

## Tesis Doctoral

# Búsqueda de genes candidatos que controlen QTLs involucrados en la resistencia al estrés hídrico mediante el análisis de perfiles transcripcionales en especies silvestres de Solanum

Lucca, A. María Florencia

2011

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Lucca, A. María Florencia. (2011). Búsqueda de genes candidatos que controlen QTLs involucrados en la resistencia al estrés hídrico mediante el análisis de perfiles transcripcionales en especies silvestres de Solanum. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Lucca, A. María Florencia. "Búsqueda de genes candidatos que controlen QTLs involucrados en la resistencia al estrés hídrico mediante el análisis de perfiles transcripcionales en especies silvestres de Solanum". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2011.

**EXACTAS** UBA

Facultad de Ciencias Exactas y Naturales



**UBA**

Universidad de Buenos Aires



**UNIVERSIDAD DE BUENOS AIRES**

**FACULTAD DE CIENCIAS EXACTAS Y NATURALES**

*Búsqueda de genes candidatos que controlen QTLs  
involucrados en la resistencia al estrés hídrico mediante el  
análisis de perfiles transcripcionales en especies silvestres de  
Solanum*

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires  
en el área **CIENCIAS BIOLÓGICAS**

**Ing. Agr. A. María Florencia Lucca**

Director de tesis: H. Esteban Hopp

Consejero de Estudios: H. Esteban Hopp

Lugar de trabajo: Instituto Vasco de Investigación y Desarrollo Agrario, NEIKER-  
Tecnalia, España; EEA INTA Balcarce e Instituto de Biotecnología, CICVyA, INTA  
Castelar.

Junio de 2011

# AGRADECIMIENTOS

Es increíble pero cierto, cierro la etapa pendiente de la Tesis!

Agradezco a mi familia que me apoyó en todo, toda mi vida.

Merece un agradecimiento super-especial, Apu, a quien le debo su apoyo particularmente en esta última etapa. Quien me facilitó todo lo que pude necesitar, para que finalmente este escribiendo esta sección de agradecimientos. Me dio mucha mucha serenidad, contención, sabiduría, afecto y un hogar en estos últimos meses. MUCHAS GRACIAS! Y a sus hijos que hicieron mis días más alegres y que se preocupaban por verme escribiendo sin parar y colonizando su cuarto. Marti, Joaquín y Santi, van a recuperar su habitación muy pronto!

A mis amigos por su apoyo siempre y que saben cada detalle de los pase durante esta Tesis. A Mery y su familia, que adoro y que me llamaban hasta último momento antes de entregar. A Guada y su familia y las agradables cenas que disfrutamos en su casa. A mis amigas de Euskadi, las hermanas Albaina, Josune y Esti que estimo muchísimo y a nuestra gran amiga Marta, que aunque no estés con nosotros ahora, te vas a poner super feliz. Josune, andá poniéndole fecha a tu pasaje para venir a recorrer Argentina, ya estoy libre!

A **Rocío**, que tendría que figurar con negritas (más en negritas aún) en esta Tesis por el apoyo que me brindó todo este tiempo con los *microarrays*. Era fácil terminar delirando sobre todo lo que podíamos y podemos hacer aún con las matrices de datos, pero salió bonito, no? y todavía nos quedan ideas en el tintero. A Cristina y Coral que conocí por su intermedio que imagino les hará mucha ilusión que presente este Trabajo.

A Esteban Hopp por su apoyo, a pesar de ser un electrón libre.

A toda la gente de Biotecnología de Castelar, que me acompañó durante este camino, estando aquí o en Euskadi, a especialmente a Marisa Farber y su flia

que han sido una familia para mí en Baires, a Lucila, a Paula F., Marisa L. B., Virginia W., Ruth, Tere, Valeria B., Norma, Seba, Diego Lija y tantos tantos más. A la Institución que me brindó la posibilidad de realizar esta Tesis y a Lela por autorizarme trabajar fuera de hora en esta última etapa.

A Vero Nishi, que se preocupó siempre donde sea que estuviera, por saber qué era de mi vida.

Una mención especial para Luisito que lo bombardee con pedido de papers, creo que unas botellas de vino no alcanza, ya te debo una bodega con la cantidad de mails que te mandé, gracias.

A Natalia Almasia, por hacerme sentir parte del lab y por su darme ánimo con la presentación.

A la Polaca por su gran ayuda con las real time y por su buena onda.

A Ceci Vazquez por sus correcciones y gran apoyo.

A Dani Tosto por ser incondicional y por preocuparse por mi trabajo y especialmente por mi persona.

A toda la gente de Neiker que me dio una mano en todo lo que necesite y que me brindó su amistad y a la Institución por brindarme la posibilidad de realizar gran parte de esta Tesis. A Joaquín Salazar por atender mis consultas. A Enrique por recibirme en su grupo. A la incondicional de Elena B. y su familia por su amistad y que siempre estuvo apoyándome. A Mónica que siempre estimaré especialmente. A los chicos del lab Iratxe, Inés, Iñigo, Gorka, Isi, Rakel y Judith. A Ana Aragonés y su familia, Ana H., Natxo, Txekos. A las secretarias y los chicos de campo.

A Susana García por darme una mano para tener todos los slides que requería y por contactarme con Angel Dominguez.

Al Dr. Angel Dominguez por brindarme la posibilidad de utilizar las instalaciones de su laboratorio. A Nuria, Ana y Rita les agradezco su amistad y los momentos copados que pasamos juntas. A Leticia, por ser una gran amiga que me acompañó durante toda mi estancia allí.

A todos, MUCHAS GRACIAS!

# 1- ÍNDICE

# ÍNDICE

Resumen.....	8
Abstract.....	9
Abreviaturas.....	10
2 - INTRODUCCION.....	15
El cultivo de papa.....	15
El origen de la papa .....	15
La papa en poblaciones amerindias .....	16
Descripción botánica y morfológica.....	17
Distribución de la papa.....	18
El cultivo de papa en el mundo.....	19
El cultivo de papa en Argentina.....	21
Genética y recursos genéticos de la papa.....	22
Genómica de la papa.....	22
Recursos genómicos de la papa.....	23
Especies silvestres y cultivadas.....	23
Las plantas y los estreses abióticos.....	24
Estrés hídrico en plantas.....	27
Respuestas bioquímicas y fisiológicas de las plantas a la sequía.....	28
Señalización celular y regulación génica.....	29
Señalización por mensajeros secundarios.....	32
Señalización de fosfolípidos.....	32
Estrés hídrico en papa.....	32
Efecto del contenido del agua en la calidad de la papa.....	34
Interacción del estrés hídrico con otros estreses abióticos y bióticos.....	35
Control del estrés hídrico en papa.....	35
Herramientas de Biología Molecular y Bioinformática.....	36
Análisis de la expresión genética.....	36
Avances Tecnológicos.....	36
Tecnología de micromatrices.....	37
Micromatrices <i>Potato 10K</i> .....	39
Biología computacional y Bioinformática.....	39
Metodología para la identificación de perfiles de expresión diferencial.....	41
Conceptos GO.....	40
3 - OBJETIVOS E HIPÓTESIS.....	43
4 - MATERIALES Y METODOS.....	46
Material biológico.....	46
Ensayos de estrés hídrico.....	47

Herramientas de Biología molecular.....	49
Procesamiento del tejido vegetal .....	49
Extracción de ARN.....	49
Micromatrices de ADNc.....	49
Análisis de Expresión de las especies <i>S. cardiophyllum</i> (CPH) y <i>S. venturii</i> (VNT).....	51
Análisis de Expresión de la especie <i>S. pascoense</i> (PSC).....	52
Purificación del ARN para su utilización en ensayos de micromatrices.....	52
Síntesis de cDNA por marcación directa.....	52
Purificación, cuantificación y concentración del cDNA.....	54
Pre-hibridación, hibridación y lavado de micromatrices.....	55
Diseño del ensayo de micromatrices.....	57
Herramientas de Bioinformática.....	59
Escaneo de chips.....	59
Pre-procesamiento de micromatrices.....	59
Procesamiento de micromatrices.....	59
Construcción de la base de datos estructurada.....	60
Validaciones biológicas RT-PCR.....	61
Extracción de ARN total.....	63
Tratamiento de ADNasa I.....	63
Síntesis de ADNc.....	63
PCR cuantitativa o en tiempo real.....	65
Herramientas de visualización de datos de expresión génica.....	66
5 - RESULTADOS.....	69
Evaluación de especies de <i>Solanum</i> silvestres frente al estrés hídrico.....	69
Análisis de la expresión génica diferencial de especies silvestres de <i>Solanum</i> .....	76
Hibridación de micromatrices 10K de papa (TIGR).....	77
Hibridaciones en la Plataforma Genómica de la Universidad de Salamanca.....	79
Análisis bioinformático de los datos de micromatrices.....	81
Análisis exploratorio inicial de los datos de expresión génica.....	86
<i>Solanum pascoense</i> .....	104
Análisis bioinformático con el programa Limma.....	106
<i>Solanum cardiophyllum</i> .....	112
Metabolismo de hidratos de carbono.....	115
Genes asociados a La Pared celular.....	115
Glucólisis.....	115
Gluconeogénesis.....	116
Metabolismo de aminoácidos.....	116
Transporte de iones metálicos.....	116
Metabolismo secundario.....	116

Proteínas de respuesta al estrés.....	117
Procesos Rédox Celulares.....	117
Regulación de la transcripción.....	117
Metabolismo de proteínas.....	118
Señalización celular.....	119
Estructura y división celular.....	119
Transporte celular.....	119
Desarrollo y almacenamiento.....	119
<i>Solanum venturri</i> .....	121
Fotosíntesis.....	125
Metabolismo de proteínas.....	126
Pared celular.....	128
Metabolismo de hidratos de carbono.....	129
Glucólisis.....	131
Metabolismo de hormonas.....	132
Transporte de iones metálicos.....	133
Proteínas LEA.....	134
Proteínas de respuesta al estrés.....	134
Señalización.....	134
Procesos Rédox Celulares.....	135
Validaciones mediante qPCR.....	136
6 - DISCUSIÓN.....	146
7 - CONCLUSIONES.....	164
Perspectivas futuras.....	165
8 - BIBLIOGRAFÍA.....	167

## RESUMEN

La papa (*Solanum tuberosum* ssp. *tuberosum*) es el tercer cultivo más importante y la dicotiledónea de mayor importancia en la alimentación. El estrés hídrico es la mayor limitante de la producción. En respuesta a la sequía las plantas activan estrategias que involucran adaptaciones morfológicas, fisiológicas y bioquímicas controladas por un número de genes aditivos y en muchas ocasiones sinérgicos.

Esta tesis se focalizó en el análisis de los cambios transcripcionales de las especies silvestres de *Solanum*: *S. venturii* (VNT) y *S. cardiophyllum* (CPH) que presentan comportamientos contrastantes frente al estrés hídrico mediante hibridación con micromatrices de ADNc. Se analizaron 1.033.250 datos de expresión de 32 soportes de cristal de papa, que monitorearon 10 días de evolución del estrés. El análisis bioinformático de los datos de expresión de estas especies arrojó diferencias significativas en los perfiles transcripcionales. El análisis comparativo de estos perfiles reveló que en CPH se ven afectados por el tratamiento de sequía un total de 235 genes, mientras que en VNT la diferenciación involucra 4133 genes. Ambas especies presentan una expresión de 151 genes comunes. Los genes expresados diferencialmente no solo variaron en el número sino también en sus perfiles temporales de expresión. Los genes diferencialmente expresados como respuesta al tratamiento tendieron a agruparse mostrando comportamientos particulares en ambas especies. En CPH, los perfiles de expresión permitieron identificar al menos 5 conglomerados que resumen las principales tendencias del comportamiento de los transcritos frente al estrés, mientras que en VNT la dinámica de expresión mostró expresiones contrastantes y agrupadas, que sugieren una regulación común sincrónica, confirmando la diversidad genética de respuestas en el germoplasma de especies silvestres de papa. Estas diferencias de expresión entre las especies silvestres de papa sugieren nuevas vías para explotar la variabilidad genética del germoplasma silvestre para el mejoramiento del cultivo. La información generada brinda interesantes genes candidatos involucrados en respuesta al estrés hídrico a la vez que permite identificar secuencias para el desarrollo de marcadores funcionales útiles en programas de mejoramiento asistido.

**Palabras claves:** *papa, sequía, microarrays, MapMan, germoplasma.*

# ABSTRACT

The potato is the third most important crop and the most relevance dicot in feeding. Water stress is the greatest limitation of production. In response to drought, plant triggers strategies that involve morphological, physiological and biochemical changes controlled by a number of additive genes and often synergistic.

This Thesis focused on the analysis of transcriptional profiling on *Solanum* wild species: *S. venturii* (VNT) and *S. cardiophyllum* (CPH) that present contrasting behavior under water stress using cDNA microarray hybridizations. We analyzed 1,033,250 expression data of 32 potato chips, which monitored 10 days of stress treatment. Bioinformatic analysis of expression data of wild species showed significant differences in transcriptional profiles. The comparative analysis of these profiles revealed that CPH differentially expressed under drought 235 genes, while VNT involves 4133 differentially genes. Both species have a common expression of 151 genes. Differentially expressed genes vary not only in number but also in their temporal expression profiles. The genes differentially expressed in response to stress tended to cluster following a particular behaviors in both species. In CPH, the expression profiles helped to identify at least 5 clusters that summarize the main trends of behavior of the transcripts under stress, while the dynamics of expression in VNT showed contrasting clustered patterns, suggesting a common regulation synchronously, confirming the genetic diversity of wild potato species responses. These differences in expression between wild potato species suggest new ways to exploit the genetic variability of wild germplasm for crop breeding. The information generated provides interesting candidate genes involved in response to water stress as well as to identify sequences for the development of useful functional Marker Aided Selection (MAS) in Plant Breeding programs.

**Key words:** *potato, drought, microarrays, MapMan, germoplasm*

## ABREVIATURAS

µg: microgramo

µl: microlitro

ABA: ácido abscísico (del inglés, *abscisic acid*)

ABA-dependientes

ABA-independientes

ABRE: elemento de respuesta a ABA (del inglés, *abscisic acid responsive*)

ABRE y DRE/CRT: del inglés *dehydration responsive element/C-repeat*

Ac: Acetato

ADN: ácido desoxirribonucleico

ADNc: ADN complementario

AIP: Año Internacional de la Papa

ARN: ácido ribonucleico

ARNm: ARN mensajero

ARNr: ARN ribosomal

BER: *Solanum berthaultii*

BLAST: herramienta de búsqueda de alineamiento de secuencias de tipo local, del inglés *Best Local Alignment Search Tool*

BLB: *Solanum bulbocastanum*

BSA: Seroalbúmina bovina.

bZIP: del inglés basic region/leucine zipper motif

CBP: CREB binding protein

CIP: Centro Internacional de la Papa

CD: Disco compacto

CPH: *Solanum cardiophyllum*

CREB/CBP: del inglés CREB: cAMP response element-binding

CTAB: Bromuro de Hexadeciltrimetilamonio (Cetrimonium Bromide)

DAG: diacilglicerol

DDS: *Solanum doddsii*

DEPC: dietilpirocarbonato

dNTP: desoxirribonucleótidos trifosfato

DO: densidad óptica

DRE: elemento de respuesta a la deshidratación (del inglés drought)  
DREB: proteína de unión a DRE  
DTT 1,4 ditiotreitól  
EDTA: ácido etilendiaminotetraacético  
ERD: respuesta temprana a la deshidratación, del inglés *Early-responsive to dehydration*  
ESTs: Marcador de secuencia expresada (Expressed sequence tags)  
familias MAP (del inglés mitogen activated protein)  
Factores de transcripción (FTs)  
FOI: Frecuencia de Incorporación (del inglés *Frequency of Incorporation*)  
GND: *Solanum glandulosum*  
H<sub>2</sub>O agua destilada  
Ha: Hectárea  
HD-ZIP  
IFD: *Solanum infundibuliforme*  
Interconexión: del inglés *crosstalk*  
IP3: Inositol fosfato 3  
JA: ácido jasmónico  
Kb: kilobase  
LEA: del inglés, *late embryogenesis abundant*  
LiCl: Cloruro de Litio  
LIMMA: Limma: del inglés: *Linear Models for Microarray Data*  
Log: logaritmo  
LogFC: del inglés *Log Fold Changes*  
LPH: *Solanum leptophyes*  
M: molar  
m: metro  
M: micro molar  
min: minuto  
ml: mililitro  
mM milimolar  
mm: milímetro  
mM: milimolar  
MOPS 3-N-morfolino ácido propanesulfónico

MRN: *Solanum marinasense*

nm nanómetro

P/V: Peso/Volumen

pb pares de bases

PCA: análisis de componentes principales, del inglés *Principal Component Analysis*

PCR: Reacción en Cadena de la Polimerasa, del inglés *Polimerase chain Reaction*

PLT: *Solanum polytrichon*

PSC: *Solanum pascoense*

PTA: *Solanum papita*

qPCR: Reacción en Cadena de la Polimerasa Cuantitativa, del inglés *Quantitative Polimerase Chain Reaction*

QTL: Loci de Caracter Cuantitativo, del inglés *Quantitative Trait Loci*

ROS: especies reactivas de oxígeno, del inglés *Reactive Oxygen Species*

rpm: revoluciones por minuto

RT: Transcripción Reversa

RT-PCR: Reacción de la Cadena de Polimerasa en Transcripción Reversa

RWC: Contenido relativo de agua, del inglés *Relative Water Content*

SA: ácido salicílico

SCT: *Solanum sanctae-rosae*

SDS: dodecilsulfato sódico

Seg: segundo

SOS: extrema sensibilidad a la sal, del inglés *salt-overlyensitive*

SPL: *Solanum sparsipilum*

StGI: *Solanum tuberosum Gene Index*

TIGR: El Instituto de Investigación Genómica del inglés *The Institute for Genomic Research*, hoy parte del *J. Craig Venter Institute*

Tris: tris (hidroximetil) aminometano

USA: Estados Unidos de América

VNT: *Solanum venturii*



## **2- INTRODUCCIÓN**

## 2- INTRODUCCIÓN

### **El cultivo de papa**

La papa (*Solanum tuberosum*) es el tercer cultivo más importante en el mundo tras el cultivo del trigo y arroz (FAO, 2008) por el alto contenido de almidón de sus tubérculos. La especie *S. tuberosum* se divide en dos subespecies, *andigena*, adaptada a condiciones de días cortos que se cultiva principalmente en los Andes, y *tuberosum*, la variedad que hoy se cultiva en el resto del mundo.

La papa es miembro clave de la familia Solanaceae (Juss.), junto con el pimiento, el tomate, la berenjena y el tabaco. Asimismo, es la dicotiledónea de mayor importancia en la alimentación humana, constituyendo un alimento ideal en países en vías de desarrollo y más aún en las regiones con altitudes superiores a 2000 metros, en las cuales hay pocas opciones de cultivos y donde el hombre y los animales requieren un aporte calórico mayor.

La Asamblea General de las Naciones Unidas declaró el 2008 como el Año Internacional de la Papa (AIP), dando un marco ideal para focalizar la atención del mundo hacia el rol que puede desempeñar la papa en la lucha contra el hambre y la pobreza (FAO, 2008). Las papas constituyen un recurso decisivo para la seguridad alimentaria de cientos de millones de personas del mundo en desarrollo.

Este cultivo tiene un valor social importante, se ha considerado a la papa como un “Tesoro enterrado”, ya que las personas de bajos ingresos de las zonas urbanas y rurales, encuentran en la papa un cultivo de rápido crecimiento, gran plasticidad y que tiene una buena producción aún con pocos insumos, siendo ideal para lugares donde hay limitante de tierras y donde abunda la mano de obra. Además, tiene un gran potencial de producción y productividad aún inexplorado, que hacen que este tesoro adquiera un valor especial en tierras marginales, inadecuadas para la producción de otros cultivos.

### **El origen de la papa**

La papa se originó hace unos 8000 años. Milenios de cultivo tornaron oscuro su origen. El origen de las papas cultivadas se remonta a las variedades nativas

que los agricultores precolombinos desarrollaron a partir de especies que crecían en estado silvestre. Los botánicos rusos Juzepcczuk y Bukasov demostraron que entre los antepasados de la papa se encontraban diferentes especies silvestres (Rouselle *et al*, 1999).

El género *Solanum* contiene numerosas especies, distinguiéndose alrededor de 5500 variedades (FAO, 2008). Su distribución es muy amplia en todo el mundo, evidenciándose una fuerte concentración de especies en América del Sur y América Central. De este amplio género, sólo el 10% está constituido por especies tuberosas. Están reportadas alrededor de 200 especies de este género distribuidas en 21 series taxonómicas (Hawkes, 1990).

El hábitat de estas especies es muy variable. Crecen a diferentes altitudes, desde el nivel del mar hasta más de 4000 metros y lo hacen bajo una gama extensa de temperaturas, fotoperíodos y suministros de agua. Asimismo, los genotipos silvestres se adaptan a condiciones ecológicas muy diversas que van desde tierras altas, trópicos hasta regiones desérticas (Rouselle *et al*, 1999).

El centro de máxima variabilidad de éstas especies se localiza en los Andes, principalmente en Perú y Bolivia, dónde han sido detectadas 100 especies silvestres y más de 400 cultivares de papa andígenas (Rouselle *et al*, 1999; Burton, 1996).

Las papas cultivadas son originarias de América del Sur. El número básico de cromosomas es 12. Se presentan distintos grados de ploidía, desde diploides ( $2n=24$ ) hasta hexaploides ( $2n=74$ ), siendo la papa cultivada, *Solanum tuberosum*, tetraploide ( $2n=48$ ).

El cultivo de papas diploides se conserva en la actualidad en tierras altas de Perú y Bolivia. Estas especies están en constante evolución cruzándose entre sí y con otras *Solanum* silvestres, conservando una gran variabilidad genética. Se cree que la domesticación se ha iniciado en esta región, y se ha difundido a lo largo de la cordillera andina (Correll, 1962).

### **La papa en poblaciones amerindias**

Se han encontrado vestigios de la alfarería de la cultura Mochica (Siglos I a VII) que presentaban imágenes de papa, que era considerada un elemento de importancia religiosa en esta civilización. En particular, en las civilizaciones preincaicas, el sentimiento religioso era muy profundo. Entre las deidades a las

cuales rendían culto se encontraba la diosa Tierra o *Pachamama*. Las plantas adquirirían un renombre, al punto de ser consideradas seres divinos. A la madre de la papa se la denominada *Axomama*, a la del maíz *Zaramama* y recibía el nombre de *Quinuamama*, la madre de la quínoa (Rouselle *et al*, 1999).

### **Descripción botánica y morfológica**

La papa (*Solanum tuberosum* L.) fue descrita en 1753 por Linneo (<http://www.botanicus.org/title.aspx?bibid=b12069590#>). Es una especie herbácea dicotiledónea, perenne, pero cultivada frecuentemente como planta anual.

Presenta varios tallos principales, que inicialmente se presentan erguidos y que con el progreso del ciclo vegetativo pueden volverse parcial o totalmente rastreros, brindando un porte extendido (Figura 1).

Las hojas presentan una disposición alterna helicoidal (índice filotáxico 5/13), que gira hacia la izquierda (sinistrosa). Las hojas son compuestas. Presentan un pecíolo seguido por grandes folíolos, dispuestos de a pares a lo largo del raquis, que culmina en un folíolo único. La nerviación de las hojas es de tipo reticulado.

Las flores están agregadas en inflorescencias cimosas o definidas, ubicadas en el extremo terminal del tallo y sostenida por el pedúnculo. Las flores son actinomorfas y pentámeras (Plaisted, 1982).

El fruto es una baya verde o pardo-violácea, que generalmente presenta varias semillas de pequeño tamaño contenidas en una malla mucilaginosa. Las bayas son esféricas u ovoidales, de escaso diámetro, ausente en varios cultivares (Rouselle *et al*, 1999) y bicarpelados. La semilla no se utiliza en la producción agrícola, sino que principalmente es un recurso utilizado en la obtención de variedades.

Los tubérculos conforman un sistema subterráneo originado a partir de estolones y como consecuencia de la acumulación de reservas debidas al rápido crecimiento y división celular (Burton, 1989). El 75-85% de la materia seca total se acumula en los tubérculos. Es no sólo un órgano de reserva, sino que sirve para la multiplicación vegetativa. Se distinguen por su forma, color de carne, color y textura de su piel y por los ojos.

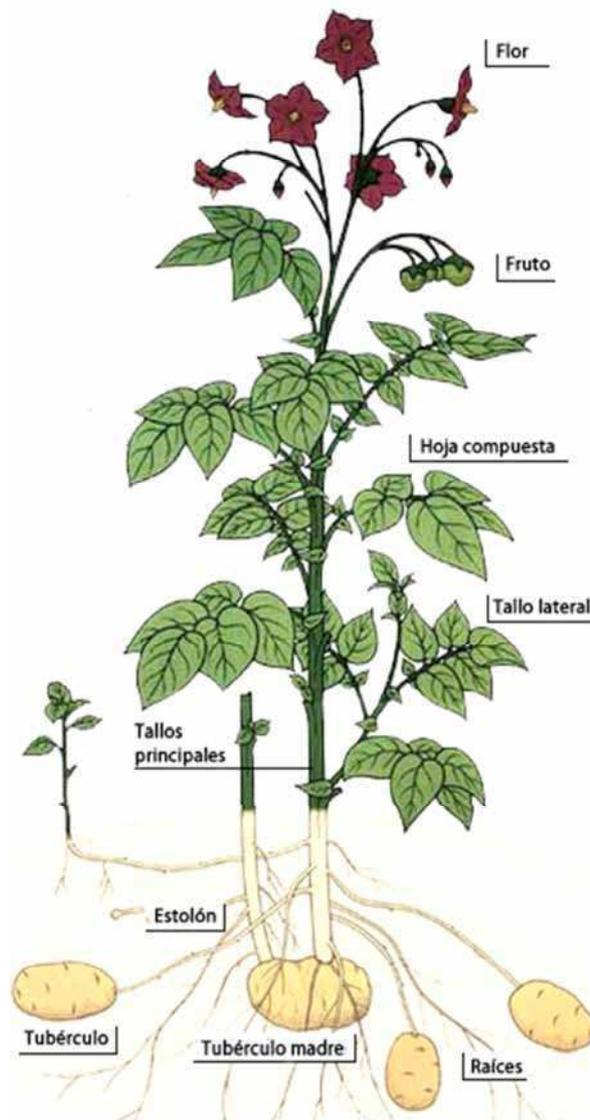


Figura 1. Esquema de la planta de papa. Diseño: © CIP.

