ARNm transduplicado
N86	-	-	-	-	-	-	Posible contaminación
N87	ath-MIR169k	Precursor de miARN	0,41	-	-	-	Posible pre-miARN, no confirmado
N99	emb AJ309824.2	Retrotransposón de <i>Zea mays</i>	$8e^{-3}$	Os02g0681900	Factor de transcripción GRAS	0,3	Posible retrotransposón con ARNm transduplicado
N106	YP_235263.1	Proteína bacteriana	$1e^{-43}$	-	-	-	Posible Contaminación
N119	gi 38678182.1	Transposasa de Rim2-M47	0,01	-	-	-	Posible transposón

<sup>(1)</sup> Se detalla la identidad de los sectores homólogos internos.

A modo de resumen, del total de 22 transcriptos diferenciales por modo de reproducción aislados inicialmente por Display Diferencial (Laspina *et al.*, 2008), uno de ellos presentó similitud con secuencias proteicas (N116), 8 pudieron ser extendidos mediante 5' y 3' RACE (N7, N11, N13, N22, N43, N56 y N104) y 4 fragmentos fueron clasificados como posibles contaminaciones (N2, N49, N77 y N103). De las secuencias

amplificadas por 5' y 3' RACE, 3 de ellas resultaron similares a precursores de miARN de plantas (N11, N13 y N56), una resultó similar a una secuencia proteica (N104) y dos resultaron similares a elementos repetitivos (N17 y N22), uno centromérico y otro perteneciente a una familia compleja de retrotransposones que transportan segmentos génicos. Los 9 transcritos restantes (N38, N58, N70, N82, N86, N87, N99, N106 y N119) fueron comparados con las bases de datos. Cinco de estas secuencias (N58, N70, N82, N99 y N119) resultaron similares a elementos repetitivos de las cuales dos incluyeron fragmentos génicos (Tabla 7), uno correspondiente a un ARNm no caracterizado de maíz (N82) y otro a un factor de transcripción de tipo GRAS relacionado con procesos del desarrollo. Una de las 9 secuencias (N87) resultó similar a un pre-miARN, sin embargo la predicción *in silico* no resultó positiva. Finalmente, 3 de las 9 secuencias estudiadas no mostraron similitudes relevantes en las bases de datos de plantas. Sin embargo, para inferir funcionalidad significativa se consideró el criterio utilizado por Larrinua y Belmar (2008). Estos autores postulan que si bien valores de eV de 0,01 pueden indicar la funcionalidad de la secuencia incógnita, generalmente es necesario obtener valores del orden de  $e^{-10}$  o menores para que estas similitudes puedan ser consideradas como funcionalmente significativas en ausencia de informaciones adicionales o de confirmaciones experimentales (Larrinua and Belmar, 2008). En este trabajo se trató como similares a los alineamientos que arrojaron eV mayores al orden  $e^{-10}$  y como homologías con asignaciones de funcionalidad positiva a los alineamientos que arrojaron eV del orden de  $e^{-10}$  o menores. Así, de las 22 secuencias diferenciales según el modo de reproducción, pudo atribuirse asignación funcional positiva a 6 de ellas, N104 y N116 homólogas a secuencias proteicas no caracterizadas de especies relacionadas, N11 y N13 como precursores de miARN y N17 y N22 como elementos repetitivos de *P. notatum*. La información hallada para cada una de las 22 secuencias incógnitas diferenciales según el modo de reproducción se encuentra compilada en la sección A del Anexo II. La figura XIX constituye una representación esquemática proporcional de los resultados de la clasificación preliminar de todas las secuencias en estudio asociadas a la ocurrencia de la aposporía.

**Gráfico proporcional de las funciones putativas asignadas a los transcritos diferenciales por modo de reproducción**



**Figura XIX:** Clasificación preliminar del subgrupo de secuencias asociadas al modo de reproducción, aisladas inicialmente por Laspina *et al.* (2008), que no habían podido ser clasificadas.

#### 4.6.2- Estudio bioinformático de los transcritos diferenciales por nivel de ploidía

Del total de 36 transcritos que carecían de anotación, diferenciales por nivel de ploidía entre plantas diploides y tetraploides de *P. notatum* aislados inicialmente por Display Diferencial (Martelotto *et al.*, 2005), 15 (L20, L24, L47, L53, L49, L66, L75, L76, L87, L94, L95, L774, L802, L854, L884) presentaron homologías con secuencias génicas de especies relacionadas durante las búsquedas iniciales. Estos genes correspondieron principalmente a proteínas hipotéticas sin función caracterizada de maíz y arroz (Tabla 4). Además, uno de los transcritos (L832) presentó homologías significativas lo largo de toda su secuencia con vectores de clonado (vector pBR322 gnl|uv|J01749.1:1-4361-49, eV: 2e-51). Y los 20 transcritos restantes permanecieron sin anotación funcional.

Con la finalidad de incrementar la caracterización y detectar patrones o dominios conservados en las secuencias proteicas, se desarrollaron búsquedas exhaustivas en las bases de datos ProSite y Pfam y de dominios conservados de NCBI. Los datos obtenidos en las búsquedas fueron resumidos en la Tabla 8. Tres de las secuencias (L75, L87 y L802) contaron con asignación funcional específica en sus proteínas homólogas. La secuencia L75 (366nt) presentó homologías con una proteína de tipo hidrolasa de enlaces éster de *O. sativa* donde se identificó el dominio perteneciente a la superfamilia de fosfoesterasas (cl10627) involucrado en el catabolismo de los ácidos grasos. La secuencia L87 (176 nt) mostró homologías con la subunidad 1 de la enzima NADHdH de *Bambusa olshamii*. La NADH: ubiquinona oxido reductasa (NADHdH: IPR00164) es una proteína de la cadena respiratoria que interviene en los procesos de transferencia de protones a través de la

membrana plasmática. Por otra parte, la secuencia L802 (158nt) resultó similar a una proteína de maíz de tipo ankirina kinasa perteneciente a una familia de metiltransferasas putativas (DUF248) donde se identificó además un dominio metiltransferasa dependiente de S-adenosil metionina (cl12011). Este tipo de enzimas se encuentran involucradas en la regulación de la expresión génica por medio de mecanismos epigenéticos.

Se identificaron dominios conservados en las proteínas hipotéticas homólogas de 5 de los fragmentos: L20 (372 nt), L66 (234 nt), L95 (241 nt), L844 (133 nt) y L854 (110 nt) (Tabla 8). En el primer caso (L20) correspondió al dominio catalítico Ser/Thr kinasa perteneciente a la superfamilia PKC (cl09925). También, L66 presentó un dominio de tipo bucle-A, uno de unión a ATP, dos sitios de unión a sustrato y un dominio tipo PKc, propios de proteínas tipo kinasa (cd00180). Las enzimas kinasas catalizan la transferencia del grupo fosforil gama del ATP a grupos oxidrilo de sustratos específicos, como son los restos aminoacídicos Serina, Treonina y Tirosina. La proteína hipotética homóloga a L95 contuvo un dominio de tipo ubiquitina (cl00155) la cual participa directamente en el catabolismo proteico. Por otro lado se identificó un dominio conservado para L844: Josephin (cl02099) y dos tipos de dominios conservados para L854: DPBB\_1 (cl04011) y polen\_alerg\_1 (cl08320) los cuales no contuvieron ninguna anotación funcional asociada.

No se identificaron posibles patrones o dominios conservados en el resto de las secuencias génicas similares/homólogas a los transcritos L24, L47, L49, L53, L76, L94 y L774.

**Tabla 8: Transcritos diferenciales por nivel de ploidía que presentaron homologías con proteínas**

Clon <sup>(1)</sup>	ID	Anotación principal	Características	eV
L20 (4x) <sup>(2)</sup>	gb EU957444.1	Proteína hipotética de maíz	Dominio Ser/Thr kinasa. Superfam. PKC (cl09925)	8e <sup>-20</sup>
L24 (2x) <sup>(2)</sup>	gb EU966467.1	Proteína hipotética de maíz	-	3e <sup>-55</sup>
L47 (2x) <sup>(2)</sup>	dbj AK224581	cDNA <i>Oryza officinalis</i>	ARNm expresado en estadíos tempranos en panoja	4e <sup>-10</sup>
L49 (4x) <sup>(2)</sup>	gb EU951333.1	mRNA de proteína de maíz	-	1e <sup>-21</sup>
L53 (2x)	gb EU964722.1	Proteína hipotética de maíz	-	1e <sup>-5</sup>
L66 (4x)	ref NM_001139432.1	mRNA de proteína de maíz	4 dominios catalíticos de tipo PKC (cd00180)	1e <sup>-7</sup>

<b>L75 (2x)<sup>(2)</sup></b>	Os03g0826600	Proteína de <i>Oryza sativa</i> tipo fosfoesterasa	Superfam. Fosfoesterasas (cl10627)	2e <sup>-26</sup>
<b>L76 (2x)<sup>(2)</sup></b>	gb EU970386.1	Proteína hipotética de maíz	-	1e <sup>-18</sup>
<b>L87 (2x)<sup>(2)</sup></b>	gb EU365401.1	NADHdH de <i>B. oldhamii</i>	NADHdH expresada en mitocondria	1e <sup>-54</sup>
<b>L94 (2x)<sup>(2)</sup></b>	gb BT018512.1	mRNA de proteína de maíz	-	2e <sup>-24</sup>
<b>L95 (2x)<sup>(2)</sup></b>	gb BT038322.1	Proteína hipotética de maíz	Proteína tipo Ubiquitina (cl00155)	7e <sup>-49</sup>
<b>L774(4x)<sup>(2)</sup></b>	LOC100192571	mRNA de proteína de maíz	-	2e <sup>-10</sup>
<b>L802 (2x)</b>	gb EU975659.1	mRNA de ankirina kinasa de maíz	Presenta un dominio Ankirin K y uno metiltransferasa (PF03141)	9e <sup>-6</sup>
<b>L844(4x)<sup>(2)</sup></b>	gb EU960848.1	mRNA de proteína de maíz	Familia Josephin (PF02099)	1e <sup>-15</sup>
<b>L854(4x)<sup>(2)</sup></b>	Os02g0267000	Proteína de <i>Oryza sativa</i>	Dominios conservados: DPBB_1(cl04011 y polen_alerg_1 (cl08320)	2e <sup>-23</sup>

<sup>(1)</sup>Se detalla el genotipo del cual fue aislado cada uno de los clones por DD, 2x: diploide, 4x: tetraploide.

<sup>(2)</sup>Asignaciones funcionales positivas (eV del orden de e-10) de los transcritos diferencialmente expresados.

Los 20 transcritos restantes que no habían resultado similares a secuencias proteicas (L14, L29, L45, L48, L55, L61, L72, L73, L86, L86, L90, L784, L794, L812, L812-1, L834-1, L834-3, L852, L852-1, L884) fueron analizados mediante búsquedas bioinformáticas exhaustivas en las bases de datos de elementos repetitivos de gramíneas (TIGR) y de precursores de miARN (Sanger). Además se indagó la presencia de dominios conservados en las bases de datos ProSite y Pfam (Sanger) y se realizaron las predicciones de blancos regulatorios en los casos de las secuencias similares a precursores de miARN (Tabla 9).

Ocho secuencias mostraron similitudes con elementos repetitivos de plantas (L48, L55, L61, L85, L90, L794, L852 y L884). Solamente una secuencia, L884 (106 nt) presentó homologías significativas (eV: 2e<sup>-30</sup>) con elementos repetitivos de gramíneas: la inserción del transposón *Gypsy* 21 de *A. thaliana*. Otros 6 transcritos (L48, L61, L85, L90, L794 y L852 de 191, 91, 229, 197, 73 y 105 nt respectivamente) presentaron similitudes en toda su longitud con elementos repetitivos de plantas, sin embargo los valores esperados no resultaron suficientemente negativos. L48 resultó similar al retrotransposón de maíz *ji\_17BA11-1*, L85 al retrotransposón Josephine de *O. sativa* y L90 al sector codificante para los genes Gag y Pol del retrotransposón Ty3 *Gypsy* de *Hordeum*

*vulgare*. También, se identificaron tres fragmentos similares a secuencias repetitivas: L61 resultó similar a repeticiones invertidas de elementos transponibles de maíz, L794 a repeticiones trinucleotídicas de la misma especie y L852 a repeticiones del transposón p-SINE de *O. sativa*. Una secuencia (L55, 173 nt) mostró similitudes con un retrotransposón en el 45% de su secuencia (posición 2-67) e identidad completa a lo largo de 26 nt con una proteína de *A. thaliana* perteneciente a la familia F-box (posición 96-121), lo que parecería corresponder a un retrotransposón portador de ARNm transduplicado (Tabla 9).

Ocho transcriptos resultaron similares a precursores de miARN durante las búsquedas realizadas en las bases de datos de Sanger (L14, L29, L45, L72, L73, L812, L834-1 y L834-3 de 512, 287, 290, 288, 431, 133, 254 y 353 nt respectivamente). Se estudiaron las posibles estructuras secundarias con el programa mFold 2.3 y se implementaron las estrategias de predicción de ARNm blanco regulatorio. Las búsquedas en las bases de datos de NCBI permitieron identificar pequeños sectores conservados (desde 1 a 38) similares a ARN mensajeros de varias especies de plantas para 7 de ellos (L14, L29, L72, L73, L812, L834-1 y L834-3). L43 no mostró similitudes con secuencias génicas, aunque la estructura secundaria de la hebra (-), obtenida mediante el programa mFold, resultó estable ( $\Delta G = -163$ ).

Se identificó un sector conservado de 27 nt para L14, correspondiente a una proteína de tipo aspartil proteasa de *A. thaliana* (identidad 25/27 nt, 92%). Las estructuras secundarias adoptadas por el transcripto resultaron estables ( $\Delta G = -163$ ) pero sólo el plegamiento (-) contuvo el sector conservado en una estructura tipo horquilla con cada una de las burbujas características de pre-miARN de plantas. Los estudios realizados con el programa miRU permitieron identificar la secuencia de *A. thaliana* (At2g42980.1: proteína aspartil proteasa con dominio aspartil proteasa PF00026) que cumplió con cada uno de los parámetros establecidos por Schwab et al. (2005) para la predicción de blancos regulatorios de miARN. Se detectaron sólo dos desapareamientos en las posiciones 3 y 16 del miARN maduro y la energía de hibridización resultó del 81,2% de la energía de hibridización de complementariedad perfecta. Estos resultados permitieron deducir una predicción positiva de At2g42980.1 como ARNm blanco de L14 (Tabla 9).

Por otro lado, durante las búsquedas realizadas en NCBI se detectó un sector conservado de 21 nt, para L29, similar a una proteína serin-treonin quinasa de *Malus domestica* (identidad 21/21, 100%). Los estudios de las estructuras secundarias permitieron obtener plegamientos moderadamente estables para ambas hebras ( $\Delta G = -109$  y  $-109,7$ ) y el

sector conservado se halló en una estructura tipo horquilla, sin embargo sólo dos de las burbujas características de los pre-miARN de plantas (Rhoades *et al.*, 2002) se encontraron en dicha estructura. Los estudios realizados con el programa miRU no arrojaron blancos putativos de especies modelo que cumplieren con las condiciones de predicción (Schwab *et al.*, 2005). Para la secuencia L72 se identificó un sector conservado correspondiente a 39 nt de un ARNm sin asignación funcional de *O. sativa* (identidad 33/39 nt, 84%), cuya posición en la estructura secundaria ( $\Delta G = -99$ ) de la hebra (+) se dispuso en una estructura de horquilla ramificada (no corresponde a una estructura característica de pre-miARN de plantas). Los estudios realizados con el programa miRU no arrojaron blancos putativos de *O. sativa* que cumplieren con las condiciones de predicción. En el caso de L73 se identificó un sector conservado de 28 nt presente en varias especies de plantas, siendo la similitud más relevante la obtenida con la subunidad F de la enzima NADHdH de *Mascagnia anisopétala* (identidad 26/28 nt, 92%). El plegamiento de la hebra (-) resultó estable ( $\Delta G = -131,54$ ) y localizó el sector conservado en el centro de una estructura tipo horquilla que contó con las burbujas típicas más una burbuja extra en el sector inferior. Aquí tampoco se pudieron identificar genes blanco regulatorios en las especies modelo con el programa miRU. Por otra parte, para L812 se detectó un sector conservado de 33 nt correspondiente a un ARNm no caracterizado de *O. sativa* (identidad 29/33 nt, 87%). El plegamiento de la hebra (+) resultó estable ( $\Delta G = -48,04$ ), sin embargo la longitud de la secuencia impidió obtener valores de  $\Delta G$  suficientemente negativos. El sector conservado se localizó en una horquilla típica que contó con las tres burbujas típicas más una burbuja extra en el sector superior. Sin embargo, tampoco se identificaron secuencias blanco en *O. sativa* que cumplieren con las características postuladas por Schwab *et al.* (2005), el supuesto miARN presentó al menos 9 desapareamientos con cada uno de los genes blanco predichos por el programa. Finalmente, tanto para la secuencia L834-1 como para L834-3 se lograron identificar pequeños sectores conservados de 28 y 37 nt respectivamente, homólogos al gen *cytP450* en diferentes especies. En el primer caso (L834-1), los mejores resultados se obtuvieron con el gen *cytP450* de *Populus trichocarpa* (identidad 25/27, 92%), y en el segundo (L834-3) de *A. thaliana* (identidad 33/37 nt, 88%). Se desarrollaron los plegamientos de las estructuras secundarias *in silico* y ambos transcritos presentaron disposiciones estables que contuvieron el sector conservado en estructuras de tipo horquilla con las burbujas características de los pre-miARN de plantas. Para el primer caso (L834-1) el plegamiento más estable correspondió a la hebra (+) ( $\Delta G = -158,74$ ) y en el segundo

(L834-3) a la hebra (-) ( $\Delta G = -100,62$ ). Los análisis realizados con el programa miRU no permitieron identificar secuencias blanco en las especies modelo ya que los miARN presentaron al menos 8 desapareamientos con cada uno de los genes blanco predichos por el programa (Tabla 9).

Por último, 4 de las 20 secuencias estudiadas no mostraron similitudes relevantes en las bases de datos de plantas (Tabla 9). Las búsquedas realizadas con L86 (187 nt), L784 (97 nt), L852-1 (122 nt) y L852-2 (84 nt) no permitieron identificar homologías significativas con genes codificantes en los sitios NCBI, TAIR y Gramene. Las comparaciones con las bases de elementos repetitivos de gramíneas y de precursores de miARN no arrojaron resultados significativos manteniendo los parámetros preestablecidos en los programas de las plataformas utilizadas.

**Tabla 9: Resultados de los análisis bioinformáticos de los transcritos diferenciales por nivel de ploidía que no presentaron homologías con proteínas conocidas.**

Clon <sup>(1)</sup>	ID principal	Anotación principal	eV/ Score	ID de gen <sup>a</sup>	Anotación del gen <sup>(3)</sup>	eV	Conclusión
L14(2x) <sup>(2)</sup>	osa-MIR419	Precursor de miARN	Sc= 120	AT2G42980	Aspartil proteasa	0,8	Posible pre-miARN. Pred. positiva
L29(2x)	osa-MIR166b	Precursor de miARN	Sc= 124	emb Z17313.1	Ser/Thr quinasa	1,6	Posible pre-miARN
L45(2x)	osa-MIR812i	Precursor de miARN	Sc= 125	-	-	-	Posible pre-miARN
L48(2x)	ZRSiTERTO OT00206	Retrotransp. ji_178A11-1	eV= 0,18	-	-	-	Posible Retrotransp.
L55(2x)	ORSiTERTO OT00437	Retrotransp. rn_525-269	eV= 0,3	AT4G11580	Proteína de la flia. F-box	1,1	Posible retrotransp. con ARNm transduplicado
L61(2x)	ZRSiTEMT05 900025 gi 16225181	IR miniatura de elemento transponible.	eV= 0,66	-	-	-	Posible repetición invertida
L72(4x)	ath-MIR863	Precursor de miARN	Sc= 160	Os06g0334600	ARNm	0,46	Posible pre-miARN
L73(2x)	osa-MIR809e	Precursor de miARN	Sc= 100	gb AF351044.1	Subuni-dad F de NADHdH	2,4	Posible pre-miARN
L85(4x)	TRSiTERTO OT00028	Retrotransp. Josephine	eV= 0,42	-	-	-	Posible Retrotransp.
L86(2x)	-	-	-	-	-	-	-

<b>L90(4x)</b>	HRSiTERT00 200014 gi 15340829	Genes Gag- Pol de retrotransp. Ty3- <i>Gypsy</i>	eV= 0,085	-	-	-	Possible Retrotransp.
<b>L784(4x)</b>	-	-	-	-	-	-	-
<b>L794(2x)</b>	ZRSiOTOT00 000019	Repetición CAG y TAG de <i>Zea mays</i>	eV= 0,24	-	-	-	Possible repetición trinucleotídica
<b>L812(2x)</b>	tae-MIR1138	Precursor de miARN	Sc= 96	dbj AK1 10092.1	ARNm	0,57	Possible pre- miARN
<b>L834-1(4x)</b>	ptc-MIR476a	Precursor de miARN	Sc= 122	7475047 CYP705 B4P-1	<i>cyt P450</i>	1,4	Possible pre- miARN
<b>L834-3(4x)</b>	osa-MIR319b	Precursor de miARN	Sc= 111	827199 CYP705 A1	<i>cyt P450</i>	4	Possible pre- miARN
<b>L852(2x)</b>	ORSiTEMT01 100018	Repetición de Retrotransp. p-SINE	eV= 0,99	-	-	-	Possible Transposón
<b>L852-1(2x)</b>	-	-	-	-	-	-	-
<b>L852-2(2x)</b>	-	-	-	-	-	-	-
<b>L884(4x)<sup>(2)</sup></b>	gb EU558522	Inserción de transposón tipo <i>Gypsy21</i>	eV= $2e^{-30}$				Transposón

<sup>(1)</sup>Se detalla el genotipo del cual fue aislado cada uno de los clones por DD, 2x: diploide, 4x: tetraploide.

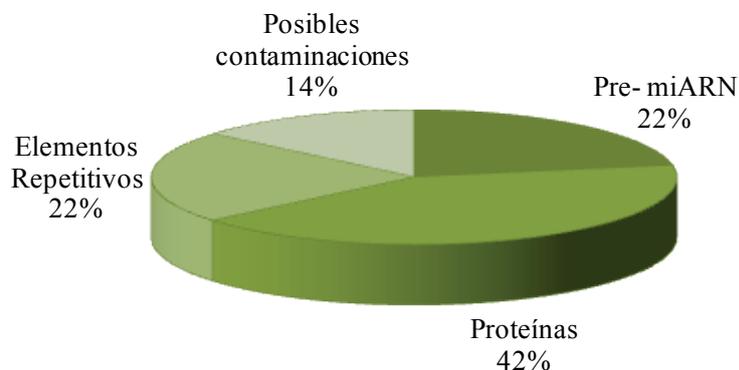
<sup>(2)</sup>Se señalan las asignaciones funcionales positivas de los transcritos diferencialmente expresados.

<sup>(3)</sup>Se detalla la identidad de los sectores homólogos internos.

Por lo tanto, a modo de resumen, del total de 36 transcritos diferenciales por el nivel de ploidía aislados inicialmente por Display Diferencial (Martelotto *et al.*, 2005), 15 (L20, L24, L47, L53, L49, L66, L75, L76, L87, L94, L95, L774, L802, L854, L884) presentaron homologías con secuencias génicas de especies relacionadas durante las búsquedas iniciales. Muchas de estas secuencias se incluyeron en pocos grupos funcionales como son las vías de señalización intracelular (quinasas), proteínas de procesamiento proteico (ubiquitinas) y proteínas de la vía bioenergética (NADHdH, *cyt P450*). Ocho secuencias mostraron similitudes con elementos repetitivos de plantas (L48, L55, L61, L85, L90, L794, L852 y L884). Ocho transcritos resultaron similares a precursores de miARN durante las búsquedas realizadas en las bases de datos de Sanger (L14, L29, L45, L72, L73, L812, L834-1 y L834-3), sin embargo se pudo desarrollar una predicción positiva sólo para una de ellas (L14). Por último, 5 de las secuencias estudiadas no

mostraron similitudes relevantes en las bases de datos de plantas. Sin embargo, pudo atribuirse asignación funcional positiva a 14 de las 36 secuencias, 12 transcritos (L20, L24, L47, L49, L75, L76, L87, L94, L95, L774, L844, L854) resultaron homólogos a secuencias proteicas de las cuales 5 correspondieron a proteínas no caracterizadas de especies relacionadas, un transcripto fue homólogo a una proteína tipo Ser/Thr kinasa de maíz, otro a una fosfoesterasa de arroz, uno a una proteína NADHdH de bambú, otra resultó homóloga a una ubiquitina de maíz y tres transcritos presentaron homologías significativas con dominios proteicos no caracterizados presentes en proteínas de maíz y arroz. Pudo asignarse funcionalidad como precursor de miARN a la secuencia L14 y a L884 como un transposón tipo *Gypsy21* conservado en *P. notatum*. La información hallada para cada una de las 36 secuencias incógnitas diferenciales según nivel de ploidía se encuentra compilada en la sección B del Anexo II. La figura XX constituye la representación esquemática de la clasificación preliminar de estas secuencias.

**Gráfico proporcional de las funciones putativas asignadas a los transcritos diferenciales por nivel de ploidía**



**Figura XX:** Clasificación *in silico* del subgrupo de secuencias asociadas al nivel de ploidía aisladas inicialmente por Martelotto, *et al.* (2005) y por Silvina Pessino (comunicación personal), que no habían podido ser clasificadas inicialmente.

## 6- DISCUSIÓN

La identificación y caracterización funcional de genes expresados diferencialmente en inflorescencias de plantas sexuales y apomícticas podría permitir el aislamiento de genes candidatos a ser utilizados en experimentos destinados a la introducción de la apomixis en cultivos sexuales. Del mismo modo, el estudio de los transcritos de expresión diferencial en inflorescencias de plantas sexuales de diferente nivel de ploidía podría arrojar luz sobre cuál es el paisaje molecular que diferencia a las plantas diploides de las poliploides, revelando las bases que rigen la estrecha asociación entre la poliploidía y la apomixis.

Nuestro grupo de investigación desarrolló un análisis del transcriptoma de *Paspalum notatum*, que permitió la identificación de 65 secuencias génicas diferencialmente expresadas entre plantas sexuales y apomícticas de esta especie (Laspina *et al.*, 2008). Varias de las secuencias identificadas pertenecían a cascadas de señalización de tipo ERK (Extracelular Receptor Kinase), con subcascadas controladas por Ras y Fosfolipasa C (Laspina *et al.*, 2008). También se identificaron transcritos pertenecientes a procesos celulares como control del ciclo celular, procesamiento de proteínas, señalización intercelular, actividad de transposones, regulación transcripcional y biosíntesis mediada por el retículo endoplasmático (Laspina *et al.*, 2008). Además, se llevó a cabo un análisis del transcriptoma de inflorescencias en citotipos sexuales diploides y tetraploides, lo que permitió la identificación de numerosos transcritos regulados en tejidos reproductivos durante las variaciones en el nivel de ploidía. Muchos de los genes expresados diferencialmente se encontraron relacionados a la modificación de la estructura de la cromatina, regulación de la transcripción, proteólisis, plegamiento de proteínas y metabolismo de lípidos (Martelotto *et al.*, 2005).

