

UCH-FC
Biotecnología
P116
c.1



“APLICACIÓN DE MACROARRAYS DE cDNA EN ESTUDIOS DE EXPRESIÓN GÉNICA EN FRUTOS DE *Prunus persica*”

Seminario de Título entregado a la Universidad de Chile, en cumplimiento parcial de los requisitos para optar al Título de Ingeniero en Biotecnología Molecular.

IGOR ALEJANDRO PACHECO CRUZ

Dra. Verónica Cambiazo Ayala
Directora Seminario de Título

Comisión Revisora

Dr. Nicolás Guiliani
Presidente

Dr. Michael Handford
Corrector



Santiago, Viernes 15 de 2006



Nacido en Santiago el 7 de Agosto de 1980, hijo de la profesora de Francés Juana Cruz Segovia y el profesor de Educación Física Igor Pacheco Blanco, Igor Pacheco Cruz desde pequeño mostró curiosidad y sensibilidad por algunos fenómenos naturales, haciendo más de alguna travesura en su afán de responder “¿qué pasa si...?”. En su adolescencia despertó una de sus grandes pasiones: la música, que lo sigue siendo hasta ahora en su desempeño como baterista autodidacta. Descubre su otra pasión, las ciencias biológicas, tardíamente a sus 17 años, cuando su formación de matemático en el Instituto Nacional ya casi terminaba. A los 18 años comienza uno de los períodos más ricos de

su vida: entra a la Facultad de Ciencias de la Universidad de Chile, a la carrera de Ingeniería en Biotecnología Molecular, disciplina que le vino como anillo al dedo a sus aspiraciones académicas. Sin duda el período de mayor aprendizaje profesional se concretó entre los años 2003 y 2004, cuando efectuó su tesis de pregrado en el Laboratorio de Bioinformática y Expresión Génica en el INTA, lugar que le abrió las puertas al sector profesional, conduciéndolo a Biosigma S.A., donde hasta el día de hoy se desempeña en el cargo de profesional investigador.



*A mis a la vez padres y maestros,
Juany e Igor,
por ponerme en los caminos correctos...
Y a Juana Segovia,
cuyo cuidado sigo recibiendo.*

AGRADECIMIENTOS

Tengo el ambicioso deseo de agradecer en una página a quienes que sin su apoyo de cualquier índole, esta tesis no hubiera dado frutos:

A Verónica, por su dedicación en entregar y hacer que se crearan conocimientos; por mostrarme que la disciplina y el amor por la ciencia siempre resultan en experiencias que van más allá de un dato o una publicación, y que trascienden en tu vida profesional; y sobre todo por su tiempo y paciencia.

A Mauricio González, por su confianza en mí, por su voz eterna de aliento, tácita en todas nuestras conversaciones, y sobre todo por sus consejos, tanto científicos y de vida.

A Mauricio González-Agüero, de quien recibí sus enseñanzas en cada uno de los días de efectuado este trabajo, y de quien aprendí la rigurosidad en esta profesión; y por su amistad y su apoyo técnico y emocional en momentos difíciles.

A todos mis compañeros del LBEG 2003-2004, a Chodar, Jano, Pulgar, Felipe, Talía, Dr. Marco Méndez, Myriam, Freddy, Angélica, Morgan, ya que el tener su amistad me hace sentir afortunado hasta el día de hoy. Muchas gracias a Leodio por su ayuda en el trabajo y por su generosa camaradería. Gracias también a Pablo su ayuda en la parte bioinformática de este trabajo.

A todo el equipo de Biosigma S.A., sobre todo a las doctoras Pilar Parada y Nicole Ehrenfeld, ya que la experiencia que he creado en la empresa gracias a su apoyo, sin duda ha contribuido en la maduración de esta tesis.

A Morphonica, Dody y todos sus integrantes, por ser mis cables a tierra en este período. Sin arte no hay ciencia y sin ciencia no hay arte.



ÍNDICE DE CONTENIDOS



AGRADECIMIENTOS.....	iii
ÍNDICE DE CONTENIDOS.....	iv
ÍNDICE DE TABLAS.....	ix
ÍNDICE DE FIGURAS.....	x
LISTA DE ABREVIATURAS.....	xi
RESUMEN.....	xiv
ABSTRACT.....	xv
INTRODUCCIÓN.....	1
1.- Herramientas de la Genómica Funcional.....	2
1.1.- Técnicas basadas en PCR y fraccionamiento en geles: Muestreo diferencial y cDNA-AFLP.....	2
1.2.- Métodos basados en el recuento de <i>tags</i> de cDNA: Análisis en Serie de la Expresión Génica (SAGE).....	3
1.3.- Técnicas basadas en la hibridación de sondas con genes blanco: arreglos de cDNA.....	5
1.3.1.- Principios técnicos de los arreglos de DNA más usados.....	5
1.3.2.- Algunos ejemplos de la aplicación de arreglos de DNA en análisis de expresión génica a gran escala.....	8
2.- El problema de la harinosidad en duraznos (<i>Prunus persica</i>).....	11

3.- Hipótesis.....	13
3.- Objetivo General.....	13
4.- Objetivos Específicos.....	14
MATERIALES Y MÉTODOS.....	15
1.- Obtención de los DNAs blanco.....	15
1.1.- Extracción de DNA plasmidial.....	15
1.2.- Amplificación y purificación de insertos a partir de DNA plasmidial.....	15
1.3.- Selección de clones a partir de una genoteca de ESTs de durazno de la Universidad de Clemson.....	16
1.4.- Amplificación de insertos a partir de cultivos bacterianos.....	19
1.5.- Verificación y cuantificación de los productos de PCR.....	19
1.6.- Obtención de genes controles.....	20
2.- Generación de <i>macroarrays</i>.....	21
2.1.- Siembra de los DNA blanco en las membranas.....	21
2.2.- Tratamiento de las membranas sembradas.....	23
2.3.- Prehibridación, hibridación y lavado de las membranas.....	23
3. Obtención de sondas.....	24
3.1.- Muestras biológicas.....	23
3.2.- Extracción de RNA total de duraznos.....	24

3.3.- Extracción de mRNA poli-(A) desde RNA total.....	25
3.4.- Transcripción <i>in vitro</i> del gen <i>dapB</i>	27
3.5.- Marcación de sondas.....	27
4.- Obtención de Datos.....	29
4.1.- Obtención de imágenes.....	29
4.2.- Cuantificación de las imágenes.....	29
4.3.- Tratamiento y selección de datos.....	30
4.4.- Análisis de dominios proteicos en secuencias de función desconocida.....	32
RESULTADOS.....	33
1.- Optimización de parámetros de siembra.....	33
1.1.- Cuantificación de las señales.....	33
1.1.1.- Obtención de datos.....	33
1.1.2. Medición de la intensidad cruda.....	34
1.1.3. Medición de la intensidad del ruido y cálculo de la intensidad neta.....	36
1.2.- Determinación de los parámetros de siembra.....	37
1.3.- Determinación del tipo de DNA blanco a sembrar.....	41
1.4.- Reproducibilidad en la medición de la señal.....	43

1.5.- Sensibilidad.....	45
1.6.- Conclusiones.....	47
2.- Estudio de la expresión de genes relacionados con el fenómeno de harinosidad en <i>Prunus persica</i>.....	48
2.1.- Selección de genes.....	49
2.2.- Siembra de los clones seleccionados en membranas de nylon.....	50
2.3.- Obtención de RNA para experimentos de hibridación.....	52
2.4.- Ensayos de hibridación.....	54
2.5.- Adquisición y análisis de datos.....	56
2.5.1.- Comparación entre réplicas.....	57
2.5.2.- Selección de datos de buena calidad.....	60
2.5.3.- Normalización de los datos de intensidad.....	62
2.5.4.- Determinación de cambios significativos en la expresión de genes.....	64
DISCUSIÓN.....	71
1.- Puesta a punto del método de confección de membranas de <i>macroarrays</i>, y de la obtención y análisis de datos de hibridaciones.....	71
2.- Determinación de los cambios en la expresión de genes en duraznos harinosos.....	77
2.1.- Análisis de los genes diferencialmente expresados.....	79
2.1.1.- Genes reprimidos en el fruto harinoso.	79

2.1.1.1.- Biosíntesis, Estructura y Metabolismo de la Pared Celular.....	79
2.1.1.2.-Secreción de proteínas.....	80
2.1.1.3.- Transducción de señales mediadas por Ca^{+2} en respuesta al frío.....	81
2.1.1.4.- Protección al estrés oxidativo.....	83
2.1.1.5.- Genes asociados a hormonas vegetales.....	84
2.1.1.6.- Genes de respuesta a estrés térmico.....	85
2.1.2.- Genes inducidos en el fruto harinoso.....	86
CONCLUSIÓN.....	88
REFERENCIAS BIBLIOGRÁFICAS.....	89





ÍNDICE DE TABLAS

Tabla I.- Criterios de búsqueda de palabras claves o secuencias consenso utilizados para la selección manual de clones.....	18
Tabla II.- Genes de <i>Prunus persica</i> amplificados desde cDNA como control de calidad de los mRNA extraídos.....	26
Tabla III.- Concentración y calidad de los RNA totales obtenidos en dos extracciones independientes de RNA.....	53
Tabla IV.- Coeficientes de correlación de Pearson para las réplicas de hibridaciones.....	59
Tabla V. Filtros aplicados a los datos obtenidos y su impacto el número de datos recuperados.....	62
Tabla VI. Genes cuyo cambio de expresión fue significativo ($p < 0,05$) entre frutos NH y H.....	66

ÍNDICE DE FIGURAS

Figura 1. Obtención de datos desde membranas de <i>macroarrays</i>	35
Figura 2. Parámetros de siembra.....	38
Figura 3. Cantidad óptima sembrado y distancia entre <i>spots</i>	40
Figura 4. Variabilidad atribuible a la siembra repetida.....	42
Figura 5. Determinación del tipo DNA blanco.....	44
Figura 6. Reproducibilidad en los ensayos de hibridación.....	45
Figura 7. Sensibilidad en las hibridaciones en <i>macroarrays</i>	47
Figura 8. Proceso de selección de los genes que fueron sembrados en las membranas de <i>macroarrays</i>	50
Figura 9. Verificación de la calidad de los productos de PCR seleccionados.....	51
Figura 10. Verificación de la integridad de los RNAs extraídos.....	53
Figura 11. Verificación de la calidad de los mRNA.....	54
Figura 12. Hibridación en membranas de <i>macroarrays</i>	56
Figura 13. Comportamiento de las señales de hibridación.....	58



LISTA DE ABREVIATURAS

°C	grados Celsius
AFLP	Amplified Fragment Length Polymorphism
Blast	Basic Local Alignment Search Tool
Blast-n	Nucleotide-nucleotide Blast
cDNA	DNA complementario
CDS	secuencia codificante
cols	colaboradores
CV	Coefficiente de Variación
dATP	desoxiadenina
dCTP	desoxicitidina
dGTP	desoxiguanina
DMSO	Dimetilsulfóxido
DNA	ácido desoxiribonucleico
dNTPs	desoxinucleotidos trifosfatos
DTT	ditiotreitól
dTTP	desoxitimidina
dUTP	desoxiuracilo
EDTA	Ácido etilen-diamino-tetraacético
EST	expressed sequence tag
g	gramo
GO	Gene Ontology

H	Harinoso
ID	identificador
INIA	Instituto de Investigación Agropecuaria
INTA	Instituto de Nutrición y Tecnología de los Alimentos
IPTG	isopropiltiogalactósido
Kb	kilo base
L	litro
LB	medio de cultivo Luria broth
M	molaridad
MCS	Sitio de múltiple colnamiento
mRNA	RNA mensajero
M.W.	Estándar de peso molecular
N	No Harinoso
pb	pares de bases
PCR	reacción en cadena de la polimerasa
RNA	ácido ribonucleico
rpm	revoluciones por minuto
rRNA	RNA ribosómico
RT	transcripción reversa
s	segundos
SAGE	Serial Analysis of Gene Expression
SAR	Systemic Acquired Resistance
SD	Desviación estándar

SDS	dodecil sulfato de sodio
SQL	Structured Query Language
tBlastX	Nucleotide-nucleotide 6-frame translation Blast
TAE	Tampón Tris-Acetato-EDTA
TAIR	The Arabidopsis Institute Research
TE	Tampón Tris - EDTA
Tris	Tris-(hidroximetil)-aminometano
U	unidad internacional de enzima, μ moles de sustrato procesado por minuto
UAI	Unidades Arbitrarias de Intensidad
UV	ultra violeta

RESUMEN

La técnica de arreglos de DNA ofrece varias ventajas sobre otros métodos de estudio de la expresión génica, ya que permite cuantificar en forma relativamente rápida y reproducible la abundancia relativa de los transcritos expresados en un tejido sometido a una condición experimental determinada. Sin embargo, para obtener resultados confiables y reproducibles, es necesario estandarizar cada uno de los pasos en la ejecución de este tipo de experimentos, incluyendo el análisis de los datos generados en éstos. Para ello se han realizado una serie de experimentos que tuvieron por objetivo optimizar distintos parámetros de siembra de muestras de DNA en membranas de Nylon (*macroarrays*). Además se diseñó y aplicó un protocolo de tratamiento, limpieza y análisis de los datos provenientes de experimentos de *macroarrays*. Posteriormente, y con el fin de poner a prueba este método, se ejecutó un conjunto de experimentos de hibridación utilizando muestras complejas de mRNA en *macroarrays* que contenían una genoteca de *Prunus persica* (duraznos). En estos experimentos se compararon los perfiles de expresión de un grupo de 880 genes involucrados directa y/o indirectamente en el metabolismo de pared celular y en los mecanismos de respuesta a estrés por frío, en frutos no harinosos (NH) y harinosos (H) de duraznos de la variedad O'Henry, identificándose finalmente 3 genes significativamente inducidos y 34 significativamente reprimidos. Con ello se demostró que el método diseñado y optimizado presenta una alta sensibilidad y reproducibilidad, lo que le confiere una gran confiabilidad en su aplicación.

SUMMARY

The technique of DNA arrays offers several advantages with respect to other methods of study of gene expression, since it allows in a relatively fast and reproducible manner, the relative abundance of the expressed transcripts in a tissue or an experimental condition to be quantified. Nevertheless, to obtain reliable and reproducible results, it is necessary to standardize each one of the steps in the execution of this type of experiments, including the analysis of the data generated in them. For it, a series of experiment aimed at optimizing different parameters from complex DNA samples in Nylon membranes (*macroarrays*) have been made. Moreover, a protocol of treatment, cleaning and analysis of the generated data was designed and applied. Later, and with the purpose of testing this method, a set of hybridization experiments was executed using complex samples in nylon *macroarrays* spotted with genes from a library of *Prunus persica* (peach fruits). In these experiments the profiles of expression of a group of 880 genes involved directly and/or indirectly in cellular wall metabolism and mechanisms of cold stress were compared in sane (NH) and wooly (H) peach fruits of the O'Henry variety, finally identifying 3 overexpressed and 34 repressed genes. In this way it was demonstrated that the designed and optimized *macroarray* method displays high sensitivity and reproducibility, and confers a great confidence in its application.

INTRODUCCIÓN

Los esfuerzos invertidos durante los últimos 25 años en aumentar la eficiencia y el rendimiento de sistemas automatizados de secuenciación han resultado en la completa anotación de más de 250 genomas (Roth y cols., 2004). Entre estos genomas, se encuentran los de organismos que constituyen modelos de estudio, como el de la mosca del vinagre (*Drosophila melanogaster*), el nemátodo *Caenorhabditis elegans*, el ratón (*Mus musculus*), la planta *Arabidopsis thaliana* e incluso el humano (*Homo sapiens*), cuya secuenciación marcó el comienzo de la era post-genómica. Tales sistemas automatizados han permitido también secuenciar numerosas genotecas de cDNA.

Por otro lado, la secuenciación de genomas completos ha hecho evidente que un gran porcentaje de los genes no tiene función asignada, más aún, muchos de ellos no tienen similitud significativa con ningún gen conocido, y por lo tanto carecen de una función putativa. Es claro entonces que la función y las condiciones en las que se expresa un gen no quedan siempre explicadas por su secuencia o la de sus elementos reguladores. Nace así uno de los mayores desafíos de la comunidad científica en la era post-genómica: asignar a cada uno de los genes presentes en un genoma una función biológica específica. La tarea no es menor, ya que requiere, en primer lugar, dilucidar las redes de regulación transcripcional de una célula frente a un estímulo o una condición genética dada. De este modo, surge una nueva disciplina, la

Genómica Funcional, cuyo principal objetivo es caracterizar el transcriptoma (conjunto de mRNAs sintetizados en una célula, tejido u organismo, en una condición biológica particular) y su dinámica. Para ello, es necesario contar con métodos que nos permitan medir o cuantificar el nivel de expresión de múltiples genes y correlacionar cambios en su expresión con las variables de interés.

1.- Herramientas de la Genómica Funcional

En los últimos años, una importante cantidad de trabajos (revisado en Horak y Zinder, 2002; Bucca y cols. 2004), han demostrado que existen herramientas eficientes para examinar la expresión simultánea de numerosos genes. En general, estas técnicas pueden dividirse en tres grandes grupos, descritos a continuación.

1.1- Técnicas basadas en PCR y fraccionamiento en geles: Muestreo diferencial y cDNA-AFLP.

Este método fue utilizado por primera vez en 1992 por Pardee y Liang para aislar genes diferencialmente expresados en células tumorales. Los autores sintetizaron cDNA por transcripción reversa a partir de las muestras de mRNA en estudio, para luego amplificar el cDNA mediante PCR, en presencia de un nucleótido radiactivo o ligado, utilizando partidores que hibridan al azar con el cDNA. Este proceso genera fragmentos pequeños y representativos de un RNA, o *tags*, que posteriormente son fraccionados en geles de poliacrilamida. Después de un análisis comparativo de los patrones

electroforéticos de las muestras es posible visualizar bandas diferencialmente representadas, las cuales corresponden a transcritos de expresión diferencial entre las muestras comparadas (Liang y Pardee, 1992).

Si bien esta técnica ha servido para aislar genes diferencialmente expresados en distintos sistemas (Qin y cols., 2000), sufre de problemas tales como un alto nivel de ruido, baja reproducibilidad y una tasa de falsos positivos que asciende al 50%. Estos problemas han sido resueltos casi por completo usando enzimas de restricción que generan trozos específicos de cDNA, los cuales son amplificados por PCR y analizados mediante electroforesis. Estas innovaciones recientes del muestreo diferencial son llamadas comúnmente cDNA-AFLP (revisada en Breyne y Zabeau, 2001), y han sido aplicadas en procesos biológicos particulares, tales como, la respuesta de células de tabaco al ataque de patógenos fúngicos (Durrant y cols., 2000). Esta metodología no necesita de un conocimiento previo de las secuencias de los transcritos analizados, sin embargo, subsisten algunos problemas en lo que concierne al tiempo de ejecución de los experimentos (demanda varias y largas reacciones de PCR) y a la capacidad de comparar datos provenientes de distintos experimentos (Breyne y Zabeau, 2001).

1.2.- Métodos basados en el recuento de *tags* de cDNA: Análisis en Serie de la Expresión Génica (SAGE).

Esta técnica fue reportada en Velculescu y cols. (1995) en un intento por optimizar la eficiencia de la secuenciación de *expressed sequence tags* (ESTs).

El método SAGE consiste en generar fragmentos pequeños del extremo 3' del cDNA, los cuales serán ligados entre sí para formar concatémeros que contienen hasta 1000 fragmentos de cDNAs. Los concatémeros son clonados para dar lugar a una genoteca SAGE. El patrón de expresión de cualquier población de transcritos puede ser evaluado cuantitativamente al determinar la abundancia de cada fragmento e identificar el gen que corresponde a cada uno de ellos por su homología con genes anteriormente secuenciados y que están presentes en las bases de datos.

Esta técnica, sin embargo, contempla varios pasos de reacciones enzimáticas, pasos de purificación y de selección mediante cromatografía de afinidad en columnas de estreptavidina-agarosa, lo que demanda la utilización de grandes cantidades de RNA (Fujita y Silver, 1993). Además, las secuencias generadas tienen un largo promedio de 10 pares de bases, de modo que la probabilidad de falsos positivos es alta. No obstante, en los últimos años, la técnica se ha perfeccionado permitiendo disminuir 500 a 5000 veces la cantidad de RNA a utilizar (Datson y cols., 1999) y aumentar el largo de las secuencias generadas a 17 pb (Saha y cols., 2002). Respecto a esto último, se espera que un 99,8% de las secuencias del largo mencionado se encuentren sólo una vez dentro del genoma humano, con lo cual se ha disminuido la probabilidad de tener falsos positivos. Pese a los antecedentes señalados, actualmente la técnica continúa siendo de difícil manejo, ya que requiere muchos pasos experimentales, y el análisis de sólo una condición fisiológica implica construir

una genoteca SAGE. De este modo, es difícil comparar perfiles de expresión entre varias condiciones experimentales.

1.3.- Técnicas basadas en la hibridación de sondas con genes blanco: arreglos de cDNA.

Sin duda alguna, en la mayoría de las publicaciones recientes que abordan la caracterización de perfiles de expresión génica a gran escala se ha utilizado las técnicas llamadas genéricamente "arreglos de DNA" o *microarrays* (Schena y cols., 1995; Shalon y cols., 1996). Su implementación, en 1996 por P.O. Brown y cols. fue uno de los mayores hitos de la era post-genómica, ya que esta nueva técnica permitía medir cambios cualitativos y cuantitativos en la abundancia de transcritos como resultado de la expresión diferencial de genes y/o de la estabilidad diferencial de sus correspondientes mRNAs (Van Hal y cols., 2000).

1.3.1.- Principios técnicos de los arreglos de DNA más usados.

La técnica consiste en inmovilizar secuencias de DNA conocidas (DNA blanco) sobre una matriz sólida. Dependiendo del formato empleado, productos de PCR provenientes de genotecas de cDNA u oligonucleótidos (50-70 mer) son depositados en volúmenes pequeños sobre una superficie de nylon o vidrio, para generar *macroarrays* y *microarrays*, respectivamente. Alternativamente, en el caso de los *chips* de oligonucleótidos, secuencias de oligonucleótidos (20-40 mer) son sintetizadas *in situ* sobre una superficie de cerámica/silicio (Lockhart y

cols., 1996). En cada punto de la matriz sólida, se depositan numerosas moléculas de DNA específicas de un gen particular. De este modo, el sistema es miniaturizado, pudiendo sembrarse hasta decenas de miles de genes en superficies no mayores a cuatro centímetros cuadrados (revisado en Freeman y cols., 2000).

En segundo lugar, las muestras de RNA o mRNA a analizar son sometidas a un proceso de transcripción reversa, en el cual los cDNAs son marcados, para su posterior detección y cuantificación. Existen esencialmente dos formas de marcación: radioactiva con ^{32}P o ^{33}P -[dCTP] y fluorescente con Cy3 y Cy5-[dUTP]. En un experimento típico de *microarrays*, las dos muestras de RNA cuyo patrón de expresión nos interesa determinar, son marcadas separadamente con cada fluoróforo. Posteriormente, las poblaciones de cDNA marcado son hibridadas simultáneamente contra la lámina de vidrio en la que se encuentran depositados los DNAs blanco. Después de efectuar la lectura del *microarray* en un escáner láser y generar una imagen digital, se estima el nivel de expresión de cada gen a partir de la intensidad de los puntos de hibridación (generados por el apareamiento de cada especie de cDNA con su DNA blanco). Mientras mayor la intensidad de un fluoróforo en un punto del arreglo, es mayor la cantidad del correspondiente mRNA en la población analizada. Para el caso de los chips de oligonucleótidos, el procedimiento es similar al descrito anteriormente, pero es posible analizar mas de 40.000 genes simultáneamente, en cambio, con los *microarrays* el número se reduce a 10.000 genes (Freeman y cols., 2000). Esto se debe a que el sistema de síntesis *in situ*

de oligonucleótidos permite una concentración más eficiente de estas secuencias de DNA blanco en un espacio reducido.

En el caso de los *macroarrays*, las muestras de RNA de cada condición a analizar son marcadas radiactivamente, y luego hibridadas en membranas independientes. Posteriormente, las señales de hibridación deben ser traducidas en una imagen usando un escáner de foto-fosforescencia (*phosphor-scanner*). Si bien esta plataforma de arreglos de DNA permite analizar entre 500 y 2000 genes simultáneamente, se ha reportado que la sensibilidad de este método (definida como la cantidad necesaria de un transcrito presente en una muestra de RNA para que éste sea detectado de forma confiable) es hasta 100 veces mayor que la descrita para los otros métodos existentes (Bertucci y cols., 1999). Por otra parte, la sonda radiactiva posee un rango lineal de detección que es dos órdenes de magnitud mayor que en el caso de las sondas fluorescentes, y los arreglos en membranas pueden ser utilizados varias veces sin pérdida de calidad en los ensayos de hibridación (Baldi y Hatfield, 2002). Otra ventaja ofrecida por esta plataforma de arreglos de DNA es su menor costo respecto de los *chips* de oligonucleótidos y de los *microarrays* en matrices de vidrio.