### Distribución de la papa

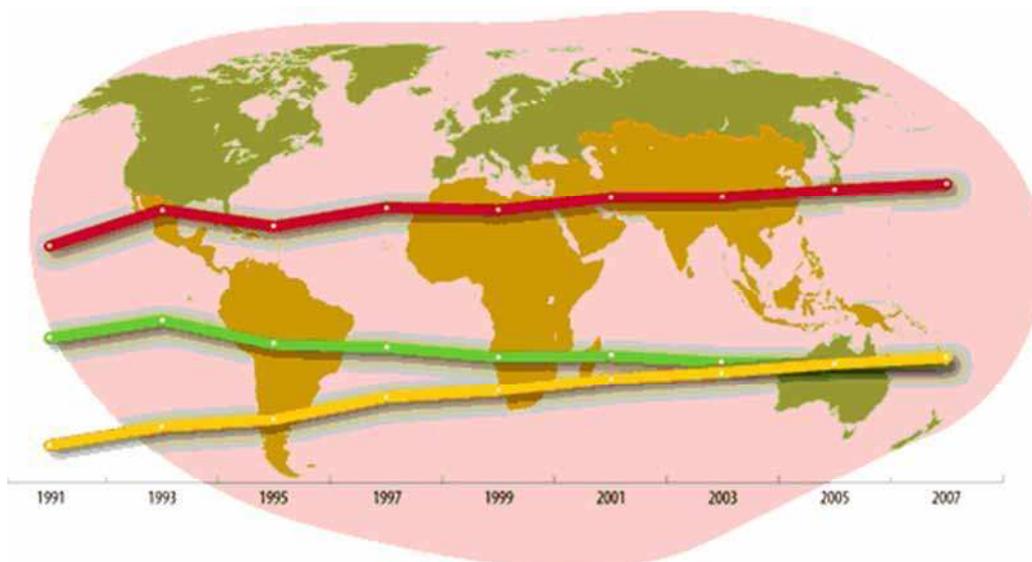
La papa fue llevada desde las tierras altas de los Andes de Sudamérica a Europa por los conquistadores españoles alrededor de 1570, distribuyéndose en numerosas áreas del norte de Europa, en especial Irlanda, donde a mediados de 1700 gran parte de la población basaba su dieta exclusivamente en esta hortaliza y leche. Desde allí fue introducida en América del Norte y cultivada en las colonias. El avance desde Europa siguió hacia otras partes del mundo y hoy se producen papas en una superficie estimada de 180.000 kilómetros cuadrados, desde la planicie de Yunnan en China y las tierras bajas

subtropicales de la India, hasta las montañas ecuatoriales de Java y las estepas de Ucrania. Es por tanto, la dicotiledónea de mayor importancia en la alimentación humana, constituyéndose en una fuente energética entre dos a cuatro veces superior al maíz y entre seis, diez o más veces que los cereales finos. Por su parte, el rendimiento en proteína por unidad de área excede al trigo, arroz y maíz (Hooker, 1980). Hasta un 85% de la planta es comestible para las personas, en comparación con el 50% en el caso de los cereales. Cientos de plagas y patógenos han sido reconocidos en el cultivo de papa. Algunos de estos patógenos fueron transportados a nuevos lugares junto con el cultivo y otros se encontraban presentes en lugares donde la papa fue introducida y proliferaron en ella convirtiéndola en un nuevo huésped. La propagación vegetativa por tubérculos hace de la papa un vehículo de plagas y patógenos que han seguido al cultivo dondequiera que crezca. Entre las principales enfermedades que afectan al cultivo de papa encontramos el Tizón Tardío del papa causado por *Phytophthora infestans*, también el cultivo es afectado por bacterias como *Ralstonia solanacearum* y numerosas virosis (PVY, PVX, PVS, PLRV) que afectan su producción y calidad, (Hooker, 1980; CIP, 1999).

### **El cultivo de papa en el mundo**

La superficie cultivada de papa a nivel mundial ha ido disminuyendo en los últimos años. La superficie cultivada de papa en los países más desarrollados se ha reducido debido a la demanda intensiva de la labor para la siembra, la recolección, el transporte y el cambio en los hábitos de consumo.

No obstante, el rendimiento y la producción mundial han aumentado, en parte debido a las mejoras en las técnicas de cultivo. En Europa la producción de papa ha disminuido a razón del 1% anual, mientras que China y la India han mostrado un incremento muy significativo (5% anual) que ha llevado a que la producción mundial casi duplique su producción en los últimos 20 años (FAO, 2010). En 2005, por primera vez, la producción de papa de los países en desarrollo superó a la de los países desarrollados (Figura 2, Tabla 1).



**Figura 2. Producción mundial de papa entre 1991-2007.** (FAO, 2010). En rojo se muestra la Producción mundial de papa, en verde, la producción en países desarrollados y en amarillo la de países en desarrollo según la Tabla 1.

**Tabla 1.** Producción mundial de papa, 1991-2007 (FAO, 2010)

	1991	1993	1995	1997	1999	2001	2003	2005	2007
Países	millones de toneladas								
<b>Desarrollados</b>	183,13	199,31	177,47	174,63	165,93	166,93	160,97	159,97	159,89
<b>En desarrollo</b>	84,86	101,95	108,50	128,72	135,15	145,92	152,11	160,01	165,41
<b>MUNDO</b>	267,99	301,26	285,97	303,35	301,08	312,85	313,08	319,98	325,30

Fuente: FAOSTAT

Hace 20 años, los países en desarrollo producían apenas por encima del 20 por ciento de la producción actual. Aun así, la producción y el consumo mundial de papas crecen a tasas inferiores que la que lo hace la población (FAO, 2010). Desde principios de los años 60, dice la FAO, el aumento de la superficie dedicada al cultivo de papa en el mundo en desarrollo ha superado la de todos los demás productos alimentarios básicos, y se prevé que para 2020 la demanda de papa duplique la de 1993.

Europa y Asia se presentan como las principales regiones productoras de papa del mundo, concentrando en 2007 el 80% de la producción mundial (Tabla 2). Por su parte en África y América Latina aunque con un volumen menor, presentaron producciones extraordinarias. En cuanto al consumo de papa, Asia concentra casi la mitad del consumo mundial de papa, mientras que África

y América Latina, presentan los menores consumos, aunque con tendencia creciente (FAO, 2010).

**Tabla 2. Principales productores mundiales de papa, 2007 (FAO, 2010)**

	Cantidad (t)
1.  China	72 040 000
2.  Fed. de Rusia	36 784 200
3.  India	26 280 000
4.  Estados Unidos	20 373 267
5.  Ucrania	19 102 300
6.  Polonia	11 791 072
7.  Alemania	11 643 769
8.  Belarús	8 743 976
9.  Países Bajos	7 200 000
10.  Francia	6 271 000

Fuente: FAOSTAT

### **El cultivo de papa en Argentina**

La producción de papa en Argentina se estima alrededor de 2.013 millones de toneladas por año (FAO, 2010). El principal mercado de la papa es para consumo fresco, seguido en orden de importancia por la producción destinada a papa-semilla y la exportación (Corvo Dolcet, S, 2005).

Debido a la distribución regional de la producción de papa, Argentina históricamente ha abastecido la demanda interna y eventualmente ha exportado a países vecinos. Cerca del 50% de la producción nacional se centra en Balcarce, región sudeste de Buenos Aires. Esta región tiene el mayor impacto sobre los precios durante el año. Los últimos años señalan rendimientos promedios de 40 tons/ha. Las exportaciones anuales de 47.000 toneladas comprenden solo el 2% del total de la producción nacional. Argentina exporta principalmente a Uruguay, Brasil, Bolivia y Paraguay. Desde la creación del Mercosur, las exportaciones argentinas se incrementaron un 300% y continúan con tendencia positiva (Corvo Dolcet, S, 2005; Huarte, M. comunicación personal).

Un reciente relevamiento demuestra que la superficie estimada de papa para consumo fresco e industrial se ubica en 68.000 Ha (FAO, 2008). El consumo aparente de papa fresca del país se ubica en 55 kg de papa/hab/año ([www.argenpapa.com.ar](http://www.argenpapa.com.ar)), resultando un total de 2,13 millones de toneladas.

Del total de superficie plantada con papa para consumo, entre un 11 y un 15 % se destina a industrias, las cuales consumen cerca de 450.000 tn/año para

distintos productos, siendo el principal bastones pre-fritos congelados.

### **Genética y recursos genéticos de la papa**

Los primeros estudios genéticos sobre la papa fueron reportados hace aproximadamente 100 años (Salaman, 1926), quien intentó aplicar los principios mendelianos. En ese tiempo no se tenía conocimiento respecto de la tetraploidía y de la herencia tetrasómica de la papa, que obstaculizaron los avances en éste área.

Para sortear estos problemas, con el tiempo se desarrollaron mapas genéticos sobre la base de la manipulación de los niveles de ploidía. Así fue como se generaron mapas en diploides, y se obtuvo un mapa de ligamiento ultradenso, sentando los avances realizados en la caracterización del genoma de papa (Bonierbale *et al*, 1988, Tanksley *et al*, 1992, Gebhardt *et al*, 1989, 1991, 2001). Los estudios avanzaron al nivel tetraploide (Hackett *et al*, 2003; Bradshaw *et al*, 2004) y de ellos se generaron numerosos marcadores moleculares de diverso tipo, que facilitaron la comparación de distintos mapas genéticos de papa y posibilitaron la comparación entre la papa y su pariente cercano, el tomate (*Solanum lycopersicum*).

### **Genómica de la papa**

La papa es un miembro clave en la familia Solanaceae. Su pariente cercano, el tomate, tiene un genoma similar en tamaño (Arumuganathan y Earke, 1991), y sus mapas genéticos han mostrado un alto grado de co-linealidad.

El genoma de papa que se encuentra en estado avanzado de secuenciación donde de manera complementaria se han secuenciado una planta diploide desarrollada a partir de la papa cultivada y un doble monoploide proveniente de otra papa cultivada del grupo phureja. El tamaño del genoma es de aproximadamente 850-1000 Mpb, el doble del tamaño del genoma de arroz (Yu, J. *et al*, 2002) y el séxtuple del de Arabidopsis (Greilhuber, *et al*, 2006). Por su parte el tomate presenta un genoma similar al de esta especie de papa, con el que muestra un alto nivel de colinealidad (Tanksley *et al*, 1992) y esta próxima la publicación de su genoma. Su diploidía y su tipo de reproducción promovieron su utilización como sistema modelo para estudios genéticos y genómicos.

El desarrollo de herramientas moleculares similares para otras especies solanáceas en el marco de iniciativas de investigación en solanáceas a nivel mundial y regional, como lo es la iniciativa LatSol, brindan un espacio promisorio para el futuro de la investigación en papa (<http://cniia.inta.gov.ar/lat-sol/>).

La información respecto de la estructura y organización del genoma de tomate parece ser aplicable a la papa, así como el bajo número de copias de secuencias repetidas, encontrado inicialmente en tomate y confirmado más tarde en papa (Bryan y Hein, 2008).

### **Recursos genómicos de la papa**

El índice de genes de la papa disponible (<http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=potato>) cuenta con 236740 ESTs contenidos en 31972 “contigs” y con algo más de 62000 secuencias únicas. La estructura y organización de los índices de genes de TIGR (Quackenbush *et al*, 2000) constituye un recurso importante para la investigación genómica de la papa. Se estima que comprende entre el 50-70% del total de genes de la papa (Bryan y Hein, 2008). Este recurso es una fuente importante para el descubrimiento de genes candidatos, de marcadores genéticos y para el desarrollo de micromatrices de ADN, hasta tanto la secuencia del genoma entero de la papa se depure y esté disponible.

### **Especies silvestres y cultivadas**

La papa cultivada tiene un acervo génico secundario muy rico conformado por especies silvestres relacionadas. Se han descrito más de 200 especies silvestres de la sección *Tuberarium*, que son el principal foco de atención en las expediciones destinadas a coleccionar germoplasma de *Solanum*. El Centro Internacional de la Papa (CIP), con sede en el Perú, tiene la colección más abundante de germoplasma de papa, que incluye 3.800 muestras de papas andinas tradicionales cultivadas y unas 1.500 muestras de cerca de 100 especies silvestres ([www.cipotato.org](http://www.cipotato.org)). A partir de este banco de nutrida diversidad genética se buscan determinar nuevos genes de importancia agronómica. Estos estudios tienen como objetivo incorporar rasgos nuevos a los cultivares elite usados en la producción y crear nuevas variedades mediante

una combinación de técnicas tradicionales de mejoramiento y métodos biotecnológicos.

La taxonomía de la papa cultivada y sus parientes silvestres ha sido objeto de estudio durante muchos años, principalmente por medio de observaciones morfológicas y en menor medida a través de métodos experimentales como la citogenética y experimentos de hibridación.

La base genética de la papa cultivada fuera de los Andes es muy estrecha, lo que hace difícil que el desarrollo de nuevas variedades de mayor rendimiento, o con mayor resistencia a los principales estreses bióticos y abióticos que sufre este cultivo. El uso estratégico de recursos genéticos basado en parientes silvestres originarios es fundamental para mejorar del cultivo. Esta fuente de diversidad genética puede ser incorporada mediante programas de mejoramiento genético a líneas y variedades comerciales.

### **Las plantas y los estreses abióticos**

El crecimiento y desarrollo de las plantas depende de factores genéticos y medio-ambientales (Levitt, 1980). Se habla de estrés cuando por exceso o por carencia de alguno/s de estos factores, el desarrollo y el crecimiento de las plantas se ve afectado.

Los estreses abióticos son la principal causa de pérdidas de rendimiento a nivel mundial, reportándose mermas por encima del 50% en los principales cultivos (Bray *et al*, 2000; Boyer, 1982). Los factores de estrés medioambiental como la sequía, la elevada la temperatura, la salinidad y el aumento de CO<sub>2</sub> se han convertido en una creciente amenaza para la agricultura sostenible. La gran preocupación actual se centra en los efectos que el cambio climático puede provocar sobre los recursos vegetales, la biodiversidad y la seguridad alimentaria mundial (Ahuja *et al*, 2010). El cambio climático impacta en todos los aspectos de la arquitectura de la planta y representan un serio desafío para el desarrollo de una agricultura sostenible ante el mantenimiento de las tasas de crecimiento de la población mundial (Meehl, 2007).

La adaptación de las plantas al estrés implica la activación de diversas respuestas fisiológicas, metabólicas y del desarrollo y son rasgos poligénicos, controlados por un número de genes aditivos y en muchas ocasiones sinérgicos (Bohnert *et al*, 1995).

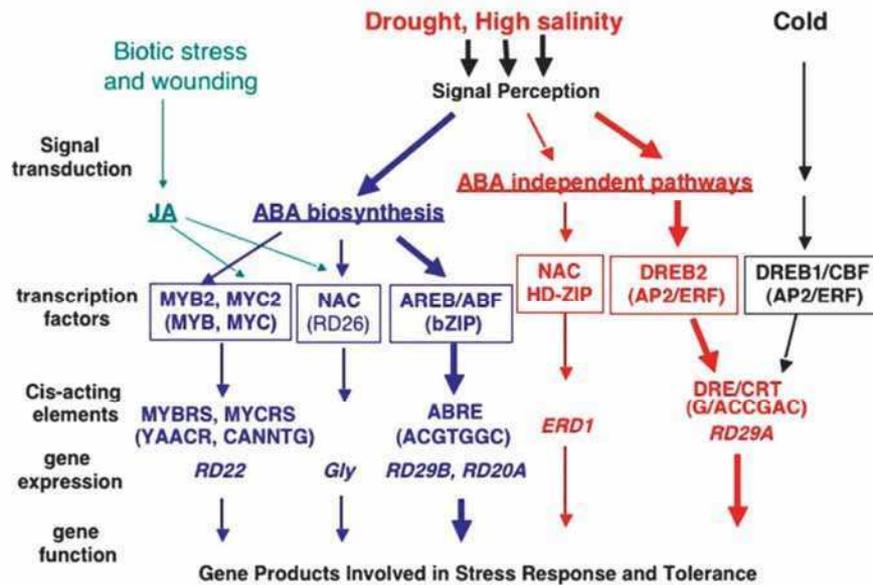
Estudios transcriptómicos, proteómicos y estudios de expresión de genes (Valliyodan y Nguyen, 2006; Shinozaki y Yamaguchi-Shinozaki, 2007, Pradeep *et al* 2006; Seki *et al.* 2001; Zhu, 2001, Shinozaki *et al*, 2003) identificaron dos grandes grupos de genes inducidos en respuesta a estreses abióticos.

Un grupo constituido por “proteínas funcionales” que incluye proteínas que protegen la membrana, proteínas de membrana que intervienen en el movimiento de agua a través de ésta (proteínas de canales de agua y transportadores de membrana), enzimas claves de la biosíntesis de osmo-protectores (prolina, betaína y azúcares), enzimas antioxidantes (hidrolasas, catalasas, superóxido dismutasa, etc.) y otras proteínas con función protectora de macromoléculas (proteínas LEA), osmotinas, chaperonas y proteínas asociadas a ARN, etc. Un segundo grupo está involucrado en la cascada de señales y en el control transcripcional y lo conforman proteínas reguladoras como factores de transcripción (bZIP, del inglés basic region/leucine zipper motif), CREB/CBP (del inglés CREB: cAMP response element-binding); CBP (CREB binding protein), HD-ZIP, proteínas Kinasas y Proteinasas, involucradas en transducción de señales y expresión génica.

Por su parte dentro de los genes inducidos frente a distintos estreses abióticos se encuentran los genes ABA-dependientes y los genes ABA-independientes, cuyos promotores contienen elementos ABRE y DRE/CRT (del inglés dehydration responsive element/C-repeat) respectivamente, habiéndose detectado asimismo genes con ambos tipos de elementos (Yamaguchi-Shinozaki y Shinozaki, 2005).

Diferentes fuentes de estrés (sequía, heladas, salinidad) disparan una respuesta en cierto modo única, que posee a la vez elementos comunes e idiosincrásicos respecto a otras respuestas y a vías metabólicas y morfo-genéticas del organismo. Las investigaciones realizadas hasta la fecha, estudian las vías de transducción de señales en respuesta a estreses (bióticos y abióticos) con relativa independencia, sobre todo porque era difícil considerar varios estímulos al mismo tiempo. Se han formulado modelos de una única vía lineal o con vía principal, con ramificaciones. Sin embargo, considerando que las plantas están expuestas a múltiples estreses, las señales de interconexión (del inglés *crosstalk*) entre las vías de transducción deberían permitir que las plantas coordinen y prioricen sus reacciones para sobrevivir con recursos

limitados. La superposición e interconexión entre tales redes de decisiones puede resultar sinérgica o antagónica. Estudios recientes sobre la interconexión entre estas cascadas de señales se están convirtiendo en un tema central en varios campos de investigación (Yoshioka y Shinozaki, 2009).



**Fig. 3. Redes de regulación transcripcional de las señales de estrés abiótico y la expresión génica.** (Shinozaki y Yamaguchi-Shinozaki, 2007) Al menos seis vías de transducción de señales existentes en la respuesta a la sequía, alta salinidad y al frío: tres son dependientes de ABA y tres son independientes de ABA. En la vía dependiente de ABA, ABRE funciona como un elemento importante que responde a ABA. AREB / ABF son factores de transcripción AP2 involucrados en este proceso. MYB2 y MYC2 funcionan en la expresión génica del gen rd22, ABA-inducible. MYC2 también actúa en la expresión génica JA-inducible. El factor de transcripción RD26 NAC está implicado en la expresión génica en respuesta a ABA y JA en respuesta al estrés. Los factores de transcripción MYC2 y NAC pueden actuar en la interconexión durante el estrés abiótico y el estrés en respuesta a heridas. En una de las vías independientes de ABA, DRE está principalmente implicado en la regulación de los genes no sólo por la sequía sino de estrés por sal y frío. DREB1/CBFs están involucrados en la expresión génica en la respuesta al frío. DREB2s son factores de transcripción importantes en la de la expresión de genes en respuesta a la deshidratación y la alta salinidad. Otra vía independiente de ABA está controlada por la sequía y la sal, pero no por el frío. Los factores de transcripción NAC y HD-ZIP están implicados en la expresión génica de ERD1.

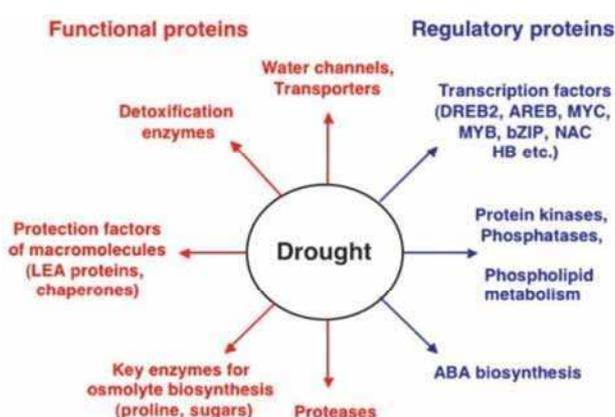
Estudios recientes reportaron que el promotor del gen inducible por sequía, alta salinidad y frío, *RD29A/COR78/LT178* contiene dos elementos mayores que

actuán en cis, ABRE y DRE/CRT y funcionan en la expresión dependiente e independiente de ABA.

### Estrés hídrico en plantas

De los numerosos factores abióticos naturales causantes de estreses en las plantas, el estrés hídrico constituye uno de los factores más limitantes de la producción mundial de cultivos (Bray, 1997, 2002; Yordanov *et al*, 2003, Boyer, 1982). La sequía o estrés por déficit hídrico, causa deshidratación celular por remoción de agua hacia los espacios extracelulares y ello induce respuestas fisiológicas y bioquímicas que incluyen el cierre de estomas, inhibición del crecimiento, disminución de la tasa fotosintética, acumulación de ácido absícico y de osmolitos, activación de la respiración y síntesis de nuevas proteínas (Shinozaki y Yamaguchi-Shinozaki, 2007; Yordanov *et al*, 2003). Además de estas respuestas, las plantas sufren cambios morfológicos. Las plantas responden también a nivel celular y molecular, modificándose la expresión génica de genes de enzimas claves en la síntesis de osmolitos, enzimas antioxidantes, proteínas con acción protectoras, factores de transcripción entre otras (Shinozaki y Yamaguchi-Shinozaki, 2007).

Tras el análisis transcriptómico de genes inducidos por sequía en *Arabidopsis* basado en micromatrices, clasificaron en dos grupos los genes que se inducen en la respuesta al estrés (Figura 4) (Shinozaki *et al*, 2003).



**Fig. 4. Funciones de los genes inducidos por sequía en la tolerancia y la respuesta al estrés.** (Shinozaki *et al* 2003). En un primer grupo se incluyen las proteínas que probablemente actúen en la tolerancia al estrés (proteínas funcionales, en rojo), y el segundo grupo contiene los factores de la proteína implicados en la regulación ulterior de la transducción

de señales y la expresión génica que probablemente actúen en respuesta al estrés (proteínas reguladoras, en azul).

El primer grupo incluye la mayoría de las proteínas que probablemente actúen en la tolerancia al estrés. Estos incluyen moléculas como chaperonas, proteínas LEA, osmotinas, proteínas anticongelantes, proteínas de unión, las enzimas clave para la biosíntesis de osmolitos, proteínas de canales de agua, transportadores de azúcar y prolina, las enzimas de detoxificación, y varias proteasas. El segundo grupo está compuesto por proteínas reguladoras, es decir, los factores de proteína implicados en la regulación de transducción de señales y en la expresión de genes de respuesta al estrés. Estos incluyen diversos factores de transcripción, Proteín Kinasas, Fosfatasas, enzimas que participan en el metabolismo de los fosfolípidos y otras moléculas de señalización como la proteína de unión a calmodulina. Un estudio y clasificación equivalente realizaron los mismos autores analizando el transcriptoma del arroz en respuesta a estrés hídrico (Shinozaki y Yamaguchi-Shinozaki, 2007; Moreno, 2009, Valliyodan y Nguyen, 2006).

### **Respuestas bioquímicas y fisiológicas de las plantas a la sequía**

Como respuestas asociadas a la sequía, se incluyen la pérdida de turgencia, cambios en la permeabilidad y fluidez de la membrana, cambios en la concentración de solutos e interacciones proteína-proteína y proteína-lípidos.

La limitación de agua, promueve un cambio de la relación entre el crecimiento de la parte aérea y el desarrollo de raíces, que lleva a una limitación de la expansión foliar al tiempo que las raíces continúan su desarrollo en busca de agua en zonas más profundas (Moreno, 2009).

El control estomático que regula las pérdidas de agua es un evento temprano en la respuesta al déficit hídrico, que lleva a una limitación en la absorción de carbono por las hojas (Chaves *et al*, 2002). Es un proceso gobernado por ABA que se inicia cuando el mesófilo comienza a sufrir deshidratación. En tales condiciones, el contenido de ABA en las hojas se incrementa promoviendo una pérdida de iones  $K^+$  y aniones  $Cl^-$  o malato en las células guarda, que acarrea una salida de agua hacia el citoplasma y en consecuencia se cierran los estomas.

Este cierre de los estomas no sólo tiene como consecuencia la disminución de la pérdida de agua por la hojas, sino también la reducción en la entrada del CO<sub>2</sub>, lo cual repercute directamente en el proceso fotosintético y, por tanto, en la formación eficiente de fuentes carbonadas necesarias para la nutrición vegetal. La reducción de la fotosíntesis se produce por el cierre en sí de los estomas así como por una reducción de la actividad de las enzimas fotosintéticas. Por su parte, la asimilación de carbono a nivel de planta entera disminuye como consecuencia de la disminución de la difusión de CO<sub>2</sub> en la hoja, la desviación de la asignación de carbono a los órganos no fotosintéticos y las moléculas de defensa o por cambios a nivel bioquímico en la hoja (Chaves *et al*, 2002).

A nivel celular se produce un ajuste osmótico que busca mantener el potencial osmótico durante el período de estrés utilizando azúcares u otros solutos compatibles (Ingram y Bartels, 1996) y por la acumulación de iones, particularmente K<sup>+</sup> (Moreno, 2009). La acumulación de iones se da en la vacuola mientras que la acumulación de solutos se observa en el cloroplasto. Las moléculas que actúan como agentes de equilibrio osmótico se acumulan en respuesta al estrés en la planta y se degradan una vez que éste concluye. También se producen cambios en el metabolismo de carbohidratos de las plantas que están expuestas a estrés hídrico, en donde los compuestos polihidroxilados pueden establecer interacciones hidrofóbicas con los fosfolípidos de la membrana promoviendo la estabilidad de ésta.

