

VCH-FC
Bioteología
T622
C.1



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS**

**Caracterización del patrón de expresión génica de
cuatro estadios de postcosecha de frutos de duraznos
(*Prunus persica*) mediante análisis de ESTs**

Andrés Tittarelli Elaskar

Director de Tesis: Dr. Herman Silva

Laboratorio de Genética Molecular Vegetal

Departamento de Biología

Santiago de Chile, Agosto 2005



Memoria de Título entregada a la Universidad de Chile en cumplimiento parcial de los requisitos para optar al Título de Ingeniero en Biotecnología Molecular

“Caracterización del patrón de expresión génica de cuatro estadios de postcosecha de frutos de duraznos (*Prunus persica*) mediante análisis de ESTs.”

ANDRÉS TITTARELLI ELASKAR

Dr. Herman Silva Ascencio
Director de Memoria de Título

A handwritten signature in black ink, appearing to be 'HSA', written over a horizontal line.

Comisión de Evaluación de la Memoria de Título

Dra. Liliana Cardemil Oliva
Presidente de Comisión

A handwritten signature in blue ink, appearing to be 'Liliana Cardemil', written over a horizontal line.

Dr. Michael Handford
Profesor Corrector

A handwritten signature in blue ink, appearing to be 'Michael Handford', written over a horizontal line.

Santiago de Chile, 18 de Agosto de 2005



"La madurez del hombre es haber vuelto a encontrar la seriedad con que jugaba cuando era niño". Federico Nietzsche.

Mi nombre es Andrés Tittarelli, el 6 de mayo de 1979 nací en la ciudad de Godoy Cruz, Mendoza, Argentina. Mis padres, también mendocinos, son Beatriz Elaskar y Jorge Tittarelli. Junto con mi hermana menor, Georgina, vivimos en Mendoza hasta que cumplí siete años de edad, luego nos fuimos a vivir con mi madre a Mar del Plata, una hermosa ciudad de la provincia de Buenos Aires. Allí realicé mis estudios de enseñanza media, en la Escuela de Educación Agropecuaria N°1 Nicanor Ezeiza, egresando el año 1996 con el título de Bachiller Agrario. En este período de mi vida generé amistades excelentes que hasta la fecha perduran.

En 1997 llegué a Chile para vivir con mi padre. Mis intenciones primeras fueron estudiar medicina, hasta que descubrí esta interesante carrera: Ingeniería en Biotecnología Molecular. Sonaba bien. Pues entonces, en 1998 comencé a estudiar esta carrera en la Facultad de Ciencias de la Universidad de Chile. Feliz y nostálgicamente, después de varios años de esfuerzo esta etapa ha concluido. Me queda, un asomo de conocimiento, grandes amigos, respetables colegas y un intrigante pero apasionante futuro.

A mi familia y amigos

AGRADECIMIENTOS

Un especial agradecimiento a mi director de tesis por su constante dedicación y apoyo, tanto a nivel profesional como personal.

A la doctora Lee Meisel por su visión crítica y ayuda en la realización de esta Tesis.

A todos los integrantes y ex integrantes del laboratorio de Genética Molecular Vegetal, especialmente a los doctores: Liliana Cardemil, Michael Handford, Lorena Marchant, Ana Mercado, Lorena Norambuena, Viviana Órdenes y Ariel Orellana. Además a los no doctores (y no por eso menos importantes): Rodrigo Álvarez, Enrique Armijo, Roberto Bastias, Andrea Bello, Manuel Bugueño, Rodrigo Caroca, Francisco Duarte, Beatriz Fonseca, Daniela Gonzalez, Mariela Gonzalez, Susana Gonzalez, Mariano Latorre, Nicolás Loira, Jonathan Maldonado, Luis Milla, Andrea Morales, Ricardo Nilo, Enzo Pantoja, Francisca Reyes, Bárbara Ribeck, Cecilia Rodríguez, Margarita Santiago, Carlos Saffie, Ricardo Tejos, Rafael Torres, Daniela Urbina, Sebastián Urbina, Jorge Verbusch, Alexander Vergara, Angélica Vega, Natalia Vergara y Paula Vizoso.

Esta tesis se desarrolló con el apoyo de:

- Iniciativa Genoma Chile, FDI G02 P1001
- Millenium Nucleus in Plant Cell Biology P02-009-F

ÍNDICE DE MATERIAS

TÍTULO	i
AUTOBIOGRAFÍA	ii
DEDICATORIA	iii
AGRADECIMIENTOS	iv
ÍNDICE DE TABLAS	viii
ÍNDICE DE FIGURAS	viii
LISTA DE ABREVIATURAS	ix
RESUMEN	xi
ABSTRACT	xiii
1.- INTRODUCCIÓN	1
2.- HIPÓTESIS	10
3.- OBJETIVOS	10
3.1.- OBJETIVO GENERAL	10
3.2.- OBJETIVOS ESPECÍFICOS	10
4.- MATERIALES Y MÉTODOS	11
4.1.- Genotecas de cDNA	11
4.2.- Agrupamiento de ESTs en contig	11

4.3.- Identificación de contigs expresados diferencialmente entre dos condiciones, utilizando el método estadístico de Audic y Claverie (método AC)	12
4.4.- Análisis de expresión génica multicondicional, agrupamiento jerárquico correlacional	14
4.5.- Anotación de 1.801 contigs deducida por homología contra secuencias de bases de datos públicas	16
4.6.- Análisis de expresión mediante transcripción reversa con reacción en cadena de la polimerasa (RT-PCR)	16
4.6.1.- Purificación de RNA mensajero	16
4.6.2.- Cuantificación de RNA mensajero	17
4.6.3.- Construcción de DNA complementario	17
4.6.4.- Reacciones en cadena de la polimerasa (RT-PCR)	18
4.7.- Clonamiento de promotores de genes inducidos por frío (taumatina, endoquitinasa y deshidrina) y harinosidad (lipoxigenasa)	19
4.7.1- Extracción de DNA desde hojas de durazno variedad O'Henry	19
4.7.2- Generación de genotecas genómicas	21
4.7.3.- Amplificación de putativas secuencias promotoras o Caminata Genómica (Genome Walker)	22
4.8.- Material microbiológico, medios y condiciones de cultivo	24
4.9.- Preparación de células de <i>E. coli</i> DH5 α competentes	24
4.10.- Transformación de <i>E. coli</i> DH5 α por estrés térmico	25
4.11.- Extracción de DNA plasmidial desde <i>E. coli</i> DH5 α	25
4.12.- Electroforesis en geles de agarosa para análisis de DNA	26
4.13.- Digestión de DNA con endonucleasas de restricción	27
4.14.- Ligación de DNA utilizando T4 DNA ligasa	27
 5.- RESULTADOS	 28
5.1.- Genotecas de cDNA, número de ESTs secuenciados y generación de contigo	28
5.2.- Identificación y análisis de contigs con niveles de expresión diferencial utilizando el método estadístico de Audic y Claverie	30

5.3.- Análisis de perfiles de expresión digital: identificación de contigs coexpresados	36
5.4.- Análisis de la expresión de taumatina, endoquitinasa, deshidrina y lipoxigenasa por RT-PCR	45
5.5.- Clonamiento de las secuencias de los posibles promotores de taumatina, endoquitinasa, deshidrina y lipoxigenasa	50
5.6.- Análisis de las secuencias de los posibles promotores de taumatina, endoquitinasa, deshidrina y lipoxigenasa	52
6.- DISCUSIÓN	57
7.- CONCLUSIONES	73
8.- PROYECCIONES	74
9.- REFERENCIAS	75

ÍNDICE DE TABLAS

Tabla 1. Descripción de las genotecas de cDNA de frutos de durazno, variedad O'Henry	29
Tabla 2. Número de ESTs de buena calidad por genoteca de cDNA	30
Tabla 3. Número de contigs formados por distintas cantidades de ESTs	30
Tabla 4. Factores de transcripción expresados preferencialmente en distintas genotecas	44
Tabla 5. Elementos regulatorios en <i>cis</i> (sitios de unión de factores de transcripción) encontrados en las secuencias promotoras aisladas	56

ÍNDICE DE FIGURAS

Figura 1. Distribución de contigs con niveles de expresión diferencial (método AC) según categorías funcionales	33
Figura 2. Agrupamiento jerárquico de contigs y genotecas de cDNA según sus perfiles de expresión	40
Figura 3. Patrones de expresión génica de los 10 clusters principales obtenidos mediante agrupamiento jerárquico	41
Figura 4. Distribución de los contigs analizados según sus perfiles de expresión y categorías funcionales asignadas	43
Figura 5. Niveles de expresión génica digital y análisis de expresión de los genes taumatina, endoquitinasa, deshidrina y lipoxigenasa mediante RT-PCR	49
Figura 6. Amplificación de secuencias promotoras de deshidrina, taumatina, endoquitinasa y lipoxigenasa mediante la técnica de Genome Walker	53
Figura 7. Secuencias de los promotores obtenidos mediante la técnica de Genome Walker para los genes deshidrina, taumatina, endoquitinasa y lipoxigenasa	54

LISTA DE ABREVIATURAS

ABRE	Elemento de respuesta a ácido abscísico
AC	Audic y Claviere (método estadístico)
AP	"Adaptor primer" (Partidor del adaptador)
BLAST	"Basic local alignment search tool" (herramienta básica de búsqueda de alineamientos locales)
CDNA	DNA complementario
CTAB	Bromuro de Cetil Trimetil Amonio
DMSO	Dimetil sulfóxido
DRE	Elemento de respuesta a deshidratación
dNTPs	Deoxi nucleotidos trifosfatos
dTTPs	Deoxi timidina trifosfato
E1	Estadio 1, frutos de duraznos no maduros
E2	Estadio 2, frutos de duraznos maduros jugosos
E3	Estadio 3, frutos de duraznos no maduros almacenados a 4°C
E4	Estadio 4, frutos de duraznos maduros harinosos
EDTA	Ácido Etilen-diaminotetra-acético
EGAD	"Expressed gene anatomy database" (Base de datos de expresión génica anatómica)
EST	"Expressed sequence tags" (Secuencias parciales de cDNA)
GSP	"Gene specific primer" (Partidor gen específico)
Kb	Kilobases
Mb	Megabases
MIPS	"Munich information center for protein sequences" (Centro de información de Munich para secuencias protéicas)
MRNA	RNA mensajero
MRNA poli A	mRNA poliadenilado
ODEPA	Oficina de estudios y políticas agrarias
p	Probabilidad de variación en expresión génica por azar (método AC)
pb	Pares de bases

pg	Picogramos
PCR	"Polimerase Chain Reaction" (Reacción en cadena de la polimerasa)
RNA	Ácido ribonucleico
RNAse	Ribonucleasa
rpm	Revoluciones por minuto
RT	"Reverse transcription" (Transcripción reversa)
<i>r_{uc}</i>	Coefficiente de correlación lineal de Pearson no centrado
SDS	"Sodium dodecyl sulphate" (Dodecil sulfato de sodio)
TAIR	"The Arabidopsis information resource" (Fuente de información de Arabidopsis)
UTR	"Untranslated region" (Región no traducida)

RESUMEN

Chile es uno de los principales países exportadores de frutas de carozo del Hemisferio Sur, siendo su principal mercado países del Hemisferio Norte, por lo cual las frutas de exportación son sometidas a prolongados períodos de almacenamiento y transporte refrigerado. Este tratamiento genera problemas en su calidad, siendo el más serio el daño por frío que sufren duraznos y nectarines. Uno de los principales signos de este problema es la harinosidad de los frutos. La mayoría de las investigaciones realizadas en este ámbito, sólo se han basado en el estudio de un grupo de enzimas participantes en la degradación de la pared celular, por lo tanto aún no se conocen otros componentes moleculares que puedan estar asociados a esta patología. En este trabajo de tesis, se realizó un análisis del perfil transcripcional global de cuatro estadíos de postcosecha de frutos de duraznos de una variedad susceptible a generar síntomas de harinosidad (O'Henry), dichos estadíos simulan el proceso normal de exportación de la fruta. El análisis se realizó a través de la construcción de cuatro genotecas de cDNA y la secuenciación de aproximadamente 50.000 ESTs. Se obtuvieron 43.491 ESTs de buena calidad que fueron ensamblados en un total de 4.983 secuencias de genes únicas (contigs). Se realizaron análisis de expresión digital utilizando dos aproximaciones estadísticas distintas y complementarias: el método estadístico de Audic y Claverie (AC) para detectar genes con diferencias significativas en expresión entre dos estadíos y el método de Agrupación Jerárquica ("Hierarchical Clustering") para agrupar genes con patrones de expresión similares en los cuatro estadíos. Estos análisis permitieron identificar varios genes relacionados al proceso de maduración normal (frutos jugosos), genes inducidos y reprimidos en frutos almacenados a 4°C por 21 días (genes que responden transcripcionalmente al estrés por frío) y genes relacionados al proceso de maduración anormal (generación de frutos harinosos). Los niveles de

expresión digital de tres genes expresados mayormente en frutos almacenados en frío (taumatina, deshidrina y endoquitinasa) y de un gen expresado mayormente en frutos harinosos (lipoxigenasa), fueron comprobados por RT-PCR. Utilizando la técnica de "Genome Walker" se aislaron las secuencias promotoras de estos genes y se identificaron *in silico* diferentes elementos regulatorios en *cis* que podrían explicar el modo de regulación transcripcional de estos genes.

ABSTRACT

Chile is one of the main exporters of stone fruits of the Southern Hemisphere, with countries in the Northern Hemisphere being the main market. Thus the exported fruits are put under prolonged periods of storage and cooled transport, generating problems in their quality, being chilling injury the most serious in peaches and nectarines. One of the main signs of this problem is the woolliness of the fruits. Most of the investigations in this area have only been based on the study of a group of enzymes involved in the degradation of the cell wall, therefore other molecular components that could be associated with this pathology are unknown. In this Thesis, the global transcriptional profile of a peach cultivar susceptible to wooliness symptoms (O' Henry) was analyzed at four postharvest stages simulating the normal process of fruit export. The analysis was made through the construction of four cDNA libraries and the sequencing of approximately 50,000 ESTs. 43,491 ESTs of good quality were obtained and these were assembled in a total of 4,983 sequences of unique genes (contigs). Analyses of digital expression were made using two different and complementary statistical approaches: the Audic and Claverie statistical method (AC) to detect genes with significant differences in level of expression between two stages and the hierarchical clustering method to group genes with similar patterns of expression in the four stages. These analyses allowed us to identify several genes related to the process of normal ripening (juicy fruits), genes induced and repressed in fruits stored at 4°C for 21 days (cold stress responsive genes) and genes related to the process of abnormal ripening (generation of woolly fruits). The levels of digital expression of three genes expressed mainly in fruits stored in the cold (taumatin, dehydrin and endochitinase) and of a gene expressed mainly in woolly fruits (lipoxygenase), were verified by RT-PCR. The promoters of these genes were isolated using the Genome Walker strategy. *In silico* analysis of

the promoter sequences identified *cis* acting regulatory elements that could explain the transcriptional regulation of each gene.

1.- INTRODUCCIÓN

Chile es un importante proveedor de frutas de carozo de nivel mundial, ocupando en varios casos y particularmente en duraznos y nectarines, el primer lugar como oferente del Hemisferio Sur (ODEPA, www.odepa.gob.cl). Chile es favorecido por inmejorables condiciones agro-ecológicas que facilitan la producción de este tipo de frutas, como son, infraestructura e inversiones establecidas, disponibilidad de suelos, condiciones sanitarias propicias, ausencia de precipitaciones durante el desarrollo de la fruta, suficiente frío invernal y régimen de temperaturas, entre otras. Si a esto le sumamos la perecibilidad intrínseca de los duraznos y nectarines, se logra explicar la superioridad de las exportaciones chilenas a mercados de contra-estación. Sin embargo, la distancia de Chile a los principales mercados consumidores de fruta (Estados Unidos, Asia y países de la Unión Europea) implica un prolongado período de transporte marítimo refrigerado (mínimo de 15 días), necesario para prevenir la sobremadurez y senescencia de la fruta, aumentando de esta manera el tiempo de vida de mercado. Este manejo puede causar serios desórdenes fisiológicos, denominados genéricamente como daño por frío o descomposición interna. En duraznos, el síntoma más pronunciado del daño por frío se conoce como harinosidad, caracterizado principalmente por falta de jugo y textura harinosa de la fruta cuando se deja madurar a temperatura ambiente posterior al período de almacenamiento (Lill y col., 1989). Este problema conlleva a una disminución de la calidad final de la fruta que, desafortunadamente, no puede ser detectada durante el proceso de cosecha, transporte o distribución, sino que básicamente por el consumidor. Este hecho está haciendo que Chile esté perdiendo liderazgo

como exportador de duraznos, debido a un creciente rechazo de los mercados consumidores más exigentes, lo que se traduce en una disminución en los precios y limitación de los volúmenes a exportar.

La susceptibilidad a desarrollar harinosidad está determinada por una serie de factores, siendo la variedad uno de los componentes más decisivos (Von Mollendorf, 1987; Crisosto y col., 1999), indicando que el trasfondo genético puede ser un factor clave en el desarrollo de estos síntomas. Igualmente, el régimen de temperatura al que es sometida la fruta resulta crucial en la inducción de la expresión de estos desórdenes fisiológicos. Por ejemplo, manteniendo las frutas a 0°C se asegura un mayor tiempo de almacenamiento sin la ocurrencia de sobremadurez, sin embargo, favorece la posterior ocurrencia de los síntomas de la harinosidad. Comparativamente, un rango de temperatura más elevado, entre 2-8°C, determina incluso una mayor expresión de los síntomas de harinosidad luego de un tiempo más breve de exposición (Ben-Arie y Lavee, 1971). Temperaturas de 10°C y superiores no inducen la expresión de daño por frío (Anderson, 1979; Ju y col., 2000), pero no permiten un adecuado manejo de la fruta para un almacenamiento/transporte como el requerido para exportaciones a mercados distantes. Además, se ha observado que el grado de harinosidad incrementa sustancialmente a medida que se aumenta el tiempo de almacenamiento en frío, siendo mínimo después de una a dos semanas y máximo después de tres y cuatro semanas (Brummell y col., 2004a).

Desde hace tiempo se ha postulado que la harinosidad en duraznos y nectarines estaría relacionada con alteraciones en la actividad de enzimas pectolíticas (hidrolasas) responsables de la degradación de la pared celular

que acompaña normalmente el ablandamiento y maduración en este tipo de frutos (Ben-Arie y Lavee, 1971; Ben-Arie y Sonego, 1980). Estas alteraciones generarían pectinas insolubles con un tamaño mayor al normal y con bajos niveles de esterificación. Este tipo de pectinas sería capaz de producir geles que retendrían fuertemente el agua libre, provocando los síntomas de harinosidad en la pulpa del fruto.

Existen ciertas discrepancias entre los estudios que sugieren esto último como la causa de la harinosidad. Al compararse con frutos jugosos, la harinosidad ha sido asociada con una reducción (Buescher y Furmanski, 1978), un aumento (Ben-Arie y Sonego, 1980; Brummell y *col.*, 2004a) o una invarianza (Obenland y Carroll, 2000; Zhou y *col.*, 2000a) en la actividad de la enzima pectina metil esterasa. Similarmente, hay estudios que muestran una disminución (Zhou y *col.*, 2000a) o una invarianza (Artes y *col.*, 1996; Brummell y *col.*, 2004a) en la actividad de la enzima exo-poligalacturonasa en frutos harinosos en comparación a frutos jugosos. Estos antecedentes, más otras publicaciones (Zhou y *col.*, 2000b, 2001; Dong y *col.*, 2001) sugieren que el problema de la harinosidad no es sólo atribuible a un desbalance en la degradación de pectinas, sino que podría tener una génesis más compleja, en la cual otros componentes aún no descritos podrían estar actuando.

Una manera de descubrir múltiples componentes que pudieran estar relacionados al problema de la harinosidad, sería realizar un análisis global de los cambios moleculares asociados con este fenómeno, esto es, niveles de expresión génica y acumulación de proteínas. Existen diferentes aproximaciones experimentales para realizar análisis de expresión génica globales. Las tecnologías de macro y microarreglos permiten medir la

expresión de cientos a miles de genes simultáneamente en un único experimento (Shalon y *col.*, 1996; Phimister, 1999). Una de las principales limitaciones de estas técnicas es la tendencia de los clones de cDNA y oligos a generar artefactos de hibridización, especialmente hibridización cruzada entre miembros de familias génicas altamente relacionados (Kuo y *col.*, 2002; Lipshutz y *col.*, 1999). Otra limitación de estas técnicas es el desarrollo de estructuras secundarias estables de las sondas (Southern y *col.*, 1999).

Una aproximación experimental alternativa es el uso de datos de ESTs (expressed sequences tags, o secuencias parciales de cDNAs; Adams y *col.*, 1991), que además de ser un eficiente método para el descubrimiento de genes, puede proveer información sobre niveles de expresión génica globales (Ewing y *col.*, 1999; Ewing y Claverie, 2000; Ohlrogge y Benning, 2000; Ojihara y *col.*, 2003). La deducción de niveles de expresión génica utilizando datos de ESTs, está basada en la suposición de que si el gen X es altamente expresado, es decir existen altos niveles de mRNA_x, entonces los cDNAs contruidos a partir del mRNA_x debieran ser abundantes en una genoteca de cDNA no normalizada hecha a partir del tejido donde el gen X es expresado. Luego de la secuenciación al azar de un gran número de clones de la genoteca de cDNA, el conteo del número de ESTs que corresponden al mRNA del gen X puede proveer una estimación de la abundancia del mRNA_x en la población original. Por lo tanto, las variaciones en la frecuencia de ESTs en distintas condiciones, estadios de desarrollo, tejidos u órganos (especificadas por el origen de las genotecas de cDNA), son usadas como base para la determinación de los niveles de expresión génica de uno o varios genes de un genoma.

El punto crítico para inferir diferencias en expresión génica utilizando las frecuencias relativas de ESTs, es la eficiencia del método estadístico que se utilice. Debido a que los experimentos de secuenciación al azar de gran cantidad de ESTs de diferentes genotecas de cDNA no son realizados en triplicados, los errores estándar asociados con cada una de las medidas de expresión (número de ESTs por gen) no pueden ser estimados por su dispersión y ninguno de los test estándar que requieren estimación de la varianza pueden ser usados. La naturaleza de este tipo de experimentos permite el uso de la distribución de Poisson como modelo de estructura de datos (Audic y Claverie, 1997). En base a este último punto, para identificar genes expresados diferencialmente entre dos condiciones, el método desarrollado por Audic y Claverie (método AC; Audic y Claverie, 1997) es uno de los más adecuados (Claverie, 1999; Romualdi y col., 2001). Este método ha sido ampliamente utilizado para detectar diferencias en expresión génica en diversos sistemas biológicos. Por ejemplo, White y col., 2000 utilizando datos de ESTs provenientes de genotecas de cDNA de *Arabidopsis thaliana*, reconstruyeron la vía metabólica responsable de convertir los carbohidratos en los aceites generados en el transcurso del desarrollo de las semillas. Prade y col., 2001 utilizando ESTs de dominio público, originados de diferentes genotecas de cDNA de diversas especies fúngicas, identificaron grupos de genes importantes para el desarrollo de tipo asexual en *Aspergillus nidulans*. En un trabajo más actual, Fei y col., 2004, utilizaron el método estadístico AC para detectar grupos de genes expresados diferencialmente en el proceso de maduración de frutos de tomate, comparando colecciones de ESTs de frutos de estadio maduro verde (mature green) y maduro rojo (breaker). Estos

trabajos y otros similares (Andrews y col., 2000; Iyengar y col., 2002; Ronning y col., 2003), sugieren un alto grado de credibilidad del método AC para detectar genes expresados diferencialmente en dos condiciones distintas, utilizando como información una gran cantidad de ESTs generados de diferentes genotecas de cDNA.