Los formatos de arreglos de DNA mencionados anteriormente han entregado valiosa información para el estudio de la regulación de la expresión génica en varios sistemas. Sin embargo, la técnica es susceptible de ser optimizada en términos de su sensibilidad, especificidad, reproducibilidad, y costo. Un gran número de publicaciones han revisado y descrito nuevas

tecnologías aplicadas en arreglos de cDNA (Heller, 2002). Además, numerosos trabajos han reportado estrategias para optimizar el análisis de datos, incluyendo el ajuste de un área de lectura de puntos en un experimento particular, hasta el estudio de algoritmos matemáticos para identificar genes diferencialmente expresados, (revisado en Leung y Cavalieri, 2003).

1.3.2.- Algunos ejemplos de la aplicación de arreglos de DNA en análisis de expresión génica a gran escala.

Desde la creación de esta técnica, son numerosos los trabajos que han entregado a la comunidad científica las primeras aproximaciones para establecer redes de regulación transcripcional. Estos trabajos han sido aplicados a un sinnúmero de sistemas biológicos y han demostrado que la técnica de arreglos de DNA entrega resultados confiables al compararlos con aquellos obtenidos de técnicas de cuantificación a menor escala, tales como PCR en tiempo real. En general, los experimentos reportados a la fecha caen dentro de dos categorías distintas: i) Estudios de casos, en los que se compara los cambios en la expresión de genes entre una situación control y una experimental y ii) Series de tiempo, experimentos en los cuales se establece una dinámica temporal (a intervalos de tiempo definidos con respecto a un tiempo de referencia) del transcriptoma en el sistema biológico que se está analizando.

Trabajos que ejemplifican los estudios de casos son aquellos reportados por Alizadeh y cols. (2000) y Bittner y cols. (2000) quienes utilizan *microarrays*

de cDNA para caracterizar varios tipos de tumores sobre la base de sus patrones de expresión. El uso de datos provenientes de estos experimentos ha permitido, posteriormente, desarrollar un método para clasificar muestras provenientes de tejidos tumorales y determinar a qué tipo de cáncer corresponde en base a sus patrones de expresión. En la segunda categoría se encuentra el trabajo de De Risi y cols. (1997). En él se examina el programa temporal de la expresión génica que acompaña al cambio metabólico cuando las células de *Saccharomyces cerevisiae*, inoculadas en un medio con glucosa, comienzan a utilizar en aerobiosis el etanol producido durante la fermentación después de agotar el azúcar fermentable. Este estudio también da luces acerca de la posible función de genes desconocidos, a partir de la similitud de sus patrones de expresión con aquellos definidos para genes de función conocida.

En plantas, son numerosos los estudios que han utilizado esta técnica para examinar los patrones de expresión en *Arabidopsis thaliana*. Por ejemplo, Schaffer y cols. (2001) utilizaron *microarrays* de cDNA que contenían 7800 genes únicos (30% de los genes anotados para *A. thaliana*) y analizaron los patrones de expresión en plantas sometidas a ciclos de luz y oscuridad. En este estudio se determinó que un 2% de los genes analizados se expresaba diferencialmente en un patrón característico dentro de cada ciclo, sugiriendo que ellos tienen como función regular el funcionamiento del ciclo circadiano. Por otra parte, Maleck y cols. (2000) utilizaron arreglos de cDNA conteniendo 7000 genes para analizar los cambios en la abundancia de transcritos en plantas sometidas a ácido salicílico, molécula necesaria y suficiente para producir

resistencia sistémica adquirida (SAR) en la planta. En este estudio se identificaron alrededor de 300 genes cuyo cambio de expresión era significativo durante la SAR y un elemento promotor común para un grupo de genes inducidos en SAR (Maleck y cols., 2000).

En relación con la aplicación de esta tecnología en modelos vegetales de importancia biotecnológica, Brinker y cols (2004) utilizaron *microarrays* para analizar la expresión génica durante el desarrollo, inducido por auxinas, de las raíces secundarias de *Pinus contorta*. En este trabajo se hibridó RNA extraído en distintos tiempos durante el desarrollo de las raíces secundarias contra un *microarray* que contenía 2178 genes de *Pinus taeda*. De estos 2178 genes, se identificaron 220 genes expresados diferencialmente que pertenecían a distintas categorías funcionales, como síntesis y degradación proteica, fotosíntesis, replicación celular, síntesis y degradación de la pared celular, respuesta a estrés, metabolismo hormonal y transducción de señales, lo cual permitió a los autores generar un modelo acerca del desarrollo la raíz de una gimnosperma (Brinker y cols., 2004). Además se han estudiado mediante *macroarrays* de cDNA los perfiles de expresión de la caña de azúcar (*Saccharum spp.*) en respuesta a estrés por bajas temperaturas en distintos tiempos, dilucidando parcialmente la cascada metabólica que se activa en respuesta a este estímulo ambiental (Nogueira y cols., 2003). Este último trabajo es uno de tantos que intenta, a través del un análisis del transcriptoma, identificar genes clave en un problema relevante para la producción vegetal y de aplicación biotecnológica.

Estos estudios nos demuestran que los arreglos de DNA constituyen una poderosa herramienta para la caracterización del transcriptoma en diversos procesos biológicos, contribuyendo a la asignación de funciones y la identificación de posibles genes marcadores del proceso analizado, los que pueden ser claves para caracterizar disfunciones de organismos de importancia económica y biotecnológica. En este contexto, surge la necesidad de implementar y aplicar esta tecnología al estudio de problemas relevantes para la economía del país, uno de los cuales menciono a continuación.

2.- El problema de la harinosidad en duraznos (*Prunus persica*).

Los frutos de durazno que se exportan desde Chile a Europa y Norteamérica están sometidos a prolongados períodos de transporte a bajas temperaturas, debido a esto sufren alteraciones fisiológicas que dan como resultado una pérdida de calidad y de competitividad. Una de estas alteraciones fisiológicas más importante es la harinosidad, la que se manifiesta en una falta de jugo en el fruto. Se ha demostrado que esta alteración ocurre en dos etapas: una de inducción por el mencionado almacenamiento en frío (4 °C) y una etapa posterior de expresión, alcanzada luego de un período de maduración a temperatura ambiente (Crisosto y cols., 1999). Así, el problema es detectado sólo cuando el producto llega al consumidor.

Uno de los mecanismos más aceptados para explicar el fenómeno de la harinosidad, consiste en la gelificación del agua de la fruta, producto de la liberación de pectinas con un bajo grado de esterificación, altamente solubles

en agua, que provienen de la degradación de la lámina media de la pared celular en el mesocarpo de los frutos. Se ha probado que esta degradación es producida por la exposición al frío, lo cual generaría una pérdida del balance en la expresión de genes que codifican para enzimas que degradan estas pectinas, en particular, pectinmetilesterasa y exopoligalacturonasa (Crisosto y cols., 1999). Sin embargo, la complejidad del metabolismo de la pared celular, la diversidad de elementos que participan en éste, y la conexión de ellos con otros procesos, tales como la maduración y la respuestas a estrés por bajas temperaturas, sugieren que el origen de la harinosidad podría ser más complejo que el desbalance entre dos actividades enzimáticas (Ben-Arie y Sonogo, 1980; Giovannoni, 2001). En particular, proponemos que otros componentes del metabolismo de pared celular podrían encontrarse alterados en este desorden, como por ejemplo, elementos de respuesta a estrés, entre otros.

Existe información disponible acerca del control transcripcional en frutos, incluyendo el trabajo de Aharoni y cols. (2000) quienes utilizaron *microarrays* de cDNA para identificar genes responsables de la biogénesis de flavonoides asociados con el sabor en frutillas, y el trabajo de Trainotti y cols. (2003) que caracterizó los perfiles de expresión génica en diferentes etapas del desarrollo del fruto de durazno. Tomando en cuenta lo anterior y considerando que el problema de la harinosidad podría originarse por alteraciones en la maquinaria de reacción al estrés por bajas temperaturas, y en la expresión de proteínas claves para la estructura de la pared celular de estos frutos, en el Laboratorio de Bioinformática y Expresión Génica del INTA nos hemos propuesto

implementar y realizar un análisis de expresión génica a gran escala en duraznos para conocer el efecto del almacenamiento a bajas temperaturas sobre la adquisición de la condición harinosa, a través de un diseño experimental que contempla la caracterización de varios parámetros de la siembra de membranas con librerías de genes. Estos análisis se realizaron en el contexto del proyecto "Genómica Funcional en Duraznos y Nectarines: Plataforma para potenciar la competitividad de frutas de exportación en Chile" y durante esta Memoria de Título se implementó un sistema automatizado de siembra para arreglos de DNA y un protocolo que permite la extracción de datos a partir de membranas hibridadas con sondas radiactivas (*macroarrays* de cDNAs).

3.- Hipótesis.

El análisis de expresión génica usando la técnica de *macroarrays* de cDNAs, permitirá detectar genes asociados con el fenómeno de harinosidad en *Prunus persica*.

4.- Objetivo general.

Implementar un protocolo para la producción de *macroarrays* de cDNA en membranas de nylon y su hibridación con sondas complejas de cDNAs.

5.- Objetivos específicos.

Para lograr el objetivo planteado anteriormente se han propuesto los siguientes objetivos:

5.1- Optimizar parámetros para la siembra automatizada de *macroarrays*, entre ellos: densidad de puntos en el arreglo, cantidad de DNA a sembrar, modalidad de siembra, tipo de DNA sembrado y adquisición de datos,

5.2- Examinar la sensibilidad y reproducibilidad de los experimentos de hibridación,

5.3- Aplicar el método optimizado al análisis de un problema fisiológico: harinosidad en duraznos de la variedad O'Henry, variedad que presenta una alta propensión a este problema.

MATERIALES Y MÉTODOS.

1.- Obtención de los DNAs blanco.

1.1- Extracción de DNA plasmidial:

Para la extracción de DNA plasmidial se utilizaron las siguientes cepas transformadas de *E. Coli* DH5- α :

- i) Transformantes conteniendo el vector *pGEM-t Easy* (Promega®) sin inserto.
- ii) Transformantes que contienen el producto de PCR del gen *exp1* (expansina 1) de *Prunus persica* ligado en el sitio de clonamiento del vector anteriormente nombrado.
- iii) Clones seleccionados al azar provenientes de una genoteca de sustracción estado específica de *Drosophila melanogaster*.

Las bacterias fueron inoculadas en 5 mL de medio LB, conteniendo ampicilina en una concentración de 0,1 mg/mL. Los cultivos fueron incubados por 18 horas a 37 °C y se extrajo DNA plasmidial de 5 mL del cultivo utilizando el método de lisis alcalina (Birnboim, 1983), descrito en Sambrook, 1989.

1.2.- Amplificación y purificación de insertos a partir de DNA plasmidial.

Para la amplificación de los insertos se utilizaron los partidores T7 y SP6 (concentración final de 0,2 μ M), los que hibridan con secuencias del vector *pGEM-T Easy* (Promega®), situadas a ambos lados del inserto. Los sustratos de la reacción de PCR fueron 200 μ g de DNA plasmidial, *buffer* 1X (Invitrogen

®), $MgCl_2$ 1,5 mM, dNTPs 0,2 mM y 2,5 U de Taq DNA polimerasa (Invitrogen®). Se utilizó un programa de 29 ciclos (desnaturalización 30 s a 94 °C, alineamiento 30 s a 55 °C y extensión de 2 minutos. a 72 °C). Las reacciones tenían un volumen final de 100 μ L. Los productos de PCR fueron precipitados utilizando isopropanol y acetato de sodio, según lo descrito en Diehl y cols., 2002, obteniéndose una concentración de entre 0,1 y 2,5 μ g/ μ L de producto.

1.3.- Selección de clones a partir de una genoteca de ESTs de durazno de la Universidad de Clemson.

Los DNAs a sembrar fueron seleccionados desde una genoteca de ESTs de durazno maduro de la variedad Loring, construida y parcialmente secuenciada en la Universidad de Clemson (South Carolina, EUA), y adquirida por el Laboratorio de Bioinformática y Expresión Génica del INTA, Universidad de Chile. Esta librería fue construida en los sitios *EcoRI* y *XhoI* del vector pBluescriptSK-, el cual posee un marcador de resistencia a ampicilina. La genoteca consta de 16.512 clones, distribuidos en 43 placas de 384 pocillos. De estos 16.512 clones disponemos de la secuencia de sólo 9.984, de las cuales 3.843 secuencias representan a genes únicos.

Para la selección de los ESTs presentes en la librería se realizó el siguiente procedimiento:

- a) Con el fin de seleccionar aquellas secuencias de la genoteca que presentan similitud (tanto nucleotídica como aminoacídica) con los genes anotados en la base de datos de TAIR (*The Arabidopsis Institute Research*), se elaboraron informes mediante los algoritmos Blast-n y tBlastX, a partir del alineamientos de las 3.843 secuencias de unigenes con la base de datos de CDS (secuencia codificante) de TAIR (www.arabidopsis.org), los que se ejecutaron bajo plataforma Linux. Cada informe incluye el número de clon asignado por Clemson, el mejor alineamiento realizado para cada Blast, la descripción de este alineamiento (ID) según el código de acceso (locus At para la base TAIR y código NM para GenBank), los valores del alineamiento (*e-value* y *score*) y una posible descripción según la clasificación de GO (*Gene Ontology*) en sus tres categorías, Función Molecular, Componente Celular y Proceso Biológico.
- b) Los resultados del tBlastX y del Blast-n fueron almacenados en distintas bases de datos y sus entradas fueron comparadas mediante consultas en lenguaje SQL. Se obtuvieron 2.055 informes de ESTs que, utilizando ambos algoritmos, alineaban con el mismo CDS de *A. thaliana* y 1.788 informes de ESTs que alineaban con diferentes CDS. De estos últimos, sólo se consideraron aquellos que presentaron el mejor alineamiento (menor *e-value* y mayor *score*; n= 1.622).

- c) Los ESTs resultantes fueron seleccionados manualmente de acuerdo con los criterios descritos en la Tabla I, lo cual dio como resultado 627 alineamientos obtenidos mediante Blast-n y 325 mediante tBlastX.
- d) Finalmente, se seleccionaron 880 clones que en sus informes presentaron un *e-value* menor a 10^{-10} , los cuales fueron organizados en 10 placas de 96 pocillos para su utilización en la construcción de las membranas de *macroarrays*.

Tabla I. Criterios de búsqueda de palabras claves o secuencias consenso utilizados para la selección manual de clones. Los genes sometidos a este análisis habían sido previamente seleccionados por su similitud con los genes presentes en la base de datos del TAIR (The *Arabidopsis* Information Resource, www.arabidopsis.org), usando los algoritmos Blastn y tBlastx

Criterios de búsqueda	Palabras clave/secuencias consenso seleccionadas
A) Búsqueda de palabras claves en la descripción del ID	Enzimas de pared celular, elementos del metabolismo de etileno, maduración, estrés por bajas temperaturas, etc.
B) Búsqueda de términos GO (The Gene Ontology Consortium. 2001)	Función molecular: actividad celulosa sintetasa, transporte de carbohidratos, actividad HSP, factor de transcripción, etc. Proceso biológico: Metabolismo de carbohidratos, síntesis de carotenoides y flavonoides, respuesta a estrés y etileno, etc. Componente celular: Pared celular, membrana plasmática, sistema de endomembrana, extracelular, etc.
C) Búsqueda en PubMed	Descripción en publicaciones atinentes a maduración y procesos afines.
D) Búsqueda de secuencias consenso	Secuencias de péptido señal mediante programas: Target P y PSORT (Emanuelsson y cols., 2000; Nakai y Norton, 1999)

1.4.- Amplificación de insertos a partir de cultivos bacterianos.

Los clones seleccionados a partir de la genoteca de ESTs de durazno fueron crecidos en placas de 96 pocillos conteniendo 200 μL de medio LB y ampicilina 50 $\mu\text{g}/\text{mL}$ a 37°C por 22 horas. Posteriormente 5 μL de este cultivo fueron agregados a 33 μL de agua libre de nucleasas y fueron calentados a 95°C por 15 minutos. Luego se agregó: *buffer* 1X (Invitrogen®), dNTPs 0,25 mM, MgCl_2 2 mM, partidor T3 y partidor T7 1 pM y 2,5 U de Taq DNA polimerasa (Proporcionada por el Laboratorio de Genética Molecular Vegetal, Facultad de Ciencias, Universidad de Chile), en un volumen final de 50 μL . Estos partidores hibridan con secuencias del vector de clonamiento que flanquean al inserto, permitiendo su amplificación. El programa consta de 30 ciclos que incluyen desnaturalización a 94°C por 30 segundos, alineamiento a 55 °C por 1 minuto y extensión a 72°C por 3 minutos e incluye una extensión final a 72°C por 10 minutos. Estas reacciones de PCR fueron llevadas a cabo en placas de PCR de 96 pocillos (Termowell Costar 6511, Corning Incorporated).

1.5.- Verificación y cuantificación de los productos de PCR.

Para verificar el tamaño de los productos de PCR, 5 μL de las reacciones de PCR fueron fraccionados en geles de agarosa al 1% en buffer TAE 1X (40 mM Tris, 1 mM EDTA disódico, 20 mM ácido acético), y las bandas visualizadas con Bromuro de Etidio (Ausubel y cols., 1998).

En el caso de los productos obtenidos a partir de la amplificación directa de alícuotas de cultivo bacteriano se realizaron los siguientes controles de calidad: 1) Si un clon presentaba más de una banda de amplificación, el PCR se repetía, 2) Si el resultado seguía siendo más de una banda, se amplificaba el inserto usando DNA plasmidial de los clones como sustrato. Solamente productos únicos de amplificación fueron utilizados en la siembra de membranas. De este modo se logró amplificar un 99% de los clones inicialmente escogidos. La cuantificación de todos los DNA blanco se llevó a cabo midiendo la absorbancia a 260 nm en un espectrofotómetro Perkin Elmer® MBA-2000.

1.6.- Obtención de genes controles.

Las membranas destinadas a los experimentos de expresión génica en frutos de durazno contenían los siguientes controles:

1) Gen heterólogo, *dapB*: Se sembraron tres diluciones del producto de PCR de un cultivo del clon *dap* (ATCC no. 87486), cuyo vector consiste en un cDNA del gen *dapB* de *Bacillus subtilis* (1.840 pb) clonado entre los sitios *XhoI* y *BamHI* de un plásmido pBluescriptII-KS+ en el cual se encuentra una cola de poli(A) río abajo del inserto. Inóculos de 18 horas de este clon fueron sometidos a reacciones de PCR similares a las descritas en la sección 1.2.1. Las concentraciones de DNA usadas fueron de 1,33 µg/µL para la primera dilución

de *dapB* (*dap1*), 0,42 $\mu\text{g}/\mu\text{L}$ para la segunda dilución (*dap2*), y 0,12 $\mu\text{g}/\mu\text{L}$ para la tercera (*dap3*).

2) Genes de posible expresión constitutiva y gen *exp1*: estos tres genes fueron obtenidos mediante RT-PCR desde mRNA de *Prunus persica* (detallado en sección 3.3). Los genes fueron *act1*, codificante para la actina 1 de *P. persica* (0,61 $\mu\text{g}/\mu\text{L}$) y *tub1*, codificante para la α -tubulina de *P. persica* (0,58 $\mu\text{g}/\mu\text{L}$). El gen que codifica para la expansina (*exp1*) fue ubicado en las placas con una concentración de 0,63 $\mu\text{g}/\mu\text{L}$.

3) Controles Negativos: en cada placa se ubicaron 2 controles negativos de hibridación: Vector pGEM-t Easy sin inserto y agua libre de nucleasas.

Estos controles fueron dispuestos en cada una de las 10 placas de 96 pocillos utilizadas. Así, cada una de ellas contenía 88 genes seleccionados de la genoteca Clemson, las tres diluciones del gen *dapB*, los dos genes de expresión constitutiva (*act1* y *tub1*), el gen *exp1* y los dos controles negativos (agua y vector sin inserto).

2.- Generación de *macroarrays*.

2.1.- Siembra de los DNA blanco en las membranas.

El DNA plasmidial, los productos de PCR y/o productos purificados de PCR fueron depositados en placas de 96 pocillos (Orange Scientific) en volúmenes de 15 μl y se agregó un volumen igual de DMSO. El DNA contenido en las placas fue sembrado en membranas de Nylon de distintas marcas,

mediante el sistema automatizado de siembra ChipWriter Compact (BioRad®), equipado con un número variable (entre 3 y 8) de agujas sólidas de sembrado (*solid pins*, ArrayIt® mod. SSP015), repitiendo entre una y seis veces la siembra de muestras en un mismo punto, con el fin de variar la concentración del el DNA sembrado. Se usaron distancias de 1200, 1500 y 1800 μm entre *spots*. Los DNAs fueron sembrados en duplicado o en tetraplicado según correspondiera a cada experimento. Cada programa de siembra generó un archivo de extensión "vrk", que relaciona el contenido de cada pocillo de la placa fuente con el lugar de la membrana en que éste fue depositado.

En el caso de los productos de PCR obtenidos a partir de la amplificación de la genoteca de Clemson, se diseñó un programa que permitiera la siembra del contenido de 10 placas de 96 pocillos en una membrana de Nylon de 8,5 x 11,5 cm^2 (GeneScreen Plus, NEN®). En este programa se ocuparon cuatro agujas sólidas de sembrado, y cada producto de PCR fue sembrado en una sola posición en la membrana con una separación de 1800 μm . Con el fin de tener un exceso de DNA blanco en la membrana, cada punto se sembró 6 veces. En promedio se sembró aproximadamente 76.86 ng de cada producto de PCR, valor estimado con la concentración promedio de los productos de PCR (1,83 $\mu\text{g}/\mu\text{L}$), el número de siembras por punto (6 siembras/punto) y el volumen promedio depositado por cada aguja de sembrado (7 nL).

2.2.- Tratamiento de las membranas sembradas.

Todas las membranas fueron sometidas a un baño de 7 minutos en solución de desnaturalización (0,5 M NaOH, 1,5 M NaCl), sin agitación. Luego las membranas fueron lavadas dos veces por tres minutos en solución de neutralización (1,5 M NaCl, 0,5 M Tris HCl pH 7,5, 1 mM EDTA) con agitación. Posteriormente, el DNA desnaturado fue unido covalentemente a las membranas mediante irradiación UV por 36 segundos a $120000 \text{ } [\mu\text{J/s}\cdot\text{cm}^2]$ en un horno UV (*UV-crosslinker* CL-100, UltraViolet Products®).

2.3.- Prehibridación, hibridación y lavado de las membranas.

Las membranas fueron pre-hibridadas en 5 mL de *buffer* de hibridación (SSC 5X, solución de Denhardt 5X, SDS 10 g/L, formamida 50%(v/v)) mas 250 μg de DNA de salmón, a 42 °C por un tiempo mínimo de 2 horas en un horno de hibridación. Una vez purificada la sonda ésta fue desnaturalizada a 100 °C por 15 minutos y luego puesta en hielo por 5 minutos. Se agregó la totalidad de la sonda purificada y desnaturalizada (no menos de 10^6 cpm) a las membranas pre-hibridadas. La hibridación se extendió por 18 horas a 42 °C. Las membranas hibridadas fueron sometidas a 4 lavados de 20 minutos a 65 °C en 20 mL de buffer de lavado (0,1% SDS, 1x SSC).

3. Obtención de sondas.

3.1.- Muestras biológicas

Las muestras utilizadas para extraer RNA correspondieron a: 1) mesocarpo de duraznos (*Prunus persica*, variedad O'Henry) cosechados y madurados a 20 °C por 8 días (no harinoso; NH; 38% de jugo), y 2) mesocarpo de duraznos cosechados, almacenados por 21 días a 4 °C, y madurados a 20 °C por 5 días (harinoso; H; 0% de jugo). El nivel de harinosidad de estas muestras fue medido a través de un procedimiento que determina el porcentaje de jugo presente en cada muestra (Crisosto y Labavitch, 2002). Este procedimiento fue realizado en el laboratorio de Post-cosecha del INIA La Platina. Las muestras utilizadas en este estudio, así como los datos de harinosidad fueron gentilmente proporcionados por los doctores Julio Retamales y Reinaldo Campos, del Laboratorio de Post-Cosecha del INIA La Platina.

3.2.- Extracción de RNA total de duraznos.