Un grupo importante de los transcritos identificados por Martelotto *et al.* (2005) y Laspina *et al.* (2008) no pudieron ser clasificados funcionalmente, porque no guardaban homologías significativas con las secuencias depositadas en las bases de datos. Justamente, el presente trabajo se enfocó en la caracterización de estas secuencias expresadas, asociadas a la apomixis y la poliploidía en *P. notatum*, que no presentaban homologías con secuencias proteicas. Se analizó un total de 58 fragmentos, originalmente aislados por Display Diferencial, de los cuales 22 habían sido asociados a la ocurrencia de aposporia (Laspina *et al.*, 2008) y 36 a la de poliploidía (Martelotto *et al.*, 2005).

A través de estos estudios bioinformáticos iniciales se pudo detectar que, a diferencia de lo informado en los estudios anteriormente mencionados, algunos de estos transcritos guardaban similitud con proteínas de especies de plantas. Este hallazgo puede explicarse fácilmente, ya que durante los últimos cinco años las bases de datos de genes se enriquecieron con abundante información proveniente de análisis de transcriptómica y genética reversa. Por lo tanto, la posibilidad de detectar homologías con genes es actualmente mucho mayor que la de años atrás. Quince fragmentos diferenciales por nivel de ploidía y una secuencia diferencial por modo de reproducción resultaron similares a secuencias proteicas, principalmente proteínas hipotéticas de gramíneas (ver Tabla 4). Además, en estas primeras aproximaciones, una secuencia diferencial por nivel de ploidía fue clasificada como posible contaminación por presentar homología completa con vectores de clonado comerciales.

Luego de esta actualización inicial de las búsquedas de homología, 41 secuencias permanecieron sin anotación. Por lo tanto, se decidió intentar una caracterización funcional de algunas secuencias selectas. Con el objeto de asegurar que no se trataba de simples contaminaciones, se realizaron amplificaciones a partir de ADN genómico de *P. notatum*. Se analizaron 16 fragmentos, todos ellos relacionados con el modo de reproducción. Doce transcritos pudieron ser amplificados a partir de ADN genómico de *P. notatum* y 4 fueron descartados por considerarse posibles contaminaciones. Del total de fragmentos presentes en el genoma de *P. notatum*, 9 presentaron patrones monomórficos y 3 mostraron polimorfismos entre el genotipo sexual y el apomíctico, que podrían ser utilizados para localizar los genes en futuros experimentos de mapeo genético.

Se decidió intentar ampliar las secuencias de varios candidatos selectos, amplificando los extremos 5' y 3' de los fragmentos mediante la técnica de RACE (Rapid Amplification of cDNA Ends). Contar con secuencias más extensas de los transcritos permitiría identificar nuevas homologías en las bases de datos. Se obtuvieron amplificaciones satisfactorias para 8 de los 12 transcritos diferenciales para el modo de reproducción estudiados (N7, N11, N13, N17, N22, N43, N56 y N104). Cuatro de las secuencias pudieron ser extendidas hacia al menos uno de los extremos (N7, N13, N17, N104). Se aislaron pequeños fragmentos internos de la secuencia N11 y se identificaron 5 secuencias derivadas de N22 (A43, A44, A45, A46 y A47) y 5 secuencias derivadas de N43 (A11, A12, A13, A15 y A16) que no pudieron ser tratadas como un consenso.

Los análisis bioinformáticos de cada uno de los transcriptos extendidos permitieron identificar similitudes con precursores de miARN de especies de plantas para tres de ellas (N11, N13 y N56). Una de las secuencias (N17) mostró homologías con elementos repetitivos de posición centromérica. Los fragmentos aislados a partir de N22 y N43 resultaron similares a elementos repetitivos LTR. A su vez, algunos de ellos portaron fragmentos génicos de secuencias proteicas cuya función había sido asociada previamente con el desarrollo apomítico. Finalmente, dos secuencias resultaron similares a proteínas (N7 y N104). Estos datos permitieron inferir que algunos de los genes investigados podrían guardar homologías con retroelementos, con precursores de miRNAs y con proteínas, por lo que la segunda etapa del trabajo de tesis estuvo apuntada a la caracterización casuística y exhaustiva de algunos representantes de cada uno de estos grupos.

### **6.1- Análisis de los transcriptos homólogos a retrotransposones**

Como se mencionó anteriormente, los resultados de las búsquedas bioinformáticas permitieron inferir que los transcriptos N17 y N22 podrían corresponder a retrotransposones transportadores de segmentos génicos. N17 mostró homologías con una secuencia repetida de retrotransposones de arroz localizada en una región centromérica. Se realizaron experimentos de mapeo con marcadores moleculares de tipo SCAR que revelaron que al menos una de las copias de N17 se localizó en una región central no recombinante del grupo de ligamiento F3c de *P. notatum*, cercano a un locus previamente asociado a la aposporía, la kinesina Mc10. La localización central y la baja recombinación en la zona sugieren que podría tratarse de una región cercana al centrómero. Los estudios de detección del número de copias genómicas por medio de experimentos de PCR en tiempo real, permitieron determinar que este retroelemento (N17) se encontraba mayormente representado en genoma de una planta sexual respecto a una apomítica (460 y 50 respectivamente). Los experimentos de PCR en tiempo real desarrollados para determinar los perfiles de expresión en los diferentes estadios del desarrollo reproductivo permitieron detectar una significativa sobreexpresión del transcripto en el estadio meiótico y el postmeiótico, que sólo ocurría en el genotipo sexual. Particularmente, a través de los análisis de hibridización *in situ* de tejidos reproductivos se identificó una mayor representación de la hebra sentido de N17 en óvulos inmaduros de plantas sexuales respecto a las apomíticas. Todos estos estudios sugieren que N17 correspondería a un retrotransposón centromérico portador de un segmento génico transduplicado, representado

de manera incrementada en el genoma de la planta sexual y sobreexpresado en óvulos de esta planta respecto a la apomíctica durante el estadio meiótico hasta el final del desarrollo reproductivo (Ochogavía *et al.*, 2011).

Por otra parte, las secuencias aisladas a partir de N22 resultaron homólogas en algunas secciones internas a retrotransposones de tipo *Gypsy* y curiosamente, en otros sectores resultaron altamente homólogas a segmentos génicos previamente asociados a la apomixis, como *serk2* y *cytP450* (Albertini *et al.*, 2005; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008). Los estudios de detección del número de copias genómicas por medio de experimentos de PCR en tiempo real permitieron identificar que este transcripto (N22) también se encontraba representado con un número de copias significativamente mayor en genoma de la planta sexual respecto a la apomíctica (53 y 16 respectivamente). Los análisis de hibridización *in situ* de tejidos reproductivos permitieron identificar una mayor representación de la hebra sentido de N22 en óvulos inmaduros de plantas sexuales respecto a las apomícticas. Los análisis de expresión desarrollados para N22 por PCR en tiempo real sugirieron que el transcripto se encontraría activo en el genotipo sexual durante el estadio meiótico y continuaría de este modo durante el estadio postmeiótico, a diferencia del genotipo apomíctico donde sólo se induciría hacia el final del desarrollo reproductivo. Se realizaron experimentos de Display Diferencial específicos para aislar y caracterizar miembros adicionales de la familia de transcriptos relacionados con N22 que se hallaran expresados diferencialmente entre plantas sexuales y apomícticas. Se identificaron 13 secuencias que presentaron similitudes con elementos repetitivos relacionados, todos ellos diferencialmente expresados en el genotipo sexual y el apomíctico. Las distancias genéticas entre los sectores homólogos a elementos repetitivos de cada uno de los miembros de la familia N22 permitieron detectar una distribución en dos grupos principales: uno relacionado con retroelementos *Ty3 Gypsy*, donde se agrupó el fragmento original N22, y otro con *Gypsy 20*. Estos fragmentos se encontraron mayormente representados en el genotipo sexual y muchos de ellos eran portadores de fragmentos génicos transduplicados que se encontraron agrupados en pocas funciones celulares (biosíntesis proteica, transporte proteico y lipídico, metabolismo de oxidoreducción y vías de transducción de señales hormonales, particularmente recepción de la señal). Algunas de estas funciones también habían sido reportadas como asociadas al desarrollo de la apomixis en *P. notatum* y otras especies (Albertini *et al.*, 2004; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008). Estas observaciones sugieren que N22 sería un miembro de una

familia de retrotransposones relacionados, portadores de segmentos génicos transduplicados, cuya representación se encontraría aumentada en el genoma del genotipo sexual y cuyos transcritos se encontrarían sobreexpresados principalmente durante el estadio meiótico en óvulos del genotipo sexual respecto al apomíctico. La existencia de fragmentos transduplicados de genes relacionados con la apomixis sugiere que estos transcritos podrían cumplir un rol regulatorio de la expresión génica, pero esto deberá demostrarse en análisis funcionales posteriores (Ochogavía *et al.*, 2011).

Resultados similares fueron obtenidos para el transcrito N43. Al igual que para N22, las amplificaciones por 3' y 5' RACE permitieron identificar 5 secuencias relacionadas con N43 que no pudieron ser tratadas como un consenso. Todas ellas, a excepción de A16, presentaron sectores conservados pertenecientes a diferentes tipos de retrotransposones. Dos de ellos, A11 y A12 contaron además con sectores similares a kinasas y asparagininas, funciones proteicas previamente asociadas a la apomixis y/o poliploidía (Albertini *et al.*, 2004; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008).

Como habíamos mencionado en la introducción de este trabajo de tesis, en los últimos años se han publicado diversos trabajos reportando la identificación de transcritos expresados en tejidos reproductivos de plantas sexuales y apomícticas (Rodrigues *et al.*, 2003; Albertini *et al.*, 2004; Albertini *et al.*, 2005; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008; Yamada-Akiyama *et al.*, 2009; Polegri *et al.*, 2010; Sharbel *et al.*, 2010; Zeng *et al.*, 2011). Estos artículos se enfocaron principalmente en la caracterización de secuencias génicas codificantes para proteínas, y su funcionalidad en las vías centrales celulares. En unos pocos estudios se informaron secuencias expresadas diferencialmente entre plantas sexuales y apomícticas correspondientes a retrotransposones (Albertini *et al.*, 2004; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008), sin embargo su caracterización no fue profundizada. Sólo durante los últimos dos años (2010 y 2011), a partir de estudios de mutantes de las especies modelo, surgieron nuevos enfoques respecto a las vías regulatorias involucradas en el control de la apomixis (García-Aguilar *et al.*, 2010; Olmedo-Monfil *et al.*, 2010; Armenta-Medina *et al.*, 2011; Baroux *et al.*, 2011; Eckardt, 2011; Singh *et al.*, 2011) que insinúan una posible participación de los retrotransposones en el desarrollo reproductivo. Estos estudios indican que los retrotransposones podrían intervenir en mecanismos de silenciamiento mediados por ARN pequeños, y que guardan relación con la metilación y alteración de la estructura de la cromatina en sectores específicos del ADN durante el desarrollo reproductivo (García-Aguilar *et al.*, 2010;

Olmedo-Monfil *et al.*, 2010; Armenta-Medina *et al.*, 2011; Baroux *et al.*, 2011; Eckardt, 2011; Singh *et al.*, 2011).

Un artículo recientemente publicado reportó el rol central del mecanismo de silenciamiento génico en la determinación del destino gamético en *Arabidopsis thaliana* (Olmedo-Monfil *et al.*, 2010). Aquí se demostró que los retrotransposones estarían involucrados en el proceso de regulación tanto del desarrollo reproductivo sexual como del asexual. La inactivación del gen *ago9*, que codifica para una proteína con dominio PAZ-PIWI de *A. thaliana*, provocó la disminución abrupta de la concentración de pequeños ARN en la nucela, y la diferenciación concomitante de numerosas células acompañantes no reducidas que expresaron marcadores gaméticos. En algunos casos, estas células acompañantes dieron origen a sacos embrionarios adicionales no reducidos, que coexistieron con el megagametofito meiótico dentro del mismo óvulo. Debido a la generación de múltiples sacos embrionarios no reducidos, estas mutantes *ago9* presentaron un fenotipo similar a la aposporia (Olmedo-Monfil *et al.*, 2010). Curiosamente, la fracción de pequeños ARN producidos a partir de la actividad de la proteína AGO9 resultó provenir principalmente de retrotransposones y, en menor medida de miARN y genes codificantes. Dentro de las diferentes familias de retrotransposones procesadas por AGO9, la mejor representada fue la familia *Gypsy* (Olmedo-Monfil *et al.*, 2010). Fue demostrado además que AGO9 puede unirse *in vitro* a pequeños ARN de 24 nucleótidos correspondientes a retrotransposones *Athila* que se encuentran expresados en el óvulo en un estadio previo a la polinización (Duran-Figueroa and Vielle-Calzada, 2010). AGO9 es requerida para inactivar una amplia proporción de retrotransposones de tipo LTR en el óvulo. Los elementos transponibles blanco de esta proteína se encuentran localizados en las regiones pericentroméricas de los cinco cromosomas de *A. thaliana*, sugiriendo una relación entre una vía de silenciamiento por pequeños ARN dependiente de AGO9 y la formación de la heterocromatina (Duran-Figueroa and Vielle-Calzada, 2010).

Además surgieron nuevas evidencias del rol regulatorio de los ARN pequeños en el desarrollo apomítico (Mallory and Vaucheret, 2010; Olmedo-Monfil *et al.*, 2010). La caracterización de nuevas mutantes *AGO104* en maíz, con un desarrollo reproductivo similar al apomítico, permitió detectar la expresión de los transcritos del gen *ago104* confinada principalmente en ovarios y anteras en el estadio cercano a la meiosis (Singh *et al.*, 2011). Además los perfiles de expresión génica generales estudiados en ovarios indicaron una interrupción masiva de la expresión génica tanto de regiones repetitivas como

de genes codificantes para proteínas en los ovarios mutantes y la demetilación de regiones centroméricas y repeticiones de ADN knob (Eckardt, 2011; Singh *et al.*, 2011). Otros autores han reportado altos niveles de *Dmc102* (proteína homóloga a la metiltransferasa *CMT3* de *Arabidopsis*) en las células de la nucela de estas mutantes de maíz (García-Aguilar *et al.*, 2010). Se conoce (en *Arabidopsis* y en *Oryza*) que las proteínas metiltransferasas participan en la regulación de la expresión génica, particularmente de elementos repetitivos, por medio de mecanismos de metilación del ADN en la vía regulada por siARN (Chan, 2008; Wu *et al.*, 2010). Y mutantes simultáneas *Dmc2* y *Dmc3* (una proteína homóloga muy cercana a *Dmc2*) de maíz producen la formación de sacos embrionarios ectópicos, algunos de ellos provenientes de CMM múltiples (García-Aguilar *et al.*, 2010). Estos primeros estudios plantean que una reprogramación epigenética a gran escala, involucrando metilación del ADN dirigida por ARN y modificaciones de la cromatina, contribuyen al establecimiento de fondos epigenéticos que activan o reprimen la expresión de los genes (Armenta-Medina *et al.*, 2011; Baroux *et al.*, 2011). Si bien, estas evidencias plantean un rol regulatorio de los ARN pequeños en la apomixis, relacionándolo principalmente con el procesamiento de elementos repetitivos y su función en la modificación epigenética (Mallory and Vaucheret, 2010; Olmedo-Monfil *et al.*, 2010; Armenta-Medina *et al.*, 2011; Baroux *et al.*, 2011), representan los primeros acercamientos a la/s función/nes de estos elementos en el desarrollo apomítico.

En cuanto a los informes sobre la existencia de retrotransposones portadores segmentos génicos transduplicados que participan de fenómenos de regulación relacionados con el silenciamiento, varios artículos interesantes se encuentran disponibles en la literatura. Tanto los elementos transponibles (TE) de clase 1 (retroelementos) como los de clase 2 (elementos móviles de ADN) fueron asociados con la transduplicación o la resíntesis de fragmentos génicos (Bennetzen, 2005). Con respecto a los elementos repetitivos de clase 1, estudios previos realizados en maíz mostraron que las regiones LTR de los retrotransposones podían adquirir secuencias génicas específicas y amplificarlas a lo largo de todo el genoma (Jin and Bennetzen, 1994). Se han descrito además numerosos casos en los que los retrovirus de animales, un ADN móvil estrechamente relacionado a los retroelementos, lograban la adquisición de segmentos génicos contiguos (Malik *et al.*, 2000). La actividad transcripcional de estos retrotransposones quiméricos portadores de segmentos génicos fue investigada en trigo y arroz (Kashkush *et al.*, 2003; Kashkush and Khasdan, 2007). Estos estudios permitieron detectar que durante la activación de estas

especies repetitivas, las regiones terminales LTR dirigen la síntesis de transcritos a partir de las secuencias adyacentes incluyendo tanto la secuencia sentido como la antisentido de genes conocidos (Kashkush *et al.*, 2003; Kashkush and Khasdan, 2007).

Respecto a los elementos transponibles de clase 2 (elementos móviles de ADN), han sido reportados numerosos eventos de transduplicación en muchas especies. En el genoma de arroz fueron identificados más de 3000 subtipos de elementos *Pack-MULE* (elementos tipo *Mutator* portadores de segmentos génicos) donde aproximadamente el 5% de los mismos se encontraron expresados en esta especie (Jiang *et al.*, 2004). Fue demostrado además que en más del 90% de los elementos *Pack-MULE* expresados, la transcripción era iniciada dentro del retroelemento (Jiang *et al.*, 2004).

Otro trabajo apuntado a la caracterización de 8274 elementos *MULE* de maíz mostró que 1337 de los mismos contuvieron segmentos génicos duplicados en su interior (Juretic *et al.*, 2005). El estudio detallado de estos segmentos duplicados permitió detectar que el 5% de los segmentos que resultaban transcritos de manera efectiva presentaban los elementos de una típica conformación pseudogénica, como dominios proteicos conservados fragmentados, pequeños marcos de lectura y múltiples codones de stop. Además, la relación en la sustitución de aminoácidos sinónimos o no sinónimos en las duplicaciones resultó acorde a la distribución esperada para los pseudogenes. Estos resultados sugieren que las duplicaciones génicas mediadas por los elementos *MULE* serían las responsables de la generación de pseudogenes (Juretic *et al.*, 2005). Sin embargo, fue comprobado que estas secuencias duplicadas presentan ciertas características que refieren a un rol potencial en la regulación de la expresión génica de la secuencia interna (Juretic *et al.*, 2005). Fue demostrado además que los *Helitrones*, un subtipo de elementos móviles de ADN distribuidos a lo largo de todo el genoma de maíz, a menudo capturan fragmentos génicos en su interior y que en el 60% de los casos corresponden a fragmentos de genes nucleares codificantes para proteínas (Bennetzen, 2005; Yang and Bennetzen, 2009). La mayoría de estos fragmentos adquiridos son provenientes de procesos de deriva génica al azar, pero se calcula que un 4% de ellos se encuentra bajo una presión adaptativa de selección. Al resultar ampliamente beneficioso para el organismo hospedador, este fenómeno sería el responsable de la retención y transposición de los elementos *Helitrones* (Yang and Bennetzen, 2009).

Algunos investigadores propusieron que los transposones quiméricos que portan pequeños segmentos génicos podrían intervenir en la creación de nuevos genes (Jiang *et*

*al.*, 2004; Bennetzen, 2005). Sin embargo, aún no ha surgido evidencia experimental de que la actividad de elementos transponibles en plantas pueda dirigir la generación de nuevos genes (Bennetzen, 2005). Por el contrario, Juretic y colaboradores (2005) informaron que todos los segmentos transduplicados que contenían dominios conservados presentaban irregularidades que impedían la codificación de una proteína funcional. Además, la relación entre sus aminoácidos sustituidos sinónimos y no sinónimos no resultaron acordes a las relaciones generadas comúnmente bajo presión de selección. Estos segmentos transduplicados constituyen una enorme fuente de pseudogenes que presentan características particulares que sugieren un posible rol regulatorio de la expresión génica (Juretic *et al.*, 2005). Se propuso además que muchos de los ARN generados a partir de *Helitrones* o elementos *Pack-MULE* quiméricos no sólo inducirían su propia regulación epigenética, por ejemplo por medio del mecanismo de silenciamiento, sino que contribuirían a regular del mismo modo la expresión de los genes íntegros transduplicados (Morgante *et al.*, 2005).

Los resultados presentados en este trabajo de tesis constituyen el primer estudio de retrotransposones portadores de segmentos génicos transduplicados de posible función regulatoria durante el desarrollo apomítico y se encuentra en completa concordancia con los reportados por Olmedo-Molfil y colaboradores (2010) y por Duran-Figueroa y Vielle-Calzada (2010). Por una parte, en las mutantes *ago9*, la ausencia de procesamiento de los retrotransposones debida a la inactivación del gen *ago9* generó un fenotipo similar al apospórico y por otro lado, en las plantas apospóricas naturales analizadas aquí se detectó la regulación negativa de N17 y de los retrotransposones relacionados a N22. La capacidad de procesamiento de los retrotransposones se encontraría afectada tanto en las mutantes sintéticas de *Arabidopsis* (*ago9*) como en las plantas naturales apospóricas, en el primer caso provocado por una alteración en la maquinaria de procesamiento por interferencia, y en el segundo porque los propios retrotransposones no se encontrarían expresados. El efecto de esta regulación diferencial sobre los mecanismos de formación de heterocromatina y sobre las vías de regulación por interferencia en plantas apomíticas y sexuales debería ser investigado en profundidad.

El artículo publicado durante el presente año, en base a estos estudios, representa el primer antecedente de caracterización de retrotransposones que contuvieron segmentos génicos previamente asociados con la aposporia sugiriendo un posible rol regulatorio de los mismos en el desarrollo apomítico (Ochogavía *et al.*, 2011).

## 6.2- Análisis de transcriptos homólogos a precursores de miARNs

Los microARNs (miARNs) son pequeños ARN (21-25nt) expresados de manera endógena, no traducidos, que regulan la expresión génica en animales y plantas. Estos ARN son generados por medio de enzimas de procesamiento denominadas DICER (RNAsa III) a partir de secuencias precursoras. Las enzimas reconocen estructuras endógenas plegadas en forma de horquilla. Luego las moléculas de miARN maduro son incorporadas a los complejos proteicos de silenciamiento, que contienen proteínas denominadas ARGONAUTA, para dirigir la represión de los genes blanco (Jones-Rhoades *et al.*, 2006). Los miARN son química y funcionalmente similares a los siARN (pequeños ARN de interferencia) los cuales pueden mediar los mecanismos de ARN de interferencia (ARNi), silenciamiento génico post-transcripcional (PTGS) y el silenciamiento génico transcripcional (TGS) (Jones-Rhoades *et al.*, 2006).

Las funciones de los miARN y los siARN se encuentran ampliamente conservadas en plantas y animales, sin embargo existen algunas diferencias relativas a las vías de procesamiento entre estos reinos. En animales, los miARNs son procesados a partir de estructuras de mayor longitud denominadas precursores primarios (pri-miARN) (Lee *et al.*, 2002). Aquí, el procesamiento inicial es producido por la enzima DROSHA, una RNAsa III de localización nuclear, por medio de dos incisiones (Lee *et al.*, 2002; Han *et al.*, 2006). De esta manera la enzima logra liberar el pre-miARN de las regiones flanqueantes del pri-miARN (Lee *et al.*, 2003b). Luego el pre-miARN es exportado al citoplasma y la enzima DICER (RNAsa III de localización citoplasmática) produce cortes adicionales, separando el miARN maduro con su secuencia reversa complementaria del resto del pre-miARN (Lee *et al.*, 2003b). En plantas, la enzima requerida para la acumulación de miARN es la denominada DCL1 (RNAsa III de plantas de localización principalmente nuclear), la cual sería responsable del primer procesamiento (Jones-Rhoades *et al.*, 2006). Se conoce que en *Arabidopsis* ninguna de las enzimas DICER es requerida para la biogénesis de los miARNs, además fue demostrado que la formación del miARN maduro doble hebra se genera en el núcleo, sugiriendo que DCL1 cumpliría ambas funciones en el núcleo, la de DROSHA y la de DICER (Gascioli *et al.*, 2005; Park *et al.*, 2005). Los pre-miARN comúnmente encontrados en animales, son detectados en raras ocasiones en las plantas, lo que indicaría que ambos procesamientos se producirían de manera casi simultánea en estas últimas (Jones-Rhoades *et al.*, 2006). Aunque existe mucha heterogeneidad entre los miARN maduros de plantas, es claro que DCL1 corta preferencialmente en posiciones

específicas de la horquilla del precursor (Reinhart *et al.*, 2002). Esto resulta en la acumulación de los miARN maduros apropiados (Reinhart *et al.*, 2002). Si bien no se conocen aún los mecanismos moleculares por los cuales DCL1 reconocería la secuencia a escindir, sería la estructura secundaria adoptada por la secuencia precursora la particularidad principal para dirigir este proceso (Reinhart *et al.*, 2002). La caracterización de las estructuras secundarias de los precursores de miARN se encuentra ampliamente desarrollada en animales, sin embargo los primeros trabajos apuntados a la identificación de los determinantes estructurales para el procesamiento de miARN en *Arabidopsis thaliana* han surgido recientemente (Mateos *et al.*, 2010; Song *et al.*, 2010; Werner *et al.*, 2010). Si bien existe una mayor variabilidad en cuanto a los determinantes estructurales del procesamiento en plantas que en animales, se conoce por medio de estudios de mutantes que la incapacidad de formar ciertas estructuras impide el procesamiento del pri-miARN (Song *et al.*, 2010). Además, mientras que en animales las estructuras de los precursores de miARN son de tamaño y forma uniformes, en plantas contienen una enorme colección de estructuras tipo horquilla diferentes (Mateos *et al.*, 2010). Algunas estructuras encontradas en plantas presentan tres zonas de pequeñas burbujas en el sector del miARN maduro y sus regiones flanqueantes: la zona anterior, la zona media y la zona posterior (Song *et al.*, 2010). Sin embargo se ha detectado que la conservación de la estructura secundaria de la sección inferior de la horquilla resulta indispensable para el procesamiento, mientras que una alteración en la sección media y superior resulta menos importante o neutra (Mateos *et al.*, 2010; Song *et al.*, 2010).

Los precursores de miARN se encuentran agrupados en familias. Las familias de miARNs son designadas por un número y los miembros individuales de las familias por números seguidos de una letra (Ambros *et al.*, 2003; Griffiths-Jones *et al.*, 2008). Las tecnologías de clonado, genética directa, bioinformática y las nuevas tecnologías de secuenciación de próxima generación han permitido la anotación de más de 3343 secuencias de potenciales miARN en plantas (miRbase, publicación N° 17: Abril 2011) (Kozomara and Griffiths-Jones, 2011). Muchas de estas secuencias, agrupadas en familias, se encuentran conservadas entre varias especies, sin embargo algunos precursores de miRNAs, se localizan una sola vez en un único genoma. La extensa homología presente entre estos precursores de miARNs y sus secuencias blanco sugieren que algunos de estos miARNs jóvenes podrían provenir de duplicaciones de segmentos génicos del gen blanco (Allen *et al.*, 2004).