Los datos derivados de un gran número de ESTs de distintas genotecas de cDNA, además de permitir la identificación de genes expresados diferencialmente entre dos condiciones, permiten inferir relaciones funcionales entre genes mediante la identificación matemática de similitudes significativas entre los patrones de expresión globales de cada uno de los genes en varias condiciones. Es decir, los datos de ESTs derivados de diversas genotecas de cDNA aportan información relativa a la coexpresión génica y por lo tanto, revelan genes interrelacionados funcionalmente. Existen varias aproximaciones estadísticas para identificar grupos de genes coexpresados en múltiples condiciones (Claverie, 1999). Eisen y col., 1998, desarrollaron un método para detectar grupos de genes coexpresados, utilizando datos de ESTs de genotecas de cDNA de la levadura *Saccharomyces cerevisiae*. En dicho método se calcularon los coeficientes de correlación lineal de Pearson como medida de similitud para cada uno de los pares de perfiles de expresión en diferentes genotecas de cDNA. Posteriormente se realizó el agrupamiento jerárquico de los perfiles de expresión según la distancia calculada y a continuación los resultados son mostrados como un mapa de colores y un dendrograma representando las relaciones de similitud entre los perfiles de expresión de los genes y las genotecas de cDNA. En muchos casos, los genes implicados en el mismo proceso biológico o codificantes para

subunidades de complejos macromoleculares, poseían perfiles de expresión correlacionados, y por lo tanto, fueron agrupados en un mismo "cluster" jerárquico. Utilizando este tipo de método, Ogiwara y *col.*, 2003 identificaron genes coexpresados en diferentes etapas del ciclo de vida del trigo. Utilizando datos de ESTs de 27 genotecas de cDNA de diferentes tejidos/tratamientos de tomate y colecciones de ESTs de *Arabidopsis* y uva, Fei y *col.*, 2004 realizaron un análisis de agrupación jerárquica y observaron que es posible utilizar este tipo de análisis para comparar perfiles de expresión entre diferentes especies. Una vez identificados los perfiles de expresión génica globales, el paso siguiente es la interpretación de los resultados, esto es, relacionar los perfiles de expresión de los genes con las funciones de las proteínas que estos codifican. La función putativa de un gen puede ser deducida utilizando algoritmos computacionales (BLAST, Basic Local Alignment Search Tool, Altschul y *col.*, 1990), que permiten la comparación de secuencias de genes desconocidos con grandes bases de datos públicas de genes con función conocida. En plantas, la disponibilidad de la secuencia completa del genoma de *Arabidopsis thaliana* (<http://www.arabidopsis.org/cgi.html>; *Arabidopsis* Genome Initiative, 2000), está permitiendo aumentar ampliamente el conocimiento de las funciones de genes de esta especie y de otras especies vegetales de mayor importancia económica (Mysore y *col.*, 2001).

Por otro lado, los genes que tienen un patrón de expresión similar, pueden estar siendo regulados transcripcionalmente por mecanismos moleculares similares. Este tipo de coordinación es determinado por secuencias regulatorias, como son los promotores, los cuales están localizados en el genoma, generalmente río arriba de la región codificante del mRNA. Se ha

observado que promotores de genes que son expresados simultáneamente contienen en su secuencia elementos de unión a factores de transcripción en común (Ettwiller y col., 2003; Rombauts y col., 2003; Maruyama y col., 2004). Un ejemplo de esto es la conservación de los elementos de respuesta a deshidratación (DRE) y a ácido abscísico (ABRE) en promotores de muchos genes inducidos por frío en *Arabidopsis* y en otras plantas como tabaco, *Brassica* spp. y arroz (Yamaguchi-Shinozaki y Shinozaki, 1994; Bonetta y McCourt, 1998; Kasuga y col., 1999; Maruyama y col., 2004). Por lo tanto, la combinación del análisis de expresión génica más la identificación y análisis de promotores puede contribuir a la elucidación de la participación de un gen en determinada función biológica (Werner, 2002).

Como se mencionó anteriormente, los duraznos (*Prunus persica*) son económicamente importantes para Chile. Adicionalmente, el durazno es considerado como modelo de estudio de las especies de la familia de las Rosaceae, entre las que se encuentran manzana (*Malus domestica*), pera (*Pyrus communis*) y frutilla (*Fragaria* spp.), entre otras. Esto debido a que el tamaño de su genoma es relativamente pequeño (tamaño haploide de 300 Mb, Baird y col., 1994), aproximadamente el doble que el de *Arabidopsis*, tiene un período juvenil relativamente corto (2-3 años) y existe gran cantidad de información de tipo genética y genómica, como son mapas de marcadores moleculares (Georgi y col., 2002). La suma de estas características han planteado la necesidad de la formación de un consorcio internacional para realizar, en un futuro cercano, la secuenciación del genoma de durazno (Abbott, G.; www.genome.clemson.edu/gdr/conference/).

Por estos motivos, la secuenciación y análisis de gran cantidad de ESTs de diferentes genotecas de cDNA de frutos de durazno sometidos a diferentes tratamientos de postcosecha, no sólo puede dar indicios de las posibles causas moleculares del problema de la harinosidad, sino que también estaría contribuyendo a aumentar la información de genómica funcional de una especie de importancia científica y económica.

Con el objeto de identificar grupos de genes que pudiesen estar relacionados a la generación de la harinosidad en duraznos, en el presente trabajo de tesis, se analizaron los perfiles de expresión digital de cuatro genotecas no normalizadas de cDNA. Estas genotecas representan cuatro estadios de postcosecha de frutos de la variedad O'Henry (variedad muy susceptible a generar harinosidad). Los cuatro estadios simulan el proceso normal de exportación de la fruta a mercados distantes y por lo tanto, el proceso de generación de los síntomas de la harinosidad. Los análisis de expresión digital permitieron identificar genes con niveles de expresión diferencial entre los cuatro estadios evaluados. Además, para estudiar posibles eventos de correulación transcripcional, se clonaron y analizaron las secuencias de promotores de tres genes coexpresados preferencialmente en frutos durante el almacenamiento en frío (taumatina, deshidrina y endoquitinasa) y de un gen expresado preferencialmente en frutos harinosos (lipoxigenasa), evaluando la presencia de motivos regulatorios conservados entre estos promotores.

Por lo tanto, según los antecedentes planteados, en este trabajo se planteó la siguiente hipótesis:

2.- HIPÓTESIS

- Los niveles globales de expresión génica durante la vida de postcosecha de frutos de duraznos revelan genes candidatos relacionados al proceso de la harinosidad causada por frío.

3.- OBJETIVOS

3.1.- Objetivos Generales

- Analizar los perfiles de expresión globales en cuatro estadios de postcosecha en frutos de duraznos.
- Aislar promotores de genes de duraznos regulados por frío y expresados en frutos harinosos.

3.2.- Objetivos Específicos

1. Identificar grupos de genes expresados diferencialmente entre cuatro estadios de postcosecha de durazno utilizando los perfiles de expresión digital mediante el método estadístico AC.
2. Identificar grupos de genes expresados coordinadamente en cuatro estadios de postcosecha de durazno realizando el agrupamiento jerárquico de los perfiles de expresión digital.
3. Aislar y analizar *in silico* promotores de genes regulados positivamente por frío y expresados en frutos de durazno harinosos.

4.- MATERIALES Y MÉTODOS

4.1. - Genotecas de cDNA

En el presente trabajo de tesis se analizaron ESTs provenientes de cuatro genotecas de cDNA, generadas desde mRNA aislado de mesocarpo de frutos de duraznos (variedad O'Henry), de cuatro estadios de postcosecha distintos: E1, frutos de cosecha salidos de empaque ("packing"); E2, frutos salidos de empaque más cuatro días a 20°C (frutos maduros y jugosos); E3, frutos salidos de empaque y almacenados por 21 días a 4°C (empaque + 4°C) y E4, frutos salidos de empaque, almacenados por 21 días a 4°C más cuatro días a 20°C (frutos maduros y harinosos). Se utilizaron frutos de la variedad O'Henry debido a que fue la variedad con mayor índice de harinosidad bajo nuestras condiciones experimentales, es decir, mayor número de frutos con porcentajes de jugo extraíble bajo el 10% (p/v), de ocho variedades analizadas en la temporada del año 2003 (Campos y col., datos no publicados).

En nuestro laboratorio se purificó el RNA total de frutos de cada uno de los estadios descritos en el punto anterior y se generaron cuatro "pooles" formados por cantidades equitativas de RNA de 10 frutos por estadio (Meisel y col., 2005). El aislamiento del mRNA, la construcción de las genotecas y la secuenciación de los cDNAs, fueron realizados por la empresa Agencourt Bioscience Corporation (www.agencourt.com; USA).

4.2.- Agrupamiento de ESTs en contigs

En total se secuenciaron 50.625 ESTs de las cuatro genotecas de cDNA, aproximadamente 12.500 ESTs por cada genoteca. Luego del

enmascaramiento de las secuencias provenientes de vectores y la eliminación de secuencias de mala calidad, se obtuvieron 43.491 secuencias de buena calidad (definidas por un valor Phred Q > 20; Ewing y Green, 1998). Para definir las secuencias de ESTs que provienen de un mismo gen, los 43.491 ESTs de buena calidad fueron agrupados en contigs utilizando el programa Phrap bajo las condiciones de minmatch 50 y minscore 50 (Green, P., <http://bozeman.mbt.washington.edu/Phrap.docs/Phrap.html>). Este proceso generó 4.983 contigs constituidos por 36.598 ESTs (Loira y col., datos no publicados). Las secuencias que no ensamblaron en ningún contig fueron llamados singletons y estos no fueron considerados en este trabajo. Utilizando el programa de visualización de ensamblajes Consed (Gordon y col., 1998), se revisó el detalle de aproximadamente 200 contigs, para descartar cualquier tipo de artefactos en los ensamblajes, como por ejemplo, ensamblajes quiméricos.

4.3.- Identificación de contigs expresados diferencialmente entre dos estadíos utilizando el método estadístico de Audic y Claverie (método AC)

Los 4.983 contigs generados como se describe en 4.2., están formados por el ensamblaje de ESTs provenientes de las cuatro genotecas de cDNA. Cada contig puede estar formado por distintas cantidades de ESTs de cada genoteca, como se muestra de forma general en el siguiente esquema:

	Condición 1 (Genoteca 1) N_1	Condición 2 (Genoteca 2) N_2
Gen A (contig A)	X_{A1}	Y_{A2}
Gen B (contig B)	X_{B1}	Y_{B2}
Gen ... (contig ...)	$X_{...1}$	$Y_{...2}$
Gen n (contig n)	X_{n1}	Y_{n2}

En donde se ve que el hipotético contig A está formado por X_{A1} ESTs provenientes de la genoteca 1 (constituida por un total de N_1 ESTs) y por Y_{A2} ESTs de la genoteca 2 (constituida por un total de N_2 ESTs). X_{A1} e Y_{A2} describen la "intensidad" de expresión del gen A en la condición 1 y en la condición 2, respectivamente.

Para identificar genes o contigs de duraznos expresados diferencialmente entre dos estadios de postcosecha (dos genotecas de cDNA), se usó el método estadístico AC (Audic y Claverie, 1997). Dicho método define que la probabilidad de observar X e Y ESTs para un gen igualmente expresado en las condiciones N_1 y N_2 , respectivamente, está dada por:

$$p(y|x) = \left(\frac{N_2}{N_1}\right)^y \frac{(x+y)!}{x! y! \left(1 + \frac{N_2}{N_1}\right)^{(x+y+1)}}$$

Valores bajos de p corresponden a diferencias entre X e Y , que improbablemente se deban a variaciones por azar si el gen a analizar es expresado al mismo nivel en las condiciones N_1 y N_2 . Este método fue utilizado realizando todas las comparaciones posibles entre las genotecas, es

decir, E1 vs. E2, E1 vs. E3, E1 vs. E4, E2 vs. E3, E2 vs. E4 y E3 vs. E4. En dichas comparaciones, se definieron los contigs expresados diferencialmente cuando el valor p , definido en la fórmula anterior, fue menor a 0,01. Los valores p fueron calculados utilizando el programa Winflat (<http://igs-server.cnrs-mrs.fr>).

4.4.- Análisis de perfiles de expresión digital: identificación de contigs coexpresados mediante agrupamiento jerárquico correlacional

Para identificar aquellos contigs que tienen un patrón de expresión génica similar en las cuatro genotecas de cDNA, se analizaron 1.801 contigs constituidos por 5 o más ESTs (Ewing y col., 1999). Para ello se generó una tabla representando los perfiles de expresión de los 1.801 contigs (números de ESTs por genoteca); cada contig es colocado en una fila y cada genoteca en una columna diferente, por lo tanto, se obtiene una matriz de expresión de 1.801×4 . Si dos genes son expresados de manera coordinada, se espera que sus perfiles de expresión tengan forma similar, es decir, que dos series de números de EST sigan la misma tendencia de inducción-represión en las cuatro condiciones. Como medida de similitud o distancia entre perfiles de expresión de contigs y genotecas de cDNA, se utilizó el coeficiente de correlación lineal de Pearson no centrado (r_{uc}).

Siendo x e y vectores de n -componentes, para pares de cantidades (x_i, y_i) , $i = 1, \dots, n$, el coeficiente de correlación lineal de Pearson no centrado está dado por la siguiente fórmula:

$$r_{uc} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

\bar{x} media del vector **x**
 \bar{y} media del vector **y**

Valores de r_{uc} cercanos a 0 indican un bajo grado de correlación (o mayor distancia), mientras que valores cercanos a 1 indican un alto grado de correlación entre x e y (o menor distancia). En donde x e y son diferentes genes asociados con sus correspondientes cantidades de ESTs (x_1, \dots, x_n) e (y_1, \dots, y_n) medidas en las genotecas de cDNA 1, ..., n (con $n = 4$). Lo mismo es calculado para medir el grado de correlación entre las genotecas de cDNA, siendo ahora x e y diferentes genotecas asociadas con sus cantidades de ESTs (x_1, \dots, x_n) e (y_1, \dots, y_n) para los genes 1, 2, ..., n (con $n = 1.801$).

Luego, los contigs y genotecas de cDNA fueron agrupados jerárquicamente utilizando agrupamiento promedio ("average linkage"), en donde la distancia entre dos grupos (clusters) es calculada como la distancia promedio entre todos los pares de r_{uc} en los dos grupos diferentes (Eisen, 1999). De esta manera, se genera una nueva tabla en donde los genes y las genotecas con perfiles de expresión similares se agrupan juntos.

Los coeficientes de correlación lineal de Pearson no centrado fueron calculados utilizando el programa Cluster 3.0 (Eisen y col., 1998; <http://rana.lbl.gov/EisenSoftware.html>).

Luego, los resultados fueron visualizados y analizados utilizando el programa Java TreeView (<http://jtreeview.sourceforge.net>).

4.5.- Anotación de 1.801 contigs deducida por homología contra secuencias de bases de datos públicas

Para asignar una función putativa a cada uno de los 1.801 contigs utilizados en el análisis de expresión digital, estos fueron anotados electrónicamente por búsqueda de secuencias similares en tres bases de datos distintas, GenBank nr (<http://www.ncbi.nlm.nih.gov/>), PlantGDB (<http://plantgdb.org/pgrop/pgrop.php>) y TAIR (<http://www.arabidopsis.org>), utilizando el programa BLASTx. Un grado de similitud significativo entre las secuencias fue definido por un valor $e < 10^{-10}$ (Fei y col., 2004). Luego, los contigs fueron clasificados en diferentes categorías funcionales según los términos de procesos biológicos de Gene Ontology (Ashburner y col., 2000; www.geneontology.com).

4.6. - Análisis de expresión mediante transcripción reversa con reacción en cadena de la polimerasa (RT-PCR)

4.6.1. - Purificación de RNA mensajero

Para analizar los patrones de expresión de los genes codificantes para taumatina, endoquitinasa, deshidrina y lipoxigenasa, mediante la técnica de RT-PCR, se aisló mRNA de los cuatro "pooles" de RNA total que se utilizaron para construir las genotecas de cDNA (E1, E2, E3, E4), empleando el kit "Oligotex mRNA Spin-Column" de Qiagen (catálogo 70022, Oligotex Handbook, 05/2002). Se utilizaron 29 µg, 250 µg, 157 µg y 69 µg de RNA total de los "pooles" E1, E2, E3 y E4, respectivamente. Se utilizaron cantidades diferentes de RNA total por motivos de disponibilidad. Estas cantidades de

RNA total se llevaron a un volumen final de 250 μ L con agua destilada libre de RNAsa y se siguió el procedimiento descrito por el fabricante.

4.6.2. - Cuantificación de RNA mensajero

Para utilizar cantidades iguales de mRNA de los cuatro estadios (E1, E2, E3 y E4) en las reacciones de transcripción reversa, fue necesario cuantificar de manera precisa el mRNA aislado. Para ello se utilizó el sistema de detección de mRNA poli A ("Poly(A) mRNA Detection System") de Promega (catálogo K4040, www.promega.com/tbs/tb282/tb282.htmL), el cual es un sistema basado en la utilización de la enzima luciferasa diseñado para detectar específica y sensiblemente mRNA poliadenilado. La detección está basada en una serie de reacciones enzimáticas que resultan en una emisión de luz que es proporcional a la cantidad de mRNA poli A en la muestra. Se utilizaron triplicados de concentraciones conocidas de un mRNA poli A estándar (mRNA sintético de 1,2 Kb codificante para kanamicina), para la construcción de una curva de calibración, según las indicaciones del fabricante. Así mismo, se utilizaron diluciones (por triplicado) de mRNA de E1, E2, E3 y E4 y se calculó la concentración según el procedimiento descrito por el fabricante.

4.6.3. - Construcción de DNA complementario

Se utilizaron 5ng de mRNA de E1, E2, E3 y E4 como templado para reacciones de transcripción reversa. Se utilizó "ImProm-IITM reverse Transcription System" de Promega y partidores de óligos dTTPs para realizar los ensayos según instrucciones del fabricante (catálogo A3800), realizando las reacciones de transcripción reversa por 3 horas. Además, como control

interno de estas reacciones, a cada muestra de mRNA (E1, E2, E3 y E4) utilizada se le agregó 0,25 pg de un mRNA heterólogo (mRNA de 1,2 Kb codificante para kanamicina). Para descartar una posible contaminación con DNA genómico en las cuatro muestras de mRNA, se realizó un control consistente en reacciones de transcripción reversa sin la enzima transcriptasa reversa. Para confirmar si los fragmentos amplificados desde los cDNAs correspondían a los esperados, estos fueron clonados en el vector pBluescript y secuenciados.

4.6.4. - Reacciones en cadena de la polimerasa (RT-PCR)

Los partidores utilizados para amplificar un segmento del cDNA de taumatina fueron: F-PRT (5'-ACCTTGCCATCCTCTTCTT-3') y R-PRT (5'-AGAAATCTTGACCCCGTTC-3'), generando un fragmento de 346 pb. Para endoquitinasa: F-BEC (5'-GTCAGCAGCGTCGTTAGCTC-3') y R-BEC (5'-GAGTTGGATGGGTCCTCTGC-3'), generando un fragmento de 351 pb. Para deshidrina: F-DX2 (5'-CCAAACCAAAGCCAGTTTGATTCA-3') y R-DX2 (5'-CCAGGTTTTGTATGAGTGCCGTA-3') generando un fragmento de 259 pb. Para lipoxigenasa: F-LIP (5'-AAGGAGCTCTTGACGTTGGA-3') y R-LIP (5'-TGTTAACAGGTGGGAAAACC-3') generando un fragmento de 306 pb. Para amplificar un segmento de cDNA de actina los partidores utilizados fueron: F-ACTIN7 (5'-CCTTCCAGCAGATGTGGATT-3') y R-ACTIN7 (5'-AGATTAGGCAAGGCGAGGAT-3') generando un fragmento de 325 pb. Los partidores para amplificar un fragmento de 323 pares de bases del cDNA del mRNA control (kanamicina) fueron: "Upstream Control Primer" (5'-

GCCATTCTCACCGGATTCAGTCGTC-3') y "Downstream Control Primer" (5'-AGCCGCCGTCCCGTCAAGTCAG-3').

Los cDNAs generados (descrito en 4.6.3.), se diluyeron 100 veces y en las reacciones de PCR se utilizó 1 μ L de cada una de estas diluciones como templado. Las reacciones fueron realizadas en un volumen final de 20 μ L, conteniendo 0,5 μ M de partidores, dNTPs 0,2 mM, MgCl₂ 1,5 mM, 5 U de Taq DNA polimerasa y 1X de tampón correspondiente. Las condiciones de las reacciones fueron: denaturación inicial a 93°C por 5 minutos, seguido por número variable de ciclos (de 26, 28, 30, 32 y 34) de 93°C por 30 segundos, 55°C por 1 minuto y 1 minuto de elongación a 72°C, finalizando con un paso de elongación final a 72°C por 10 minutos.