La extracción de RNA total se hizo utilizando el protocolo descrito por Meisel y cols. (2005), a partir de 3 gramos de tejido de mesocarpo de durazno. El precipitado final de RNA fue resuspendido en 35 μ L de agua libre de nucleasas. La integridad del RNA se verificó por electroforesis en geles de agarosa al 1%, usando TAE preparado en agua libre de nucleasas como buffer de corrida y visualizado con bromuro de etidio. Sólo se utilizaron aquellas

muestras de RNA total que mostraran un patrón electroforético en el que la banda correspondiente al rRNA 28S fuera de una intensidad dos veces mayor que la banda del rRNA 18S. La concentración de RNA, el contenido de contaminantes proteicos y de contaminantes fenólicos en las muestras extraídas fueron determinados midiendo la absorbancia a 260 nm, y calculando las razones de absorbancia $A_{260/280}$ y $A_{260/240}$, respectivamente. Se utilizó un criterio de calidad de RNA tal que si la extracción final presentaba una concentración mayor o igual que 70 $\mu\text{g}/\mu\text{L}$, una razón $A_{260/280}$ entre 1,8 y 2,1, y una razón $A_{260/240}$ entre 1,7 y 2,1, la muestra era utilizada en la síntesis de sondas complejas.

3.3.- Extracción de mRNA poli-(A) desde RNA total.

Las muestras de RNA total fueron sometidas a una purificación del mRNA poli-(A), mediante el *kit Oligotex mRNA Mini* (Qiagen®), según las instrucciones del fabricante, obteniendo un volumen final de 40 μL . La concentración de las muestras fue estimada mediante mediciones de absorbancia a 260 nm.

La integridad de los transcritos presentes en la muestra de mRNA fue evaluada mediante reacciones de RT-PCR utilizando partidores específicos de seis genes de durazno. Para esto se sintetizó cDNA con la enzima M-MLV RT (Promega®) usando mRNA como sustrato. Se incubaron 100 ng de mRNA y partidador oligo-dT en una concentración final de 30 $\text{ng}/\mu\text{L}$, a 70°C por 5 minutos,

en un volumen final de 15 μ L. Posteriormente, a cada tubo se agregaron 5 μ L de *buffer* 5X de la enzima, 15 mmoles de cada dNTP y 200 U de la transcriptasa reversa en un volumen final de 25 μ L. La reacción procedió a 42°C por dos horas y finalmente fue incubada a 70°C por 15 minutos para inactivar la enzima.

Los cDNAs sintetizados sirvieron como sustrato para la amplificación de los genes mostrados en la Tabla II. La reacción de PCR contenía 100 ng de cDNA, 5 pmoles de cada partidor, MgCl₂ 1,5 mM, dNTPs 0,2 mM y 2,5 U de Taq DNA polimerasa (Invitrogen®) en un volumen final de 25 μ L. Se utilizó un programa de 29 ciclos (desnaturalización 30 s a 94 °C, alineamiento 30 s a 55 °C y extensión de 2 minutos a 72 °C) y una extensión final a 72 °C por 6 minutos. Los productos de las amplificaciones fueron examinados en geles de agarosa al 1% para verificar la presencia de bandas de los tamaños esperados. Si alguno de los genes no presentaba una amplificación satisfactoria, entonces el mRNA no calificaba para posteriores experimentos.

Tabla II. Genes de *Prunus persica* amplificados desde cDNA como control de calidad de los mRNA extraídos. Las secuencias de los partidores se muestran en sentido 5' \rightarrow 3'.

Nombre del Gen.	Nº Acceso Genbank	Partidor-5'	Partidor-3'	Tamaño Esperado (pb)
aminociclopropano-1-carboxilato oxidasa (<i>aco1</i>)	AF532976	aatcatcaactggagggc	cgaatttcgggtacactt	827
proteína de stress por ácido abscísico (<i>SRprot</i>)	AF317032	gaaacctcagactaccctca	cttcttccttgcctcctt	478
Expansina 1 (<i>exp1</i>)	AB029083	tgactatggaggaggatg	cattgctggtcacagtcct	632
Receptor de etileno (<i>etr</i>)	AF396830	gtcaacgaaacgctatgc	caagtcactaccaccag	848

3.4.- Transcripción *in vitro* del gen *dapB*.

El mRNA del gen *dapB* fue sintetizado *in vitro* con el kit *Riboprobe*® *in vitro Transcription Systems* (Promega®) usando la enzima T3 RNA polimerasa. El templado para esta reacción fue DNA plasmidial del clon *dapB* (ATCC no. 87486) previamente linearizado en su sitio *NofI*, y se usó como partidor de la transcripción el oligonucleótido T3. El RNA obtenido fue cuantificado por absorbancia a 260 nm.

3.5.- Marcación de sondas.

El oligonucleótido T7 fue marcado incubando una mezcla de 30 pmoles del óligo, 10 pmoles de γ^{32} -dATP, 1 μ L de buffer T4 polinucleótido quinasa 10X, 5 U de la enzima T4 polinucleótido quinasa (Promega®) (10 μ L finales) por 30 minutos a 37°C. La purificación de esta sonda se llevó a cabo mediante precipitación con glicógeno (20 μ g) y etanol 100%, resuspendiendo finalmente en 100 μ L de buffer TE.

Las sondas simples (gen *exp1*) fueron obtenidas desde DNA plasmidial del clon *exp1* digerido con la enzima *EcoRI*. El inserto fue purificado desde geles de agarosa utilizando el *Concert Rapid Gel Extraction System* (Life Technologies). La marcación con α^{32} -dCTP del fragmento purificado se llevó a cabo usando el kit *Random Primers Labelling System* (Promega®), según las recomendaciones del fabricante. Para purificar esta sonda se utilizó el kit

Wizard DNA Clean-up de Promega. Se obtuvo un volumen final de 100 μL de sonda.

Las sondas complejas de cDNA de hebra simple fueron sintetizadas usando la transcriptasa inversa Superscript II en presencia de $\alpha^{32}\text{-dCTP}$, teniendo como templado mezclas de RNA total de *Drosophila melanogaster* (2 μg) o mRNA de *Prunus persica* (70 o 150 ng), las que además contenían cantidades conocidas de mRNA *dapB*, correspondientes al 1%, 0,1% y 0,01% de la cantidad de mRNA presente en las muestras. Las reacciones fueron realizadas en un volumen de 15,5 μL y consistían de los mRNAs, 2 μg de partidor óligo dT, 12,5 mmoles de dATP, dGTP y dTTP, 125 μmoles de dCTP, y 50 μCi de $\alpha^{32}\text{-dCTP}$. Las reacciones fueron sometidas a una temperatura de 70°C por 10 min, seguida una disminución de temperaturas entre 65 y 45 °C (5 °C por minuto) y una etapa final a 42 °C por 10 minutos. Inmediatamente después se agregó 12.5 μL del buffer 2x de la enzima, 0,25 mmoles de DTT, 40 U de RNAsina (Promega®) y 200 U de Superscript II. La mezcla final (25 μL) fue incubada por 2 horas a 42°C. Luego se inactivó la enzima a 70 °C por 15 minutos. El RNA fue degradado agregando 5,3 μL de buffer M1 (0,2% SDS, 0,1 M EDTA, 1,7 M NaOH) a la reacción e incubando ésta por 30 minutos a 68 °C. La solución se neutralizó agregando 26 μL de buffer M2 (0,4 M Tris pH 7,4, 0,02 M HCl), después de lo cual la sonda fue purificada con el kit *Qiaquick Purification System* (Qiagen®) y resuspendida en un volumen de 100 μL .

4.- Obtención de Datos.

4.1.- Obtención de imágenes.

Las membranas hibridadas fueron expuestas en una pantalla *Phosphor* (Kodak®) por tiempos variables para obtener imágenes en formato TIFF a una resolución de 50 μm , con el equipo Personal Molecular Image FX (BioRad®), mediante la utilización del software Quantity One® (BioRad®).

4.2.- Cuantificación de las imágenes.

Las imágenes generadas fueron analizadas con el software VersArray Analyzer (BioRad®), para lo cual se construyó una grilla de cuantificación usando el modo de búsqueda automática de grillas. Los nombres de cada punto de siembra en estas grillas (*.gal) fueron asignados con el programa controlador del sembrador, ChipWriter Compact (BioRad®), combinando archivos de contenido de placa (*.txt) con archivos de coordenadas de siembra (*.vrk). Las grillas etiquetadas permitieron asignar las mediciones de intensidad a las muestras sembradas en los correspondientes puntos de siembra (de aquí en adelante, llamados "*spots*"). Los siguientes parámetros fueron cuantificados: intensidad cruda, intensidad de ruido, intensidad neta, desviación estándar del ruido y área del *spot*. Las tablas obtenidas en formato VersArray fueron transformadas a formato Excel, para su posterior análisis.

4.3.- Tratamiento y selección de datos.

Cuando las replicas de los *spots* fueron sembradas en una misma membrana, se calculó la intensidad neta promedio, la desviación estándar y el coeficiente de variación (CV) de las réplicas. Sólo se aceptaron los datos cuyas réplicas tuvieran un coeficiente de variación menor a 0,5. El coeficiente de variación se define como el cuociente entre la desviación estándar de los valores de intensidad para un grupo de réplicas, y el valor promedio de éstas.

En el caso de los experimentos de expresión génica en frutos de durazno, las réplicas correspondieron a 2 grupos de tres membranas idénticas hibridadas dos veces contra sondas del mismo origen. De este modo se obtuvieron datos provenientes de doce hibridaciones, seis correspondientes a la condición no harinosa (NH) y seis a la condición harinosa (H). Los datos fueron sometidos a los siguientes análisis:

- 1.- Comparación entre réplicas: se compararon las membranas de cada grupo de acuerdo a su distribución de intensidades netas. Para ello se elaboraron histogramas de distribución de intensidad neta para cada una de las 12 hibridaciones en el programa SYSTAT v.10.0. Luego, se comparó cada uno de los valores de hibridación de las membranas con los de sus respectivas réplicas, con el fin de verificar que el valor de intensidad de cada punto en las distintas membranas fuera similar. Ésta comparación fue realizada mediante una correlación de *Pearson* (SYSTAT v.10.0). En conjunto con este análisis se determinó que el valor de corrección de Bonferroni para cada comparación

fuera cercano a cero, es decir, que las similitudes detectadas no se debían al azar.

2.- Determinación de la calidad de los puntos: en cada membrana, se aplicaron los siguientes filtros de calidad:

- Se eliminaron aquellos datos cuyo valor de intensidad neta fuera menor o igual a cero.
- Se consideró como genes expresados a aquellos puntos cuyo valor de intensidad neta fuera mayor a la desviación estándar de ruido local:
 $q_{\text{sig-noise}} = 1 - [\text{bkg1} / (\text{sig} + \text{bkg1})]$, donde sig es la señal y bkg1 es el valor del ruido local medido en la periferia del área del punto de siembra (Wang y cols, 2001).
- Se asignó un puntaje de calidad (Qcom) a cada uno de los puntos, de acuerdo a lo reportado en Wang y cols. (2001). Este puntaje considera parámetros como el tamaño del punto (q_{tam}), la relación señal-ruido ($q_{\text{S-R}}$), la variabilidad del ruido local (q_{R1}) y el exceso del ruido local (q_{R2}). El puntaje de calidad está definido por la fórmula:

$$Q_{\text{com}} = (q_{\text{tam}} \times q_{\text{S-R}} \times q_{\text{R1}} \times q_{\text{R2}})^{0,25}$$

3.- Normalización: con el propósito de comparar las señales de hibridación recuperadas entre distintas membranas, los datos fueron normalizados usando

el valor promedio del control heterólogo *dapB*, ya que éste fue adicionado en cantidades iguales en la sonda provenientes de frutos no harinosos y harinosos.

4.- Calidad de réplicas normalizadas: se calculó el CV de los datos normalizados en las tres réplicas de cada tratamiento, aceptando a aquellos genes con un valor de CV menor a 0,5. En aquellos grupos de réplicas que presentaron un CV mayor a 0,5 se determinó si existían réplicas cuyos valores estuvieran alejados del valor promedio. Estos valores fueron eliminados y se calculó el CV sólo con las réplicas restantes.

5.- Determinación de cambios significativos de expresión: se calculó el estadístico t con el objetivo de detectar los genes que presentan cambios significativos en su expresión entre las condiciones no harinoso y harinoso. Los genes cuyo cambio de expresión fue significativo ($p < 0,05$) fueron incluidos en la Tabla VI (página 64).

4.4.- Análisis de dominios proteicos en secuencias de función desconocida.

En el caso de encontrar secuencias diferencialmente expresadas, y que hacían “match” con genes que codificaran para proteínas de función desconocida, estas secuencias fueron sometidas a un análisis de dominios proteicos utilizando la base de datos InterPro (Apweiler y cols., 2001), mediante la interfaz de búsqueda InterProScan (<http://www.ebi.ac.uk/InterProScan>).

RESULTADOS.

1.- Optimización de parámetros de siembra.

En esta primera parte del trabajo se describen los resultados obtenidos durante la puesta a punto de los procedimientos de construcción de *macroarrays* de DNA y obtención de datos.

1.1.- Cuantificación de las señales.

1.1.1.- Obtención de datos.

La lectura de las membranas se efectuó mediante el uso de una pantalla *Phosphor* (Kodak®), previamente expuesta a una fuente de radiación (membrana hibridada con la sonda radiactiva), en un *Phosphor Scanner* (Personal Molecular Imager FX, BioRad®) a través del programa Quantity One (BioRad®). El *scanner* tiene una resolución máxima de 50 μm y el rango lineal de intensidades que detecta es de cinco órdenes de magnitud. La cuantificación de señales en membranas de Nylon se realizó mediante el programa de cuantificación VersArray Analyzer (BioRad®), a partir de imágenes en formato TIFF.

El programa VersArray Analyzer tiene la capacidad de ubicar en la imagen todos los *spots* del arreglo, identificarlos gracias a una matriz de texto delimitado por tabulaciones que indica el contenido de cada pocillo de la placa de siembra, asociar una celda de cuantificación a cada uno de ellos y cuantificar

la intensidad de los píxeles que se encuentran dentro de esta celda. Sin embargo, el usuario debe contar con los archivos de siembra (*.gal) elaborados por el programa controlador del robot sembrador (ChipWriter Compact, BioRad®), el cual combina un archivo de texto (*.txt) que porta información acerca del contenido de cada pocillo en la placa fuente, con la información de un archivo de siembra (*.vrk) que indica el lugar de la membrana donde se ha sembrado el contenido de cada pocillo. Este archivo (*.gal) debe ser incorporado en el programa de cuantificación en el momento de construir la "grilla de cuantificación" (conjunto ordenado de las celdas de cuantificación correspondientes a todos los *spots* del arreglo), ya que con él cada celda de cuantificación estará debidamente identificada en relación con el origen de la muestra sembrada en el punto correspondiente. Este procedimiento se resume en el diagrama de flujo mostrado en la Figura 1.

Luego de alinear todas las celdas de la grilla con todos los *spots* del arreglo, el programa cuantificador arroja una tabla de datos que contiene los valores de intensidad cruda, intensidad de ruido e intensidad neta para cada celda de cuantificación (previamente etiquetada). Esta tabla de datos puede ser exportada desde VersArray Analyzer hasta Microsoft Excel (Microsoft®), donde los datos pueden ser analizados.

1.1.2.- Medición de la intensidad cruda.

Si bien el software VersArray Analyzer ofrece diferentes modalidades de cuantificación de las intensidades de hibridación, se decidió medir la intensidad

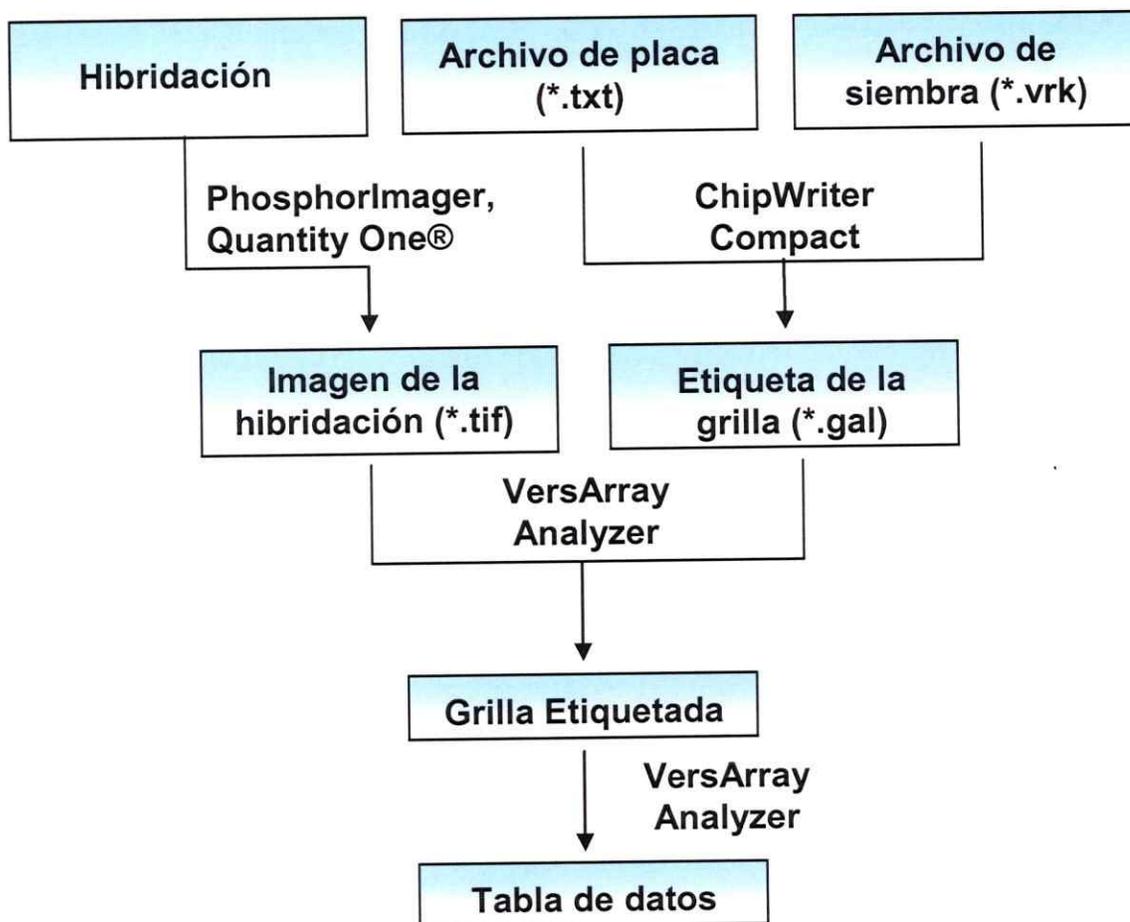


Figura 1. Obtención de datos desde membranas de *macroarrays*. Las imágenes de las membranas de *macroarrays* fueron obtenidas mediante el programa Quantity One (Biorad). Previo a la lectura de las membranas hibridadas, se identificó la posición de cada DNA sembrado en la membrana mediante la generación de un archivo *.gal. Posteriormente se ajustó una grilla a cada imagen obtenida en la hibridación, lo que permitió la cuantificación adecuada de la intensidad de cada *spot*. Los datos generados mediante esta cuantificación fueron exportados a Microsoft Excel, para su análisis.

del *spot* como la mediana de los valores de intensidad para todos los píxeles dentro de la celda de cuantificación, ya que de este modo se descartarían los píxeles cuyos valores se escaparon de la tendencia promedio en la celda de cuantificación. Esta forma de medir la intensidad cruda dentro del *spot* ha ganado aceptación en trabajos recientes (Herwig y cols., 2001; Nogueira y cols, 2003).

1.1.3.- Medición de la intensidad del ruido y cálculo de la intensidad neta.

Para cuantificar la intensidad de hibridación en los *spots*, es necesario determinar en primer término, la intensidad del ruido. El ruido puede originarse por hibridaciones inespecíficas de la sonda con la membrana, señales residuales de *spots* vecinos e irregularidades de la superficie de la membrana, las que en conjunto resultan en un valor de ruido que varía en cada punto del arreglo, lo que hace inconveniente utilizar un mismo valor de intensidad de ruido para todos los *spots*. Una solución más adecuada consiste en corregir la intensidad cruda con el valor de intensidad de los píxeles que se encuentran en el área comprendida entre la celda de cuantificación y un cuadrado circunscrito a ella. Esta modalidad, llamada "área de esquinas locales", tiene la ventaja de considerar sólo espacios "entre *spots*". Con ello, nuestro sistema de cuantificación de las intensidades netas de los *spots* consistirá en restar el valor de ruido local (mediana de los valores de intensidad de los píxeles que se

encuentran en el área de esquinas locales) al valor de intensidad cruda (mediana de los valores de intensidad de todos los píxeles de la celda).

1.2.- Determinación de los parámetros de siembra

Se utilizaron 24 diluciones de un DNA plasmidial para obtener 0 a 90 ng de DNA por *spot* (2.12×10^7 a 1.4×10^{10} moléculas/*spot*, Fig. 2, panel A). Estas diluciones fueron sembradas en 6 grupos o supergrillas (Fig. 2, panel B) con una separación entre *spots* de 1200 ó 1800 μm . Además, cada dilución fue depositada 1, 3 y 5 veces en el mismo lugar con el propósito de variar la masa del DNA. De este modo, el conjunto de diluciones fue sembrado de 6 maneras distintas en 6 lugares diferentes de la membrana. Cada dilución fue sembrada en duplicado para examinar el nivel de variación que entrega cada modalidad de sembrado (Fig. 2, panel C). Estas membranas fueron hibridadas con 330 ng de una sonda generada a partir del oligonucleótido T7 (20 pb) marcado con $\gamma^{32}\text{dATP}$ utilizando el fragmento Klenow de la DNA polimerasa de *E. coli*. Esta sonda tiene un 100% de identidad con la secuencia del promotor T7, presente en el vector sembrado (pGEM-T Easy, Promega®). Utilizando este diseño experimental evaluamos:

a) Cantidad de DNA a sembrar: En la Figura 3 podemos observar una relación lineal entre la Intensidad Neta y la cantidad de DNA sembrado. En ambos gráficos, el valor de la intensidad neta supera al valor de la intensidad de ruido a

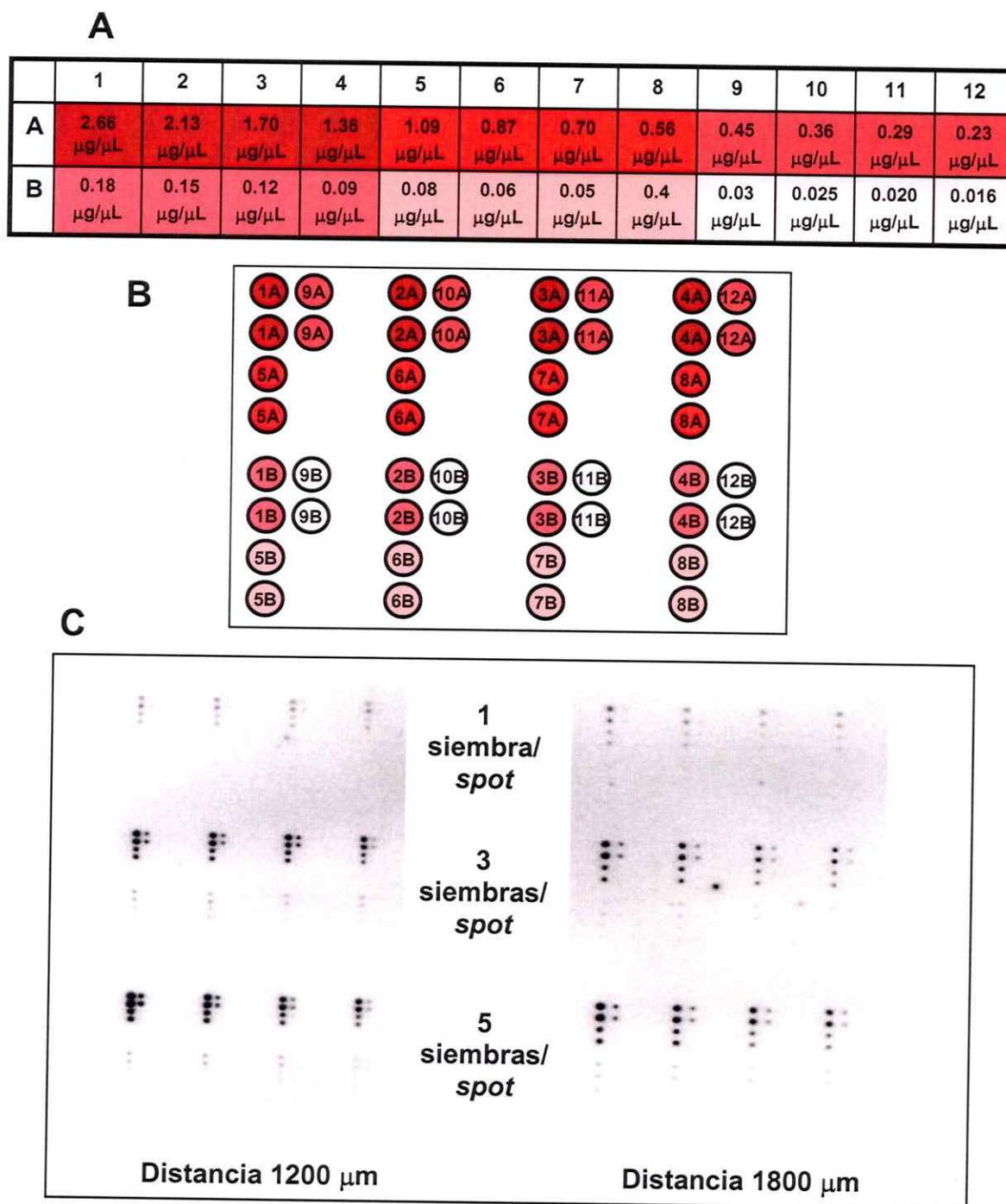


Figura 2. Parámetros de siembra. Veinticuatro diluciones (2.12×10^7 a 1.4×10^{10} moléculas/*spot*) del vector pGEM-T Easy (Promega) dispuestas en una placa de 96 pocillos (A) fueron sembradas en duplicado en la membrana de nylon para generar una supergrilla (B), la cual fue sembrada en 6 modalidades distintas en cuanto a separación entre *spots* y número de siembras por *spot* (C). Estas membranas fueron hibridadas con 330 ng del oligonucleótido T7 marcado con γ^{33} -dATP.

partir de 6×10^9 moléculas de DNA blanco (20 ng de DNA plasmidial), correspondiente a la cantidad de DNA que permite distinguir claramente (razón intensidad bruta / intensidad del ruido > 2) el valor de la intensidad del *spot* del valor del ruido local.

b) Distancia entre los spots: Un punto importante a considerar en la elaboración de *macroarrays* de cDNA es determinar la distancia entre los *spots*. Si esta distancia es pequeña, la membrana tendrá capacidad para un mayor número de *spots*. Sin embargo una distancia muy reducida puede dar origen a interferencias entre las señales de *spots* vecinos, lo cual aumentaría el ruido y resultaría en una pérdida de señal. El efecto de la distancia entre *spots*, se ilustra en la Figura 3.