Por otro lado, una característica distintiva de los miARN de plantas es que son altamente complementarios a su ARNm blanco (Jones-Rhoades *et al.*, 2006). Es por esto que la predicción bioinformática de los miARNs y sus blancos regulatorios resulta una manera rápida, confiable y ampliamente utilizada para la identificación de las secuencias potencialmente reguladas por los miARNs, resultando necesaria su posterior validación experimental (Berezikov *et al.*, 2006). Esta tecnología requiere de la consideración de criterios clave, como son la información obtenida a partir de la estructura secundaria de la secuencia y la evaluación de la estabilidad termodinámica de dicha estructura (Schwab *et al.*, 2005; Berezikov *et al.*, 2006). Así también, para la identificación bioinformática de las secuencias blanco, han sido establecidos algunos criterios que implican las observaciones de determinados parámetros relativos a la energía libre de la hibridación y al número y tipo de desapareamientos de las bases entre el miARN y su potencial secuencia blanco (Schwab *et al.*, 2005). De cualquier modo, las secuencias blanco predichas de manera bioinformática requieren de una validación experimental. Este tipo de análisis, en plantas, resulta más sencillo que en animales porque la mayoría de los miARNs de plantas regulan la expresión de sus genes blanco por clivado y posterior degradación de sus ARNm (Jones-Rhoades *et al.*, 2006). La validación por 5' RACE es la forma más utilizada para comprobar la función regulatoria del miARN sobre su ARNm (Llave *et al.*, 2002; Palatnik *et al.*, 2003; Allen *et al.*, 2004; Mallory *et al.*, 2004; Jones-Rhoades *et al.*, 2006; Mica *et al.*, 2006; Allen *et al.*, 2010; Liu *et al.*, 2010).

En este trabajo de tesis, dos secuencias que carecían de anotación funcional (N11 y N13) y que presentaron homologías con precursores de miARNs de plantas, fueron estudiadas exhaustivamente. Los análisis aplicados *in silico* a cada una de las secuencias detectaron características particulares concordantes con las comúnmente halladas en este tipo de estructuras de plantas. Estas mostraron que el sector correspondiente al miRNA maduro putativo se hallaba localizado en una estructura de tallo típico para una de las hebras simples (hebra -). Además, en ambos casos, las horquillas formadas contaban con las burbujas inferiores, centrales y superiores típicas en la región del microARN maduro predicho.

Las predicciones bioinformáticas de las secuencias blanco putativas de los posibles miARN resultaron positivas. Se asignó el ortólogo de Os03g0199100 (codificante para una proteína hipotética no caracterizada de *O. sativa*) como el gen blanco del miARN incluido en el precursor N11. Mientras, el ortólogo del gen gb|BT068773.1| (codificante para una

proteína no caracterizada con dominio ETR de *Zea mays*) fue predicho como blanco posible del miARN incluido en el precursor N13.

Se llevaron a cabo experimentos de PCR en tiempo real para el transcripto N11 (precursor sin procesar), los que arrojaron una activación en el genotipo apomíctico en meiosis y antesis, y una sobreexpresión respecto al sexual durante el estadio meiótico. En concordancia con estos hallazgos, los análisis de hibridización *in situ* de tejidos reproductivos para N11 permitieron detectar una mayor representación de la hebra (-) en óvulos inmaduros (estadio I, premeiótico tardío) únicamente en las plantas apomícticas. Al indagar los perfiles de expresión para su gen blanco regulatorio predicho, el ortólogo de Os03g0199100, se detectó que éstos fueron similares a los de N11, mostrando una mayor representación en el genotipo apomíctico principalmente durante el estadio meiótico. Esto podría sugerir que una falta de procesamiento del precursor de miRNA en las plantas apomícticas durante el estadio meiótico se correlacionaría con altos niveles de su gen blanco en el mismo estadio. Además, los experimentos de 5' RACE, desarrollados con la finalidad de detectar el producto de procesamiento del blanco regulatorio inducido por el miARN maduro correspondiente, permitieron identificar productos secundarios de degradación del gen blanco predicho para N11 sólo a partir de la biblioteca de ADNc del genotipo apomíctico. Estos transcriptos presentaron diferentes tamaños, los cuales variaban en 23 nt (o múltiplos de 23 nt) desde el sitio de corte calculado. No se encontraron productos de clivado en el sitio de corte predicho. Si bien este resultado no representa una prueba definitiva de la función regulatoria de N11 sobre el gen blanco predicho, resulta llamativa la presencia de fragmentos de clivado que difieren en 23 nt, que sugieren la existencia de un clivado específico señalado por el miARN seguido de la generación de siARN por una enzima análoga a DICER. Para confirmar definitivamente las predicciones desarrolladas *in silico* se deberán desarrollar nuevos experimentos de Northern blot con pequeños ARN y evaluar la presencia del miARN maduro.

Por otra parte, los estudios de PCR en tiempo real desarrollados para el transcripto N13 (precursor sin procesar) permitieron detectar una marcada sobreexpresión durante los estadios premeiótico y postmeiótico en el genotipo sexual. En concordancia con estos hallazgos, los análisis de hibridización *in situ* de tejidos reproductivos para N13 mostraron una mayor representación de la hebra (-) en óvulos inmaduros (estadio I, premeiótico tardío) únicamente en las plantas sexuales. Al indagar los perfiles de expresión de su gen blanco regulatorio predicho, el ortólogo de gb|BT068773.1|, se detectó su sobreexpresión en

el genotipo sexual en estadio meiótico y de antesis. Estos resultados sugieren que una falta de procesamiento del precursor de miRNA se correlacionaría con altos niveles de su gen blanco en el estadio inmediato posterior en plantas sexuales. Asimismo, los experimentos de 5' RACE desarrollados para aislar la secuencia procesada del gen blanco predicho para N13, permitieron identificar productos secundarios de degradación únicamente a partir del genotipo sexual. Estos fragmentos, todos del mismo tamaño, se encontraron a 221 nt del sitio de clivado predicho. El sector que diferencia al producto de clivado del sitio de corte puede ser considerado un múltiplo de 22, sobre todo teniendo en cuenta que pueden existir variaciones puntuales entre las secuencias de maíz y *Paspalum*. De manera análoga al caso de N11, el gen blanco de N13 podría estar siendo procesado por una enzima análoga a DICER en fragmentos de 22 nt. Si bien el hallazgo de los productos de degradación no representa una prueba definitiva de la función regulatoria de N13 sobre su blanco predicho, el hecho de contar con secuencias de menor tamaño que las esperadas podría explicarse considerando que una activa cinética de degradación dificultaría aislar el fragmento de procesamiento inicial.

Los resultados hallados en este trabajo de tesis indican que tanto los precursores de miRNA como sus blancos putativos se expresarían en tejidos reproductivos de *P. notatum*, y lo harían de manera diferencial entre plantas apomícticas y sexuales. También permiten inferir una posible disminución en el procesamiento del precursor N11 en la planta apomíctica en estadio meiótico y de antesis. Un aumento en la concentración del precursor completo se encontraría correlacionado con un aumento en la concentración de los transcritos de su gen blanco. Por otra parte, en las plantas sexuales se produciría una disminución en el procesamiento de N13 en estadio premeiótico y postmeiótico, que se correlacionaría con un aumento de la concentración de los transcritos del gen blanco en el estadio inmediato posterior. Estos resultados constituyen un valioso estudio de posibles precursores de miARN diferencialmente expresados en óvulos de plantas sexuales y apomícticas, los cuales regularían genes que demostraron expresión diferencial en estos genotipos.

### **6.3- Análisis *in silico* del resto de los transcritos bajo estudio**

En este trabajo de tesis se desarrolló además una clasificación del resto de los transcritos que no pudieron ser asignados funcionalmente en trabajos anteriores, que permitiese agruparlos de manera inferencial en torno a su potencial funcionalidad.

Los análisis de similitud de secuencias fueron realizados por medio de comparaciones con las diferentes bases de datos públicas como la de NCBI, TAIR, Gramene, de elementos repetitivos TIGR, de miARN de Sanger y de dominios proteicos ProDom. Se utilizaron las herramientas del programa Blast de cada uno de estos sitios. De acuerdo a las similitudes halladas en las bases de datos, los transcriptos fueron clasificados de manera preliminar en 4 grupos. Sin embargo, se consideró el criterio utilizado por Larrinua y Belmar (2008) para inferir funcionalidad significativa. Estos autores postulan que si bien valores de eV de 0,01 pueden indicar la funcionalidad de la secuencia incógnita, generalmente es necesario obtener valores del orden de  $e^{-10}$  o menores para que estas similitudes puedan ser consideradas como funcionalmente significativas en ausencia de informaciones adicionales o de confirmaciones experimentales (Larrinua and Belmar, 2008). En este trabajo se trató como similares a los alineamientos que arrojaron eV mayores al orden  $e^{-10}$  y como homologías con asignaciones de funcionalidad positiva a los alineamientos que arrojaron eV del orden de  $e^{-10}$  o menores.

Del total de 18 transcriptos diferenciales por modo de reproducción aislados inicialmente por Display Diferencial (Laspina *et al.*, 2008), excluyendo los analizados en los apartados anteriores, uno de ellos presentó homología con una secuencia proteica (N116), 4 pudieron ser extendidos mediante 5' y 3' RACE (N7, N43, N56 y N104) y 4 fragmentos fueron clasificados como posibles contaminaciones (N2, N49, N77 y N103) (ver Anexo II). De las secuencias amplificadas por 5' y 3' RACE, una de ellas pareció corresponder a una secuencia proteica portadora de un péptido señal (N7), una de ellas resultó similar a un precursor de miARN de plantas (N56), una resultó homóloga a una secuencia proteica (N104) y se lograron aislar 5 transcriptos a partir de N43 (A11, A12, A13 y A15) que no pudieron ser tratados como un consenso y resultaron similares a retrotransposones portadores de fragmentos génicos. Los 9 transcriptos restantes (N38, N58, N70, N82, N86, N87, N99, N106 y N119) fueron analizados únicamente de manera bioinformática (Anexo II). Cinco de estas secuencias (N58, N70, N82, N99 y N119) resultaron similares a elementos repetitivos de las cuales dos incluyeron fragmentos génicos, uno correspondiente a un ARNm no caracterizado de maíz (N82) y otro a un factor de transcripción de tipo GRAS relacionado con procesos del desarrollo. Una de las 9 secuencias (N87) resultó similar a un pre-miARN, sin embargo la predicción *in silico* no resultó positiva. Finalmente, 2 de las 9 secuencias estudiadas (N38 y N106) no mostraron similitudes relevantes en las bases de datos de plantas (Anexo II). Por lo tanto, una

clasificación preliminar permitió agrupar al total de los 22 transcriptos diferenciales por modo de reproducción en cuatro grupos funcionales: el 35% pudo clasificarse dentro de elementos repetitivos, el 20% como precursores de miARN, el 10% como secuencias proteicas y el 35% como posibles contaminaciones. Sin embargo, pudo atribuirse asignación funcional positiva a 6 de estas secuencias, N104 y N116 homólogas a secuencias proteicas no caracterizadas de especies relacionadas, N11 y N13 como precursores de miARN y N17 y N22 como elementos repetitivos de *P. notatum*.

Por otro lado, del total de 36 transcriptos diferenciales por el nivel de ploidía aislados inicialmente por Display Diferencial (Martelotto *et al.*, 2005), 15 (L20, L24, L47, L53, L49, L66, L75, L76, L87, L94, L95, L774, L802, L844, L854) presentaron homologías con secuencias génicas de especies relacionadas durante las búsquedas iniciales. Muchas de estas secuencias se incluyeron en pocos grupos funcionales como son las vías de señalización intracelular (quinasas), proteínas de procesamiento proteico (ubiquitinas) y proteínas de la vía bioenergética (NADHdH, *cyt* P450). Ocho secuencias mostraron similitudes con elementos repetitivos de plantas (L48, L55, L61, L85, L90, L794, L852 y L884). Ocho transcriptos resultaron similares a precursores de miARN durante las búsquedas realizadas en las bases de datos de Sanger (L14, L29, L45, L72, L73, L812, L834-1 y L834-3), sin embargo se pudo desarrollar una predicción positiva sólo para una de ellas (L14). Por último, 5 de las secuencias estudiadas no mostraron similitudes relevantes en las bases de datos de plantas. Una clasificación preliminar permitió agrupar al total de los 36 transcriptos diferenciales por modo de reproducción en cuatro grupos funcionales: el 42% pudo clasificarse como secuencias proteicas, el 22% dentro de elementos repetitivos, el 22% como precursores de miARN, y el 14% como posibles contaminaciones. Sin embargo, pudo atribuirse asignación funcional positiva a 14 de estas secuencias. Doce transcriptos (L20, L24, L47, L49, L75, L76, L87, L94, L95, L774, L844, L854) resultaron homólogos a secuencias proteicas de las cuales 5 correspondieron a proteínas no caracterizadas de especies relacionadas, un transcripto fue homólogo a una proteína tipo Ser/Thr kinasa de maíz, otro a una fosfoesterasa de arroz, uno a una proteína NADHdH de bambú, otra resultó homóloga a una ubiquitina de maíz y tres transcriptos presentaron homologías significativas con dominios proteicos no caracterizados presentes en proteínas de maíz y arroz. Pudo asignarse funcionalidad como precursor de miARN a la secuencia L14 y a L884 como un transposón tipo *Gypsy21* conservado en *P. notatum*.

Los resultados informados aquí permitieron obtener una potencial clasificación de las secuencias desconocidas que habían sido asociadas al carácter en los estudios previos. Puede observarse que las 14 secuencias proteicas a las que pudo asignarse homología (12 diferenciales según el nivel de ploidía y 2 diferenciales según el modo de reproducción) se encontraron relacionadas funcionalmente con las informadas en trabajos anteriores (Albertini *et al.*, 2004; Cervigni *et al.*, 2008a; Laspina *et al.*, 2008; Yamada-Akiyama *et al.*, 2009; Polegri *et al.*, 2010; Sharbel *et al.*, 2010; Zeng *et al.*, 2011). Aunque secuencias homólogas a elementos repetitivos fueron reportadas en los trabajos anteriores, como es el ejemplo de los retrotransposones de tipo *Gypsy* también informados por Cervigni y colaboradores (Cervigni *et al.*, 2008a; Cervigni *et al.*, 2008b), su análisis nunca fue profundizado.

En este trabajo de tesis se pudo demostrar que los retrotransposones a menudo pueden ser portadores de segmentos génicos transduplicados. Varios trabajos recientemente publicados (en 2010 y 2011) en la literatura internacional, realizados en plantas sexuales modelo mutantes (*arabidopsis* y maíz), permitieron predecir que el desarrollo sexual estaría regulado por RNA pequeños originados a partir de retrotransposones, y que la apomixis podría surgir como consecuencia de una alteración de mecanismos de silenciamiento génico a través de la vía de metilación del ADN mediada por ARN (RdDM). Los resultados obtenidos aquí, que fueron publicados en enero de 2011, constituyeron la primera contribución aparecida en la literatura internacional a la caracterización de varios retrotransposones que estarían involucrados en la regulación del desarrollo apomítico. La detección de segmentos génicos transduplicados en estos retrotransposones sugiere que la actividad de estos elementos podría no estar limitada a dirigir un silenciamiento por una vía de metilación del ADN mediada por ARN (RdDM), sino que además podrían directamente dirigir la degradación de transcritos blancos a través de un mecanismo de interferencia.

Por otro lado, la identificación de posibles miARN regulados diferencialmente tanto en función del modo de reproducción como a las modificaciones en el nivel de ploidía representa el primer antecedente en la bibliografía internacional de la asociación de estas moléculas con el desarrollo apomítico. Los resultados obtenidos aquí parecen indicar que la actividad de los miRNA no se limita a la simple degradación de los genes blanco por exonucleasas, sino que existiría un mecanismo de generación de siRNAs

asociado a la actividad del miRNA, que podría estar controlando la transcripción del blanco a través de un proceso de metilación del ADN mediado por ARN (RdDM).

## 7- CONCLUSIONES

Este trabajo de tesis estuvo orientado a lograr la anotación funcional de un grupo de 58 transcriptos desconocidos cuya expresión había sido asociada con la apomixis y/o la poliploidía en la especie apospórica *Paspalum notatum*. Mediante un análisis preliminar, estas secuencias pudieron ser clasificadas como posibles proteínas (31%), posibles elementos repetitivos (27,5%), posibles precursores de miARN (21,5%) y posibles contaminaciones (20%). En base a las homologías, se logró asignar funcionalidad positiva a 20 de las secuencias analizadas y 6 secuencias fueron clasificadas como contaminaciones, mientras que el resto presentaron similitudes inferiores al orden de  $e^{-10}$  que sólo permitieron inferir una función putativa.

A partir de los resultados presentados aquí se puede deducir que: 1) durante el desarrollo reproductivo se expresan retrotransposones portadores de segmentos génicos transduplicados correspondientes a genes previamente asociados con la formación de sacos embrionarios, sugiriendo un posible rol regulatorio en el desarrollo reproductivo. Estos retrotransposones se encuentran sobrerrepresentados tanto en el genoma como en el transcriptoma de las plantas sexuales respecto a las apomícticas; 2) durante el desarrollo reproductivo se expresan precursores de miARN que se encuentran diferencialmente representados entre plantas sexuales y apomícticas. Los genes blanco predichos para esos precursores también presentan expresión diferencial en órganos reproductivos de plantas sexuales y apomícticas, observándose una correlación entre su concentración y la del precursor. La presencia de fragmentos de degradación de los genes blanco sugiere una posible actividad de enzimas implicadas en el fenómeno de interferencia.

Este trabajo de tesis permitió la clasificación funcional de varias secuencias asociadas a la apomixis y la poliploidía. Los resultados obtenidos aportan información original acerca de la actividad de los retrotransposones y los miRNAs durante el desarrollo apomíctico. Además estos hallazgos proveen datos novedosos sobre las vías de regulación de la apomixis en gramíneas, abriendo un interesante panorama en un campo de estudio de desarrollo incipiente. Actualmente, en nuestro laboratorio, se encuentra puesto a punto un sistema de transformación estable de *Paspalum notatum* que permite la generación de plantas transgénicas con expresión alterada de transcriptos específicos. La información contribuida por este trabajo será utilizada en el futuro por nuestro grupo de investigación

para producir transformantes y evaluar los efectos y modificaciones sobre el fenotipo reproductivo, provocados por las alteraciones de la expresión de estas secuencias.

## 8- REFERENCIAS

- Akiyama, Y., Hanna, W. W. and Ozias-Akins, P. (2005). High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in *Cenchrus ciliaris* is located on a heterochromatic and hemizygous region of a single chromosome. *Theor Appl Genet* 111(6): 1042-1051.
- Albertini, E., Barcaccia, G., Porceddu, A., Sorbolini, S. and Falcinelli, M. (2001a). Mode of reproduction is detected by Parth1 and Sex1 SCAR markers in a wide range of facultative apomictic Kentucky bluegrass varieties. *Mol Breed* 7: 293-300.
- Albertini, E., Porceddu, A., Ferranti, F., Reale, L., Barcaccia, G., Romano, B. and Falcinelli, M. (2001b). Apospory and parthenogenesis may be uncoupled in *Poa pratensis*: a cytological investigation. *Sex Plant Reprod* 14: 213-217.
- Albertini, E., Marconi, G., Barcaccia, G., Raggi, L. and Falcinelli, M. (2004). Isolation of candidate genes for apomixis in *Poa pratensis* L. *Plant Mol Biol* 56(6): 879-894.
- Albertini, E., Marconi, G., Reale, L., Barcaccia, G., Porceddu, A., Ferranti, F. and Falcinelli, M. (2005). SERK and APOSTART. Candidate genes for apomixis in *Poa pratensis*. *Plant Physiol* 138(4): 2185-2199.
- Allen, E., Xie, Z., Gustafson, A. M., Sung, G. H., Spatafora, J. W. and Carrington, J. C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat Genet* 36(12): 1282-1290.
- Allen, R. S., Li, J., Alonso-Peral, M. M., White, R. G., Gubler, F. and Millar, A. A. (2010). MicroR159 regulation of most conserved targets in *Arabidopsis* has negligible phenotypic effects. *Silence* 1(1): 18.
- Ambros, V., Bartel, B., Bartel, D. P., Burge, C. B., Carrington, J. C., Chen, X., Dreyfuss, G., Eddy, S. R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G. and Tuschl, T. (2003). A uniform system for microRNA annotation. *RNA* 9(3): 277-279.
- Armenta-Medina, A., Demesa-Arevalo, E. and Vielle-Calzada, J. P. (2011). Epigenetic control of cell specification during female gametogenesis. *Sex Plant Reprod*.
- Asker, S. E. and Jerling, L. (1992). Apomixis in plants. *FL, USA: CRC Press*.
- Baroux, C., Raissig, M. T. and Grossniklaus, U. (2011). Epigenetic regulation and reprogramming during gamete formation in plants. *Curr Opin Genet Dev* 21(2): 124-133.

- Bennett, M. D. (2003). Comparisons with *Caenorhabditis* (100Mb) and *Drosophila* (175Mb) using flow cytometry show genome size in *Arabidopsis* to be 157 Mb and thus 25% larger than *Arabidopsis* genome initiative estimate of 125Mb. *Ann Bot* 91: 547.
- Bennetzen, J. L. (2005). Transposable elements, gene creation and genome rearrangement in flowering plants. *Curr Opin Genet Dev* 15(6): 621-627.
- Berezikov, E., Cuppen, E. and Plasterk, R. H. (2006). Approaches to microRNA discovery. *Nat Genet* 38 Suppl: S2-7.
- Berstein, E., Caudy, A. A. and Hammond, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409: 363-366.
- Blanc, G. and Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* 16(7): 1679-1691.
- Boutillier, K., Offringa, R., Sharma, V. K., Kieft, H. and Ouellet, T. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14(1737-1749).
- Burton, G. W. (1948). The method of reproduction in common Bahiagrass, *Paspalum notatum*. *J Am SOC Agron* 40: 443-452.
- Burton, G. W. (1967). A search for the origin of *Pensacola bahiagrass*. *Econ Bot* 21: 379-382.
- Burton, G. W. (1974). Recurrent restricted phenotypic selection increases forage yields of *Pensacola Bahiagrass*. *Crop Sci* 14: 831-835.
- Capron, A., Gourgues, M., Neiva, L. S., Faure, J. E., Berger, F., Pagnussat, G., Krishnan, A., Alvarez-Mejia, C., Vielle-Calzada, J. P., Lee, Y. R., Liu, B. and Sundaresan, V. (2008). Maternal control of male-gamete delivery in *Arabidopsis* involves a putative GPI-anchored protein encoded by the LORELEI gene. *Plant Cell* 20(11): 3038-3049.
- Cervigni, G. D., Paniego, N., Diaz, M., Selva, J. P., Zappacosta, D., Zanazzi, D., Landerreche, I., Martelotto, L., Felitti, S., Pessino, S., Spangenberg, G. and Echenique, V. (2008a). Expressed sequence tag analysis and development of gene associated markers in a near-isogenic plant system of *Eragrostis curvula*. *Plant Mol Biol* 67(1-2): 1-10.

- Cervigni, G. D., Paniego, N., Pessino, S., Selva, J. P., Diaz, M., Spangenberg, G. and Echenique, V. (2008b). Gene expression in diplosporous and sexual *Eragrostis curvula* genotypes with differing ploidy levels. *Plant Mol Biol* 67(1-2): 11-23.
- Comai, L., Tyagi, A. P., Winter, K., Holmes-Davis, R., Reynolds, S. H., Stevens, Y. and Byers, B. (2000). Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* 12: 1551-1567.
- Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nat Rev Genet* 6(11): 836-846.
- Conner, J. A., Goel, S., Gunawan, G., Cordonnier-Pratt, M. M., Johnson, V. E., Liang, C., Wang, H., Pratt, L. H., Mullet, J. E., Debarry, J., Yang, L., Bennetzen, J. L., Klein, P. E. and Ozias-Akins, P. (2008). Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of *Pennisetum* and *Cenchrus*. *Plant Physiol* 147(3): 1396-1411.
- Cooper, D. C. and Dick, R. A. (1949). The endosperm-embryo relationship in the autonomous apomict, *Taraxacum officinale*. *Bot Gaz* 11: 139-152.
- Correns, C. (1905). Gregor Mendels Briefe an Carl Naegeli, 1866-1873. Leipzig, Teubner.
- Cronquist, A. (1988). The evolution and classification of flowering plants. *New York Botanical Garden* Ed 2.
- Crow, J. F. and Kimura, M. (1965). Evolution in sexual and asexual populations. *Amer Nat* 99: 439-450.
- Chan, S. W. (2008). Inputs and outputs for chromatin-targeted RNAi. *Trends Plant Sci* 13(7): 383-389.
- Chase, A. (1929). North American species of *Paspalum*. *Cont US Natl Herb* 28: part 1.
- Chaudhury, A. M. and Peacock, J. W. (1993). Approaches towards isolating apomictic mutants in *Arabidopsis thaliana*: prospects and progress. In Apomixis: Exploiting Hybrid Vigor in Rice. *GS Khush* Manila: IRRI: 66-71.
- Chaudhury, A. M., Craig, S., Dennis, E. and Peacock, W. (1998). Ovule and embryo development, apomixis and fertilization. *Curr Opin Plant Biol* 1(1): 26-31.
- Chaudhury, A. M., Koltunow, A., Payne, T., Luo, M., Tucker, M. R., Dennis, E. S. and Peacock, W. J. (2001). Control of early seed development. *Annu Rev Cell Dev Biol* 17: 677-699.

- Chen, L., Miyazaki, C., Kojima, A., Saito, A. and Adachi, T. (1999). Isolation and characterization of a gene expressed during early embryo sac development in apomictic Guinea grass (*Panicum maximum*). *J Plant Physiol* 154: 55-62.
- Chen, L., Guan, L., Seo, M., Hoffmann, F. and Adachi, T. (2005). Developmental expression of ASG- 1 during gametogenesis in apomictic guinea grass (*Panicum maximum*). *J Plant Physiol* 162(10): 1141-1148.
- Chenchik, A., Diachenko, L., Moqadam, F., Tarabykin, V., Lukyanov, S. and Siebert, P. D. (1996). Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. *Biotech* 21(3): 526-534.
- Daurelio, L., Espinosa, F., Quarín, C. L. and Pessino, S. C. (2004). Genetic diversity in sexual diploid and apomictic tetraploid populations of *Paspalum notatum* situated in sympatry or allopatry. *Plant System and Evol* 244: 189-199.
- Duarte Silveira, E. D., Alves-Ferreira, M., Guimaraes, L. A., da Silva, F. R. and Carneiro, V. T. (2009). Selection of reference genes for quantitative real-time PCR expression studies in the apomictic and sexual grass *Brachiaria brizantha*. *BMC Plant Biol* 9: 84.
- Duran-Figueroa, N. and Vielle-Calzada, J. P. (2010). ARGONAUTE9-dependent silencing of transposable elements in pericentromeric regions of *Arabidopsis*. *Plant Signal Behav* 5(11): 1476-1479.
- Dusi, D. M. A. (2001). Apomixis in *Brachiaria decumbens* Stapf. *University of Wageningen, Netherlands*.
- Eckardt, N. A. (2011). A Role for ARGONAUTE in Apomixis. *Plant Cell* 23(2): 430.
- Felitti, S. A., Seijo, J. G., González, A. M., Podio, M., Laspina, N. V., Siena, L., Ortiz, J. P. A. and Pessino, S. C. (2011). Expression of lorelei-like genes in aposporous and sexual *Paspalum notatum* plants. *Plant Mol Biol* DOI 10.1007/s11103-011-9814-9.
- Forbes, I. and Bourton, G. W. (1961). Cytology of diploids, natural and induced tetraploids, and intraspecies hybrids of bahiagrass, *Paspalum notatum* Flüggé. *Crop Sci* 1: 402-406.
- Garcia-Aguilar, M., Michaud, C., Leblanc, O. and Grimanelli, D. (2010). Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes. *Plant Cell* 22(10): 3249-3267.