### **Señalización celular y regulación génica**

Los estudios transcriptómicos han permitido identificar genes involucrados en la respuesta al estrés hídrico (Shinozaki y Yamaguchi-Shinozaki, 2007; Bray *et al*, 2000; Seki *et al*, 2001) que sin dudas juegan un rol relevante en la adaptación de las plantas al estrés.

Por ejemplo, se sugiere que las proteínas LEA (*Late Embryogenesis Abundant Proteins*), cumplen una acción protectora sobre las estructuras celulares específicas así como también contribuyen a minimizar los efectos del estrés (Ingram y Bartels, 1996). Dentro de las proteínas LEA, hay varias familias, de acuerdo con la homología en la secuencia de aminoácidos y fueron descritas

en las vías de síntesis de osmolitos, así como antioxidantes. La acción de estas proteínas está mediada por la fitohormona ABA. Son proteínas altamente hidrofílicas (Ingram y Bartels, 1996).

El agua mantiene la estructura de proteínas y fosfolípidos. La deshidratación lleva a la cristalización de los componentes celulares, que daña las estructuras celulares. Se postula que las proteínas LEA, actuarían sustituyendo el agua y manteniendo la estructura de las proteínas y membranas en ausencia de ésta.

Algunas de las proteínas LEA, podrían ser consideradas solutos compatibles, que apoyen el posible papel de los azúcares en el mantenimiento de la estructura del citoplasma en ausencia de agua (Ingram y Bartels, 1996).

Los osmolitos son compuestos claves en el ajuste osmótico, que a su vez protegen a las células de las especies reactivas de oxígeno. Se acumulan en las plantas en respuesta al estrés hídrico (Valliyodan y Nguyen, 2006) por sobreexpresión de las enzimas clave de su biosíntesis. Son algunos de ellos: prolina, poliaminas, glicin-betaína, sacarosa, azúcares alcohólicos, y oligosacáridos (Moreno, 2009). Su acumulación, usualmente citoplasmática, promueve el descenso del potencial osmótico de la célula y mantiene la turgencia. La acumulación de prolina, por ejemplo, confiere un ajuste osmótico a la vez que promueve el aumento de los niveles de otros osmolitos. Asimismo actuaría desintoxicando especies reactivas de oxígeno. Hay ejemplos de plantas transgénicas de petunia y soja, que sobreexpresan este osmolito, confiriéndoles cierto grado de tolerancia al estrés (Valliyodan y Nguyen, 2006). Por su parte los azúcares, en condiciones de estrés, actuarían como solutos compatibles que contribuyan a un ajuste osmótico. Se ha visto que la sacarosa se acumula en las células en respuesta a la sequía (Ingram y Bartels, 1996).

Se han propuesto al menos 4 sistemas regulatorios independientes para la expresión de genes en respuesta al estrés hídrico: dos son ABA-dependientes (una dependiente de la síntesis de proteínas y otra independiente de la síntesis de proteínas) y otros dos son ABA-independientes (Shinozaki y Yamaguchi-Shinozaki, 2000).

Muchos de los cambios en los niveles de ARNm que se observan en respuesta al estrés hídrico, reflejan una activación transcripcional (Ingram y Bartels, 1996).

Los factores de transcripción (FTs) juegan un rol central en la elaboración de la respuesta y el programa morfogénico de la planta. Se trata de proteínas que actúan en *trans*, capaces de reconocer blancos de secuencias específicas de ADN (elementos *cis*) localizadas en las regiones promotoras de determinados genes. En plantas, se han identificado y caracterizado numerosos genes que codifican FTs. Los FTs vegetales se clasifican en familias y subfamilias de acuerdo al grado de conservación de la secuencia de aminoácidos, al tamaño, y a la composición estructural de los genes codificantes. Algunos FTs fueron caracterizados funcionalmente y, en casi todos los casos, se observó que intervienen en varias vías de señalización. Así por ejemplo, las proteínas de las familias MYB, MYC, b-Zip y HD-Zip participan en las respuestas a distintos tipos de estrés abiótico; mientras que las proteínas de la familia WRKY se relacionan tanto con la respuesta al ataque por organismos patógenos tanto como a la respuesta frente a estrés abiótico.

Dentro de los FTs específicamente vinculados a estreses ambientales, existe un grupo de genes bien caracterizados cuya acción se ubica en el inicio de la cascada de señalización en respuesta a estrés abiótico. Este grupo se encuentra representado por Rd29A de *A. thaliana*. En los promotores de los genes de este grupo se identificaron las secuencias denominadas ABRE (del inglés *ABA-responsive element*, elemento de respuesta a ABA) y DRE/CRT (del inglés *dehydration-responsive element/C-repeat*, elemento de respuesta a deshidratación) (Yamaguchi-Shinozaki y Shinozaki, 1994).

Plantas transgénicas de *A. thaliana* sobreexpresando DREB1 y DREB2 resultaron tolerantes a déficit hídrico. Por otra parte, la expresión heteróloga de DREB1B/CBF1 de *A. thaliana* en tomate, además de conferir una elevada tolerancia a sequía, también lo hizo frente a estrés oxidativo y al generado por frío. Apparently el mecanismo de tolerancia ocurre vía el cierre estomático rápido, el incremento en la concentración de prolina y el aumento de la actividad catalasa, que provoca la reducción en la acumulación de H<sub>2</sub>O<sub>2</sub>.

## **Señalización por mensajeros secundarios**

Varios genes inducibles por estrés cuyos productos intervienen en la transmisión de señales de la respuesta han sido identificados y caracterizados exhaustivamente, incluyendo Fosfolipasas y Kinasas de proteínas pertenecientes a las familias MAP (del inglés *mitogen activated protein*), y SOS (del inglés *salt-overlyensitive*) (Ingram y Bartels, 1996, Zhu, 2002).

La activación de una cascada de MAPK representa una de las primeras respuestas celulares frente a señales ambientales. Una cascada de MAPK típica consta de tres proteínas Kinasas MAPK-Kinasa Kinasa (MAPKKK), MAPK Kinasa (MAPKK), y MAPK que fosforilan secuencialmente los sustratos intermedios correspondientes. Una MAPK fosforilada se activa y fosforila varias proteínas como factores de transcripción, Proteín Kinasas, enzimas metabólicas, y las proteínas del citoesqueleto (Yoshioka y Shinozaki, 2009).

En la cascada de señalización de MAP Kinasas, la señal de entrada podría ser el estrés osmótico y la de salida, la acumulación de osmolitos que contribuyan a establecer la homeostasis osmótica, protejan del daño del estrés, o contribuyan a reparar mecanismos celulares (Zhu, 2002).

Puesto que operan en un nivel alto de la cascada de decisiones de la respuesta, constituyen por definición excelentes puntos de intervención para obtener tolerancia generalizada. Varios de estos genes han sido introducidos en plantas resultando en líneas con fenotipos tolerantes a diversas fuentes de estrés. El estrés osmótico induce la señalización del calcio en la que las proteínas Kinasas dependientes de calcio son los principales candidatos que enlazan la señal de calcio a la respuesta río abajo (Zhu, 2002).

### **Señalización de fosfolípidos**

Los fosfolípidos de membrana constituyen un sistema dinámico que genera una gran cantidad de moléculas señal (IP<sub>3</sub>, DAG, PA, etc) que cumplen funciones estructurales en respuesta al estrés (Zhu, 2002). Los sistemas de señalización de fosfolípidos se agrupan según las Fosfolipasas que catalizan la formación de lípidos y otros mensajeros que no son el producto directo de las fosfolipasas, como diacilglicerol pirofosfato y fosfatidilinositol 3,5 bifosfato. Se ha observado en plantas sometidas a estrés hiperosmótico, que el nivel de IP<sub>3</sub> se incrementaba rápidamente, de igual manera que con el agregado de ABA exógeno. En *A. thaliana*, el gen de PLC, también es inducido por sequía. Se ha

visto que en las células guarda, el IP3 induce un aumento de  $\text{Ca}^{+2}$  en el citoplasma y provoca cierre de los estomas (Zhu, 2002).

### **Estrés hídrico en papa**

La papa es uno de los cultivos más sensibles al déficit hídrico en el suelo, aunque el uso del agua es mucho más eficiente que el de los cereales, con índices de cosecha cercanos al 0,75% en relación al 0,5% de los cereales. Las variedades modernas de papa necesitan un riego superficial frecuente.

El estrés hídrico es reconocido mundialmente como la mayor limitante a nivel mundial de la producción de papa (Schapendonk *et al*, 1989). Van Loon (1981), menciona que las altas temperaturas y el estrés hídrico son los principales condicionantes de la calidad y el rendimiento de las papas.

El agua es el principal componente de las papas, 90-95% en los tejidos verdes y 75-85% en los tubérculos. Un cultivo de papa en activo crecimiento demanda altas dosis de agua y bajo condiciones óptimas de crecimiento, las plantas de papa transpiran con una tasa promedio, que puede reemplazar su contenido hídrico total, cerca de 4 veces al día.

Cuando la tasa de transpiración es alta y la humedad del suelo es baja, el contenido de agua de la planta decrece rápidamente. Si el contenido de agua decrece un 10%, el crecimiento se vuelve más lento y puede, eventualmente cesar. Si decrece un 20%, la planta se marchita y los principales procesos fisiológicos se vuelven más lentos o se detienen y si llega a decrecer un 40%, la planta usualmente muere (Coleman, 2008).

La mayoría de los déficits de agua en la planta se producen cuando la acumulación de agua en el suelo desciende por debajo de un nivel crítico. Para el cultivo de la papa, el estatus de humedad en el suelo se vuelve crítico cuando el porcentaje de agua cae por debajo de 60-65%. Esta sensibilidad probablemente se deba a su sistema radicular y sus respuestas fisiológicas frente a déficits hídricos moderados.

La primera de las respuestas observada ante el estrés hídrico es el cierre de los estomas de las hojas, un estrés prolongado lleva al marchitamiento, debido a la pérdida de turgencia de las hojas.

El déficit hídrico causa el cierre de estomas, la reducción de la absorción de  $\text{CO}_2$  para la fotosíntesis, la reducción de crecimiento de las plantas y el

rendimiento, ya que disminuye el desarrollo de raíces y tallos y la tasa de incremento en el área foliar motivando un menor desarrollo del canopeo (Steyn *et al*, 1992). Una disminución de la expansión foliar limita la capacidad fotosintética total, mientras que una disminución del desarrollo radicular limita la capacidad de la planta de capturar agua y nutrientes del suelo. La reducción de la producción de fotosintatos restringe el transporte desde las hojas a los tubérculos. El rendimiento del cultivo de papa depende de la acumulación de esos productos fotosintéticos en los tubérculos.

El rendimiento del cultivo de papa está medido por el número y el tamaño de los tubérculos. La acumulación de productos fotosintéticos durante el período de desarrollo de estolones es necesaria para promover la tuberización (Vos y Haverkort, 2007).

Un déficit hídrico en este estado de crecimiento limita marcadamente el rendimiento a través de una reducción en el número y tamaño de los tubérculos que se forman. Durante la tuberización (o *bulking*) (estado de crecimiento IV), los productos fotosintéticos en exceso son traslocados para el rápido desarrollo de los tubérculos. Este estadio es considerado el período crítico del cultivo, por los efectos en la reducción del crecimiento potencial de la planta.

### **Efecto del contenido del agua en la calidad de la papa**

Hay dos indicadores que reflejan el grado de estrés en las plantas de papa: las concentraciones de  $\text{Ca}^{2+}$  y  $^{13}\text{C}$ .

Las concentraciones de calcio disminuyen a medida que aumenta el estrés hídrico, ya que la planta tiene menos agua por unidad de materia seca producida. Las plantas de papas tienen la capacidad de discriminar el  $^{13}\text{CO}_2$  durante su difusión desde el aire al espacio intercelular. Cuando hay estrés, esta capacidad disminuye marcadamente y el  $^{13}\text{C}$  se acumula en los espacios intercelulares (Vos y Groenwold, 1989).

El déficit de agua también provoca desórdenes en el desarrollo de los tubérculos que afectan su comercialización al disminuir su calidad. Las típicas deformaciones en los tubérculos sometidos a estrés hídrico (o *growth cracks*) se originan por fluctuaciones en el contenido de agua por interrupción y rápida reanudación del crecimiento, que provoca un rápido incremento en el volumen de los tejidos internos, originando malformaciones (Struik *et al*, 1989).

### **Interacción del estrés hídrico con otros estreses abióticos y bióticos**

Las plagas y las enfermedades pueden afectar el rendimiento y la producción de cultivo por su impacto en la economía hídrica de la planta.

Estudios de la relación entre el estrés hídrico y los efectos de *Verticillium dahliae* (Haverkork *et al*, 1990, Bowden y Rouse, 1991) demuestran que la inoculación de plántulas promovió el cierre de estomas y una reducción en la tasa de transpiración y fotosíntesis, síntomas similares a los del estrés hídrico.

Otro ejemplo lo constituye el nematode del quiste de la papa, que afecta la capacidad de captar agua, ya que infesta raíces. Esto promueve una disminución del potencial hídrico de la hoja y un incremento de la resistencia estomática (Evans *et al*, 1975).

### **Control del estrés hídrico en papa**

La forma más eficiente y económica del control del déficit hídrico es mediante la utilización de variedades resistentes y mediante la implementación de prácticas culturales de manejo como la implementación de un adecuado sistema de irrigación.

Los métodos de mejora tradicionales se ven dificultados debido a la existencia de patrones de segregación tetrasómicos, la fertilidad incompleta de muchos cultivares comerciales de papa y la importancia de los efectos heteróticos en caracteres importantes, como son el rendimiento, el peso seco, distintos parámetros de calidad industrial, etc. (Ross, 1986). A esto se suma el carácter esencialmente conservador de muchos productores e industrias, que suelen preferir los cultivares tradicionales ya establecidos, aunque las nuevas variedades puedan presentar ventajas comparativas. Esta situación se observa, por ejemplo, con el cultivar Russet Burbank en Estados Unidos, con Bintje en Europa y con Spunta en Argentina.

En muchos países en desarrollo se suma a las dificultades mencionadas una falta de conocimiento, por parte de los agricultores que permita un adecuado manejo cultural del cultivo bajo estos estreses.

### **Herramientas de Biología Molecular y Bioinformática**

## **Análisis de la expresión genética**

La búsqueda de resistencia genética por medio del mejoramiento tradicional es lenta y laboriosa en el caso de la papa. El análisis de la expresión génica supone una estrategia que posibilita la identificación de los genes responsables de la mejora agronómica del cultivo dado que refleja los cambios en la transcripción de los genes ante una situación específica del tejido, del tratamiento y del estado de desarrollo. Los cambios en la expresión de genes están inducidos por una serie compleja de eventos de transducción de señales, muchos de las cuales no están esclarecidas (Zhu *et al*, 2002). La revolución existente en el campo de las herramientas y técnicas aplicables al estudio de la expresión génica posibilita la implementación de diferentes estrategias para el análisis transcriptómico génico diferencial de la respuesta de *Solanum* a estreses bióticos y abióticos.

En respuesta al estrés, en la planta se producen alteraciones en los niveles de expresión de un conjunto determinado de genes. La expresión génica diferencial generalmente se analiza comparando diferentes condiciones experimentales (normal vs. estresado). Aquellos genes que no presentan alteraciones durante el tratamiento, por codificar proteínas cuya función típica es parte del metabolismo celular básico, se denominan genes de mantenimiento (del inglés *housekeeping*) (Coker y Davies, 2003).

A través del análisis de la expresión génica en diferentes momentos de la evolución del estrés, es posible comprender mejor la dinámica de expresión de un gran número de genes en respuesta al estrés, así como establecer relaciones funcionales entre los genes involucrados.

## **Avances Tecnológicos**

Los avances en biología molecular y las nuevas técnicas computacionales nos permiten investigar de manera sistemática los complejos procesos moleculares que subyacen a los sistemas biológicos (Durbin *et al.*, 1998). Una amplia gama de tecnologías de la expresión de genes han sido utilizadas por investigadores de la papa (Bryan y Hein, 2008). El análisis de expresión de alto rendimiento es una disciplina que está en transición y es probable que logre un desarrollo significativo en el futuro, en particular con la evolución reciente de "próxima generación" de tecnologías de secuenciación (NGS), que tienen el potencial de

ampliar radicalmente las capacidades para descubrir genes novedosos y encontrar sus funciones.

### **Tecnología de micromatrices**

Entre las herramientas más poderosas y versátiles de la genómica se encuentran las matrices de alta densidad de oligonucleótidos o ADN complementario. También conocida como “*microarrays*”, revolucionaron la investigación biológica moderna por su capacidad de controlar el nivel de expresión de miles de genes simultáneamente (Brown y Botstein, 1999; Scott *et al*, 2009), mientras que los métodos tradicionales sólo podían manejar la de un gen en un criterio temporal.

Las micromatrices explotan la capacidad de una sonda derivada de una molécula de ARNm de unirse específicamente, o a hibridarse con un templado de ADN complementario. La tecnología de micromatrices consigue inmovilizar una alta densidad de material biológico en un soporte por lo que pueden determinarse, en un solo experimento, los niveles de expresión de cientos o miles de genes. Con la ayuda de herramientas informáticas, estas uniones son cuantificadas con precisión, generando un perfil de expresión génica (Scott *et al*, 2009).

Los primeros pasos en los estudios de expresión génica con micromatrices se obtuvieron hacia finales de la década de los ochenta de manos de los científicos Fodor, Pirrung, Read y Stryer, que desarrollaron una revolucionaria técnica para la determinación y cuantificación de ADN en una muestra. Posteriormente se convertiría en la primera plataforma de micromatrices denominada GeneChip (Affymetrix).

Esta tecnología de “*genechips*” abrió un nuevo camino para el estudio de los organismo con un enfoque genómico funcional más amplio. Esta técnica permite comprender mejor la dinámica de expresión de un gran número de genes en respuesta al estrés impartido, así como establecer relaciones funcionales entre los genes involucrados. También ha sido reportada para la detección de mutaciones y polimorfismos, secuenciación, seguimiento de terapia, medicina preventiva, *prospección* y toxicología de fármacos y diagnóstico molecular.

Técnicamente una micromatriz de ADN consiste en un gran número de moléculas de ADN ordenadas sobre un sustrato sólido de manera que formen una matriz de secuencias de dos dimensiones. Los fragmentos de material genético que se depositan sobre la superficie del arreglo pueden ser fragmentos o genes completos procedentes generalmente de bibliotecas génicas. Las sondas que conforman la micromatriz pueden ser secuencias cortas llamadas oligonucleótidos, o de mayor tamaño, cDNA, o bien productos de PCR, replicados *in vitro* mediante una reacción en cadena de la polimerasa. El tamaño de las sondas utilizadas puede ser de pocas decenas de pares de bases hasta kb. Los productos de PCR, consisten en fracciones de genes generados por PCR procedentes de clones de cDNA o librerías genómicas. La longitud de estos fragmentos oscila entre 200 a 500 pb. Los oligonucleótidos tienen una longitud más limitada de 20-25 pares de bases hasta 100 bp.

Los ácidos nucleicos de las muestras a analizar se deben marcar por diversos métodos (enzimáticos, fluorescentes, etc) y se incuban sobre la micromatriz, permitiendo la hibridación (reconocimiento y unión de moléculas complementarias) de secuencias homólogas.

Durante la hibridación, las dos muestras de material genético etiquetados con dos diferentes fluoróforos (por ejemplo Cy3 – verde y Cy5 – rojo) se mezclan y se hibridan con una sola micromatriz que luego se escanea en un escáner específico para visualizar la fluorescencia de los fluoróforos. La intensidad relativa de cada fluoróforo se utiliza para identificar los genes reprimidos o inducidos bajo una condición dada.

Las micromatrices comerciales disponibles para papa son de ADNc (micromatrices 10K, TIGR) y de oligonucleótidos (POCI, Agilent) (Kloosterman *et al* 2008).

El análisis de micromatrices genera una acumulación enorme de conocimiento en repositorios de datos digitales. Esta gran dimensionalidad de datos promueve desafíos en el procesamiento estadístico y en el análisis de los datos con la finalidad de extraer información útil de los mismos, facilitando la identificación de aquellos genes candidatos que estén involucrados en la respuesta a estreses bióticos y abióticos de los miles de genes analizados en cada micromatriz.

### **Micromatrices *Potato 10K***

Al momento de realizar esta tesis, el genoma de papa no estaba disponible y sólo existía información sobre ESTs. Un conjunto no redundante de 10.000 ESTs se utilizó por el Instituto de Investigación Genómica (TIGR) para el desarrollo de una micromatriz de ADNc de papa. Esa misma organización ofrecía un servicio de transcripción que permitía la evaluación de estas matrices por una amplia gama de usuarios que trabajan en diferentes especies de solanáceas y bajo diferentes condiciones biológicas. Esto ha permitido la generación de datos masivos surgidos de las micromatrices 10K de papa, que están disponibles públicamente. Actualmente no se están comercializando estos chips.

Sin embargo, esta plataforma tiene el inconveniente de contener sólo una proporción del repertorio de genes de la papa (1/4 parte del total del genoma, aproximadamente).

### **Biología computacional y Bioinformática**

La Bioinformática surgió en los últimos años como disciplina auxiliar de la genómica y otras x-ómicas, que necesitan del procesamiento de enormes cantidades de datos generados en las últimas décadas merced a los avances en la biología molecular y el equipamiento disponible para la investigación en este campo, tales como la secuenciación sistemática, la proteómica, las matrices de expresión, la prospección de genes, la metabolómica, etc.

Es aquí donde la Informática se integra con la Biología para manejar, estudiar y procesar los volúmenes de datos en crecimiento exponencial que se utilizan en la investigación, en un nuevo campo denominado Bioinformática que permita cubrir la necesidad de un alto nivel de organización (Attwood y Parry-Smith, 2002).

Una de las principales tareas de esta disciplina es la creación y mantenimiento de bases de datos de información biológica. Las secuencias nucleotídicas (y las secuencias proteicas que derivan de las mismas) son los componentes mayoritarios de la información almacenada en estos repositorios. Pero la Bioinformática tiene muchísimo potencial que va más allá de ayudar a resolver con sus herramientas, preguntas fundamentales de la biología a través de la secuenciación, el ensamblado de genes y genomas, el plegado de las

proteínas entre otros y que incluye aspectos de gran complejidad biológica como la simulación de sistemas celulares y redes de interacción molecular.

### **Conceptos de Biología computacional y Bioinformática**

Es importante distinguir tres acepciones en las que se unen la biología y la informática, cada una con objetivos y metodologías bien diferenciadas:

1. **Bioinformática o Biología Molecular Computacional:** investigación y desarrollo de la infraestructura y sistemas de información y comunicaciones que requiere la biología molecular y la genética (redes y bases de datos para el genoma, microarreglos, etc.). Por ejemplo, informática aplicada a la biología molecular y la genética.

2. **Biología Computacional:** ciencia de la computación que se aplica al entendimiento de cuestiones biológicas básicas, no necesariamente en el nivel molecular, mediante la modelización y simulación (ecosistemas, modelos fisiológicos). Por ejemplo, informática y matemáticas aplicadas a la biología.

3. **Biocomputación:** desarrollo y utilización de sistemas computacionales basados en modelos y materiales biológicos (biochips, biosensores, computación basada en ADN, redes de neuronas, algoritmos genéticos). Por ejemplo, biología aplicada a la computación.

El desafío de la Bioinformática es extraer los datos de interés de los experimentos, organizarlos y analizarlos a partir de la interacción con la información contenida en los repositorios.

La fase *in silico* de análisis e interpretación de las señales de hibridación obtenidas, es un proceso denominado *Data Mining* o minería de datos. La tecnología informática ha proporcionado una gran solución a este tipo de problemas.

Las computadoras, ya no solo acumulan y organizan la información de secuencias en bases de datos, sino que también pueden analizar los datos de las secuencias muy rápidamente. La evolución del poder computacional y la capacidad de almacenamiento son muy importantes para la investigación, ya que posibilita hacer frente a la creciente cantidad de información de secuencias que está siendo obtenida. Se han desarrollado sofisticados algoritmos que permiten comparar secuencias basados en la teoría de probabilidades. Estas

comparaciones se han convertido en la base de la determinación de la función de genes, desarrollando relaciones filogenéticas y simulando modelos de proteínas (Attwood y Parry-Smith, 2002).

La Bioinformática también se encuentra con algunos problemas por superar, como por ejemplo el de la regulación de genes. Ya que dado un conjunto de genes de un genoma, el desafío está en obtener la red de interacciones entre este conjunto, determinando cómo, cuándo y qué genes se activan o reprimen. Las redes de regulación de genes son básicamente interruptores de encendido - apagado de una célula que funcionan a nivel de genes. Estos coordinan dinámicamente el nivel de la expresión para cada gen en el genoma, controlando cuándo y cómo ese gen será transcrito en ARN.

La comprensión de los mecanismos del funcionamiento celular requiere un estudio exhaustivo del complejo comportamiento de un conjunto de genes en interacción. A pesar de la simplificación que realizan las redes, las redes de regulación genética son procesos complejos, en los cuales forman parte muchos componentes que afectan directa o indirectamente la regulación de la expresión.

## **Metodología para la identificación de perfiles de expresión diferencial**

### **Conceptos GO.**

El proyecto GO (Gene Ontology) (Consortio, 2000) congrega una de las caracterizaciones más potente de genes. Utiliza un vocabulario dinámico estructurado en ontologías para describir los productos génicos en términos de sus procesos biológicos asociados, los componentes celulares y funciones moleculares de forma independiente de las especies (Consortio, 2000). Los términos GO están organizados como redes jerárquicas, donde cada nivel corresponde a una definición (o término) de distinta especificidad y donde a mayor nivel de términos, su información es más general que la del término de nivel inferior. El empleo de GO, permite automatizar la búsqueda de las funciones génicas en las bases de datos (QuickGo).

### **3. HIPÓTESIS Y OBJETIVOS**

### 3- HIPÓTESIS Y OBJETIVOS

#### Hipótesis de Trabajo

Las especies silvestres de papa procedentes de regiones geográficas con distintas condiciones de estrés hídrico presentarán respuestas adaptativas particulares y diferenciadas por inducción de la expresión de genes factibles de ser detectados mediante evaluación de perfiles transcriptómicos basados en el análisis de hibridación de ARNs mensajeros con micromatrices conteniendo una muestra de 10.000 genes de papa.

#### Objetivos

Identificar vías metabólicas y procesos fisiológicos diferencialmente inducidos y asociados con la resistencia del cultivo de papa al estrés hídrico mediante el análisis transcriptómico de la expresión génica diferencial.