Con una distancia de $1800 \mu\text{m}$, el nivel de ruido se mantiene constante (Fig. 3B), mientras que con $1200 \mu\text{m}$, el ruido aumenta con la intensidad neta (Fig. 3A). La correlación Intensidad Neta/DNA blanco es superior cuando los *spots* se encuentran separados por una distancia de $1800 \mu\text{m}$ ($R = 0,96$) comparada con la obtenida para los *spots* sembrados cada $1200 \mu\text{m}$ ($R = 0,88$). Al calcular la desviación estándar (σ) de todos los datos de ruido para ambas configuraciones de sembrado, vemos que la σ para la separación de $1200 \mu\text{m}$ es dos veces mayor que la obtenida para la separación de $1800 \mu\text{m}$. En conclusión, una distancia de $1800 \mu\text{m}$ permite que los datos obtenidos sean de mejor calidad, ya que existe un mejor ajuste lineal, una menor variación y una

mayor razón intensidad bruta/intensidad del ruido, aun cuando se disminuya la densidad de *spots*.

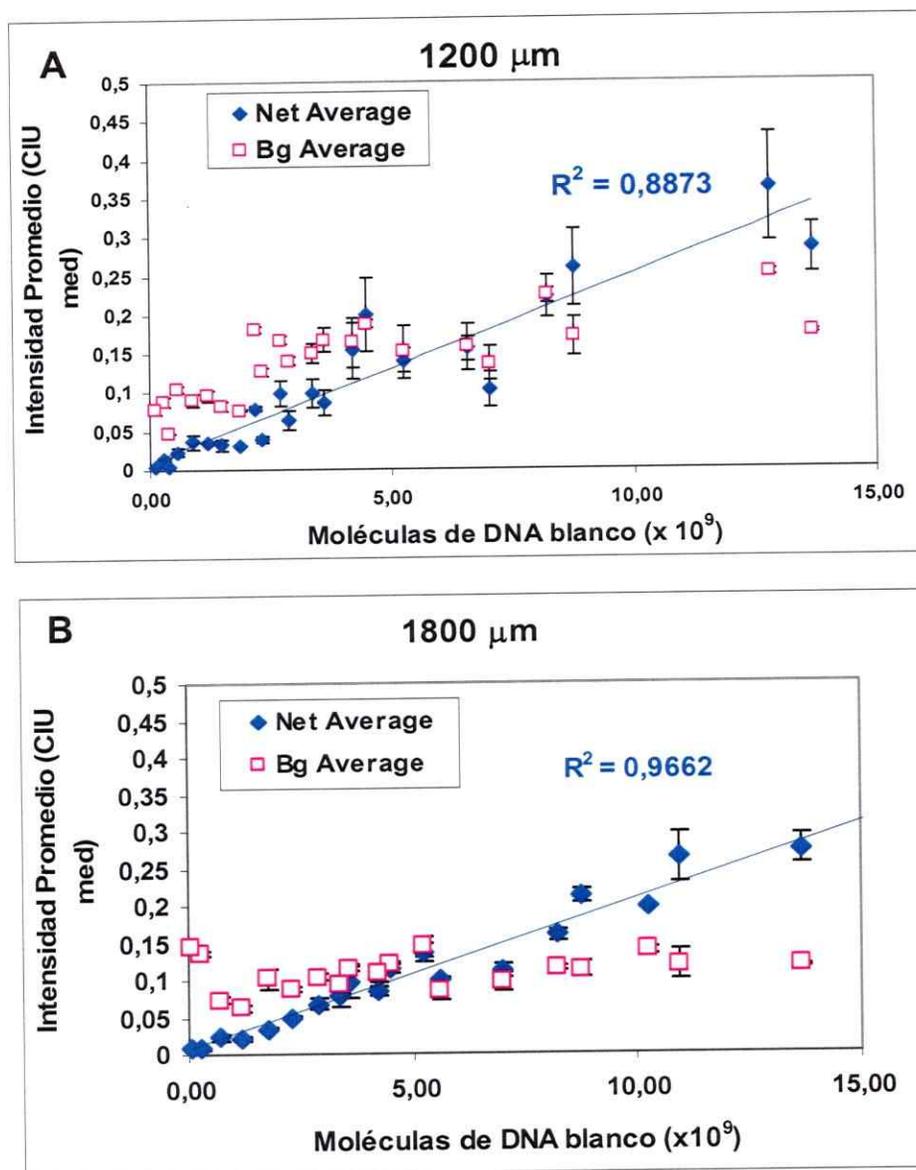


Figura 3. Cantidad óptima sembrado y distancia entre *spots*. Con el propósito de determinar la cantidad de DNA que permite obtener una intensidad neta detectable sobre la intensidad del ruido local se cuantificaron las intensidades netas y las intensidades de ruido de las señales de hibridación mostradas en la **Figura 2**. Se correlacionó la cantidad de DNA sembrada en cada *spot* con el promedio de los valores de intensidad de cada par de *spots* (replicas) y se calculó el coeficiente de correlación lineal. Panel **A**: datos obtenidos a partir de membranas en las cuales la distancia entre *spots* fue de 1200 μm ; **B**: datos obtenidos a partir de membranas en las cuales la distancia entre *spots* fue 1800 μm . En el gráfico los rombos azules (◆) indican el valor de intensidad neta y cuadrados borde fucsia (◻) muestran la intensidad del ruido local.

c) **Número de siembras por spot (*prints*):** Como se demostró anteriormente, la intensidad de señal aumenta linealmente con la cantidad de DNA blanco sembrado, de modo que aumentando la cantidad de DNA blanco, podemos aumentar nuestro límite de detección. Sin embargo, obtener muestras concentradas de DNA es costoso, de manera que una opción es aumentar la masa de DNA mediante siembras repetidas en el mismo punto. Con el fin de determinar cómo influye la siembra repetida en la variación de las intensidades de hibridación se calculó el Coeficiente de Variación (CV) entre *spots* depositados en duplicado. Mientras mayor el valor de CV, menor la calidad del dato, se ha descrito en la literatura que un valor aceptable de CV para las réplicas es 0,5 a 0,25 (Nogueira y cols., 2003; Tseng y cols., 2001). En los gráficos de la Figura 4 se comparan los valores de CV obtenidos para *spots* en los cuales la muestra se ha depositado una, tres ó cinco veces en el mismo punto de la membrana. En general, a mayor número de siembras por punto, aumenta la calidad de los duplicados (CV mayor a 0,25), es decir, disminuye la variabilidad de las señales.

1.3.- Determinación del tipo de DNA blanco a sembrar.

Con el propósito de determinar qué tipo de DNA blanco es el más adecuado para los ensayos de hibridación en membranas de Nylon, se sembraron, por duplicado, cinco diluciones de: a) DNA plasmidial conteniendo el inserto correspondiente al cDNA de expansina 1 (*exp1*) de *Prunus persica*; b) producto de PCR del cDNA de *exp1* (obtenido mediante la amplificación del

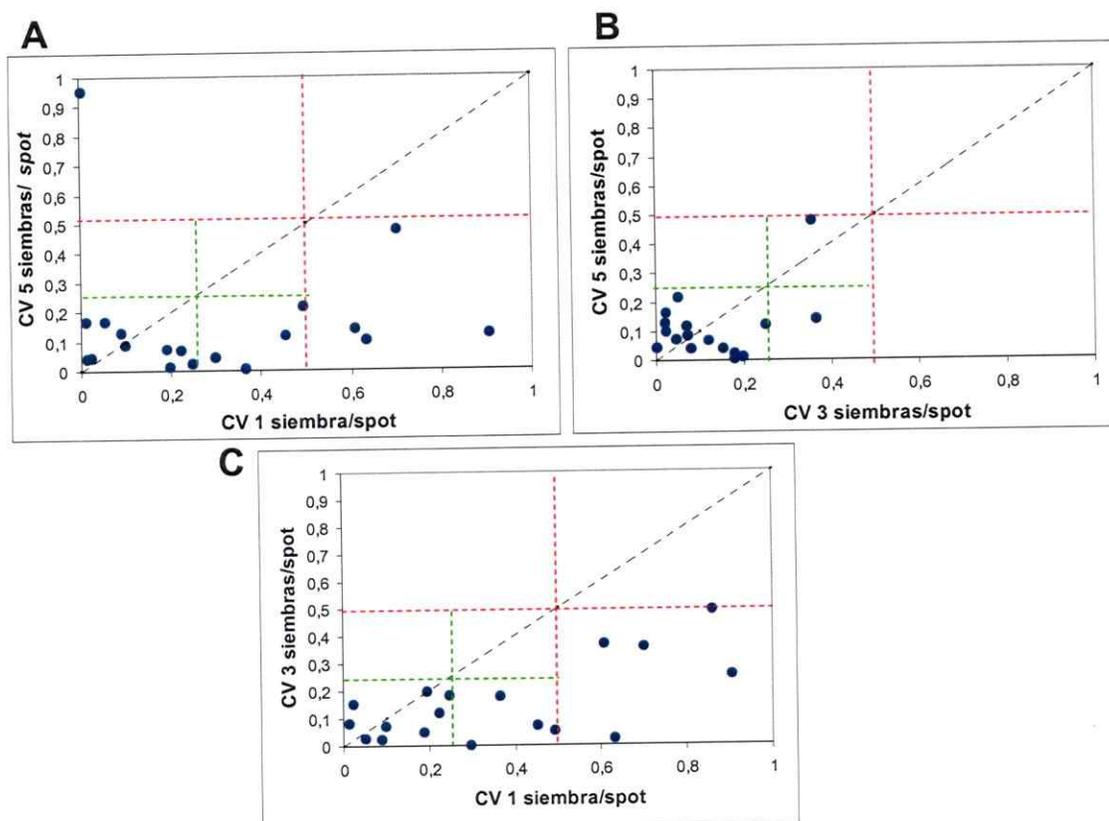


Figura 4. Variabilidad atribuible a la siembra repetida. Con el propósito de comparar la variación en la intensidad neta atribuible al número de siembras en el mismo punto, se calculó el coeficiente de variación (CV) de las intensidades netas medidas en *spots* sembrados en duplicado (**Figura 2**). Panel **A**: comparación de los CV obtenidos para 1 y 5 siembras/punto; Panel **B**: comparación de los CV obtenidos para 3 y 5 siembras/punto; Panel **C**: comparación de los CV obtenidos para 1 y 3 siembras/punto. Las líneas rojas indican un valor de CV igual a 0,5, en tanto que las líneas verdes indican un valor de CV igual a 0,25.

inserto); y c) producto purificado del PCR del cDNA de *exp1*, en tres regiones de una misma membrana. Las membranas fueron hibridadas con 200 ng de un fragmento de DNA correspondiente al transcrito de *exp1* (650 pb), marcado con $\alpha^{32}\text{dCTP}$.

Como se observa en la Figura 5, el producto purificado de PCR presenta una mayor recuperación de señal, en función del aumento en su concentración,

mientras que el PCR no purificado presenta valores intermedios. La menor recuperación de señal obtenida con el DNA plasmidial (menor pendiente) se debe a que sólo el 14,25% de la masa sembrada corresponde al inserto, en este caso, nuestro DNA blanco.

Estos resultados indican que la mayor recuperación de señal por mol de moléculas sembradas se obtiene sembrando el producto purificado de PCR. Sin embargo, en términos prácticos, la obtención de estos productos purificados es muy costosa en tiempo y recursos, por lo que decidimos sembrar el producto no purificado de PCR, que presenta una adecuada recuperación de señal.

1.4.- Reproducibilidad en la medición de la señal.

Con el propósito de evaluar la reproducibilidad de los eventos de hibridación, se utilizaron tres membranas idénticas obtenidas bajo las mismas condiciones de siembra. Estas membranas contenían 96 productos de PCR obtenidos desde una genoteca de *Drosophila melanogaster*, sembrados en duplicado. Cada una de las membranas fue hibridada con sondas complejas provenientes de RNA total de *D. melanogaster*, marcadas con $\alpha^{32}\text{dCTP}$, las que fueron sintetizadas en tres eventos independientes. Una inspección visual de las señales, indica que a tiempos iguales de exposición, los niveles de intensidad en los tres experimentos son equivalentes (Fig. 6B).

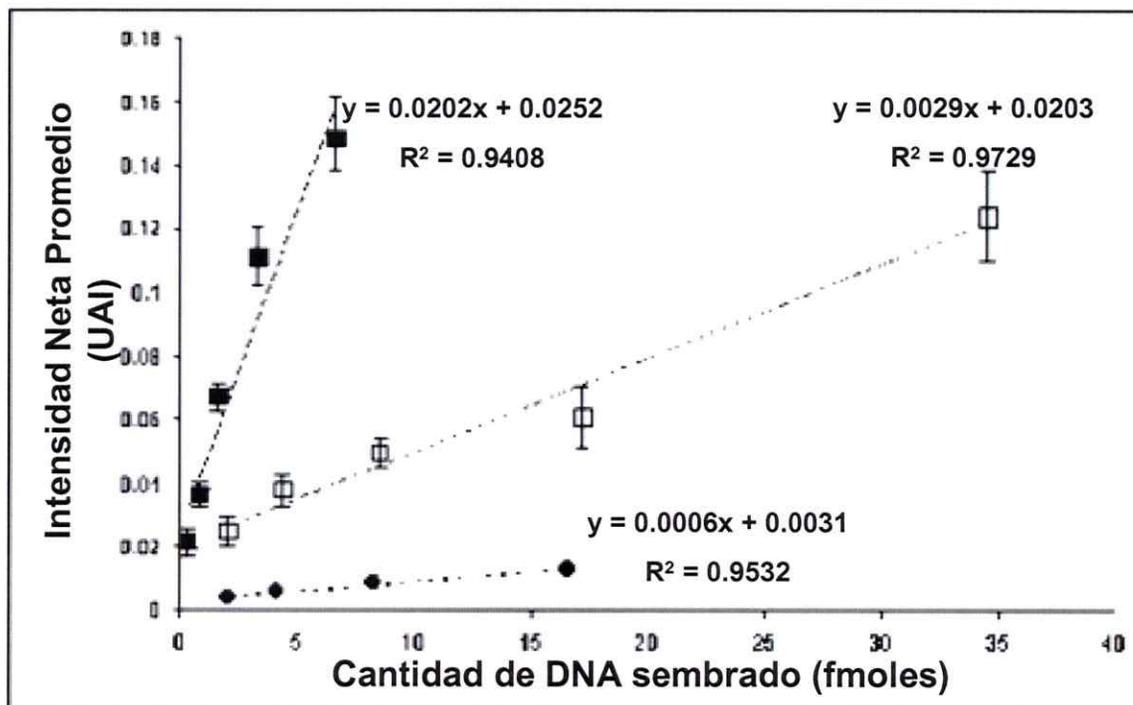


Figura 5. Determinación del tipo DNA blanco. Se sembraron en duplicado 5 diluciones de tres tipos de DNA: DNA plasmidial conteniendo un inserto de 650 pb deL gen *exp1*, producto de PCR de *exp1*, y producto purificado de PCR de *exp1* en membranas de nylon. Estas fueron hibridadas con el cDNA correspondiente a *exp1*, marcado con α^{32} dCTP. Los cuadrados negros (■) indican los valores de intensidad neta obtenidos a partir de productos purificados de PCR, en tanto que los cuadrados blancos (□) y los círculos negros (●) representan a los valores de intensidad obtenidos sembrando productos de PCR sin purificar y DNA plasmidial, respectivamente.

Esto último queda demostrado al realizar un análisis estadístico de la desviación estándar de los valores logarítmicos de intensidad neta para cada experimento ($\sigma_1=0,38$, $\sigma_2=0,40$, $\sigma_3=0,40$). En la Figura 6A, se encuentra graficada la comparación entre los valores logarítmicos de la intensidad de cada *spot* para cada experimento, se observa una distribución espacial similar en las tres comparaciones realizadas. Además, existe una alta correlación entre los

niveles de intensidad de los distintos experimentos (mediana=0,84), lo cual indica una buena reproducibilidad de los experimentos de hibridación.

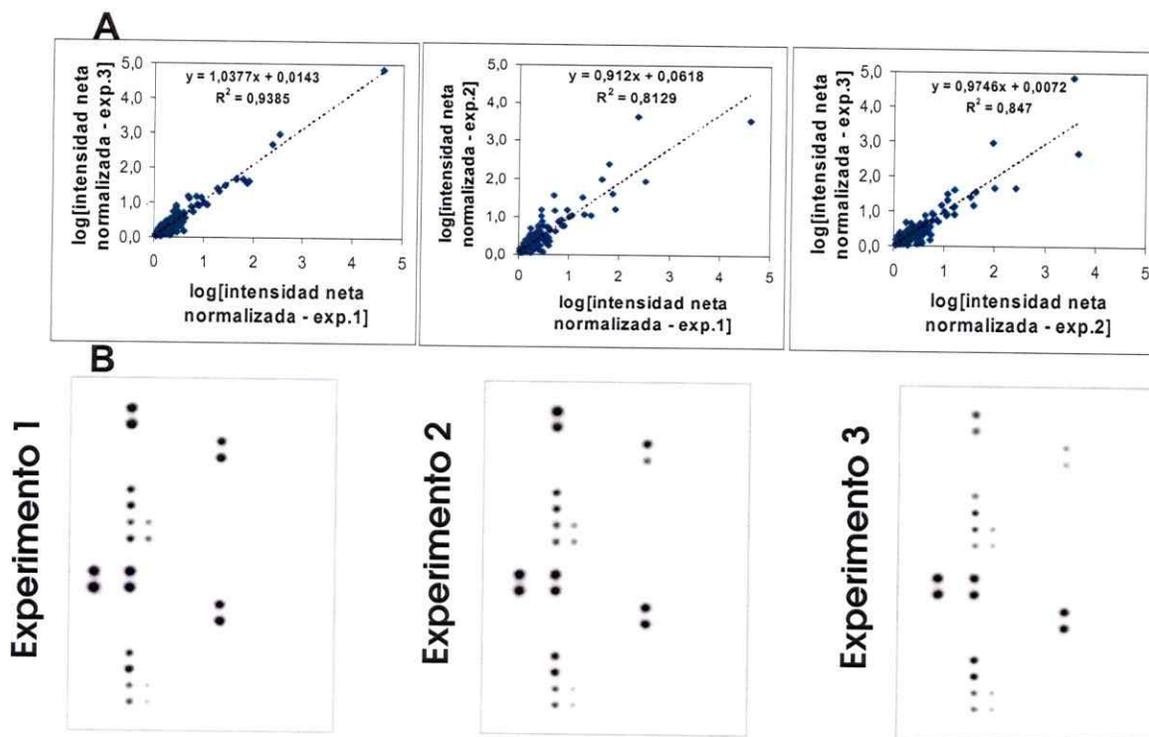


Figura 6. Reproducibilidad en los ensayos de hibridación. Productos de PCR provenientes de 96 clones de una genoteca de cDNA fueron sembrados en duplicado en tres membranas de nylon idénticas, las cuales fueron hibridadas, de forma independiente, contra una muestra compleja de cDNA marcado con α^{32} dCTP. Las imágenes de las hibridaciones se muestran en el panel **A**. Con el propósito de evaluar la reproducibilidad de los ensayos de *macroarrays*, los valores logarítmicos de intensidad neta para cada hibridación fueron comparados entre sí, determinándose las correlaciones entre los datos de experimentos independientes, las cuales se muestran en el panel **B**.

1.5.- Sensibilidad.

La sensibilidad en un experimento de hibridación en *macroarrays* de DNA es un parámetro importante, pues indica el nivel mínimo de abundancia de un transcrito que puede ser detectado. Con el objeto de determinar la sensibilidad de nuestros ensayos de hibridación, se elaboraron tres membranas en las que se sembraron en triplicado doce productos de PCR provenientes de una

genoteca de *D. melanogaster* y doce diluciones del cDNA bacteriano *dapB*, el cual no tiene homología con transcritos de eucariontes.

Por otra parte, se elaboraron tres mezclas de RNA, conteniendo 10 μg de RNA total de *D. melanogaster* y distintas cantidades del mRNA de *dapB*, que correspondían al 1%, 0,1% y 0,01% de la cantidad estimada de mRNA presente en la muestra de RNA total. Se asume que la fracción del mRNA es un 2% del RNA total, de manera que la cantidad de mRNA de *D. melanogaster* presente en 10 μg de RNA total es 200 ng. Así, las masas del mRNA bacteriano adicionadas al RNA eucarionte corresponden a 2 ng, 0,2 ng y 0,02 ng. Cada mezcla fue marcada con $\alpha^{32}\text{-dCTP}$ y luego hibridada con cada una de las membranas anteriormente descritas.

En el gráfico de la Figura 7 podemos observar que la intensidad de hibridación en los *spots* correspondientes al gen *dapB* aumenta linealmente en función de la cantidad de DNA sembrado. Este resultado nos indica que en nuestros ensayos de hibridación podemos detectar transcritos con una abundancia mínima de un 0,01% (0,02 ng del transcrito). De acuerdo con Kane y cols. (2000) existen 360.000 copias o especies de mRNA por célula eucarionte, lo que corresponde a 0,001 ng de mRNA. Si la cantidad total de mRNA usada en nuestro ensayo fue de 200 ng (cantidad mRNA que proviene de 200.000 células) y nuestra cantidad mínima de mRNA detectable es de 0,02 ng (correspondiente a 7.200.000 copias), nuestro método nos permite detectar 7.200.000 copias mRNA en 200.000 células, es decir, 36 copias mRNA/célula.