4.7. - Clonamiento de promotores de genes inducidos por frío (taumatina, endoquitinasa y deshidrina) y harinosidad (lipoxigenasa)

4.7.1.- Extracción de DNA desde hojas de durazno variedad O'Henry

El procedimiento para aislar DNA genómico desde hojas de durazno (variedad O'Henry), se realizó adaptando el método descrito por Manubens *et al.*, 1999. Se recolectaron hojas de árboles de durazno O'Henry, ubicados en el Instituto de Investigaciones Agropecuarias (INIA) La Platina. Las hojas fueron lavadas con agua destilada, congeladas en nitrógeno líquido y almacenadas a -80°C hasta el momento de su utilización. Cuatro a seis hojas fueron molidas utilizando un mortero y nitrógeno líquido hasta obtener un polvo fino. Luego, el material vegetal pulverizado fue transferido a varios tubos de microcentrífuga de 2 mL (aproximadamente 0,1 g de tejido por tubo). Posteriormente se

agregó a cada tubo 0,6 mL de tampón de lisis (Tris-HCl 100 mM, pH 8,0; EDTA 50 mM, pH 8,0; SDS 4% (p/v); NaCl 0,3 M; β -mercaptoetanol 0,14 M) y 0,2 mL de tampón CTAB (Tris-HCl 100 mM, pH 8,0; EDTA 20 mM, pH 8,0; CTAB 8,4% (p/v)). Se incubaron las suspensiones a 65°C por 30 minutos mezclando ocasionalmente por inversión suave. Al finalizar la incubación, los tubos fueron puestos a temperatura ambiente y después de aproximadamente 5 minutos se agregó a cada uno 160 μ L de acetato de potasio 5 M y se dejaron en hielo por 10 minutos. A continuación se agregó a cada tubo 960 μ L de solución fenol:cloroformo:alcohol isoamílico en proporción 25:24:1 y se homogenizó la mezcla vigorosamente por inversión. Luego los tubos se centrifugaron a 14.000 rpm por 10 minutos para posteriormente recuperar 900 μ L de la fase acuosa (superior) y depositarla en nuevos tubos de microcentrífuga de 2 mL. Se agregó a cada tubo 900 μ L de solución fenol:cloroformo:alcohol isoamílico en proporción 25:24:1 y se mezclaron vigorosamente por inversión para finalmente centrifugarlos a 14.000 rpm por 10 minutos. Posteriormente, 750 μ L de la fase superior obtenida fueron transferidos a tubos de microcentrífuga de 1,5 mL nuevos. Para precipitar el DNA se agregó a cada tubo 750 μ L de isopropanol previamente enfriado a -20°C, se mezcló invirtiendo los tubos 7 a 10 veces y se incubaron por 30 minutos a -20°C. Inmediatamente se centrifugaron por 10 minutos a 14.000 rpm y 4°C, se descartó el sobrenadante y los pellets de DNA se lavaron con etanol 70%, posteriormente se centrifugaron a 14.000 rpm por 5 minutos, se eliminó el sobrenadante y se secaron los pellets de DNA por centrifugación al vacío (Speed Vac) por aproximadamente 5 minutos. Los pellets fueron resuspendidos en 50 μ L de tampón TE (Tris-HCl 10 mM, pH 8,0; EDTA 1 mM,

pH 8,0), esto fue facilitado incubando la solución a 65°C por 10 minutos. El RNA fue eliminado adicionando 0,15 mg de Ribonucleasa A e incubando por 40 minutos a 37°C. A continuación se agregó 0,2 mg de Proteinasa K y se incubó a 50°C por 20 minutos. Las fracciones contenidas en cada tubo se juntaron en un único tubo de microcentrífuga de 1,5 mL. A este se le agregaron 800 µL de fenol, se mezcló vigorosamente por inversión y se centrifugó a 14.000 rpm por 10 minutos. Luego se recuperó la fase acuosa y el DNA fue precipitado por la adición de 1/10 de volumen de 3 M acetato de sodio y 1 volumen de etanol a -20°C por una hora. Una vez precipitado el DNA, los tubos se centrifugaron a 14.000 rpm a 4°C por 10 minutos, descartando posteriormente el sobrenadante y lavando el pellet de DNA con etanol 70%. Una vez seco, el DNA se resuspendió en 100 µL de agua desionizada estéril. La calidad y concentración del DNA genómico fueron confirmadas al correr una alícuota de 1 µL en geles de agarosa al 0,7%.

4.7.2. - Generación de genotecas genómicas

Para aislar promotores de genes de duraznos inducidos por frío y en la condición de harinosidad que sean funcionales en frutos, se utilizó la técnica de "Genome Walker" o caminata genómica descrito por Siebert y *col.*, 1995. Para ello se construyeron genotecas genómicas, constituidas por fragmentos de DNA genómico de extremos romos fusionados a un adaptador de secuencia conocida. Estas genotecas representan el genoma completo del organismo en estudio. En este trabajo se construyeron 4 genotecas genómicas diferentes, utilizando cuatro endonucleasas de restricción de corte romo distintas (Mlsl, Pvull, EcoRV y Pael). Se digirió con cada una de estas

enzimas 0,25 µg de DNA genómico de durazno de la variedad O'Henry por 18 horas a 37°C en un volumen final de 500 µL y utilizando 80 unidades enzimáticas en cada reacción. A las dos horas de incubación, se agitaron las reacciones por 10 segundos utilizando un vortex a velocidad suave. Una vez digerido el DNA, se purificó mediante una extracción orgánica con 1 volumen de fenol, luego una extracción con 1 volumen de cloroformo y finalmente el DNA se precipitó con 2 volúmenes de etanol al 95%, 0,1 volumen de acetato de sodio 3 M (pH 4,5) y 20 µg de glicógeno. Luego de centrifugar el DNA a 14.000 rpm por 20 minutos a 4°C, el pellet obtenido se lavó con 100 µL de etanol 80% y finalmente fue resuspendido en 20 µL de tampón TE (Tris-HCl 10 mM, pH 8,0; EDTA 0,1 mM, pH 8,0; llevado a pH 7,5).

4.7.3. - Amplificación de putativas secuencias promotoras o Caminata Genómica (Genome Walker)

Para amplificar la secuencia 5' río arriba del inicio de la traducción de los genes inducidos por frío (endoquitinasa, taumatina y deshidrina) y en harinosidad (lipoxigenasa), se diseñaron los siguientes partidores específicos: para endoquitinasa, BEC-GSP1 (5'-TGCATTTCCAGCTTGCCTCCCACATTG-3') y BEC-GSP2 (5'-CTGAGATCCCTAACAGCAAAGCTAGGGATA-3'); para taumatina, PRT-GSP1 (5'-TGATTTTAGCTGCATGTGCACCTGAGAA-3') y PRT-GSP2 (5'-CGTCATGGAAATGTCTTAATTGGCTTGCTG-3'); para deshidrina, DX2-GSP1 (5'-ACCGGTTCCGGTGGTGGTGTGATGAACC-3') y DX2-GSP2 (5'-ACTCATCAGTCTTAGTAGGCTCGGGTGTT-3'); para lipoxigenasa, LIP-GSP1 (5'-GAAGAAAACAAATTGGGAGGAGGAGAA-3') y LIP-GSP2 (5'-GCGTGTTCCAAAGAACACAATTCAGTGCCTT-3'). Utilizando

como templado 1 μL de las cuatro genotecas genómicas, se realizaron sucesivas reacciones de PCR. Para el primer grupo de reacciones de PCR, se utilizó el partidor específico para el adaptador 1, AP1 (5'-GGATCCTAATACGACTCACTATAGGGC-3'), y los partidores específicos 1 (GSP1) para cada uno de los genes de interés; la concentración final de cada partidor en las reacciones de PCR fue de 0,2 μM en un volumen final de 50 μL . Se utilizó 5 U de la enzima "Synergy DNA polymerase" de Genecraft, haciendo "hot start" manualmente, es decir, se adicionó la enzima a la reacción en el momento que se alcanzó el primer pico de temperatura (93°C). Las condiciones para esta amplificación fueron: 1 ciclo a 93°C por 10 minutos, seguido de 7 ciclos de 93°C por 30 segundos y 72°C por 15 minutos, luego, 37 ciclos de 93°C por 30 segundos y 67°C por 15 minutos. Para el segundo grupo de reacciones de PCR ("nested PCR"), se utilizó el partidor específico para el adaptador 2, AP2 (5'-ACTATAGGGCACGCGTGGT-3'), y los partidores gen específicos 2 (GSP2). Como DNA templado se utilizó 1 μL de una dilución 1/50 de las primeras reacciones. Igual que en las reacciones anteriores, la concentración final de cada partidor fue de 0,2 μM en un volumen final de 50 μL y se utilizó 5 U de la enzima "Synergy DNA polymerase" haciendo "hot start" manualmente. Las condiciones para esta amplificación fueron: 1 ciclo a 93°C por 10 minutos, seguido de 5 ciclos (7 ciclos en el caso de deshidrina), de 93°C por 30 segundos y 72°C por 15 minutos, luego 20 ciclos (30 ciclos en el caso de deshidrina), de 93°C por 30 segundos y 67°C por 15 minutos. En los casos señalados, los amplificados se clonaron en el vector pBluescript o pGEMT-easy y fueron secuenciados.

4.8. - Material microbiológico, medios y condiciones de cultivo

La bacteria *Escherichia coli* cepa DH5 α , se creció en medio LB líquido estéril (1% triptona; 0,5% extracto de levadura; 1% NaCl) y en LB sólido estéril (LB líquido más 1,5% de agar). Las incubaciones se realizaron a 37°C por 16 horas y cuando se trató de medio líquido la velocidad de agitación fue de 250 rpm. El antibiótico de selección utilizado en este trabajo fue Ampicilina (100 μ g/mL), ya que los vectores utilizados poseían el gen que codifica para la enzima β -lactamasa, la cual confiere resistencia a este antibiótico.

4.9. - Preparación de células de *E. coli* DH5 α competentes

Las células competentes utilizadas en esta tesis se prepararon según Inoue y *col.*, 1990. A partir de una colonia individual de *E. coli* DH5 α , se inocularon 3 mL de medio LB y se incubó a 37°C por 16 horas con agitación de 250 rpm. Se agregó 1 mL de este cultivo a un volumen de 250 mL de medio LB. Las bacterias se crecieron a 25°C hasta una densidad óptica igual a 0,6 medida a 600 nanómetros de longitud de onda. Luego, las bacterias se incubaron por 10 minutos en hielo y se traspasaron a un tubo de centrifuga previamente frío. Las células se centrifugaron a 4.000 rpm por 10 minutos a 4°C. Se descartó el sobrenadante y el pellet se resuspendió en 20 mL de tampón TB frío (preparado con una solución PIPES 10 mM; CaCl₂ 15 mM; KCl 250 mM ajustado a pH 6,7 con KOH 5 N y agregando MnCl₂ hasta una concentración de 55 mM). Las bacterias se centrifugaron nuevamente a 4.000 rpm por 10 minutos a 4°C. Posteriormente, se descartó el sobrenadante y el pellet se resuspendió en tampón TB frío, se adicionó DMSO hasta una concentración final de 7% y se separaron alícuotas de 200 μ L en tubos de microcentrifuga de

1,5 mL previamente enfriados a -20°C. Las bacterias se congelaron en nitrógeno líquido y se almacenaron a -80°C hasta su uso.

4.10. - Transformación de *E. coli* DH5α por estrés térmico

Para transformar las células de *E. coli* DH5α con los vectores generados en este trabajo, se incubaron en hielo 50 µL de células competentes con aproximadamente 10 ng de DNA plasmidial por 30 minutos. Posteriormente, las células se traspasaron a 42°C por 1,5 minutos para producir el estrés térmico y luego se enfriaron en hielo por 2 minutos. A continuación se agregó 1 mL de medio LB y se incubaron las bacterias a 37°C con agitación de 250 rpm por 1 hora. Al finalizar el tiempo de incubación, se colectaron las células por centrifugación por un período de 20 segundos a 14.000 rpm descartándose el sobrenadante mediante inversión suave del tubo, lo cual deja un remanente de aproximadamente 160 µL. Las células se resuspendieron por pipeteo suave y se plaquearon en LB-agar, suplementado con ampicilina (100 µg/mL).

4.11. - Extracción de DNA plasmidial desde *E. coli* DH5α

El procedimiento se realizó según Xiang *et al.*, 1994. Se crecieron los clones de *E. coli* que portaban los plasmidios de interés en 3 mL de medio LB suplementado con ampicilina (100 µg/mL) a 37°C por 16 horas, con agitación de 250 rpm. Luego de la incubación, se traspasaron 1,6 mL del cultivo a un tubo de microcentrífuga y se centrifugó a 14.000 rpm por 20 segundos. Se descartó todo el sobrenadante y se resuspendió completamente el pellet de bacterias en 100 µL de GTE (glucosa 50 mM; Tris-HCl 25 mM, pH 8,0; EDTA

10 mM, pH 8,0) mediante una mezcla por vortex por aproximadamente 30 segundos. Se añadieron 200 μ L de tampón de lisis (NaOH 0,2 M; SDS 1%) y se invirtió el tubo dos veces para mezclar. Luego se adicionaron 150 μ L de solución acetato de potasio 5M (preparado mezclando 60 mL de acetato de potasio 5M; 11,5 mL de ácido acético glacial y 28,5 mL de agua destilada) y se agitó vigorosamente. Esta mezcla se centrifugó por 10 minutos a 14.000 rpm. El sobrenadante se recuperó cuidadosamente con una micropipeta, evitando remover el precipitado, y se depositó en un tubo de microcentrifuga nuevo. Los ácidos nucleicos se precipitaron utilizando 1 mL de etanol 95% y se mezcló invirtiendo un par de veces el tubo. Se centrifugó durante 1 minuto para obtener el precipitado que corresponde a DNA plasmidial y RNA. El sobrenadante fue descartado y el pellet se secó a temperatura ambiente o en "Speed Vac". Posteriormente el pellet se resuspendió en un volumen de agua desionizada variable (de 30 a 100 μ L), conteniendo 10 μ g/mL de Ribonucleasa A. El DNA obtenido en este punto pudo ser rutinariamente utilizado en reacciones de PCR y digestiones por endonucleasas.

4.12. - Electroforesis en geles de agarosa para análisis de DNA

Se utilizaron geles de agarosa desde 0,7% a 2,5%, según cada caso, en tampón TAE 1X (Tris-acetato 40mM; EDTA 1mM pH 8,0) con bromuro de etidio (0,1 mg/mL). Como estándar de peso molecular se utilizaron "1kb DNA Ruler", "1Kb Plus DNA Ladder", "GeneRuler 100pb DNA Ladder" y λ Hind III de Fermentas. Los geles fueron corridos en una cámara BioRad Minisub Cell G1, utilizando una fuente de poder BioRad Powerpac 300 de 50 a 100 V de potencia.

4.13. - Digestión de DNA con endonucleasas de restricción

Las digestiones se realizaron utilizando 2 unidades de enzima por cada 0,5 µg de DNA, aproximadamente, en un volumen final mínimo de 15 µL e incubaciones de a lo menos 2 horas a la temperatura recomendada para cada enzima de restricción.

4.14. - Ligación de DNA utilizando T4 DNA ligasa

Las ligaciones de fragmentos de DNA purificados con vectores linearizados, se realizaron con proporciones de aproximadamente 8:1, respectivamente. Se utilizaron entre 2 a 4 unidades de T4 DNA ligasa de New England BioLabs, en volúmenes finales entre 7 a 15 µL a 16°C por 18 horas, utilizando baños termoregulados de agua o termocicladores.

5.- RESULTADOS

5.1.- Genotecas de cDNA, número de ESTs secuenciados y generación de contigs

En el proyecto dentro del cual se enmarca este trabajo de tesis (www.genomavegetal.cl), se secuenciaron 50.625 ESTs provenientes de cuatro genotecas de cDNA no normalizadas de frutos de duraznos de la variedad O'Henry sometidos a diferentes tratamientos de postcosecha. Dichos tratamientos simulan el proceso de exportación de las frutas (almacenamiento en frío y generación de los síntomas de la harinosidad) y el proceso de maduración normal, es decir, cuando las frutas no son sometidas a una cadena de frío. Se utilizó la variedad O'Henry debido a que fueron frutos muy susceptibles a tornarse harinosos (ver materiales y métodos). En la tabla 1 se describen cada una de las genotecas de cDNA.

De cada genoteca se secuenciaron cantidades similares de ESTs (aproximadamente 12.500). Al eliminar las secuencias de mala calidad utilizando el programa Phred, se obtuvieron en total 43.491 ESTs de buena calidad (ver materiales y métodos). Como se observa en la tabla 2, la proporción de ESTs de buena calidad siguió siendo similar para cada una de las genotecas de cDNA utilizadas. Estos ESTs fueron ensamblados utilizando el programa Phrap bajo las condiciones de minmatch 50 y minscore 50 (Green, P., <http://bozeman.mbt.washington.edu/Phrap.docs/Phrap.html>). Estos estrictos criterios, permiten ensamblar en un mismo contig ESTs provenientes del mismo transcrito previniendo la formación de ensamblajes quiméricos. Además, la estrictez de los criterios del ensamble permite separar ESTs que

sean de genes de la misma familia génica y formas alternativas del mismo gen, ya sean casos de empalmes alternativos o polimorfismos (Burke y col., 1998). En la tabla 3 se muestra el número de contigs formados agrupando los ESTs de las cuatro genotecas por el método de Phrap. Un total de 4.983 contigs fueron generados. Debido a las propiedades del ensamble, cada uno de los contigs representa un único gen (Burke y col., 1998; Ohihara y col., 2003). Además, en la tabla 3 se muestra el número de contigs formados por distintas cantidades de ESTs, indicando de cierta manera, genes con diferentes niveles de expresión (Ewing y col., 1999).

Tabla 1. Descripción de las genotecas de cDNA de frutos de durazno, variedad O'Henry.

Nombre	Descripción
E1	Construida a partir de mRNA de mesocarpo de 10 frutos de cosecha procesados en "packing" (empaque; frutos no maduros).
E2	Construida a partir de mRNA de mesocarpo de 10 frutos procesados en "packing" y dejados madurar a 20°C por 4 días (frutos maduros y jugosos).
E3	Construida a partir de mRNA de mesocarpo de 10 frutos procesados en "packing" y sometidos a una cadena de frío a 4°C por 21 días (frutos no maduros almacenados a 4°C).
E4	Construida a partir de mRNA de mesocarpo de 10 frutos procesados en "packing", sometidos a una cadena de frío a 4°C por 21 días y dejados madurar a 20°C por 4 días (frutos maduros y harinosos).

Tabla 2. Número de ESTs de buena calidad por genoteca de cDNA.

TOTAL		ESTs E1 (%)	ESTs E2 (%)	ESTs E3 (%)	ESTs E4 (%)
ESTs analizados (Q > 20)	43.491	10.961 (25,2) ^a	10.728 (24,7) ^a	11.142 (25,6) ^a	10.660 (24,5) ^a

^aPorcentaje de ESTs de cada genoteca con respecto a los 43.491 ESTs de buena calidad (valor Phred Q > 20).

Tabla 3. Número de contigs formados por distintas cantidades de ESTs.

# ESTs por Contig	<5	≥ 5	5 < X ≤ 10	10 < X < 25	≥ 25
# Contigs	3.182	1.801	674	540	286

5.2.- Identificación y análisis de contigs con niveles de expresión diferencial utilizando el método estadístico de Audic y Claverie

Un número relativamente alto de ESTs provenientes de un tejido o tratamiento específico, pueden indicar los niveles de expresión de varios genes en dichos tejidos o tratamientos, tomando en cuenta la abundancia relativa de un determinado cDNA dentro de la población total de ESTs (Rafalski y col., 1998; Ewing y col., 1999; Mekhedov y col., 2000). El problema básico relacionado con este tipo de análisis, es la eficiencia de los tests estadísticos utilizados para detectar cuáles genes están siendo expresados diferencialmente en las condiciones estudiadas (Claverie, 1999). Romualdi y colaboradores, 2001, compararon diferentes métodos estadísticos usados comúnmente en la detección de genes expresados diferencialmente, y sus resultados sugirieron

que el método estadístico de Audic y Claverie (método AC; Audic y Claverie, 1997), es el método más eficiente en detectar diferencias en expresión génica al comparar dos condiciones diferentes. El método AC ha sido utilizado por varios investigadores y con datos de ESTs de diversos sistemas biológicos (Fei y col., 2004; Ronning y col., 2003; White y col., 2000; Andrews y col., 2000; Iyengar y col., 2002).

En este trabajo, se utilizó esta aproximación experimental (también conocida como "Northern electrónico") para identificar genes expresados diferencialmente entre dos estadios de postcosecha distintos, analizando estadísticamente (método AC) las frecuencias de ESTs por cada uno de los genes (contigs) en cada genoteca.

Al comparar el estadio E1 con E2, se están observando aquellos genes relacionados al proceso de maduración normal. Se identificaron 39 contigs inducidos en este proceso (mayor expresión en E2, $p < 0,01$) y 44 contigs que son reprimidos (mayor expresión en E1, $p < 0,01$). Comparando el estadio E1 con E3, se observan los genes cuya expresión es modificada durante el almacenamiento de los frutos por 21 días a 4°C. En este caso, se identificaron 63 contigs que se inducen (mayor expresión en E3, $p < 0,01$) y 60 contigs que se reprimen (mayor expresión en E1, $p < 0,01$). Cuando se compara el estadio E3 con E4 se están observando los genes relacionados al proceso de maduración "anormal", el cual genera frutos harinosos. En este proceso se identificaron 57 contigs que son inducidos (mayor expresión en E4, $p < 0,01$) y 84 contigs que son reprimidos (mayor expresión en E3, $p < 0,01$). Todos estos contigs fueron subdivididos en categorías funcionales de acuerdo a la función putativa asignada por homología utilizando Blastx (figura 1).

De los genes inducidos en el proceso de maduración, identificados a través de esta aproximación experimental (figura 1A), varios han sido asociados previamente a la maduración de frutos, como son los casos de ACC oxidasa (biosíntesis de etileno; Barry y *col.*, 1996), endopoligalacturonasas (desensamble de la pared celular; Trainotti y *col.*, 2003), proteína E6 (estructura de la pared celular; Trainotti y *col.*, 2003), proteínas similares a alérgenos (estructura de la pared celular; Brenna y *col.*, 2004), entre otros. Como se puede observar en la figura 1A, la mayoría de los genes que son inducidos en frutos maduros están relacionados al proceso de degradación de la pared celular (14 contigs similares a endopoligalacturonasas), lo que guarda relación con el proceso de ablandamiento que experimentan los frutos cuando maduran (Trainotti y *col.*, 2003).

La mayoría de los genes que son reprimidos en el proceso de maduración (figura 1B), corresponden a genes similares a proteínas relacionadas a diversos procesos de transporte y genes que responden a estímulos de ácido abscísico (ABA). Interesantemente, 6 de estos genes están relacionados al desensamble de la pared celular (4 similares a pectina metil esterases, Micheli, 2001 y 2 similares a pectato liasas, Marín-Rodríguez y *col.*, 2002).