- Gascioli, V., Mallory, A. C., Bartel, D. P. and Vaucheret, H. (2005). Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr Biol* 15(16): 1494-1500.
- Goldblatt, P. (1980). Polyploidy in angiosperms: Monocotyledons. In: Lewis, W. H. . *ed Polyploidy: biological relevance* New York, USA:Plenum Press: 219-239.
- Grant, V. (1981). Plant speciation. New York, USA: Columbia University.
- Grant, V. (2002). Frequency of spontaneous amphiploids in *Gilia* (Pomeraceae) hybrids. *Am J Bot* 89: 1197-1202.
- Griffiths-Jones, S., Saini, H. K., van Dongen, S. and Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36(Database issue): D154-158.
- Grimanelli, D., Leblanc, O., Espinosa, E., Perotti, E., Gonzalez de Leon, D. and Savidan, Y. (1998a). Non-Mendelian transmission of apomixis in maize-Tripsacum hybrids caused by a transmission ratio distortion. *Heredity* 80 ( Pt 1): 40-47.
- Grimanelli, D., Leblanc, O., Espinosa, E., Perotti, E., Gonzalez de Leon, D. and Savidan, Y. (1998b). Mapping diplosporous apomixis in tetraploid Tripsacum: one gene or several genes? *Heredity* 80 ( Pt 1): 33-39.
- Grimanelli, D., Leblanc, O., Perotti, E. and Grossniklaus, U. (2001). Developmental genetics of gametophytic apomixis. *Trends Genet* 17(10): 597-604.
- Guo, M., Davis, D. and Birchler, J. A. (1996). Dosage effects on gene expression in maize ploidy series. *Genetics* 142: 1349-1355.
- Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., Sohn, S. Y., Cho, Y., Zhang, B. T. and Kim, V. N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125(5): 887-901.
- Hanna, W. and Bashaw, E. C. (1987). Apomixis: its identification and use in plant breeding. *Crop Sci* 27: 1136-1139.
- Hecht, V., Vielle Calzada, J., Hartog, M. V., Schmidt, E. D. and Boutilier, K. (2001). The *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiol* 127: 803-816.
- Jarret, L., Ozias-Akins, P., Phatac, S., Nadimpalli, R., Duncan, R. and Hiliard, S. (1995). DNA contents in *Paspalum* spp. determined by flow cytometry. *Genet Resour Crop Evol* 42: 242-273.

- Jeffrey Chen, Z., Wang, J., Tian, L., Lee, H. S., Wang, J. J., Chen, M., Lee, J. J., Josefsson, C., Madlung, A., Watson, B., Lippman, Z., Vaughn, M., Chris Pires, J., Colot, V., Doerge, R. W., Martienssen, R. A., Comai, L. and Osborn, T. C. (2004). The development of an *Arabidopsis* model system for genome-wide analysis of polyploidy effects. *Biol J Linn Soc Lond* 82(4): 689-700.
- Jiang, N., Bao, Z., Zhang, X., Eddy, S. R. and Wessler, S. R. (2004). Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431(7008): 569-573.
- Jin, Y. K. and Bennetzen, J. L. (1994). Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the Bs1 retroelement of maize. *Plant Cell* 6(8): 1177-1186.
- Jones-Rhoades, M. W., Bartel, D. P. and Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57: 19-53.
- Juretic, N., Hoen, D. R., Huynh, M. L., Harrison, P. M. and Bureau, T. E. (2005). The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res* 15(9): 1292-1297.
- Kantama, L., Sharbel, T. F., Schranz, M. E., Mitchell-Olds, T., de Vries, S. and de Jong, H. (2007). Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proc Natl Acad Sci U S A* 104(35): 14026-14031.
- Kashkush, K., Feldman, M. and Levy, A. A. (2003). Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet* 33(1): 102-106.
- Kashkush, K. and Khasdan, V. (2007). Large-scale survey of cytosine methylation of retrotransposons and the impact of readout transcription from long terminal repeats on expression of adjacent rice genes. *Genetics* 177(4): 1975-1985.
- Koltunow, A. M. (1993). Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules. *Plant Cell* 5(10): 1425-1437.
- Koltunow, A. M., Bicknell, R. A. and Chaudhury, A. M. (1995). Apomixis: molecular strategies for the generation of genetically identical seeds without fertilization. *Plant Physiol* 108(4): 1345-1352.
- Koltunow, A. M., Johnson, S. D. and Bicknell, R. A. (1998). Sexual and apomictic development in *Hieracium*. *Sex Plant Reprod* 11: 213-230.

- Kosambi, D. D. (1944). The estimation of map distance from recombination values. *Ann Eugen* 12: 172-175.
- Kozomara, A. and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 39(Database issue): D152-157.
- Kraitshtein, Z., Yaakov, B., Khasdan, V. and Kashkush, K. (2010). Genetic and epigenetic dynamics of a retrotransposon after allopolyploidization of wheat. *Genetics* 186(3): 801-812.
- Lakshmanan, F. L. and Ambegaokar, K. K. (1984). Polyembryony. In: Embriology of Agiosperms (ed BM Johri). *Springer Verlag* Berlin 445-474.
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E. and Newberg, L. A. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1(2): 174-181.
- Larrinua, I. and Belmar, S. (2008). Bioinformatics and its relevance to weed science. *Weed Sci* 56: 297-305.
- Laspina, N. V., Vega, T., Seijo, J. G., Gonzalez, A. M., Martelotto, L. G., Stein, J., Podio, M., Ortiz, J. P., Echenique, V. C., Quarín, C. L. and Pessino, S. C. (2008). Gene expression analysis at the onset of aposporous apomixis in *Paspalum notatum*. *Plant Mol Biol* 67(6): 615-628.
- Leblanc, O., Grimanelli, D., Gonzalez de León, D. and Savidan, Y. (1995). Detection of the apomictic mode of reproduction in maize *Trypsacum* hybrids using maize RFLP markers. *Theor Appl Genet* 90(198-203).
- Lee, H. S. and Chen, Z. J. (2001). Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. *Proc Natl Acad Sci U S A* 98: 6753-6758.
- Lee, Y., Jeon, K., Lee, J. T., Kim, S. and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* 21(17): 4663-4670.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. and Kim, V. (2003a). A nuclear RNase III Drosha initiates microRNA processing. *Nature* 425: 415-419.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. and Kim, V. N. (2003b). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425(6956): 415-419.
- Levin, D. A. (1983). Polyploidy and novelty in flowering plants. *Am Nat* 122: 1-25.

- Lewis, W. H. (1980a). Polyploidy in angiosperms. In: Lewis, W.H. *Polyploidy: biological relevance* New York, USA: Plenum Press: 241-268.
- Lewis, W. H. (1980b). Polyploidy in species populations In: Lewis, W.H. *Polyploidy: biological relevance* New York, USA:Plenum Press: 103-144.
- Liang, P. and Pardee, A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257(5072): 967-971.
- Liu, C., Zhang, L., Sun, J., Luo, Y., Wang, M. B., Fan, Y. L. and Wang, L. (2010). A simple artificial microRNA vector based on ath-miR169d precursor from *Arabidopsis*. *Mol Biol Rep* 37(2): 903-909.
- Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R. and al., e. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93: 1195-1205.
- Llave, C., Xie, Z., Kasschau, K. D. and Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297(5589): 2053-2056.
- Ma, X. F. and Gustafson, J. P. (2005). Genome evolution of allopolyploids: a process of cytological and genetic diploidization. *Cytogenet Genome Res* 109(1-3): 236-249.
- Malik, H. S., Henikoff, S. and Eickbush, T. H. (2000). Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* 10(9): 1307-1318.
- Mallory, A. and Vaucheret, H. (2010). Form, function and regulation of ARGONAUTE proteins. *Plant Cell* 22: 3879-3889.
- Mallory, A. C., Dugas, D. V., Bartel, D. P. and Bartel, B. (2004). MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Curr Biol* 14(12): 1035-1046.
- Martelotto, L. G., Ortiz, J., Espinoza, F., Quarin, C. and Pessino, S. (2005). A comprehensive analysis of gene expression alterations in a newly sythetized *Paspalum notatum* autotetraploid. *Plant Sci* 169: 211-220.
- Martelotto, L. G., Ortiz, J. P., Stein, J., Espinoza, F., L., Q. C. and Pessino, S. (2007). Genome rearrangements derived from autopolyploidization in *Paspalum notatum* sp. *Plant Sci* 172: 970-977.
- Martinez, E. J., Urbani, M. H., Quarin, C. L. and Ortiz, J. P. (2001). Inheritance of apospory in bahiagrass, *Paspalum notatum*. *Hereditas* 135(1): 19-25.

- Martinez, E. J., Acuna, C. A., Hojsgaard, D. H., Tcach, M. A. and Quarin, C. L. (2007). Segregation for sexual seed production in *Paspalum* as directed by male gametes of apomictic triploid plants. *Ann Bot* 100(6): 1239-1247.
- Mateos, J. L., Bologna, N. G., Chorostecki, U. and Palatnik, J. F. (2010). Identification of microRNA processing determinants by random mutagenesis of *Arabidopsis* MIR172a precursor. *Curr Biol* 20(1): 49-54.
- Materson, J. (1994). Stomatal size in fossil plants: evidence of polyploidy in majority of angiosperms. *Sci* 264: 421-423.
- Mecchia, M. A., Ochogavia, A., Selva, J. P., Laspina, N., Felitti, S., Martelotto, L. G., Spangenberg, G., Echenique, V. and Pessino, S. C. (2007). Genome polymorphisms and gene differential expression in a 'back-and-forth' ploidy-altered series of weeping lovegrass (*Eragrostis curvula*). *J Plant Physiol* 164(8): 1051-1061.
- Meier, M., Zapacosta, D., Selva, J. P., Pessino, S. and Echenique, V. (2011). Evaluation of different methods of assessing the reproductive mode of weeping lovegrass plants *Eragrostis curvula* (Schrad.) Ness. *Aust J Bot* 59: 253-261.
- Mica, E., Gianfranceschi, L. and Pe, M. E. (2006). Characterization of five microRNA families in maize. *J Exp Bot* 57(11): 2601-2612.
- Michaels, D. and Amasino, R. M. (1999). Flowering locus C encodes a novel MADs domain protein that acts as a repressor of flowering *Plant Cell* 11: 949-956.
- Miklos, G. L. and Rubin, G. M. (1996). The role of the genome project in determining gene function: insights from model organisms. *Cell* 86(4): 521-529.
- Mishra, N. S., Tuteja, R. and Tuteja, N. (2006). Signaling through MAP kinase networks in plants. *Arch Biochem Biophys* 452(1): 55-68.
- Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A. and Rafalski, A. (2005). Gene duplication and exon shuffling by helitron-like transposons generate intraspecific diversity in maize. *Nat Genet* 37(9): 997-1002.
- Naumova, N. T. (1993). Apomixis in angiosperms: nucellar and integumentary embryony. *CRC Press* Boca Raton - Florida.
- Naumova, T. N., Hayward, M. and Wagenvoort, M. (1999). Apomixis and sexuality in diploid and tetraploid accessions of *Brachiaria decumbens*. *Sex Plant Reprod* 12: 43-52.

- Nogler, G. A. (1984). Gametophytic Apomixis. In: *Johri BM (ed) Embryology of Angiosperms*. Springer-Verlag, Berlin.
- Normann, G. A., Quarín, C. L. and Burson, B. L. (1989). Cytogenetics and reproductive behaviour of different chromosome races in six *Paspalum* species. *Jour Hered* 80: 24-28.
- Noyes, R. D. and Rieseberg, L. H. (2000). Two independent loci control agamospermy (Apomixis) in the triploid flowering plant *Erigeron annuus*. *Genetics* 155(1): 379-390.
- Noyes, R. D. (2005). Inheritance of apomeiosis (diplospory) in fleabanes (*Erigeron*, Asteraceae). *Heredity* 94(2): 193-198.
- Noyes, R. D. (2006). Apomixis via recombination of genome regions for apomeiosis (diplospory) and parthenogenesis in *Erigeron* (daisy fleabane, Asteraceae). *Sex Plant Reprod* 19: 7-18.
- Ochogavía, A. C., Cervigni, G., Selva, J. P., Echenique, V. C. and Pessino, S. C. (2009). Variation in cytosine methylation patterns during ploidy level conversions in *Eragrostis curvula*. *Plant Mol Biol* 70(1-2): 17-29.
- Ochogavía, A. C., Seijo, J. G., Gonzalez, A. M., Podio, M., Duarte Silveira, E., Machado Lacerda, A. L., Tavares de Campos Carneiro, V., Ortiz, J. P. and Pessino, S. C. (2011). Characterization of retrotransposon sequences expressed in inflorescences of apomictic and sexual *Paspalum notatum* plants. *Sex Plant Reprod* 24(3): 231-246.
- Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., Grimaneli, D., Slotkin, R. K., Martienssen, R. A. and Vielle-Calzada, J. P. (2010). Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* 464(7288): 628-632.
- Ortiz, J. P., Pessino, S. C., Leblanc, B., Hayward, M. D. and Quarín, C. L. (1997). Genetic fingerprinting for determining the mode of reproduction in *Paspalum notatum*, a subtropical apomictic forage grass. *Theor Appl Genet* 95: 850-856.
- Osborn, T. C., Pires, J. C., Birchler, J. A., Auger, D. L., Chen, Z. J., Lee, H. S., Comai, L., Madlung, A., Doerge, R. W., Colot, V. and Martienssen, R. A. (2003). Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* 19(3): 141-147.

- Ouyang, S. and Buell, C. R. (2004). The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res* 32(Database issue): D360-363.
- Ozias-Akins, P., Roche, D. and Hanna, W. W. (1998). Tight clustering and hemizyosity of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus that may have no allelic form in sexual genotypes. *Proc Natl Acad Sci U S A* 95(9): 5127-5132.
- Ozias-Akins, P. (2006). Apomixis: developmental characteristics and genetics. *Crit. Rev. Plant Sci.* 25: 199-214.
- Ozias-Akins, P. and van Dijk, P. J. (2007). Mendelian genetics of apomixis in plants. *Annu Rev Genet* 41: 509-537.
- Ozkan, H., Levy, A. A. and Feldman, M. (2001). Allopolyploidy-induced rapid genome evolution in the wheat (Aegilops-Triticum) group. *Plant Cell* 13(8): 1735-1747.
- Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C. and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature* 425(6955): 257-263.
- Park, M. Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H. and Poethig, R. S. (2005). Nuclear processing and export of microRNAs in *Arabidopsis*. *Proc Natl Acad Sci U S A* 102(10): 3691-3696.
- Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L. and Poethig, R. S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev* 18(19): 2368-2379.
- Pessino, S., Ortiz, J. P., Leblanc, O., del Valle, C. B., Evans, C. and Hayward, M. D. (1997). Identification of a maize linkage group related to apomixis in *Brachiaria*. *Theor Appl Genet* 94: 439-444.
- Pessino, S., Ortiz, J. P., Hayward, M. D. and Quarin, C. L. (1999). The molecular genetics of gametophytic apomixis. *Hereditas* 130: 1-11.
- Pessino, S. and Martelotto, G. (2006). Genome structure and gene expression in polyploid plants. In: . *Floriculture, Ornamentals and Plant Biotechnology: Advances and Tropical Issues* Teixeira da Silva, Global Science Books, London, UK.
- Pessino, S. C., Espinoza, F., Martinez, E. J., Ortiz, J. P., Valle, E. M. and Quarin, C. L. (2001). Isolation of cDNA clones differentially expressed in flowers of apomictic and sexual *Paspalum notatum*. *Hereditas* 134(1): 35-42.

- Pessino, S. C. and Ortiz, J. P. A. (2010). Caracterización molecular de la apomixis y su aplicación en la agricultura. En: *Biotecnología y Mejoramiento Vegetal*, 2da. Edición, Levitus G., Echenique V, Rubinstein C, Hopp E and Mroginski L (eds),. *Ediciones Instituto Nacional de Tecnología Agropecuaria (INTA)* ISBN 987-521-138-9: 403-420.
- Polegri, L., Calderini, O., Arcioni, S. and Pupilli, F. (2010). Specific expression of apomixis-linked alleles revealed by comparative transcriptomic analysis of sexual and apomictic *Paspalum simplex* Morong flowers. *J Exp Bot* 61(6): 1869-1883.
- Pupilli, F., Caceres, M. E., Arcioni, S. and Quarín, C. L. (1997). Segregation analysis of RFLP markers reveals a tetrasomic inheritance in apomictic *Paspalum simplex*. *Genome* 40(6): 822-828.
- Pupilli, F., Martínez, E. J., Busti, A., Calderini, O., Quarín, C. L. and Arcioni, S. (2004). Comparative mapping reveals partial conservation of synteny at the apomixis locus in *Paspalum* spp. *Mol Genet Genomics* 270(6): 539-548.
- Quarín, C. L. and Norman, G. A. (1987). Relaciones entre el número de cromosomas, su comportamiento en la meiosis y el sistema reproductivo del género *Paspalum*. *Anales del IV Congreso Latinoamericano de Botánica- Colombia* 3: 25-35.
- Quarín, C. L., Normann, G. A. and Urbani, M. (1989). Polyploidization in aposporous *Paspalum* species. *Apomixis Newsl* 2(44-46).
- Quarín, C. L. and Normann, G. A. (1990). Interspecific hybrids between five *Paspalum* species. *Bot Gaz* 15: 366-369.
- Quarín, C. L. and Burson, B. L. (1991). Cytology of sexual and apomictic *Paspalum* species. *Cytologia* 56: 223-228.
- Quarín, C. L. (1992). The nature of apomixis and its origins in Panicoid grasses. *Apomixis Newsletter* 5: 8-15.
- Quarín, C. L., Espinosa, F., Martínez, E., Pessino, S. and Novo, O. A. (2001). A rice of ploidy level induces the expression of apomixis in *Paspalum notatum*. *Sex Plant Reprod* 13: 243-249.
- Quarín, C. L., Urbani, M. H., Blount, A. R., Martínez, E. J., Hack, C. M. and Quesenberry, K. H. (2003). Registration of Q4188 and Q4205, sexual tetraploid germplasm lines of Bahiagrass. *Crop Sci* 43: 243-249.
- Ramsey, J. and Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Ann Rev of Ecol and System* 29(467-501).

- Ramsey, J. and Schemske, D. W. (2002). Neopolyploidy in flowering plants. *Annu Rev Ecol Systm* 33: 589-639.
- Rehmsmeier, M., Steffen, P., Hochsmann, M. and Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. *RNA* 10(10): 1507-1517.
- Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. and Bartel, D. P. (2002). MicroRNAs in plants. *Genes Dev* 16(13): 1616-1626.
- Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B. and Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell* 110(4): 513-520.
- Roche, D., Cong, P., Chen, Z., Hanna, W. W., Gustine, D. L., Sherwood, R. T. and Ozias-Akins, P. (1999). Short Communication: An apospory-specific genomic region is conserved between Buffelgrass (*Cenchrus ciliaris* L.) and *Pennisetum squamulatum* Fresen. *Plant J* 19(2): 203-208.
- Rodrigues, J. C., Cabral, G. B., Dusi, D. M., de Mello, L. V., Rigden, D. J. and Carneiro, V. T. (2003). Identification of differentially expressed cDNA sequences in ovaries of sexual and apomictic plants of *Brachiaria brizantha*. *Plant Mol Biol* 53(6): 745-757.
- Rodriguez, M. P., Cervigni, G. D., Quarin, C. L. and Ortiz, J. P. (2011). Frequencies and variation in cytosine methylation patterns in diploid and tetraploid cytotypes of *Paspalum notatum* assessed by MSAP markers. *Biol Plant* (in press).
- Rosenberg, O. (1908). Cytological studies on the apogamy in *Hieracium*. *Bot Tidsskr* 28: 143-170.
- Rosenberg, O. (1930). Apogamie and parthenogenesis. In: Baur, E. Hartmann, M. Johri BM ed *Handbuch Vererbungswiss II* 12 (Borntraeger, Berlín).
- Sahgai-Marroof, M. A., Soliman, K. M., Jorgensen, R. A. and Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci U S A* 81: 8014-8018.
- Savidan, Y. (2000). Apomixis: genetics and breeding. *Plant Breed. Rev.* 18(J. Janik (Ed.) John Wiley & Sons, Inc. London).
- Schmidt, E. D., Guzzo, F., Toonen, M. A. and de Vries, S. C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* 124: 2049-2062.

- Schranz, M. E., Kantama, L., de Jong, H. and Mitchell-Olds, T. (2006). Asexual reproduction in a close relative of *Arabidopsis*: a genetic investigation of apomixis in *Boechera* (Brassicaceae). *New Phytol* 171(2): 425-438.
- Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M. and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev Cell* 8(4): 517-527.
- Sharbel, T. F., Voigt, M. L., Corral, J. M., Thiel, T., Varshney, A., Kumlehn, J., Vogel, H. and Rotter, B. (2009). Molecular signatures of apomictic and sexual ovules in the *Boechera holboellii* complex. *Plant J* 58(5): 870-882.
- Sharbel, T. F., Voigt, M. L., Corral, J. M., Galla, G., Kumlehn, J., Klukas, C., Schreiber, F., Vogel, H. and Rotter, B. (2010). Apomictic and sexual ovules of *Boechera* display heterochronic global gene expression patterns. *Plant Cell* 22(3): 655-671.
- Siena, L. A., Sartor, M., Espinoza, F., Quarín, C. L. and Ortíz, J. P. (2008). Genetic and embryological evidences of apomixis at the diploid level in *Paspalum notatum*. *Sex Plant Reprod* 21: 205-215.
- Singh, M., Burson, B. L. and Finlayson, S. A. (2007). Isolation of candidate genes for apomictic development in buffelgrass (*Pennisetum ciliare*). *Plant Mol Biol* 64(6): 673-682.
- Singh, M., Goel, S., Meeley, R. B., Dantec, C., Parrinello, H., Michaud, C., Leblanc, O. and Grimanelli, D. (2011). Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. *Plant Cell* 23(2): 443-458.
- Smith, J. (1841). Notice of a plant which produces seeds without any apparent action of pollen. *Transactions of the Linnaean Society of London (meeting of June 18 1839)* 18.
- Soltis, D. E. and Soltis, P. S. (1995). The dynamic nature of polyploid genomes. *Proc Natl Acad Sci U S A* 92(18): 8089-8091.
- Soltis, D. E. and Soltis, P. S. (2003). The role of phylogenetics in comparative genetics. *Plant Physiol* 132(4): 1790-1800.
- Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., Sankoff, D., Depamphilis, C. W., Wall, P. K. and Soltis, P. S. (2009). Polyploidy and angiosperm diversification. *Am J Bot* 96(1): 336-348.
- Song, K., Lu, P., Tang, K. and Osborn, T. C. (1995). Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc Natl Acad Sci U S A* 92(17): 7719-7723.

- Song, L., Axtell, M. J. and Fedoroff, N. V. (2010). RNA secondary structural determinants of miRNA precursor processing in *Arabidopsis*. *Curr Biol* 20(1): 37-41.
- Soreng, R. G., Davidse, G., Peterson, P., Zuloaga, F., Judziewicz, E., Filgueiras, M. and Morrone, O. (2009). CATALOGUE OF NEW WORLD GRASSES (Poaceae). Data base <http://www.tropicos.org/project/cnwg> (Missouri Botanical Garden - Saint Louis - USA).
- Spillane, C., Curtis, M. D. and Grossniklaus, U. (2004). Apomixis technology development-virgin births in farmers' fields? *Nat Biotechnol* 22(6): 687-691.
- Stebbins, G. L., Jr. (1947). Types of polyploids; their classification and significance. *Adv Genet* 1: 403-429.
- Stein, J., Quarín, C. L., Martínez, E. J., Pessino, S. C. and Ortiz, J. P. (2004). Tetraploid races of *Paspalum notatum* show polysomic inheritance and preferential chromosome pairing around the apospory-controlling locus. *Theor Appl Genet* 109(1): 186-191.
- Stein, J. (2006). Estudios de generación y herencia de genes relacionados con la apomixis en razas tetraploides de *paspalum notatum*. Tesis Doctoral presentada en F.C.B. y F.-U.N.R. Rosario- Argentina.
- Stein, J., Pessino, S., Martínez, E., Rodríguez, M. P., Siena, L., Quarín, C. L. and Ortiz, J. P. (2007). A genetic map of tetraploid *Paspalum notatum* Flüggé (bahiagrass) based on single-dose molecular markers. *Mol Breed* 20: 153-166.
- Stone, S. L., Kwong, L. W., Yee, K. M., Pelletier, J. and Lepiniec, L. (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci U S A* 98: 11806–11811.
- Stupar, R. M., Bhaskar, P. B., Yandell, B. S., Rensink, W. A., Hart, A. L., Ouyang, S., Veilleux, R. E., Busse, J. S., Erhardt, R. J., Buell, C. R. and Jiang, J. (2007). Phenotypic and transcriptomic changes associated with potato autopolyploidization. *Genetics* 176(4): 2055-2067.
- Toenniessen, G. H., The flowering of apomixis: from mechanisms to genetic engineering (2001). Feeding the world in the XXI century: plant breeding, biotechnology and the potential role of apomixis. In: Savidan, Y.; Carman, J.G.; and Dresselhaus, T. eds. *Mexico DF: CIMMYT IRD, European Commission, DG VI.*

- Van Dijk, P. J. and Bakx-Schotman, J. M. (2004). Formation of unreduced megaspores (diplospory) in apomictic dandelions (*Taraxacum officinale*, s.l.) is controlled by a sex-specific dominant locus. *Genetics* 166(1): 483-492.
- Vielle-Calzada, J. P., Nuccio, M. L., Budiman, M. A., Thomas, T. L., Burson, B. L., Hussey, M. A. and Wing, R. A. (1996). Comparative gene expression in sexual and apomictic ovaries of *Pennisetum ciliare* (L.) Link. *Plant Mol Biol* 32(6): 1085-1092.
- Vijverberg, K., Van Der Hulst, R. G., Lindhout, P. and Van Dijk, P. J. (2004). A genetic linkage map of the diplosporous chromosomal region in *Taraxacum officinale* (common dandelion; Asteraceae). *Theor Appl Genet* 108(4): 725-732.
- Warmke, H. E. (1954). Apomixis in *Panicum maximum*. *Am J Bot* 41: 5-11.
- Wendel, J. F. (2000). Genome evolution in polyploids. *Plant Mol Biol* 42(1): 225-249.
- Werner, S., Wollmann, H., Schneeberger, K. and Weigel, D. (2010). Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. *Curr Biol* 20(1): 42-48.
- Winge, O. (1917). The chromosomes: their number and general importance. In: Jackson, R.C; Hauber D. *Polyploidy*. Stroudsbery, PA, USA: Hutchinson Ross Press: 131-275.
- Winkler, H. (1916). Über die experimentelle erzeugung von Pflanzen mit abweichenden chromosomenzahlen. *Zeitschrift für Botanik* 8(417-531).
- Wu, K. K., Burnquist, W., Sorrells, M. E., Tell, T. L., Moore, P. H. and S.D., T. (1992). The detection and estimation of linkage in polyploids using single-dose restriction fragments. *Theor Appl Genet* 83: 194-300.
- Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C. and Qi, Y. (2010). DNA methylation mediated by a microRNA pathway. *Mol Cell* 38(3): 465-475.
- Yaakov, B. and Kashkush, K. (2011a). Massive alterations of the methylation patterns around DNA transposons in the first four generations of a newly formed wheat allohexaploid. *Genome* 54(1): 42-49.
- Yaakov, B. and Kashkush, K. (2011b). Methylation, transcription, and rearrangements of transposable elements in synthetic allopolyploids. *Int J Plant Genomics* 2011: 569826.
- Yamada-Akiyama, H., Akiyama, Y., Ebina, M., Xu, Q., Tsuruta, S., Yazaki, J., Kishimoto, N., Kikuchi, S., Takahara, M., Takamizo, T., Sugita, S. and Nakagawa, H. (2009).