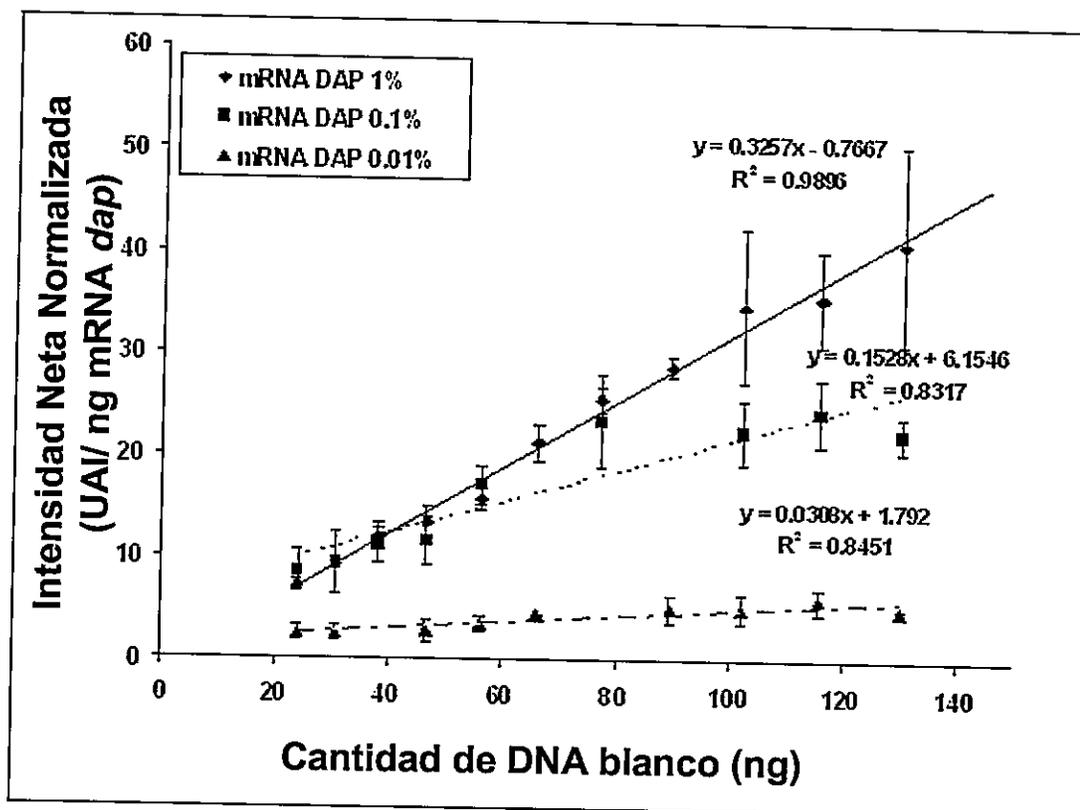


Figura 7. Sensibilidad en las hibridaciones en *macroarrays*. Para evaluar la cantidad mínima de transcritos detectable en nuestros ensayos, se llevaron a cabo 3 hibridaciones utilizando membranas conteniendo doce productos de PCR provenientes de una genoteca de *D. melanogaster* y doce diluciones del gen heterólogo *dapB*. Estas membranas fueron hibridadas por separado con tres cDNA sintetizados a partir de mezclas de RNA que contenían distintas proporciones del mRNA *dapB*. Los datos de intensidad neta fueron cuantificados y graficados contra la cantidad de DNA depositada en el *spot* correspondiente. (◆), 1% *dapB*; (■), 0,1% *dapB*; y (▲), 0,01% *dapB*.

1.6.- Conclusiones.

De este modo, con este conjunto de pruebas de optimización y eficiencia de la técnica implementada, logramos definir aspectos básicos en la manufactura de arreglos de DNA en membranas de nylon y en la ejecución de los experimentos de hibridación. En resumen, acerca de los objetivos 1 y 2 podemos decir que:

- Se definió el método de adquisición de datos:
$$\text{Int. Neta (med)} = \text{Int. Cruda (med)} - \text{Ruido local (med)}$$
- Se determinaron las siguientes condiciones de siembra:
 - ✓ la distancia entre *spots* será de 1800 μm .
 - ✓ la cantidad mínima total a sembrar por *spot* es de 6×10^9 copias de DNA blanco.
 - ✓ en cada *spot* la siembra se repetirá 5 veces al menos.
 - ✓ se sembrará DNA en forma de productos de PCR sin purificar.
- Encontramos que la reproducibilidad de método implementado es de un nivel aceptable.
- Se logró determinar la sensibilidad del método, pudiéndose detectar transcritos con una abundancia mayor a un 0.01%.

Una vez conocido el desempeño del método, estamos en condiciones de ejecutar experimentos de expresión génica en un sistema biológico de interés.

2.- Estudio de la expresión de genes relacionados con el fenómeno de harinosidad en *Prunus persica*.

Con el fin de aplicar la técnica de *macroarrays* de cDNA a un problema de interés biotecnológico, se diseñaron experimentos de hibridación para identificar genes cuya expresión se ve afectada en frutos de durazno cuando éstos son expuestos a bajas temperaturas. A continuación se describe el

proceso de selección de genes a analizar, de muestras biológicas y los resultados obtenidos de los experimentos de *macroarrays*.

2.1.- Selección de genes.

Con el propósito de centrar nuestro estudio de expresión génica en el problema de la harinosidad en frutos expuestos a bajas temperaturas, se diseñó una estrategia de selección de genes relacionados con los procesos de: metabolismo de pared celular, metabolismo de etileno, maduración y stress por frío en durazno. La selección se hizo a partir de una genoteca de ESTs generada en la Universidad de Clemson (South Carolina, EUA). Esta genoteca estaba compuesta de 16.512 clones, de los cuales 3.843 corresponden *unigenes*. Las secuencias de los 3.843 ESTs fueron alineadas contra la base de datos de CDS de *Arabidopsis thaliana* (<ftp://ftp.arabidopsis.org/home/tair/Sequences>) utilizando los algoritmos BlastN y tBlastX, de acuerdo con el procedimiento descrito en Materiales y métodos, y luego seleccionadas manualmente de acuerdo con los criterios descritos en la Tabla I. De este modo se seleccionaron 880 clones de la genoteca, según su similitud con genes de *Arabidopsis thaliana* y su posible participación en procesos biológicos particulares. El método de búsqueda utilizado se resume en el diagrama de flujo de la Figura 8.

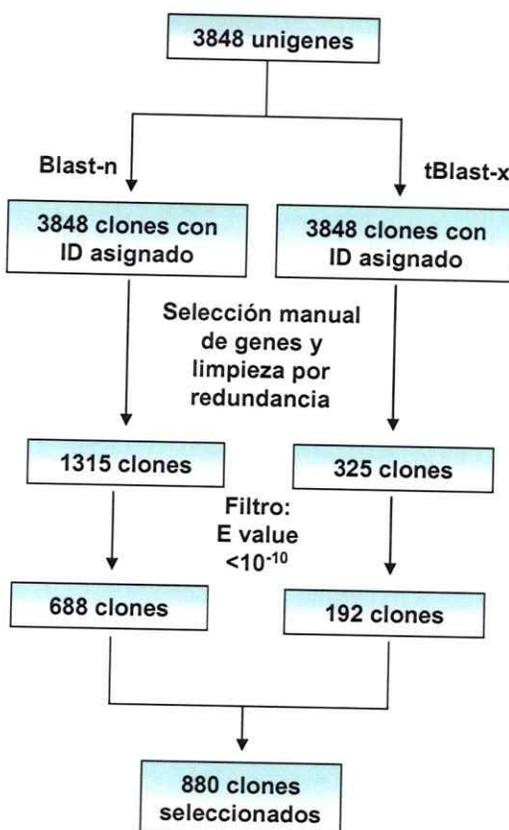


Figura 8. Proceso de selección de los genes que fueron sembrados en las membranas de *macroarrays*.

2.2.- Siembra de los clones seleccionados en membranas de nylon.

Los 880 clones seleccionados fueron crecidos en medio líquido, y alícuotas de $5 \mu\text{L}$ de cada cultivo fueron usadas en la amplificación por PCR de los respectivos insertos. Estas reacciones fueron analizadas mediante geles de agarosa y sólo aquellas reacciones que presentaran una banda única fueron seleccionadas para la siembra en membranas. Un ejemplo de estos geles se muestra en la Figura 9. Si la amplificación generaba más de una banda, el clon volvía a ser inoculado en medio líquido y la reacción de PCR era repetida. Si el clon nuevamente amplificaba más de una banda, el inserto era amplificado

desde el DNA plasmidial del clon. De este modo se logró amplificar casi la totalidad de los clones seleccionados desde la genoteca Clemson (99%).

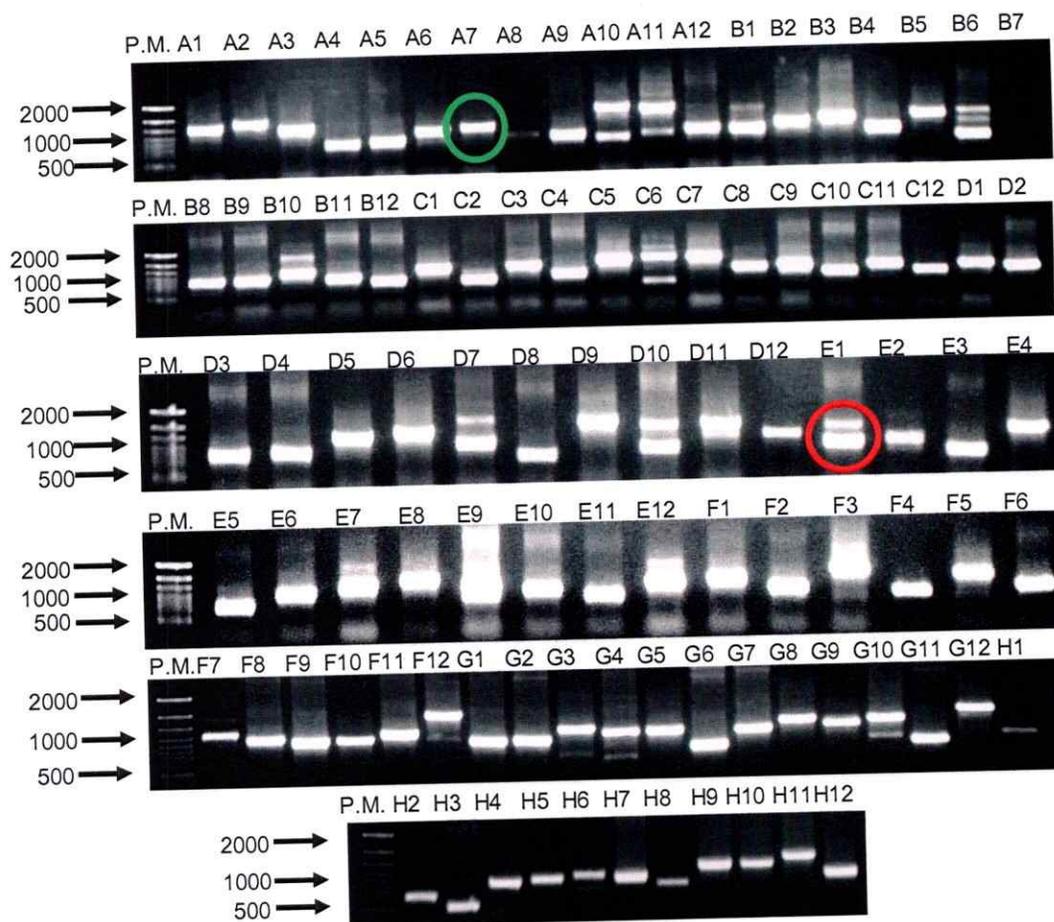


Figura 9. Verificación de la calidad de los productos de PCR seleccionados. Los clones seleccionados fueron amplificados mediante PCR a partir de cultivos bacterianos. Estos productos (5 μ L) fueron fraccionados en geles de agarosa al 1% y teñidos con bromuro de etidio para verificar la especificidad de la reacción de PCR. Aquellos productos que presentan una banda única calificaron para ser sembrados en las membranas (círculos verdes, \odot), mientras que los productos que presentaron más de una banda (círculos rojos, \circ) no calificaron, y sus reacciones fueron repetidas hasta lograr obtener una banda única. Los números sobre cada carril indican la ubicación del producto de PCR en la placa fuente. P.M.: marcador de peso molecular de 100 pb.

Los productos de PCR tanto de los genes control como los que fueron seleccionados de la genoteca Clemson fueron transferidos a 10 placas de 96 pocillos, de manera que cada placa contuviera 88 genes y 8 secuencias

controles (controles de expresión constitutiva *tub2* y *act1*, control positivo *exp1*, tres diluciones del control heterólogo *dapB*, controles negativos pGEM-t easy sin inserto y agua). Con el fin de disminuir la variación en la cantidad de DNA sembrado, cada producto de PCR fue depositado seis veces en el mismo punto. El producto de cada gen fue sembrado en un solo lugar de la membrana y los genes controles fueron depositados 10 veces en ella. De este modo se elaboraron seis membranas que contenían 960 *spots* cada una.

2.3.- Obtención de RNA para experimentos de hibridación.

Utilizamos el método de Meisel y cols. (2005), que permite extraer RNA total desde tejidos vegetales que poseen una alta concentración de polisacáridos y compuestos fenólicos. De los RNA obtenidos por este método, sólo fueron aceptados aquellos que presentaran una concentración final sobre los 70 $\mu\text{g}/\mu\text{L}$ y que tuvieran un grado de pureza aceptable de acuerdo con los criterios reportados en la bibliografía (Ojaniemi y cols., 2003).

Las muestras presentaran valores de razones de absorbancia $A_{260/280}$ entre 1,8 y 2,1 y de $A_{260/240}$ entre 1,7 y 2,0 (Tabla III). Adicionalmente se examinó la integridad de las muestras de RNA extraídas según el patrón electroforético de los rRNA 28S y 18S en geles de agarosa, considerándose que la integridad de la muestra era aceptable cuando la banda correspondiente al rRNA 28S era al menos 2 veces mas intensa que la banda del rRNA 18S. Los patrones electroforéticos de los RNA total utilizados en las hibridaciones se muestran en la Figura 10.

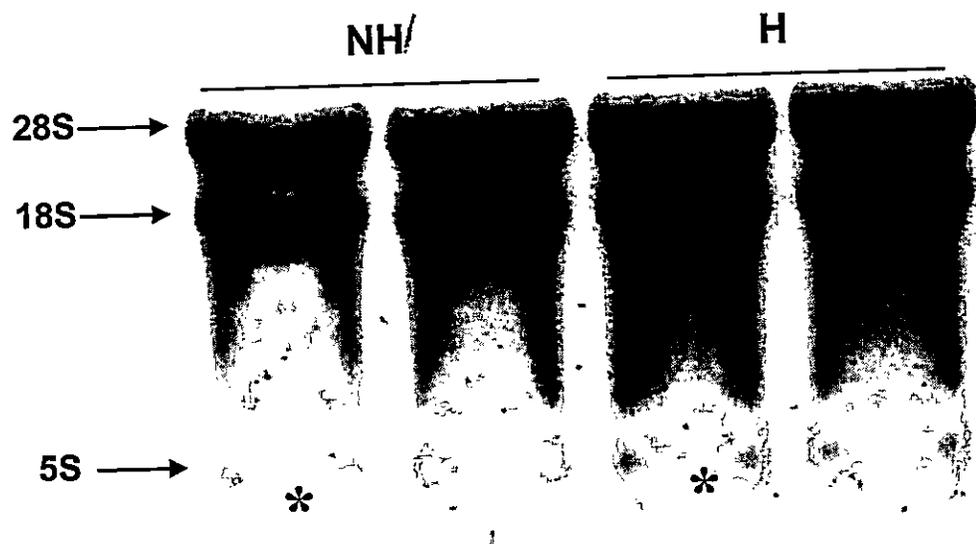


Figura 10. Verificación de la integridad de los RNAs extraídos. Cuatro muestras de duraznos O'Henry, dos correspondientes a un fruto NH y dos correspondientes a un fruto H fueron sometidos a extracciones de RNA total por el método descrito en Meisel y cols. (2005). Estos extractos fueron resueltos en geles denaturantes de agarosa al 1% y teñidos con bromuro de etidio para visualizar su integridad. Las flechas indican las bandas correspondientes al rRNA ribosomal 28S, 16S y 5S. Un asterisco (*) señala a aquellos extractos utilizados en los posteriores ensayos de hibridación.

Tabla III.- Concentración y calidad de los RNA totales obtenidos en dos extracciones independientes de RNA, para las muestras de duraznos no harinoso (NH) y harinoso (H). Un asterisco (*) señala a aquellos extractos utilizados en los posteriores ensayos de hibridación.

Muestra	[RNA] (mg/mL)	A 260/240	A 260/280
NH*	2,48	2,05	1,82
NH	4	1,94	1,75
H*	1,2	1,69	2,1
H	6,5	2,1	1,75

El mRNA poli-(dA)⁺ fue purificado a partir de las muestras de RNA total usando una resina comercial que funciona en base a la afinidad de oligonucleótidos dC₁₀dT₃₀ por las colas de poli-(dA) presentes en los mRNA. El motivo de extraer RNA poli-(dA)⁺ desde RNA total es aumentar la proporción

estadística de transcritos de mRNA en la muestra. Para cerciorarnos de que no existía degradación del mRNA, se sintetizó cDNA desde cada una de las muestras de mRNA, y mediante PCR se amplificó un grupo de genes presentes en los *macroarrays* y que deberían estar expresados en los duraznos estudiados (ver Materiales y Métodos, Tabla II). Si la amplificación de uno o más genes fallaba en alguna muestra, entonces ésta no era utilizada para la síntesis de la sonda. Los resultados de las reacciones de PCR se muestran en la Figura 11, donde se puede observar que todas las bandas obtenidas son del tamaño esperado, lo que nos indica a su vez que al menos los cuatro transcritos analizados estarían íntegros en los mRNA purificados.

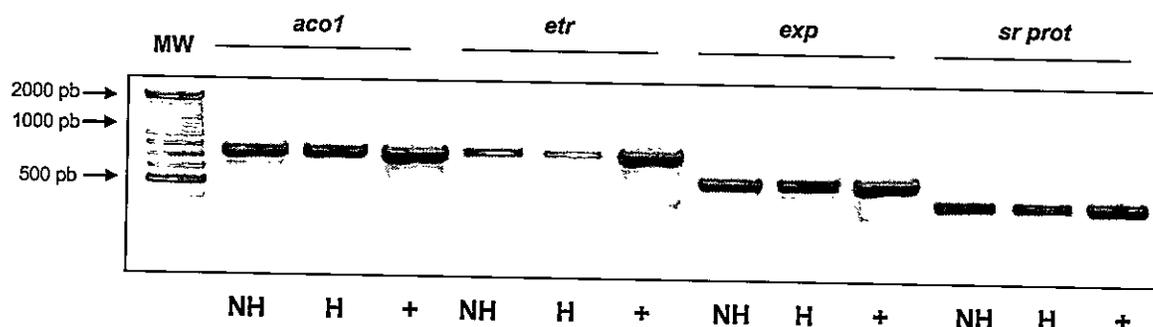


Figura 11. Verificación de la calidad de los mRNA. cDNA de hebra simple fue sintetizado a partir de los mRNA de frutos NH y H y utilizado como sustrato en reacciones de PCR con partidores de específicos para cuatro genes de *P. persica*. *aco1*: 1-aminociclopropano -1-carboxilato oxidasa, AF532976; *etr*: receptor de etileno, AF396830; *exp*: Expansina 1, AB029083; *SRprot*: proteína de stress por ácido abscísico, AF317032). MW: estándar de peso molecular 200 pb. Los carriles indican la muestra utilizada como sustrato para la reacción por PCR: NH, cDNA de durazno no harinoso; H, cDNA de durazno harinoso; Carril +, control positivo de DNA plasmidial cuyo inserto corresponde a cada gen analizado.

2.4.- Ensayos de hibridación.

Estos experimentos tienen como objetivo comprobar que el método de confección de arreglos de DNA es útil para detectar cambios significativos en la

expresión de genes involucrados en un proceso biológico particular. En nuestro caso el proceso a analizar es el fenómeno de la harinosidad, el cuál -como se mencionó anteriormente- se genera por una exposición de los frutos de durazno a bajas temperaturas por un tiempo prolongado (15-30 días). Para ello se obtuvieron muestras de mRNA a partir de un fruto que no fue expuesto a frío (fruto maduro y no harinoso; **NH**) y de un fruto expuesto a frío (fruto maduro y harinoso; **H**), las que sirvieron como templado en dos eventos de síntesis de cDNA de hebra simple. En el primer evento se marcaron por separado 70 ng de mRNAs extraídos de frutos NH y H, a los que se adicionó 0,2 ng de mRNA *dapB*, sintetizado *in vitro*. De manera similar en el segundo evento se marcó por separado 150 ng del mRNA de frutos NH y H. El mRNA externo *dapB* sirvió como mRNA de referencia ya que su proporción en la muestra a marcar es conocida y constante.

Las sondas sintetizadas a partir de 70 y 150 ng de mRNA de frutos NH y H fueron hibridadas, en eventos independientes, con tres membranas idénticas, obteniéndose doce imágenes. En la figura 12 se muestran dos membranas representativas, una hibridada con la muestra de durazno NH y la otra con la sonda proveniente de frutos H. En esta figura podemos distinguir algunos puntos de hibridación cuyas intensidades cambian entre las hibridaciones. Se señalan, además, algunos de los controles sembrados en la membrana, incluyendo al control interno *dapB* y controles negativos como agua y producto de PCR de vector sin inserto (llamado MCS).

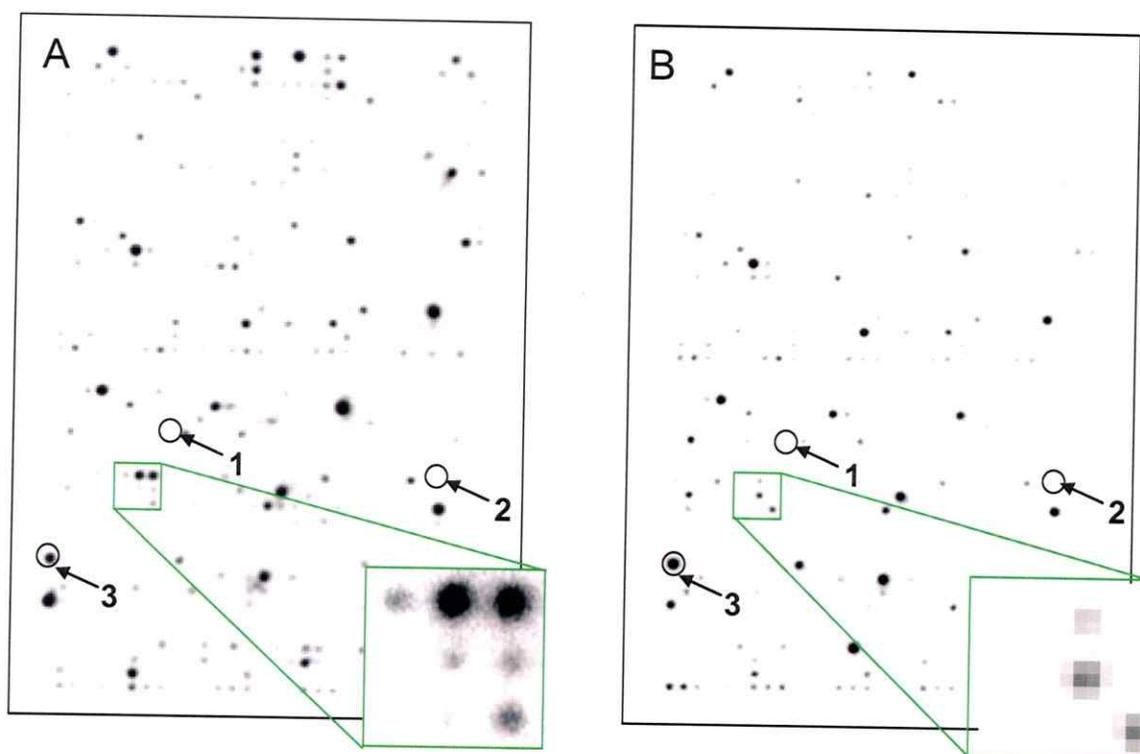


Figura 12. Hibridación en membranas de *macroarrays*. En los ensayos de hibridación, 70 o 150 ng de mRNA de muestras no harinosas y harinosas fueron utilizados para sintetizar sondas complejas de cDNA marcadas con $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$, las que fueron hibridadas con membranas idénticas. El **panel A** corresponde a una imagen representativa obtenida con el programa Quantity One (BioRad) luego de la hibridación con la sonda proveniente de una muestra de fruto **NH**. El **panel B** corresponde a la hibridación con la sonda de un fruto **H**. Los recuadros dentro de las imágenes, que se encuentran aumentados en la parte inferior muestran 9 puntos de siembra que presentan claras diferencias de intensidad entre las muestras no harinosas y harinosas. La hibridación no específica fue examinada utilizando puntos que representan al vector vacío pBluescript II SK (-) (flechas en círculo 1) y agua (flechas en círculo 2). Las flechas en el círculo 3 representan la señal para el cDNA del gen *dap* que hibrida con el *spike* RNA.

2.5.- Adquisición y análisis de datos.