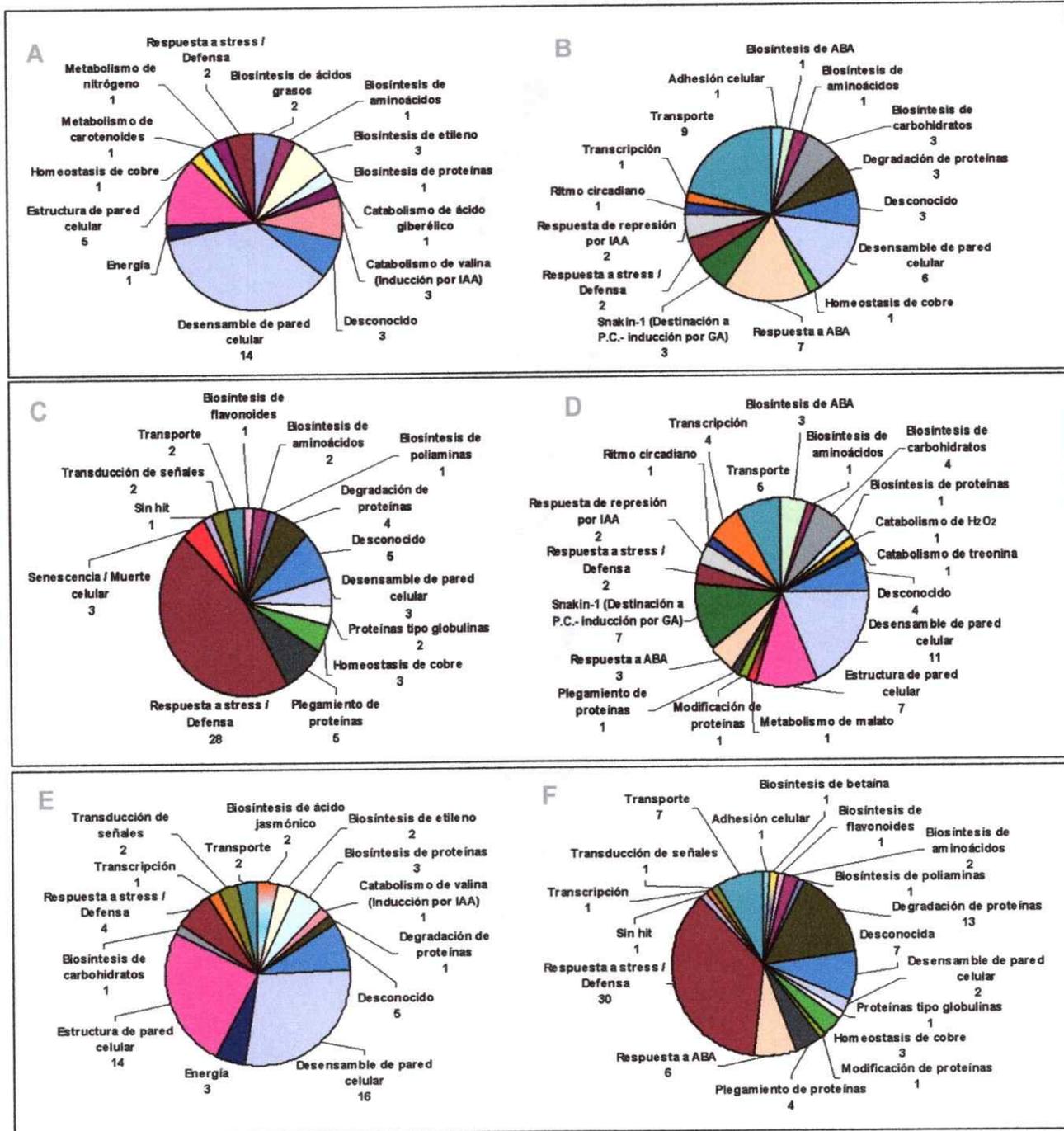


Figura 1. Distribución de contigs con niveles de expresión diferencial (método AC) según categorías funcionales. En los diferentes gráficos de tortas se muestran el número de contigs expresados diferencialmente ($p < 0,01$) en tres comparaciones y los procesos biológicos asignados por homología (BLASTx). A: clasificación de los 39 contigs inducidos en frutos maduros normales (E2) comparativamente con frutos de empaque (E1), B: clasificación de los 44 contigs reprimidos en frutos maduros normales (E2) comparativamente con frutos de empaque (E1), C: clasificación de los 63 contigs inducidos en frutos almacenados en frío (E3) comparativamente con frutos de empaque (E1), D: clasificación de los 60 contigs reprimidos en frutos almacenados en frío (E3) comparativamente con frutos de empaque (E1), E: clasificación de los 57 contigs inducidos en frutos harinosos (E4) comparativamente con frutos almacenados en frío (E3), F: clasificación de los 84 contigs reprimidos en frutos harinosos (E4) comparativamente con frutos almacenados en frío (E3). En la parte inferior de la figura se muestra un esquema que describe las tres comparaciones analizadas. (IAA, auxina; ABA, ácido absísico; GA, ácido giberélico; P.C., pared celular).

Como se puede observar en la figura 1C, la mayoría de los genes inducidos en los frutos de duraznos cuando son almacenados a 4°C por 21 días, tienen homología a genes relacionados en respuestas a estrés, en total son 28 contigs, dentro de los cuales se encuentran: 6 similares a taumatinas (Rupert y *col.*, 2002), 7 similares a quitinasas (Gijzen y *col.*, 2001) y 6 similares a deshidrinas (Thomashow, 1998). La mayoría de los genes que son reprimidos en este proceso (figura 1D), son genes relacionados al desensamble de la pared celular, lo que guarda relación con la detención del proceso de maduración cuando los frutos son almacenados a 4°C.

En la figura 1E se puede observar que la mayoría de los genes que son inducidos en E4 con respecto a E3, están relacionados al proceso de desensamble de la pared celular, como es en el caso de los genes inducidos en el proceso de maduración normal (figura 1A). Sin embargo, el número de genes relacionados a la estructura de la pared celular es mayor en el proceso de maduración anormal con respecto al proceso de maduración normal (14 y 5

contigs, respectivamente). Además se observan 2 contigs similares a una enzima (lipoxigenasa; contigs 4373 y 4733) que participa en la biosíntesis de ácido jasmónico (Heitz y *col.*, 1997). Este compuesto es una hormona vegetal que participa en diversos procesos biológicos en casi todo el reino vegetal, dentro de los que se encuentran respuesta a estrés, desarrollo y maduración de frutos (Creelman y Mullet, 1997). La mayoría de los genes reprimidos en esta comparación (E3 vs E4) corresponden a genes relacionados a defensa y respuesta a estrés (30 contigs, figura 1 F), incluyendo 4 contigs similares a taumatinas, 4 similares a quitinasas y 8 similares a deshidrinas. Estos resultados sugirieron que la inducción de la expresión de este tipo de genes en frutos a 4°C, es reprimida al colocar nuevamente los frutos a temperatura ambiente (E4).

Interesantemente, al comparar los estadíos E2 y E4 (datos no mostrados), únicamente se detectaron por este método 7 genes que se inducen (mayor expresión en E4, $p < 0,01$) y 9 genes que son reprimidos (mayor expresión en E2, $p < 0,01$). Los 7 contigs inducidos corresponden a: 3 similares a proteína E6 (contigs 4950, 4952 y 4975), 1 a pectina metil esterasa (contig 4957), 1 a expansina (contig 4933; Hayama y *col.*, 2000), 1 a ACC sintetasa (contig 4557; biosíntesis de etileno; Nakano y *col.*, 2003) y 1 similar a lipoxigenasa (contig 4733; biosíntesis de ácido jasmónico; Heitz y *col.*, 1997). Los 9 contigs reprimidos son los siguientes: 4 similares a proteínas alergenas (contigs 4749, 4863, 4976 y 4983), 2 similares a endopoligalacturonasas (contigs 4702 y 4981), 2 similares a proteínas inducidas por ABA (contigs 4959 y 4973) y 1 contig similar a una proteína de función desconocida (contig 4851).

5.3.- Análisis de perfiles de expresión digital: identificación de contigs coexpresados

Un primer paso lógico en la extracción de información de valores de expresión génica a escala genómica, es examinar los casos extremos, por ejemplo, aquellos genes con diferencias significativas en sus niveles de expresión en dos muestras individuales, la cual es una técnica eficiente para buscar posibles marcadores moleculares de algún fenotipo en particular. Este objetivo, fue abordado en esta tesis, comparando estadísticamente las cantidades de ESTs de cada uno de los genes (contigs) en dos genotecas de cDNA distintas, utilizando el método AC ("northern electrónico").

Sin embargo, este tipo de análisis no puede ser utilizado para identificar grupos de genes que tengan un patrón de expresión similar en múltiples condiciones y por defecto, grupos de genes que tengan un patrón de expresión diferente en múltiples condiciones (Stekel y *col.*, 2000). Este tipo de enfoque, nos aporta información acerca de genes que son coexpresados (algunos de los cuales pueden estar siendo corregulados) y por lo tanto, podrían estar interrelacionados funcionalmente (Ohlrogge y Benning, 2000). Además, los perfiles de expresión, aportan información adicional y útil para realizar la anotación funcional de genes que no exhiben similitud estructural significativa con ortólogos previamente caracterizados (Claverie, 1999).

Para abordar este objetivo, se realizó un análisis de expresión génica coordinada de los ESTs de las genotecas E1, E2, E3 y E4. Para ello se seleccionaron los contigs con un mayor nivel de expresión, que basado en trabajos previos (Ewing y *col.*, 1999), son aquellos que están formados por 5 o más ESTs (tabla 3). Con estos 1.801 contigs se construyó una tabla matriz

con los perfiles de expresión, como se describe en materiales y métodos. Como medida de similitud de los perfiles de expresión de los contigs y de las genotecas se utilizó el coeficiente de correlación de Pearson no centrado para luego agruparlos jerárquicamente de acuerdo al método de Eisen *y col.*, 1998. En la figura 2 se observa la representación gráfica de los resultados obtenidos. Los perfiles de expresión son simbolizados con una escala colorimétrica, yendo de amarillo, cuando no hay ESTs, hasta rojo cuando hay 25 o más ESTs. Los contigs (filas) y las genotecas (columnas) son reordenados y agrupados según la similitud en sus perfiles de expresión. El grado de similitud es representado por un árbol o dendrograma, en donde el largo de cada rama representa los coeficientes de correlación de Pearson calculados para cada par de genes.

Claramente se puede observar en la figura 2A, grupos de genes altamente expresados en cada una de las cuatro genotecas como bloques de color rojo intenso. Además, se pueden identificar 10 grupos o clusters jerárquicos principales (indicados como A, B, C, D, E, F, G, H, I y J). Cada uno de estos clusters tienen un coeficiente de correlación $r_{uc} > 0,8$, lo que indica un alto grado de similitud (Eisen, 1999). También se puede observar que las genotecas con perfiles de expresión más similares son E2 y E4 ($r_{uc} = 0,7$). Esto es consistente con los resultados obtenidos con el método AC, en donde sólo se detectaron 16 contigs con diferencias significativas en sus niveles de expresión.

En la figura 2B se muestran dos imágenes ampliadas correspondientes a grupos de contigs expresados preferencialmente en el estadio E3 (cluster B,

imagen superior) y en el estadio E4 (cluster G, imagen inferior). Puede observarse que genes con funciones similares se agrupan juntos.

Los patrones de expresión de cada uno de los cluster (A hasta J), en promedio, son mostrados en la figura 3. Cuatro de los diez clusters muestran patrones de expresión de tipo genoteca-específicos: los clusters B (293 contigs), D (140 contigs), G (103 contigs) y el cluster I (194 contigs) contienen contigs expresados preferencialmente en los estadios E3, E1, E4 y E2, respectivamente. El cluster J está constituido por 259 contigs expresados preferencialmente en los estadios de frutos maduros, tanto en jugosos como en harinosos (E2 y E4).

Los 1.801 contigs analizados en este ensayo, fueron anotados funcionalmente como se describe en materiales y métodos. En la figura 4 se muestra la distribución de estos contigs según las categorías funcionales asignadas por homología y según los patrones de expresión, es decir, a cuál cluster jerárquico pertenecen. Puede observarse en la figura 4A, que gran cantidad de los contigs pertenecientes al cluster B tienen homología a genes relacionados en respuestas a estrés y defensa (61 contigs). Dentro de este grupo de contigs se encuentran 11 similares a endoquitinasas, 11 deshidrinas y 10 taumatinas. Además en este cluster se observa una gran cantidad de contigs relacionados con los procesos de degradación de proteínas (27 contigs) y plegamiento de proteínas (10 contigs), procesos muy activos en respuestas a estrés por frío y otros tipos de estrés (Sabehat y *col.*, 1998; Iba, 2002; Spear y Ng, 2003). También puede observarse, un importante número de contigs relacionados a la estructura y a procesos de degradación de la pared celular en el cluster J (13 y 27 contigs, respectivamente). Esto último es esperado, ya que como se

mencionó antes, son procesos relacionados con la maduración y ablandamiento de los frutos. Los resultados guardan relación con lo observado con el método AC (figura 1A y 1E).

Nótese también en la figura 4A, que existe una categoría llamada snakin-1, este es un grupo de contigs que muestran homología con una proteína anti-patógena de papa (*Solanum tuberosum*), llamada snakin-1 (Segura y col., 1999). Estos contigs están clasificados independientemente porque en el trabajo de Trianotti y col., 2003, se comenta que esta proteína es destinada a la pared celular y el patrón de expresión detectado en ese trabajo es similar al observado por este procedimiento (cluster D), es decir, alta expresión en frutos fisiológicamente maduros (madurez de cosecha) y baja en frutos con madurez de consumo, debido a que su expresión es regulada negativamente por etileno.

En las figuras 4B y 4C se muestra la distribución de otros grupos de contigs que participarían en distintos procesos metabólicos y en diversos procesos de transporte, respectivamente. Es interesante notar que un número importante de contigs relacionados a procesos de transporte de proteínas se agrupan en el cluster H (figura 4C) y esto se relaciona con la proporción de contigs que participan en procesos de biosíntesis de proteínas en este mismo cluster (figura 4A).

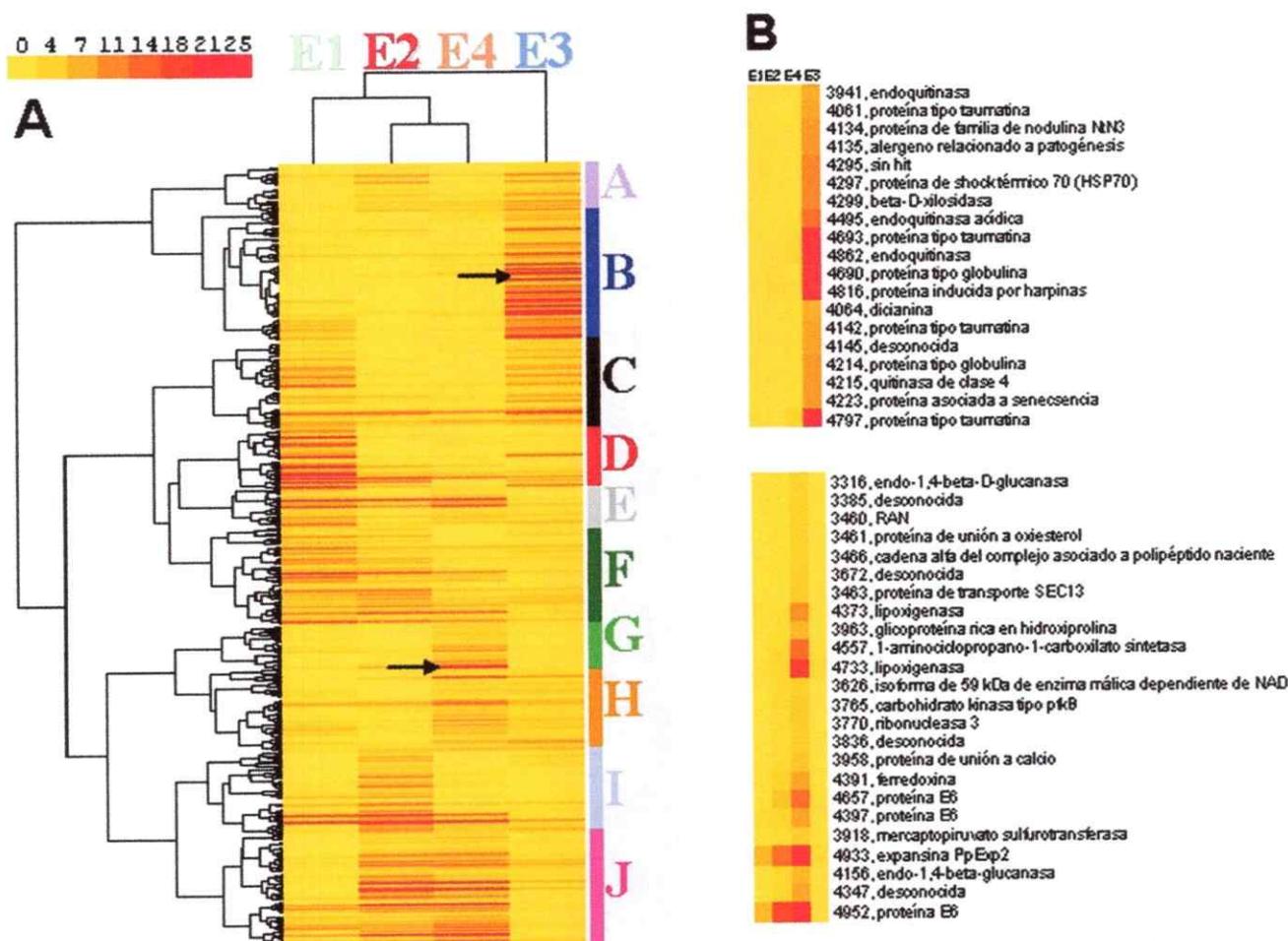


Figura 2. Agrupamiento jerárquico de contigs y genotecas de cDNA según sus perfiles de expresión. A: Un total de 1.801 contigs, formados por 5 o más EST's, fueron seleccionados para construir un mapa de agrupamiento correlacional, utilizando los coeficientes de correlación lineal de Pearson no centrado y agrupamiento promedio (ver materiales y métodos). El perfil de expresión de cada gen o contig es representado por una única fila de cajas coloreadas y cada una de las genotecas (E1, E2, E3 y E4) es representada por una única columna. Diferencias en la expresión génica son representadas con una escala colorimétrica, el rango es mostrado en la esquina superior izquierda de la figura, siendo amarillo 0 EST's y rojo 25 o más EST's. Se indican en la figura los diez clusters jerárquicos principales (A-J). B: Se muestran dos regiones expandidas del mapa (indicadas con flechas negras en 2A), en donde se describe el código numérico de cada contig y la anotación deducida por homología (Blastx); la región expandida superior corresponde a una parte del cluster B (contigs mayormente expresados en E3), y la región expandida inferior muestra una parte del cluster G (contigs mayormente expresados en E4).

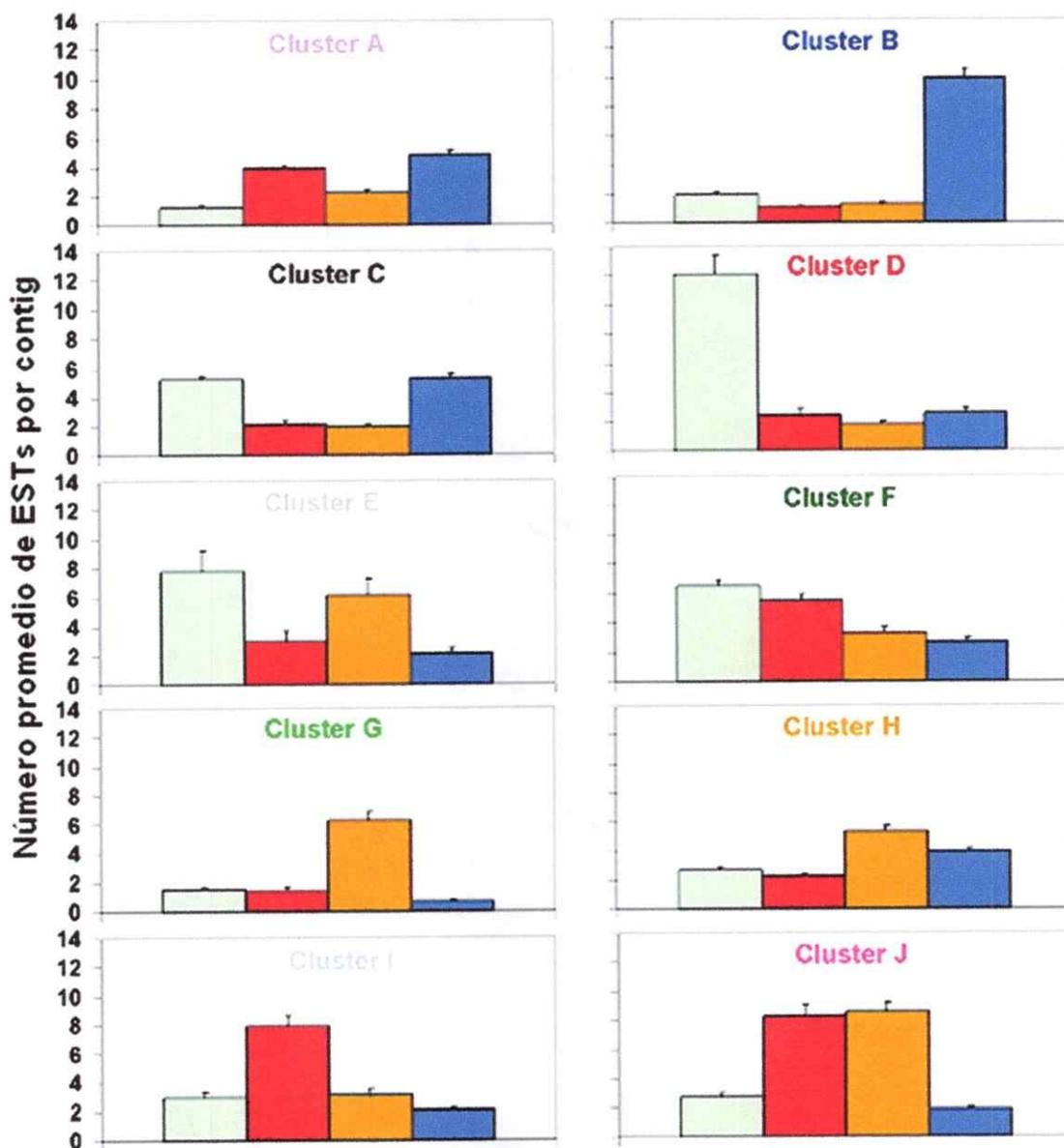


Figura 3. Patrones de expresión génica de los 10 clusters principales obtenidos mediante agrupamiento jerárquico. Clusters (A-J) se corresponden con (A-J) de la figura 2. En los histogramas se grafica el número promedio de EST's constituyentes de cada contig (ejes Y), con el error estándar para cada una de las genotecas de cDNA, (verde: E1; rojo: E2; anaranjado: E4; azul: E3).

Dentro de los contigs expresados preferencialmente en el estadio E4 (agrupados en el cluster G) se pueden observar dos contigs similares a lipoxigenasa (contigs 4373 y 4733; biosíntesis de ácido jasmónico, figura 4B). Este resultado guarda relación con lo obtenido mediante el método AC (figura

1E), en donde se observó que estos dos contigs son expresados significativamente más en el estadio E4 en comparación a E3. Hay que notar que, aunque estos dos contigs son expresados coordinadamente, solo el contig 4733 es expresado significativamente más (método AC) en frutos del estadio E4 que en frutos del estadio E2 (ver página 35).