Este estudio será encarado desde dos puntos de vista, el biológico y el bioinformático, a fin de lograr una mayor comprensión de la respuesta molecular de la papa frente al estrés hídrico, posibilitando la identificación de genes que proporcionen una mayor *performance* agronómica del cultivo.

Para la concreción de esa meta se plantearon una serie de objetivos particulares:

- Clasificar 14 especies silvestres de *Solanum* pertenecientes al Banco de Germoplasma de Neiker-Tecnalia por su comportamiento frente al estrés hídrico.
- Obtener ARNm en cantidad y calidad de seis momentos de tratamiento de sequía para el análisis de la expresión génica diferencial.

- Aplicar micromatrices como estrategia principal para monitorear los cambios en los perfiles transcripcionales de miles de genes al mismo tiempo de *Solanum* en respuesta al tratamiento de sequía.
- Utilizar herramientas bioinformáticas para el análisis de las matrices de expresión generadas por los estudios transcriptómicos
- Establecer patrones de expresión génica, según cada especie que facilite la interpretación de los datos.
- Utilizar herramientas de identificación de genes con expresión génica diferencial que faciliten el análisis de datos de expresión génica y a la vez posibilite establecer relaciones funcionales entre los mismos.
- Extraer de los resultados obtenidos, vías metabólicas y procesos fisiológicos diferencialmente inducidos y en respuesta a los estreses estudiados.
- Validar este conocimiento mediante análisis de RT-PCR (tiempo real) en dos momentos de estrés donde el cambio de expresión haya sido más marcado.

## **4. MATERIALES Y MÉTODOS**

## 4. MATERIALES Y MÉTODOS

### Material biológico

Como punto de partida en la selección de especies de interés para realizar los estudios de estrés hídrico que sustentan este trabajo, se recurrió a la reserva genética de especies de *Solanum* del Banco de Germoplasma de Neiker – Tecnalia, Instituto Vasco de Investigación y Desarrollo Agrario (<http://www.neiker.net/neiker/germoplasma/default.htm>).

Se evaluaron 14 especies de *Solanum* (Tabla 3) que forman parte de la colección *in vitro* y que provenían de colecciones mundiales de germoplasma de papa: GERMAN-Dutch Potato Collection (BGRC), IPK Genbank Groß Lüsewitz (GLKS), Commonwealth Potato Collection (CPC). La totalidad del material utilizado en este ensayo se evaluó previamente por el Grupo de Investigación de Neiker frente a patógenos de interés agronómico de papa como *Phytophthora infestans*, *Globodera pallida* y varios virus de la papa (Ruiz de Galarreta *et al*, 1998).

**Tabla 3. Especies silvestres de *Solanum* empleados en el estudio de estrés hídrico y su nomenclatura abreviada.**

<b>Especies silvestres de <i>Solanum</i></b>	<b>Abrev. Nombre</b>
<i>S. berthaultii</i>	BER 3607
<i>S. infundibuliforme</i>	IFD 1704
<i>S. polytrichon</i>	PLT 53650
<i>S. sparsipilum</i>	SPL 959
<i>S. cardiophyllum</i>	CPH 18325
<i>S. pascoense</i>	PSC 2877
<i>S. marinasense</i>	MRN 2278
<i>S. gandarillasii</i>	GND 2423
<i>S. leptophyes</i>	LPH 27215
<i>S. sanctae-rosae</i>	SCT 18315
<i>S. bulbocastanum</i>	BLB 8008
<i>S. venturii</i>	VNT 8239
<i>S. doddsii</i>	DDS 2880
<i>S. papita</i>	PTA 15442

Las plantas fueron cultivadas en condiciones de invernadero bajo un fotoperíodo de 18 horas de luz y 6 horas de oscuridad y una humedad relativa de 65-70%, utilizando un sustrato especial para semillero detallado en la Tabla 4 (Terra Nature, Bas Van Buuren, Holanda) de empleo rutinario en la producción de papa bajo cubierta en Neiker-Teknalia. Se utilizaron bandejas rígidas de polietileno de uso forestal, que constaban de 6 celdas cada una (dimensiones: 170 x 295 x 54mm).

**Tabla 4. Sustrato profesional *Terra Nature* utilizado en los ensayos de estrés hídrico.**

Composición / propiedades del sustrato	Detalle
Turba Rubia Báltico 0 – 20	80%
Turba Negra Holanda 0 – 10	20%
Homogeneidad	100% 0 – 10mm
Fertilizante PG 12 – 14 – 24	1,2Kg/m <sup>3</sup>
Cal. (CaCO <sub>3</sub> )	6Kg/m <sup>3</sup>
Ph	5,4 – 5,6.
Conductividad	0,8 – 1,0 ms.
NH <sub>4</sub>	1,85mmol/L
K	1,9mmol/L
Na	0,9mmol/L
Mg	0,8mmol/L
Ca	1,3mmol/L
NO <sub>3</sub>	3,5mmol/L
Cl	1,6mmol/L
SO <sub>4</sub>	1,1mmol/L
HCO <sub>3</sub>	< 0,2mmol/L
P (H <sub>2</sub> PO <sub>4</sub> )	1,43mmol/L
Si	0,06mmol/L
Fe	9,3ymol/L
Mn	2,6ymol/L
Zn	0,5ymol/L
B	9,0ymol/L
Cu	0,3ymol/L

### Ensayos de estrés hídrico

El tratamiento de estrés hídrico se inició cuando el material biológico proveniente de cultivo *in vitro* tenía entre 3 y 5 semanas crecido en condiciones de invernadero.

Para cada especie se formaron dos grupos de 20-25 plantas. Un grupo de plantas fue sometido a estrés hídrico mediante la supresión del riego durante

10 días y el otro grupo fue regado normalmente durante ese período. Todas las plantas fueron irrigadas hasta capacidad de campo 72 hs antes de comenzar el tratamiento. El estado del agua de las plantas fue supervisado mediante el empleo de una cámara de presión digital tipo “Scholander” y el método de contenido relativo de agua RWC (del inglés *Relative Water Content*) obtenido según:

$$\text{RWC / CRA} = \frac{(\text{Peso Fresco} - \text{Peso Seco})}{(\text{Peso Túrgido} - \text{Peso Seco})} \times 100$$



**Figura 5. Cámara de presión Scholander** utilizada para medir el potencial de agua (Mpa) de plantas de papa silvestres evaluadas bajo estrés hídrico y sus controles irrigados. A la izquierda se muestra cómo se monitorea la salida de savia del tallo para detectar el estado hídrico de la planta.

La unidad de muestreo consistió en 3-4 plantas de cada grupo (grupo estresado y control respectivamente) que fueron colectadas en seis momentos puntuales del tratamiento (días 0, 1, 3, 5, 7, y 10 posteriores al inicio del tratamiento de sequía) siguiendo el protocolo propuesto por TIGR ([ftp://ftp.tigr.org/pub/data/s\\_tuberosum/SGED//055\\_TIGR\\_drought1/055\\_Study\\_Description.pdf](ftp://ftp.tigr.org/pub/data/s_tuberosum/SGED//055_TIGR_drought1/055_Study_Description.pdf)). Las hojas cortadas en cada tiempo de muestreo fueron agrupadas, congeladas inmediatamente en nitrógeno líquido y almacenadas a -80°C para su posterior procesamiento.

## **Herramientas de Biología molecular**

### **Procesamiento del tejido vegetal.**

El material vegetal conservado a -80°C fue molido con nitrógeno líquido hasta obtener un polvo fino en un mortero pre-enfriado y se reservaron las muestras molidas en tubos estériles de polipropileno a -80°C .

### **Extracción de ARN**

Aproximadamente 1 g de material vegetal en polvo se homogeneizó con 10ml de reactivo TRIzol® (Invitrogen), siguiendo las instrucciones del proveedor. Una vez obtenido el precipitado de ARN, éste se lavó con 1 ml de etanol al 70% y una vez seco se resuspendió en 500 µl de H<sub>2</sub>O-DEPC y 500 µl de tampón de cloruro de litio (LiCl 4 M, TrisHCl pH 7,4 20 mM y EDTA 0,5 M 10 mM). La mezcla se dejó precipitando a -20°C durante una noche. Posteriormente las muestras se centrifugaron durante 30 minutos a la máxima velocidad y se lavaron dos veces con etanol 70% frío. El ARN fue resuspendido en 50µl de H<sub>2</sub>O DEPC.

La integridad del ARN fue verificada en geles de agarosa al 1%, sometiendo 1µl de cada muestra a electroforesis en una solución tampón de corrida ME (400 mM MOPS, 100 mM acetato de Na, 10 mM EDTA pH 8,0, todo solubilizado en agua DEPC). Por su parte, cada muestra fue tratada con una solución desnaturalizante que contenía formamida deionizada (60%), ME (1,2X), formaldehído (8,5%), azul de bromofenol/xilen-cianol (0,01% P/V). Las muestras se incubaron 5 min a 65 °C e inmediatamente fueron transferidas a hielo, previo a la siembra en el gel.

La concentración fue ajustada a 1 µg/ml (mediante medición de absorbancia a 260nm) por precipitación con etanol y resuspensión en H<sub>2</sub>O DEPC.

Del total de muestras procesadas, se utilizó el ARN extraído de las especies CPH 18325, PSC 2877, VNT 8239 para los ensayos de micromatrices.

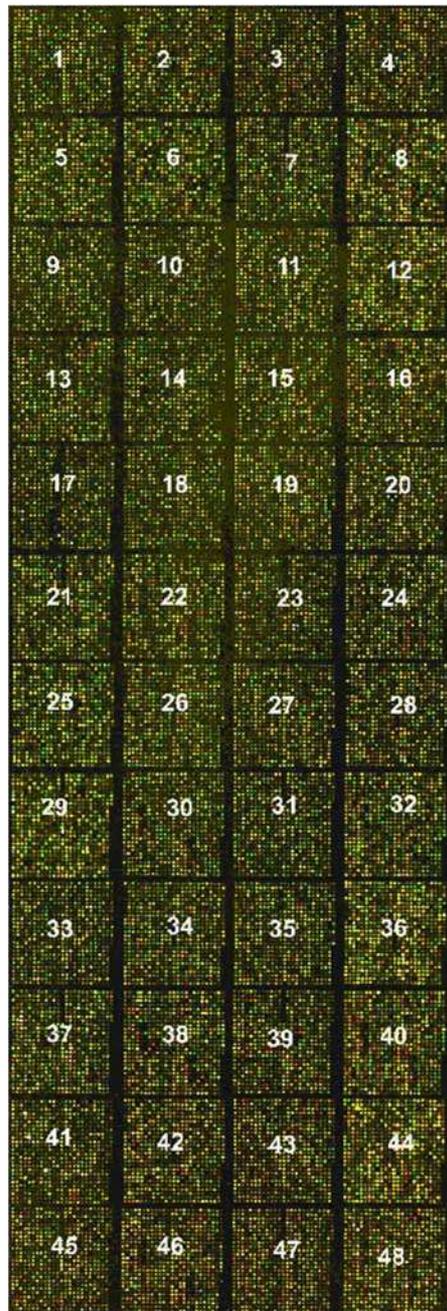
### **Micromatrices de ADNc**

Se utilizaron para este estudio micromatrices de ADNc de papa diseñadas por el Instituto TIGR. Se utilizó la versión 4 de las micromatrices 10K de papa.

Para la construcción de esta micromatriz se utilizaron estolones, raíces, microtubérculos, tubérculos con latencia o dormición, hojas sanas y bibliotecas obtenidas de la interacción de *Phytophthora infestans* (compatible e incompatible) con papa ([http://jcvl.org/potato/sol\\_ma\\_layout.shtml](http://jcvl.org/potato/sol_ma_layout.shtml)).

Están depositados sobre la superficie del chip 15.264 secuencias de ADNc dispuestas por duplicado, entre las que se incluyen clones seleccionados como control (Ver Tabla en Material Complementario). De los 32.448 puntos de hibridación del chip, 22.824 se encuentran validados por medio de resecuenciación y geles de agarosa.

Las secuencias contenidas en el array están organizadas en 48 bloques (Figura 6). Los duplicados no siguen un esquema particular de espaciamiento dentro del arreglo.



BARCODE



**Figura 6. Micromatriz 10K de papa (TIGR).** Imagen de la disposición de los ADNc de la micromatriz de papa 10K en 48 bloques.

### **Análisis de Expresión de las especies *S. cardiophyllum* (CPH) y *S. venturii* (VNT)**

Dos de las especies (CPH y VNT) fueron seleccionados por el Servicio de Perfiles de Expresión de TIGR de papa para ser procesados en su Plataforma.

El ensayo tiene asignado el número 102 dentro de los Proyectos del Servicio de análisis de TIGR. Los detalles del experimento, así como todos los datos se encuentran disponibles públicamente en la Base de Datos de Expresión Génica de Solanáceas (SGED: Solanaceae Gene Expression Database) en el sitio: [ftp://ftp.tigr.org/pub/data/s\\_tuberosum/SGED//102\\_Lucca/](ftp://ftp.tigr.org/pub/data/s_tuberosum/SGED//102_Lucca/) y también disponible *on line* en GEO (Gene Expression Omnibus) accesible en el vínculo [www.ncbi.nlm.nih.gov/gds/](http://www.ncbi.nlm.nih.gov/gds/), que es una base de datos que almacena conjuntos de datos de expresión génica curados, utilizando el término de búsqueda “Lucca”, o eventualmente el concepto de búsqueda deseado, donde se localizan los experimentos de interés.

### **Análisis de Expresión de la especie *S. pascoense* (PSC)**

El ensayo de micromatrices con la especie PSC 2877 fue procesado en la Universidad de Salamanca, utilizando las facilidades con las que cuenta el Instituto de Microbiología Bioquímica (IMB) y el Centro de Investigación del Cáncer. Se siguió el protocolo propuesto por TIGR (que se correspondía con el utilizado con las otras especies en estudio) y que se detalla en los puntos siguientes.

### **Purificación del ARN para su utilización en ensayos de micromatrices**

Previo a su utilización como templado en el proceso de Síntesis de cDNA, el ARN fue purificado utilizando el Kit comercial RNeasy Plant Mini Kit (QIAGEN). Se cuantificó y determinó su calidad mediante el sistema Bioanalyzer 2100 – Bio-sizing A.02.12.SI292 (Agilent), servicio ofrecido por el Centro de Investigaciones del Cáncer (CIC), que permite determinar en forma más confiable y rigurosa que en los sistemas tradicionales. Se basa en un sistema de electroforesis mediante el uso de nanocapilares.

### **Síntesis de cDNA por marcación directa**

Para la síntesis de cDNA con Cianina marcada con Cy3 o Cy5-dUTP (Amersham Biosciences), que son fluorocromos que absorben y emiten fluorescencia a diferentes longitudes de onda, se siguió el método de marcación directa sugerido por TIGR y disponible en el sitio: [http://jcvi.org/potato/images/SGED\\_SOP\\_5.1.1.pdf](http://jcvi.org/potato/images/SGED_SOP_5.1.1.pdf). Fue el mismo protocolo con

el que se analizaron las muestras de los especies CPH y VNT enviadas al Servicio de Expresión de TIGR.

El procedimiento se basa en una reacción de retro-transcripción en la cual se incorporan nucleótidos marcados durante la síntesis de las moléculas de cDNA, finalizando con un proceso de purificación de la sonda sintetizada.

Así es como el proceso se inicia con la formulación de una mezcla de marcación que contenía dCTP, dATP, dGTP, dTTP, ditiotreitól (DTT) y el buffer de la enzima transcriptasa reversa (RT). Esta mezcla se formuló para 15 reacciones según se muestra en la Tabla 5. De la mezcla formulada, se utilizó una alícuota en el siguiente paso de síntesis de cDNA marcado, reservándose el resto a -20°C para marcaciones futuras.

**Tabla 5. Marcación de ADNc para micromatrices.** Mezcla principal para la reacción de marcación de cDNA

<b>Reactivos</b>	<b>Volumen (µl)</b>
5 x Buffer RT (Invitrogen)	120
DTT (5 mM)	60
dATP (100 mM) (Sigma)	3
dCTP (100 mM) (Sigma)	3
dGTP (100 mM) (Sigma)	3
dTTP (100 mM) (Sigma)	0,6
AGUA DEPC	110,4
Volumen final	300

Seguidamente, 130 µg de RNA total se combinaron en un tubo de polipropileno con 2 µg de oligo (dT) (Invitrogen) y con agua tratada con DEPC hasta un volumen total de 20µl. Se mezcló la preparación y se incubó a 70°C durante 10 minutos, enfriándola posteriormente en hielo durante un minuto. Se centrifugó la mezcla brevemente a 4°C. A esta mezcla de ARN se añadieron 20 µl de la mezcla principal de marcación preparada en el paso anterior en las proporciones que se muestran en la Tabla 6.

**Tabla 6.** Reacción de marcación del cDNA

<b>Reactivos</b>	<b>Volumen (µl)</b>
Mezcla para el marcado RT	20µl
Cy3 ó Cy5-dUTP (1nmole/µl) (Amersham Pharmacia)	4µl
Superscript™ II RT (200U/µl) (Invitrogen)	2,66µl

Se mezclaron los componentes de la reacción exhaustivamente en tubos de polipropileno cubiertos con papel de aluminio para evitar la exposición a la luz de los fluorocromos y se incubaron a 42°C durante 2 horas. Seguidamente se centrifugó brevemente la reacción.

La reacción se detuvo añadiendo 2,0 µl de EDTA pH 8.0 20 mM. A continuación se añadieron 2,0 µl de NaOH 500 mM y se calentó la mezcla a 70°C durante 10 minutos. Se neutralizó la reacción con la incorporación de 2,0 µl de HCl 500 mM.

#### **Purificación, cuantificación y concentración del cDNA**

Los nucleótidos fluorescentes no incorporados durante la síntesis de cDNA fueron eliminados con un kit de purificación de productos de PCR, QIAquick PCR purification kit (QIAGEN), provisto de columnas con membranas de sílice, siguiendo el protocolo descrito por la casa comercial.

La elución de los productos purificados se realizó con 50 µl de H<sub>2</sub>O Milli Q. Se cuantificó el cDNA sintetizado, la cantidad de fluorocromo incorporado y la frecuencia de incorporación del mismo según las fórmulas que se indican a continuación:

$$\text{Sonda (ng de cDNA)} = \text{Absorbancia a 260nm} \times 37 \times \text{volumen final de sonda (µl)}$$

$$\text{Fluorocromo incorporado Cy3 (pmoles)} = \frac{\text{Absorbancia a 550nm} \times \text{volumen final de muestra}}{0,15}$$

$$\text{Fluorocromo incorporado Cy5 (pmoles)} = \frac{\text{Absorbancia a 650nm x volumen final de muestra}}{0,25}$$

$$\text{Frecuencia de incorporación (f.o.i)} = \frac{\text{fluorocromo incorporado (pmoles) x 324,5}}{\text{cantidad de cDNA (ng).}}$$

Dada la baja concentración de cDNA que se obtiene, la absorbancia de las muestras se midió directamente, sin diluciones, utilizando ultramicrocubetas (Hellma).

Los valores óptimos de f.o.i se encuentran entre 20 y 50, considerando como umbral para nuestros experimentos la utilización de sondas con una f.o.i.  $\geq 16$ , con el fin de asegurar buenos resultados. El volumen de muestra correspondiente a 65 pmoles de fluorocromo se concentró utilizando filtros microcon-30 (Amicon) hasta conseguir un volumen final de 5 a 10  $\mu\text{l}$ .

### **Pre-hibridación, hibridación y lavado de micromatrices**

Siguiendo el protocolo propuesto por TIGR (*Solanaceae Genomics Resource*, [http://jcvi.org/potato/images/SGED\\_SOP\\_6.1.1.pdf](http://jcvi.org/potato/images/SGED_SOP_6.1.1.pdf)), se preparó una solución tampón de prehibridación que contenía 5 x SSC, 0,1 % SDS, 1% de albúmina de suero bovino (BSA, Sigma) y 10 mM EDTA. Esta solución se filtró a través de un microfiltro de 0,2  $\mu\text{m}$  antes de su utilización. La solución fue precalentada a 42°C durante al menos 30 minutos.

Pasado este tiempo, se colocó el/los portaobjetos a pre-hibridar en un vaso de Coplin (VWR) que fue llenado con la solución de pre-hibridación precalentada y se incubó a 42°C durante 45 minutos. Finalizada la incubación se procedió a realizar dos lavados por agitación con agua Milli Q a temperatura durante 5 minutos cada uno.

Se realizó un último lavado de los portaobjetos con isopropanol a temperatura ambiente durante 2 minutos. Los portaobjetos se secaron mediante centrifugación a 500 rpm durante 5 minutos. Esta centrifugación se realizó en tubos de polipropileno de 50 ml, cuya base cónica estaba perforada.

Los portaobjetos se utilizaron inmediatamente después de la prehibridación. Por su parte para la hibridación se preparó una solución tampón según la Tabla 7.

**Tabla 7.** Solución tampón de hibridación

<b>Reactivos</b>	<b>Volumen (µl)</b>
Formamida	20 µl
20X SSC	10 µl
10% de SDS	0,4 µl

Se combinaron las muestras de cDNA procedentes de cada una de los tratamientos marcadas cada una de ellas con un fluorocromo distinto (correspondiente al control y al estrés respectivamente de un momento del tratamiento de sequía) en un mismo tubo, se secó la sonda de hibridación en una bomba de vacío SpeedVac y se resuspendió en 16,25µl de agua MilliQ. Para resuspender la sonda se añadieron 2µl de 0,5 M de EDTA y se mezcló, seguidamente se añadió 1 µl de poli (A) de ADN (20µg/µl) (Pharmacia) y 0,75 ul de esperma de salmón (10 mg / ml) (Invitrogen) a fin de bloquear hibridaciones inespecíficas.

Para cada hibridación se mezclaron las muestras de cDNA procedentes de cada una de las dos cepas a comparar (marcada una con Cy3 y otra con Cy5) y 5 µl de DNA de esperma de salmón (Sigma-Aldrich) previamente desnaturalizado por calor.

El cDNA se desnaturalizó manteniendo la muestra 2 minutos a 95°C y enfriándola inmediatamente en hielo, seguidamente se añadió tampón de hibridación (Dig Easy Hybridization buffer, Roche) hasta un volumen final de 60µl.

La mezcla se depositó entonces sobre un cubreobjetos de plástico (60x24 mm, Sigma-Aldrich), la micromatriz se puso en contacto con la muestra por la cara que contenía el DNA y una vez que la muestra se extendió por aproximadamente la mitad del cubreobjetos se dio la vuelta al micromatriz y se

colocó en una cámara de hibridación (Corning), que se sumergió en un baño a 42°C.

Después de 12 horas de hibridación, se retiró el cubreobjetos introduciendo la micromatriz en una cubeta con la primera solución de lavado (SDS 1%/SSC 2X), se transfirió a un tubo Falcon con la misma solución donde se mantuvo 5 minutos a temperatura ambiente con agitación ocasional y protegido de la luz. Para eliminar los restos de SDS, la micromatriz se pasó sucesivamente por dos tubos de polipropileno de 50ml con SSC 2X, en el último de ellos se mantuvo 5 minutos con agitación ocasional. Por último, la micromatriz se secó mediante centrifugación a 1100 rpm durante 5 minutos a temperatura ambiente.

### Diseño del ensayo de micromatrices

A fin de evaluar el efecto del estrés hídrico sobre la expresión del genoma en tres especies silvestres de *Solanum* (PSC 2877, CPH 18325 y VNT 8239), se utilizaron micromatrices de ADNc de dos colores provistas por TIGR. Los tratamientos fueron Testigo (no estresado) (T) y estresado (E). Esta comparación se realizó para una secuencia temporal: 0, 5, 7 y 10 días. Por su parte la especie PSC 2877 también incluyó los días 1 y 3. El esquema detallado del experimento para cada especie se muestra en las Tablas 8, 9 y 10.

**Tabla 8. Esquema del diseño de micromatrices utilizado para la especie I- PSC 2877**

Especie	Descripción slides	Cy3	Cy5
I – PSC 2877	1 - Día 0	Control	Estrés
I – PSC 2877	2 - Día 1	Control	Estrés
I – PSC 2877	3 - Día 3	Control	Estrés
I – PSC 2877	4 - Día 5	Control	Estrés
I – PSC 2877	5 - Día 7	Control	Estrés
I – PSC 2877	6 - Día 10	Control	Estrés
I – PSC 2877	7 - Día 0_1	Estrés	Control
I – PSC 2877	8 - Día 1_1	Estrés	Control
I – PSC 2877	9 - Día 3_1	Estrés	Control
I – PSC 2877	10 - Día 5_1	Estrés	Control
I – PSC 2877	11 - Día 7_1	Estrés	Control
I – PSC 2877	12 - Día 10_1	Estrés	Control

*Dye-swap*

**Tabla 9. Esquema del diseño de micromatrices utilizado para la especie II, CPH 18325**

Especie	Descripción slides	Cy3	Cy5	
II - CPH 18325	1 - Día 0_1	Control	Estrés	
II - CPH 18325	2 - Día 0_2	Control	Estrés	
II - CPH 18325	3 - Día 0_3	Estrés	Control	] <i>Dye-swap</i>
II - CPH 18325	4 - Día 0_4	Estrés	Control	
II - CPH 18325	5 - Día 5_1	Control	Estrés	
II - CPH 18325	6 - Día 5_2	Control	Estrés	
II - CPH 18325	7 - Día 5_3	Estrés	Control	] <i>Dye-swap</i>
II - CPH 18325	8 - Día 5_4	Estrés	Control	
II - CPH 18325	9 - Día 7_1	Control	Estrés	
II - CPH 18325	10 - Día 7_2	Control	Estrés	
II - CPH 18325	11 - Día 7_3	Estrés	Control	] <i>Dye-swap</i>
II - CPH 18325	12 - Día 7_4	Estrés	Control	
II - CPH 18325	13 - Día 10_1	Control	Estrés	
II - CPH 18325	14 - Día 10_2	Control	Estrés	
II - CPH 18325	15 - Día 10_3	Estrés	Control	] <i>Dye-swap</i>
II - CPH 18325	16 - Día 10_4	Estrés	Control	

**Tabla 10. Esquema del diseño de micromatrices utilizado para la especie III, VNT 8239**

Especie	Descripción slides	Cy3	Cy5	
III - VNT 8239	17 - 1' - Día 0_1	Control	Estrés	
III - VNT 8239	18 - 2' - Día 0_2	Control	Estrés	
III - VNT 8239	19 - 3' - Día 0_3	Estrés	Control	] <i>Dye-swap</i>
III - VNT 8239	20 - 4' - Día 0_4	Estrés	Control	
III - VNT 8239	21 - 5' - Día 5_1	Control	Estrés	
III - VNT 8239	22 - 6' - Día 5_2	Control	Estrés	
III - VNT 8239	23 - 7' - Día 5_3	Estrés	Control	] <i>Dye-swap</i>
III - VNT 8239	24 - 8' - Día 5_4	Estrés	Control	
III - VNT 8239	25 - 9' - Día 7_1	Control	Estrés	
III - VNT 8239	26 - 10' - Día 7_2	Control	Estrés	
III - VNT 8239	27 - 11' - Día 7_3	Estrés	Control	] <i>Dye-swap</i>
III - VNT 8239	28 - 12' - Día 7_4	Estrés	Control	
III - VNT 8239	29 - 13' - Día 10_1	Control	Estrés	
III - VNT 8239	30 - 14' - Día 10_2	Control	Estrés	
III - VNT 8239	31 - 15' - Día 10_3	Estrés	Control	] <i>Dye-swap</i>
III - VNT 8239	32 - 16' - Día 10_4	Estrés	Control	

## Herramientas de Bioinformática

### Escaneo de chips

El escaneo de los chips fue realizado en la plataforma de genómica del Instituto de Investigación del Cáncer de Salamanca, que consta de un escanner GenePix 4000B, cuyas imágenes fueron procesadas con el paquete informático GenePix Pro 5.0.0.49 (Axon).