Luego de cuantificar las intensidades de los puntos de hibridación para cada una de las doce imágenes obtenidas, se llevó a cabo un análisis de los datos obtenidos con el objetivo de descartar datos poco confiables. Este análisis comprendió cuatro etapas: 1) Comparación entre réplicas; 2) Selección

de datos de buena calidad; 3) Normalización; y 3) Determinación de cambios significativos en la expresión de genes.

2.5.1.- Comparación entre réplicas.

Con el propósito de examinar la distribución de las señales en cada una de las membranas analizadas, se confeccionaron histogramas de valores de intensidad neta para cuatro grupos de imágenes, estos correspondían a los triplicados de las membranas que fueron hibridadas con: sondas generadas a partir de 70 ng de mRNA de la muestra de durazno NH, 150 ng de mRNA de la muestra de durazno NH, 70 ng de mRNA de la muestra de durazno H, y 150 ng de mRNA de la muestra de durazno H. En cada grupo de hibridaciones el conjunto total de señales se comporta de manera similar en las tres membranas (Fig. 13). Para las hibridaciones con muestras provenientes de frutos NH, las señales obtenidas usando 70 ng de mRNA (Fig.13a) muestran una distribución casi idéntica al conjunto de las señales provenientes de las hibridaciones con 150 ng de mRNA (Fig.13b), observándose para ambas que el mayor número de puntos de hibridación posee un nivel de intensidad de alrededor de 0,06 unidades arbitrarias de intensidad. Si bien ambos grupos de membranas hibridadas con mRNA de la muestra de durazno H se comportan de forma similar (Figs. 13c y 13d), la mayoría de las señales se agrupa alrededor de 0 unidades arbitrarias de intensidad, sugiriendo que en la condición harinosa puede existir un menor nivel en la expresión de genes. Podemos concluir que las distintas muestras de mRNA mostraron diferentes distribuciones de señales

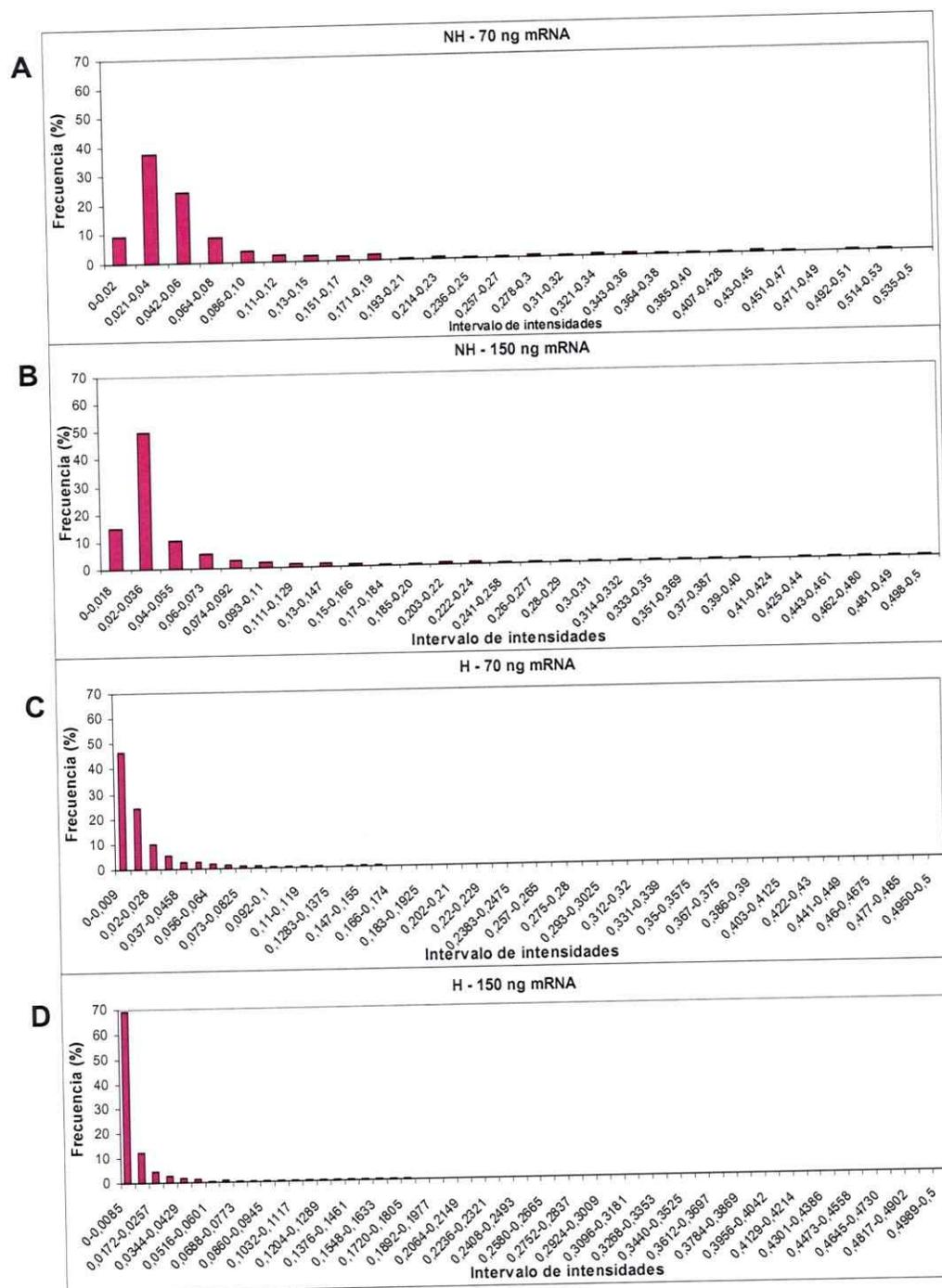


Figura 13. Comportamiento de las señales de hibridación en *macroarrays*. Con el propósito de evaluar la distribución de las intensidades de señal obtenidas en los experimentos de hibridación, los datos obtenidos fueron agrupados en distintos rangos de intensidad, para luego generar histogramas utilizando el programa SYSTAT 2.0. Panel **A**, Réplicas O2A_5, usando 70 ng de mRNA; Panel **B**, Réplicas O2A_5, usando 150 ng de mRNA; Panel **C**, Réplicas O4Y_8, usando 70 ng de mRNA; Panel **D**, Réplicas O4Y_8, usando 150 ng de mRNA.

y que además el comportamiento de las intensidades en las membranas hibridadas con distintas cantidades de sonda se mantuvo uniforme. Si bien las réplicas tuvieron un comportamiento adecuado en cuanto a la distribución de sus señales, debemos comprobar si cada punto tiene un valor comparable en las distintas réplicas. Para saber si los datos dentro de cada membrana podrían ser comparados con los datos de las demás membranas, se calculó el coeficiente de correlación de Pearson entre los 960 datos de cada una de las membranas y las restantes cinco que habían sido hibridadas con sondas provenientes de un mismo fruto (Tabla IV).

Tabla IV. Coeficientes de correlación de Pearson para las réplicas de hibridaciones. Estas se obtuvieron comparando cada una de las 6 membranas y sus réplicas. Las membranas fueron hibridadas con sondas provenientes de frutos NH (A) o H (B).

A	M1 70 ng	M2 70 ng	M3 70 ng	M1 150 ng	M2 150 ng	M3 150 ng
M1 70 ng	1					
M2 70 ng	0,857	1				
M3 70 ng	0,693	0,791	1			
M1150 ng	0,788	0,79	0,715	1		
M2 150 ng	0,637	0,758	0,695	0,671	1	
M3 150 ng	0,709	0,806	0,726	0,773	0,748	1
B	M1 70 ng	M2 70 ng	M3 70 ng	M1 150 ng	M2 150 ng	M3 150 ng
M1 70ng	1					
M2 70 ng	0,551	1				
M3 70 ng	0,624	0,651	1			
M1150 ng	0,486	0,692	0,655	1		
M2 150 ng	0,684	0,505	0,498	0,505	1	
M3 150 ng	0,8	0,641	0,641	0,719	0,753	1

Existe una mejor correlación entre los datos obtenidos de las hibridaciones con sondas provenientes del mRNA de la muestra de durazno

NH, ya que el valor mínimo de correlación es de 0,637 (Membrana 5 vs. Membrana 1). Para el caso de las membranas hibridadas con sondas provenientes del mRNA de la muestra de durazno H existe un valor mínimo de correlación de 0,486 (Membrana 4 vs. Membrana 1). Sin embargo, en la mayor parte de las comparaciones se obtiene un valor de correlación mayor a 0,5, lo que nos indica que el valor de intensidad para un *spot* cualquiera en una membrana es comparable con el mismo *spot* en las demás membranas. Adicionalmente se calculó el coeficiente de corrección de Bonferroni para cada comparación (n=960). La corrección de Bonferroni es un ajuste estadístico para múltiples comparaciones usada cuando se hacen varios cálculos estadísticos simultáneamente y se utiliza para evitar falsos positivos en análisis de una gran cantidad de datos. En este estudio se obtuvo para todas las comparaciones un valor cercano a cero, lo cual nos indica que las similitudes entre las señales de las distintas réplicas, reflejadas en los coeficientes de Pearson, no se deben al azar.

En conclusión, en base a los resultados mostrados anteriormente, podemos afirmar que es posible considerar a cada conjunto de membranas hibridadas con el mismo RNA como réplicas válidamente comparables.

2.5.2.- Selección de datos de buena calidad.

El primer filtro empleado para seleccionar los datos fue descartar aquellos cuyo valor de intensidad neta fuera igual o menor que cero. Así, estamos desechando en cada réplica a todos los genes cuya intensidad de

hibridación estuvo bajo el nivel de lectura detectable por el equipo. Luego de aplicar este filtro se recuperaron 942 datos para las hibridaciones utilizando sonda sintetizada a partir de la muestra de durazno NH y 925 para las hibridaciones utilizando una sonda sintetizada a partir de la muestra de durazno H, lo cual representa un 98% y un 96% de los datos iniciales, respectivamente (Tabla V).

El segundo filtro está destinado a descartar a aquellos datos con intensidad neta positiva que presentan un nivel cercano al ruido local, y que por lo tanto son poco confiables. Previo a ello, debemos determinar el nivel mínimo de intensidad neta que debe presentar una señal de hibridación para ser aceptada como confiable (punto de corte). Para ello se midió la cantidad de datos recuperados en cada hibridación después de variar el valor del punto de corte entre 0 y 2 veces la desviación estándar del ruido local. Se observó que un 55% y un 30% de los datos obtenidos a partir de hibridaciones con muestras de duraznos NH y H respectivamente, presentaron valores de intensidad superiores a la desviación estándar del ruido local. Los porcentajes de datos recuperados usando este filtro se indican en la Tabla V.

Para el tercer y último filtro de datos se aplicó un protocolo descrito por Wang y cols., 2001, el cual asigna un valor de calidad compuesta (Q_{com}) que depende del tamaño del punto, la relación señal-ruido, la variabilidad del ruido local y del nivel de ruido local. De este modo, este tercer filtro tiene como objetivo descartar aquellos datos de mala calidad presentes en nuestros resultados. En nuestro análisis se descartaron todos aquellos datos cuyo valor

de Q_{com} fuera menor a 0,8. De este modo sólo consideramos el 78% y el 58%, respectivamente, de los datos inicialmente obtenidos con las muestras de durazno NH y H ($n=960$, Tabla V).

Tabla V. Filtros aplicados a los datos obtenidos de las hibridaciones con sondas sintetizadas a partir de muestras de duraznos NH y H y su impacto el número de datos recuperados.

FILTROS	Muestra de durazno NH		Muestra de durazno H	
	Nº de datos filtrados	% datos recuperados	Nº de datos filtrados	% datos recuperados
Int. Neta > 0	942	98,1	925	96,4
Int. Neta >1SD	822	85,6	571	59,5
$Q_{com} > 0.8$	756	78,8	560	58,3
$CV < 0.5$	509	53	315	32,81

2.5.3.- Normalización de los datos de intensidad.

El proceso de normalización de los datos de expresión génica considera las diferencias globales de intensidades entre las membranas a comparar, las cuales pueden originarse por diferencias en la cantidad de RNA marcado, en la eficiencia de la marca, en la siembra del DNA, en la eficiencia de la hibridación, etc. Durante este proceso los valores de intensidad de los puntos dentro de las distintas membranas son ajustados para permitir su comparación. Con el objetivo de normalizar los valores de intensidad, primero se normalizaron las intensidades netas de acuerdo a la cantidad de DNA sembrado en cada punto, y luego se normalizó el valor obtenido por el valor de intensidad del gen *dapB*, el cuál fue usado como estándar externo, adicionado en cantidad conocida y equivalente en todas las muestras hibridadas. El primer ajuste se hizo para

equilibrar las diferencias de cantidades de DNA blanco presente en el *spot*, ya que es sabido que la intensidad de un punto varía linealmente con la cantidad de DNA sembrado, y que la concentración de los productos de PCR sembrados no era la misma. El segundo ajuste se efectuó para equilibrar las intensidades según las diferencias en las distintas preparaciones de mRNA, en las síntesis de sonda, o en las eficiencias de hibridación en experimentos independientes. Adicionalmente, dado que la intensidad varía linealmente con la cantidad de sonda, si conocemos la abundancia e intensidad neta de nuestro gen referencia (*dapB*) podemos estimar el nivel de abundancia de los transcritos de intensidad neta conocida (Bernard y cols., 1996; Nguyen y cols., 1995). Así, además de llevar a cabo la normalización de nuestros datos, se hizo un cálculo estimativo de la expresión absoluta de los genes que pasaron los filtros descritos en la sección anterior. Esta expresión absoluta se tradujo al porcentaje de abundancia y el número de copias de transcrito detectadas en la muestra.

Luego de tener los valores absolutos de expresión, y con el fin de considerar sólo aquellos datos que mostraron una buena reproducibilidad en las réplicas efectuadas, se descartaron todos aquellos datos que como conjunto de réplicas mostraran un valor de CV menor a 0,5. De este modo se recuperaron 509 genes para la muestra de durazno NH y 315 genes para la muestra de durazno H (53% y 32% de los 960 datos iniciales, respectivamente), de ellos, 262 fueron comunes a ambas frutas (27%).

El valor mínimo de abundancia detectado en este ensayo correspondió a un gen que codifica para una proteína desconocida con un dominio de anclaje

GPI (At3g02210.1), que presentó un nivel abundancia de 21 copias/célula, 15 copias por célula menos que el nivel determinado durante la medición de sensibilidad del método.

2.5.4.- Determinación de cambios significativos en la expresión de genes.

En un análisis de expresión génica en que se comparan dos condiciones, un paso clave es la identificación de genes que se expresan diferencialmente entre frutos NH y H. Para determinar aquellos genes que cambiaron significativamente su expresión se realizó un *test* de significancia para dos variables (*test* t de student), utilizando los datos de expresión recuperados para los dos frutos. Se aceptaron como genes diferencialmente expresados aquellos que presentaran un valor de $p < 0,05$, es decir, que las diferencias observadas tuvieran una probabilidad de error inferior a un 5%. Los genes que mostraron una diferencia estadísticamente significativa y sus respectivos valores de cambio se muestran en la Tabla VI. Como se puede observar, sólo 3 genes se encuentran inducidos en los frutos harinosos y corresponden al gen de la pectilmetilesterasa (*pme*, Pp_05K11), el gen *gad2* (Pp_03I05), y el gen *grf4* (Pp_02A14). Cabe destacar que *pme* tuvo un aumento de 5,4 veces, lo que sugiere que este gen puede estar involucrado en el fenómeno de harinosidad.

Por otra parte, dentro de los 34 genes reprimidos se encuentran genes asociados con síntesis de componentes de la pared celular (Pp_10L02, Pp_28D07, Pp_03F22 y Pp_30I06), reguladoras de la homeostasis de calcio (Pp_29H01, Pp_37P20 y Pp_26H15), destinación de proteínas a la pared

celular (Pp_14G04, Pp_09B20 y Pp_33A18), estrés térmico (Pp_11M19, Pp_27I04 y Pp_09B20). Además, encontramos proteínas inducidas por hormonas vegetales como auxinas y ácido abscísico (Pp_25G17, Pp_17A10, Pp_09B20 y Pp_07A24), y proteínas involucradas en la síntesis de flavonoides (Pp_19B20).

Tabla VI. Genes cuyo cambio de expresión fue significativo ($p < 0,05$) entre frutos NH y H. Se indica en la primera fila: el ID del clon sembrado proveniente de la genoteca Clemson, el código Genbank de la secuencia que presentó una mayor similitud con el clon seleccionado, el *e value* para el mejor alineamiento (Blast-n o tBlastX) entre el clon Clemson y la base de datos Genbank, la función putativa de cada producto génico basada en similitudes de secuencias, una breve descripción de la función de estos productos génicos, las veces de cambio en la expresión de estos genes desde una condición NH a una condición H (los valores de cambio en la expresión para los genes reprimidos han sido multiplicados por -1), y por último las referencias bibliográficas en las que se describe la función de cada uno de los productos génicos señalados. Un asterisco (*) señala aquellas secuencias que hicieron match con proteínas de función desconocida, o que comparten algunos dominios proteicos conocidos, los cuales fueron encontrados en la base de datos InterPro.

Clon ID	Código Genbank	e value	Función Putativa	Descripción	Veces de cambio	Referencias
Pp_05K11	NM_124531	1,1E-40	Pectinsterasa	Importante rol en la pared celular durante la maduración del fruto. Desmetila el polímero de pectina a pectato y metanol.	5.4	Crisosto y cols., 1999
Pp_02A14	NM_103213	1,2E-38	Proteína GF14 Φ 14-3-3 (grf4)	Familia de proteínas que comparten la función de completar la activación de proteínas fosforiladas y que participan en diversos procesos de transducción de señales.	2.1	Sehnke y cols., 2002; Wu y cols., 1997
Pp_03I05	NM_105268	1,6E-26	Glutamato decarboxilasa 2 (GAD 2)	Enzima inducida por un aumento en la concentración intracelular de Ca ²⁺ , catalizando la formación de GABA.	1.8	Turano y Fang, 1998.
Pp_19B20	NM_127791	1,0E-111	Glutaredoxina putativa	Oxidoreductasas ubicuas que participan en la homeostasis redox de la célula, reduciendo enlaces disulfuro glutatión-proteína.	-1.8	Rouhier y cols., 2005.
Pp_27O06	NP_568535	1,4E-11	Proteína de transmembrana	Análisis de secuencia aminoacídica revela la existencia de dominios de transmembrana, péptido señal de destinación a membrana plasmática y tres dominios no caracterizados.	-1.8	Koch, 1990
Pp_01E24	NM_111375	1,3E-74	Bomba Na ⁺ /H ⁺ putativa	Intercambiador Na ⁺ /H ⁺ de membrana plasmática, encargado de mantener la homeostasis de sodio en la célula. Se ha descrito un aumento en su actividad en presencia de altas concentraciones extracelulares de sodio.	-1.9	Qiu y cols., 2003

Clon ID	Código Genbank	e value	ID Putativo	Descripción	Veces de cambio	Referencias
Pp_29H01	NM_100791	1,1E-88	Calreticulina 2 (CRT2)	Proteína cuya inducción produce un aumento en la concentración de Ca+2 en el retículo endoplásmico.	-1.9	Persson y cols, 2001.
Pp_11M19	NM_111656	7,3E-12	Similar a la proteína de choque térmico HSP90.	Secuencia que presenta una alta similitud con HSP's. Descrita en 2 perfiles Pfam: PF02518, ATPasa, histidina kinasa, DNA gyrase B, y dominio proteico similar a HSP90; PF00183: proteína Hsp90.	-2.0	*
Pp_24P19	NM_121228	1,2E-19	Proteína expresada	Análisis de dominios proteicos (IPR001950) indican la existencia de dominios que le asocian con familias de alta similitud con el factor de iniciación eIF-1 y con la familia del factor de elongación eIF-1.	-2.1	*
Pp_10L02	NM_126246	1,3E-46	Xiloglucano endo-transglucosilasa	Enzima clave en la regulación de la expansión de la pared celular, y se cree que es la enzima responsable de la integración de xiloglucano a la matriz de la pared celular.	-2.1	Bourquin y cols., 2002; Johanson y cols., 2004
Pp_30I13	NM_101278	1,5E-67	Dioxigenasa dependiente de 2-oxoglutarato	Enzima que cataliza la reacción de hidroxilación del terpenoide indólico taberosina, dentro de la vía de síntesis del alcaloide vindolina.	-2.2	Turnbull y cols., 2004; Biemelt y cols., 2004.
Pp_20I08	NM_120224	1,5E-85	Proteína expresada	Contiene un dominio similar a la proteína integral de membrana LMBR-1, de función desconocida.	-2.3	*
Pp_28D07	NM_124561	1,7E-53	Fosfo-glucmutasa	Cataliza la interconversión de la Glucosa 6-P a la Glucosa 1-P, dentro de la vía de síntesis de la pared celular.	-2.3	Brown, 1953
Pp_27I04	NM_123613	1,3E-12	Similar a la familia DnaJ	Proteína inducida por choque térmico que interactúa con la Hsp70	-2.3	*
Pp_33G20	NM_119381	1,5E-11	Proteína expresada	Contiene un dominio Interpro IPR001026 de Epsina en su extremo N terminal.	-2.3	*
Pp_37P20	NM_120569	1,5E-82	Proteína kinasa dependiente de calcio (CDPK)	Kinasa activada por Ca+2 que aumenta su expresión en respuesta a estímulos hormonales y ambientales.	-2.3	Mitra y Johri, 2000.

Clon ID	Código Genbank	e value	ID Putativo	Descripción	Veces de cambio	Referencias
Pp_21J19	NM_113506	1,2E-32	Subunidad regulatoria B de la proteína fosfatasa	Sub-unidad reguladora de proteínas fosfatasa Serina/Treonina, cuyos transcritos se acumulan en respuesta a estrés por temperaturas, sugiriendo que esta estarían involucradas en los mecanismos de respuesta a estrés en plantas.	-2.3	Terol y cols., 2002.
Pp_25G17	NM_106112	6,8E-12	Proteína inducida por ácido abscísico (HVA22)	Gen inducido en distintos tipos de estrés ambiental, como deshidratación, salinidad y temperaturas extremas. Se ha demostrado que este gen regula el tráfico vesicular en células expuestas a estrés, para facilitar el recambio de la membrana plasmática, o disminuir la secreción innecesaria.	-2.4	Brands y Ho, 2002.
Pp_30C13	NM_118911	1,1E-48	Proteína expresada	Posee dominio de la familia DUF719 (IPR008509), de función desconocida	-2.4	*
Pp_03F22	NM_114187	1,5E-67	Sucrosa sintasa	Enzima clave en las vías de síntesis de azúcares que corresponden a monómeros del almidón y hexosas que conforman la pared celular. Cataliza la conversión reversible de sucrosa en fructosa y UDP-glucosa.	-2.4	Chopra y cols., 1992.
Pp_17A10	NP_171949	1,3E-33	Proteína 12 inducida por ácido indolacético (IAA12) expressed protein	Proteína expresada en las fases tempranas de la exposición a ácido indolacético (IAA). Miembro de la familia de reguladores transcripcionales AUX/IAA.	-2.4	Abel y cols., 1995.
Pp_02D19	NM_121870	1,3E-40		Contiene un dominio RING finger (IPR001841, IPR008913). Muchas proteínas que contienen este dominio juegan un rol central en la ubiquitinación dependiente de enzimas E2 conjugantes de ubiquitina. Se ha demostrado que este dominio es responsable en algunas interacciones proteína-proteína. Además presenta el dominio DUF917.	-2.4	Andersen y cols., 2004.
Pp_30I06	NM_112675	1,3E-18	aldose 1-epimerasa	Enzima responsable de la interconversión de D-Glucosa y otras aldosas en sus formas a y b.	-2.6	Heese-Peck y Raikhel, 1998.