Es de conocimiento general que los factores de transcripción son importantes en regular las respuestas fisiológicas frente a factores ambientales, como también en distintos eventos de desarrollo en todos los organismos vivos (Chen y col., 2002). Por lo tanto, para poder dilucidar de manera más detallada posibles vías regulatorias que puedan estar activas en los distintos estadios de postcosecha de durazno aquí evaluadas, se analizaron los contigs que presentaban similitud a factores de transcripción y que estén expresados preferencialmente en las genotecas E1, E2, E3 y E4, como aquellos expresados principalmente tanto en las genotecas E2 y E4 (clusters D, I, B, G y J, respectivamente). En la tabla 4 se muestran los contigs similares a factores de transcripción encontrados exclusivamente en cada uno de los clusters mencionados. Varios de los factores de transcripción han sido caracterizados funcionalmente en otras especies vegetales y para algunos de ellos se conocen las secuencias de los sitios de unión a promotores. En la tabla 4 se describen brevemente las funciones asignadas a cada uno de los factores de transcripción y las referencias.

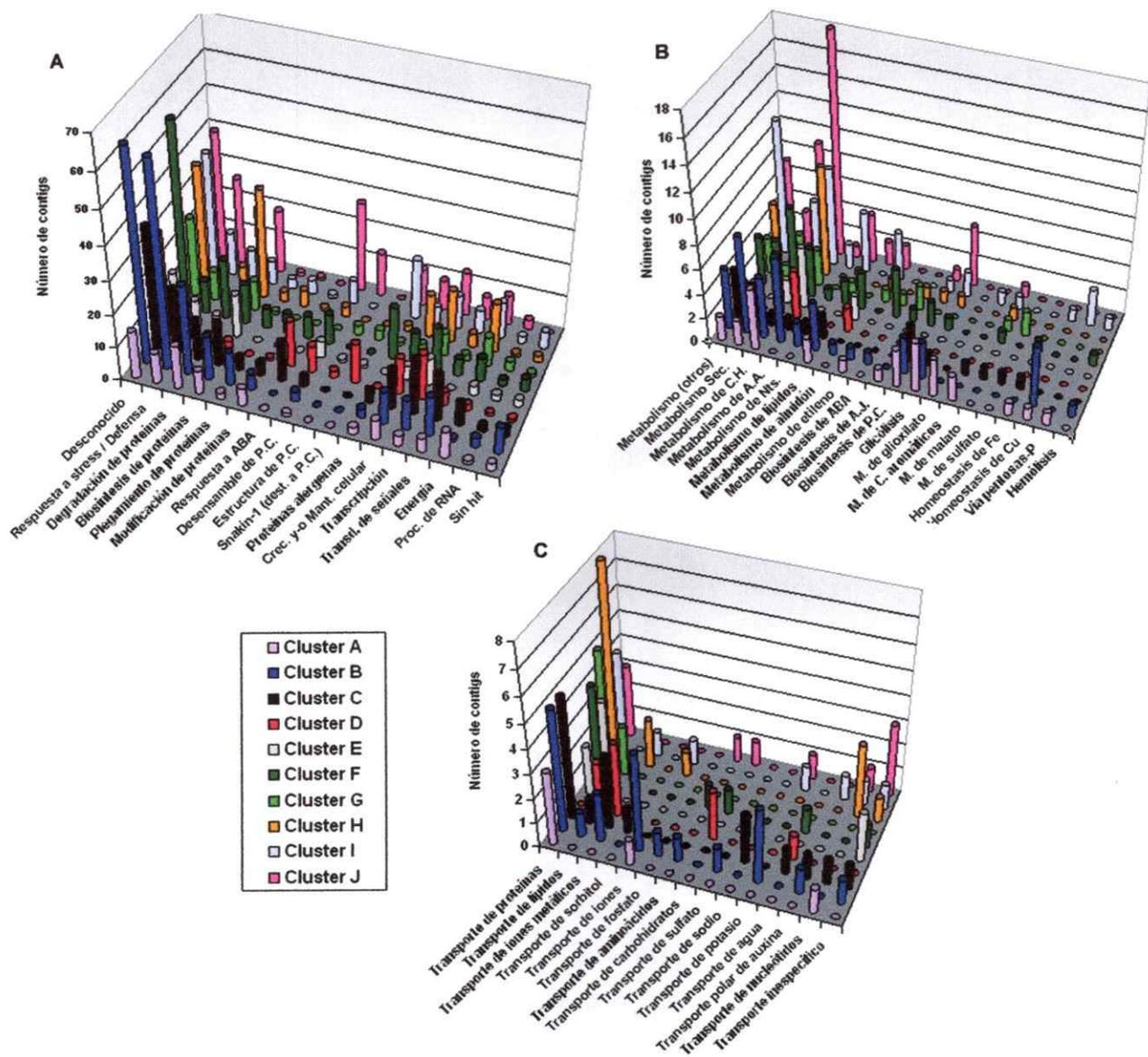


Figura 4. Distribución de los contigs analizados según sus perfiles de expresión y categorías funcionales asignadas. Los 1.801 contigs formados por 5 o más EST's fueron anotados utilizando BLASTx contra tres bases de datos diferentes (GenBank nr, TAIR y PlantGDB); luego fueron clasificados en las categorías funcionales correspondientes de acuerdo a su función putativa (ejes X). En los gráficos se muestra el número de contigs pertenecientes a cada una de estas categorías funcionales (eje Z) y a qué grupo de perfiles de expresión pertenecen (clusters A, J; eje Y). En A se muestran los contigs que participarían en procesos biológicos distintos a metabolismo o transporte como también aquellos a los que se les desconoce una función y los contigs que no son homólogos a ninguna secuencia publicada (sin hit). En B se describe el número de contigs que participarían en diferentes procesos metabólicos y en C se muestran aquellos contigs que participarían en procesos de transporte. (P.C., pared celular; dest. a P.C., destinación a pared celular; Crec. y-o Mant. celular, crecimiento y-o mantenimiento celular; Transd. de señales, transducción de señales; Proc. de RNA,

procesamiento de RNA; Metabolismo Sec., metabolismo secundario; Metabolismo de C.H., metabolismo de carbohidratos; Metabolismo de A.A., metabolismo de aminoácidos; Metabolismo de Nts., metabolismo de nucleótidos; Biosíntesis de A.J., biosíntesis de ácido jasmónico; Biosíntesis de P.C., biosíntesis de pared celular; M., metabolismo; M. de C. aromáticos, metabolismo de compuestos aromáticos; Vía pentosas-P, vía de las pentosas fosfato).

Tabla 4. Factores de transcripción expresados preferencialmente en distintas genotecas.

CLUSTER	Factor de transcripción (contig)	Descripción	Referencia
D	ERF1 (3235)	Factor de transcripción con dominio de unión a GCC-box, regula la inducción de genes en respuesta a estrés biótico y abiótico, de manera dependiente o independiente de etileno.	Fujimoto y col., 2000
	Tipo DREB, superfamilia A6 (3489)	Factor de transcripción con dominio de unión a DRE. Activador de la expresión génica en respuesta a estrés de frío o de deshidratación.	Shinozaki y Yamaguchi - Shinozaki, 2000
	Tipo WRKY (3877)	Factor de transcripción con dominios de unión a elementos W box. Regulación de la expresión de genes en respuesta a elicitores.	Dellagi y col., 2000
	bZIP tipo O2 (3724)	Control de la expresión específica en endosperma de varios genes de proteínas de almacenamiento en semillas de cereales a través del motivo en <i>cis</i> GCN4.	Onodera y col., 2001
	Homeobox-cierres de leucina 22 (2851)	Regulación de varios procesos relacionados al desarrollo (floración, pigmentación, elongación de hipocótilos, morfología de hojas).	Schena y Davis, 1994
	Proteína tipo NAM (3304)	Regulador de transcripción expresado en varios tejidos y estados de desarrollo.	Ooka y col., 2003
	ERF 5 (4319)	Factor de transcripción con dominio de unión a GCC-box, regula la inducción de genes en respuesta a estrés biótico y abiótico, de manera dependiente o independiente de etileno.	Fujimoto y col., 2000
	Proteína tipo NAM (4173)	Regulador de transcripción expresado en varios tejidos y estados de desarrollo.	Ooka y col., 2003
I	Proteína de respuesta a IAA (3917)	Represor de genes de respuesta temprana a IAA a bajas concentraciones de IAA. Es inducido por IAA.	Liscum y Reed, 2002
	Homeodominio 1 tipo BEL-1 (4191)	Control de la fertilidad y desarrollo de frutos.	Dong y col., 2000
	Familia Myb (3563)	Regulación de varios procesos metabólicos.	Baranowskij y col., 1994
	Familia Myb-PHANTASTICA (3312)	Establecimiento de identidad celular durante el desarrollo.	Koltai y McKenzie, 2000
	Familia Myb con dominio de unión a motivo AG-1 (3498)	Inducción de genes relacionados a defensa en respuesta a daño y elicitores.	Sugimoto y col., 2003
B	Bzip tipo TGA (4778)	Regula la expresión de genes en zonas de abscisión, es regulado positivamente por etileno y negativamente por IAA. Se une a elementos ACGTCA.	Tucker y col., 2002
	Proteína tipo DREB (3742)	Factor de transcripción con dominio de unión a DRE. Activador de la expresión génica en respuesta a estrés de frío o de deshidratación.	Shinozaki y Yamaguchi - Shinozaki, 2000
	Proteína tipo MADS-box (4451)	Regulación del desarrollo de órganos vegetativos (raíces y hojas) y reproductivos (flores, semillas y frutos).	De Bodt y col., 2003

	Proteína con dedos de zinc tipo CCCH (3660)	Factor de transcripción envuelto en respuestas a estrés hídrico. Se une a elementos con core AGT y ACT-box.	Sakamoto y col., 2000
	ERF3 (4586)	Regulador negativo de la vía de señalización mediada por etileno durante transducción de señales de estrés. Contiene dominio de unión a los sitios GCC-box.	Ohta y col., 2001
	Tipo WRKY (3201)	Factor de transcripción con dominios de unión a elementos W box. Regulación de la expresión de genes en respuesta a estrés bióticos y abióticos.	Eulgem y col., 2000
G	Familia TCP (3690)	Regulación de genes relacionados a crecimiento y desarrollo.	Kosugi y Ohashi, 2002
	Proteína con homeodominio de cierres de leucina (4230)	Probable participación en la activación de la expresión génica en respuesta a JA y azúcares como en la represión en respuesta a fosfato y auxina.	Tang y col., 2001
	Tipo bHLH (3952)	Amplio espectro de funciones en regulación de proliferación y diferenciación celular.	IEM y col., 2003
	Unión a CCAAT-box (3931)	Participación en la regulación ambiental y de desarrollo de la expresión génica.	Edwards y col., 1998
J	Relacionado a AP2.2 (4534)	Proteína de unión a elementos de respuesta a etileno. Expresado fuertemente en flores en Arabidopsis.	Okamuro y col., 1997
	bZIP (4550)	Participaría en la inducción de la expresión de genes controlados por luz.	Rugner y col., 2001
	bZIP (4122)	Relacionado al desarrollo de flores en tabaco.	Strathmann y col., 2001
	Proteína IAA16 (4400)	Factor de transcripción inducido por IAA.	Liscum y Reed, 2002
	Proteína de respuesta a IAA (4189)	Factor de transcripción inducido por IAA.	Liscum y Reed, 2002
	Proteína de respuesta a IAA (4726)	Factor de transcripción inducido por IAA.	Liscum y Reed, 2002
	EREBP (4857)	Relacionado al término de la dormancia de hayucos. Responde a etileno.	Calvo y col., 2003
	Proteína conteniendo dominio AP2 (4625)	Similar a cDNA con dominio AP2 aislado de <i>Prunus armeniaca</i> en estado de maduración.	Sin publicar
	Proteína AUX22 (4734)	Factor de transcripción inducido por IAA.	Ainley y col., 1988
Proteína AUX28 (4846)	Factor de transcripción inducido por IAA.	Ainley y col., 1988	

Cluster D, I, B, G y J: contigs expresados mayoritariamente en E1, E2, E3, E4 y E2-E4, respectivamente.

5.4.- Análisis de la expresión de taumatina, endoquitinasa, deshidrina y lipoxigenasa por RT-PCR

Los análisis de expresión digital realizados en este trabajo, sugieren la existencia de genes expresados diferencialmente en los cuatro estadios estudiados. Muchos de los genes inducidos en el proceso de maduración ya

habían sido asociados a este proceso, tanto en duraznos (Trainotti y col., 2003) como en otros frutos, especialmente en tomate (Fei y col., 2004). Además, resulta muy interesante que muchos de los genes inducidos en el estadio E3, o sea, en frutos almacenados a temperaturas bajas, tengan homología a genes relacionados con procesos de defensa. Se ha observado que los niveles de expresión de los mRNA de este tipo de genes se ven aumentados en respuesta a diversos tipos de estrés, incluido estrés por frío (Schenk y col., 2000; Kreps y col., 2002; Mahalingam y col., 2003).

Dentro de los genes altamente expresados en el estadio E3, detectados por el método AC y que funcionalmente están asociados a procesos de defensa, se encuentran 6 taumatinas, 7 endoquitinasas y 6 deshidrinas. Los perfiles de expresión digital sugieren que estos genes están siendo coexpresados, ya que todos se agrupan en el cluster B. Para obtener evidencias del posible mecanismo de regulación de la transcripción de estos genes, en este trabajo se aislaron las secuencias promotoras de un representante de cada tipo de gen (taumatina, contig 4797; endoquitinasa, contig 4862; deshidrina, contig 4889) y se analizaron *in silico* los dominios de unión de factores de transcripción. Estos representantes fueron escogidos debido a que sus niveles de expresión digital mostraron mayores diferencias de expresión entre el estadio E3 y el resto de los estadios (datos no mostrados). Previo a realizar la clonación de las secuencias promotoras y para comprobar los resultados obtenidos por los análisis de expresión digital por otra técnica, se analizaron por el método de RT-PCR los niveles de expresión de estos genes en los cuatro estadios de postcosecha evaluados.

Este mismo procedimiento se realizó para un gen similar a lipoxigenasa (contig 4733), enzima clave en la ruta biosintética del ácido jasmónico (Heitz y *col.*, 1997). El análisis por el método AC sugiere que este gen es inducido en E4 con respecto a E2. Además, mediante el método de agrupación jerárquica, este gen mostró tener un patrón de expresión preferencial para el estadio E4 (cluster G).

En la figura 5A se muestran los niveles de expresión digital de los cuatro genes escogidos; además se incluye en esta figura los niveles de expresión digital del contig 4881 (similar a actina), el cual fue elegido como potencial gen "housekeeping" debido a que no presenta diferencias significativas de expresión entre los cuatro estadios.

Para realizar el ensayo de RT-PCR se utilizó mRNA de 10 frutos representativos de cada estadio, el mRNA se aisló desde los "pooles" de RNA total utilizados para la construcción de las genotecas de cDNA (ver materiales y métodos). En la figura 5B se muestran los resultados obtenidos en el ensayo de RT-PCR. Se puede observar que los niveles de expresión tanto del gen de taumatina (346 pb), endoquitinasa (351 pb) y deshidrina (259 pb), son mayores en E3 que en el resto de los estadios. Se observan bandas más tenues en E1, E2 y E4 en las amplificaciones de endoquitinasa y taumatina, indicando que estos genes también son expresados (aunque a menor nivel) en los otros estadios. En cambio, sólo se detecta expresión de deshidrina en el estadio E3. Para los tres genes, al disminuir el número de ciclos de 34 a 26, las bandas en E1, E2 y E4 desaparecen (datos no mostrados). Para lipoxigenasa (306 pb), los resultados se acercan a lo esperado, es decir,

mayor expresión en E4 con respecto a los otros estadios. Se puede observar una banda de mayor intensidad en E4 comparativamente a E1, E2 y E3.

Se puede apreciar que los niveles de expresión de actina (325 pb) son similares en los cuatro estadios. Como se mencionó anteriormente, los análisis de expresión digital de este contig, sugieren que no hay diferencias significativas en los niveles de expresión en las cuatro genotecas (método AC), por lo que este método también podría ser útil para identificar posibles genes del tipo "housekeeping" en las condiciones estudiadas.

Los niveles de amplificación del control interno de la transcripción reversa (kanamicina, 323 pb), son similares en las cuatro muestras de cDNA, sugiriendo que las diferencias encontradas en las amplificaciones de los genes de durazno son reflejo de la cantidad de mRNA de cada gen y no de artefactos en las transcripciones reversas.

Para descartar contaminación con DNA genómico en las muestras de mRNA, se realizaron controles negativos sin transcriptasa reversa para las cuatro muestras de mRNA (ver materiales y métodos). En estos controles, como era lo esperado, no se obtuvo amplificación de ninguno de los genes (datos no mostrados). Para confirmar si los fragmentos amplificados desde los cDNAs correspondían a los esperados, estos fueron clonados en el vector pBluescript y secuenciados; las secuencias obtenidas fueron 100 % idénticas a las esperadas (datos no mostrados).

Los resultados del análisis de expresión digital y el ensayo de RT-PCR sugieren que los genes que codifican para taumatina (contig 4797), endoquitinasa (contig 4862) y deshidrina (contig 4889), son expresados

diferencialmente en frutos almacenados a 4°C por 21 días con respecto a las otras 3 condiciones evaluadas.

Además el gen que codifica para lipoxigenasa (contig 4733) tiene niveles de expresión mayores en frutos maduros harinosos comparativamente con frutos maduros jugosos.

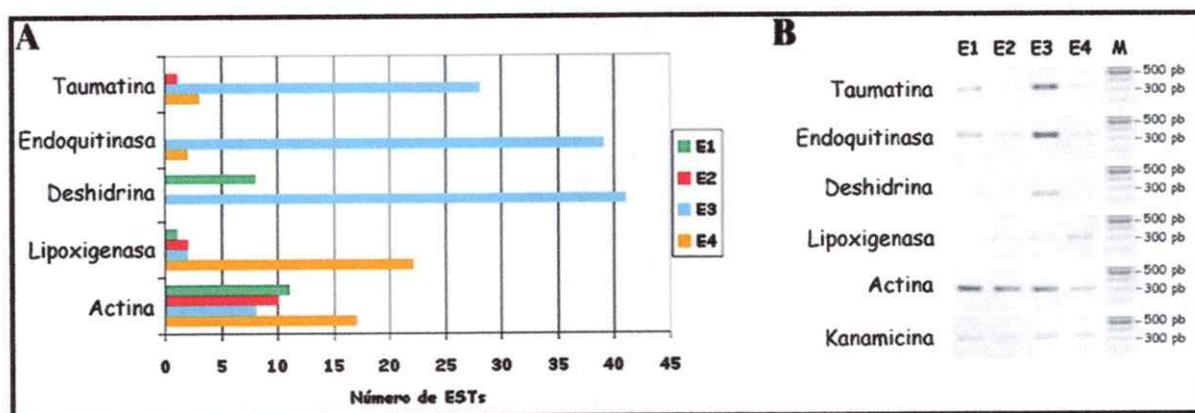


Figura 5. Niveles de expresión génica digital y análisis de la expresión de los genes taumatina, endoquitinasa, deshidrina y lipoxigenasa mediante RT-PCR. A: Niveles de expresión génica digital basados en el número de ESTs por cada estadio para los genes taumatina (contig 4797), endoquitinasa (contig 4862), deshidrina (contig 4889), lipoxigenasa (contig 4733) y actina (contig 4881). E1, barras verdes; E2, barras rojas; E3, barras celestes y E4, barras anaranjadas. B: Análisis de la expresión de los genes descritos en A mediante RT-PCR. Se utilizó mRNA aislado de un pool de RNA total de 10 frutos por cada uno de los estadios. Como control se utilizó actina y como control de la transcripción reversa se utilizó concentraciones similares de mRNA heterólogo (kanamicina). Los resultados son representativos de tres ensayos independientes (utilizando el mismo RNA total). M: marcador de peso molecular, GeneRuler™ 100pb DNA Ladder.

5.5.- Clonamiento de las secuencias de los posibles promotores de taumatina, endoquitinasa, deshidrina y lipoxigenasa

Para la obtención de los promotores putativos de las secuencias que codifican para taumatina, endoquitinasa, deshidrina y lipoxigenasa, se utilizó la técnica de "Genome Walker" (Siebert y *col.*, 1995), la cual permite aislar un fragmento desconocido de DNA genómico, adyacente a una zona conocida, en este caso, las secuencias de los contigs 4797, 4862, 4889 y 4733, respectivamente.

Se purificó DNA genómico desde hojas de durazno (variedad O'Henry), con el propósito de construir genotecas genómicas. Estas se generan digiriendo el DNA con enzimas de restricción que dan como resultado fragmentos con extremos romos, para luego realizar una ligación con una secuencia adaptadora. La secuencia adaptadora utilizada presenta sitios de reconocimiento para las enzimas de restricción Sall y SmaI, además de sitios de unión a partidores para realizar PCR. Se generaron 4 genotecas genómicas diferentes, utilizando las enzimas Mlsl, Pvull, EcoRV y PaeI (figura 6A). Una vez que se cuenta con las genotecas genómicas, el siguiente paso es aislar los putativos promotores utilizando 2 reacciones sucesivas de PCR por cada gen. Las primeras reacciones (PCR1) se realizaron utilizando un partidador específico de la secuencia adaptadora (AP1) y partidores específicos de las secuencias de los cDNAs (GSP1's); estos últimos fueron diseñados a 82, 87, 84 y 86 pares de bases de distancia río abajo del codón de inicio de la traducción (ATG) predicho para los cDNA de taumatina, endoquitinasa, deshidrina y lipoxigenasa, respectivamente. En la figura 6B se muestra el

resultado de este primer grupo de reacciones. Sólo se muestran las reacciones que dieron resultados positivos, esto es, cuando se utilizaron como templado las genotecas realizadas con las enzimas Mlsl y EcoRV. La segunda reacción de PCR (PCR2) se realizó utilizando como templado los productos de los primeros PCRs y partidores más internos (PCR anidado), uno para la secuencia del adaptador (AP2) y el otro para la secuencia del cDNA (GSP2's). Estos últimos fueron diseñados a 54, 46 y 50 pares de bases de distancia río abajo del ATG para los genes endoquitinasa, deshidrina y lipoxigenasa, respectivamente. Para taumatina el partidor más interno fue diseñado en la región 5' no traducida (5' UTR) inmediatamente río arriba del ATG, debido a que la secuencia entre el ATG y el partidor más externo no era propicia para el diseño de partidores. Los resultados de los PCR2 se muestran en la figura 6C. Las bandas obtenidas fueron purificadas y clonadas en los vectores pBluescript (taumatina y endoquitinasa) y en pGEM-t easy (deshidrina y lipoxigenasa). Luego los insertos fueron secuenciados y como se puede observar en la figura 7, todos poseen en un extremo las secuencias del partidor del adaptador AP2 y en el otro las secuencias de los partidores GSP2's más las secuencias específicas de cada gen, por lo tanto, se puede concluir que los amplificados corresponden a la parte 5' río arriba de cada uno de los genes seleccionados (posibles promotores). Se obtuvo una secuencia de promotor de aproximadamente 1.500 pares de bases desde el ATG para taumatina, 850 pares de bases para endoquitinasa, 1.350 pares de bases para deshidrina y 2.500 pares de bases para lipoxigenasa.