### Pre-procesamiento de micromatrices

Los datos fueron procesados mediante fgStatistics y la librería limma de R. Las micromatrices fueron leídas mediante la función `read.maimage` utilizando el formato `genepix`. Para la corrección por background se utilizó la función `backgroundCorrect` con el método "normexp" y `offset=50`. Para la normalización dentro de micromatrices se aplicó la función `normalizeWithinArrays` con todas sus opciones por defecto. La normalización entre micromatrices se realizó mediante la función `normalizeBetweenArrays` utilizando el método "Aquantile". Dentro de cada micromatriz se disponía de dos réplicas técnicas de cada gen. Los log<sub>2</sub> de los fold changes de la condición estresante respecto del control (LogFC) dentro de cada micromatriz fueron promediados. Durante la lectura y normalización cada lectura es evaluada por su calidad. Aquellas a las cuales se les asignó peso (`weight`) cero fueron descartadas (se les asignó valor faltante). Los genes que no tenían datos completos para todos los tiempos y repeticiones fueron descartados. También fueron descartados las entradas identificadas con ID nulo o 0, que corresponden a lecturas nulas o genes no validados (`Name=null` ó `Do not use. Not validated` respectivamente).

### Procesamiento de micromatrices

El diseño del experimento con respecto a las micromatrices fue de tipo *dye swap* alternando sólo los tratamientos T y E. El número de repeticiones para cada tiempo varió de 2 en el especie PSC 2877 a 4 para los especies CPH 18325 y VNT 8239 (Tablas 8, 9 y 10).

El análisis se basó en una estrategia en dos etapas. En la primera se realizó un ajuste de un modelo lineal en cada instancia de la evaluación (tiempo) y para cada gen. El propósito fue estimar el efecto del estrés y la asignación de un p-

valor a ese efecto para cada instancia de evaluación. La siguiente etapa: la búsqueda, dentro del conjunto de genes con p-valor significativo en alguna de las instancias de tiempo consideradas, de patrones de respuesta en el tiempo. Para la primera etapa se utilizó la función *lmFit* del paquete *limma*. Para compensar el escaso número de repeticiones disponible para cada gen (4 ó 2 según especie), las pruebas de hipótesis fueron recalculadas mediante la función *eBayes*, con sus parámetros por defecto.

La segunda etapa parte de una matriz para cada especie que resume el resultado del ajuste de los modelos lineales. Estas matrices tienen tantas filas como genes considerados (en cada especie) y en las columnas, para cada tiempo considerado, el logFC y el p-valor ajustado por inferencia simultánea mediante el procedimiento de Benjamini-Hochberg. Para cada matriz se realizó un análisis de componentes principales de los logFC, en los tiempos: 0, 5, 7 y 10 días. En el espacio generado, los genes aparecen como puntos alrededor del origen (punto cero). Cuanto más distantes están del origen más diferenciados están en uno o varios tiempos de evaluación. Luego, los genes representados en la periferia de la nube de puntos representan los genes que se tienen los fold changes más extremos. Teniendo en cuenta este principio, se seleccionaron aquellos genes que estuvieran por encima del percentil 0.95 de la distribución de distancias al origen del plano de representación seleccionado y que cumplieran simultáneamente la condición de tener al menos un p-valor ajustado (Benjamini-Hochberg, 1995) menor que 0,05 para algún tiempo de evaluación.

En el Material Complementario (CD) se adjunta el código utilizado en el programa de Bioconductor para el procesamiento de micromatrices.

### **Construcción de la base de datos estructurada**

Para generar una base de datos que congregue toda la información inherente a la constitución de la micromatriz 10K de papa de TIGR, se combinó la información contenida en el archivo tipo GAL (del inglés *GenePix Array List*), en particular la versión 4, que corresponde a los micromatrices utilizados y la información actualizada de los Índices de Genes de Papa, disponible en el sitio web <http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=potato>. Este último paso busca ampliar y actualizar la información contenida, en particular

los conceptos GO. Seguidamente se muestra una salida acotada del archivo GAL que resume los códigos de los clones depositados en la micromatriz y su correspondiente anotación biológica que serán de utilidad para interpretar los resultados del análisis estadístico implementado en las secciones previas (Tabla en Material Complementario).

La base de datos esta estructura de la siguiente manera:

**Nombre del Clon:** identificador único para este clon de papa que tienen la siguiente estructura STMXX## (códigos XX: letras, ##: numérico).

**Anotación:** hace referencia a la anotación putativa del gen que ha sido asignada en la base a un análisis de secuencias tipo BLASTX donde se estableció como límite mínimo un valor E de 1e-5.

**TC actual:** el consenso provisional (o tentative consensus TC) actual, es el identificador de secuencia asignado a ese clon por el Instituto de Cancer Dana-Farber en el marco del Proyecto de Índice de Genes de Papa.

**TC retirados:** corresponden a anotaciones anteriores de la Base de Datos de Índice de Genes de Papa (DFCI) cuyo identificador (dada las nuevas tecnologías) ha sido asignado a un nuevo TC.

**Número de la base de datos del Genbank:** El número/s de introducción/es del Genbank (s) para el EST 5' 3' de este clon.

**EST:** corresponde a los identificadores de ESTs utilizado para crear el TC asociado a este clon.

**GO y GO-Slim:** los conceptos de GO se obtuvieron en base al alineamiento de las secuencia de los ESTs de papa con bases de datos de Arabidopsis (TIGR y TAIR), utilizando el algoritmo BLASTX, con un corte de rigor de <1e-10.

Esta base fue actualizada con la herramienta BLAS2GO disponible en <http://www.blast2go.org/>, a fin de contar con información actualizada sobre las ontologías génicas de estos clones.

## **Validaciones biológicas. RT-PCR**

Ensayos biológicos de estrés hídrico

Las validaciones de los ensayos de micromatrices fueron realizadas en Argentina y a fin de contar con cantidad suficiente de ARN para la síntesis de

ADNc y las posteriores validaciones por PCR cuantitativa, se realizaron dos nuevos ensayos biológicos bajo las mismas condiciones en las que se realizó el ensayo original, en un invernadero del Grupo Propapa del EEA Balcarce.

A fin de validar los genes de las micromatrices, se seleccionaron dos momentos del tratamiento de estrés, en los cuales, los cambios de expresión fueron más evidentes. Para la especie CPH 18325 se seleccionaron los tiempos 05 y 07 y para la especie VNT 8239, los tiempos 07 y 10 (Tabla 11).

**Tabla 11. Ensayos biológicos de sequía realizados en la EEA Balcarce-INTA con las especies silvestres CPH y VNT.** Cada tiempo tiene asignado el nombre de la especie, del ensayo, del tiempo, del tratamiento y la repetición

<b>Especie_Ensayo_Tiempo_Tratamiento_Repetición</b>	
CPH_1_5_ C_1	CPH_2_5_ C_1
CPH_1_5_ C_2	CPH_2_5_ C_2
CPH_1_5_ S_1	CPH_2_5_ S_1
CPH_1_5_ S_2	CPH_2_5_ S_2
CPH_1_7_ C_1	CPH_2_7_ C_1
CPH_1_7_ C_2	CPH_2_7_ C_2
CPH_1_7_ S_1	CPH_2_7_ S_1
CPH_1_7_ S_2	CPH_2_7_ S_2
VNT_1_7_ C_1	VNT_2_7_ C_1
VNT_1_7_ C_2	VNT_2_7_ C_2
VNT_1_7_ S_1	VNT_2_7_ S_1
VNT_1_7_ S_2	VNT_2_7_ S_2
VNT_1_10_ C_1	VNT_2_10_ C_1
VNT_1_10_ C_2	VNT_2_10_ C_2
VNT_1_10_ S_1	VNT_2_10_ S_1
VNT_1_10_ S_1	VNT_2_10_ S_1

*Especies:* CPH 18325 y VNT 8239

*Ensayos:* 1 y 2 constituyen réplicas biológicas del tratamiento de estrés hídrico.

*Tiempo:* momentos del tratamiento de estrés hídrico analizado. Tiempos: 05, 07 y 10, indican días después de iniciado el estrés hídrico.

*Tratamiento:* se asignó una C al control sin estresar y una S, al sometido a estrés hídrico.

*Repeticiones:* 1 y 2 constituyen réplicas técnicas del control y estresado.

### **Extracción de ARN total**

A partir de 100mg de tejido molido según se indica en la sección 2.2.1., se procedió a la extracción de ARN total mediante el Kit comercial RNAqueous (Ambion), siguiendo las instrucciones del fabricante. La concentración y la calidad del ARN en solución se estimaron a partir de las absorbancias a 260 y a 280nm. Considerando que D.O (260 nm) = 1 equivale a 40µg/ml de ARN y que a 280nm, se da el pico máximo de absorción de las proteínas, se puede verificar que las relación de Abs.260/Abs.280 cercanas a 2, indican que el ARN analizado es de alta calidad y prácticamente libre de proteínas. Para las mediciones se utilizó el espectrofotómetro Nano Drop (Thermo Scientific).

### **Tratamiento de ADNasa I**

A fin de eliminar el posible ADN contaminante de las muestras de ARN, se hace necesario tratar las mismas con DNasa I. Este tratamiento es especialmente requerido en experimentos de RT-PCR, ya que pueden obtenerse falsos positivos debido a la presencia de ADN contaminante.

Previo a la síntesis de ADNc, 1,5µg de RNA se combinó con la enzima ADNasa I y una solución tampón para la misma (ambas provistas por Invitrogen) como se muestra en la Tabla 12:

**Tabla 12 Tratamiento de ARN total con la enzima endonucleasa ADNasa I.**

<b>Reactivos</b>	<b>Volumen (µl)</b>
10 × Buffer DNasa I (Invitrogen)	1
DNasa I Amp. Grade (Invitrogen)	1
AGUA DEPC	Llevar a volumen final (10µl)

La mezcla fue incubada 15 minutos a temperatura ambiente. Para frenar la reacción se agregó 1µl de EDTA 25mM, provisto en el Kit comercial de la

enzima y posteriormente se inactivó la reacción por tratamiento con calor (10 minutos a 65°C).

### Síntesis de ADNc

Para la síntesis de ADNc, se utilizó como molde inicial el ARN tratado en el paso anterior (10µl), en un proceso que se resume en dos etapas detalladas a continuación (Tablas 13 y 14):

**Tabla 13. Síntesis de ADNc: Primera etapa**

Reactivos	Volumen (µl)
hexanucleótidos al azar (150ng/ml) (Invitrogen)	1
10mM dNTPs (Invitrogen)	1
H <sub>2</sub> O	1

La reacción se activó calentando la mezcla 5 minutos a 65°C. Se frenó la misma con una incubación en hielo durante 1 minuto (se sugiere hacer una centrifugación de corto tiempo o *spin*, de la mezcla antes de continuar). A fin de completar la síntesis de la primera y la segunda hebra del ADNc, en cada tubo de reacción se incluyó:

**Tabla 14. Síntesis de ADNc: Segunda etapa**

Reactivos	Volumen (µl)
5 x Buffer First Strand (Invitrogen)	4
0,1 M DTT (Invitrogen)	1
RNAsa out (Invitrogen)	1
SuperScript III	1

Se incubó la reacción a 25°C durante 10 minutos, luego de la cual se centrifugó brevemente. Se continuó con una incubación a 55°C durante 30 minutos. Finalmente, la mezcla se incubó durante 15 minutos a 70°C para inactivar la reacción.

A modo de prueba de la calidad del ADNc sintetizado, se realizó una reacción de amplificación con oligonucleótidos de Actina, incluyendo entre las muestras,

controles de ADNc y ADN. La reacción de prueba se formuló como se indica en la Tabla 15.

**Tabla 15. Componentes de la mezcla maestra de reacción de PCR con oligonucleótidos de Actina.**

<b>Reactivos</b>	<b>Volumen (µl)</b>
10 X PCR Buffer (Invitrogen)	2.5
Oligonucleótido Actina UP 0,5 uM	0.5
Oligonucleótido Actina LOW 0,5 uM	0.5
dNTPs 10mM	0.5
ADNc 60 ng	1
Cl <sub>2</sub> Mg 50mM	0.75
H2O cromasol	Llevar a 25µl

La reacción se llevó a cabo en un termociclador Biorad, utilizando el siguiente programa: 10 min a 94°C, seguido de 35 ciclos de 1 min a 94°C, 1 min a 50°C y 30 seg a 72°C, con una extensión final de 10 min a 72°C.

Se realizaron diluciones del ADNc obtenido (1:10, 1:50 y 1:100) para las pruebas de templado en la reacción de PCR cuantitativa del paso posterior.

### **PCR cuantitativa o en tiempo real**

Las reacciones de PCR cuantitativa se realizaron utilizando el kit comercial QuantiTect™ SYBR® Green PCR (QIAGEN), siguiendo las recomendaciones del fabricante, modificando el volumen final de la reacción a 25µl. Los oligonucleótidos utilizados en este proceso fueron sintetizados utilizando el programa Primer 3 (<http://frodo.wi.mit.edu/primer3/>), teniendo en cuenta la guía de recomendaciones del proveedor del kit. Las reacciones se llevaron a cabo según se describe en la Tabla 16.

**Tabla 16. Componentes de la mezcla maestra de reacción de PCR cuantitativa**

<b>Reactivos</b>	<b>Volumen (<math>\mu</math>l)</b>
2X QuantiTect SYBR Green PCR 1X (QUIAGen)	12,5
Oligonucleótido A 0,5 $\mu$ M	2
Oligonucleótido B 0,5 $\mu$ M	2
ADNc 60 ng (dilución 1:10)	1
H2O bidestilada libre de RNAs (QIAGen)	Llevar a 25 $\mu$ l

Las reacciones se llevaron a cabo en un equipo 7500 Real-Time PCR System (Applied Biosystems) utilizando el siguiente programa: 15 seg a 95°C, 1 min a 60°C y 30 seg a 72°C por 40 ciclos. Una vez termina da la amplificación se creó la curva de disociación, medida como los cambios en los niveles de fluorescencia en función de la temperatura, de modo de detectar los productos no específicos. El programa de disociación utilizado fue el siguiente: 15 seg a 95°C, 60 seg a 60°C.

Se realizaron 3 réplicas de cada reacción. Se utilizó como gen control de normalización, el factor de elongación 1 alfa (EF1 $\alpha$ ), cuya expresión constitutiva fue reportada para el cultivo de papa (Nicot *et al*, 2005) frente a estreses bióticos y abióticos.

La cuantificación de los cambios relativos en la expresión génica se realizó utilizando el método 2- $\Delta\Delta$ CT según lo descrito por Livak y Schmittgen (2001).

### **Herramientas de visualización de datos de expresión génica**

A fin de graficar las rutas metabólicas utilizamos una herramienta que muestra grandes conjuntos de datos, como lo son los datos de expresión de genes, bajo la forma de diagramas de las rutas metabólicas y otros procesos. Esta herramienta informática se denomina MapMan (Thimm *et al*, 2004) y se encuentra accesible en el sitio <http://mapman.gabipd.org/web/guest/mapman>. El programa desarrollado inicialmente para trabajar con datos de expresión de *Arabidopsis* fue actualizando (Usadel *et al*, 2005) y se logró adaptar para su utilización en solanáceas (Urbanczyk-Wochniak, *et al*, 2006; Rotter *et al*, 2007).

Para este trabajo se utilizó la versión 3.5.1 (19.11.2010) del programa. La información con la que trabaja el sistema esta estructurada en tres bases de datos: los mapas, las rutas metabólicas y los datos de expresión.

Los mapas y las rutas metabólicas se descargaron del sitio <http://mapman.gabipd.org/web/guest/mapmanstore>, en particular accediendo al vínculo denominado "Stu\_TIGR download".

Este vínculo, presenta la clasificación de los contigs provisional (TC) de papa obtenidos de la base de datos del Índice de Genes de la Papa (StGI) disponibles en Dana-Farber Cancer Institute (DFCI) (<http://compbio.dfci.harvard.edu/cgi-bin/tgi/qimain.pl?gudb=potato>). De esta manera puede utilizarse esta herramienta para el análisis particular de micromatrices de papa. En el desarrollo de las bases que utiliza el MapMan, pusieron especial atención en el mapeo de genes que no podrían ser anotados en base a su similitud con los genes de *Arabidopsis*, dado que posiblemente representarían genes únicos de la papa. 97 de estos genes se han clasificado en "BIN" o clases funcionales, luego de su anotación manual.

El programa consta de dos módulos separados, el módulo SCAVENGER que recoge y clasifica los parámetros medidos en un conjunto de categorías jerárquicas funcionales (BINs, subBINs...enzima individual) y el módulo IMAGEANNOTATOR que importa las clasificaciones y las utiliza para organizar y mostrar datos en diagramas.

## **5. RESULTADOS**

## 5. RESULTADOS

Con el propósito de monitorear los cambios globales de transcripción génica implicados en las respuestas defensivas de especies silvestres de papa frente al estrés hídrico, se llevó a cabo un análisis comparativo de los perfiles de transcripción, mediante la técnica de micromatrices entre plantas de *Solanum cardiophyllum* (CPH 18325), *S. venturii* (VNT 8239) y *S. pascoense* (PSC 2877). Se compararon plantas que crecieron en condiciones óptimas y plantas expuestas a sequía y se estudiaron los principales mecanismos puestos en marcha para mantener el metabolismo y limitar el daño bajo estas condiciones. Este estudio requirió de una serie de etapas llevadas a cabo en colaboración con diferentes centros de investigación, que partieron desde la selección y evaluación del germoplasma de papa en el Instituto Vasco de Investigación y Desarrollo Agrario Neiker-Tecnalia, de Vitoria-Gasteiz (España), las hibridaciones de micromatrices realizadas en la Universidad de Salamanca (España) y las realizadas por el Servicio de Análisis de expresión de TIGR (USA), el posterior análisis bioinformático de los datos de expresión realizado con el grupo de Biología Computacional del Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada (España) y el Instituto alemán DKFZ, el análisis estadístico de los datos de expresión en la cátedra de Estadística y Biometría Universidad Nacional de Córdoba (Argentina), concluyendo con ensayos de validaciones biológicas y ensayos biológicos complementarios en Argentina, en el Instituto de Biotecnología-CICVyA, del INTA Castelar y del Grupo de Investigación en Papa, de la Estación Experimental Agropecuaria de INTA, respectivamente.

### **Evaluación de especies silvestres de papa frente al estrés hídrico.**

El comportamiento frente al estrés hídrico se evaluó según el protocolo propuesto por TIGR, quien implementó este procedimiento en estudios de perfiles de transcripción realizados con Solanáceas ([ftp://ftp.tigr.org/pub/data/s\\_tuberosum/SGED//055\\_TIGR\\_drought1/](ftp://ftp.tigr.org/pub/data/s_tuberosum/SGED//055_TIGR_drought1/)).

El tratamiento de estrés, que consistió en la supresión del riego por un período de 10 días, permitió clasificar las 14 especies de papa silvestres evaluadas (Tabla 3) sobre la base del contenido relativo de agua (RWC), un indicador útil para establecer el balance hídrico de una planta. Estas mediciones fueron contrastadas asimismo con evaluaciones de la reducción del potencial de agua de las hojas en las plantas de las especies estudiadas con una cámara de presión de Scholander (Scholander. *et al*, 1964) que arrojaron una clasificación equivalente.

Al cabo del período de estrés, entre las 14 especies de papa silvestres analizadas, hubo materiales cuyo aspecto fenotípico no fue muy diferente al control irrigado, mientras que otros materiales mostraban características de susceptibilidad al déficit hídrico, que reflejaba diferencias de adaptación al estrés impartido. En la Figura 7 A-B y C puede observarse el aspecto fenotípico de las plantas control y las sometidas a estrés (recuadro en azul) al finalizar el ensayo de sequía, con los correspondientes valores de RWC.

A

Control

Sequía



RWC

*S. bulbocastanum*  
(BLB): 83%

*S. sparsipilum*  
(SPL): 82%

*S. leptophyes*  
(LPH): 79%

*S. gandarillasii*  
(GND): 62%

Figura 7A. Fenotipo de las especies silvestres de papa BLB, SPL, LPH y GND al finalizar el tratamiento de sequía. A la derecha, con recuadro azul se muestran las plantas al finalizar el tratamiento de sequía con sus valores de RWC y a la izquierda las plantas control.

Control

Sequía

B



RWC

*S. doddsii*  
(DDS): 92%

*S. polytrichon*  
(PLT): 89%

*S. sanctae-rosae*  
(SCT): 83%

*S. infundibuliforme*  
(IFD): 62%

Figura 7B. Fenotipo de las especies silvestres de papa DDS, PLT, SCT y IFD al finalizar el tratamiento de sequía. A la derecha, con recuadro azul se muestran las plantas al finalizar el tratamiento de sequía con sus valores de RWC y a la izquierda las plantas control.

C

Control

Sequía



RWC

*S. berthaultii*  
BER: **96%**

*S. marinasense*  
MRN: **89%**

*S. papita*  
PTA: **68%**

Figura 7B. Fenotipo de las especies silvestres de papa VER, MRN y PTA al finalizar el tratamiento de sequía. A la derecha, con recuadro azul se muestran las plantas al finalizar el tratamiento de sequía con sus valores de RWC y a la izquierda las plantas control.

De todos los materiales evaluados, se seleccionaron tres especies que presentaron comportamientos divergentes frente al estrés. En particular, se buscó cubrir todos niveles de sensibilidad al estrés, seleccionando especies con RWC alto, intermedio y bajo, a fin de ser útiles en la identificación de genes candidatos involucrados en la respuesta a la sequía mediante análisis de micromatrices. La Tabla 17 resume los valores de RWC y de potencial agua de las hojas de las especies *S. cardiophyllum* (CPH 18325), *S. venturii* (VNT 8239) y *S. pascoense* (PSC 2877) y su relación con el *Solanum tuberosum* cv Kennebec, a los 10 días de ausencia de irrigación, a fin de tener un valor de referencia del estado hídrico de la papa cultivada en relación a las especies silvestres de *Solanum* sometidas a estrés.

**Tabla 17. Evaluación del tratamiento de estrés hídrico.** RWC y potencial de agua de las hojas de las especies de *Solanum* seleccionados para los ensayos de micromatrices. Las mediciones se realizaron luego de someter a las plantas al tiempo máximo de sequía evaluado (10 días).

Especie	PF (g)	PH (g)	PS (g)	RWC (%)	RWC promedio (%)	MPa (Scholander)	
						C	S
CPH 18325	0,5109	0,5851	0,0685	85,64	87,54	-0.9	- 1.5
	0,9942	1,0968	0,1243	89,45			
VNT 8239	1,2642	1,7415	0,1636	69,75	68,60	-0.8	- 2.5
	1,0016	1,442	0,0889	67,45			
PSC 2877	0,0615	0,062	0,0085	99,07	99,08	-0.8	- 1.2
	0,0645	0,065	0,0093	99,10			
<i>S. tuberosum</i> cv. Kennebec	1,9372	2,8838	0,2565	63,97	74,09	-0.9	- 2.1
	1,907	2,2401	0,1313	84,20			

PF: peso fresco, PH: peso hidratado, PS: peso seco, expresado en gramos (g). Los valores de contenido relativo de agua (RWC) están expresados en porcentaje (%). Las unidades de medición con la cámara Scholander están expresadas en MPa (MegaPascuales). C: plantas control, S: plantas estresadas

El aspecto que presentaron las especies seleccionadas, CPH, VNT y PSC, tras ser sometidas a estrés hídrico puede verse en la Figura 8. En el recuadro de la

izquierda se identifican las plantas estresadas y a la derecha sus controles irrigados.

**Control**

**Sequía**



Figura 8. Fenotipo de las especies *S. cardiophyllum* (CPH), *S. pascoense* (PSC) y *S. venturii* (VNT) al finalizar el tratamiento de sequía (10 días). A la derecha de la imagen,

recuadradas en azul se muestran las plantas sometidas a estrés hídrico, a la derecha sus controles irrigados.

### **Análisis de la expresión génica diferencial de especies silvestres de *Solanum***

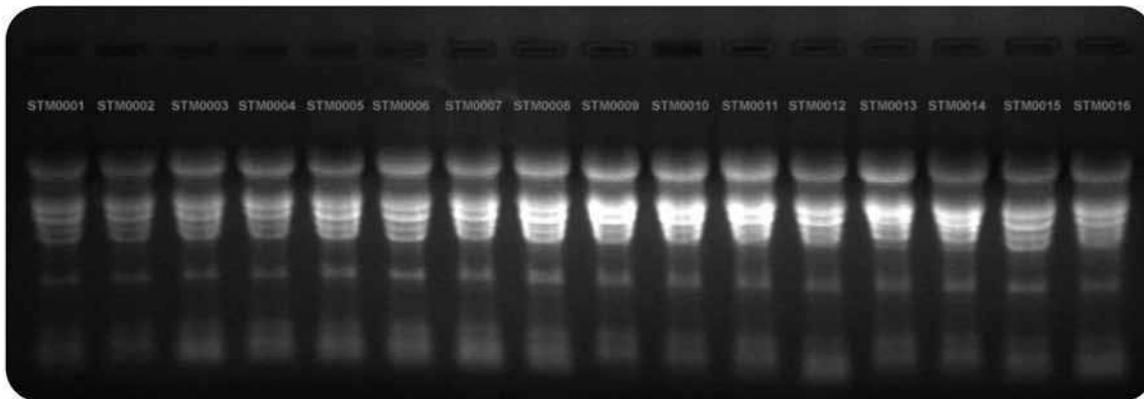
Para realizar el análisis transcripcional, se cortaron las hojas de 3 plantas de cada especie y tratamiento (control y estresada) en 6 momentos del estrés hídrico: T00, T01, T03, T05, T07 y T10, que correspondieron a los días posteriores a iniciar el tratamiento de sequía. Las hojas cortadas de las 3 plantas fueron agrupadas e inmediatamente congeladas en nitrógeno (Figura 9).