- Analysis of expressed sequence tags in apomictic guineagrass (*Panicum maximum*). *J Plant Physiol* 166(7): 750-761.
- Yang, L. and Bennetzen, J. L. (2009). Distribution, diversity, evolution, and survival of Helitrons in the maize genome. *Proc Natl Acad Sci U S A* 106(47): 19922-19927.
- Yoshikawa, M., Peragine, A., Park, M. Y. and Poethig, R. S. (2005). A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev* 19(18): 2164-2175.
- Yu, Z., Haberer, G., Matthes, M., Rattei, T., Mayer, K. F., Gierl, A. and Torres-Ruiz, R. A. (2010). Impact of natural genetic variation on the transcriptome of autotetraploid *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 107(41): 17809-17814.
- Zeng, Y., Conner, J. A. and Ozias-Akins, P. (2011). Identification of ovule transcripts from the Apospory-Specific Genomic Region (ASGR)-carrier chromosome. *BMC Genomics* 12(1): 206.
- Zhang, B. H., Pan, X. P., Wang, Q. L., Cobb, G. P. and Anderson, T. A. (2005). Identification and characterization of new plant microRNAs using EST analysis. *Cell Res* 15(5): 336-360.
- Zhang, Y. (2005). miRU: an automated plant miRNA target prediction server. *Nucleic Acids Res* 33(Web Server issue): W701-704.
- Zou, C., Lehti-Shiu, M. D., Thibaud-Nissen, F., Prakash, T., Buell, C. R. and Shiu, S. H. (2009). Evolutionary and expression signatures of pseudogenes in *Arabidopsis* and rice. *Plant Physiol* 151(1): 3-15.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13): 3406-3415.

## ANEXO I- Sección A

**Secuencias de bandas diferenciales según el modo de reproducción que no presentaron anotaciones funcionales. Fueron obtenidas por la Dra. Natalia Laspina (Tesis Doctoral, 2008) mediante Display Diferencial**

**N2**

AGGCCGCTTAGTGTGAGCAGGTAAGTCTGTCGTTAACTTATAGTTGGAGCGG  
GCAAAGTAAGAAAGGAACGAAAAGTCTGACTCTGATGAAGAAGCATCCGC  
CTTGGTGGCAGCGCTTGCCAGCTTTTTATAAGCCGAAGATGGGAAAGCTTCC  
GCTTCAGCATAAATGCTGTAAACGCGCTGGTTTTGATAACTCATGCCACCA  
CGCCATCAATGTCGTGCCTGTCGGCAAACGAAGTGGCATAACCGGAAAAAGT  
TGTTGGTGGTAATGTTTCAGGATCTGGTTCGGAGGAAGCAAGGCCACCGCCTT  
TGGGATATCCCGAGTTGCGTTCGGTAAGCGGCCTA

**N7**

AGGAAGGTGCTAACCAGTCAGGAAAAAGTTCTATCCCAGACGCCAACCCCTG  
ATATTGCTGATGCGATGCAATTCTAAGTGAAGACAAATAGGCTTATTTGGCG  
TGGTTTGCAGGTGTTTCAGATGTAGGCTTTTATTTTGCAAATGTAGCACGATG  
GCTCCCCTACGCCACAGTGCTTGATGGTGTAAAAGGAACATTCAAAGAAG  
GTACCCTCAGCTCACTACGAAGAAGTATCTATGTTACCGCAGTGCAGGTGTA  
AAAGCAGGTGCGTCCGCCAAGTGCAGGTGTAAAACAATGTTTGATGGTTCG  
CACAGAAATTACATCAGTAACCTTCCTAATGTGCTACTACCCCTCAGCCTGC  
CACAATGTTTATTTGTGTTTGTAACTCAATGTTATACGTCATGTTCTGTGT  
ATGAACAATCTTCAGCTGTAATATATGCTAAGTGTGTAGCTCATTCTGTTGT  
AACCGCCCTTGTCTATTGCACCTTCCTA

**N11**

CGAACGGCGGTAGAAACAATGTTAACGGACTCTGCCGTTTGAGCCATACCT  
GAAATCGTGAACAATGCCAATAAAGCAGATAGCCCAAATTTGGGGAAATAC  
TTCATGTAGGGATTTTTAAGTAGATTTTCATCTCATCAAGCAGACCATTAAA  
ATAAGATCATCTGCCCCGCACCTGCATAGCGCTCACCCACCTAAGGCAATGA  
TAACCCGGTAAATATAAATACCTTTTTCTAATTTCTGCCACTCGCATCGGC  
ACCATCCCAACTTATTTGCGCATTACGTGATCCCGCCGATTGCACAATCGAT  
TGCAACTGGCGTACCAGCTGTCCGTTAGACGTGTAGATACCAACCATCACTG  
AAAATGCTGCCCCGGGTTGGTTGTGTTTCGAAAGAAAAATTGGTATGTCCCG  
AAAACGGGTTGGGATAATTCATTAAATGGGTGATACTAAGTTTCTCCCGCTT  
AACTACCATAAAGTCCAGGCTGGCAGTAGCCGATTGGTTGGCCACATTCCA  
GGCCCGCAAGCGGATCGTATGGCGGCCCGGCGTAAGGCCTGATAACTGGTA  
CTGCAAAGACCCGCGTTGGTAATTGTCTTTATCGGCCGTATAAAAATCGTTC  
AACACCAGGGTATTATTTTCATTGCCGTCGACCACCGCAATGATATCATGAC  
CAATACCATTACCCGATGTGTTGATGCCCGATGTATCGAATATCCGCGCAAT  
GAGCAAAGGATTTTCATTGGTCAATCCGCCGTTTCGA

**N13**

CTAACGGCGGTCCAGAAGACCAACTAAGCCCCAAACATGACCTAGAGGGCC  
CATCAGTCATGTTTGGAAAGGAGGAGTTGCCCAAGTATCAAGGGGTTGTCT  
GGGACAAAGCCTCACAACCCACCAAACTAAGTTTATAGATTAGATCGTTCAAG  
TTTTCGACAACAAGGTAGGCCCTTAGCCATTTACTTATTTAGGGAGTTC

CACCACCCTATATTTGTTTGTGCATTACGTTTAAGAGTTGTATGAAACCCCA  
 CATGCATAAACCCCTGATATCCTTTGTAGTATGAAAAGATCGAAGTCTATAC  
 TTATCCATGCCTAAACGGTACAAGTCGAGTGAAGATATCACTCGCGAGAT  
 GATATGGCTCAAAAATGTTTTACAAAAGTTTAATTTGAAAAGCAATATGAT  
 ATGATATGATGTGATGGAAAAATTGAGACCGGGTGGAGATATGGGTATATT  
 ATATTGAAAACGACGAGCATGGTTTGGTGGACTCCGCCCAAGTCATTTAA  
 GACCCCGTCCGCCGTTCTGA

**N17**

ACGGCCGACCTCTTCACCCCTTAGTGCTATAAATACTCCCCTTGAGAGCTCA  
 TTTGAGTGTGTGGATAGATGTAGTATAGTGGTATACTAGCACTTGAGAGTGT  
 GAAGTTGTATAGTGTGGAGAAGTGAAGCTTCCTTAGCTTGTAGAGAGAGG  
 CATTCTCTTGTATCAAGGAGAGTTGTGAGTATAGAGTGCGGGAGTCGGATT  
 TCTCTTCTCTCCAGACTTGTGTTCTCAAAGGAAGGATATTCAGTTCTAGGA  
 TTGAGTAAAGTTCCTTATCTCTTACCTTTGTTATTTTCATCCTTACTTTACTTTC  
 AACTTAGTGATTTACTTGTCTGTTTACGGGAACCTTGTGCATGCTTATAGTG  
 TTGGCTATAAAGTAGTGCTTGTGTTTTCTGGTTCATTACCAGTAAGCAGTAAA  
 CTCTCTGGGTGGCACGATGGTGTTCATACTTAGAGTTTGCCTTGGTTGGTCG  
 GCCGTA

**N22**

CAGCGAACTACACNGGTTGCACTAAACAAAGAGGAGCAAAAACCTCTACTAA  
 CCAACAGAAAGAGGGGCACTGGACACAATAACGGAGGACCACACGTCTT  
 CACTGCTGGAGGGAGCTGGTAACTTCATCCGGGGATGTCTCCCTAGAGCTG  
 CTGTTGATGATTTCTGGCGACGGCTGATTCGAGGCGGCTGAACCCTCTTTTC  
 CTTCACTGGGTGAGGCAGGGCTTGTGGCTCTGAGCTAGATACTGAGAGGG  
 GTAATGCTGCTTGCCTCCCTAGCAGCTCTTCCAAGGCTAGTTCGCTGA

**N38**

CTGGCATGCACTTTTACTTTTATCGCTCTCCCTCTTGCTTTATTATTACCTTTT  
 GTCATCTCCACGTCATGTCAACTATATTTTTATATCTCTTTCTCTCATCTCTCC  
 TACTTCTTTCTANNNNNNAATTTTATTATATTATGGGAACTGTAGGTCCTN  
 CCCGATACGCATGACCACGCCTTTCCCGGCGCGGGGATATGAACGTGCC  
 TTCGATTTAGATTCGTCATCTCCCGCCTCAGTCGCGAATAATGAGAGCGCCC  
 GGTCTCAGGGCAAACANCTGCGAGTTGTGTCAGGCGACACCACCCGGTCT  
 GTCCGTCTTCAAGTGCGTTCCTTGTGTCAGAGCGAGTATATATCGTATTCTAGTG  
 GACCTGTCGCGTGTGATGATAATCACTGTGGTCATTGCCAGCAGGTTCTAGA  
 GGCAGCCTAGCCATCATAAAGTATAGTTCTGTGTCCACNTATCGTCTTATGC  
 GCGGTCATATAAGAGAACGCTTCTTCGCGTGCGTTACATCGTCCTGACATAC  
 TAGTCAGTAGGTTCTATGACTGCGTATCTGTGGTACTAGAGACTGATTGTCT  
 CTCACTTGCAGTCTGACGGCGTACGTGCGAACAATCAATTGATCTTACAGCT  
 CTACGTCCTTAGCTCCATTATGACCGAGCGAGAGGATGTTGAGTATAGAAC  
 GNACGTGTCACAGTTCTTCGATCAGGGATAGAACTATTCGATCCGCGCCTC  
 AGGCACATTGTCATGTGCCCTGCCACTCCAGTGTATCGCGGACACGCAATC  
 ATCACTCAATATTGTGTCGACTGTGTGCAANCTGCCTACTTCATGGGCTTG  
 TAGTTAGCTAACATTCGCGTTAGCTGATGACTCGAGTACGATCACTACACAT  
 ATGTGACGCTGTTGGTGAAGCTCTGATCTAGACGTCCATATGTATGAATGAC  
 AACTCACGACTGTTTCGTGTATTTACCCTGCGGCTAGAATCGGACACACAT  
 ATTTGCTGGCTGTTTACGACAGGCCATTTCTGTCTCTATTTGTCTCTCCCCGCT  
 GTATCATAACTTCATTCTTATCTGTCATATATCATCCCGCTCTATAGCCTCAT

ATTATATATAATGTAGCTGCCCTCAGAATATAGTTTACCTGTAGCGCANGTG  
CATGCGTGCTACCTCTATGAATTGATAAGATATTCGTAAGCTACTCTTGACA  
CCTGTATCCTCT

**N43**

CCTCCCAGAGGGGCTACCAGCAAAGCTACACGGAGAGGATGGAAAGCTAT  
AAACATGTTCAAGCGTAGCTTAAACAACGATAGATCTTCCTGGATCAGAGG  
AGTGGATCTAGCCAGAGCCTAGCGTCTAAGTCTAGGGTTGAGTTATAACAC  
CCGCAAGAATAACAAGACTTGAATAACTAAGTGAAATACAATCTAAGCCTA  
GCAAGTATGCAAAATAAAGGTACAACCTACTTGCAGGAGATAGAAGTAAA  
GTAAAATATAACTGAATCATATTGAAGATAATCTAAATCCAAAAGATACAA  
GAGTTTTACCACAGACTCGAAGATAACTCCGGAGACTCCGAGGCAATGCTC  
GACTCAACTCTAACTCCCAGGACTATCCTAACTACTAAGATAACAGTGAGG  
CACTCTAAGATAGAACTCTCTAACCAACTCCACACTTTACAAATGAGCTCTC  
GGTGCAATCACTAATGAATTCGGCCGCTGTAGGTGACCATATGGGAG  
AGTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGGCCACCTAA  
ATAAGCTTGGCGTAACCCTGGGCCTAGCTGTCTCCCCGTGGGAAATTGTTAT  
TCGCTCAACAATTTTCCCCACCATACCAAACCTGAAAGCATAAAGTGGTA  
AAGCCTGGGTGTGCTCAAATGGACTGAGCTCAACTTAACTTAAATTTCCGT  
TCGCGACTTACTGGCCCCGCCTTTTCCACTCCGGGCAAACCCGTTTTTGGC  
CACATCTGCAACTTA

**N49**

GGGCCTCTATGGGAGATTTAGAATTATTTAAGAAATAGTTTGCAGAACAAA  
ATTAGGACCACCCCATGAGCAATGTATGTGAATTTTATGTAGGCAAATTAAT  
TTGTCCAATTAAGGATTGTTAAGAATCATGTAGGTGATAATCTTTTTAGG  
GTTTCAACTAAATTAATTAGAGAAGTTTTCAAGCTAAACACTAACTATTATA  
CCCATTAGAAAATTGATATGGAGGATTTCCAAGCCAGATATGATGTATAGA  
GGCCCA

**N56**

CGTCGCCCATCGTTTTATAGGCTTGTGTACAAAGATATTGGGCAAGTCTCGG  
TCGAAGGGCTTGGCCATTAACCTACCGGATCACCGGCAAACATGAACCAA  
AGGCAGCCACACGGCTGTCTTGGACAGCAGCAATAGCGGTCATATCAACCA  
ATGAAAAGCTGGTCTCCATGGCTCCGATGGCAAACCGTTTTATGGGTATGT  
GTTCAAGGACAAGATACAAGATTTTGTGTAAGGGCATGACCATCAGGTCAC  
TGGAATAATGGAAAACACCCGTGCCAACTAGGGCAACGTGGCTGTCCATGA  
CTGCAGCGAGACGTTATATGGGTAAACCTTGCCGGAACCTTGATGGGATTCG  
ATCTATGCAACATGATTTGCTGTGGGAAACAAGCGATCTACATAGGGAGGT  
ATTAAGGGGTTCTCACCATGCAGTTGATGGCAGAAGAGATGACATGGGCGA  
CGA

**N58**

CGTCGCCCATCTATCTGGCCCCACGACGCTGATCCATCGTCCTACAAAAGAG  
ATAAAAAGATTTATTACATGCGAGAGATTTTCATAGTTGTTAGTTTAAAAATT  
ACCATGTTGTGATCTAGATAATTAACCACTATTTTGTAATTTATTACCATCTG  
TTTTTGTGTTCTTTCATGAAACTCCAATCATTCTCTAAACATGGGCGACGA

**N70**

CGACCAGAGCACAGTTGTTGCAGCGCAATGATCAGGCTATGAAGATATCGG  
ATGCGATGGCATCATTGAATCCAATATTGAAGAATGAAGTGGATAAGTTCA  
AGGCAGCGCCAGAAGGACCTGCAAAACAATTTGCTGGCGCTCGAGTTGTT  
TCAAATATTTTGGCATGACCCCTTACCTTGAAGGTGGTCTAGAGCGGCATGG  
TGAGCGCGTTGATGAATTCGATTGGTACAACGGAAATTTCTGGGTGCCGTTT  
CCGCTTGTGAAAAAGCCGAAGACCGAAGAAGATATGTGGTCTTATTATGCA  
ACTGTGCTTTTAAACGGTAGTAATTCGATGAGTGAGCGGCTTAAGGGCTATG  
GCGAGGATGCGTTGAAGAGCCGACTTCCACTGAAGAAAAGAATCGAGCCG  
CTGCTGAGCGCGCTTTGATTCTCAAAAATAATCCATCTCGATATTATGGACA  
GACGGTGTAGGTTGGGCGAAGTCGCATCCTACTGATCCAGATGTTCCAGA  
GATGCTTTATCGAATCGTGAAATTGCCGAAGTGGACAGCGGAATCAGAAGT  
CGGCAGTGAGTATTCGCGAAAGGCTTATATGGCGTTGCATGCAAATTATCCC  
GACAGCAAGTGGGCTAAGAAGGCGGTTTGGTACTACTAGGCAATGCTCGTC  
TTGTTCTTCTAAGCAATGCTCGTTTTGGCTTATTTAAGCCTCGCTGTTTTGGC  
TCGTTTAAGACATATTGTTTGGGCTCGTTAAGCCATACTCGTTTTGGCTAACT  
GCTGGTAGGTGTTGCCTGCATTAGCGGCACGATCGCCGAGTGACGCGCTTC  
GCTCATATTTGACT

**N77**

GCCCGACGCGCTAGTGCAGGAGTTATCCCAGCATCAACTGCCGGAGACCGG  
TTATTTACTGCTGGCACATTACCGCTACCTGGCCAGCGAATTTTTATTGATCT  
GTCTGCTGGGCAGCAAAGATCACTTTGAATTAAGTGAAGCAGCTGGAGCTGG  
CAACTTCCCGCCATCTGGATATCAGCCGGATGCAGCTAGCAGCGCGGTTGG  
ATCTGACCGAATATGCGGTAATCGTGAGCAGGAGCATTACATCAGCTTTA  
TCCGTGGTTCGCGTCGGGCA

**N82**

CGGTGACATCCAAGAACCTTGCTCTTTTACCATGCTAGAAAGGAACTGGAAG  
TCTGGAATGTGCGATTGCATTCAATACTAGGGACTATTCACATATTCACACA  
CATGTACATATAACAACAAGTGGGCTATTTTGGGCCATATGGGCCATAA  
TAGAGAATACCTACTGTAAGCCACCTAGAACTCAAGTCAGCGTAGAGTCGA  
CAAGAGGCGAAAGTTTGGGTTTGGAGAGGAGAATTAGGGTTTGGATGAAATAG  
AATTGCATAAAAGTAGATAAAAGTAATGTATTTGGTTCGATTGGCGATACCT  
CAATCGGCTTGGCCCCTTATATTTATAGTGGAGGGTGGTCTTGACCTTTGT  
ACAAGTCAAACCCCGGAGATCTCGCACAAAACACTAGATTCGTCTTGAAC  
TCGACCACACCGGATAGTCTGGCATAGGCAAACGGATTGTCTGGTGCCACT  
GGAGCGATTGTCTGGTGCCACTGGAGCGATACATGCCACCAGAACCGGTGC  
AAGCCTACCTAAATTGGTGCTCAACACGATGTCACCGA

**N86**

ACGGCAGTGGAGGTAGGCCCTATTGGGAGAGTGGCACATGCGGCGACTGCG  
TCCATTGATTATTTGATTATCGGGACAAGTGCCTACTAGGCACGCTTGTC  
GGTCTTTTATTGCAAGTTATACTGCGGAGGATGTATTGCTATGGCTCTTAG  
CGATGTAAGCATTAAAGAGTTGTGTGTGGGCGAGCGCCCCCTCATCGAGAA  
TTTCATCGACCACTGCCGTA

**N87**

TAATTTTAAGAATAGTTTGTAAATTTTATATTTACTATAACCATACTTTTGTCT  
AATTATCACAAATTTATATCTTATATCNCCATAATGATATCTAAGGGAGAA

ATTGGGCCACGACTTTAGGATTGCCTCGGCGCCTGCGAGGGGGGGCCGGGG  
 AAAATAATTATCCCCTTGCACAGGCGATGTTAGGATTTATCTAAGTGAGTA  
 GGACTCCGGATAGCCGGTGAACAGTTAATATATCGATGTATTTATATCGCCA  
 TACCACGGGCAGTAAAGGCATCACGTAGATACACCAAAGACCGAACACCCT  
 ACAAATGCCTAGCTACTCTAGGTCAGAGAGACGCTCTTGCGTACAGGTGTT  
 GTCATCAGATATTAGAATTCGTGAGACTCCCTGCGCGGGTAGAAGATACAG  
 GACAGCAGCTTTCGGAATCACGTCTGATATAGCTTCAAATAGTCAATACTT  
 CGTATAGATGCTCGTGCTTATAAAAGATTTGATCAATAATTGCTGGGTTTAT  
 TACTGCAATTCCTGATGTAAAAAGATGTGTATTGACGTTACATCAGACT  
 TCCAGTATTCACCATAACTAGATCCAGGTAAGCTATACAAGATGACACAAG  
 TCCCGGAGAATGTGCTTAATTGATTTAGAGCTTATAGCTCTCTCTAAACATA  
 TGTCGCTATGTGTCATCAATTGCCTCAGAGTGTTCAGCGACGTTAAATA  
 AAGCTAGCTCGTTGTACTTGCATAGCAGTGTATATAGATATCTGACGCTATA  
 AATTACTCGATAGAAGAAAGTTTGTACTGTGATTATATGCAGTATTATAT  
 TTTTAGATGTATCCTCCGGCTCTATCTAGACCTACAGATGTCATCTTCGATTC  
 GATATCAAGCCTGGCATGATAAATTCGGTTTAAATCAATATCTATTCATCTCA  
 ACAGGCTCGAGTTATTACCTATGCATACGCCGAGGAG

**N99**

GGAGCCCCCTATGGACCAGCTCTTGGGCTATTTATACCTAGACTCAAATAGT  
 GGCAGTTTTACACATAGACCCTTAACACTAGAAATGCCCTGAGGGGCTCAA  
 ACGAAAAGATACAAGATATTTTCGGGACGGACGACCACGATGAAGTCTTGAT  
 CTTCTAGGTTCCACGCAATCTTCAGATGTTCTCCATGTGGACAGTAGTCCA  
 ATCCGCGATCCACCCCGGTTTTGAAGTCAAAAACCGGAAATTTTAGGAAG  
 TCTCCATCTGGGCCGGATTTGGCCATCTTCTCTATCTTGCAAGTTGTGGCC  
 CACTCACCAAGGCGTGGAACGGAGAGGTGATTTTACGGCTCGCGACACCG  
 TCTCGCACCCATGGTACACTTGTGGGCCGAGCCTAACTTGGGCCGCAACCC  
 AAGCCATCTGCCAGTCCACTCGGCAGGGATTTTGGCACGCCAAGTCCGCTCT  
 GCGCCGCCTTTCAGTCCACTCGAGAGGGCTTTCCTCCAGGTCTCACTTAGG  
 AGGTGCTCGCTCGTCCATCAGCAAGTCTTCGGTGTGACAAGCTCTCATCAC  
 AGTCAGGACGAAAGACCAGACTCGGTACACTTGAGTCCCTCAATGTCCATC  
 CATCCGTGAACCTCCCCCTTGATCTTTGCTTCAACTTCTTCAACTTCTCCCCCT  
 TTGGCATCAATTAACCAAAAAGACACAATCACTAAGCATCATCATCTTCTG  
 CCGGAGGGGGCTCCA

**N103**

GGTCAGGCCCGATGGGGTATCCATGCAATATCCTCTGCGCGCCAGTCATGC  
 ATTTTCAAAGGATTATGAAATTTTTTGTGTGAAGACCGTTACAAGTAATATT  
 AGCGGGACGGAAGATATGCCTGTTTCGGACAACAATATTCATTTCTGGCCTG  
 AGCTTGCAGGATTTAATCATGTAGTCGAATACAGCGGTCAGCGTGAAGTAA  
 TTTACTGGCACCATCAAAGCCGAAGCTCTTTGTTGCACCGGGGCTGACCA

**N104**

GGTCAGGCCCCCCCTGCATATTTAGAAGTCGGTCTTAGAAGGCCATTTTGACG  
 GGTGAGACTTGCTGAGTATTAGTTGCTCAGGGCTTCGTTGTTTAAACTTTA  
 CCGCAGGTATTCCTGATTGCAACTTCTGCGGTTGCTGTGGAAAAGTAGTGTG  
 GGAGGGAGAACCGCCCCAGGATTGGTCCATTGAGTGCGTTCCTCTTCTCCC  
 CGCCAGCATCAAGGCGAGAGAGATCCTGTCTACGGGTCTGATCTCTTCTTAC  
 CGTACCCGCGTCTGCCTGCCTGTTTCATCATATGCCACAGCTCCCCTCACCTTC  
 GATGCTAACTTTGACGCTTCTGTAATGACTCTGCAATCTGCTCGCGTGGGCC

CAGTCGTATGCTCCACAGCTGAGAGATTGTTTTTCCTCACACCTCCTTCTATC  
TCGCCACCGGATTCTACTTGGCTTTATCGGTTTCTCACTATCCCAATAGGTC  
GTCTTCCCTCTACGTGCTCTCTGACGAACCTCCCCCAAATAGAGACCTTCAT  
TTTCTAATCACTTTTCCATTCTTCTCTACATCTACCTAAACTGTTTCGCACAA  
CCCCCTTTACCTCCTTTTCGGATCCTAGCCTATCATGTGTGTTTTTCCCAAC  
CTCCTATCCTGATTCCATGCTTTCTTGCAATATATTACTGTCTCCCCAGTCT  
AACACATTCTGTTCCGATACTCTGAACACCCGTC

**N106**

GGTCAGGCCCCAGTCAATCCTGTTTCCTTGCTGATCTGACGCTCCCCTTTTC  
ACGCATTGGCGTGAGGTGGCTCCGCCGCTGAGGCTGAGTTAGAGAGGGCTA  
TTGGAGCTTTGCGATTCGGTTCAGCTACTATCGCGTATCGCGCAATCAATT  
TACGTATGGATGCACTGATGCTCAGCGATTCCGATACCTTGGGAAGTGCTTT  
CAAGCGTGCCTTTTACCGGGTTGACGGTATAACGATGTATGCGTGTTGGATG  
GTGTGGGCGGGGATACTGATCTGGGACCTGCTTGGTTCGAAAGCTCAGGT  
ATCCACACGGTTGTCCTGATTCTGATAGGCCTTTTAAATCCTTTCCTGTTTCT  
GCTGCTTGGCCTTTGGCGTTTACCCGGTTTGCTGACGGCGTTGATTATCATC  
GGTATCAATATCAGGTTTTTATTCGCTTGGCTTTGGGCCTGACCA

**N116**

GAATTTCGATTTTCGGGCCGTGGCAAAGAAAGTGATATACTGGAGCTATTG  
CTCAAGAGTTGCCCGCTGATAGAGCGGTTTTAGGGCTAAAATAGCTGTAAA  
ACTTACTAGGTCCTAGTCTCAGATGTGATATGAACTGCAAATATTTATCTTA  
CTGTTAAAATTTATTTATTTACTTCTATTATACCTTCCCTTTTTTGGTGTGCTGC  
GCAATTGCTGTTTTTCTTTGAATTGGCCGGGCTAGTTCTAGTTTGTTTTATGG  
AATGGTGACTTAGATGAGTTGAGATTAACCTTGGCCTGCGAATAAACTGGTC  
CCTGTACGGCCCGAAAATCACTAGTGAATTC

**N119**

TGCACTGGAGCTAGAGTGAAAAATATCAAGTATATTATAGCTAGCTAGGT  
ACATATACTGCAGATATGCAGGGAGATGAACTCTGAAATAATTTATGGGAT  
TAAGAGCATGGAGTTGAGCWWGACAGGCTCATGCATGTACTGTTTTGTATA  
AGACAGACGACAGCCTTTACATTACAGGCATGGATATATATATACACGA  
ACGGCTTTATTATTGGTTA

**Secuencias de bandas diferenciales según el nivel de ploidía que no presentaron anotaciones funcionales. Fueron obtenidas por el Dr. Luciano Martelotto (Tesis Doctoral, 2006) y por la Dra. Silvina Pessino (comunicación personal) mediante Display Diferencial.**

**L14**

TATGTCACAAATGCACAACAGCGCATCAGTTTTAAGTTTTAACAGAGCAGC  
AAAGCGCGCTTATCTTCCTAGTTTCTATATATGCTACTAAAATTTTGCTAGG  
CTCAAGCCAGGAGCTCGAGCTTCTCTCATACTGAGGTGGACTACTCTATATT  
ATTAGCTCATGAGTCATGATAGGGTTACGATCGCCACCCACTAGACGTGCTA  
TAACATGAGCAATATATGTGCCCTCAGTCTAACACAGCTGACAACATCTCTT  
TACAAGCCATTTGTGCTAAATGGTGTGCCCAACATTTCCGCGATCAACAGCG  
ATCAATATACCTTGCAACTTGCATGCCTCCACTGGATTCTCCTTCCCGTAAA  
CCTTTCCGGCTGGAGACGAGACGATGGTGTGATGGCAGGGCGTTTCAGCCTGC

GATTTTCTCCTCTCGCATTGGATCTGGCCAGGTACAAGTACAACCTAGATATG  
AGCCAACGGGAGAGGTTTCGATAGCTCCACCCTCTCGCTTGCCCGGA

**L20**

TTGAAACATTAATGTTCAAGGTTTACCTGCCACAGGCTGTAAACATCAGATT  
CCATATATGAAGTCCAGTACTGATCTTACTAATTTTCAGCTTTTAGACAATAA  
CGTTCAATCCCAAATATCGAAACAAAGCTTCAGCATAACATTTAGAATATCA  
GAAACACAAAACTGTGGGAAGGACAATCAATACCTACCTGAGAATTTTTT  
TTCCCATGTAAAAAGGAAGACTGTAGGGTTCCAACAGAGCTCATGAGCCCC  
AACTGTGAATAAACTATCTAAGTTCAGCTAATGATCTGGACTTCTAGATCT  
TCCATAGCCATGCCTTCTTGGAGAGCGTGATCTGCTCCTCATTCTCCCCGA  
CCCGTTGTGA.