Clon ID	Código Genbank	e value	ID Putativo	Descripción	Veces de cambio	Referencias
Pp_14G04	NM_119206	1,2E-16	Proteína SRP	Contiene alta similitud con la sub-unidad alfa de partículas reconocedoras de señal (SRP).	-2.6	Schunemann, D., 2004.
Pp_26K18	NM_101612	1,1E-95	Hidrolasa similar a haloácido dehalogenasa	Grupo de diversas proteínas entre las cuales están la L-2-haloácido dehalogenasa, epóxido hidrolasas y fosfatasa.	-2.7	*
Pp_26H15	NM_123137	1,6E-24	Calmodulina 4 (CAM1/4)	Calmodulina activada por Ca ²⁺ que juega un papel crucial en la respuesta al choque térmico	-3.0	Liu y cols., 2004
Pp_33A18	NM_112368	1,6E-60	Proteína de unión a GTP perteneciente a la familia Ras	Superfamilia que incluye proteínas involucradas en diversos procesos celulares como proliferación celular, ensamble y organización del citoesqueleto y tráfico de membranas intracelular.	-3.4	Vernoud, V. y cols., 2003.
Pp_09B20	NM_112093	1,1E-91	Proteína de choque térmico hsp70	Familia de chaperonas moleculares que comparten alta homología, tienen un peso molecular de aproximadamente 70KDa, actividad ATPasa, y tienen la capacidad de unir y liberar dominios hidrofóbicos en proteínas no plegadas.	-3.5	Anderson y cols., 1994
Pp_29D07	NM_116121	1,5E-23	Proteína de unión a GTP	Proteína similar a SAR1, que corresponden a proteínas GTPasas involucradas en el transporte vesicular del Aparato de Golgi y el Retículo endoplásmico.	-3.7	Kuge y cols., 1994.
Pp_07A24	NM_126400	1,1E-21	Factor de transcripción de la familia myb	Familia de factores transcripcionales en la que algunos miembros son inducidos por la hormona ácido abscísico y estrés por bajas temperaturas y deshidratación.	-3.9	Martin y Paz-Ares, 1997.
Pp_15O16	NM_105632	1,1E-42	Fosfatasa ácida putativa	Proteínas cuya función consiste en liberar fosfatos desde compuestos presentes en el ambiente. Se ha visto un aumento en su expresión en raíces frente a una disminución de fósforo y nitrógeno en el suelo.	-4.0	Olczak y cols., 2003.
Pp_29E03	NM_116532	1,3E-21	Proteína 1 de unión a auxinas (ABP1)	Receptor de auxinas que induce un aumento en la expansión celular.	-4.1	Napier, R., 2004.

Clon ID	Código Genbank	e value	ID Putativo	Descripción	Veces de cambio	Referencias
Pp_24H21	NM_129030	1,4E-21	fatty acid hydroxylase (FAH1)	Enzima responsable de la síntesis de ácidos grasos de cadena muy larga (very-long-chain fatty acids).	-4.3	Aguade, 2001
Pp_22D10	NM_114287	1,1E-37	Piridina nucleotido-disulfuro oxidorreductasa	Oxidoreductasas responsables de reducir FAD desde NADH.	-4.7	Perry y cols., 1991
Pp_25P11	NM_116279	1,3E-12	galactosa:MGDG galactosiltransferasa (DGD2)	Enzima que cataliza la síntesis del galactolípido digalactosilglicerol (DGDG) a partir de monogalactosilglicerol (MGDG) y UDP galactosa. DGDG es uno de los lípidos del cloroplasto más importantes en plantas superiores. Se ha reportado que los niveles de transcrito DGD2 aumentan considerablemente en una disminución de la concentración de fósforo. Proteína que unida a la droga ciclosporin, inhibe la actividad fosfatasa de la calcineurina y bloquea la inactivación de los canales de potasio dependientes de calcio en las células de guarda. Además, su actividad peptidilprolisomerasa puede acelerar el plegamiento de algunas proteínas.	-5.0	Kelly y Dormann, 2002
Pp_11C02	NM_120034	1,2E-29	Peptidilproli isomerasa (ROC1)		-5.4	Lee y cols, 2002; Yao y cols., 2001.

DISCUSIÓN.

En este trabajo de memoria de título se han definido algunos aspectos técnicos de la siembra de DNA en membranas de nylon (*macroarrays*) y de la ejecución de experimentos de hibridación en esta plataforma. Además, se logró diseñar un método de estudio de expresión génica a gran escala que demostró ser útil para identificar genes involucrados en vías relacionadas con la harinosidad en duraznos expuestos a bajas temperaturas.

1.- Puesta a punto del método de confección de membranas de *macroarrays*, y de la obtención y análisis de datos de hibridaciones.

Al iniciar este estudio nos planteamos definir una manera de cuantificar las señales de hibridación obtenidas a partir de imágenes en formato TIFF. Consideramos que el valor de intensidad de señal en un *spot* es igual al valor de la mediana de las intensidades de los píxeles dentro de su correspondiente celda de cuantificación, de área constante a lo largo de toda la membrana. En cada uno de los *spots*, el valor de intensidad neta fue corregido por el valor de la intensidad del ruido, que se definió como área comprendida entre la celda de cuantificación y un cuadrado circunscrito a ella, y que corresponde a señales producidas por hibridaciones inespecíficas e irregularidades de la superficie de la membrana. Existen numerosas formas de adquisición de datos desde las imágenes obtenidas a partir de experimentos de *macroarrays*, y se ha sugerido

que el modo de estimación de la intensidad de señal debe ser evaluado según las características de la plataforma utilizada (Baldi y Hatfield, 2002). Los métodos utilizados con mayor frecuencia para estimar las intensidades de la señal y del ruido corresponden al cálculo del promedio y de la mediana de los píxeles dentro del área definida para tales estimaciones. Hedge y cols. (2000) postula que los métodos para cuantificar la señal dentro de un *spot* son equivalentes ya que el análisis final de los datos de *macroarrays* o *microarrays* utiliza razones de intensidades para las dos condiciones a comparar. Sin embargo, Leung y Cavalieri (2003) recomiendan el uso de la mediana como estimador de intensidades, sobre todo en casos donde existen señales extremas de píxeles dentro del área de cuantificación que pueden interferir con la correcta medición de la señal.

Una vez establecido el método para extraer los datos de intensidad neta desde las imágenes de las hibridaciones, quisimos optimizar algunos parámetros de la siembra de DNA en membranas de nylon. El primer paso fue determinar la cantidad de DNA blanco por *spot*. Usando una cantidad de sonda constante y cantidades crecientes de DNA blanco, pudimos determinar que las intensidades netas aumentaban linealmente al aumentar la cantidad de DNA depositado en la membrana, y que se obtenían señales por sobre el nivel de ruido local a partir de 20 ng de DNA plasmidial (6×10^9 copias de DNA blanco), cantidad que está dentro del rango reportado por Bernard y cols. (1996), quienes propusieron un método de estudio de la expresión génica a mediana escala similar al desarrollado en este trabajo. Los autores comprobaron que

existe una relación lineal de la intensidad neta frente a una variación tanto de la cantidad de DNA blanco como de la cantidad de sonda usada. Ellos observaron linealidad entre 0,3 y 20 ng de DNA en la forma de producto de PCR, lo que equivale a 3×10^8 y 2×10^{10} moléculas de DNA blanco, respectivamente, estos resultados apoyan el uso de 10^9 copias de DNA blanco por *spot*.

Frente a eventuales problemas en la obtención de DNA blanco en concentraciones suficientes, esta cantidad de DNA puede ser alcanzada en las membranas concentrando la muestra en los *spots* mediante múltiples siembras. No obstante, antes de utilizar esta modalidad de siembra, se evaluó la variación (CV) presente en los datos obtenidos desde membranas en las que se depositó 1, 3 y 5 veces el DNA blanco en cada *spot*. Los resultados muestran que al sembrar 5 veces cada punto se disminuye considerablemente el CV de spots sembrados en duplicado, estos resultados confirman los obtenidos por Nogueira y cols. (2003), quienes determinaron que la siembra repetida de DNA blanco reduce en un 50% el CV entre los *spots*. De este modo, además de aumentar la cantidad de DNA presente en cada *spot*, estamos disminuyendo la variación en la cantidad de DNA sembrado a lo largo de la membrana.

En este trabajo el llamado "efecto de alrededores", definido como el solapamiento de las señales de puntos de hibridación vecinos (Machl y cols., 2002) fue minimizado aumentando las distancias entre los *spots*. Se determinó que una distancia de 1800 μm disminuye en seis veces la variabilidad que tienen los valores de ruido local, en comparación con la variabilidad obtenida usando una separación de 1200 μm . De acuerdo con ello, cuando los puntos se

encontraban separados por una distancia de 1800 μm , se logró una mejor correlación entre los valores de intensidad neta vs. cantidad de DNA blanco. De este modo estamos disminuyendo el nivel del ruido local de las señales al disminuir la interferencia que tiene un punto de hibridación sobre los *spots* adyacentes a él. Sin embargo, también estamos disminuyendo la densidad de puntos de nuestro arreglo. Un procedimiento alternativo es el descrito por Machl y cols. (2002) quienes propusieron aumentar la densidad de la siembra en el arreglo y minimizar el “efecto de alrededores” mediante un factor de corrección. Este factor fue determinado experimentalmente y relaciona las diferencias de las señales entre *spots* sembrados en duplicado con las diferencias de las señales de sus respectivos *spots* “vecinos”. Machl y cols. (2002) demostraron que esta relación es lineal y su pendiente describe el efecto de las intensidades de las señales de los *spots* adyacentes sobre un *spot* particular.

Analizamos, además, la calidad de las señales obtenidas de *macroarrays* en los que un mismo conjunto de genes fue sembrado utilizando tres DNAs obtenidas con métodos distintos: DNA plasmidial, producto de PCR y producto purificado de PCR. Diehl y cols. (2002), determinaron que en *microarrays* de cDNA sembrados en vidrio es posible usar productos de PCR sin purificar como DNA blanco, ya que la intensidad de señales que ellos obtienen es levemente menor (94%) a la intensidad obtenida con la siembra de productos de PCR purificados. Además comprueban que los productos purificados de PCR muestran una mayor fluctuación en la cantidad de DNA presente en los *spots*, y

que la presencia de partidores y DNA templado en el producto de PCR sin purificar no afecta mayormente a la intensidad de la señal obtenida. Sin embargo, en el trabajo citado no se cuantifica la recuperación de señal de intensidad neta en función de la cantidad de moléculas de DNA sembradas. En nuestro trabajo mostramos que si bien, el uso de productos sin purificar tiene una recuperación de señal aceptable, el producto purificado tiene una recuperación de señal por femtomol de DNA blanco que es 10 veces mayor, lo cual indica que el proceso de purificación tiene un mayor efecto en la señal obtenida que lo reportado en Diehl y cols. (2002). En este análisis es necesario considerar que las técnicas de purificación a gran escala de productos de PCR permiten recuperar sólo entre un 20% y un 50% del producto inicial (Diehl y cols., 2002) y poseen desventajas adicionales como un alto costo en tiempo y recursos, en contraste se ha estimado que la amplificación de productos de PCR a partir de colonias bacterianas (Fig. 10) genera en promedio $1,83 \mu\text{g}/\mu\text{L}$ una concentración suficiente para llevar a cabo la siembra de numerosas membranas. Así, podemos concluir que a pesar de obtener una mayor recuperación de señal utilizando los productos de PCR purificados, los productos sin purificar presentan la ventaja de tener un costo mucho menor, en condiciones que otorgan una intensidad de señal aceptable, y una especificidad y homogeneidad similar a los productos purificados de PCR.

Luego de determinar algunos parámetros de la siembra de DNA en membranas, examinamos dos características de los arreglos generados mediante nuestro protocolo: la reproducibilidad y la sensibilidad en ensayos de

hibridación. Pudimos determinar que los datos provenientes de tres hibridaciones independientes tuvieron una buena correlación entre sí, reflejada en un ajuste de un 90% a una relación lineal de pendiente cercana a 1 e intercepto igual a cero (Fig. 6, página 44). En otros estudios (Donovan y Becker, 2002; Herwig y cols., 2001) se han reportado valores similares al comparar los valores logarítmicos de intensidad neta entre ensayos independientes de hibridación en membranas de nylon.

En relación con la sensibilidad de esta técnica, nuestros resultados indican que usando 200 ng de mRNA para la síntesis de sondas marcadas, la señal de menor intensidad neta detectada, correspondía a un gen cuya expresión alcanzaba las 36 copias mRNA/célula (una abundancia de 0,01% según Kane y cols., 2000). En los ensayos de expresión génica en frutos utilizamos 150 ng y 70 ng de mRNA de durazno, pudiendo detectar señales de hibridación correspondientes a transcritos con un nivel de abundancia de 0,0045%. Esto puede deberse a que en el ensayo de expresión génica se utilizaron cantidades mayores de mRNA, en comparación con el estudio de la sensibilidad de nuestro ensayo, donde la cantidad señalada de mRNA utilizada corresponde a una estimación hecha en base a la proporción teórica de mRNA contenido en una muestra de RNA total eucarionte, correspondiente a un 2% (un 2% de 10 μ g corresponden a 200 ng). En el estudio de Kane y cols. (2000) se reportó una detección de tres veces de cambio, para abundancias mínimas de transcrito de 30 y 10 copias de mRNA por célula, sensibilidad que se

encuentra dentro del mismo orden a la obtenida en nuestro estudio. Bertucci y cols. (1999) determinaron que su ensayo (hibridaciones en membranas de nylon usando sondas radioactivas) puede detectar transcritos expresados en una abundancia de 0,01% utilizando sólo 100 ng de RNA total, una cantidad bastante baja de material inicial (2 ng de mRNA), con lo cual podemos considerar al método descrito como una buena alternativa para ensayos de expresión génica en que las cantidades de material inicial son limitadas. Sin embargo, el equipo de lectura y obtención de imágenes que se utilizó en dicho trabajo tiene una resolución de 25 μm , dos veces mayor a la resolución del equipo de lectura utilizado en nuestros estudios. A pesar de esto, la sensibilidad medida en nuestro ensayo es bastante alta (dos veces mayor que la de los experimentos de Bertucci y cols., 1999), indicando que nuestro protocolo es eficiente en la detección de transcritos poco expresados.

2.- Determinación de los cambios en la expresión de genes en duraznos harinosos.

Con el propósito de relacionar funcionalmente la generación de la harinosidad en frutos de duraznos con su almacenamiento a bajas temperaturas, quisimos comparar los perfiles de expresión génica entre frutos que maduraron normalmente y frutos que fueron sometidos a 4 °C por 30 días. Para ello se realizó la selección de 880 genes relacionados con funciones o procesos biológicos, tales como: 1) Metabolismo de pared celular, existen

varias líneas de evidencia que señalan diferencias en la expresión y actividad de enzimas remodeladoras de la pared celular en frutos expuestos a bajas temperaturas y una correlación entre estas diferencias y la adquisición del fenotipo harinoso (Brummell y cols., 2004; Crisosto y Labavitch, 2002); 2) Estrés térmico, por estar directamente involucradas en la reacción de la célula frente a una baja en la temperatura exterior, los genes que codifican para proteínas de estrés térmico fueron incluidos en este análisis. Se ha postulado que, en frutos de tomate, estas proteínas ejercen un posible efecto protector contra las alteraciones, causadas por el estrés por frío (Sabehat y cols., 1998; Kadyrzhanova y cols., 1998); 3) Maduración, se han seleccionado genes involucrados con procesos de desarrollo de frutos, respuestas a hormonas vegetales y metabolismo de etileno (Brands y Ho, 2002; Trainotti y cols., 2003); 4) otros procesos, vías de transducción activadas por bajas temperaturas y estrés oxidativo.

Estos criterios de selección permitieron reducir el volumen de genes a analizar y centrar nuestro estudio de expresión génica en aquellas vías metabólicas posiblemente involucradas en el fenómeno de la harinosidad. Como se discute más adelante, este proceso nos permitió obtener una cantidad considerable de genes diferencialmente expresados entre las condiciones estudiadas, a diferencia de otros estudios de expresión génica en gran escala, en los cuales los genes analizados han sido seleccionados al azar (Nogueira y cols., 2003).

Nuestro procedimiento de filtrado y selección de datos confiables nos permitió rescatar 262 señales de hibridación que superaron todos los filtros de calidad en las dos condiciones estudiadas y que correspondieron al 30% de los genes iniciales. De ellos sólo 37 mostraron un cambio significativo de expresión (14,12% de los datos recuperados). Esta cantidad es alta en comparación con la cantidad de genes significativamente expresados en el trabajo de Nogueira y cols. (2003). Estos autores utilizaron membranas de nylon en las cuales depositaron manualmente 1.536 ESTs de caña de azúcar (*Saccharum* spp.), de los cuales 59 mostraron un cambio significativo en su expresión (3,84%). Estos ESTs correspondían a clones escogidos al azar desde cuatro genotecas de que comprendían un total de 43.141 clones, es decir, no hubo una selección de genes relacionados con el proceso biológico a investigar, y por lo tanto es de esperar encontrar en el análisis citado una gran cantidad de genes cuyo nivel de expresión no cambió significativamente en las condiciones estudiadas.

2.1.- Análisis de los genes diferencialmente expresados.

2.1.1.- Genes reprimidos en el fruto harinoso.

En nuestro estudio se obtuvieron 34 genes reprimidos desde una condición NH a H, los cuales pueden agruparse en 6 grupos, que se detallan a continuación

2.1.1.1.- Biosíntesis, Estructura y Metabolismo de la Pared Celular.

Algunos de los genes que disminuyen su expresión en el durazno harinoso codifican para presuntas proteínas involucradas en la biosíntesis y

metabolismo de componentes de la pared celular, tales como: la sacarosa sintasa (Pp_03F22) y la fosfoglucomutasa (Pp_28D07). Se detectó también una disminución en la expresión del gen que codifica para xiloglucano endotransglicosilasa (*exgt-A3*, Pp_10L02), la cual ha sido reportada como un elemento de remodelación de la pared celular y desarrollo de la pared secundaria (Bourquin y cols., 2002; Johansson y cols., 2004). Estos resultados sugieren, que en el durazno harinoso, existiría una expresión disminuida de genes que codifican para proteínas responsables de la síntesis de componentes de la pared celular, así como proteínas involucradas en el remodelamiento estructural de la pared. Adicionalmente, un aumento en la expresión del gen que codifica para la enzima pectinmetilesterasa estaría contribuyendo a aumentar la des-metilación de las pectinas, generando monómeros de pectato altamente hidrosolubles que en conjunto con el Ca^{+2} son responsables de la retención de agua en los frutos en forma de gel. Esto último es causa de la textura seca y harinosa de los frutos almacenados a bajas temperaturas.

2.1.1.2.-Secreción de proteínas.

Las enzimas encargadas del metabolismo, estructura y modificación de la pared celular se movilizan a través del sistema de endomembrana de la célula y probablemente serían secretadas al apoplasto (Staehein y Moore, 1995). De la misma forma, algunos carbohidratos no celulósicos de la pared, tales como, las hemicelulosas y pectinas también son destinadas por la vía secretoria previa síntesis en las cisternas de Golgi (Dupree y Sherrier, 1998).

Estos antecedentes nos indican que el sistema de secreción de proteínas es fundamental para un adecuado metabolismo de la pared celular. Un grupo de genes que codifica para proteínas involucradas en las vías de secreción en plantas, entre ellos: elementos de reconocimiento de péptido señal, proteína de unión a GTP perteneciente a la familia Ras y proteínas de unión a GTP SAR1A (Schunemann, 2004; Vernoud y cols., 2003; Kuge y cols., 1994), disminuyen significativamente su expresión en el durazno harinoso. Análisis de PCR en tiempo real efectuados en el Laboratorio de Bioinformática y Expresión Génica, INTA, dan cuenta de cambios en la expresión de un grupo mayor de genes que codifican para proteínas involucradas en los sistemas de tráfico intracelular y secreción proteica (n=13) en duraznos harinosos (González, 2005). En conjunto, estos resultados sugieren que existiría una relación entre la alteración en la expresión de estos genes y el desarrollo de la harinosidad, la cual se produciría debido a una modificación directa o indirecta del tráfico de proteínas a la pared celular.

2.1.1.3.- Transducción de señales mediadas por Ca^{+2} en respuesta al frío.

Entre otros procesos biológicos, el ión calcio (Ca^{+2}) actúa como segundo mensajero en las vías de transducción de señales involucradas en las respuestas al frío, activando proteínas quinasas, las cuales fosforilan a su vez a proteínas que estarían participando activamente en el control de la expresión de genes inducibles por frío y en el desarrollo de la tolerancia a las bajas temperaturas (Monroy y cols., 1993). Además se ha reportado que existe un

aumento transitorio en la concentración intracelular del ión frente a una exposición a frío en plantas de *A. thaliana* y alfalfa (Thomashow, 1999), y que este ión participa en la formación de uniones entre los polímeros de pectina que constituyen la pared celular (Knight y cols., 1991). En nuestro estudio, tres genes cuyos niveles de transcrito disminuyen entre frutos no harinosos y frutos harinosos, codifican para proteínas que participan en la homeostasis de Ca^{+2} y en vías de señalización mediadas por Ca^{+2} . Uno de ellos corresponde a *crt2*, que codifica para la calreticulina 2, proteína responsable de regular la concentración de Ca^{+2} en el retículo endoplásmico (Persson y cols., 2001). Otro gen que disminuye frente a un tratamiento prolongado a bajas temperaturas es la proteína quinasa dependiente de calcio (CDPK). Contrario a lo observado en nuestro análisis, Mitra y Johri (2000) reportaron que la expresión de este gen aumenta en respuesta al estrés por deshidratación y bajas temperaturas. Una posible explicación de estos resultados contradictorios proviene de la naturaleza del tratamiento empleado, es así como la muestra harinosa correspondía a un durazno cosechado, sometido a frío por 20 días y luego madurado a temperatura ambiente por 6 días, período durante el cual la abundancia de los transcritos de CDPK podría haber disminuido luego del período de estrés. Por otra parte, encontramos que el gen que codifica para la calmodulina 4 (*cam1/4*) disminuye su expresión en el durazno harinoso. Se ha descrito que esta proteína es el transductor de calcio más prominente en las células eucariontes, regulando la actividad de una variedad de proteínas quinasas involucradas en procesos tales como la reacción frente a tratamientos con hormonas vegetales,

el desarrollo de algunos tejidos, y la respuesta a estrés biótico y abiótico (Bouche y cols., 2005). En relación con el estrés por bajas temperaturas, se ha demostrado que en *A. thaliana* el gen *cam1/4* actúa como regulador negativo del gen *COR15a* de respuesta al frío (Townley y Knight, 2002). Esto último nos sugiere que posiblemente en nuestro durazno harinoso, al disminuir la abundancia del transcrito *cam1/4*, *COR15a* podría estar acumulado en la célula. Sin embargo, este último transcrito no fue incluido en nuestro análisis, siendo interesante llevar a cabo un experimento en el que se mida la abundancia del transcrito *COR15a* junto con la del transcrito *cam1/4* en un fruto tratado con bajas temperaturas.

2.1.1.4.- Protección al estrés oxidativo.

No sólo las funciones relacionadas con la estructura y el metabolismo de la pared celular se ven alteradas en los frutos harinosos. También se ha encontrado una represión del gen que codifica para una dioxigenasa dependiente de 2-oxoglutarato (Pp_19B20), responsable de la síntesis de flavonoides. Estos compuestos, además de brindar características organolépticas agradables en los frutos (Turnbull y cols., 2004), las que estarían disminuyendo en los duraznos expuestos a bajas temperaturas, entregan protección contra el daño oxidativo, causado entre otras razones por el estrés a bajas temperaturas.

2.1.1.5.- Genes asociados a hormonas vegetales.

Nuestros resultados indican una disminución en la expresión del gen *iaa12* (Pp_17A10), activado por auxinas y de un gen que codifica para un precursor de proteínas ABP1, receptoras de estas hormonas (Pp_29E03), las que son importantes en procesos relacionados con el crecimiento celular y la maduración de frutos (Abel y cols., 1995). Además, se encontró una disminución en la abundancia del transcrito "Pp_07A24", similar a genes pertenecientes a la familia *myb*. Se ha reportado que estos genes codifican para factores de transcripción inducidos por la hormona ácido abscísico (ABA), cuya participación en sistemas de señalización en respuesta a estrés ambientales ha sido bastante estudiada (Xiong y cols., 2002; Martin y Paz-Ares, 1997). Una baja en la cantidad de transcrito correspondiente a *myb* nos sugiere que el durazno harinoso presenta una alteración frente a las bajas temperaturas. Por otra parte encontramos una disminución en la expresión del gen *hva22* (Pp_25G17), codificante para la proteína HVA22, inducida por ABA, y responsable de regular el tráfico de vesículas en células vegetales sometidas a estrés para facilitar el recambio de membranas, o para disminuir la secreción excesiva, hecho comprobado por Brands y Ho (2002), mediante experimentos de complementación en levaduras. En relación con esto último, se ha descrito que *hva22* interactúa físicamente con GTPasas Rab, a las cuales se les atribuye un importante papel en varios aspectos del transporte vesicular (Yang y cols., 1998). Esta disminución en la abundancia de transcrito *hva22*, nos sugiere una vez más que en el fruto harinoso existe una alteración en la

respuesta a estrés en el contexto del transporte vesicular y la secreción de proteínas.

2.1.1.6.- Genes de respuesta a estrés térmico.