5.6.- Análisis de las secuencias de los posibles promotores de taumatina, endoquitinasa, deshidrina y lipoxigenasa

Una vez obtenidas las secuencias de los promotores clonados, se buscaron elementos de regulación conocidos en ellas. Para esto se utilizaron bases de datos de sitios de unión a factores de transcripción conocidos. Las bases de datos utilizadas fueron, PlantCARE (Lescot y col., 2002; <http://sphinx.ruq.ac.be:8080/PlantCARE/>) y PLACE (Higo y col., 1999; <http://www.dna.affrc.go.jp/htdocs/PLACE/>). Estas bases de datos contienen información sobre elementos regulatorios en cis específicos de plantas. En la tabla 5 se describen los elementos regulatorios más interesantes encontrados en cada una de las secuencias de los promotores.

Interesantemente, se encontraron varios elementos relacionados en respuestas a estrés, en particular en respuestas a frío y deshidratación (ABRE, DRE, LTR, GCC-box, AP2, GT1), en las secuencias de los promotores de los cuatro genes. Algunos de estos dominios se encontraron únicamente en los promotores de los genes inducidos en el estadio E3 (AP2 y GT1). Además, algunos elementos regulatorios se encontraron exclusivamente en la secuencia del promotor del gen que codifica para lipoxigenasa (GARE y QAR), el cual tiene un patrón de expresión diferente (inducción en E4).

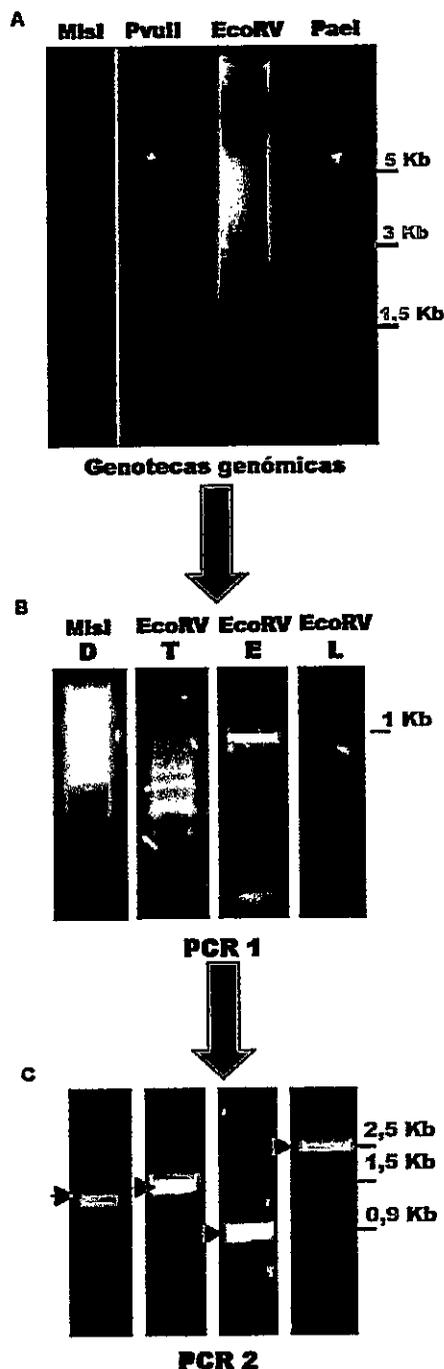


Figura 6. Amplificación de secuencias promotoras de deshidrina, taumatina, endoquitinasa y lipoxigenasa mediante la técnica de Genome Walker. A: Genotecas genómicas construidas con enzimas de restricción de corte como (MisI, PvuII, EcoRV y PaeI). A estas genotecas se les ligó un adaptador como se describe en materiales y métodos y luego fueron utilizadas como templado para reacciones de PCR. **B:** Resultados positivos fueron obtenidos con dos de las cuatro genotecas, MisI y EcoRV, en donde se aprecia un amplificado inespecífico para los genes deshidrina

Debido a que en los análisis de expresión digital se observaron factores de transcripción específicos con patrones de expresión similares a taumatina, deshidrina, endoquitinasa (tabla 4, cluster B) y similares al de lipoxigenasa (tabla 4, cluster G), se buscaron en las secuencias de los promotores aislados elementos de unión conocidos para estos tipos de factores de transcripción. Por ejemplo, en el promotor del gen de deshidrina se encontraron 4 tipos de dominios de unión de los 6 factores de transcripción agrupados en el cluster B (DRE, dominio de unión para proteína tipo DREB, contig 3742; GCC-box, dominio de unión para ERF3, contig 4586; W-box, dominio de unión para factor de transcripción tipo WRKY, contig 3201 y el motivo ACGTCA, dominio de unión para el factor de transcripción bZIP tipo TGA, contig 4778).

Otro elemento regulatorio interesante encontrado en la secuencia de los promotores de los genes de deshidrina y taumatina es el motivo TGTCACA. Este motivo está relacionado con la expresión fruto específica del gen de la cucumisina en melón (Yamagata y col., 2002). Rupert y col., 2002, observaron que un gen codificante para taumatina en durazno (99% idéntico al estudiado en este trabajo), es inducido por estrés de manera fruto y flor específica, lo que sugeriría que el promotor de taumatina aislado en este trabajo podría tener el mismo tipo de actividad.

Tabla 5. Elementos regulatorios en *cis* (sitios de unión de factores de transcripción) encontrados en las secuencias promotoras aisladas.

Nombre	Promotores (#)*	Función	Referencia
ABRE	E(2); D(7); T(5); L(2)	Elemento de respuesta a ABA.	Baker y col., 1994 Yamaguchi y col., 1994
Tipo AP2	E(1); D(2); T(1)	Sitio de unión de dominio AP2, elemento de respuesta a etileno.	Shinozaki y Yamaguchi - Shinozaki, 2000
TGACG- motif	E(1); T(3); L(2)	Elemento de respuesta a metil jasmonato.	Rouster y col., 1997
DRE	D(3); L(1)	Elemento de respuesta a deshidratación y frío.	Shinozaki y Yamaguchi - Shinozaki, 2000 Maruyama y col., 2004
ERE	E(3); D(2); T(3); L(2)	Elemento de respuesta a etileno.	Itzhaki y col., 1994
GCC-box	D(1)	Elemento de respuesta a etileno.	Fujimoto y col., 2000
LTR	T(1); L(1)	Elemento de respuesta a bajas temperaturas.	Baker y col., 1994
MYB	E(2); D(6); T(3); L(3)	Sitio de unión de factores Myb, encontrado en promotores de varios genes regulados por deshidratación.	Abe y col., 2003 Solano y col., 1995
MYC	E(1); D(6); T(9); L(3)	Sitio de unión de factores Myc, encontrado en el promotor de CBF3.	Abe y col., 2003 Chinnusamy y col., 2003
CCAAT box	E(3); D(2); T(4); L(3)	Sitio de unión de factores de transcripción tipo CCAAT-box.	Tasanen y col., 1992
ACGTCA motif	E(1); D(1); T(1); L(1)	Sitio de unión de factores de transcripción bZIP tipo TGA.	Tucker y col., 2002
W-box	E(4); D(4); T(8); L(3)	Sitio de unión de factores WRKY, presente en la secuencia de promotores de varios genes regulados por diferentes tipos de estrés bióticos y abióticos.	Rushton y Somssich, 1998
TGTCACA	D(1); T(2)	Elemento <i>enhancer</i> necesario para expresión fruto-específica.	Yamagata y col., 2002
GT1- motif	E(2); D(1); T(5)	Juega un rol en la inducción por patógenos y estrés salino en la expresión del gen <i>SCaM-4</i> de <i>Glycine max</i> .	Park y col., 2004
GARE	L(4)	Elemento de respuesta a ácido giberélico.	Ogawa y col., 2003 Gubler y col., 1995
QAR	L(1)	Elemento encontrado en promotor de gen extensina de <i>Brassica napus</i> el cual controla la activación en respuesta a daño y estrés de tensión.	Elliot y Shirsat, 1998

* Número de veces que está presente el elemento regulatorio en cada promotor (E, endoquitinasa; D, deshidrina; T, taumatina y L, lipoxigenasa)

6.- DISCUSIÓN

Prunus persica es considerado como organismo modelo de estudio de miembros de las Rosaceae, una familia económicamente importante que incluye plantas frutales y ornamentales como manzana, pera, ciruela, cereza, ciruela, almendra y rosa. Esto debido, principalmente, al tamaño relativamente pequeño de su genoma (300 Mb, Baird y col., 1994), a su breve período juvenil (2-3 años) y a la existencia de una gran cantidad de información genética, como son los mapas de marcadores moleculares (Georgi y col., 2002). La información de tipo genética y genómica de duraznos puede jugar un papel significativo en el descubrimiento de genes y en el entendimiento de la genética de especies relacionadas (Jung y col., 2004). Teniendo en cuenta esta necesidad, en mayo del 2004, se llevó a cabo en la Universidad de Clemson la "International Rosaceae Genome Mapping Conference" (Abbott, G.; www.genome.clemson.edu/gdr/conference/), en dicho evento, entre otros puntos, se establecieron las bases de una iniciativa internacional para realizar la secuenciación del genoma completo de durazno, en donde nuestro grupo de trabajo (www.genomavegetal.cl) también fue invitado a participar (Silva y col., 2004).

Chile es el principal exportador de duraznos del Hemisferio Sur, siendo uno de sus principales mercados Estados Unidos y países de la Comunidad Europea (ODEPA, www.odepa.gob.cl). Para prevenir la sobremadurez y senescencia durante el transporte a estos países, la fruta es almacenada a bajas temperaturas. Este tratamiento puede inducir un tipo de daño por frío llamado harinosidad, caracterizado por falta de jugo y textura harinosa de la fruta (Lill y col., 1989). Varios estudios sobre las bases bioquímicas de la harinosidad en

duraznos han identificado factores que pudieran ser importantes en el desarrollo de estos síntomas, aunque existe bastante discrepancia entre dichos estudios (Buescher y Furmanski, 1978; Ben-Arie y Sonego, 1980; Obenland y Carroll, 2000; Zhou y *col.*, 2000a; Artes y *col.*, 1996; Brummell y *col.*, 2004a). En estos artículos se analizan las diferencias en las actividades de enzimas pectolíticas responsables de la degradación de la pared celular entre frutos normales y harinosos. Otros estudios relacionan a la hormona etileno con el problema de la harinosidad, en ellos se muestra que la aplicación de un inhibidor de la acción de etileno (1-metilciclopropeno) en frutos de duraznos y nectarines, incrementa la generación de los síntomas (Dong y *col.*, 2001; Liguori y *col.*, 2004; Fan y *col.*, 2002). Estos antecedentes sugieren la existencia de un mayor número de componentes o un mayor grado de complejidad génica relacionados a la harinosidad en duraznos.

Para identificar posibles factores moleculares asociados a este problema, en este trabajo de tesis se realizó un análisis de expresión génica a gran escala en cuatro diferentes etapas de postcosecha de frutos de durazno de la variedad O'Henry (variedad muy susceptible a generar harinosidad), utilizando la información que genera la secuenciación de un gran número de ESTs de cuatro diferentes genotecas de cDNA. Se analizaron los perfiles de expresión globales de 1.801 genes (contigs) de duraznos en cuatro diferentes estadíos (E1, E2, E3 y E4; tabla 1). Estos contigs corresponden a genes con niveles de expresión moderada y alta, ya que están formados por 5 o más ESTs (Ewing y *col.*, 1999). Se utilizaron dos aproximaciones estadísticas para analizar los datos, la primera fue la utilización del método AC para detectar genes expresados diferencialmente entre dos condiciones (Audic y Claverie, 1997),

la segunda fue realizar una agrupación jerárquica de los genes y genotecas según la similitud (coeficiente de correlación de Pearson) de sus perfiles de expresión (Eisen y *col.*, 1998). Con el primer método se detectaron los genes con niveles de expresión significativamente diferentes entre dos estadios ($p < 0,01$) y con el segundo método se observaron no sólo estos genes, sino que también los genes que son expresados de forma correlacionada a ellos. Utilizando este último método, los 1.801 contigs analizados fueron clasificados en 10 grupos o clusters de expresión diferentes (figuras 2 y 3). De estos 10 clusters, cinco muestran patrones de expresión bien definidos. Los clusters D, I, B y G están formados por genes expresados principalmente en los estadios E1, E2, E3 y E4, respectivamente y el cluster J por genes expresados en frutos maduros, tanto jugosos como harinosos. En estos cinco clusters, se agrupan la mayoría (90%, aproximadamente), de los genes con diferencias significativas de expresión entre cualquiera de los estadios (AC, $p < 0,01$), por lo tanto, estos son los clusters que más se analizaron.

Debido a la gran cantidad de información obtenida, en la siguiente parte de este capítulo se discutirán sólo los resultados más significativos obtenidos de los perfiles de expresión génica globales. La discusión se centrará, principalmente, en algunos de los genes relacionados a los procesos de maduración, harinosidad y respuesta a estrés por frío.

Análisis de expresión génica digital relacionada a la maduración

La mayor parte del conocimiento de las bases moleculares de la maduración de frutos climatéricos, es decir, frutos que aumentan la tasa de respiración y biosíntesis de etileno durante la maduración, ha sido generado utilizando

como organismo modelo al tomate (*Lycopersicon esculentum*; Giovannoni, 2001, 2004; Fei y col., 2004; Seymour y col., 2002). Menor es el número de investigaciones relacionadas a la maduración de frutos realizada en duraznos (también fruto climatérico) y la mayoría se ha enfocado sólo en estudiar los cambios que ocurren en el metabolismo de la pared celular (Trianotti y col., 2003; Brummell y col., 2004b).

En este trabajo, se han identificado genes expresados diferencialmente en la maduración de frutos de duraznos, al comparar los niveles de expresión digital (frecuencias de ESTs) de una genoteca de cDNA construida con mRNA de 10 frutos no maduros (madurez de cosecha o "mature", E1) y una genoteca de cDNA construida con mRNA de 10 frutos maduros (madurez de consumo o "ripening", E2).

Utilizando el método AC se identificaron 39 genes inducidos significativamente ($p < 0,01$; figura 1A) en el proceso de maduración. Según los perfiles de expresión identificados mediante el método de agrupación jerárquica (figuras 2 y 3), 27 de estos genes se agrupan junto a otros 232 genes que tienen niveles de expresión preferencial tanto en E2 como en frutos maduros harinosos (E4, cluster J, figura 3), por lo tanto, pueden ser relacionados al proceso de maduración pero no a la harinosidad, ya que se encuentran altos niveles de transcritos en los dos tipos de frutos maduros.

Dentro de este grupo de genes expresados mayormente en frutos maduros (cluster J), se encontraron 27 genes relacionados a procesos de degradación de la pared celular (figura 4A), dentro de ellos se incluyen: 20 endopoligalacturonasas, 4 pectato liasas y 2 expansinas. Los niveles de expresión génica detectados para estos genes son esperados debido a que las proteínas

que codifican han sido asociadas a la maduración en muchos frutos, incluido duraznos (Trainotti y *col.*, 2003; Brummell y *col.*, 2004b; Brummell y Harpster, 2001). Estas proteínas son responsables de algunas de las modificaciones estructurales que sufre la pared celular durante el ablandamiento y cambios en la textura de frutos cuando maduran (Giovannoni, 2001; Brummell y Harpster, 2001). Interesantemente, dentro del cluster J también se identificaron 13 genes que codifican para proteínas estructurales de la pared celular siendo 12 de ellos similares a la proteína E6. Se ha observado que esta proteína se acumula en fibras de algodón durante los estadios tardíos de síntesis de la pared celular primaria y estados tempranos de síntesis de la pared celular secundaria (John y Crow, 1992), así como también durante el ablandamiento de frutos de duraznos y se le ha asignado una posible función como estabilizador de la estructura de la pared celular durante las modificaciones que ocurren en ella en el transcurso del ablandamiento (Trainotti y *col.*, 2003), por lo tanto no es raro que exista coexpresión de este tipo de genes con aquellos responsables de la degradación de la pared celular.

Otros genes expresados preferencialmente en frutos maduros son dos contigs codificantes para la enzima 1-aminociclopropano-1-carboxilato oxidasa (ACC oxidasa, contigs 4785 y 4945), enzima que participa en la biosíntesis de etileno. Se sabe que los niveles de etileno aumentan durante la maduración de frutos de duraznos y todos los frutos climatéricos y este aumento es asociado directamente con los niveles de transcritos y actividad de esta enzima (Barry y *col.*, 1996; Giovannoni, 2001; Brummell y *col.*, 2004b).

Un número importante de genes relacionados al metabolismo de aminoácidos también fueron identificados como relacionados a la maduración (18 contigs,

cluster J, figura 4B). Con relación a esto último, varios reportes indican que existe un aumento en el contenido de aminoácidos libres durante la maduración en uvas (fruto no climatérico; Lamikanra y Kassa, 1999; Nassar y Kliwer, 1966; Kliwer, 1968) y en bananas (fruto climatérico; Peumans y *col.*, 2002). El aumento en el número de transcritos de genes relacionados al metabolismo de aminoácidos en frutos de duraznos durante el proceso de maduración sugiere que los niveles de estos compuestos también podrían estar aumentados en frutos maduros. Dentro de estos genes, cuatro contigs codifican para enzimas que participan en la biosíntesis de metionina (cistationina gama sintetasa, contig 3912; 5-metiltioribosa kinasa, contig 4360; S-adenosilmetionina sintetasa, contigs 4656 y 4947). El aminoácido metionina es utilizado como precursor en la biosíntesis de etileno (Bleecker y Kende, 2000), por lo tanto no es sorprendente que el perfil de expresión de estos genes este correlacionado con el de la enzima ACC oxidasa. Estos resultados indican que podría existir una coexpresión de genes que participan en la ruta biosintética del etileno, durante la maduración de los frutos de duraznos.

Dentro de los genes expresados mayoritariamente en frutos maduros también se encontraron contigs similares a proteínas cuyos mRNAs son inducidos por la hormona auxina en *Arabidopsis thaliana*, ellos son los contigs 4855, 4751, 4575, 4576 y 4645 (similares a 3-hidroxiisobutil-coenzima A hidrolasa; función en catabolismo de valina; Zolman y *col.*, 2001) y 5 contigs codificantes para factores de transcripción que también son inducidos por auxina en *Arabidopsis* y soja (tabla 3, Liscum y Reed, 2002; Ainley y *col.*, 1988). La regulación transcripcional por auxina de este tipo de genes es conservada

entre especies como *Arabidopsis* y soja (*Glycine max*), lo que sugiere que también podría existir este tipo de regulación en duraznos (perteneciente a la misma subclase que soja) indicando que, posiblemente, los niveles de auxina en frutos maduros serían mayores que en frutos no maduros. No existen evidencias del papel de auxina en el control de la expresión génica durante las etapas tardías de la maduración de frutos climatéricos, pero esta hormona cumple un papel regulador en el proceso de maduración de frutos no climatéricos como la frutilla (Aharoni y col., 2002). Además, existe evidencia de que durante el proceso de abscisión de frutos y flores, la auxina regula la sensibilidad de las células por etileno (van Doorn y Stead, 1997), regulando de esta manera la acción de enzimas modificantes de la pared celular. Este tipo de regulación, podría también estar presente durante la maduración de frutos de duraznos. Sería interesante realizar, en trabajos futuros, mediciones de niveles de auxina y etileno en frutos de duraznos y evaluar su función regulatoria durante la maduración.

Análisis de expresión génica digital relacionada a la maduración y a la harinosidad

Como se mencionó en el punto 5.2, al comparar los perfiles de expresión de los estadios E2 y E4 utilizando el método AC, sólo se detectaron 16 genes con niveles de expresión significativamente diferentes ($p < 0,01$), 7 genes expresados mayoritariamente en E4 y 9 genes en E2. Sin embargo, los 7 genes expresados significativamente más en E4 con respecto a E2 se agruparon con 96 genes con un perfil de expresión similar (cluster G, figura 3). Los 9 genes expresados significativamente más en E2 con respecto a E4 se

agruparon junto a 185 genes más con el mismo perfil de expresión (cluster I, figura 3). En el cluster I se observan aquellos genes que son inducidos en frutos maduros jugosos y son reprimidos en frutos maduros harinosos. En el cluster G, se agrupan los genes que no son expresados fuertemente en frutos maduros jugosos pero si en frutos maduros harinosos.

Dentro de los genes presentes en el cluster I se encuentran 18 que codifican a proteínas alergenas (estructurales de la pared celular), 4 de estos 18 genes muestran diferencias significativas entre E2 y E4 utilizando el método AC ($p < 0,01$). Como se mencionó anteriormente, se piensa que este tipo de proteínas juega un papel en la estabilización de la estructura de la pared celular durante la etapa tardía del ablandamiento del fruto. No se conocen antecedentes de cambios en los niveles de expresión de estas proteínas en frutos de duraznos harinosos, pero el hecho de encontrar 18 genes homólogos a este tipo de proteínas que tienen un nivel de expresión menor en frutos harinosos, más antecedentes que sugieren que en frutos harinosos existe una alteración en las propiedades de la pared celular primaria (Brummell y *col.*, 2004a), sugiere una relación entre la regulación negativa de estas proteínas y los síntomas de la harinosidad. Sin embargo, como también se observaron 12 contigs similares a proteínas alergenas en el cluster J (genes expresados tanto en E2 como en E4), es arriesgado asociar estas proteínas a cambios de la estructura de la pared celular en frutos harinosos sin realizar una investigación más detallada. Existen 5 contigs que codifican para factores de transcripción que se encuentran exclusivamente en el cluster I (tabla 3). En este grupo se encuentra un factor de transcripción homeótico tipo BEL-1 (contig 4191). Este factor de transcripción ha sido caracterizado en manzana, también miembro

de la familia Rosaceae, (Dong y *col.*, 2000) y se lo asocia al control del desarrollo temprano de los frutos. Un proceso que es activo tanto en este estadio así como en la maduración de frutos, es la expansión celular (Westwood, 1978). Por lo tanto, una explicación posible del rol que este factor de transcripción pudiera estar jugando en la maduración de frutos de duraznos es que sea un regulador de la expansión celular, entre otras cosas, y este proceso estaría activo durante la maduración normal de frutos de durazno, no así en frutos harinosos. Existen evidencias que muestran que la forma de las células del mesocarpo de frutos harinosos es diferente a la de frutos jugosos, siendo más esféricas y pequeñas en el primer caso (Brummell y *col.*, 2004a). Teniendo en cuenta estas evidencias, BEL1 se convierte en un buen candidato para futuros estudios, como por ejemplo, sobreexpresar este gen durante el almacenamiento en frío en frutos de duraznos, utilizando promotores con este tipo de actividad (posiblemente como los aislados en este trabajo) y observar si disminuyen los síntomas de la harinosidad al dejar madurar posteriormente los frutos.