**L24**

TAATACGGCGGCGAATAACGGGACAGTGGGGGCGCCGACAGTGCGCCAGA  
GAATGACAGCTTACAAGATCACACACACCTTGATAATTTACATGCAAATAA  
GCAAGAGAATATAGGAAAGCATAATGCTATTCATGCTTCACTGGCGACGAA  
ACACAGCAATAGATAAGGTACTGGCCAACATTGTCGACAACAATCGGTGGA  
TGCTTCTCTCGCGCAAGGAATCTGAGTGATAGGGATCCAACCAGATAGAGA  
AGCAAACCTGTCGGCAAGACTAAGCATCATGCAGTGAGAGGGCCGCA

**L29**

TTCAGGATTCAGCTAGTTTAGCTTCTTTGTCCTTAAGCGCTTGGTGTGTTAGG  
GAGTGGCTATTCTTCAAGGACTCCAAGGAACTGGCAGTCTCACGAAACGAC  
TTCTCAAAGCCGTCTCGCGATTTTTGCAGCCGCTGATACTCGTCATTTCTTTG  
GCGCATCTTTTTCTATGGAATGATGCAATTCATATGCTTCTGTAAGCAAAGA  
ATAGGCTAGTTGGGATTCTTGACGGGTTACCTCATGGTTTTTCTTCTCAATAT  
CCAGCTGTCTACGAAGGTCTGCCTA

**L45**

CATATGGATACAAGCATGGGCACTTTTATGTGTACAAGTAGCCTTTGTTCTA  
CCCCATGCGCTGAACAGAACGTGAGCCTTTGTCCGGTTGCATGGCGTGAAG  
TTCAGAATCAGTATACTTTTCATCAGTAACTATTCAGTATGAGCAAAC  
CAGTAGCCTGGTGGTTCGGTGGAGATATTAGTTTGAACCTGTATATGCTAGCT  
GATGCACCTGTTATCTTGTGAGATAAAAAAGTGTTTTTATTTTATGAATGCAC  
TTAAATTAGCTCTTGGGCAAAAAAAAAAAAAA

**L47**

CTGCTGGTTCAGATGTTTCTGAAGTAGATAGTTGAAGATTTAGTTGTCTGAG  
CTGTTTCAGATTGCTATCAAGGCATCAGAGATTGGTTCCGTTGAAACTGCGG  
CTGCACAGCTGTTCCAGACGCAGCTGCTCATCCACTCCTCACACTGTAGT  
CGCAGCAGCTGCACATGAACTTGTAGCTGTGGCGAGCAGGCACATATTCAG  
GGCGTGTTAGTTTGGTAGAAGTAGCTGTTGAGTTCAATGTATTTAACAAACA  
GTGGTGCTTGTGCTGCAGCATTTTGTAGTATGATGATTGCTGGATGGCACCA  
ATCTTGGTTTTGTTTTCTTTCTTATATAGCATGTGTTTTTATTGTTGCATCACA  
GAAATCAGGATGAAAGCATTCTACATTATGCTAGCACATTTGTCCGTGAGA  
GCTGTTTACTTGTTTCATTGTGCAAGTATCGACTTCTTGGTGGCAAAAAAAAA  
AAAAA

**L48**

GTACTTTACACATACTCCGTACAGGACACCAAGTACCTAGAAGCCTAGCTA  
AAACACGGAACCAACGATGACGAGCTCACTTAGAGTCACCAAAGAGTGAA  
AATGGATGTAGAAAACGCTATTTGTGTACCGTATACAGAATGAAAAATCAT  
TGTTTGCACGGATGCACTACCTGGTGACCCTAGCCCA

**L49**

CTGAAAATCCTTTTATTTTGCCATTGGATGCCAAGGCATGGCCCCGGTGGCA  
AAAAAAAAGTTCCAATCGGGATTTCACTTCTTAATGGCAAAAACAGAAACA  
GTCCATTCGTGGCGTTACTTACAGCCGCACGGACACGCAGCCTAAGATTGTA  
AGCAAAAGTAGATGGCTTCAAGTTTCAGGTGCTACCGTGCTAAAAGATCCT  
ATAGTAAACCTAATAACCTAGCATAGCTAAGGTTCCAAGTTGAGAGTGGCA  
TCAGCGTTTGCCCTCAATTTCCATTTTGCTCCGGATGATGCACCTTCTTAGT  
TCCTTCATCTAGGTACACCAGGTCTTCTGCAACCGCGACATGCCATAAAGC  
AACTCTTCCAGATTGTCGGTAGGAGCTGCCTTCTTACCCTAGCCCA

**L53**

TGCTACCTTGGTCGGCTACGCCTTCGTTAGGATAGTTGTCATGCCTTCAATG  
ATCTGCCTTGGTTCGGCGCACATAAGAGGGGATCTAGCACTCAAATGATG  
CAACTTCAAAGCTAAGATCTGAGCGACATGCCTCCATGATCACTCATTTTTT  
CTAGAGGGA

**L55**

AGTTGTCTGAGTTATTACACCAGTGCCTAGGTTAGGGCAACCTAAAACAT  
GCAACTCTTGGAGATTGGTGAGTAGCTGGATGCTCCCCGGCAAAGTTTGGA  
TGCCATTGCAATTGATAATCTCGAGACTCTCGAGGCTTGTCAGGTCTCCGAA  
CCAGTCCGGCAGCGCCTCA

**L61**

AACTATGGTTCTCAACAGTAGTTTTGTATGAAATACTAGTACATTTTCACAC  
AAAAATCAGATTCCAATGTTCCAGCGGGGCGACAGCGA

**L66**

AAACAAGGACAAGTGACGACTGCAGCATCAGAATGTAGATGTACCATGAA  
AACGTGTCGTCAATCCAATATTGTGCATACAGAAGTTGCTTCCTCATATGTA  
CTCTATACTCCTTTGGCTCAACAATGATGTTTCAGACTACAAATGAAGCAAAT  
AAATCTATACATGGAACTGAAACCACCTAGTAAACTTCTGGGCAAGGTAT  
TCATTCGCCTATGCTACTTTCTGTGTACA

**L72**

CATTGTAGGAGCAGCTTGTAACCTTTCAGGGTTCACAAAATATATA  
TGTTTCATGTTCTTTTTACATTATTTTTCTCTCTGTGGATGAAATTCATCTT  
CTTAGTCCACAGTGCCAAGTTGCACAGCAGGAGACCATGCATAGTCACAAA  
ATAAAGAGTAATGCTAACCTCTCTTTGCTAGGTTCTGATTATCTATTTTACTT  
TACAGGTGGTCTATCTAGGTAATCTTATCCTCTACAGAATCTTACAGACATA  
CTATCTTTCTGTACCGGCCTCTAGA

**L73**

TGGTTTTTGACAACCACATGCTCGTATCTGCTTGATGTATTTTTATGCAATAT  
TTGTCACATCCATGATACATGCTTTACTTTCTGCGTGCTATGCCATACCCA

ATATGTTCCCTACAATTATGATGACTCGAGCAGTTGTTAAATCCTGCTCGGGT  
GCGATAATTTAAGCAGCTGTCAAATCCTGCTTCCACTAATTAATTTGATCTT  
GTATAAAAGAAATGGCTAAGGAATCAGATGCTAGTAAGCAAGCTTTGCGTC  
AGCCGTAAGCCTCATCCATTAAGTCCTATTGAATAATTAGTACTCAGGGTTT  
GCCCCACTTTTGGCTTACAGGTACGGAGTTTCTCTTCTGTGCCTGTGTTGGT  
AGAGGGTCGTGTTCCCATCAGGAATGAGGAGAGTCATGATCTGGCAAGGAG  
CGGCGTCAGTGGGA

**L75**

GAGAGCCTACCCTGGTTGGTGTGGTGCAAGTATGATAGTGTGCATGTAATGG  
ATGTAGATCTTTAAATGGTTTTGTAAATAGATGTCATGTATTCCGCCCTCA  
GGCTTACCCATCGGCTGACTGTCGTGGTAAACACTGAGATAACAGATCGAA  
ATATCAATTCATCTGATATATGCAAGTCCCTATCTAGATAATTTGTAGTGGG  
GAATTCTAAAGCTTTGAGGAATTGTACCTATTCAGTCCGTAAGTACTGTA  
TGTAAAGCAGTTGAACTGCACTTTGGAGTAATGTATATGCTACCTAAACACG  
ATGTTTGTGGTGTGCTTTGAACTATACATGTGGTCTGTTTGACAAAAAAAAA  
AAAA

**L76**

CAAGGCTTTCATGATTTGACACGCCATTGTGATATGAATGTCAAATCCTAAA  
AGGCTCGCAGATTGAGACCCCAAGTTCCTTCGCCACTGTGTTTACAGGGAA  
GTTGTGCATCCCACAAGAGAAGGGAAACACCACAACAGCCAAACTGTCTCC  
ATACTTTAAAGCTTGACTGAACTACAACACTAGTTACAATATCATGTAAGG  
C

**L 85**

CATACTAAAGCGATCGAGCGGAACAAGAGACTAAAGCAAAGGATCAGTTG  
TCTGTGAGATTCTTCATTGCGGGCTCAGAGCGATCGAGCGTTTCTAAAGATG  
TTATTAATCGAAAAGTTAAGTAGCTGTAAGTCTCTTAAAAGGAACGCTCGCT  
GCCTTTGATATGTATATGGAATCAAAGCTCCAGGTGAGTTTTATTGAACCAC  
AATTCGACAAAAAAAAAAAAAAAAA

**L86**

CGACACCGCCGGAAGGTTACGCAATGTGAAGCCTGACAAGGGAAACCACA  
CTCTGCAGATTGCTTGGCATTGGTACCGAGAGAGCATGGTTCCACCTCCTGT  
GATGCAGCCAGCATTCCCTGTGATGCAGTCCCCATTTCTGCTAGGTGTGTAG  
AAAGATGCAGAGCCGTCTCAGTGGCCTCTGGTA

**L87**

TCATCTTGGGAAGGATCGAGTAAGATTTACGGAACCTTGGACAATTTTTGGA  
TAGGAAAAATCCCCCTACACCAAGAAGTTTGTGAAAACCGGGCTGGCAT  
AAACGGCATCCCTTCGCACAGAGAACTCCCCCAGAGCACTACTCCTCGACT  
CCCGGACAAAAAAAAAAAAAAAAA

**L90**

CAGTACTATTTCATATTCGAACTAAGCAGCTTCTTCCCTGCAAAGCAGTCAA  
AAAAAGTTTTGTGGACCCTAAGTTGGCCCTTCGTTTCGATGCGATTTATATA  
TCAAGCAAAGTATCTGTTTCATACTTTCATCCAATAGTTGATGGCGACAAA  
AAAAAAAAAAAAAAAAAAGTTTTGTGGACAAAAAAAAAAAAA

**L94**

CCGTTCCGGTCATGATGCTCCAGATATGCCCTCGCTGCGTCAATCTTTTCGAG  
TTGCCTATGTACATGGTTTCATGTAGCGACGTTGGTGTAGCTATGACACTAC  
CTCATAAAGTAAGGCTAGTGAGCGTAGCGGGGTCAAAGATTTTTTTTGT  
ACATCGTCCGTTCCGTGCAACCAGTTTGTGCCCTGACGTTGCACGGTTTGT  
TAGGACCTGAAAGATGTGCAGCAGAGACGATTCATATCAGCTCAATTTTGA  
CAAAAAAAAAAAAAA

**L95**

ACAAGGGTTGGCTGTGGACAACTGATCTTGGTGAAGGAATTTATTGAATGA  
ATGAGTACGAAAGGAAGCATGCACATGTGGTGTAAATACATCTGGAAATTCT  
TAACGGTTGCTGAACAAATTCGAAAAAGTAGCACCTATACTTGCAGATTT  
GATGTCCTGAACTTACAACATCATTGTTCTATCTCAACAAACAGACTCGCCT  
CGGGATGCAAGCCTGCTTCCTCCAGCGTCAGTGGGA

**L774 (DDT32774cm1)**

TTAGGAAGGTGCCGACAAAGCCAGGGTGCTAATCCACAAGGTTGATGAGAT  
AACGGACGAGGTAGATTCAATGGTCGTCCTATAAACGGCA

**L784 (DDT32784cm1)**

GATTGGTTCCAGCTCTCAAATGAAAGGATGATATCACATCCGCTAAATGGA  
CAATCTTCCAGTTGCATGTTCACTTTATGCAAATCTAGAGTGTCTG

**L794 (DDT32794cm)**

CGATTCAGGGACAACACTTGCAACTACCAACTCGAGAAGGGGACAACCTGCT  
ACTACTGCTACTACTACTGCATTAGAGAGACAG

**L802 (DDT32802cm)**

TTATTAAGCTGCTAACATGACGAATTCTGCAACCCAACCTTGACCAGACGA  
GCAAGAATTAGCTTGGCTTGTGGATAACCGGAAAACCGCACAAATAGGTA  
TTGATGCAGGGGTTGAGGGGGATCTGAGTATTTGCCCTCTGAACACTCATA  
TATG

**L812 (DDT32812cm)**

GAAAGTGGAACAGAGATAAATTGTGCTGCAAGATTCTGTTAGCAGTATGCT  
TGAGATGGGTGTGCATGCACGGCTGTAGCAAGACAACCTATTGAACTGTTGT  
ACTACCTGCTGTCAACTGTGTTTGCGGGATT

**L832 (DDT32832cm)**

ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCC  
AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCAT  
AGGCTCCGCC

**L834-1 (DDT32834cm1)**

CGGCAACGGTGTCAAGCAGTCTAAAACCTCAAGTTATATTGAGGTCAACGG  
TTTTTTCAGTATAAAATTGAGCTGCTCTAGGTTGCTTAAGAATAGTTCGATC  
TCGATGTTTTGAGCTAAATTTGTAGACTAAAGCTGTGGCTAAACACCAAAA  
AACTCTGAAGTAGAGAT

**L834-3 (DDT32834cm3)**

CGGCCACCGTGGTGGTGTAAACACCCTACGTCCTCTTGGGGAGATCTCAATAT  
ATGGCAAAGAACCCAAGTTACAACCCAGTGGAGACTACTGTAGTGGTGAAG  
ATGCCTTGAGTGGAAGTGAATGGAGTGGGAAGTAACATGCAGAACATGTGA  
GTGAATTTTGGTTCCTCTGG

**L844-2 (DDT32844cm2)**

CCTGGATAGCATACTCAGGCAAGGTGGAGAACTCATGGTCGTGCTGCAAGA  
TGATTAACCACTAACCATGCTGCCCAAATCTCGAAAAGCCACCAAAAAGT  
CTCTTCTATGTGCTGG

**L852 (DDT32852cm)**

GCCCAGCAAAGAAAGTTTAACTGCTGCAACAGATTTTAAGTGAAGGTGGGG  
AGCATCCCACAAACAAGCGATTTGGATGTCACCTTGCGGAAGCCACTGTGT  
ATA

**L852-1(DDT32852cm1)**

GGTGCTAACCAGTCAGGAAAAAAGTTCTATCCCAGACGCCAACCCCTGATAT  
TGCTGATGCAGTGCAATTCTAGTGAAGACAAATAGGCTTATTGGCGTGGTTT  
GCAGGGTGTTCAAGTGTAGGCT

**L852-2 (DDT32852cm2)**

GCCCAGCAAAGAAAGTTTAACTGCTGCAACAGATTTTAAGTGAAGGTGGGG  
AGCATCCCACAAACAAGCGATTTGGATGTCACCTTGCGGAAGCCACTGTGT  
ATA

**L854 (DDT32854cm2)**

GAAAGTGGAACCCACCGGAGCCGTAGCGAAAGCATGGTCTTTAATAGGGCG  
ATTGTTGCACTGCTTATGGACCCAACCTGGGTGAATCATNTCCATGACCAGG  
TGAACCT

**L884 (DDT32884cm)**

ATCACTAGTGAAATTCGCGGCCGCCTGCAGGTCGACCATATGGAGAAGCTC  
CCAACGCGTTGGATGCAGTAGCTTGAGTATATCTACAGTGTTACCTAAATT  
AGT

## ANEXO I -Sección B

**Secuencias extendidas y derivadas de los experimentos de RACE de cada uno de los clones estudiados****Secuencia consenso N7**

AGGAAGGTGCTAACCAGTCAGGAAAAAGTTCTATCCCAGACGCCAACCCCTG  
 ATATTGCTGATGCGATGCAATTCTAAGTGAAGACAAATAGGCTTATTTGGCG  
 TGGTTTGCAGGTGTTTACAGATGTAGGCTTTTATTTTGCAAATGTAGCACGATG  
 GCTCCCCTACGCCACAGTGCTTGTAGGTGTAAGGAACATTCAAAGAAG  
 GTACCCCTCAGCTACTACGAAGAAGTATCTATGTTACCGCAGTGCAGGTGTA  
 AAAGCAGGTGCGTCGCCGAAGTGCAGGTGTAACAATGTTTGTATGGTTCG  
 CACAGAAATTACATCAGTAACCTTCCTAATGTGCTACTACCCCTCAGCCTGC  
 CACAATGTTTATTTGTGTTTGTAACTCAATGTTATACGTCATGTTCTGTGT  
 ATGAACAATCTTCAGCTGTAATATATGCTAAGTGTGTAGCTCATTCTGTTGT  
 AACCGCCCTTGTCTATTGCACCTTCCTAGAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAGCGGCCGCTGAATTCTGG

**Secuencia consenso N13**

ATTCGATTACTACTATAGGGCTCGAGCGGCCGCCGGCAGGTTGAGAGG  
 TACCGTTGGTACTCAGGACGCGCATTGGAGGGTATGAGTTCCTGAAGACG  
 ATGGCGACTACACCAGAGCCCGAGACGACCAGCTCTACTCCAGATACGTGA  
 TCAGGAAACAAGTACATGGAGCTACATCAATCTAACGGYGGTYCAGAAGAC  
 CAACTAARCCCAARCATGACCTRGAGKGCATCAGTCRTGTTTGGAAAGG  
 AGGAGTTGYCCAAGTGATCAAGSGTTGYCTRGGRCAAAGCCTCACACCY  
 RSCAAACTAACTTTAGATTAGWTYGTTCAGTTTTCRACAACAAAGGTAGG  
 CCCCTTAGSCATTTACACYTAYTTAGGGAGTYCACCACCCTMTAYTTGWW  
 TGTGCATTTAYGTTYARGAGKTGTWTGAAACCCACATGCATAAACCCCTG  
 ATAWCCTTTGTAGTATGAAAAGATCGAAGTCTATACTTATCCATGCCTAAA  
 CGGTACAAGTCGAGTGAAGATATCACTCGCGAGATGATATGGCTCAAAA  
 TGTTTTACAAAAGTTAATTTGAAAAGCAATATGATATGATATGATGTGATG  
 GAAAATTGAGACCGGGTGGAGATATGGGTATATTATATTGAAAACGACG  
 AGCATGGTTTGGTGGACTCCGCCAAGTCATTTAAGACCCCGTCCGCCGTT  
 GA

**Secuencia consenso N17**

ACTCACTATNCGGNTCGAGCGGCCGCCGGCAGGTATTTTTGCATGCC  
 AATTTTAATTTTGCATCAGTTTATTTTCTTAAACCATGCATTTTCGAAAAA  
 AAATTTTAGTGGCCGGTCAGCTACATCTTGTGGCCGGCCGACCAAGCCACTG  
 TTGGCGTTCCTTAGCCAATTAGATATAAATCCGCAAGCGCACGGAAAACGC  
 GTTGTAGCCTTTCACCCGAGAGTATCCAGGGTATCGAATCCACGAGGAAC  
 GAGGAGTATAATAATCACACAACACAATGGCCAACTAGCGAAGCTAGCA  
 AGTAGAGGATAAAAAGACAGGAGAGAAGATTCTCAACAAATGTTAAAGCAC  
 CTTGCTATACTAAACATGCAAAGCACTATACTTGTGTCATGTAAACATAGACAC  
 TAAAGCTGCANAGCCACTGCCNNGCACCAGAGTGCNNATAACCCAGCCGAG  
 GAAGGCGGGCGCTATAGGTTTCAGCTATATCGTCAGCGAGTCTATCTGAAC  
 GAACGGGATTGGAATACTGAGAACACAAAAGCTGTCACCNANTGTGTCTGC  
 CAACCANAGCTGCGTGGGATAGCCGACCAACCAAGGCAAACCTCTAAGTATG  
 AACACCGTCGTGCCACTCAGAGAGTTTACTACTTACTGGTANTGAGCCAGA

AAAAGCAAGCACTACTATATAGCCAACACTATAAGCATGCACAANGNTCCC  
 GTAAANTAGACNAGTACAAAACCTAAGTATGAAAGTAAAGTAAGANNGTAG  
 TAACNAAGGTGAGAGACAAGGAACCTTACTCNATCCTANAACCTGAATNACT  
 ANNGAATTCNCGGCCCGCCTGCNNGTCNACNTATGGANAANCTCCCANNGCG  
 TTGGATGCATNNNNTNGANTATNCNAATAGNGNCNCCTA

### Secuencias derivadas de N22

#### A43 (N22)

GCAGGTCTCCCTGTTTTTCGAGGTCTGTATAAACGCATGCCAGAGGAATGGAT  
 CTTTGTGGCAGCGAAGACATTCAACGATCCTGATGTTACCGGTGAAGTTCC  
 ACTGCAAACAGAAGGAGCGTTGATTGGCCTTGCTGCGAACAAGAATCTGTT  
 CCGCTTGTATGGTTTTTGCATGACCTCAAACGAGCGGGAGCTGTGCAATCTC  
 TAGAGATTTTCGGGCGCCCGCTGGTGTACCACTATATGAGAGATCTCCCAAC  
 GCGGGATGTGCATATCGTGAATACTCTATAGTGACACCAATATATCGTGGC  
 GTTCTCGTGCATATCTCTGTTTCTGTGAGAAAGTGATATGCTCTCACATTCC  
 CACAAAACACACGACGCAAAAACATAGAGTGAAGACCGTGGTGTGTACTG  
 AGAGAGAGCTAACTCACAAATAGTGTTTTGCGCACTGTGCGCTTCCCATG  
 GGGAAAACGTGGCGTGCCATGTGCTATTGTGACTCCCAAAACGCGCGGAGA  
 GACGCTGTGTGCATAGTGCGCGCTTCCCCGCCTCCACTCACAGAGACTGC  
 GCTCTCGGGCTTTCGTGTGCTGAGAGCTATCTCATCTCACTCAGGGGGGGAA  
 ATACGTTTCTCCACAGACACGGGATATAACATGAAAAAGCATATGAGCACA  
 AAAGGCCCCCAAAGTGTCCGAATCCTATAAAGCCCCGCTTTTCGGGCTTTTT  
 CCCAAGGCCCCCCCTCAGAAGAGACTCATAAATTGCCGCTCTCATCCCAA  
 ATGTGGAAAATTCCCAGGACCCTATAGATACTCCCGGGGTTTCCCCCTGAAA  
 TCTCTCGCGCGCTCTTCTCCTC

#### A44 (N22)

CGCGCTCTTCACACTCAAGGACCAATACGACCTCAGTGACGACACCGTCAT  
 CTGCCTCCTCTGGGACATGATCACCGCTGGCGCGGACACAACGGTGATAGC  
 GGTGGAGTGGGCGATAGCGGAACCTGGTGAGAAACCCCATGGTGCAAAGA  
 AGCTGCACGAGGAGCTGGACCGTGTTCATCGGCTGTGACCGAATCTTGTCCG  
 AGACGGACTTCTCCAACCTCCCCTACTTGCAAGCCGGCGTCATAGGAGACC  
 CTCCGTCTGCTCCCAGGACACCGTTGATGCTCCCGCACAATGCTGAGGCCA  
 GTTCCATATCCTCGGATATGACATCCCCAAGGGTACCAAAGTCATCATCTA  
 TGTGTGCGGAGTGGCCCGCGATCCAAACATATTGGAGATACCTCTCGAGCC  
 CTATATTGAGTAATTACTATTGAAAAACCTGCCGCCTGCACGTCGACCATAT  
 GGCAGAGCTCCCACCGCGATAGACGCATAGCTTGAATAGTGCATATTGCCT  
 CCCACACACATGACGTACTCCGGTCAGACGTGTTTCTGTGCGAAAGTGTTA  
 TCCTCTCTCAAAGGCGCTGAAACAGCGATCCCCAAAAACATAGGGTAAAGC  
 CCGGGGTGACTCATGAGAGAAATAACCTCAAATAATCGCGTTGCGCACAAC  
 AGCCCGCTTTCCTGGCGGGTTACCCTAGGATGCCCCCCCCGCTAAAAAAAAC  
 AGCAAAACCAACGCTCAAATCAGATTTAGGTAATGCAAGCCCCCTCTATTT  
 ACATCCCCCGGATTCCCTCTGTTAGGTACCCTCCGGTGGCTGTCAGTGTTG  
 AATCCTC