**Figura 9: Metodología de muestreo.** Hojas de plantas estresadas y control se cortaron en seis momentos del estrés hídrico: T00, T01, T03, T05, T07 y T10 para las 14 especies silvestres de papa sometidas a estrés hídrico. Las hojas de cada grupo (control y estresado) fueron agrupadas.

Se obtuvieron 72 muestras provenientes de 14 especies, 2 tratamientos y 6 momentos de estrés, que fueron pulverizadas con nitrógeno líquido y almacenadas a -80°C para su posterior procesamiento .

El ARN de las tres especies seleccionadas se extrajo exitosamente en cantidad y calidad para ser empleado como templado en la técnica de micromatrices. En la Figura 10 se muestra un ejemplo de la calidad e integridad de las muestras utilizadas en los experimentos.



**Figura 10. Extracción de ARN de las 14 especies silvestres de papa.** Electroforesis de ARN total en condiciones desnaturalizantes utilizando una solución tampón de corrida ME. Calles 1 a 16: muestras de ARN de las especies VNT y CPH, que se enviaron a la plataforma TIGR. La codificación de cada calle del gel (STM0001 al 0016) corresponde al código asignado por TIGR a las muestras procesadas en su Plataforma.

Las muestras de ARN de las especies silvestres CPH y VNT se enviaron al Servicio de Perfiles de Expresión de TIGR (MD, USA) en hielo seco y precipitadas con etanol y acetato de sodio 3M, con el fin de evitar la degradación del RNA durante el envío puerta a puerta. Al llegar a destino, el ARN se re-precipitó y se comprobó que la concentración y la integridad, determinada por medio de electroforesis en un gel de agarosa en condiciones desnaturalizantes seguía siendo óptima para los experimentos de hibridación de micromatrices (Willem Albert, TIGR, comunicación personal, 2006).

Las muestras de ARN de la especie silvestre PSC se procesaron en el laboratorio de Neiker – Tecnalia, Instituto Vasco de Investigación y Desarrollo Agrario.

### **Hibridación de micromatrices 10K de papa (TIGR)**

Se hibridaron 32 micromatrices 10k de papa. Si bien el límite de hibridaciones que realiza el Servicio de expresión por Proyecto es de 30, se incluyeron dos hibridaciones adicionales para cumplir con un diseño balanceado del experimento. Estas hibridaciones se realizaron con las especies CPH y VNT (Tabla 18).

**Tabla 18. Hibridaciones *Solanum cardiophyllum* y *Solanum venturii* en TIGR-** Detalle de las 32 hibridaciones realizadas en TIGR. Se indican las muestras utilizadas en cada micromatriz, como templado control y estresado respectivamente, siguiendo el código asignado por TIGR.

Nombre de cada hibridación	Nombre de la sonda problema	Nombre de la sonda de referencia	Barcode
Solanum1_ctrl_vs_drought_0_1	STP0001_G01q	STP0002_R01r	13246406
Solanum1_ctrl_vs_drought_0_2	STP0001_G02q	STP0002_R02r	13246409
Solanum1_ctrl_vs_drought_0_3	STP0001_R03q	STP0002_G03r	13246417
Solanum1_ctrl_vs_drought_0_4	STP0001_R04q	STP0002_G04r	13246418
Solanum1_ctrl_vs_drought_5_1	STP0003_G01q	STP0006_R01r	13243718
Solanum1_ctrl_vs_drought_5_2	STP0003_G02q	STP0006_R02r	13243312
Solanum1_ctrl_vs_drought_5_3	STP0003_R03q	STP0006_G03r	13243313
Solanum1_ctrl_vs_drought_5_4	STP0003_R04q	STP0006_G04r	13243314
Solanum1_ctrl_vs_drought_7_1	STP0004_G01q	STP0007_R01r	13243315
Solanum1_ctrl_vs_drought_7_2	STP0004_G02q	STP0007_R02r	13243728
Solanum1_ctrl_vs_drought_7_3	STP0004_R03q	STP0007_G03r	13243705
Solanum1_ctrl_vs_drought_7_4	STP0004_R04q	STP0007_G04r	13243318
Solanum1_ctrl_vs_drought_10_1	STP0005_G01q	STP0008_R01r	13243319
Solanum1_ctrl_vs_drought_10_2	STP0005_G02q	STP0008_R02r	13243320
Solanum1_ctrl_vs_drought_10_3	STP0005_R03q	STP0008_G03r	13243321
Solanum1_ctrl_vs_drought_10_4	STP0005_R04q	STP0008_G04r	13243323
Solanum2_ctrl_vs_drought_0_1	STP0009_G01q	STP0010_R01r	13243324
Solanum2_ctrl_vs_drought_0_2	STP0009_G02q	STP0010_R02r	13243389
Solanum2_ctrl_vs_drought_0_3	STP0009_R03q	STP0010_G03r	13243326
Solanum2_ctrl_vs_drought_0_4	STP0009_R04q	STP0010_G04r	13243211
Solanum2_ctrl_vs_drought_5_1	STP00011_G01q	STP0014_R01r	13243212
Solanum2_ctrl_vs_drought_5_2	STP00011_G02q	STP0014_R02r	13243213
Solanum2_ctrl_vs_drought_5_3	STP00011_R03q	STP0014_G03r	13243215
Solanum2_ctrl_vs_drought_5_4	STP00011_R04q	STP0014_G04r	13243216
Solanum2_ctrl_vs_drought_7_1	STP00012_G01q	STP0015_R01r	13243217
Solanum2_ctrl_vs_drought_7_2	STP00012_G02q	STP0015_R02r	13243218
Solanum2_ctrl_vs_drought_7_3	STP00012_R03q	STP0015_G03r	13243219
Solanum2_ctrl_vs_drought_7_4	STP00012_R04q	STP0015_G04r	13243220
Solanum2_ctrl_vs_drought_10_1	STP00013_G01q	STP0016_R01r	13243387
Solanum2_ctrl_vs_drought_10_2	STP00013_G02q	STP0016_R02r	13243388
Solanum2_ctrl_vs_drought_10_3	STP00013_R03q	STP0016_G03r	13243390
Solanum2_ctrl_vs_drought_10_4	STP00013_R04q	STP0016_G04r	13243395

Se escanearon las intensidades de expresión de los canales rojo y verde, obteniendo como resultado una imagen en formato TIFF, por cada una de las micromatrices, que fue la resultante de la conversión de la corriente de fotones de los tintes Cy3 y Cy5, en valores digitales. Estas imágenes se procesaron con el Software GenPix 5.0 (Axon), a fin de generar 32 archivos de extensión GPS (del inglés *GenePix Setting*) y sus correspondientes archivos de resultados de extensión GPR (del inglés *GenePix Results*). Dado que el tamaño de los archivos de imagen de las micromatrices, así como el de datos de expresión del total de los *slides* evaluados en este trabajo, dificultan su inclusión en este manuscrito, se encuentran accesibles en el vínculo: [ftp://ftp.tigr.org/pub/data/s\\_tuberosum/SGED//102\\_Lucca/](ftp://ftp.tigr.org/pub/data/s_tuberosum/SGED//102_Lucca/), y en el CD adjunto a este trabajo como Material Complementario.

### **Hibridaciones en la Plataforma Genómica de la Universidad de Salamanca**

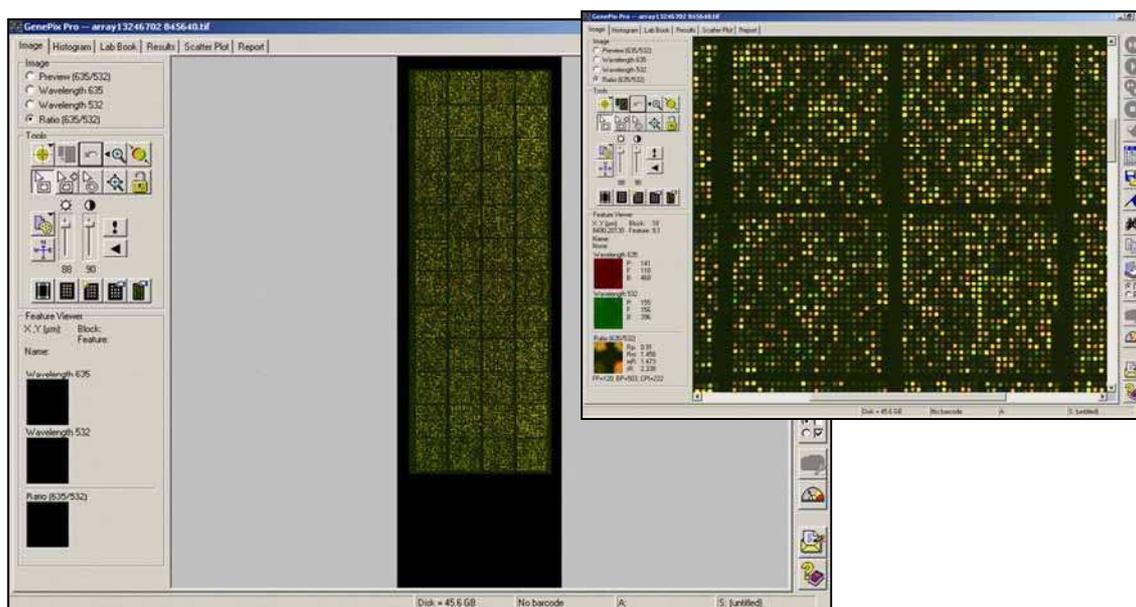
Se marcaron e hibridaron en la Plataforma de Genómica de la Universidad de Salamanca, 12 micromatrices utilizando como templado el ARN de la especie *S. pascoense*. Se siguió el protocolo propuesto por TIGR, que provee los soportes comerciales de 10K de papa y que en paralelo realizó las hibridaciones con las especies CPH y VNT, con la finalidad de realizar un posterior análisis comparativo de los perfiles de expresión de este germoplasma frente al estrés.

Las hibridaciones realizadas en Salamanca se resumen en la Tabla 19.

Una vez hibridadas las micromatrices, se obtuvieron 12 archivos de imagen en formato TIFF. Un ejemplo de las imágenes obtenidas de las micromatrices puede verse en la Figura 11, donde se muestran hibridaciones con buena señal y sin artefactos.

**Tabla 19. Detalle de las 12 hibridaciones realizadas en la Plataforma de Genómica de la Universidad de Salamanca.** Se indican las muestras utilizadas en cada micromatriz como templado control y estresado respectivamente.

Nombre de cada hibridación	Sonda marcada con Cy3	Sonda marcada con Cy5	Barcode
Solanum0_ctrl_vs_drought_0_1	PSC_0_1	PSC_0_2	0 - array13246692 805625
Solanum0_drought_vs_ctrl_0_2	PSC_0_2	PSC_0_1	0 - array13246700 900650
Solanum1_ctrl_vs_drought_0_1	PSC_1_Ctrl	PSC_1_Drought	1 - array13246704 900650
Solanum1_drought_vs_ctrl_0_2	PSC_1_Drought	PSC_1_Ctrl	1 - array13246706 930630
Solanum3_ctrl_vs_drought_0_1	PSC_3_Ctrl	PSC_3_Drought	3 - array13246695 815700
Solanum3_drought_vs_ctrl_0_2	PSC_3_Drought	PSC_3_Ctrl	3 - array13246699 850630
Solanum5_ctrl_vs_drought_0_1	PSC_5_Ctrl	PSC_5_Drought	5 - array13246697 915665
Solanum5_drought_vs_ctrl_0_2	PSC_5_Drought	PSC_5_Ctrl	5 - array13246702 830640
Solanum7_ctrl_vs_drought_0_1	PSC_7_Ctrl	PSC_7_Drought	7 - array13246694 945735
Solanum7_drought_vs_ctrl_0_2	PSC_7_Drought	PSC_7_Ctrl	7 - array13246701 900690
Solanum10_ctrl_vs_drought_0_1	PSC_10_Ctrl	PSC_10_Drought	10 - array13246698 800735
Solanum10_drought_vs_ctrl_0_2	PSC_10_Drought	PSC_10_Ctrl	10 - array13246708 850625



**Figura 11. Micromatriz de *Solanum pascoense*.** Ejemplo de una salida del Programa GenePix Pro 5.0.0.49 (Axon) que muestra una de las hibridaciones realizadas con la especie PSC frente al estrés hídrico. Cada spot refleja mediante un color, la predominancia de intensidad de las señales, siendo en color rojo (Cy5), en verde (Cy3) o amarillo, cuando las intensidades de las señales de un spot son similares.

Las 12 imágenes fueron procesadas siguiendo un procedimiento de dos etapas muy laborioso. Se identificó inicialmente la localización de cada *spot* en la matriz en forma manual, ya que la opción automática no se ajustaba correctamente a las impresiones de las micromatrices de TIGR. Las 12 micromatrices fueron recorridas *spot* por *spot* a lo largo de los 48 bloques de estructura de cada matriz. Este proceso generó 12 archivos de extensión GPS, que fueron las referencias para la exportación de datos del programa GenePix. La segunda etapa de procesamiento genera archivos GPR. Este trabajo de exportación de resultados fue realizado por personal del laboratorio del Dr. Dominguez, quienes a partir de los archivos GPS enviados, generaron 12 archivos de extensión GPR que fueron el punto de partida para el análisis bioinformático posterior.

Los datos de micromatrices de ADNc generados en la plataforma de Salamanca con el especie PSC, se encuentran disponibles en el sitio web: <http://m4m.ugr.es/Flor/Tesis.html> o en el CD adjunto a este trabajo como Material Complementario.

### **Análisis bioinformático de los datos de micromatrices**

Durante las últimas décadas, se ha estado acumulando conocimiento, proveniente de diferentes áreas de la biología, en repositorios de datos digitales. El constante aumento del tamaño de estos repositorios hace casi imposible para el investigador, la extracción de información útil de los mismos. Por esta razón, se han desarrollado diversas técnicas de minería de datos para poder revelar información oculta en grandes colecciones de datos (Hand *et al*, 2001) que puedan ayudar a los usuarios en esta tarea.

Mediante la tecnología de micromatrices estudiamos la respuesta adoptada por tres especies silvestres de *Solanum* frente a sequía monitoreando la expresión de miles de genes en un momento dado. Las mayores limitaciones en la utilización de estas tecnologías giran alrededor del análisis de datos. Aun cuando en los últimos años se han hecho muchas contribuciones en este campo, no existen soluciones ampliamente aceptadas.

El estudio bioinformático implementado en este trabajo fue realizado en colaboración con el Grupo de Biología Computacional del Departamento de

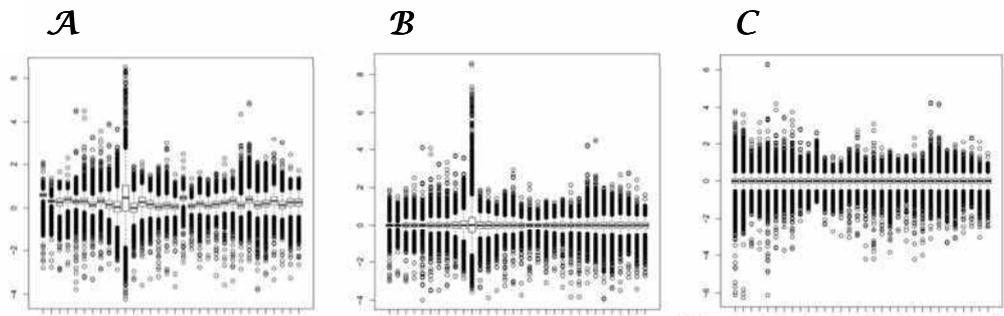
Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada con la Dra. Rocío Romero-Zaliz, con amplia experiencia en la implementación de algoritmos para el reconocimiento de perfiles de expresión génicos (Romero-Zaliz, 2005).

Asimismo los datos propios fueron procesados con diferentes herramientas de análisis bioinformático (basadas en R y BioConductor) junto con al asesoramiento de expertos que diseñaron muchas de las herramientas utilizadas, en el marco de un curso aplicado al análisis de micromatrices realizado en el Instituto DKFZ de Heidelberg, Alemania (Beissbarth, T. y Wolfgang, H. comunicación personal 2006).

En los experimentos realizados (Salamanca y TIGR) se incluyó en el diseño el procedimiento de inversión de tinte (del inglés *dye-swap*) alternando entre los tratamientos “control” y “estresado hídricamente”, a fin de compensar el sesgo que puede ocurrir en la hibridación de una matriz de dos colores (Rensink, W., comunicación personal, 2006, Yang *et al*, 2002, Quackenbush, 2002).

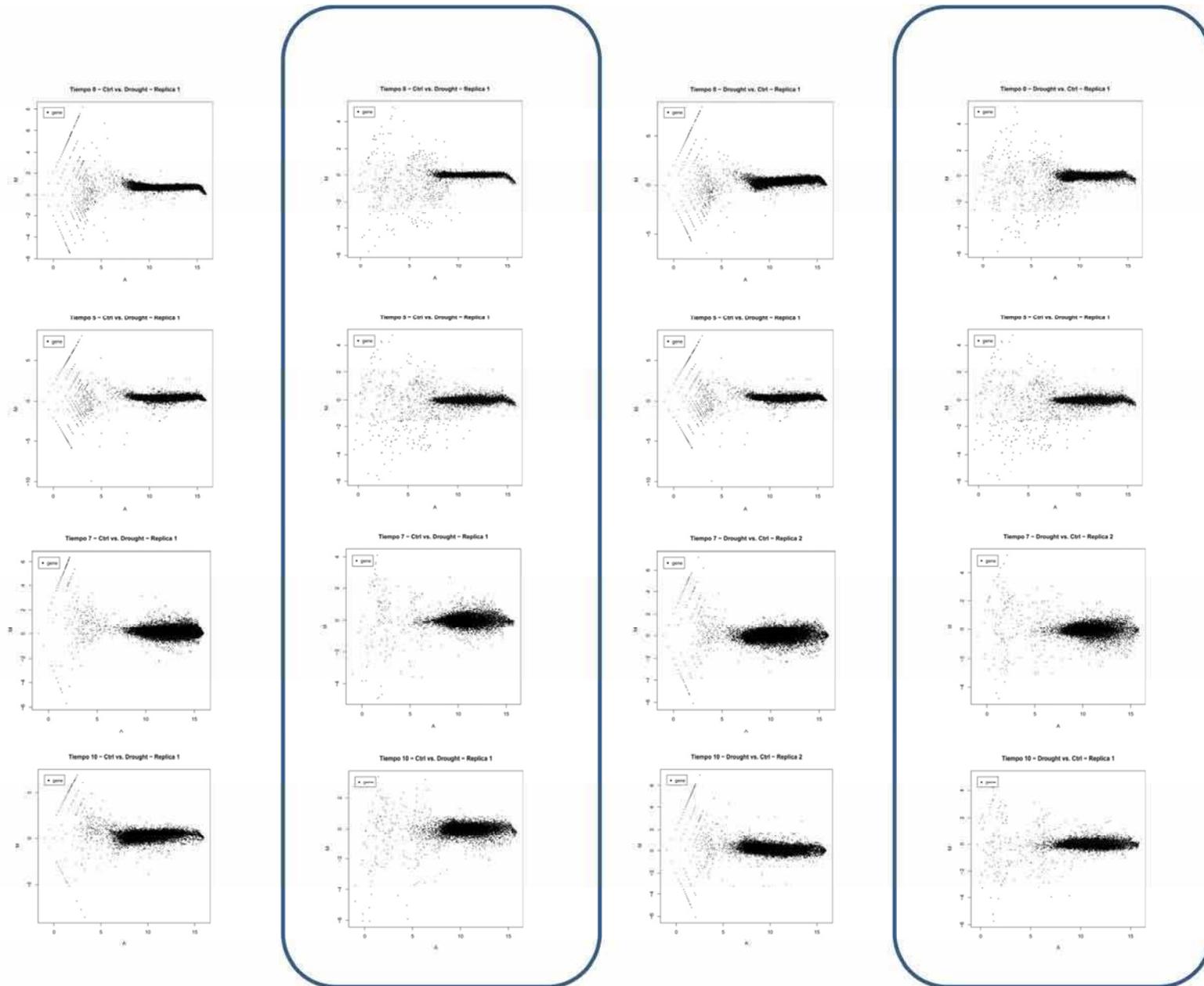
Antes de comenzar el análisis de los datos de expresión, se llevó a cabo un ajuste y normalización de la señal de fluorescencia de cada punto de la matriz para su posterior tratamiento estadístico. El objetivo de la normalización fue ajustar los efectos derivados de las variaciones propias de la técnica de micromatrices, como diferencias en las lecturas de los niveles de fluorescencia (dependiendo del escáner utilizado), variabilidad de la tinción fluorescente, etc. La literatura discute sobre la necesidad de realizar este procedimiento y sobre las estrategias para llevarlo a cabo, concluyendo en que no hay consenso respecto de las técnicas de normalización de datos de micromatrices, sino que deben evaluarse las diferentes opciones a fin de seleccionar la que mejor se adapten a las micromatrices utilizadas (Quackenbush, 2002; Smyth y Speed 2003). En los presentes experimentos se aplicó una normalización local, denominada *print-tip-lowess*, la más recomendada por los expertos para la normalización dentro de cada matriz en ensayos con micromatrices de dos colores (Beissbarth, T., comunicación personal 2006). Por su parte, para uniformar la distribución de las intensidades de los canales rojo y verde de todas las micromatrices se normalizaron con el método Aquantile (Yang *et al*, 2002, Smyth y Speed 2003, Gordon y Speed, 2003). El efecto de la

normalización aplicada a los datos de expresión de las micromatrices de las especies CPH y VNT se observa en la Figura 12.

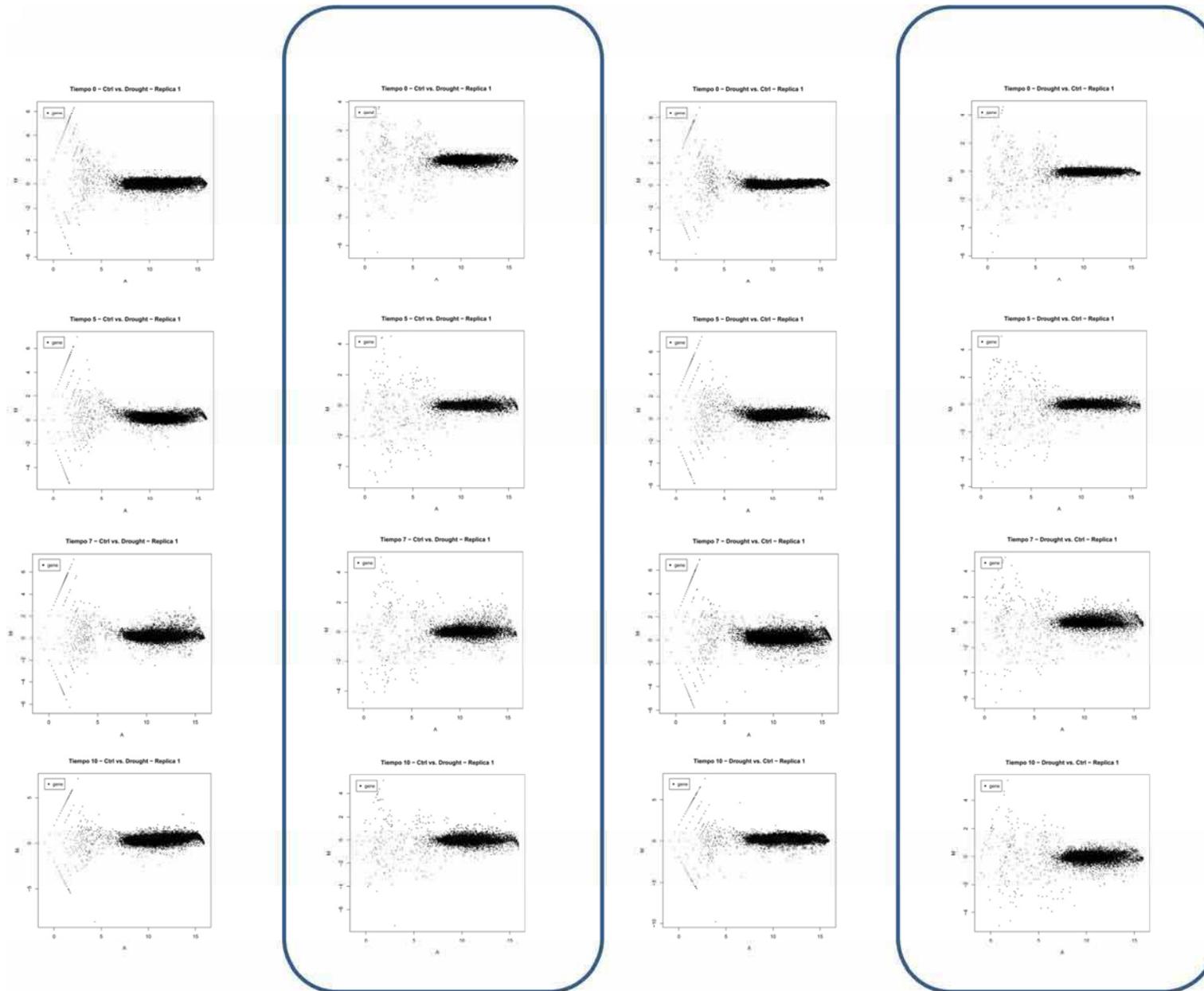


**Figura 12. Normalización de datos de expresión de micromatrices.** Efecto del proceso de normalización de los datos de expresión del experimento realizado con las especies CPH y VNT. A) micromatrices sin normalizar, B) normalización dentro de cada micromatriz con el método *print-tip-lowess* C) normalización dentro y entre micromatrices con el método de *print-tip-lowess* y *Aquantile* respectivamente de la librería “limma” de la plataforma Bioconductor.