En plantas, la expresión de genes que codifican proteínas que responden al estrés térmico (*hsp's*) aumenta durante el estrés por frío (Anderson y cols., 1994), favoreciendo una mayor tolerancia a este tipo de estrés, ya que estas proteínas actúan como chaperonas moleculares estableciendo relaciones físicas reversibles con diversos tipos de proteínas, permitiendo un correcto plegamiento de ellas o previniendo su desnaturación. Nuestros resultados señalan, que en el fruto harinoso, la expresión de dos putativas *hsp's* (Pp_11M19 y Pp_09B20), disminuye significativamente, sugiriendo una alteración en los mecanismos reguladores de la expresión de estos genes, la cual está regulada por factores de transcripción de unión a DNA denominados HSF's (*Heat Shock Factors*). Apoyando esta afirmación, y volviendo a mencionar el gen codificante para la calmodulina 1 (*cam1/4*), se ha descrito que este último participa en la activación de la unión a DNA por parte de HSF activando la expresión de distintas *hsp's* (Liu y cols., 2004), lo que nos sugiere que éstas últimas pueden estar siendo reprimidas indirectamente por la disminución de la expresión de *cam1/4*.

2.1.2.- Genes inducidos en el fruto harinoso.

Nuestros resultados indican que sólo tres genes aumentaron significativamente su expresión en el fruto harinoso. Entre ellos se encuentra el gen de la pectin-metilesterasa (*pme*, Pp_05K11), el cual tuvo un aumento de 5,4 veces en su expresión. Este aumento en la expresión de *pme* ha sido anteriormente reportado en frutos de durazno sometidos al frío (Crisosto y cols., 1999) y es la causa más directa de la textura harinosa de los duraznos tratados con bajas temperaturas. Otro gen que muestra un aumento en su expresión en la condición harinosa es el gen *gad2* (Pp_03I05), responsable de la síntesis de aminoácido no proteico GABA en plantas, frente a un stress por bajas temperaturas (Turano y Fang, 1998; Yevtushenko y cols., 2003). El tercer gen cuya expresión se vio aumentada en el durazno harinoso es *grf4* (14-3-3 ϕ), el cual es miembro de una amplia familia de genes que codifican para proteínas 14-3-3, cuyas funciones son diversas, incluyendo señales de transducción mediadas por quinasas, señalización en el crecimiento celular y desarrollo, estructura y movimiento, y respuesta a estrés biótico y abiótico (Sehnke y cols., 2002).

Si bien, este grupo de genes diferencialmente expresados nos permite inferir posibles procesos alterados por la exposición de fruto a bajas temperaturas, sin duda es necesario realizar los ensayos de hibridación utilizando replicas biológicas que nos permitan evaluar la variabilidad individual de cada fruto. Por otra parte, los datos de cambios de expresión génica deben

ser validados por un método independiente, por ejemplo mediante PCR en tiempo real, herramienta que permite cuantificar en forma rápida y precisa la cantidad de transcrito en una muestra de cDNA (Klein, 2002).

CONCLUSIÓN

En resumen, el protocolo de siembra, obtención de datos, selección de genes, limpieza de datos e identificación de genes diferencialmente expresados, permitieron asociar el fenómeno de harinosidad con vías del metabolismo de la pared celular, elementos de transducción de señales de Ca^{+2} , algunos genes que participan en vías de destinación de proteínas, proteínas de estrés térmico, genes activados por la acción de hormonas vegetales, etc. A modo de conclusión, podemos señalar que se ha implementado, diseñado y estandarizado un método de siembra de *macroarrays* de cDNA y de análisis de datos obtenidos desde hibridaciones utilizando *macroarrays* y sondas complejas marcadas radiactivamente, sintetizadas desde muestras complejas de mRNA., Este análisis permitió encontrar genes diferencialmente expresados en el contexto de un análisis transcripcional a gran escala, y permitió profundizar el conocimiento de las causas que llevan a la adquisición del fenotipo harinoso en el fruto de *P. persica*.

REFERENCIAS BIBLIOGRÁFICAS

- **Abel S., Nguyen M. y Theologis A. 1995.** The PS-IAA4/5-like Family of Early Auxin-inducible mRNAs in *Arabidopsis thaliana*. *J. Mol. Biol.* **251**:533-549.
- **Aguade M. 2001.** Nucleotide sequence variation at two genes of the phenylpropanoid pathway, the FAH1 and F3H genes, in *Arabidopsis thaliana*. *Mol. Biol. Evol.* **18**:1-9.
- **Aharoni A., Keizer L., Bouwmeester H., Sun Z., Alvarez-Huerta M., Verhoeven H., Blaas J., van Houwelingen A., De Vos R., van der Voet H., Jansen R., Guis M., Mol J., Davis R., Schena M., van Tunen A. y O'Connell A. 2000.** Identification of the SAAT gene involved in strawberry flavour biogenesis by use of DNA microarrays. *Plant Cell.* **12**:647-662.
- **Alizadeh A., Eisen M., Davis R., Ma C., Lossos I., Rosenwald A., Boldrick J., Sabet H., Tran T., Yu X., Powell J., Yang L., Marti G., Moore T., Hudson J., Lu L., Lewis D., Tibshirani R., Sherlock G., Chan W., Greiner T., Weisenburger D., Armitage J., Warnke R., Levy R., Wilson W., Grever M., Byrd J., Botstein D., Brown P. y Staudt L. 2000.** Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* **403**:503-511.
- **Andersen P., Kragelund B., Olsen A., Larsen F., Chua N., Poulsen F. y Skriver K. 2004.** Structure and biochemical function of a prototypical *Arabidopsis* U-box domain. *J. Biol. Chem.* **279**:40053-40061.
- **Anderson J., Li Q., Haskell D. y Guy C. 1994.** Structural organization of the spinach endoplasmic reticulum-luminal 70-kilodalton heat-shock cognate gene and expression of 70-heat shock genes during cold acclimation. *Plant Physiol.* **104**:1359-1370.
- **Apweiler R., Attwood T., Bairoch A., Bateman A., Birney E., Biswas M., Bucher P., Cerutti L., Corpet F., Croning M., Durbin R., Falquet L.**

- Fleischmann W., Gouzy J., Hermjakob H., Hulo N., Jonassen I., Kahn D., Kanapin A., Karavidopoulou Y., Lopez R., Marx B., Mulder N., Oinn T., Pagni M., Servant F., Sigrist C. y Zdobnov E. 2001.** The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* **29**:37-40.
- **Ausubel F., Brent R., Kingston R., Moore D., Seidman J., Smith J. y Struhl K. 1998.** Current Protocols in Molecular Biology. J. Wiley & Sons, New York, NY.
 - **Baldi P. y Hatfield W. 2002.** DNA microarrays and gene expression, from Experiments to Data Analysis and Modelling, en "DNA array formats", pp.7-16. Cambridge University Press, Cambridge, United Kingdom.
 - **Ben-Arie R. y Sonogo L. 1980.** Pectolytic enzyme activity involved in woolly breakdown of stored peaches. *Phytochemistry.* **19**:2553-2555.
 - **Bernard K., Auphan N., Granjeaud S., Victorero G., Schmitt-Verhulst A., Jordan B. y Nguyen C. 1996.** Multiplex messenger assay: simultaneous, quantitative measurement of expression of many genes in the context of T cell activation. *Nucleic Acids Res.* **24**:1435-1442.
 - **Bertucci F., Bernard K., Loriod B., Chang Y., Granjeaud S., Birnbaum D., Nguyen C., Peck K. y Jordan B. 1999.** Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples. *Hum. Mol. Gen.* **8**:1715-1722.
 - **Biemelt S., Tschiersch H. y Sonnewald U. 2004.** Impact of Altered Gibberellin Metabolism on Biomass Accumulation, Lignin Biosynthesis, and Photosynthesis in Transgenic Tobacco Plants. *Plant Physiol.* **135**:254-265.
 - **Birnboim H. 1983.** A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.* **100**:243-255.
 - **Bittner M., Meltzer P., Chen Y., Jiang Y., Seftor E., Hendrix M., Radmacher M., Simon R., Yakhini Z., Ben-Dor A., Sampas N.,**

- Dougherty E., Wang E., Marincola F., Gooden C., Lueders J., Glatfelter A., Pollock P., Carpten J., Gillanders E., Leja D., Dietrich K., Beaudry C., Berens M., Alberts D. y Sondak V. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature*. **406**:536-40.
- **Bouche N., Yellin A., Snedden W. y Fromm H. 2005.** *Annu. Rev. Plant Biol.* **56**:435-466.
 - **Bourquin V., Nishikubo N., Abe H., Brumer H., Denman S., Eklund M., Christiernin M., Teeri T., Sundberg B. y Mellerowicz E. 2002.** Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls in vascular tissues. *Plant Cell* **14**:3073-3088.
 - **Brands A. y Ho T. 2002.** Function of a Plant Stress-Induced Gene, *HVA22*. Synthetic Enhancement Screen with Its Yeast Homolog Reveals Its Role in Vesicular Traffic. *Plant Physiol.* **130**:1121-1131.
 - **Breyne P. y Zabeau M. 2001.** Genome-wide expression analysis of plant cell cycle modulated genes. *Curr. Op. Plant Biol.* **4**:136-142.
 - **Brinker M., van Zyl L., Liu W., Craig D., Sederoff R., Clapham D. y von Arnold S. 2004.** Microarray Analyses of Gene Expression during Adventitious Root development in *Pinus contorta*. *Plant Physiol.* **135**:1526-1539.
 - **Brummell D., Dal Cin V., Lurie S., Crisosto C. y Labavitch J. 2004.** Cell wall metabolism during the development of chilling injury in cold-stored peach fruit: association of mealiness with arrested disassembly of cell wall pectins. *J. Exp. Bot.* **55**:2041-2052.
 - **Bucca G., Carruba G., Saetta A., Muti P., Castagnetta L. y Smith C. 2004.** Gene expression profiling of human cancers. *Ann. N. Y. Acad. Sci.* **1028**:28-37.
 - **Chopra S., Del-favero J., Dolferus R. y Jacobs M. 1992.** Sucrose synthase of *Arabidopsis*: genomic cloning and sequence characterization. *Plant. Mol. Biol.* **18**:131-134.

- **Crisosto C., Mitchell G., y Ju Z. 1999.** Susceptibility to chilling injury of peach, nectarine, and plum cultivars grown in California. *HortScience* **34**:1116-1118.
- **Crisosto C. y Labavitch J. 2002.** Developing a quantitative method to evaluate peach (*Prunus persica*) flesh mealiness. *Post-harvest Biology and Technology*. **25**:151-158.
- **Datson N., van der Perck-de Jong J., van den Berg M., de Kloet E. y Vreugdenhil E. 1999.** MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucleic Acids Res.* **29**:1300-1307.
- **De Risi J., Iyer V. y Brown P. 1997.** Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. *Science*. **278**:680-686.
- **Diehl F., Beckmann B., Kellner N., Hauser N., Diehl S. y Hoheisel J. 2002.** Manufacturing DNA microarrays from unpurified PCR products. *Nucleic Acids Res.* **30**:e79.
- **Donovan D. y Becker K. 2002.** Double round hybridization of membrane based cDNA arrays: improved background reduction and data replication. *J. Neurosci. Methods*. **118**:59-62.
- **Dupree P. y Sherrier D. 1998.** The Plant Golgi apparatus. *Biochim. Biophys. Acta*. **1404**:259-270.
- **Durrant W., Rowland O., Piedras P., Hammond-Kosack K. y Jones J. 2000.** cDNA-AFLP reveals a striking overlap in race specific resistance and wound response gene expression profiles. *Plant Cell*. **12**:963-967.
- **Emanuelsson O., Nielsen H., Brunak S. y von Heijne G. 2000.** Predicting subcellular localization proteins based in their N-terminal amino acid sequence. *J. Mol. Biol.* **300**:1005-1016.
- **Freeman W., Robertson J. y Vrana, K. 2000.** Fundamentals of DNA Hybridization Arrays for Gene Expression Analysis. *Biotechniques*. **29**:1042-1055.

- **Fujita K. y Silver J. 1993.** Surprising lability of biotin-streptavidin bond during transcription of biotinylated DNA bound to paramagnetic streptavidin beads. *Biotechniques*. **14**:608-617.
- **Giovannoni J. 2001.** Molecular biology of fruit: maturation and ripening. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**:725-749.
- **González M. 2005.** Datos no publicados. Laboratorio de Bioinformática y expresión génica, INTA, Universidad de Chile.
- **Hedge P., Qi R., Abernathy K., Gay C., Dharap S., Gaspard R., Hughes J., Snesrud E., Lee N. y Quackenbush J. 2000.** A Concise Guide to cDNA Microarray Analysis. *Biotechniques*. **29**:548-562.
- **Heese-Pech A. y Raikhel N. 1998.** A glycoprotein modified with terminal N-acetylglucosamine and localized at the nuclear rim shows sequence similarity to aldose-1-epimerases. *Plant Cell*. **10**:599-612.
- **Heller M. J. 2002.** DNA array technologies: Devices, Systems and Applications. *Annu. Rev. Biomed. Eng.* **4**:129-153.
- **Herwig R., Aanstad P., Clark M. y Lehrach, H. 2001.** Statistical evaluation of differential expression on cDNA nylon arrays with replicated experiments. *Nucleic Acids Res.* **29**:e117.
- **Horak C.E. y Snyder M. 2002.** Global analysis of gene expression in yeast. *Funct. Integr. Genomics*. **2**:171-180.
- **Johansson P., Brumer H., Baumann M., Kallas A., Henriksson H., Denman S., Teeri T. y Jones T. 2004.** Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. *Plant Cell*. **16**:874-886.
- **Kadyrzhanova D., Vlachonasios K., Ververidis P. y Dilley, D. 1998.** Molecular cloning of a novel heat induced/chilling tolerance related cDNA in tomato fruit by use of mRNA differential display. *Plant Mol. Biol.* **36**:885-895.

- **Kane M., Jatkoe T., Stumpf C., Lu J., Thomas J. y Madore S. 2000.** Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res.* **28**:4552-4557.
- **Kelly A. y Dormann P. 2002.** DGD2, an *Arabidopsis* gene encoding a UDP-galactose-dependent digalactosyldiacylglycerol synthase is expressed during growth under phosphate-limiting conditions. *J. Biol. Chem.* **277**:1166-1173.
- **Klein D. 2002.** Quantification using real-time PCR technology: applications and limitations. *Trends Mol. Med.* **8**:257-260.
- **Knight M., Campbell A., Smith S. y Trewavas A. 1991.** Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytosolic calcium. *Nature.* **352**:524-526.
- **Kuge O., Dascher C., Orci L., Rowe T., Amherdt M., Plutner H., Ravazzola M., Tanigawa G., Rothman J. y Balch W. 1994.** Sar1 promotes vesicle budding from the endoplasmic reticulum but not Golgi compartments. *J. Cell Biol.* **125**:51-65.
- **Lee Y., Hong C., Suh Y. y Lee I. 2002.** A cDNA clone for cyclophilin from *Griffithsia japonica* and phylogenetic analysis of cyclophilins. *Mol. Cells.* **13**:12-20.
- **Leung Y. y Cavalieri D. 2003.** Fundamentals of cDNA microarray data analysis. *Trends Genet.* **19**:649-659.
- **Liang, P. y Pardee, A., 1992.** Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**:967-971.
- **Liu H., Li B., Shang Z., Li X., Mu R., Sun D. y Zhou R. 2004.** Calmodulin is involved in heat shock signal transduction in wheat. *Plant Physiol.* **132**:1186-1195.
- **Lockhart D., Dong H., Byrne M., Follettie M., Gallo M., Chee M., Mittmann M., Wang C., Kobayashi M., Horton H. y Brown E. 1996.**

Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Genet.* **21**:20-24.

- **Machl A., Schaab C. e Ivanov I. 2002.** Improving DNA array data quality by minimising 'neighbourhood' effects. *Nucleic Acids Res.* **30**(22):e127
- **Maleck K., Levine A., Eulgem T., Morgan A., Schmid J., Lawton K., Dangl J. y Dietrich, R. 2000.** The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nat. Genet.* **26**:403-410.
- **Martin C., y Paz-Ares J. 1997.** MYB transcription factors in plants. *Trends Genet.* **13**:67-73.
- **Meisel L., Fonseca B., Gonzalez S., Baeza-Yates R., Cambiazo V., Campos R., González M., Orellana A., Retamales J. y Silva H. 2005.** A Rapid and Efficient Method for Purifying High Quality Total RNA from Peaches (*Prunus persica*) for Functional Genomics Analyses. *Biol. Res.* **38**:83-88.
- **Mitra D. y Johri M., 2000.** Enhanced expression of a calcium-dependent protein kinase from the moss *Funaria hygrometrica* under nutritional starvation. *J. Biosci.* **25**:331-338.
- **Monroy A., Sarhan F. y Dhindsa R. 1993.** Cold-Induced Changes in Freezing Tolerance, Protein Phosphorylation, and Gene Expression (Evidence for a Role of Calcium). *Plant Physiol.* **102**:1227-1235.
- **Nakai K. y Horton P. 1999.** PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* **24**(1):34-36.
- **Napier R. 2004.** Plant hormone binding sites. *Ann. Botany.* **93**:227-233.
- **Nguyen C., Rocha D., Granjeaud S., Baldit M., Bernard K., Naquet P. y Jordan B. 1995.** Differential gene expression in the murine thymus assayed by quantitative hybridization of arrayed cDNA clones. *Genomics.* **29**:207-16.

- **Nogueira F., De Rosa V., Menossi M., Ulian E. y Arruda P. 2003.** RNA Expression Profiles and Data Mining of Sugarcane Response to Low Temperature. *Plant Physiol.* **132**:1811-1824.
- **Ojaniemi H., Evengard B., Lee D., Unger E., y Vernon S. 2003.** Impact of RNA extraction from limited samples on microarray results. *Biotechniques.* **35**:968-973.
- **Olczak M., Morawiecka B., Watorek W. 2003.** Plant purple acid phosphatases - genes, structures and biological function. *Acta Biochim. Pol.* **50**:1245-1256.
- **Perry A., Nibhriain N., Brown N. y Rouch D. 1991.** Molecular characterization of the *gor* gene encoding glutathione reductase from *Pseudomonas aeruginosa*: determinants of substrate specificity among pyridine nucleotide-disulphide oxidoreductases. *Mol. Microbiol.* **5**:163 - 171.
- **Persson S., Wyatt S., Love J., Thompson W., Robertson D. y Boss W. 2001.** The Ca²⁺ Status of the Endoplasmic Reticulum Is Altered by Induction of Calreticulin Expression in Transgenic Plants. *Plant Physiol.* **126**:1092-1104.
- **Qin L., Overmars H., Helder J., Popeijus H., van der Voort J., Groenink W., van Koert P., Schots A., Bakker J. y Smant, G. 2000.** An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the cyst nematode *Globodera rostochiensis*. *Mol. Plant. Microbe Interact.* **13**:830-836.
- **Qiu Q., Barkla B., Vera-Estrella R., Zhu J. y Schumaker K. 2003.** Na⁺/H⁺ exchange activity in the plasma membrane of *Arabidopsis*. *Plant Physiol.* **132**:1041-1052.
- **Roth M., Feng L., McConnell K., Schaffer P., Guerra C., Affourtit J., Piper K., Guccione L., Hariharan J., Ford M., Powell S., Krishnaswamy H., Lane J., Guccione L., Intrieri G., Merkel J., Perbost C., Valerio A., Zolla B., Graham C., Hnath J., Michaelson C.,**

- Wang R., Ying B., Halling C., Parman C., Raha D., Orr B., Jedrzkiewicz B., Liao J., Tevelev A., Mattessich M., Kranz D., Lacey M., Kaufman J., Kim J., Latimer D. y Lizardi, P. 2004. Expression profiling using a hexamer-based universal microarray. *Nat. Biotechnol.* **22**:418-426.
- Rouhier N., Villarejo A., Srivastava M., Gelhaye E., Keech O., Droux M., Finkemeier I., Samuelsson G., Dietz K., Jacquot J. y Wingsle G. 2005. Identification of plant glutaredoxin targets. *Antioxid. Redox Signal.* **7**:919-929.
 - Sabehat A., Lurie S. y Weiss D. 1998. Expression of Small Heat-Shock Proteins at Low Temperatures. *Plant Physiol.* **117**:651-658.
 - Saha S., Sparks A., Rago C., Akmaev V., Wang C., Vogelstein B., Kinzler K. y Velculescu V. 2002. Using the transcriptome to annotate the genome. *Nat. Biotechnol.* **20**:508-512.
 - Sambrook J. 1989. Molecular cloning: A laboratory manual, Second Edition (Cold Spring Harbor Laboratory, NY).
 - Schaffer R., Landgraf J., Accerbi M., Simon V., Larson M., Wisman y E. 2001. Microarray analysis of diurnal and circadian-regulated genes in *Arabidopsis*. *Plant Cell.* **13**:113-123.
 - Schena M., Shalon D., Davis R. y Brown P. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* **270**:467-470.
 - Schunemann D. 2004. Structure and function of the chloroplast signal recognition particle. *Curr. Genet.* **44**:295-304.
 - Sehnke P., DeLille J. y Ferl R. 2002. Consummating Signal Transduction: The role of 14-3-3 proteins in the completion of signal-induced transitions in protein activity. *Plant Cell*, supplement 2002:S339-S354.

- **Shalon D., Smith S. y Brown, P. 1996.** A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* **6**:639-645.
- **Staehein L. y Moore I. 1995.** The Plant Golgi apparatus: structure, functional organization and trafficking mechanisms. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* **46**:261 – 288.
- **Terol J., BARGUES M., Carrasco P., Perez-Alonso M. y Paricio N. 2002.** Molecular Characterization and Evolution of the Protein Phosphatase 2A B' Regulatory Subunit Family in Plants. *Plant Physiol.* **129**:808-822.
- **The Gene Ontology Consortium. 2001.** Creating the Gene Ontology Resource: Design and Implementation. *Genome Res.* **11**:1425-1433.
- **Thomashow M., 1999.** PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**:571–599.
- **Townley H. y Knight M. 2002.** Calmodulin as a potential negative regulator of *Arabidopsis* *COR* gene expression. *Plant Physiol.* **128**:1169-1172.
- **Trainotti L., Zanin D. y Casadoro, G. 2003.** A cell wall-oriented genomic approach reveals a new and unexpected complexity of the softening in peaches. *J. Exp. Botany.* **54**:1821-1832.
- **Tseng G., Oh M., Rohlin L., Liao J. y Wong W. 2001.** Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.* **29**:2549-2557.
- **Turano F. y Fang T. 1998.** Characterization of two glutamate decarboxilase cDNA clones from *Arabidopsis*. *Plant Physiol.* **117**:1411-1421.
- **Turnbull J., Nakajima J., Welford R., Yamazaki M., Saito K. y Schofield C. 2004.** Mechanistic studies on three 2-oxoglutarate-

dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavanone 3-beta-hydroxylase. *J. Biol. Chem.* **279**:1206-1216.

- **van Hal N., Vorst O., van Houwelingen A., Kok E., Peijnenburg A., Aharoni A., van Tunen A. y Keijer, J. 2000.** The application of DNA microarrays in gene expression analysis. *J. Biotechnol.* **78**:271-280.
- **Velculescu V., Zhang L., Vogelstein B. y Kinzler K. 1995** Serial analysis of Gene Expression. *Science.* **270**:484-487.
- **Vernoud V., Horton A., Yang Z. y Nielsen E. 2003.** Analysis of the Small GTPase Gene Superfamily of *Arabidopsis*. *Plant. Physiol.* **131**:1191-1208.
- **Wang X., Ghosh S. y Guo S. 2001.** Quantitative quality control in microarray image processing and data acquisition. *Nucleic Acids Res.* **29**:e75-5.
- **Wu K., Rooney M. y Ferl R. 1997.** The *Arabidopsis* 14-3-3 multigene family. *Plant Physiol.* **114**:1421-1431.
- **Xiong L., Schumaker K. y Zhu J. 2002.** Cell signaling during Cold, Drought, and Salt Stress. *Plant Cell. Supplement* 2002:S165-S183.
- **Yang X., Matern H. y Gallwitz D. 1998.** Specific binding to an essential Golgi membrane protein (Yip1p) functionally links the transport GTPases Ypt1p and Ypt31p. *EMBO J.* **17**:4954-4963.
- **Yao J., Kops O., Lu P. y Lu K. 2001.** Functional Conservation of Phosphorylation-specific Prolyl Isomerases in Plants. *J. Biol. Chem.* **276**:13517-13523.
- **Yevtushenko D., McLean M., Peiris S., Van Cauwenberghe O. y Shelp B. 2003.** Calcium/calmodulin activation of two divergent glutamate decarboxylases from tobacco. *J. Exp. Bot.* **54**:2001-2002.