#### A45 (N22)

AATTCACATCCCTCGGCTGAAATCGCTCGACAAGGAGAAACATCTGCACTA  
 TTCGGCCCCCTTGCGAGAGCGGCCACTGCCGAGACCCCAACTTTAACAATCA  
 CGATCCCCCTGACAACGCGATCAGTTGATGTCTGCATGGCCCCGAAATTTAC

TGCTGTTCTGCTTGCATCTATATCGCCCCCATTTGCCACAGGGAGCTAT  
 CGATACTGTAGTAAATACGGCGTCCCCTTCCACAATCCCGTATAGCAGATCA  
 CCACATTGACCTGAACTGGAATTGCGCATTTCGACACCGACACCTCAAGAG  
 CTCGGCGTCGTCATAGCCATATCCGGTTCAGAGTGTAATCGGCCCCCGAAT  
 ACTATAACCGACCGCGCCGGAACCCTAAGCCTAAAGTGTGAAATCCGCTGT  
 GAATAATCTGACACCCAAGTACGCCCAGTGCAGAGCTCCCTACCCGCT  
 TTCCAACCGGACGAGTCGTCTGCCTTAGTGCACCAATGAATCCGCCAACA  
 CACGGGGATAGGCTGTTTGCCGATCGAAAGCTTTCCCGCTTCCTCACTTCT  
 GCCACGGTGCACCTCGGTCGTAACGCTAACAGCTCTCTCAGCTCACTCAA  
 TGCGCGGATTCACCTTATCCGCGGATAACGCCGCATAACACATGAAAGCAC  
 ATGACCACCTAAATGCCCGCAACCGTTCAGAAGCCGTATAACCGCCCCTTT  
 TCCTTGAGTTTTCCCTCCCGTCCCCCTCCCTCTACAATCCCCCTCAAATC  
 CGACTCTCAN

**A46 (N22)**

ATATATATGGGGAATGGGGGCCAGTCCTATGGTCCCCGGCGCCATGGGAGG  
 CGCGGGAATTCGATTGCAGGTACCGCGGTTTTGGAGTTGTTTATAAAGGCTG  
 CCTGAGGAATGCAACTTTAGTGGCAGTGAAGAGATTAAGATCCTGATGT  
 TACCGGTGAAGTTCAATTTGAAAAAAATTTGAGTTGATTGGACTTGCTGT  
 GCACAGGAATCTGTTGCGCTCGTATGGGTTTGGCATGACCTCGGAAGAAAG  
 GGAGCTCTGCGATATTCGGTAAATCTGACGTGCCAACTGCATTATCGAACCC  
 GCCAAGGCGCGCGGAGAGCGGATAGCGTATTGAGTGCTCATCCGCTTCC  
 TCGCATCATGACTCGGTGCGCTCGGGTTCGTTTTGTTGCGGCGAGTGGTATCC  
 CCTCACTCCAAGGCGGTAATCCGGTTACCTTCAAATCAGGGGATAACGCC  
 AGGAATAACATGGTAGCATAAGGCCCCCAAAGATCTGCAACGTTACAAC  
 GCCGCTTGACGGCGTCTTTCCATAAGACCGCCCCCTACGAGCTTCAAAA  
 AATCGGCGCTCTAATCAAATTTGGGTAAACCGCACCTGAATTTTATATATCC  
 CCGCGTTTTCCCTGTAAACTCCC

**A47 (N22)**

TTTTCTGTGTCTTGGTGTGCGGGCTACCAACTTGTTCATGCAGGTGGCTCAGCTT  
 TTCATACTCGTTATCCAAGGCGAGCAGATGTTTGACCAGGTAGGCCATGGTG  
 GGATCTCCGGACTTTTTGATCTCTATCAAGATTGCAGTCCTCTTAAGCCAAG  
 TTGTGGCACATTGGTTCACGGGTGGAAAGAATCTCATAGGGCTATTTTGTAC  
 TGCCCCATGGTAGGTGAGCATAACCTTCTTACTGCACCACGAGCGACAAT  
 GTGACAGGCTTCTAAGAATCCAATGCCGACCGCTCTGCAGTACCAACCGTT  
 GAGATTGAGGTTCTCCCTGATATACAAAGTGACGGTCCATCCTCGTTCTTCC  
 CTTTCGGGGAACGCTAACCCGCAGTACTCCGGGTATCCTGGGATCCTGAGGT  
 GCCGCATGGTATCTCTTAACAGCTCGGGAAAGCCACTTTTTTCTCAGCAGTA  
 GCTGACCTGAGTGACACCTTCCATCTGATTGGGGTGCAGGACTTGCCCTGA  
 AACTCAACACGTTGACTACCACACAACGGTTACCAACTCTGCCCGTAAGG  
 AAACAACAACCACACACGCGAACTACACGGTTGCGCTAACAAAACGGAGA  
 AGAAACTCTCCTCACCTAAATAGGGGGACTGGACACCCCTAACGGATGAAC  
 AAAAGTCTTCTAATTGAAGGAGCGTGGAACCTCTCTAAGATGCCTCCCC  
 ACACCTGCCCTGATGAATTCCGGCGAGAGCTGATAATCACAATTGAAATTT  
 TCGCCCCCTCCGGTCTAACATATCCGA

**Secuencia N43 depurada de las contaminaciones con vector de clonado**

CCTCCCAGAGGGGCTACCAGCAAAGCTACACGGAGAGGATGGAAAGCTAT  
 AAACATGTTCAAGCGTAGCTTAAACAACGATAGATCTTCTGGATCAGAGG

AGTGGATCTAGCCAGAGCCTAGCGTCTAAGTCTAGGGTTGAGTTATAACAC  
 CCGCAAGAATAACAAGACTTGAATAACTAAGTGAAATACAATCTAAGCCTA  
 GCAAGTATGCAAAATAAAGGTACAACCTACTTGCAGGAGATAGAAGTAAA  
 GTAAAATATAACTGAATCATATTGAAGATAATCTAAATCCAAAAGATACAA  
 GAGTTTTACCACAGACTCGAAGATAACTCCGGAGACTCCGAGGCAATGCTC  
 GACTCAACTCTAACTCCCAGGACTATCCTAACTACTAAGATAACAGTGAGG  
 CACTCTAAGATAGAACTCTCTAACCAACTCCACACTTTACAAATGAGCTCTC  
 GGTGCAACCCTGGGCCTAGCTGTCTCCCCGTGGGAAATTGTTATTCGCTCAA  
 CAATTTTCCCCCACCATAACCAACCTGAAAGCATAAAGTGGTAAAGCCTGG  
 GTGTGCTCAAATGGACTGAGCTCAACTTAACTTAAATTTCCGTTTCGCGACT  
 TACTGGCCCCGCCTTTTCCACTCCGGGCAAACCCGTTTTTTGGCCACATCTG  
 CAACTTA

### Secuencias derivadas de N43

#### A11 (N43)

AAACTATGGCTCAACCATATTTTCGATCCTGTACCAAGTTGTGAATGCCCGAC  
 CCATTCATAATAGCTACTTAATGCACAAAGGTGCACATTCTTTTTTGGGGTG  
 GCCCTATATTGAGTGTCAATTGAGATGGAATCACTAGTGAATTCACGATCGCT  
 GCCTGTCGACCTTATGGGAGACCTCCCGTCGCCTTGGATGCATAGCTCGAAT  
 ATTCTATAATGTCACCTAAATAGCTTGTATTATGCCTGGTCATAACTGTGTC  
 CTGTGTGAGATTGTTATCCTCTCCTGTGTCCACTTGTTATACGATCCGGAATC  
 CTAAGTGTAAGACTGGGGTGCCTAATGAGTGAACCTCAGTCGTGTTAATT  
 GAGTTACCTCCCTGCCCTCTTTGCATTCCGCAAACCGGCCGTGTCCCCTGC  
 ATTAATGAAGCCGTGCCCGCTCGGTGAGAGGCTGTTTGCATATTGGGCGAT  
 ATGCCGCTTCCTCATTGACTGACTTTCTGCTCTCTCTCTCTGATGCGGTGA  
 GCGGTATCATCTCACTGCAAGGAGGTAATACAGCTATCCCCAAAATCAGGG  
 GATAACGCTCGACAGAACATGTGAATAACACGCCAACAAAAGGGTGAGAA  
 CCGTGACA

#### A12 (N43)

AACGCGTTGGATGCATAGCTTGAGGCAATGCTCGACTCAACTCTAGCACTCC  
 CACAGGCCACATCGCAGTCGCAGCAACGTCCAGGCACTCAGGCACTCCGCC  
 CGTCCGCCGTCATCACAATTCATGGCACCTCCATCCGCGGCTCCTCCGGCA  
 GCAAGCGTGCGCAATCGAGCATCCGACTCCAGGCCGTCCCGCAAGGCGGTG  
 TAGCTCTCGGTACGTTGTCGCTGAGCACGGCGCCGTCCCGCGTGTGATGG  
 CCCGCTCGAACCGCCAAGCGTCGTCGCTGTACCGGTACGCCTGCGTCGTGGC  
 GACGGACCCGTTGAGCACTTCATTGCCGCTCACGAGCATCCACCCCGCGGC  
 AGCCTGCTCGTCGGCCAGCTTCGCCTCGCGCTCGAACAGCCCCTCGGCCTCC  
 GCGTCCGTCTCCACGCTCAGGGAGTGCGTGAGGTTCCGCCCTCGTCACGTAGG  
 TGCCGTCCACGGCGGGGTCTCCGTACCATCTCCAGCCACAGCGGGTACTT  
 GGCCTCCGTCTCCACCTTGCCGACCACCTTCCGCGTCGCCGACGACCGGACC  
 ACGACCTCCGTGTCTGCTCCGCCTGCAGCCGGACGGTCTTGTTCGGCCCT  
 GCTCGGTGAACTCCACCAGGCTGCTCGCCTCAAGCTATGCATCCAACGCGTT  
 AAT

#### A13 (N43)

AACGCGTTGGATGCATAGCTTGAGGCAATGCTCGACTCAACTCTAACTCCCA  
 GAACTATCCTAACTACTAAGAAAAGAGCGAAGCACTCTAAGATAGAACTCT  
 CTACTCAAATCCACACCTTGCAAATGAGCTTCCGGGATCTATTTATAGCCAA

AAAAGTGAGCGGGTATTTATGTGTACTAGGTTCGGTGGAGGGACCAAGTCGG  
 TGGCCGGCCGACCTAGGTGGGTGGTTCGGGCGGCCACTCCCATGGCCGCCC  
 GCCATGGGATCCACCGCCTTCGTTGATCCAACGGCTGATCTTGCTCCCCCAT  
 GTCGGTTGGCAATTAGGACGGTTGGTTGCACGTGAAAAATGCGTCGAGGGG  
 TGGTCGGCCGGCCATGGCACGTGGTTCGGGCGGCCACGTGTCTGCCTCCGATT  
 TGTTGCATGTTGCTTCCTTCGGATCCTTGACATCACCAATCCTCGTCCCAACA  
 CTTTGTGACCCTCGGATACATTAGATGCTCTCTTGTATGGTTGGATCTTCACC  
 ATTTTGTGTTTGTGTTGACATCGGATCGATGATATGGCACATGAAGTGGTCAA  
 GGGACCATTCCATGTGGTTCGTGGACCCAATTCCACCCTGGAGGGTGGTCAG  
 ACGGCCACCAACCATGGTCCGGCGGCCATGGTATATGGTTCGGGCGGCCATG  
 GCCAGTCACCAATGTTGCCATCTTTGATATTTGGTGCATTTCCAGCCTTTT  
 TGAGTGATTTTTGGTACCTGGTAGTACTCAGGGGATATCAGTAGCACAATAA  
 TTAGAACTTGATGCTTCAGGTAATTATCCTCCAAGAATACCAAGACCTCAA  
 GCTATGCATCCAACGCGTTA

**A15 (N43)**

TACAGGCGGCCGGGAATTCATTAGGAAGTACATGAATAAACACCTGGAGCT  
 GGGCAGCCTATCCACATGTAACACCAAAGGTAGAAGTAAACAAGACTCCCG  
 TATTTGCCACAACAGCTTGTTTCAGCGGTCCATCAGAGGCCACCACCATGAG  
 GTCCACCACCATGCTTGTTTCAGTGACCCTTCAACATGATCAATAGAATCCA  
 TCGTACAACAATTTGCAACCCTCTACAGGACTAACAAATTGCTTTTTGCTTAAT  
 TAACTAGTTCAGCATTACAGGGAACAACACTGTCCCAAAGTCATCAACATTTA  
 AATAATGGGGCCCCATAATAAATACCAAACCAGGTGCAATTGCGAAGACT  
 CAGAACTAGATTACAAAGTATACACATATATCATGAGAACAAGAACAGTG  
 AAATAGAAAAGCACAGGTCCACAGGTCAAAACAAGGGTTTGTACTAATCGC  
 ATGTCATCGGAGCCAGCGAAGTTGAGCTAAAAGTTTAGATGCTCTCGAGTT  
 GCTTGAATGATTGGTTCATGCAGGCAAGTTTCGGAAATTGCAAAGATTGAA  
 GATTTAATTTCAATTGTGCTGGGCAAAGTGCACGCTAAATCAAAGATTGGA  
 CATCTAACTTTGATTGGACGGGAGGGATATGCGTCTCAGGGTTCTTGGGAGT  
 ATACTACTATAGAGAATATATGATCGCAGCGCAGAATAATTGAATTTGTAG  
 GTTGGATTAGCGATTGTACCTTTAGCTGGAGCCCTATAGTGAGTCGTATTAG  
 GATGGA

**A16 (N43)**

CCATCCTAATACGACTCACTATAGGGCATCACAGACCTGTTATTGCCTTAAA  
 CTTCCGTGGCCTAAACGGCCATAGTCCCTCTAAGAAGCTAGCTGCGGAGGG  
 ATGGCTCCGCATAGCTAGTTAGCAGGCTGAGGTCTCGTTCGCTAATGAATTC  
 CCGGCCGCCTGTAA

**Consenso N104**

GCGGTTCTTGCGCAAGTCCATACTCAACAGGGGTGAAAATGGAGAGTGTGA  
 CGATGAAGTCAATAATGGTCAGGCCCCCCTGCATATTTAGAAGTCGGTCTTA  
 GAAGGCCATTTTGACGGGTGAGACTTGCTGAGTATTAGTTGCTCAGGGCTTC  
 GTTGTTTAAACTTTACCGCAGGTATTCCTGATTGCAACTTCTGCGGTTGCT  
 GTGGAAAAGTAGTGTGGGAGGGAGAACCGCCCCAGGATTGGTCCATTGAGT  
 GCGTTCCCTCTTCTCCCCGCCAGCATCAAGGCGAGAGAGATCCTGTCTACGG  
 GTCTGATCTTCTACCGTACCGCGTCTGCCTGCCTGTTTCATCATATGCCA  
 CAGCTCCCCTCACCTTCGATGCTAACTTTGACGCTTCTGTAATGACTCTGCA  
 ATCTGCTCGCGTGGGCCAGTCGTATGCTCCACAGCTGAGAGATTGTTTTTC  
 CTCACACCTCCTTCTATCTCGCCCACCGGATTCTACTTGGCTTTATCGGTTTC

TCACTATCCCAATAGGTCGTCTTCCCTCTACGTGCTCTCTGACGAACTCCCC  
CAAATAGAGACCTTCATTTTCTAATCACTTTTCCATTCTTCTCTACATCTACC  
TAAACTGTTTCGCACAACCCCCCTTTACCTCCTTTTCGGATCCTAGCCTATCA  
TGTGTGTTTTTCCCAACCTCCTATCCTGATTCCATGCTTTCTTGCAATATATT  
ACTGTCTCCCCAGTCTAACACATTCTGTCCGATACTCTGAACACCCGTC  
TT

**Secuencias obtenidas a partir de N22 por medio de Display Diferencial Específico de Retrotransposones**

**ADD12-1 (N22)**

CCAGAAATCATCAACACTTTCTAAAAAAGAAGCAAAGTGTCTCTTTGAA  
TTTTAACTTACCACATTTGCCCTCTTGTCCCTTCTAGTTTATCTCGTTCATT  
TTTCTTTCAAGTCTAGCAATCTTCTTCTTATGCCCTGTTTTACCATAGTGCC  
AAAAATCTCTTGGTTCTTTGAATTGAGACCTCTTCCCCTTTGACTTTTCTCTT  
GATTTCTAGAAATTTTGGATTATTTGTCTCTCAATCATTCTGCCACTTCTTT  
GTGTGTTTAAAGCACTTCTTGATATAGATCTACTTGAGGTTTTCTACAAGTT  
TTCTCACTCAAAGAAAACCAATCACATCATCATATTTTAGGGTATTAGATC  
AGAAAATAGAATTACTACAACCGTTACCAAGCCATTCCAACCTTTAGGGA  
AAGAACGCAAATCAAGAGAGGGCTAATTTCTTCATCAAAGTTAATACTCA  
TAGAAGATAACTAAGTAGTAGTATTAATACTCATTAAATGATCCGTAAC  
AAATCCACCTTTTTCCATTTTCATAATAATAAATGCTTCATGAGAAATACC  
GTATTTGAGGTGAAGGGCA

**ADD13-2 (N22)**

CCTGAAATCATCACCAGATTGGTGAAGTACATGAAGTATTCAAAGATCTTG  
CTACGCTTGTTCATGCCCAAGGCGTTATAATTGAGGAAATCGACACAAACA  
TTGACAATTCTGCAGCAGCCACGAAAGAGGCAAAGACAGAAGTGGCCAAA  
GCAGCCAAGACTCAGAAATCAAATTCGTCCCTGCTCTGCTTAATTATGGTCA  
TCTTTGGGGTCGTGCTGCTTGTGTGATAATAGTTTTGGCAGCTTGAGACAC  
GCAATGTTTCATCTTATAAAGCTATTATATGGAGTTCAGAGTTGAAGATGGC  
ATCTTTGAAATTGGTTCTTGTCTAGCCATCATATTGTGTGCCAAAAAAAAAA  
AAAA

**ADD16-1 (N22)**

TTTTTTTTTTTTGGCAAAGACATCAAAGGGATATTATTGCACTGCACAGCAG  
GAAGTTGCATAATAAGCATCTGATGATCATGCATTATTGATTCAGTTATTCT  
TTTCTGAGATACTAGTAGCAAAGACAGTCCGAACATGCATGAAACGAAAA  
GGTGGCCGAAAGCTGGTGTGGATTGCCCTAGCGGTGAAACAGGAGCACGC  
GCGTCTTATTCCTTCTTCCCTTGCCACGCTACTGCAGCTAGGCAGCTCCTGCTGT  
TTCAAGCAGCGCAGTTGACAGGGGCGAGGTTGCATCGTCCGGGCAGGCTGT  
TGATGATTTCTGG

**ADD17-2 (N22)**

TTTTTTTTTTTTTCGCATAATAAGCATCTGATGATCATGCATTATTGATTCAGT  
TATTCCTTTCTGAGATACTAGTAGCAAAGACAGTCCGAACATGCATGAAAC  
GAAAAGGTGGCCGAAAGCTGGTGTGGATTGCCCTAGCGGTGAAACAGGA  
GCACGCGCTCTTATTCCTCTTCCCTTGCCACGCTACTGCAGCTAGGCAGCTC  
CTGCTGTTTCAAGCAGCGCAGTTGACAGGGGCGAGGTTGCATCGTCCGGGC  
AGGCTGTTGATGATTTCTGGA

**ADD2-2 (N22)**

CTGGTAACTTCATCCTAAAGTGCTATTGTGACAGATAACCTAATCGAATGTC  
CCAGATTGGTATCGAGAGTTGGCCTTATGCAATGCTTTGCAATTTTTTTAAA  
CCAGTGATTTTCGTAGCTGCCACATTTCAAAGAGAACATGAAATAAGGAGTA  
ACTGGGTGTAACATATTTTTTTCATTA AAAAGATGTTTGTCTGGAGTCAAAA  
AAAAAAAAAAAAAAAA

**ADD5-2 (N22)**

ATCTATCCCACTCACTCATGTAAGAGGAGTGCAAAGAGGCTATTACAGGA  
CAATCCTCATGTAGAGGATTTGAACA ACTGAATGAGCATTCTGAAGGTAGA  
GTTTGCCTGGCTCTTCTTGCTTGCTTGTTGGGCCATCAGATTGTTGGATCTTCC  
TATCTTCATGAAGCTGCAAGA ACTAAAAGAACTTTTGTGATGATTTCTGGA

**ADD6-2 (N22)**

TTTTTTTTTTTTTTTTGATAAACGTCNCCAGCTTTATTGGATCAGAAGCAACA  
TACAAGCTGGAGATTAATCTATCCCACTCACTCATGTAAGANGAGTGCAA  
AGAGGCTATTACAGGACAATCCTCATGTAGAGGATTTGAACA ACTGAATGA  
GCATTCTGAAGGTAGAGTTTGCCTGGCTCTTCTTGCTTGCTTGNGGGCCATC  
AGATTGTTGGATCTTCTATCTTCATGAAGCTGCAAGA ACTAAAAGAACTTT  
TGTGATGATTTCTGGA

**ADD7-1 (N22)**

CCNNAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
NGTGCTGCTAGTTACCTAGCTTCTGATCTCTGATGTACGTGTGCCGCCGTGA  
TGCGTGTGTTGTGTTGNGTCGACTGCTGCTACACTACTGTTCTGTGGCTTTCA  
CCTACTTATTCGGACAATTAATATANTAATCGTTTTGNTGCTCAAAAAAAAAA  
AAAAAAAA

**ADD7-2 (N22)**

CCAGAAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
AGTGCTGCTAGTTACCTAGCTTCTGATCTCTGATGTACGTGTGCCGCCGTGA  
TGCGTGTGTTGTGTTGTGTCGACTGCTGCTACACTACTGTTCTGTGGCTTTCA  
CCTACTTATTCGGACAATTAATATAGTAATCGTTTTGCTGAAAAAAAAAAAA  
AAAAAA

**ADD8-2 (N22)**

CCAGAAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
ACATTACATATAGTAGATCGATGTTGATGATTTCTGATGTACGTGTGCCGCC  
GTGATGCGTGTGTTGTGTTGTGTCGACTGCTGCTACACTACTGTTCTGTGGCT  
TTCACCTACTTATTCGGACAATTAATATAGTAACCGTTTTGCTGAAAAAAAAA  
AAAAAAAA

**ADD9-1 (N22)**

CCAGAAATCATCAACAGCCTGCCCGGACGATGYAACCTCGCCCCTGTCAAC  
TGCGCTGCTTGAAGCAGCGGGAGCTGGCTAGCTGCACTGGCACGGCAGGAC  
GGGGAATAAGACGCGCGTGTTCCTGTTTCACCGCGAGGGCAATCCAGCGCC  
AGCTTTCGGCCTGACCGCCGGCACCTTTTCGTTTCTTGTTTCGGGCTGTCTTTG  
CTACTAGTAGTATCTCAGGAAAAGAATAAATGAAACAATAATGCATGATCA  
TCAGATGCTTCTGCTGTGCAAAAAAAAAAAAA

**ADD10-1 (N22)**

CCAGAAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
 AGTGCTGCTAGTTACCTAGCTTCTGATCTCTGATGTACGTGTGCCGCCGTGA  
 TGCGTGTGTTGTGTTGTGTCTGACTGCTGCTACACTACTGTTCTGTGGCTTTCA  
 CCTACTTATTCGGACAATTAATATAGTAATCGTTTTGCTGCAAAAAAAAAAA  
 AA

**ADD12-2-2 (N22)**

TTTTTTTTTTTTTTGTCCGAATAAGTAGGTGAAAGCCACAGAACAGTAGTGNA  
 GCAGCAGTCGACACAACACAACACACGCATCACGGCGGCACACGTACATCA  
 GANATCAGAAGCTAGGTAAGTACTAGCAGCACTGCATGCAGAGCACGCACATTA  
 CATATAGTAGATCGATGTTGATGATTTCTGGA

**Lista de sectores homólogos a elementos repetitivos de cada una de las secuencias clonadas a partir de N22 utilizados para la construcción del árbol filogenético.**

**>N22**

CAGCGAACTACACNGGTTGCACTAAACAAAGAGGAGCAAAAACCTCTACTAA  
 CCAACAGAAAGAGGGGCACTGGACACAACCTAACGGAGGACCACACGTCTT  
 CACTGCTGGAGGGAGCTGGTAACTTCATCCGGGGATGTCTCCCTAGAGCTG  
 CTGTTGATGATTTCTGGCGACGGCTGATTCGAGGCGGCTGAACCCTCTTTTC  
 CTTCACTGGGTGAGGCAGGGCTTGTGGCTCTGAGCTAGATACTGAGAGGG  
 GTAATGCTGCTTGCCTTAGCAGCTCTTCCAAGGCTAGTTCGCTGA

**>ER-A43 (N22)**

GCAGGTCTCCCTGTTTTCGAGGTCTGTATAAACGCATGCCAGAGGAATGGAT  
 CTTTGTGGCAGCGAAGACATTCAACGATCCTGATGTTACCGGTGAAGTTCC  
 ACTGCAAACAGAAGGAGCGTTGATTGGCCTTGCTGCGAACAAGAATCTGTT  
 CCGCTTGATGGTTTTTGCATGACCTCAAACGAGCGGGAGCTGTGCAATCTC  
 TAGAGATTTCCGGCGCCCGCTGGTGTACCACTATATGAGAGATCTCCCAAC  
 GCGGGATGTGCATATCGTGAATACTCTATAGTGACACCAATATATCGTGGC  
 GTTCTCGTGCATATCTCTGTTTCTGTGAGAAAGTGATATGCTCTCACATTCC  
 CACAAAACACACGACGCAAAAACATAGAGTGAAGACCGTGGTGTGTACTG  
 AGAGAGAGCTAACTCACAAATAGTGTTTTGCGCACTGTGCGCTTTCCCATG  
 GGGAAAACGTGGCGTGCCATGTGCTATTGTGACTCCCAAAACGCGCGGAGA  
 GACGCTGTGTGCATAGTGCGCGCTTCCCCGCCTCCACTCACAGAGACACTGC  
 GCTCTCGGGCTTTCGTGTGCTGAGAGCTATCTCATCTCACTCAGGGGGGGAA  
 ATACGTTTCTCCACAGACACGGGATATAACATGAAAAAGCATATGAGCACA  
 AAAGGCCCCCAAAGTGTCCGAATCCTATAAAGCCCCGCTTTTCGGGCTTTTT  
 CCCAAGGCCCCCTCAGAAGAGACTCATAAATTGCCGCTCTCATCCCAA  
 ATGTGGAAAATTCCCAGGACCCTATAGATACTCCCGGGGTTTCCCCTGAAA  
 TCTCTCGCGCGCTCTTCTCCTTC