Dentro de los genes expresados mayormente en frutos harinosos con respecto a frutos jugosos, se encontraron 2 contigs similares a lipoxigenasa (contigs 4373 y 4733), enzima clave en la biosíntesis de ácido jasmónico (Heitz y *col.*, 1997). El ácido jasmónico es una hormona vegetal derivada del ácido linolénico a la que se le conocen varias funciones fisiológicas, como modular aspectos de la maduración de frutos, crecimiento de raíces, senescencia y resistencia a daño y patógenos, entre otros (Creelman y Mullet, 1997). El hecho de encontrar mayor número de transcritos de lipoxigenasa en frutos del estadio E4 comparativamente a frutos E2, sugiere que podría existir mayor

tasa de biosíntesis de ácido jasmónico en frutos harinosos que en frutos jugosos. Se ha observado en tomate la existencia de al menos tres isoformas de la enzima lipoxigenasa, teniendo cada una patrones de expresión diferentes, por ejemplo, transcritos de *TomLoxD* son acumulados rápida y transcientemente en plantas sometidas a daño, mientras que *TomLoxC* no es inducible por daño y es encontrada únicamente en frutos maduros (Heitz y col., 1997). Debido al patrón de expresión encontrado en este trabajo de los contigs similares a lipoxigenasa, se podría presumir que estas podrían estar siendo inducidas por un factor independiente del estado de maduración de los frutos (ya que no se observan niveles de expresión significativos en el estadio E2), sino que más bien podrían estar relacionadas a respuestas a daño o senescencia. Se ha observado que existe un rápido aumento en los niveles de ácido jasmónico debido a perturbaciones mecánicas causadas por la reducción del turgor celular inducido por déficit de agua (Creelman y Mullet, 1995). Este tipo de déficit puede estar ocurriendo en frutos harinosos debido a la retención del agua libre en geles formados por las pectinas de mayor tamaño generadas por el desbalance de la actividad de enzimas que participan en la degradación de la pared celular (Brummell y col., 2004a), explicándose de cierta manera, los mayores niveles de transcritos de lipoxigenasa encontrados en el estadio E4. Además esta posible reducción de turgor también puede explicar la disminución del tamaño celular de frutos harinosos.

En la secuencia del promotor de lipoxigenasa aislado en este trabajo, se encontraron varios elementos regulatorios en *cis* relacionados con respuestas a estrés y uno a metil jasmonato (TGACG-motif), un derivado del ácido

jasmónico (tabla 5). Además, un elemento en particular encontrado en la secuencia del promotor de lipoxigenasa, ha sido asociado con la activación génica en respuesta a estrés mecánico en *Brassica napus* (QAR, Elliot y Shirsat, 1998) y debido a que este elemento se encontró únicamente en este promotor y no en los otros tres aislados y a la actividad a la que se lo asocia, se podría especular que este elemento, junto al motivo TGACG, serían particularmente importantes en el control de la expresión de este gen.

Interesantemente, se identificaron cuatro contigs similares a factores de transcripción que son coexpresados junto a lipoxigenasa (cluster G, tabla 4), uno de ellos es similar a un factor de transcripción con dominios de cierres de leucina de soja (Tang y col., 2001). Este se une a motivos en *cis* de respuesta a jasmonatos y activa la expresión de dos genes (*VspA* y *VspB*) que son inducidos por jasmonatos en semillas de soja. Estos datos, junto con los elementos regulatorios en *cis* encontrados en el promotor de lipoxigenasa, sugieren que este tipo de factor de transcripción podría estar regulando la expresión de éste y otros genes en frutos de durazno del estadio E4.

Análisis de expresión génica digital en frutos almacenados en frío

El grado de harinosidad desarrollado en frutos de duraznos es dependiente del régimen de temperaturas y el tiempo de almacenamiento a estas, siendo máximo entre 2-8°C y después de 3 o 4 semanas (Ben-Arie y Lavee, 1971; Brummell y col., 2004a). Este hecho indica que podrían existir cambios fisiológicos en los frutos de durazno durante el almacenamiento a bajas temperaturas, que luego de un período de tiempo a temperatura ambiente (maduración), se traducirían en los síntomas de harinosidad. La mayoría de

los trabajos relacionados a este tema, se han enfocado en estudiar las diferencias existentes entre los estadios finales, es decir, entre frutos maduros jugosos y frutos maduros harinosos. Sin embargo, existen muy pocas evidencias que muestren los cambios que ocurren durante el almacenamiento en frío, o sea, el estadio que genera las condiciones fisiológicas que desencadenan la patología. En este trabajo de tesis se ha identificado un gran número de genes expresados preferencialmente en frutos de duraznos almacenados en frío (figuras 2 y 4, cluster B), al comparar los niveles de expresión digital (frecuencias de ESTs) de una genoteca de cDNA construida con mRNA de 10 frutos almacenados a 4°C por 21 días (E3) con otras tres genotecas de cDNA (E1, E2 y E4). Además, se han identificado un conjunto de genes que son reprimidos (60 contigs) o inducidos (63 contigs) durante el almacenamiento en frío (figura 1D y 1C, respectivamente), al comparar por el método AC el número de ESTs para cada gen entre E1 y E3.

La mayoría de los trabajos relacionados al análisis de respuestas génicas globales a estrés por frío en vegetales se ha llevado a cabo en la planta modelo *Arabidopsis thaliana* (Seki y col., 2001, 2002; Fowler y Thomashow, 2002; Kreps y col., 2002; Kim y col., 2003; Provar y col., 2003; Maruyama y col., 2004). Esto ha permitido la disección de la vía de transducción de señales de respuesta a estrés abiótico en plantas mejor caracterizada, la vía de aclimatación a frío CBF/DREB ("Cold binding factor/Dehydration responsive element binding"; Thomashow, 2001). Además se han identificado genes que responden al estrés por frío de manera independiente a esta vía (Fowler y Thomashow, 2002).

En este trabajo, se han encontrado 293 genes que tienen un patrón de expresión preferencial en frutos almacenados en frío (figura 2, cluster B), algunos de ellos similares a genes asociados a la vía CBF/DREB de *Arabidopsis* y otros no, sugiriendo que podría existir cierta conservación de esta vía en duraznos y que además también podría actuar en paralelo alguna vía alternativa.

Como era esperado, una gran proporción de los genes expresados mayormente en E3, o sea agrupados en el cluster B, se relacionan a respuesta a estrés y defensa (61 contigs, figura 4A). Dentro de este grupo de genes se encuentran 11 contigs similares a deshidrinas, proteínas altamente hidrofílicas pertenecientes al subgrupo D-11 de las proteínas LEA ("late-embryogenesis-abundant"; Close, 1997). Estas, son unas de las proteínas más abundantes que son inducidas durante períodos de déficit de agua provocado por sequía, salinidad y temperaturas bajas o congelantes (Ingram y Bartels, 1996; Close, 1997). La función de las deshidrinas aún es desconocida, pero existe un número de evidencias que indican que protegerían a las macromoléculas o estructuras celulares, del daño inducido por el déficit de agua (Koag y col., 2003; Sanchez-Ballesta y col., 2004). Varias deshidrinas de *Arabidopsis*, sino todas, son inducidas por frío por medio de la vía de transducción de señales CBF/DREB, es decir, su transcripción es regulada por factores de transcripción tipo DREB, los cuales interaccionan específicamente a los elementos en *cis* DRE (Maruyama y col., 2004). Interesantemente, dentro de las 1.350 pb del promotor de deshidrina aislado en este trabajo (contig 4889) se encontraron 3 elementos DRE (figura 7, tabla 5), sugiriendo la existencia de conservación entre *Arabidopsis* y

durazno en la activación transcripcional de este tipo de genes. Además, otro hecho que avala esta hipótesis, es que el contig 3742 (similar a un factor de transcripción tipo DREB) es coexpresado junto con los genes indicidos por frío (tabla 4).

Otro tipo de genes pertenecientes al grupo de los relacionados a respuestas a estrés y defensa dentro del cluster B, son 10 contigs similares a taumatina y 11 contigs similares a quitinasas, dos clases de proteínas relacionadas a patogénesis (PR). Se ha observado que este tipo de proteínas son acumuladas en los espacios apoplásticos de centeno de invierno (*Secale cereale*) durante el proceso de aclimatación a frío (inducido por bajas temperaturas) y que tienen propiedades anticongelantes que protegerían la integridad de las células de la planta de la formación del hielo (Hon y col., 1995; Yu y col., 2001). Además se ha observado que este tipo de proteínas retienen su actividad enzimática a bajas temperaturas, por lo que también son importantes en la resistencia a enfermedades, particularmente contra patógenos de bajas temperaturas como es el hongo *Typhula incarnata* (Hiilovaara-Teijo y col., 1999). La acumulación inducida por frío de proteínas PR asociada a la protección a este estrés, ha sido también observada en trigo (Ergon y col., 1998) y en tomate (Ding y col., 2002). Estas evidencias sugieren que la gran cantidad de contigs relacionados a proteínas PR, como taumatinas y quitinasas, inducidos durante el almacenamiento en frío de frutos de duraznos, puede ser asociada a un tipo de respuesta de aclimatación a frío que es conservada en varias plantas, y que protegería al fruto de un futuro daño por congelamiento y de patógenos oportunistas. En este trabajo, se comprobó el perfil de expresión digital por RT-PCR (figura 5) para un gen

similar a taumatina (contig 4797) y una endoquitinasa (contig 4862). Además, se ha mostrado que las secuencias de los promotores aislados de estos dos genes contienen varios elementos regulatorios en *cis* relacionados a respuestas a estrés y a hormonas reguladoras de respuestas de defensa como etileno y metil jasmonato (tabla 5). Al contrario de lo observado en el promotor de deshidrina, los promotores de taumatina y endoquitinasa, no contienen elementos del tipo DRE, sugiriendo que estos genes pueden ser inducidos por una vía alternativa a la de CBF/DREB. Una posible vía, sería la vía bZIP/ABRE (Shinozaki y Yamaguchi-Shinozaki, 2000). Esta vía es independiente y se entrecruza con la vía CBF/DREB durante la respuesta al estrés por frío. Se ha encontrado gran cantidad de elementos ABRE en las secuencias de los promotores de taumatina y endoquitinasa y además se ha observado que un factor de transcripción tipo bZIP (contig 4778, tabla 4) es coexpresado con estos genes. Este factor de transcripción se une a dominos similares a los ABRE en poroto y es expresado en zonas de abscisión (Tucker y col., 2002), comportamiento también observado en el gen de taumatina en durazno (Ruperti y col., 2002). Sin embargo, sin la realización de estudios más detallados, como ensayos de interacción entre factores de transcripción y promotores (geles de retardo, Yamagata y col., 2002), no se pueden emitir conclusiones.

Además del gran número de genes relacionados a respuestas de estrés y defensa encontrados en el cluster B, también se encontraron varios genes relacionados a los procesos de degradación y plegamiento de proteínas (27 y 10 contigs, respectivamente, figura 4A). Estos procesos son muy activos cuando las plantas son sometidas a condiciones de bajas temperaturas u otros

tipo de estrés y ayudan a reparar y prevenir el daño provocado por frío, estabilizando las estructuras de las proteínas y degradando las proteínas mal plegadas (Sabehat y col., 1998; Iba, 2002).

Interesantemente, se ha observado que en los frutos de duraznos almacenados a 4°C, además de un apagamiento en los niveles de transcritos de genes relacionados al desensamble de la pared celular, síntesis de etileno y otros procesos activos durante la maduración, existe un gran número de genes con alta actividad transcripcional, los cuales podrían estar actuando en la protección del daño celular por temperaturas frías.

Hasta la fecha, no existían trabajos que analicen el perfil transcripcional global en ningún tipo de frutos sometidos a bajas temperaturas, por lo que los resultados obtenidos en este trabajo pueden ser de gran utilidad no sólo dentro del contexto de la harinosidad de duraznos, sino que también como base para estudiar las respuestas a nivel de expresión génica de frutos durante el estrés por frío.

7.- CONCLUSIONES

- Existen diferencias significativas en los niveles de expresión génica globales entre los cuatro estadíos de postcosecha estudiados (Método AC).
- Se identificaron genes expresados diferencialmente entre frutos maduros jugosos y frutos maduros harinosos.
- Los niveles de expresión digital de los genes taumatina, lipoxigenasa, deshidrina, endoquitinasa y actina fueron congruentes con los niveles de expresión génica obtenidos mediante RT-PCR.
- Los promotores de los genes taumatina, deshidrina y endoquitinasa (coexpresados durante el almacenamiento en frío) contienen motivos regulatorios en *cis* conservados entre ellos y con otros genes de plantas también inducidos por frío. Explicando de cierta manera el patrón de expresión detectado, tanto por los métodos digitales como por RT-PCR.
- El promotor de lipoxigenasa (expresada preferentemente en frutos harinosos), contiene motivos regulatorios en *cis* conservados con los promotores de taumatina, endoquitinasa y deshidrina y elementos regulatorios presentes exclusivamente en su secuencia.

8.- PROYECCIONES

Los resultados de los análisis de expresión génica global pueden ser confirmados en un futuro cercano, mediante experimentos del tipo macroarreglos, ya que el proyecto, en donde se enmarca esta tesis, cuenta con una unidad encargada de hacer ese tipo de análisis.

Además, se propone medir el patrón de expresión de los genes identificados en este trabajo, en frutos de duraznos de variedades menos susceptibles a generar harinosidad, sometidos a las mismas condiciones de postcosecha evaluadas en este trabajo. Este tipo de análisis permitiría detectar con mayor certeza, genes relacionados a la generación de la harinosidad.

Se propone analizar la funcionalidad de estos promotores en frutos y hojas de duraznos (tanto a 4°C como a temperatura ambiente), transformando transientemente dichos tejidos con vectores binarios conteniendo cada uno de los promotores fusionados río arriba del gen reportero β -glucuronidasa (GUS). Debido a que las secuencias de los promotores aislados contienen elementos *in cis* conservados en genes de otras especies vegetales, se propone analizar la funcionalidad de estos promotores en *Arabidopsis thaliana*, generando plantas transgénicas con los vectores binarios mencionados anteriormente.

9.- REFERENCIAS

- Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2003. *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant Cell* **15**:63-78.
- Adams, M., Kelley, J., Gocayne, J., Dubnick, M., Polymeropoulos, M., Xiao, H., Merril, C., Wu, A., Olde, B., Moreno, F., Kerlavage, A., McCombie, W. and Venter, C. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**:1651-1656.
- Aharoni, A., Keizer, L., Van Den Broeck, H., Blanco-Portales, R., Muñoz-Blanco, J., Bois, G., Smit, P., De Vos, R. and O'Connell, A. 2002. Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climateric fruit. *Plant Physiol.* **129**:1019-1031.
- Ainley, W., Walker, J., Nagao, R. and Key, J. 1988. Sequence and characterization of two auxin-regulated genes from soybean. *J. Biol. Chem.* **263**:10658-10666.
- Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Anderson, R. 1979. The influence of storage temperature and warming during storage on peach and nectarine quality. *J. Amer. Soc. Hort. Sci.* **104**:459-461.
- Andrews, J., Bouffard, G., Cheadle, C., Lü, J., Becker, K. and Oliver, B. 2000. Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Research* **10**:2030-2043.
- Arabidopsis* Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**:796-815.
- Artes, F., Cano, A. and Fernandez-Trujillo, J. 1996. Pectolytic enzyme activity during intermittent warming storage of peaches. *Journal of Food Science* **61**:311-321.
- Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., Cherry, J., Davis, A., Dolinski, K., Dwight, S., Eppig, J. y col. 2000. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**:25-29.
- Audic, S. and Claverie, J. 1997. The significance of digital gene expression profiles. *Genome Research* **7**:986-995.
- Baird, W., Estager, A. and Wells, J. 1994. Estimating nuclear DNA content in peach and related diploid species using laser flow cytometry and DNA hybridization. *Journal Amer. Soc. Hort. Sci.* **119**:1312-1316.
- Baker, S., Wilhelm, K. and Thomashow, M. 1994. The 5'-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Mol. Biol.* **24**:701-13.

- Baranowskij, N., Frohberg, C., Prat, S. and Willmitzer, L. 1994. A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. *EMBO J.* **13**:5383-5392.
- Barry, C., Blume, B., Bouzayen, M., Cooper, W., Hamilton, A. and Grierson, D. 1996. Differential expression of the 1-aminocyclopropane-1-carboxylate oxidase gene family of tomato. *The Plant Journal* **9**:525-535.
- Ben-Arie, R. and Lavee, R. 1971. Pectic changes occurring in Elberta peaches suffering from woolly breakdown. *Phytochemistry* **10**:531-538.
- Ben-Arie, R. and Sonogo, L. 1980. Pectolytic enzyme activity involved in woolly breakdown of stored peaches. *Phytochemistry* **19**:2553-2555.
- Bleecker, A. and Kende, H. 2000. Ethylene: a gaseous signal molecule in plants. *Annu. Rev. Cell Dev. Biol.* **16**:1-18.
- Bonetta, D. and McCourt, P. 1998. Genetic analysis of ABA signal transduction pathways. *TRENDS Plant Sci.* **3**:231-235.
- Brenna, O., Pastorello, E., Farioli, L., Pravettoni, V. and Pompei, C. 2004. Presence of allergenic proteins in different peach (*Prunus persica*) cultivars and dependence of their content on fruit ripening. *Journal of Agric. and Food Chem.* **52**:7997-8000.
- Brummell, D. and Harpster, M. 2001. Cell wall metabolism in fruit softening and quality and its manipulations in transgenic plants. *Plant Mol. Biol.* **47**:311-340.
- Brummell, D., Dal Cin, V., Lurie, S., Crisosto, C. and Labavitch, J. 2004a. Cell wall metabolism during the development of chilling injury in cold-stored peach fruit: association of mealiness with arrested disassembly of cell wall pectins. *Journal of Exp. Botany* **55**:2041-2052.
- Brummell, D., Dal Cin, V., Crisosto, C. and Labavitch, J. 2004b. Cell wall metabolism during maturation, ripening and senescence of peach fruit. *Journal of Exp. Botany* **55**:2029-2039.
- Buescher, R. and Furmanski, R. 1978. Role of pectinesterase and polygalacturonase in the formation of wooliness in peaches. *Journal of Food Sci.* **43**:264-266.
- Burke, J., Wang, H., Hide, W. and Davison, D. 1998. Alternative gene form discovery and candidate gene selection from gene indexing projects. *Genome Research* **8**:276-290.
- Calvo, A., Jimenez, J., Lorenzo, O., Nicolás, C. and Rodriguez, D. 2003. Isolation and characterization of genes related to the breaking of beechnut dormancy and putatively involved in ethylene signal perception and transduction. In: *THE BIOLOGY OF SEEDS* (Nicolás, G., Bradford, K., Come, D. and Pritchard, H., Eds. CABI Publishing, Wallingford, UK). Pp:144-149.

Chen, W., Provart, N., Glazebrook, J., Katagiri, F., Chang, H., Eulgem, T., Mauch, F., Luan, S., Zou, G., Whitham, S., Budworth, P., Tao, Y., Xie, Z., Chen, X., Lam, S., Kreps, J., Harper, J., Si-Ammour, A., Mauch-Mani, B., Heinlein, M., Kobayashi, K., Hohn, T., Dangl, J., wang, X. and Zhu, T. 2002. Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative function in response to environmental stress. *The Plant Cell* **14**:559-574.

Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B., Hong, X., Agarwal, M. and Zhu, J. 2003. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes & Development* **17**:1043-1054.

Claverie, J. 1999. Computational methods for the identification of differential and coordinated gene expression. *Hum. Mol. Genetics* **8**:1821-1832.

Close, T. 1997. Dehydrins: a commonality in the response of plants to dehydration and low temperature. *Physiol. Plant.* **100**:291-296.

Creelman, R. and Mullet, J. 1995. Jasmonic acid distribution and action in plants: Regulation during development and response to biotic and abiotic stress. *Proc. Natl. Acad. Sci. USA* **92**:4114-4119.

Creelman, R. and Mullet, J. 1997. Biosynthesis and action of jasmonates in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**:355-381.

Crisosto, C., Mitchell, G. and Ju, Z. 1999. Susceptibility to chilling injury of peach, nectarine, and plum cultivars grown in California. *Hort. Science* **34**:1116-1118.

De Bodt, S., Raes, J., Van de Peer, Y. and Theißen, G. 2003. And then there were many: MADS goes genomic. *TRENDS in Plant Science* **8**:475-483.

Dellagi, A., Helibronn, J., Avrova, A., Montesano, M., Palva, BE., Stewart, H., Toth, I., Cooke, D., Lyon, G. and Birch, P. 2000. A potato gene encoding a WRKY-like transcription factor is induced in interactions with *Erwinia carotovora* subsp. *atroseptica* and *Phytophthora infestans* and is coregulated with class I endochitinase expression. *Mol. Plant Microbe Interact.* **13**:1092-1101.

Ding, C., Wang, C., Gross, K. and Smith, D. 2002. Jasmonate and salicylate induce the expression of pathogenesis-related-protein genes and increase resistance to chilling injury in tomato fruit. *Planta* **214**:895-901.

Dong, Y., Yao, J., Atkinson, R., Putterill, J., Morris, B. and Gardner, R. 2000. MDH1: an apple homeobox gene belonging to the BEL1 family. *Plant Mol. Biol.* **42**:623-633.

Dong, L., Zhou, H., Sonogo, L., Lers, A. and Lurie, S. 2001. Ethylene involvement in the cold storage disorder of "Flavortop" nectarine. *Postharv. Biol. Technol.* **23**:101-115.

Edwards, D., Murray, J. and Smith, A. 1998. Multiple genes encoding the conserved CCAAT-box transcription factor complex are expressed in *Arabidopsis*. *Plant Physiol.* **117**:1015-1022.

Eisen, M., Spellman, P., Brown, P. and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**:14863-14868.

Eisen, M. 1999. Cluster 3.0 Manual. Documentation to the program Cluster 3.0 by Michael Eisen, Stanford University.

Elliott, K. and Shirsat, A. 1998. Promoter regions of the *extA* extensin gene from *Brassica napus* control activation in response to wounding and tensile stress. *Plant Mol. Biol.* **37**:675-687.

Ergon, A., Klemsdal, S. and Tronsmo, A. 1998. Thermal hysteresis protein activity in bacteria, fungi and phylogenetically diverse plants. *Cryobiology* **30**:322-328.

Ettwiller, L., Rung, J. and Birney, E. 2003. Discovering novel *cis*-regulatory motifs using functional networks. *Genome Research* **13**:883-895.

Eulgem, T., Rushton, P., Robatzek, S. and Somssich, I. 2000. The WRKY superfamily of plant transcription factors. *Trends in Plant Science* **5**: 199-206.