A fin de visualizar el efecto de la normalización aplicada dentro de cada micromatriz, las gráficas de tipo MA plot son recursos útiles que se utilizaron en este trabajo con las tres especies en estudio. Las Figuras 13 y 14 muestran la movilización de toda la nube de puntos hacia el valor cero de M por efecto del método de *Print-tip-lowess* aplicado a cada micromatriz, diseñado para ajustar la intensidades y la tendencias espaciales, es decir, busca solucionar los errores sistemáticos que pueden surgir de la impresión de los spots en cada matriz, la falta de uniformidad de la hibridación



**Figura 13. Gráficos tipo MA del ensayo de micromatrices con la especie CPH.** Incluye hibridaciones con la combinación “control – Cy3” vs “estresado-Cy5” y su dye-swap en 4 momentos de estrés hídrico: T00, T05, T07 y T10 del tratamiento de sequía. Las columnas resaltadas en azul muestran el efecto de la normalización con el método *print-tip lowess* aplicado en cada micromatriz



**Figura 14. Gráficos tipo MA del ensayo de micromatrices con la especie VNT.** Incluye hibridaciones con la combinación “control – Cy3” vs “estresado-Cy5” y su dye-swap en 4 momentos de estrés hídrico: T00, T05, T07 y T10 del tratamiento de sequía. Las columnas resaltadas en azul muestran el efecto de la normalización con el método *print-tip lowess* aplicado en cada micromatriz

Se analizaron 1.427.712 datos de expresión correspondientes a la respuesta al estrés hídrico de 15.264 ADNcs depositados por duplicado sobre los 44 soportes de cristal de papa, que monitoreaban 10 días de evolución del estrés hídrico en las especies CPH, VNT y PSC. El análisis de las señales de hibridación confirmó la existencia de grupos de genes con expresión diferencial, inducidos y reprimidos en respuesta al estrés en el material estudiado. Se detallan a continuación los resultados obtenidos del análisis transcriptómico de las especies CPH y VNT.

### **Análisis exploratorio inicial de los datos de expresión génica**

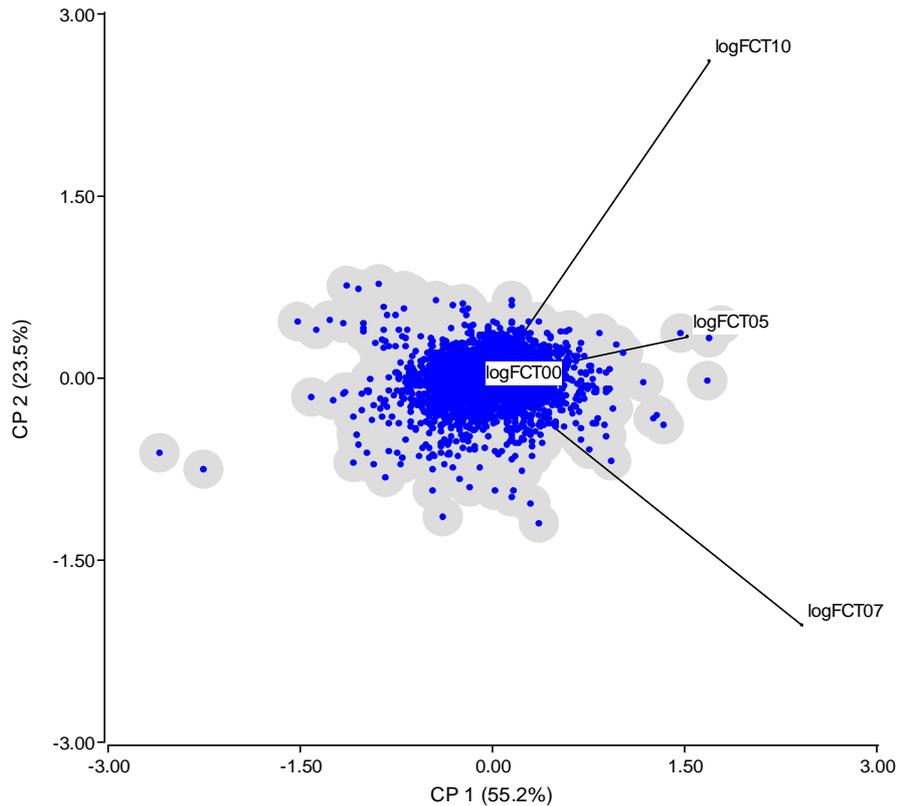
Un análisis preliminar de la diferencia en la expresión génica de las especies CPH y VNT frente a sequía mostró una coherencia de los datos de expresión y confirmó la calidad de los mismos. Desde una perspectiva multivariada se ordenaron los genes expresados diferencialmente frente a la sequía en un plano de ordenación generado por un análisis de componentes principales aplicado a la matriz de expresiones génicas. Esta aproximación tiene en cuenta la información de todos los genes simultáneamente en los tiempos evaluados a fin de ubicarlos en el plano.

El análisis de componentes principales se basó en valores de logFC, considerando 4 momentos del estrés: T00, 05, 07 y 10 correspondientes a los días de iniciada la sequía para cada especie.

La representación de los genes diferenciales en los planos definidos por los componentes principales 1, 2 y 2, 3 para la especie CPH se muestran en las Figuras 15, A y B.

En la representación espacial, los genes aparecen como puntos alrededor del origen o punto cero. Cuanto más distantes están del origen, más diferenciados están en uno o varios tiempos de evaluación. Los genes representados en la periferia de la nube de puntos representan los genes que tienen los valores de logFC más extremos.

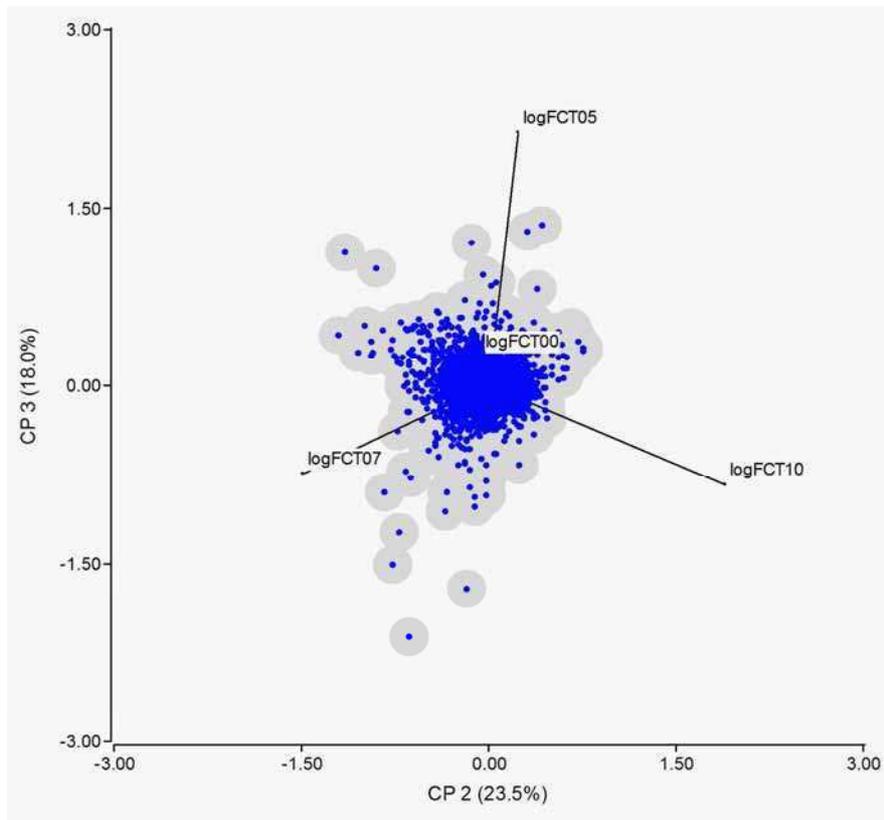
La proporción de variabilidad total retenida por estos tres ejes es del 97%. En gris aparecen todos los genes evaluados y en azul todos aquellos que resultaron significativos para alguno de los tiempos considerados.



**Figura 15: Análisis exploratorio por componentes principales para la especie CPH.** Espacio generado por los componentes principales 1 y 2 resultantes del análisis de componentes principales de la matriz de logFC de la condición estrés hídrico respecto del testigo no estresado en la especie CPH. En azul se observan los genes que presentan tener un p-valor significativo en algunas de las cuatro evaluaciones en el tiempo.

Para interpretar la relación entre la ubicación de los genes en el espacio de los componentes principales y su relación con su expresión en los distintos tiempos de evaluación hay que tener en cuenta que cuanto mayor es la proyección de estos segmentos en los ejes que representan a los componentes principales, mayor es la importancia de ese momento en la explicación de la variabilidad. Así puede observarse que la evaluación en el tiempo T00 (momento inicial del experimento) tiene un peso casi nulo en la conformación de los componentes principales 1, 2 y 3 (Figura 15 A y B). Esto es coherente con el diseño del experimento ya que no debería haber cambios en la expresión génica entre las plantas estresadas hídricamente y sus controles al inicio del experimento. Contrastando con esta situación, la Figura 15 A muestra

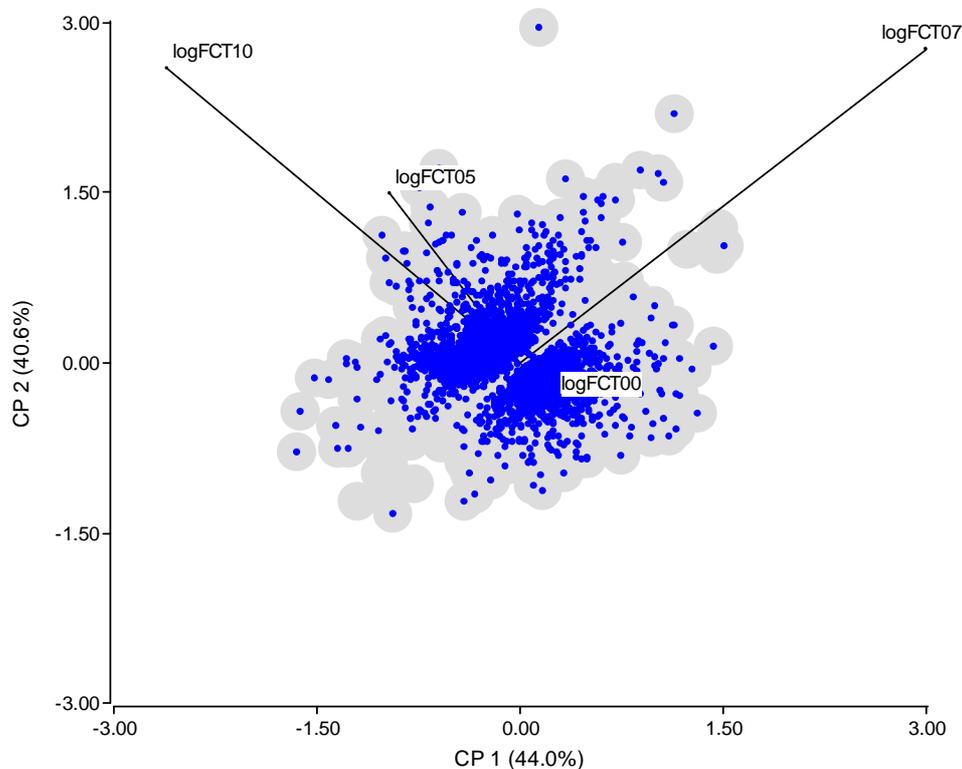
que las evaluaciones en los tiempos T07, T10 y T5 tienen una proyección similar en el Eje 1. T07 y T10 se proyectan con similar importancia sobre el Eje 2, aunque en direcciones distintas, mientras que T05 no tiene proyección sobre el Eje 2. Como el Eje 1, retiene el 55% de la variabilidad total del logFC, las instancias de tiempo T05, T07 y T10 son las que mayor proyección tienen sobre este eje y son las responsables de la mayor variabilidad observada en el logFC a lo largo de todo el experimento.



**Figura 15B: Análisis exploratorio por componentes principales para la especie CPH.** Espacio generado por los componentes principales 2 y 3 resultantes del análisis de componentes principales de la matriz de logFC de la condición estrés hídrico respecto del testigo no estresado en la especie CPH. En azul se observan los genes que presentan tener un p-valor significativo en algunas de las cuatro evaluaciones en el tiempo.

La Figura 15 B, muestra que T05 tiene más proyección sobre el Eje 3. Los genes ubicados en la parte superior del Eje 3 se sobre-expresan en T05 y los que están en la parte inferior se sub-expresan.

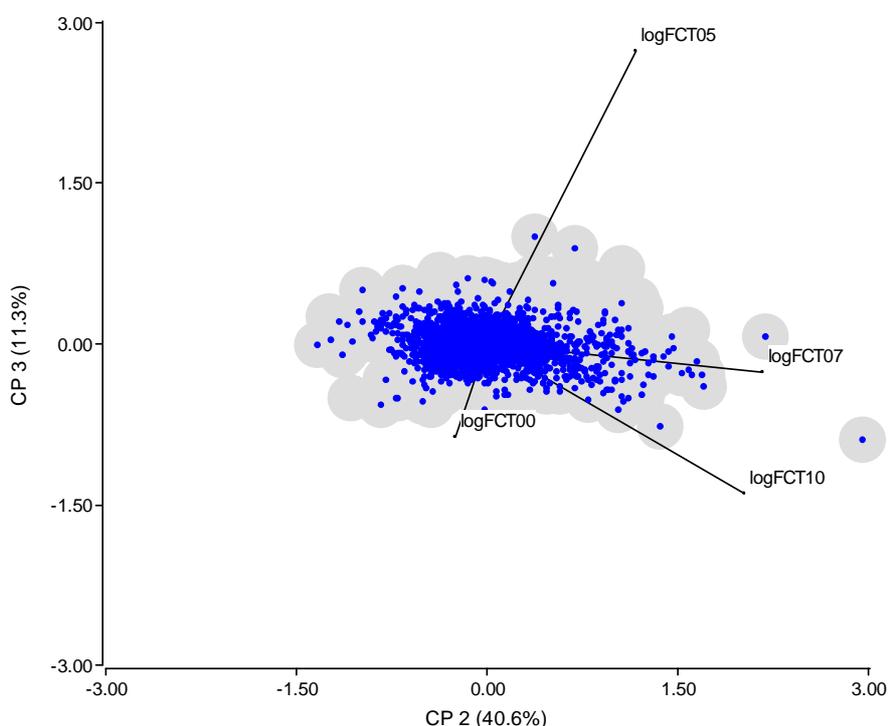
Se siguió el mismo análisis para la especie VNT y el plano de ordenación generado mediante análisis de componentes principales aplicado a la matriz de expresiones (logFC) se muestra en la Figura 16 A y B.



**Figura 16A: Análisis exploratorio por componentes principales para la especie VNT** Espacio generado por los componentes principales 1 y 2 resultantes del análisis de componentes principales de la matriz de logFC de la condición estrés hídrico respecto del testigo no estresado en la especie VNT. En azul se observan los genes presentan tener un p-valor significativo (0.05) en algunas de las cuatro evaluaciones en el tiempo.

El porcentaje total de variabilidad explicada por los tres primeros componentes principales fue del 96%. En gris aparecen todos los genes evaluados y en azul todos los que resultaron significativos para alguno de los tiempos considerados. Analizando la proyección de cada tiempo en los ejes, el tiempo T00, (momento inicial del experimento) tiene un peso muy pequeño en la conformación de los componentes principales 1, 2 y 3 (Figuras 16 A y B), que al igual que con la especie CPH, muestran ser coherente con el diseño del experimento, en donde

no debería haber cambios en la expresión génica los grupos estresado y control al inicio del experimento.

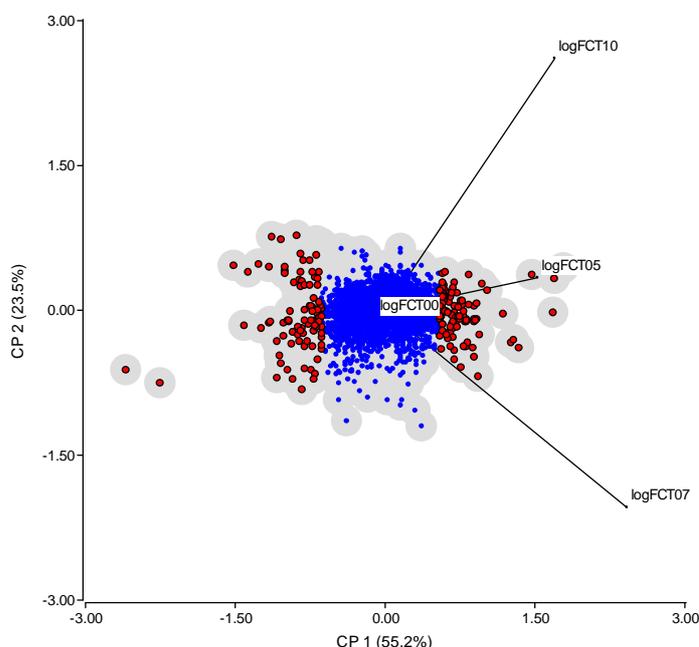


**Figura 16B: Análisis exploratorio por componentes principales para la especie VNT** Espacio generado por las componentes principales 2 y 3 resultantes del análisis de componentes principales de la matriz de logFC de la condición estrés hídrico respecto del testigo no estresado en la especie VNT. En azul se observan los genes presentan tener un p-valor significativo (0.05) en algunas de las cuatro evaluaciones en el tiempo.

Los otros tiempos, por su parte, exhiben proyecciones sobre el Eje1 (Figura 16 A), que aunque similares, presentan direcciones opuestas. T07 y T10 se proyectan con similar importancia sobre el Eje 2, con la misma orientación, mientras que T05 tiene una proyección moderada tanto en el Eje 1 como en el Eje 2. Como los dos primeros ejes retienen el 84% de la variabilidad total del logFC (44% y 40% respectivamente), las instancias de tiempo que mayor proyección tienen sobre este eje son las responsables de la mayor variabilidad observada en el logFC a lo largo de todo el experimento. Los genes que se ubican en la periferia de la nube de puntos son los genes más diferenciados en el experimento. Particularmente destacable son los genes que se ubican en los

extremos del Eje 1 ya que son los que tienen un comportamiento antagónico entre los tiempos T07 y T10. Por otra parte los genes que se ubican en los extremos del Eje 2 son los que se sobre-expresan (extremo positivo) o sub-expresan (extremo negativo) en los tiempos T07 y T10, que parecen ser los momentos en los que se producen los mayores cambios en la expresión.

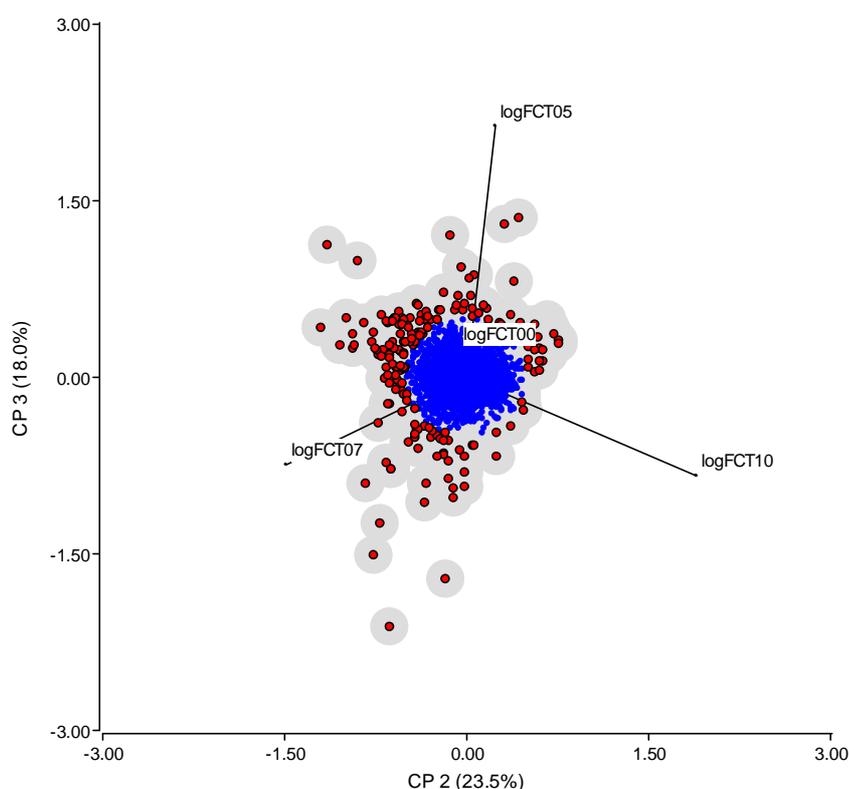
A partir del análisis realizado de componentes principales, se seleccionaron dos subconjuntos de genes dentro de aquellos significativos (azules en las Figuras 15 y 16 de las especies CPH y VNT respectivamente). Un subconjunto se derivará del ordenamiento generado por el componente principal 1. Para ello se seleccionaron aquellos genes cuya posición en el componente principal esté por debajo del percentil 1 o por encima del percentil 99, es decir estamos reteniendo el 2% de los genes más extremos proyectados sobre el primer componente principal. Estos son los genes que responden en general al estrés. Los ubicados por encima del percentil 99 son los que sobre-expresan y los que están por debajo del percentil 1 son los que se sub-expresan por efecto de la sequía. Los genes en este grupo para la especie CPH aparecen en color rojo en la Figura 17 A.



**Figura 17A. Análisis exploratorio de datos para la especie CPH.** Espacio generado por los componentes principales 1 y 2 resultantes del análisis de componentes principales de la matriz de logFC de la condición estrés hídrico respecto del testigo no estresado en la especie CPH.

Se resaltan los genes que están por debajo y por encima de los percentiles 1 y 99 de la distribución de puntaje en el componente principal 1 y tienen un p-valor ajustado (0.05) en algunas de las cuatro evaluaciones en el tiempo.

El segundo subconjunto al que hacíamos referencia, es el conformado por aquellos genes que presentan un comportamiento diferencial. La conformación de este grupo se basó en la selección de genes cuya distancia al origen, en el plano generado por los componentes principales 2 y 3, se ubica por encima del percentil 99 de la distribución de las distancias al origen y cumplen con la condición de ser significativos (azules en las Figuras 15B y 16B en las especies CPH y VNT respectivamente). Estos genes aparecen identificados en color rojo en la Figura 17B para CPH.



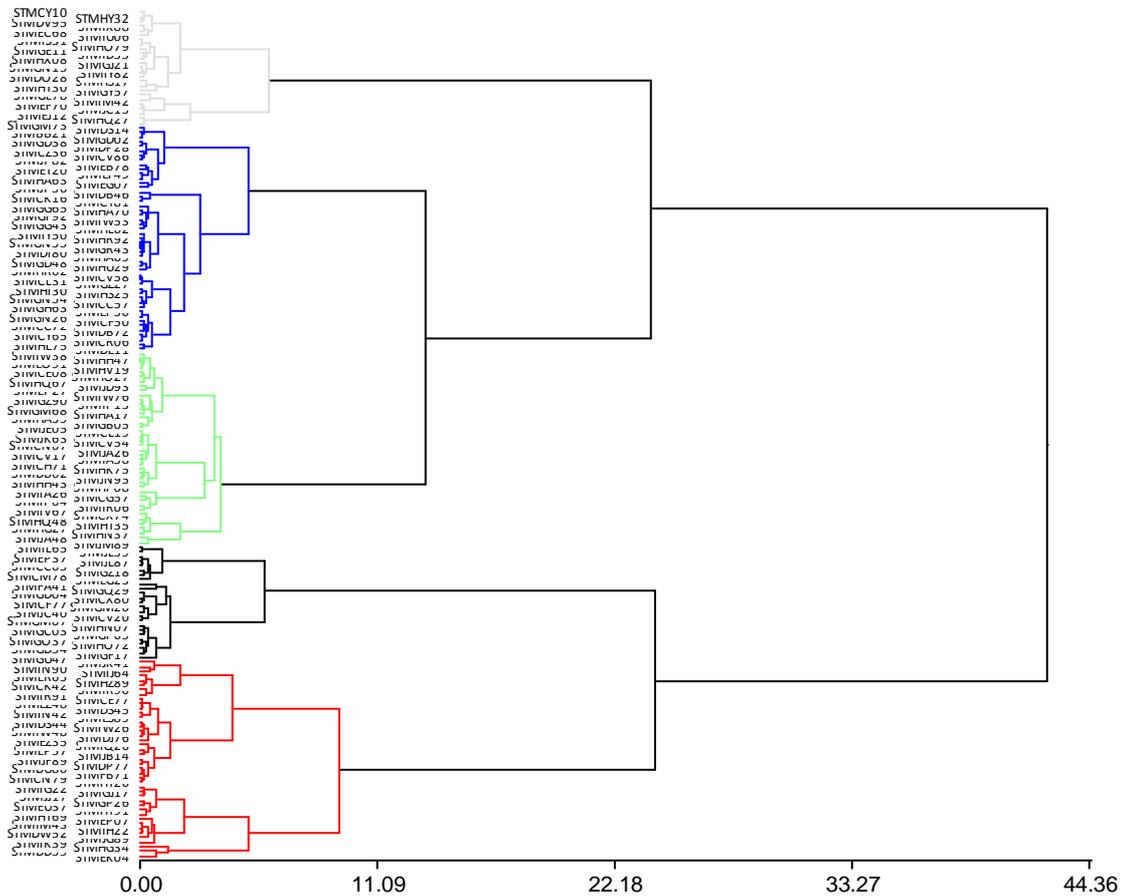
**Figura 17B. Análisis exploratorio de datos para la especie CPH.** Espacio generado por los componentes principales 2 y 3 resultantes del análisis de componentes principales de la matriz de logFC de la condición estrés hídrico respecto del testigo no estresado en la especie CPH. Se resaltan los genes que están por debajo y por encima de los percentiles 1 y 99 de la distribución de puntaje y tienen un p-valor ajustado (0.05) en algunas de las cuatro evaluaciones en el tiempo.

Los genes identificados en rojo en la Figura 17A se listan en la Tabla 20. Éstos se dividen en sobre y sub expresados en la condición “estrés hídrico” respecto del control. Se encuentran ordenados de mayor a menor por el valor absoluto de su puntaje en el componente principal 1. Por lo tanto los primeros en la lista de sobre y sub expresados son los genes más diferenciados.