**ER-A44 (N22)**

ATTACTATTGAAAAACCTGCCGCCTGCACGTGACCATATGGCAGAGCTCCC  
 ACCGCGATAGACGCATAGCTTGAATAGTGATATTGCCCTCCACACACATG  
 ACGTACTCCGGTCAGACGTGTTTCTGTGCGAAAGTGTTATCCTCTCTCAA  
 GCGCTGAAACAGCGATCCCCAAAACATAGGGTAAAGCCCGGGGTGACTC  
 ATGAGAGAAATAACCTCAAATAATCGCGTTGCGCACAACAGCCCGCTTTCC

TGGCGGGTTACCCTAGGATGCCCCCGCTAAAAAACAGCAAACCAAC  
GCTCAAATCAGATTTAGGTAATGCAAGCCCCCTCTATTTACATCCCCCGGA  
TCCCTCTGTTAGGTACCCTCCGGTGGCTGTCAGTGTTGAATCCTC

**ER-A45 (N22)**

TGCGCATTTCGACACCGACACCTCAAGAGCTCGGCGTCGTCATAGCCATATCC  
GGTTCCAGAGTGTAATCGGCCCCCGAATACTATAACCGACCGCGCCGGAAC  
CCTAAGCCTAAAGTGTGAAATCCGCTGTGAATAATCTGACACCCAACTGAC  
GCCAGTGCAGAGCTCCCTACCCGCTTTCCAAACCGGACGAGTCGTCTGC  
CTTAGTGCACCAATGAATCCGCCAACACACGGGGATAGGCTGTTTGCCGAT  
CGAAAGCTTTCCCGCTTCCCTCACTTCCCTGCCACGGTGCACCTCGGTCTACTA  
ACGCTAACAGCTCTCTCAGCTCACTCAATGCGCGGATTCACCTTATCCGCGG  
ATAACGCCGCATAACACATGAAAGCACATGACCACCTAAATGCCCGCCAAC  
CGTTCAGAAGCCGTATAACCGCCCCCTTTTCCCTTGAGTTTTCCCTCCCGTCCC  
CCC

**ER-A46 (N22)**

CGGTAAATCTGACGTGCCAACTGCATTATCGAACCCGCCAAGGCGCGCGGA  
GAGAGCGGATAGCGTATTGAGTGCTCATCCGCTTCCCTCGCATCATGACTCGG  
TGCGCTCGGGTTCGTTTTGTTGCGGCGAGTGGTATCCCCTCACTCCAAGGCGG  
TAATCCGTTACCTTCAAATCAGGGGATAACGCCAGGAATAACATGGTAG  
CATAAGGCCCCCAAAGATCTGCAACGTTACAAACGCCGCCTTGACGGCGT  
CTTCCATAAGACCGCCCCCTACGAGCTTCAAAAATCGGCGCTCTAATC  
AAAATTGGGTAAACCGCACCTGAATTTTATATATCCCCGGCGTTTTCCCTG  
TAAACTCCC

**ER-A47 (N22) (inversa complementaria)**

TCCTGCGCACCCCAATCAGATGGAAGGTGTCCTCAGGTCAGCTACTGCTG  
AGAAAAAAGTGGCTTTCCCGAGCTGTAAAGAGATACCATGCGGCACCTCAG  
GATCCCAGGATACCCGGAGTACTGCGGGTTAGCGTTCCCCGAAAGGGAAGA  
ACGAGGATGGACCGTCACTTTGTATATCAGGGAGAACCTCAATCTCAACGG  
TTGGTACTGCAGAGCGGTGCGCATTGGATTCTTAGAAGCCTGTCACATTGTC  
GCTCGTGGTGCAGTAAGAAGTTATGCTCGACCTACCATGGGGCAGTACAA  
AATAGCCCTATGAGATTCTTTCCACCCGTGAACCAATGTGCCACAACCTGGC  
TTAAGAGGACTGCAATCTTGATAGAGATCAAAAAGTCCGGAGATCCCACCA  
TGGCCTACCTGGTCAAACATCTGCTCGCCTTGGATAACGAGTATGAAAAGCT  
GAGCCACCTGCATGACAAGTTGGTAGCCCGACACCAAGACACAGAAAA

**ER-ADD12-1 (N22) (inversa complementaria)**

ATTTAAATGAGTTTAAATACTACTACTAGTTAGTTATCTTCTATGAGTATTAAC  
TTTGATGAAGAAATTAGCCCTCTCTTGATTTTGCCTTCTTTCCCTAAAAGTTG  
GAATGGCTTGGTAACGGTTGTGAGTAATTCTATTTTCTGATCTAATACCCTA  
AAATATGATGATGTGATTGGTTTTCTTTTGGAGTGAGAAAACCTGTAGGAAAA  
CCTCAAGTAGATCTATATCAAGAAGTGCTTTAAACACACAAAGAAGTGGCA  
GAATGATTGAGAGACAAAATAATCCAAAAATTCTAGGAAATCAAGAG

**ER-ADD13-2 (N22)**

AGATTGGTGAAGTACATGAAGTATTCAAAGATCTTGCTACGCTTGTTTCATG

**ER-ADD16-1 (N22)**

GGCAAAGACATCAAAGGGATATTATTGCACTGCACAGCAGGAAGTTGCATA  
ATAAGCATCTGATGATCATGCATTATTG

**ER-ADD17-2 (N22) (inversa complementaria)**

GTTTCATGCATGTTTCGGACTGTCTTTGCTACTAGTATCTCAGGAAAAGAATA  
ACTGAATCAATAATGCATGATCATCAGATGCTTATTATGCGAAAAAAAAAAAA  
AA

**ER-ADD2-2 (N22)**

CGAGAGTTGGCCTTATGCAATGCTTTGCAATTTTTTTTAAACCAGTGATTTTCG  
TAGCTGCCACATTTCAAAGAGAACATGAAATAAGGAGTAACTGGGTGTAAC  
ATATTTTTTTCATTA AAAAGATGTTTGTCTGGAGTCAAAAAAAAAAAAAAAAA  
AAA

**ER-ADD5-2 (N22)**

ATCTATCCACACTCACTCATGTAAGAGGAGTGCAAAGAGGCTATTACAGGA  
CAATCCTCATGTAGAGGATTTGAACAACCTGAATGAGCATTCTGAAGGTAGA

**ER-ADD6-2 (N22)**

TCCTCATGTAGAGGATTTGAACAACCTGAATGAGCATTCTGAAGGTAGAGTTT  
GCCTGGCTCTTCTTGCTTGCTTGNGGGCCATCAGATTGTTGGATCTTCCTATC  
TTCATGAAGCTGCAAGAACTAAAAGAACTTTTGTTGATGATTTCTGGA

**ER-ADD7-1 (N22)**

CCNNAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
NGTGCTGCTAGTTACCTAGCTTCTGATCTCTGATGTACGTGTGCCGCCGTGA  
TGCGTGTGTTGTGTTGNGTCGACTGCTGCTACACTACTG

**ER-ADD7-2 (N22)**

CCAGAAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
AGTGCTGCTAGTTACCTAGCTTCTGATCTCTGATGTACGTGTGCCGCCGTGA  
TGCGTGTGTTGTGTTGTGTCGACTGCTGCTACACTACTGTTCTGTGGCTTTC

**ER-ADD8-2 (N22)**

CCAGAAATCATCAACATCGATCTACTATATGTAATGTGCGTGCTCTGCATGC  
ACATTACATATAGTAGATCGATGTTGATGATTTCTGATGTACGTGTGCCGCC  
GTGATGCGTGTGTTGTGTTGTGTCGACTGCTGCTACACTACTGTTCTGTGGCT  
TTCACCTACTTAT

**ER-ADD9-1 (N22)**

TTCGGCCTGACCGCCGGCACCTTTTCGTTTCTTGTTTCGGGCTGTCTTTGCTAC  
TAGTAGTATCTCAGGAAAAGAATAAATGAAACAATAATGCATGATCATCAG  
ATGCTTCCTGCTGTGCAAAAAAAAAAAAAAAAA

**ER-ADD10-1 (N22) (inversa complementaria)**

ATATTAATTGTCCGAATAAGTAGGGTAAAAGCCACAGAACAGTAGTGTAGCA  
GCAGTCGACACAACAACACACGCATCACGGCGGCACACGTACATCAGAG

ATCAGAAGCTAGGTAAGTACTAGCAGCACTGCATGCAGAGCACGCACATTACAT  
ATAGTAGATCGATGTTGATGATTTCTGG

**ER-ADD12-2-2 (N22)**

GTCCGAATAAGTAGGTGAAAGCCACAGAACAGTAGTGNAGCAGCAGTCGA  
CACAACACAACACACGCATCACGGCGGCACACGTACATCAGANATCAGAA  
GCTAGGTAAGTACTAGCAGCACTGCATGCAGAGCACGCACATTACATATAGTAG  
ATCGATGTTGATGATTTCTGGA

**Secuencias obtenidas a partir de los experimentos 5'RACE de los ARNm blanco  
predichos de N11 y N13**

**A25-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACTTCCACCCGGTTGATGAGGACTCTG

**A26-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACTTCCACCCGGTTGATGAGGACTCTG

**A27-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACTTCCACCCGGTTGATGAGGACTCTG

**A28-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACTTCCACCCGGTTGATGAGGACTCTG

**A31-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACTTCCACCCGGTTGATGAGGACTCTG

**A30-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACC

**A32-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACC

**A33-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACC

**A34-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACC

**A35-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACC

**A37-N11**

ACTTGACCGCCTCCTCATCTCGCTCAGCAAACCTCTACGCAGTCTATCATTGA  
GCTGATCTCCACC

**Tar N13-9**

CCATCCTAATACGACTCACTATAGGGCCTGGAGGAAGGATCAATAGGTGCC  
TGGAGGAAGGGTGTGGCTGTGAACCTGATCA

**Tar N13-5**

CCATCCTAATACGACTCACTATAGGGCCGGAGGAAGATCAATAGGTGCCTG  
GAGGAAGAATGTGGCTGTGAACCTGATCA

**Tar N13-3**

CCATCCTAATACGACTCACTATAGGGCCGGAGGAAGATCAATAGGTGCCTG  
GAGGAAGAATGTGGCTGTGAACCTGATCA

## ANEXO II - Sección A: Tabla resumen de los resultados obtenidos para los 22 transcritos diferenciales por modo de reproducción

Nombre <sup>(1)</sup>	Ori-gen <sup>(2)</sup>	Amplif. en ADN genómico	Extensión por RACE <sup>(3)</sup>	ID principal	Anotación principal	eV	ID de sector génico	Anotación del gen	eV	Conclusiones
N2 (353nt)	A	(-)	(-)	-	-	-	-	-	-	-
N7 (494nt)	S	(+)	(+) 540nt	<a href="#">FJ014716.1</a>	Proteína hipotética de <i>Glicine max</i> con dominios protein kinasa C. Presencia de péptido señal.	1,35	-	-	-	Posible secuencia proteica con dominios kinasa
N11 (758nt)	A	(+)	(+) 758nt	<a href="#">Ath</a> <a href="#">MIR165b</a>	Precursor de miARN de <i>Arabidopsis thaliana</i>	Sc: 103	<a href="#">Os03g0199100</a>	Proteína hipotética no caracterizada de <i>Oryza sativa</i> (26nt)	0,065	Posible precursor de miARN regulatorio con blanco homólogo a <a href="#">Os03g0199100</a>
N13 (532nt)	A	(+)	(+) 718nt	<a href="#">Osa</a> <a href="#">MIR1877</a>	Precursor de miARN de <i>Oryza sativa</i>	Sc: 135	<a href="#">gb BT068773.1 </a>	Proteína hipotética de maíz con dominio ETR (22nt)	0,12	Posible precursor de miARN regulatorio con blanco homólogo a <a href="#">gb BT068773.1 </a>
N17 (476nt)	A	(+) Generación de marcador SCAR	(+) 907nt	<a href="#">ORSiCMC</a> <a href="#">M00100010</a>	Región LTR de un retrotransposón Aboov de posición centromérica de <i>Oryza sativa</i>	9e <sup>-06</sup>	<a href="#">Os11g0269700</a>	Proteína hipotética no caracterizada de <i>Oryza sativa</i> (62nt)	2e <sup>-11</sup>	Posible retrotransposón centromérico portador de gen transduplicado homólogo a <a href="#">Os11g0269700</a>

N22 (303nt)	A	(+) (+)	(+)	<a href="#">ORSiTERT 00200152</a>	Sector de protein Gag del retrotransposón Ty3 <i>Gypsy</i>	0,97	-	-	-	Posibles miembros de una familia de retrotransposones, algunos de los cuales portan segmentos de genes codificantes de plantas previamente asociados a la apomixis
			A43:849nt	<a href="#">gb EU5585 21.1 </a>	Retrotransposón <i>Gypsy</i> 20	$3e^{-65}$	<a href="#">emb AJ8416 97.1 </a>	Gen <i>serk2</i> de <i>Poa pratensis</i>	$3e^{-28}$	
			A44:827nt	<a href="#">gb EU5585 21.1 </a>	Retrotransposón <i>Gypsy</i> 20	$8e^{-35}$	<a href="#">gb EU96706 0.1 </a>	Gen Citocromo p450 de maíz	$6e^{-100}$	
			A45:784nt	<a href="#">gb EU5585 21.1 </a>	Retrotransposón <i>Gypsy</i> 20	$1e^{-16}$	-	-	-	
			A46:643nt	<a href="#">gb EU5585 21.1 </a>	Retrotransposón <i>Gypsy</i> 20	$2e^{-60}$	<a href="#">emb AJ8416 97.1 </a>	Gen <i>serk2</i> de <i>Poa pratensis</i>	$2e^{-54}$	
			A47:802nt	<a href="#">ZRSiTERT OOT00146</a>	Retrotransposón dagaf_576C20-1	$3e^{-8}$	-	-	-	
N38 (1211nt)	S	(-) (-)	(-)	-	-	-	-	-	-	-
N43 (783nt)	A	(+) (+)	(+)	<a href="#">gb EU5585 21.1 </a>	transposones de <i>A.</i> <i>thaliana</i> : CACTA2, SINE9, MITE7, <i>Gypsy</i> 20	$2e^{-11}$	-	-	-	Posibles miembros de una familia de elementos repetitivos, algunos de los cuales portan segmentos de genes codificantes de plantas
			A11:629nt	<a href="#">gb EU5585 21.1 </a>	transposones de <i>A.</i> <i>thaliana</i> : CACTA2, SINE9, MITE7, <i>Gypsy</i> 20	$5e^{-113}$	<a href="#">gb BT03538 6.1 </a>	Proteína de maíz con dominios kinasa	$2e^{-8}$	
			A12:676nt	<a href="#">gb EU5585 21.1 </a>	transposones de <i>A.</i> <i>thaliana</i> :	0,05 6	<a href="#">ref XM_0024 41009.1 </a>	Proteína hipotética de	$6e^{-156}$	

				CACTA2, SINE9, MITE7, <i>Gypsy</i> 20			sorgo con dominio asparaginasa			
				<b>A13:849nt</b>	<a href="#">TRSiTERT 00100121</a>	Retrotransposón angela_SK14_4	0,63	-	-	-
				<b>A15:775nt</b>	<a href="#">HRSiTERT OOT00038</a>	región LTR del retrotransposón Sukkula de <i>Hordeum vulgare</i>	0,47	-	-	-
				<b>A16:169nt</b>	<a href="#">ref XM_00 2488912.1 </a>	Proteína hipotética no caracterizada de sorgo	5e <sup>-55</sup>	-	-	-
<b>N49</b> (264nt)	A	(-)	(-)	-	-	-	-	-	-	Posible contaminación
<b>N56</b> (464nt)	A	(+)	(+) interno	<a href="#">ath- miR861-3p</a>	Secuencia de miARN maduro de <i>A. thaliana</i>	Sc: 82	-	-	-	Posible precursor de miARN similar a <a href="#">ath-miR861-3p</a>
<b>N58</b> (208nt)	A	(+)	(-)	<a href="#">ZRSiTERT OOT00251</a>	Retrotransposón opie_092E12-1	0,5	-	-	-	Posible retrotransposón de gramíneas
<b>N70</b> (840nt)	S	(+)	(-)	<a href="#">ORSiTERT 00200156</a>	Retrotransposón Ty3- <i>Gypsy</i>	0,7	-	-	-	Posible retrotransposón de gramíneas
<b>N77</b> (276nt)	S	(-)	(-)	-	-	-	-	-	-	Posible contaminación

<b>N82</b> (552nt)	A	-	-	<a href="#">ZRSiTERT OOT00248</a>	Retrotransposón milt_573L14-1	6,3e <sup>-5</sup>	<a href="#">gb EU94086 5.1 </a>	ARNm de <i>Zea mays</i>	5e <sup>-4</sup>	Posible retrotransposón con ARNm transduplicado similar a <a href="#">gb EU940865.1 </a>
<b>N86</b> (226nt)	A	(-)	(-)	-	-	-	-	-	-	Posible contaminación
<b>N87</b> (970nt)	S	-	-	<a href="#">ath- MIR169k</a>	Precursor de miARN	0,41	-	-	-	Posible precursor de miARN similar a ath-MIR169k
<b>N99</b> (740nt)	S	-	-	<a href="#">emb AJ309 824.2 </a>	Retrotransposón de <i>Zea mays</i>	8e <sup>-3</sup>	<a href="#">Os02g06819 00</a>	Factor de transcripción GRAS	0,3	Posible retrotransposón con ARNm transduplicado similar a <a href="#">Os02g0681900</a>
<b>N103</b> (257nt)	A	(-)	(-)	-	-	-	-	-	-	Posible contaminación
<b>N104</b> (719nt)	A	(+)	(+) 786nt	<a href="#">XM_00244 5379.1</a>	Proteína no caracterizada de <i>Sorghum bicolor</i>	6e <sup>-12</sup>	-	-	-	Secuencia proteica similar a <a href="#">XM_002445379.1</a>
<b>N106</b> (461nt)	S	-	-	<a href="#">YP_235263 .1</a>	Proteína bacteriana	1e <sup>-43</sup>	-	-	-	Posible contaminación
<b>N116</b> (343nt)	S	-	-	<a href="#">gb EU9755 38.1</a>	Proteína hipotética de Maíz	3e <sup>-10</sup>	-	-	-	Posible proteína similar a <a href="#">gb EU975538.1</a>

---

<b>N119</b> (223nt)	S	-	-	<a href="#">gi3867818</a> <a href="#">21</a>	Transposasa de Rim2-M47	0,01	-	-	-	Posible transposón
------------------------	---	---	---	---	----------------------------	------	---	---	---	--------------------

---

<sup>(1)</sup> Se detalla la nomenclatura de cada transcripto aislado por Display Diferencial seguido de su longitud original en nucleótidos.

<sup>(2)</sup> La columna refiere al genotipo a partir del cual fue aislada la banda. S: Genotipo sexual, Q4188. A: Genotipo apomítico, Q4117.

<sup>(3)</sup> Se detallan los transcriptos que fueron amplificados por 3' y 5'RACE. Los que pudieron ser extendidos cuentan con un signo (+) seguido de la longitud de la secuencia consenso en nucleótidos.

## ANEXO II - Sección B: Resumen de los resultados obtenidos para los 36 transcritos diferenciales por nivel de ploidía

Nombre (1)	Ori- gen (2)	ID principal	Anotación principal	eV	ID de sector génico	Anotación del gen	eV	Conclusiones
<b>L14</b> (564nt)	2x	<a href="#">osa-MIR419</a>	Precursor de miARN	Sc= 120	<a href="#">AT2G42980</a>	Aspartil proteasa	0,8	Posible pre-miARN Predicción <i>in silico</i> positiva
<b>L20</b> (426nt)	4x	<a href="#">gb EU95744 4.1 </a>	Proteína hipotética de maíz Dominio Ser/Thr kinasa. Superfam. PKC (cl09925)	8e <sup>-20</sup>	-	-	-	Posible proteína con dominio kinasa
<b>L24</b> (361nt)	2x	<a href="#">gb EU96646 7.1 </a>	Proteína hipotética de maíz	3e <sup>-55</sup>	-	-	-	Posible proteína no caracterizada
<b>L29</b> (335nt)	2x	<a href="#">osa- MIR166b</a>	Precursor de miARN	Sc= 124	<a href="#">emb Z17313.1  MDSTPKN</a>	Ser/Thr quinasa	1,6	Posible pre-miARN
<b>L45</b> (340nt)	2x	<a href="#">osa-MIR812i</a>	Precursor de miARN	Sc= 125	-	-	-	Posible pre-miARN
<b>L47</b> (526nt)	2x	<a href="#">dbj AK22458 1 </a>	cDNA <i>Oryza officinalis</i> ARNm expresado en estadios tempranos en panoja	4e <sup>-10</sup>	-	-	-	Posible proteína no caracterizada
<b>L48</b> (234nt)	2x	<a href="#">ZRSiTERTO OT00206</a>	Retrotransposón ji_178A11-1	0,18	-	-	-	Posible Retrotransposón

<b>L49</b> (451nt)	4x	<a href="#">gb EU95133 3.1 </a>	mRNA de proteína de maíz	$1e^{-21}$	-	-	-	Posible proteína no caracterizada
<b>L53</b> (214nt)	2x	<a href="#">gb EU96472 2.1 </a>	Proteína hipotética de maíz	$1e^{-5}$	-	-	-	Posible proteína no caracterizada
<b>L55</b> (221nt)	2x	<a href="#">ORSiTERTO OT00437</a>	Retrotransposon rn_525-269	eV= 0,3	<a href="#">AT4G11580</a>	Proteína de la flia. F-box	1,1	Posible retrotransp. con ARNm transduplicado
<b>L61</b> (91nt)	2x	<a href="#">gi 16225181 </a>	IR miniatura de elemento transponible.	eV= 0,66	-	-	-	Posible repetición invertida
<b>L66</b> (288nt)	4x	<a href="#">ref NM_0011 39432.1 </a>	mRNA de proteína de maíz, con 4 dominios catalíticos de tipo PKC (cd00180)	$1e^{-7}$	-	-	-	Posible proteína con dominios PKC
<b>L72</b> (339nt)	4x	<a href="#">ath-MIR863</a>	Precursor de miARN	Sc= 160	<a href="#">Os06g033460 0</a>	ARNm	0,46	Posible pre-miARN
<b>L73</b> (484nt)	2x	<a href="#">osa- MIR809e</a>	Precursor de miARN	Sc= 100	<a href="#">gb AF351044. 1 AF351044</a>	Subunidad F de NADHdH	2,4	Posible pre-miARN
<b>L75</b> (417nt)	2x	<a href="#">Os03g08266 00</a>	Proteína de <i>Oryza sativa</i> tipo fosfoesterasa Superfam. Fosfoesterasas (cl10627)	$2e^{-26}$	-	-	-	Posible proteína con dominio fosfoesterasa
<b>L76</b> (259nt)	2x	<a href="#">gb EU97038 6.1 </a>	Proteína hipotética de maíz	$1e^{-18}$	-	-	-	Posible proteína no caracterizada
<b>L85</b> (282nt)	4x	<a href="#">TRSiTERTO OT00028</a>	Retrotransposón Josephine	0,42	-	-	-	Posible Retrotransposón

<b>L86</b> (240nt)	2x	-	-	-	-	-	-	-
<b>L87</b> (229nt)	2x	<a href="#">gb EU36540 1.1 </a>	NADHdH de <i>B. oldhamii</i> expresada en mitocondria	$1e^{-54}$	-	-	-	Posible proteína tipo NADHdH mitocondrial
<b>L90</b> (251nt)	4x	<a href="#">gi 15340829 </a>	Genes Gag-Pol de retrotransp. Ty3- <i>Gypsy</i>	0,085	-	-	-	Posible Retrotransposón
<b>L94</b> (328nt)	2x	<a href="#">gb BT01851 2.1 </a>	mRNA de proteína de maíz	$2e^{-24}$	-	-	-	Posible proteína no caracterizada
<b>L95</b> (294nt)	2x	<a href="#">gb BT03832 2.1 </a>	Proteína hipotética de maíz tipo Ubiquitina (cl00155)	$7e^{-49}$	-	-	-	Posible proteína tipo Ubiquitina
<b>L774</b> (91nt)	4x	<a href="#">LOC100192 571</a>	mRNA de proteína de maíz	$2e^{-10}$	-	-	-	Posible proteína no caracterizada
<b>L784</b> (97nt)	4x	-	-	-	-	-	-	-
<b>L794</b> (84nt)	2x	<a href="#">ZRSiOTOTO 0000019</a>	Repetición CAG y TAG de Zea mays	0,24	-	-	-	Posible repetición trinucleotídica
<b>L802</b> (158nt)	2x	<a href="#">gb EU97565 9.1 </a>	mRNA de ankirina kinasa de maíz con un dominio Ankirina K y uno metiltransferasa (PF03141)	$9e^{-6}$	-	-	-	Posible proteína tipo Ankirina K
<b>L812</b> (133nt)	2x	<a href="#">tae-MIR1138</a>	Precursor de miARN	Sc= 96	<a href="#">dbj AK110092 .1 </a>	ARNm	0,57	Posible pre-miARN

<b>L832</b> (111nt)	2x	<a href="#">gnl uv J0174 9.1</a>	vector pBR322 1-4361-49	$2e^{-51}$	-	-	-	Posible contaminación
<b>L834-1</b> (171nt)	4x	<a href="#">ptc-MIR476a</a>	Precursor de miARN	Sc= 122	<a href="#">7475047 CYP705B4P-1</a>	cytochrome P450	1,4	Posible pre-miARN
<b>L834-3</b> (174nt)	4x	<a href="#">osa- MIR319b</a>	Precursor de miARN	Sc= 111	<a href="#">827199 CYP705A1</a>	cytochrome P450	4	Posible pre-miARN
<b>L844</b> (818nt)	4x	<a href="#">gb EU96084 8.1</a>	mRNA de proteína de maíz Familia Josephin (PF02099)	$1e^{-15}$	-	-	-	Posible proteína clasificada en la familia Josephin
<b>L852</b> (84nt)	2x	<a href="#">ORSiTEMTO 1100018</a>	Repetición de Retrotransp. p-SINE	0,99	-	-	-	Posible transposón
<b>L852-1</b> (124nt)	2x	-	-	-	-	-	-	-
<b>L852-2</b> (154nt)	2x	-	-	-	-	-	-	-
<b>L854</b>	4x	<a href="#">Os02g02670 00</a>	Proteína de <i>Oryza sativa</i> Dominios conservados: DPBB_1(cl04011 y polen_alerg_1 (cl08320)	$2e^{-23}$	-	-	-	Posible proteína con dominios conservados
<b>L884</b> (106nt)	4x	<a href="#">gb EU55852 2.1</a>	Inserción de transposón tipo <i>Gypsy21</i>	$2e^{-30}$	-	-	-	Posible Transposón

<sup>(1)</sup> Se detalla la nomenclatura de cada transcripto aislado por Display Diferencial seguido de su longitud original en nucleótidos.

<sup>(2)</sup> La columna refiere al genotipo a partir del cual fue aislada la banda. 2x: Genotipo apomítico diploide; 4x: Genotipo apomítico tetraploide.