Ewing, B. and Green, P. 1998. Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. *Genome Research* **8**:186-194.

Ewing, R., Kahla, A., Poirot, O., Lopez, F., Audic, S. and Claverie, J. 1999. Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Research* **9**:950-959.

Ewing, R. and Claverie, J. 2000. EST databases as multi-conditional gene expression datasets. *Pac. Symp. Biocomput.* 430-442.

Fan, X., Argenta, L. and Mattheis, J. 2002. Interactive effects of 1-MCP and temperature on 'Elberta' peach quality. *HortScience* **37**:134-138.

Fei, Z., Tang, X., Alba, R., White, J., Ronning, C., Martin, G., Tanksley, S. and Giovannoni, J. 2004. Comprehensive EST analysis of tomato and comparative genomics of fruit ripening. *The Plant Journal* **40**:47-59.

Fowler, S. and Thomashow, M. 2002. *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF response pathway. *The Plant Cell* **14**:1675-1690.

Fujimoto, S., Ohta, M., Usui, A., Shinshi, H. and Ohme-Takagi, M. 2000. *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *The Plant Cell* **12**:393-404.

Georgi, L., Wang, Y., Yvergnaux, D., Ormsbee, T., Inigo, M., Reighard, G. and Abbott, G. 2002. Construction of a BAC library and its application to the identification of simple sequence repeats in peach [*Prunus persica* (L.) Batsch]. *Theor. Appl. Genet.* **105**:1151-1158.

Gijzen, M., Kuflu, K., Qutob, D. and Chernys, J. 2001. A class I chitinase from soybean seed coat. *Journal of Exp. Botany* **52**:2283-2289.

Giovannoni, J. 2001. Molecular biology of fruit maturation and ripening. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**:725-749.

Giovannoni, J. 2004. Genetic regulation of fruit development and ripening. *The Plant Cell* **16**:S170-S180.

Gordon, D., Abajian, C. and Green, P. 1998. *Consed*: a graphical tool for sequence finishing. *Genome Research* **8**:195-202.

Gubler, F., Kalla, R., Roberts, J. and Jacobsen, J. 1995. Gibberellin-regulated expression of a myb gene in barley aleurone cells: evidence for Myb transactivation of a high-pl alpha-amylase gene promoter. *The Plant Cell* **7**:1879-1891.

Hayama, H., Shimada, T., Haji, T., Ito, A., Kashimura, Y. and Yoshioka, H. 2000. Molecular cloning of a ripening-related expansin cDNA in peach: evidence for no relationship between expansin accumulation and change in fruit firmness during storage. *Journal of Plant Physiol.* **157**:567-573.

Heim, M., Jakoby, M., Werber, M., Martin, C., Weisshaar, B. and Bailey P. 2003. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* **20**:735-747.

Heitz, T., Bergey, D. and Ryan, C. 1997. A gene encoding a chloroplast-targeted lipoxygenase in tomato leaves is transiently induced by wounding, systemin, and methyl jasmonate. *Plant Physiol.* **114**:1085-1093.

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

Hiilovaara-Teijo, M., Hannukkala, A., Griffith, M., Yu, X. and Pihakaski-Maunsbach, K. 1999. Snow-mold-induced apoplastic proteins in winter rye leaves lack antifreeze activity. *Plant Physiol.* **121**:665-674.

Hon, W., Griffith, M., Mlynarz, A., Kwok, Y. and Yang, D. 1995. Antifreeze proteins in winter rye are similar to pathogenesis-related proteins. *Plant Physiol.* **109**:879-889.

Iba, K. 2002. Acclimative response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance. *Annu. Rev. Plant Biol.* **53**:225-245.

Ingram, J. and Bartels, D. 1996. The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**:377-403.

Inoue, H., Nojima, H. and Okayama, H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**:23-28.

- Itzhaki, H., Maxson, J. and Woodson, W. 1994. An ethylene-responsive enhancer element is involved in the senescence-related expression of the carnation glutathione-S-transferase (*GST1*) gene. *Proc. Natl. Acad. Sci. USA* **91**:8925-8929.
- Iyengar, S., Schelling, J. and Sedor, J. 2002. Approaches to understanding susceptibility to nephropathy: from genetics to genomics. *Kidney Intern.* **61**:61-67.
- John, M. and Crow, L. 1992. Gene expression in cotton (*Gossypium hirsutum* L.) fiber: cloning of the mRNAs. *Proc. Natl. Acad. Sci. USA* **89**:5769-5773.
- Ju, Z., Duan, Y. and Ju, Z. 2000. Leatheriness and mealiness of peaches in relation to fruit maturity and storage temperature. *J. Hort. Sci. Biotechnol.* **75**:86-91.
- Jung, S., Jesudurai, C., Staton, M., Du, Z., Ficklin, S., Cho, I., Abbott, A., Tompkins, J. and Main, D. 2004. GDR (Genome Database for Rosaceae): integrated web resources for Rosaceae genomics and genetics research. *BMC Bioinformatics* **5**:130-138.
- Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. 1999. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat. Biotech.* **17**:287-291.
- Kim, H., Snesrud, E., Haas, B., Cheung, F., Town, C. and Quackenbush, J. 2003. Gene expression analyses of *Arabidopsis* chromosome 2 using a genomic DNA amplicon microarray. *Genome Research* **13**:327-340.
- Kliwer, W. 1968. Changes in the concentration of free amino acids in grape berries during maturation. *Am. J. Enol. Vitic.* 166-174.
- Koag, M., Fenton, R., Wilkens, S. and Close, T. 2003. The binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity. *Plant Physiol.* **131**:309-316.
- Koltai, H. and McKenzie, D. 2000. Epistatic repression of *PHANTASTICA* and class 1 *KNOTTED* genes is uncoupled in tomato. *The Plant Journal* **22**:455-459.
- Kosugi, S. and Ohashi, Y. 2002. DNA binding and dimerization specificity and potential targets for the TCP protein family. *The Plant Journal* **30**:337-348.
- Kreps, J., Wu, Y., Chang, H., Zhu, T., Wang, X. and Harper, J. 2002. Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiol.* **130**:2129-2141.
- Kuo, W., Jenssen, T., Butte, A., Ohno-Machado, L. and Kohane, I. 2002. Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* **18**:405-412.
- Lamikanra, O. and Kassa A. 1999. Changes in the free amino acid composition with maturity of the noble cultivar of *Vitis rotundifolia* Michx. Grape. *J. Agric. Food Chem.* **47**:4837-4841.

- Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., Rouzé, P. and Rombauts, S. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res.* **30**:325-327.
- Liguori, G., Weksler, A., Zutahi, Y., Lurie, S. and Kosto, I. 2004. Effect of 1-methylcyclopropene on ripening of melting flesh peaches and nectarines. *Postharv. Biol. Technol.* **31**:263-268.
- Lill, R., O'Donoghue, E. and King, G. 1989. Postharvest physiology of peaches and nectarines. *Horticultural Reviews.* **11**:413-452.
- Lipshutz, R., Fodor, S., Gingeras, T. and Lockhart, D. 1999. High density synthetic oligonucleotide arrays. *Nat. Genet.* **21**(Suppl. 1):20-24.
- Liscum, E. and Reed, J. 2002. Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Mol. Biol.* **49**:387-400.
- Mahalingam, R., Gomez-Buitrago, A., Eckardt, N., Shah, N., Guevara-Garcia, A., Day, P., Raina, R. and Fedoroff, N. 2003. Characterizing the stress/defense transcriptome of *Arabidopsis*. *Genome Biol.* **4**:R20.
- Manubens, A., Lobos, S., Jadue, Y., Toro, M., Messina, R., Lladser, M. and Seelenfreund, D. 1999. DNA isolation and AFLP fingerprinting of nectarine and peach varieties (*Prunus persica*). *Plant Molecular Biology Reporter* **17**:255-267.
- Marín-Rodríguez, M., Orchard, J. and Seymour, G. 2002. Pectate lyases, cell wall degradation and fruit softening. *Journal of Exp. Botany* **53**:2115-2119.
- Maruyama, K., Sakuma, Y., Kasuga, M., Ito, Y., Seki, M., Goda, H., Shimada, Y., Yoshida, S., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2004. Identification of cold-inducible downstream genes of the *Arabidopsis* DREB1A/CBF3 transcriptional factor using two microarray systems. *The Plant Journal* **38**:982-993.
- Meisel, L., Fonseca, B., Gonzalez, S., Baeza-Yates, R., Cambiazo, V., Campos, R., Gonzalez, M., Orellana, A., Retamales, J. and Silva, H. 2005. A rapid and efficient method for purifying high quality total RNA from peaches (*Prunus persica*) for functional genomic analyses. *Biol. Research*. In press.
- Mekhedov, S., Martínez de Ilárduya, O. and Ohlrogge, J. 2000. Towards a functional catalog of the plant genome: a survey of genes for lipid biosynthesis. *Plant Physiol.* **122**:389-401.
- Micheli, F. 2001. Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *TRENDS in Plant Science* **6**:414-419.
- Mysore, K., Tuori, R. and Martin, G. 2001. *Arabidopsis* genome sequence as a tool for functional genomics in tomato. *Genome Biol.* **2**:reviews 1003.1-1003.4.

Nakano, R., Ogura, E., Kubo, Y. and Inaba A. 2003. Ethylene biosynthesis in detached young persimmon fruit initiated in calyx and modulated by water loss from the fruit. *Plant Physiol.* **131**:276-286.

Nassar, A. and Kliewer, W. 1966. Free amino acids in various parts of *Vitis vinifera* at different stages of development. *Proc. Am. Soc. Hortic. Sci.* **89**:281-294.

Obenland, D. and Carroll, T. 2000. Mealiness and pectolytic activity in peaches and nectarines in response to heat treatment and cold storage. *Journal of the Amer. Soc. For Hort. Sci.* **125**:723-728.

Ogawa, M., Hanada, A., Yamauchi, Y., Kuwahara, A., Kamiya, Y. and Yamaguchi, S. 2003. Gibberellin biosynthesis and response during *Arabidopsis* seed germination. *The Plant Cell* **15**:1591-1604.

Ogihara, Y., Mochida, K., Nemoto, Y., Murai, K., Yamazaki, Y., Shin-I, T. and Kohara, Y. 2003. Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analysis of expressed sequence tags. *The Plant Journal* **33**:1001-1011.

Ohlrogge, J. and Benning, C. 2000. Unraveling plant metabolism by EST analysis. *Curr. Op. in Plant Biol.* **3**:224-228.

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. and Ohme-Takagi, M. 2001. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *The Plant Cell* **13**:1959-1968.

Okamuro, J., Caster, B., Villarroel, R., Van Montagu, M. and Jofuku, K. 1997. The AP2 domain of *APETALA2* define a large new family of DNA binding proteins in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **94**:7076-7081.

Onodera, Y., Suzuki, A., Wu, C., Washida, H. and Takaiwa, F. 2001. A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. *J. Biol. Chem.* **276**:14139-14152.

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., Matsubara, K., Osato, N., Kawai, J., Carninci, P., Hayashizaki, Y., Suzuki, K., Kojima, K., Takahara, Y., Yamamoto, K. and Kikuchi, S. 2003. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Research* **31**:239-247.

Park, H., Kim, M., Kang, Y., Jeon, J., Yoo, J., Kim, M., Park, C., Jeong, J., Moon, B., Lee, J., Yoon, H., Lee, S., Chung, W., Lim, C., Lee, S., Hong, J. and Cho, M. 2004. Pathogen- and NaCl-induced expression of the SCaM-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. *Plant Physiol.* **135**:2150-2161.

Pérez, A., Olías, R. and Sanz, C. 1999. Biosynthesis of 4-hydroxy-2,5-dimethyl-3(2H)-furanone and derivatives in *in vitro* grown strawberries. *J. Agric. Food Chem.* **47**:655-658.

Peumans, W., Proost, P., Swennen, R. and Van Damme, E. 2002. The abundant class III chitinase homolog in young developing banana fruits behaves as a transient vegetative storage protein and most probably serves as an important supply of amino acids for the synthesis of ripening-associated proteins. *Plant Physiol.* **130**:1063-1072.

Phimister, B. 1999. Chipping forecast. *Nat. Genet.* **21**:1-60.

Prade, R., Ayoubi, P., Krishnan, S., Macwana, S. and Russell, H. 2001. Accumulation of stress and inducer-dependent plant-cell-wall-degrading enzymes during asexual development in *Aspergillus nidulans*. *Genetics* **157**:957-967.

Provar, N., Gil, P., Chen, W., Han, B., Chang, H., Wang, X. and Zhu, T. 2003. Gene expression phenotypes of *Arabidopsis* associated with sensitivity to low temperatures. *Plant Physiol.* **132**:893-906.

Rafalski, J., Hanafey, M., Miao, G., Ching, A., Lee, J., Dolan, M. and Tingey, S. 1998. New experimental and computational approaches to the analysis of gene expression. *Acta Biochem. Pol.* **45**:929-934.

Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouzé, P. And Van de Peer, Y. 2003. Computational approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant Physiol.* **132**:1162-1176.

Romualdi, C., Bortoluzzi, S. and Danieli, G. 2001. Detecting differentially expressed genes in multiple tag sampling experiment: comparative evaluation of statistical tests. *Hum. Mol. Gen.* **10**:2133-2141.

Ronning, C., Stegalkina, S., Ascenzi, R., Bougri, O., Hart, A., Utterbach, T., Vanaken, S., Riedmuller, S., White, J., Cho, J., Perte, G., Lee, Y., Karamycheva, S., Sultana, R., Tsai, J., Quackenbush, J., Griffiths, H., Restrepo, S., Smart, C., Fry, W., van der Hoeven, R., Tanksley, S., Zhang, P., Jin, H., Yamamoto, M., Baker, B. and Buell, R. 2003. Comparative analyses of potato expressed sequence tag libraries. *Plant Physiol.* **131**:419-429.

Rouster, J., Leah, R., Mundy, J. and Cameron-Mills, V. 1997. Identification of methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. *The Plant Journal.* **11**:513-523.

Rügner, A., Frohnmeyer, H., Näke, C., Wellmer, F., Kircher, S., Schäfer, E. and Harter, K. 2001. Isolation and characterization of four novel parsley proteins that interact with the transcriptional regulators CPRF1 and CPRF2. *Mol. Genet. Genomics* **265**:964-976.

Ruperti, B., Cattivelli, L., Pagni, S. and Ramina, A. 2002. Ethylene-responsive genes are differentially regulated during abscission, organ senescence and wounding in peach (*Prunus persica*). *Journal of Exp. Botany* **53**:429-437.

Rushton, P. and Somsschid, I. 1998. Transcriptional control of plant genes responsive to pathogens. *Curr. Opin. in Plant Biol.* **1**:311-315.

- Sabehat, A., Lurie, S. and Weiss, D. 1998. Expression of small heat-shock proteins at low temperatures. *Plant Physiol.* **117**:651-658.
- Sakamoto, H., Araki, T., Meshi, T. and Iwabuchi, M. 2000. Expression of a subset of the *Arabidopsis* Cys(2)/His(2)-type zinc-finger protein gene family under water stress. *Gene* **248**:23-32.
- Sanchez-Ballesta, M., Rodrigo, M., Lafuente, M., Granell, A. and Zacarias, L. 2004. Dehydrin from *Citrus*, which confers *in vitro* dehydration and freezing protection activity, is constitutive and highly expressed in the flavedo of fruit but responsive to cold and water stress in leaves. *J. Agric. Food Chem.* **52**:1950-1957.
- Schena, M. and Davis, R. 1994. Structure of homeobox-leucine zipper genes suggests a model for the evolution of genes families. *Proc. Natl. Acad. Sci. USA* **91**:8393-8397.
- Schenk, P., Kazan, K., Wilson, I., Anderson, J., Richmond, T., Somerville, S. and Manners, J. 2000. Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. *Proc. Natl. Acad. Sci. USA* **97**:11655-11660.
- Schwab, W. 1998. Application of stable isotope ratio analysis explaining the bioformation of 2,5-dimethyl-4-hydroxy-3(2H)-furanone in plants by a biological Maillard reaction. *J. Agric. Food Chem.* **46**:2266-2269.
- Schwab, W., Aharoni, A., Raab, T., Pérez, A. and Sanz, C. 2001. Cytosolic aldolase is a ripening related enzyme in strawberry fruits (*Fragaria X ananassa*). *Phytochemistry*. **56**:407-415.
- Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y. and Shinozaki, K. 2001. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stress by using a full-length cDNA microarray. *The Plant Cell* **13**:61-72.
- Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A., Sakurai, M., Akiyama, K., Taji, Y., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K. 2002. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stress using a full-length cDNA microarray. *The Plant Journal* **31**:279-192.
- Seymour, G., Manning, K., Eriksson, E., Popovich, A. and King, G. 2002. Genetic identification and genomic organization of factors affecting fruit texture. *Journal of Exp. Botany*. **53**:2065-2071.
- Shalon, D., Smith, S. and Brown, P. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Research* **6**:639-645.
- Shinozaki, K. and Yamaguchi-Shinozaki K. 2000. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr. Opin. in Plant Biol.* **3**:217-223.

Siebert, P., Chenchik, A., Kellogg, D., Lukyanov, K. and Lukyanov S. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Research* **23**:1087-1088.

Silva, H., Baeza-Yates, R., Orellana, A., Latorre, M., Loira, N., Maldonado, J., Tittarelli, A., Vizoso, P. and Meisel, L. 2004. Bioinformatics in the Chilean consortium. Abstract: www.genome.clemson.edu/gdr/conference/abstractList.html.

Solano, R., Nieto, C., Avila, J., Canas, L., Diaz, I. and Paz-Ares, J. 1995. Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. *EMBO J.* **14**:1773-84.

Southern, E., Mir, K. and Shchepinov, M. 1999. Molecular interactions on microarrays. *Nat. Genet.* **21**(Suppl. 1):5-9.

Spear, E. and Ng, D. 2003. Stress tolerance of misfolded carboxypeptidase Y requires maintenance of protein trafficking and degradative pathways. *Mol. Biol. Cell* **14**: 2756-2767.

Stekel, D., Git, Y. and Falciani, F. 2000. The comparison of gene expression from multiple cDNA libraries. *Genome Research* **10**:2055-2061.

Strathmann, A., Kuhlmann, M., Heinekamp, T. and Droge-Laser, W. 2001. BZI-1 specifically heterodimerises with the tobacco bZIP transcription factors BZI-2, BZI-3/TBZF and BZI-4, and is functionally involved in flower development. *The Plant Journal* **28**:397-408.

Sugimoto, K., Takeda, S. and Hirochika, H. 2003. Transcriptional activation mediated by binding of a plant GATA-type zinc finger protein AGP1 to the AG-motif (AGATCCAA) of the wound-inducible *Myb* gene *NtMyb2*. *The Plant Journal* **36**:550-564.

Tang, Z., Sadka, A., Morishige, D. and Mullet, J. 2001. Homeodomain leucine zipper proteins bind to the phosphate response domain of the soybean VspB tripartite promoter. *Plant Physiol.* **125**:797-809.

Tasanen, K., Oikarinen, J., Kivirikko, K. and Pihlajaniemi, T. 1992. Promoter of the gene for the multifunctional protein disulfide isomerase polypeptide: functional significance of the 6 CCAAT boxes and other promoters elements. *J. Biol. Chem.* **267**:11513-11519.

Thomashow, M. 1998. Role of cold-responsive genes in plant freezing tolerance. *Plant Physiol.* **118**:1-7.

Thomashow, M. 2001. So what's new in the field of plant cold acclimation? Lots! *Plant Physiol.* **125**:89-93.

Trainotti, L., Zanin, D. and Casadoro, G. 2003. A cell wall-oriented genomic approach reveals a new and unexpected complexity of the softening in peaches. *Journal of Exp. Botany* **54**:1821-1832.

- Tucker, M., Whitelaw, C., Lyssenko, N. and Nath, P. 2002. Functional analysis of regulatory elements in the gene promoter for an abscission-specific cellulase from bean and isolation, expression, and binding affinity of three TGA-type basic leucine zipper transcription factors. *Plant Physiol.* **130**:1487-1496.
- van Doorn, W. and Stead, A. 1997. Abscission of flowers and floral parts. *Journal of Exp. Botany* **48**:821-837.
- Von Mollendorf, L. 1987. Wooliness in peaches and nectarines: a review. 1. Maturity and external factors. *Hort. Science* **5**:1-3.
- Werner, T. 2002. Promoters can contribute to the elucidation of protein function. *TRENDS in Biotech.* **21**:9-13.
- Westwood, 1997. Fruit growth and thinning. In: Temperate Zone Pomology (W.H. Freeman and Company, San Francisco). Pp:199-212.
- White, J., Todd, J., Newman, T., Focks, N., Girke, T., Martínez de Ilárdyula, O., Jaworski, J., Ohlrogge, J. and Benning, C. 2000. A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. *Plant Physiol.* **124**:1582-1594.
- Xiang, C., Wang, H., Shiel, P., Berger, P. and Guerra D.J. 1994. A modified alkaline lysis miniprep protocol using a single microcentrifuge tube. *Biotechniques* **17**:30-32.
- Yamagata, H., Yonesu, K., Hirata, A. and Aizono, Y. 2002. TGTCACA motif is a novel *cis*-regulatory enhancer element involved in fruit-specific expression of the *cucumis* gene. *J. Biol. Chem.* **277**:11582-11590.
- Yamaguchi-Shinozaki, K. and Shinozaki, K. 1994. A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *The Plant Cell* **6**:251-264.
- Yu, X., Griffith, M. and Wiseman, S. 2001. Ethylene induces antifreeze activity in winter rye leaves. *Plant Physiol.* **126**:1232-1240.
- Zhou, H., Sonogo, L., Khalchitski, A., Ben-Arie, R., Lers, A. and Lurie, S. 2000a. Cell wall enzymes and cell wall changes in 'Flavortop' nectarines: mRNA abundance, enzyme activity, and changes in pectic and neutral polymers during ripening and in woolly fruit. *Journal of the Amer. Soc. For Hort. Sci.* **125**:630-637.
- Zhou, H., Lurie, S., Lers, A., Khatchitski, A., Sonogo, L. and Ben-Arie, R. 2000b. Delayed storage and controlled atmosphere storage of nectarines: two strategies to prevent wooliness. *Postharv. Biol. Technol.* **18**:133-141.
- Zhou, H., Lurie, S., Ben-Arie, R., Dong, L., Burd, S., Weksler, A. and Lers, A. 2001. Intermittent warming of peaches reduces chilling injury by enhancing ethylene production and enzymes mediated by ethylene. *J. Hort. Sci. Biotechnol.* **76**:620-628.

Zolman, B., Monroe-Augustus, M., Thompons, B., Hawes, J., Krukenberg, K., Matsuda, S. and Bartel, B. 2001. *chy1*, an *Arabidopsis* mutant with impaired beta-oxidation, is defective in a peroxisomal beta-hydroxyisobutyryl-CoA hydrolase. *J. Biol. Chem.* **276**:31037-31046.