**Tabla 20. Listado de genes correspondientes al grupo de que sobre y sub-expresa por efecto del estrés, independientemente del momento de evaluación.**

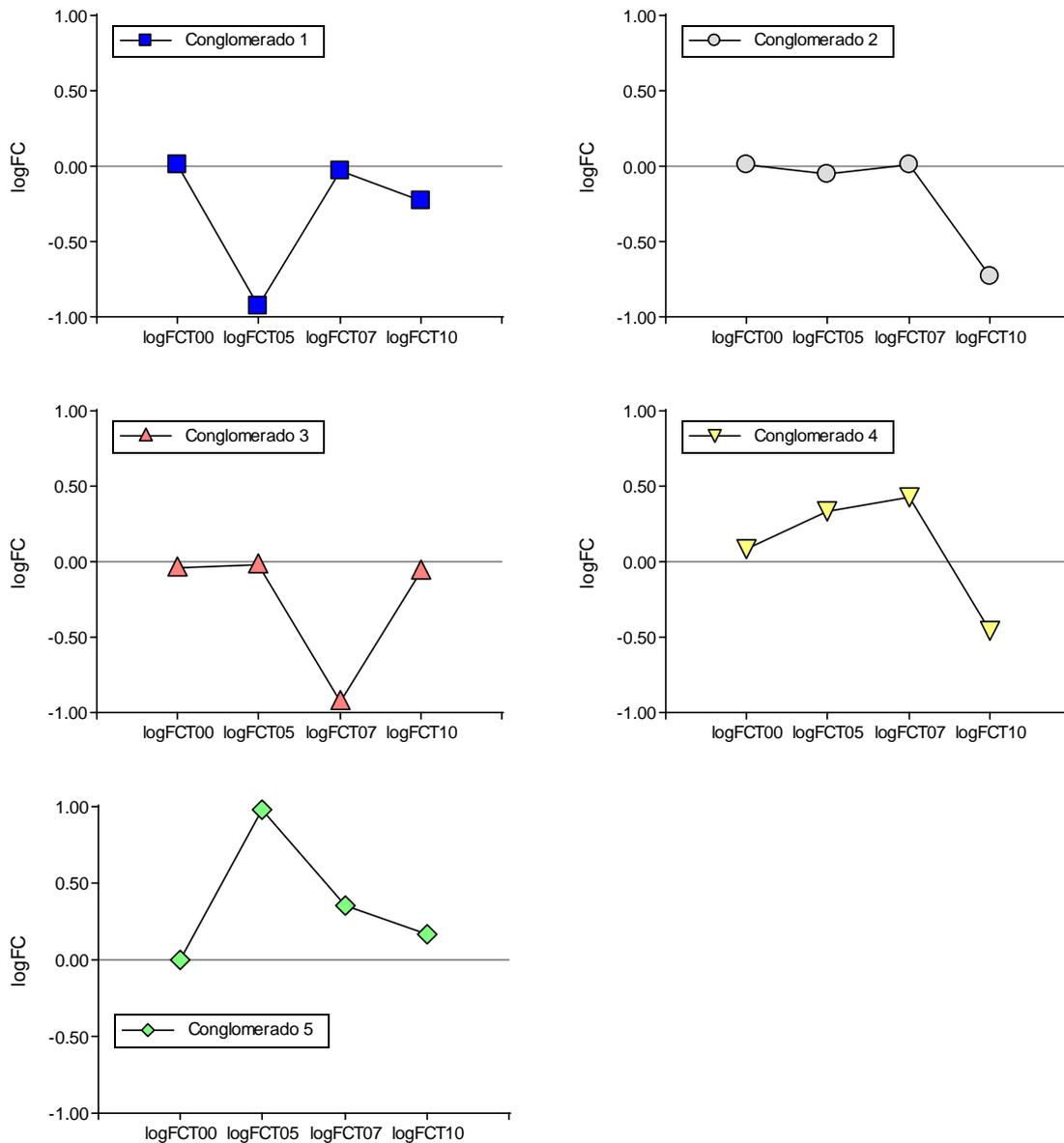
Sobre	Sub	Sobre	Sub	Sobre	Sub	Sobre	Sub
STMGM73	STMEK04	STMII64	STMJC40	STMCD31	STMJA79	STMGE11	STMGD50
STMHQ27	STMDD59	STMGN69	STMJC16	STMHX89	STMGW47	STMEN94	STMGB59
STMGE78	STMGQ29	STMII80	STMCP35	STMHX51	STMDD06	STMGJ16	
STMIM42	STMHG34	STMGG75	STM DG27	STMDO18	STMGR65	STMIS63	
STMEJ12	STMEY89	STMIR06	STMHN07	STMEO35	STMEK70	STMCO76	
STMJC15	STMFA41	STMJL39	STMGP65	STMGU06	STMCM83	STM CU13	
STMEF70	STMDF40	STMIS31	STMGM20	STMHY63	STMJP56	STMGZ28	
STMHT30	STMGD54	STMHO79	STMCG17	STMJA58	STM CV68	STMII05	
STMHO80	STMIQ06	STMHV73	STMCE17	STMJD19	STM DP57	STMIQ18	
STMGB34	STMEA11	STMIJ18	STMJG89	STMHY04	STMGE07	STMHX38	
STMGR44	STMGM07	STMIJ22	STMGQ44	STMHS26	STMIM43	STMIY82	
STMIV67	STMHY91	STMGG17	STMJA23	STMIC42	STMCR59	STMGA60	
STMGQ20	STMDB46	STMJM02	STMGD04	STMHA76	STMCF77		
STMHS17	STM EP07	STMGP79	STMCP60	STMIJ15	STMJO58		
STMHV95	STMCK16	STMHX42	STMIH22	STMEW74	STMGW15		
STMHIO2	STM CV20	STMGO56	STMDB57	STMHR07	STMJN02		
STMIA28	STMEA15	STMIP64	STMCC58	STMES47	STMEG07		
STMHP45	STM CX90	STMHX08	STMEA22	STMHW51	STMCM93		
STMIA26	STMGO37	STMGC64	STMEF73	STMGM67	STMJG06		
STMCG57	STMHO72	STMHQ21	STMER46	STMHP21	STMIQ26		
STM DI37	STMGF17	STMHY51	STMFB43	STMJF35	STMEC08		
STMGD72	STMEA27	STMIR11	STMER92	STMIV51	STMGU47		
STMCK45	STM CY81	STMIH67	STMFB12	STMHY84	STMER56		
STMBB19	STM CX38	STMGT74	STM CX80	STMIX48	STMHZ95		
STMJB50	STMEU37	STMJP63	STMJJ17	STMDS28	STMEU18		
STMHL92	STMGP26	STMJO47	STM DW52	STMHR28	STMEB61		
STMHP72	STMHT69	STMID35	STMGG84	STMHU70	STMEH63		
STMGJ39	STM IK39	STMGW72	STMER63	STMGJ21	STMDB40		
STMBB31	STMGG78	STMJJ04	STMEK44	STMHP05	STMGW82		
STMGF62	STMGC03	STMET16	STMIG22	STMJK41	STMHO84		

Para estudiar el comportamiento de los genes identificados en la Figura 15B, sus perfiles temporales de logFC fueron sometidos al análisis de conglomerados. Para ello se utilizó el método de Ward aplicado a la matriz de distancias Euclídeas entre los perfiles temporales de logFC de pares de genes. La Figura 18 muestra una representación en dendrograma de los resultados de este análisis. Una de las ventajas del método de Ward es que en el dendrograma resultante es fácil identificar grupos relativamente bien diferenciados. Se consideró que existen cinco grupos de genes.



**Figura 18. Dendrograma resultante de la aplicación de un análisis de conglomerado jerárquico** aplicando el algoritmo de Ward a la matriz de distancias Euclídeas entre perfiles de logFC en la especie CPH. Se identifican con colores 5 grupos de genes.

Los perfiles de los grupos de genes identificados por análisis de conglomerados se presentan en la Figura 19.



**Figura 19.** Perfiles temporales de los logFC para cada uno de los cinco grupos de genes identificados por análisis de conglomerados en CPH.

La Tabla 21 presenta el listado de los genes cuyos perfiles medios se observan en la Figura 19. Se encuentran ordenados en orden decreciente de acuerdo a la distancia al origen el plano de ordenación generado por los componentes principales 2 y 3.

**Tabla 21. Listado de genes por conglomerado de perfiles observados en la Figura 14 ordenados de en forma decreciente acuerdo a la distancia al origen el plano de ordenación generado por los componentes principales 2 y 3**

Cong 1	Cong 2	Cong 3	Cong 4	Cong 5	
STMEK04	STMDE11	STMEG07	STMGU47	STMJM89	STMGM73
STMHG34	STMHR62	STMHA70	STMGM07	STMJA48	STMHQ27
STMDD59	STMHL75	STMCL31	STMJC40	STMHN37	STMGY57
STMIK39	STMJP56	STMCV58	STMCV20	STMHT35	STMHT30
STMJG89	STMHU29	STMDS14	STMEG23	STMHG27	STMHS17
STMHT69	STMCR06	STMEB78	STMGM20	STMHQ48	STMIU06
STMEU37	STMGD48		STMCM78	STMCX74	STMDO28
STMDDW52	STMCY81		STMJL59	STMIV67	STMEJ12
STMIH22	STMCK16		STMIL65	STMHP08	STMIY82
STMIM43	STMHL62		STMGF17	STMJE05	STMHX08
STMHY91	STMGH63		STMGQ29	STMIR06	STMIM42
STMJJ17	STMCY65		STMGZ18	STMHA39	STMDDV95
STMGP26	STMDB46		STMCF77	STMEP27	STMGN15
STMGJ17	STMHA69		STMCC65	STMGB05	STMHO79
STMIG22	STMDDI80		STMGD54	STMHH43	STMGJ21
STMHY26	STMGR43		STMGD04	STMCH71	STMJC15
STMEP07	STMGN35		STMJL87	STMHQ67	STMID35
STMIR91	STMHA63		STMHO72	STMHA17	STMEF70
STMFB71	STMGG43		STMFA41	STMJN95	STMCY10
STMIJ64	STMGN26		STMGP65	STMIW76	STMIS31
STMCN79	STMCC57		STMCX80	STMHK75	STMGE78
STMDDU80	STMJP82		STMGC03	STMHV19	STMEC68
STMJB14	STMEF49		STMEP37	STMGM68	STMGE11
STMEZ35	STMHR92		STMGO37	STMIA38	STMHY32
STMCK42	STMCZ36		STMHN07	STMHO27	STMIX88
STMES89	STMHS25			STMDB62	
STMDDJ76	STMDB72			STMIP13	
STMIW26	STMHI30			STMGZ90	
STMIN42	STMIY50			STMJA26	

Cong 1	Cong 2	Cong 3	Cong 4	Cong 5
STMDS44	STMGN54		STMIP64	
STMHZ89	STMIW53		STMEN67	
STMEP57	STMDDP28		STMJD93	
STMDS45	STMGG65		STMCE08	
STMCE77	STMET20		STMIA26	
STMDDP77	STMEP30		STMEO91	
STMEZ48	STMGG27		STMHH47	
STMJK41	STMCF50		STMCV17	
STMIW48	STMGD02		STMCG57	
STMIQ26	STMCV86		STMJK63	
STMER65	STMGD38		STMCV54	
STMIR90	STMCC72		STMCL19	
STMIN90	STMGF92		STMIW38	
STMJF89	STMBB21			

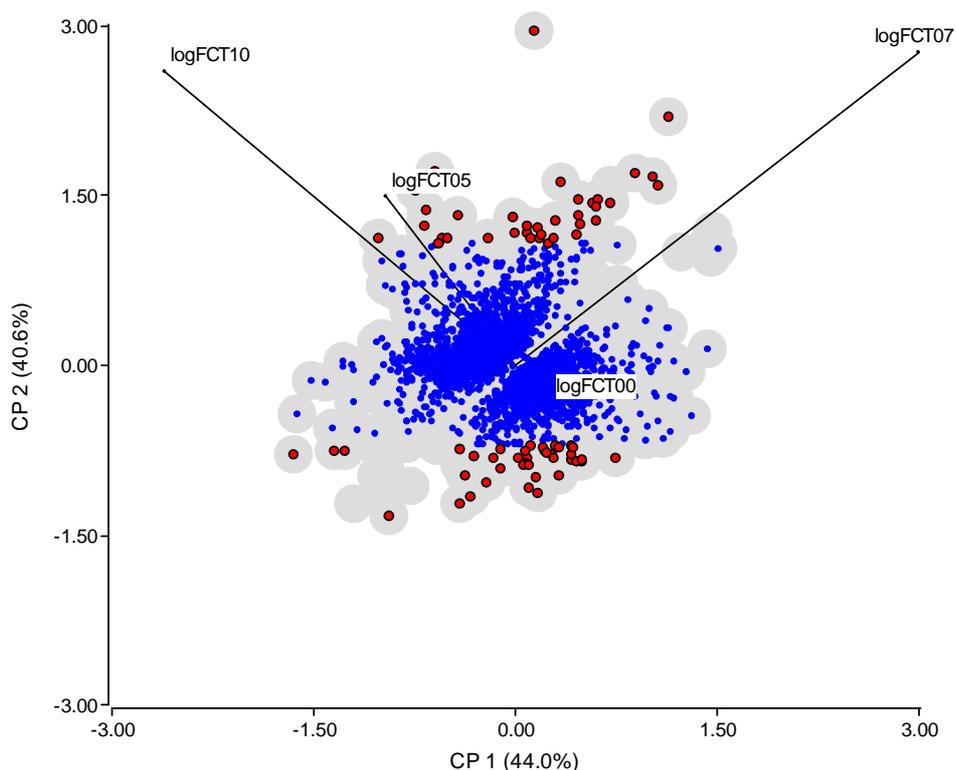
Finalmente la Tabla 22 contiene la lista de genes que simultáneamente están en la lista de la Tabla 19 y de la Tabla 20.

**Tabla 22: Listado de genes que aparecen simultáneamente en las Tablas 19 y 20.**

STMEK04	STMJP56	STMID35
STMHG34	STMJC40	STMEF70
STMDD59	STMCV20	STMGD54
STMIK39	STMIG22	STMIS31
STMGM73	STMCY81	STMGE78
STMHQ27	STMCK16	STMGD04
STMJG89	STMIR06	STMJK41
STMHT69	STMEP07	STMHO72
STMEU37	STMEJ12	STMFA41
STMDDW52	STMDB46	STMIP64
STMIH22	STMIY82	STMIQ26
STMIM43	STMHX08	STMGP65
STMHY91	STMGM20	STMGE11
STMJJ17	STMGF17	STMIA26

STMHT30	STMIM42	STMCX80
STMGP26	STMGQ29	STMEG07
STMGU47	STMHO79	STMGC03
STMIV67	STMCF77	STMCG57
STMHS17	STMGJ21	STMGO37
STMGM07	STMJC15	STMHN07

La estrategia de conformación de subconjuntos según los componentes principales 1, 2 y 3 para el ordenamiento de genes de VNT, siguió el mismo esquema utilizado en CPH. Un subconjunto proviene del ordenamiento concebido por el componente principal 1. Se seleccionó el 2% de los genes más extremos proyectados sobre el primer componente principal. Estos son los genes que responden antagónicamente al estrés a final del periodo del experimento. Los genes en este grupo aparecen en color rojo en la Figura 20.



**Figura 20. Analisis exploratorio de datos para la especie VNT.** Espacio generado por los componentes principales 1 y 2 resultantes del análisis de componentes principales de la matriz de log<sub>2</sub> (fold-changes) de la condición estrés hídrico respecto del testigo no estresado en la especie VNT. Se resaltan los genes que están por debajo y por encima de los percentiles 1 y

99 de la distribución de puntaje en el componente principal 1 y tienen un p-valor ajustado (0.05) en algunas de las cuatro evaluaciones en el tiempo.

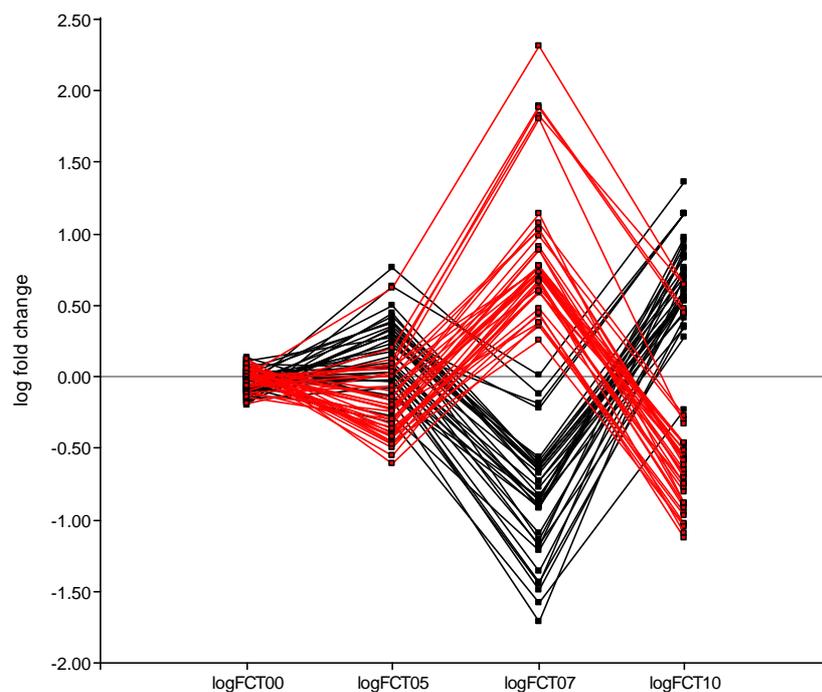
Un segundo grupo de genes que tiene comportamiento similar en T07 y T10 se conformó seleccionando aquellos genes que están por debajo y por encima de los percentiles 1 y 99 de la distribución del componente principal 2. Los genes seleccionados se listan en la Tabla 23. Estos se dividen en dos: los sobre-expresados en la condición estresante respecto del control en T07 y T10 respectivamente. Se encuentran ordenados de mayor a menor por el valor absoluto de su puntaje en el componente principal 1. Por lo tanto los primeros en la lista de cada grupo son los más diferenciados.

**Tabla 23. Listado de genes que se sobre expresan en T07 y (sub expresan en T10) y aquellos que se sobre expresan en T10 y (sub expresan en T07).**

<b>Sobre expresan en T10 sub expresan en T07</b>		<b>Sobre expresan en T07 sub expresan en T10</b>	
<b>Clon</b>	<b>Anotación StGI</b>	<b>Clon</b>	<b>Anotación StGI</b>
STMEF42	Peroxidase (EC 1.11.1.7)	STMCG17	Putative proline-rich protein (Fragment)
STMIP09	Copper amine oxidase precursor (EC 1.4.3.6)	STMIC67	Extensin like protein
STMGI29	2-oxoglutarate-dependent dioxygenase	STMJA27	Auxin-binding protein ABP19a precursor
STMIH11	Membrane channel protein (Fragment)	STMJM27	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein contains
STMCL06	2-oxoglutarate-dependent dioxygenase	STMFB55	Acyl carrier protein
STMGM17	Gamma tubulin (Fragment)	STMCX35	24K germin like protein precursor
STMCQ40	Linalool synthase-like protein (Fragment)	STMIZ16	Pathogenesis-related protein PR-1 precursor
STMGO18	2-oxoglutarate-dependent dioxygenase	STMDD41	protodermal factor 1 (PDF1) identical to protodermal factor 1 [Arabidopsis thaliana]
STMHT30	homeobox-leucine zipper protein 7 (HB-7) / HD-ZIP transcription factor 7 identical to	STMCM83	endonuclease/exonuclease/phosphatase family protein contains Pfam profile PF03372:
STMIN04	Pathogenesis related protein isoform b1 precursor	STMEN59	protodermal factor 1 (PDF1) identical to protodermal factor 1 [Arabidopsis thaliana]
STMFB48	Peroxidase precursor (EC 1.11.1.7)	STMIQ07	protodermal factor 1 (PDF1) identical to protodermal factor 1 [Arabidopsis thaliana]
STMFB69	Terpene synthase	STMCU43	ribosomal protein L34 family protein similar to plastid ribosomal protein L34
STMDJ96	expressed protein	STMDV46	chloroplast nucleoid DNA-binding protein-related contains Pfam profile PF00026:
STMEF70	Putative ankyrin protein	STMIN90	Malate synthase glyoxysomal (EC 2.3.3.9 EC 2.3.3.9 EC 4.1.3.2)
STMIC96	PilS sensor protein precursor (EC 2.7.3.-)	STMIX44	beta-1 2-N-acetylglucosaminyltransferase II identical to beta-1
STMIM87	Linalool synthase-like protein (Fragment)	STMEG32	Gibberellin 7-oxidase (EC 1.14.11.-)
STMCF39	2-oxoglutarate-dependent dioxygenase	STMCG72	plastid developmental protein DAG putative similar to DAG protein chloroplast
STMCY65	1-deoxyxylulose 5-phosphate synthase	STMHL10	Null
STMGN13	Putative cytochrome P450	STMIW10	Putative phytosulfokine peptide precursor
STMEH84	UDP-glucose glucosyltransferase (EC 2.4.1.-)	STMHY79	CR9 protein
STMGQ20	Putative non-specific lipid transfer protein StnsLTP	STMDH13	NADPH:protochlorophyllide oxidoreductase
STMCH16	Ovule/fiber cell elongation protein Ghfe1	STMHY70	expressed protein
STMGG59	expressed protein contains Pfam profile PF04396: Protein of unknown function DUF537	STMHY51	hypothetical protein
STMDH95	null	STMCL39	H-Protein (EC 2.1.2.10)
STMGB09	Putative cytochrome P450	STMDP79	expressed protein low similarity to early flowering protein 1 [Asparagus officinalis]
STMFB77	retrotransposon protein putative unclassified	STMIR79	Photosystem I reaction center subunit X psaK
STMCV17	Spermidine synthase (EC 2.5.1.16)	STMZ35	Photosystem I reaction center subunit IV A chloroplast precursor (PSI-E A)
STMFB13	Terpene synthase	STMJB96	Phosphate ABC transporter permease protein (PSTC)
STMGG42	Ovule/fiber cell elongation protein Ghfe1	STMIC48	Nonspecific lipid-transfer protein 2 precursor (LTP 2)
STMGY78	probable potassium channel protein SKT3 - potato (fragment) {Solanum tuberosum;}	STMJK18	chlorophyll A-B binding protein chloroplast (LHCB6) nearly identical to Lhcb6 protein
STMCZ51	hypothetical protein	STMJ148	immunophilin putative / FKBP-type peptidyl-prolyl cis-trans isomerase putative
STMDZ44	probable polygalacturonase 1 beta chain F1707.9 - Arabidopsis thaliana (EC 3.2.1.15)	STMGA23	plastocyanin-like domain-containing protein contains similarity to SP Q02917 Early
STMGR44	Putative non-specific lipid transfer protein StnsLTP	STMCK85	GDSL-motif lipase/hydrolase family protein similar to family II lipase EXL3

Sobre expresan en T10 sub expresan en T07		Sobre expresan en T07 sub expresan en T10	
Clon	Anotación StGI	Clon	Anotación StGI
STMCL31	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase	STMDU28	Thylakoid lumenal 16.5 kDa protein chloroplast
STMIQ86	Peroxiredoxin precursor	STMCJ47	(EC 4.1.1.65)
STMGW14	zinc finger (CCCH-type) family protein contains Pfam domain PF00642: Zinc finger		
STMDF17	(EC 5.3.4.1 EC 5.3.4.1)		
STMCK41	BYJ15 (Fragment)		

Para visualizar el comportamiento de los genes listados en la Tabla 23, los perfiles temporales de logFC se graficaron en la Figura 22. Hay claramente un patrón contrastante de perfiles de expresión, comportamiento que ya se insinúa en T05. Los genes cuyos logFC decaen en el T07, se presentan preponderantemente por encima del promedio en T05 y viceversa.



**Figura 21. Perfiles de logFC para los genes listados en la Tabla 22-.**

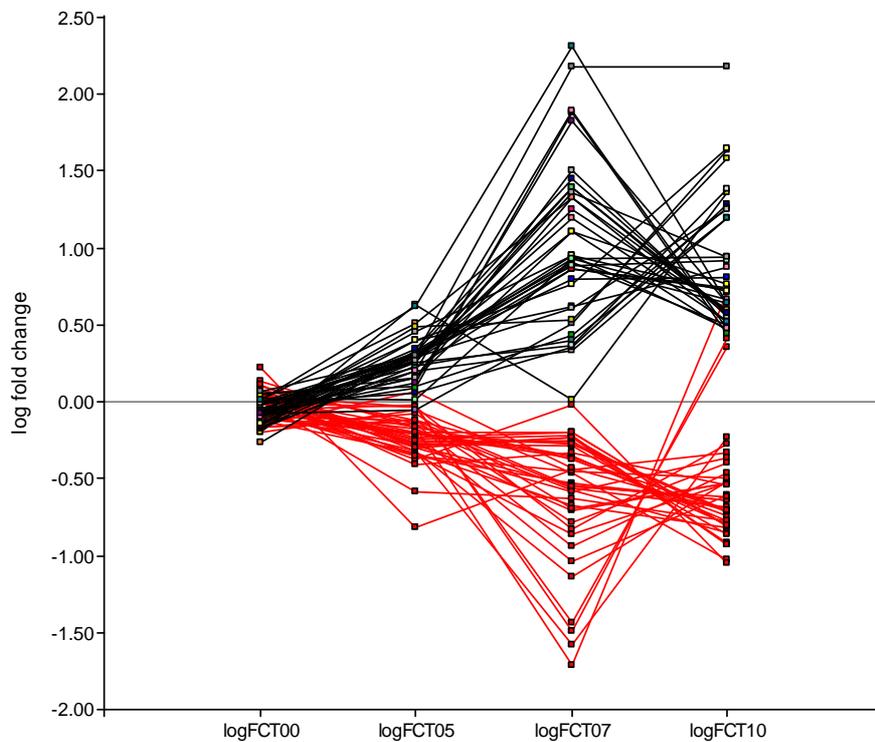
La Tabla 24 presenta el listado de los genes cuyos perfiles medios se obtuvieron de espacio generado por los componentes 2 y 3 para la especie VNT. Se encuentran ordenados de mayor a menor por el valor absoluto de su puntaje en el componente principal 2. Por lo tanto los primeros en la lista de cada grupo son los más diferenciados.

**Tabla 24.** Listado de genes que se sobre y sub expresan en T07 y en T10 y simultáneamente en la especie VNT

Sobre expresados en T07 y T10		Sub expresados en T07 y T10	
Clon	Anotación StGI	Clon	Anotación StGI
STMGM73	Hypothetical protein B1027A11.138	STMDP79	expressed protein low similarity to early flowering protein 1 [Asparagus officinalis]
STMHT30	homeobox-leucine zipper protein 7 (HB-7) / HD-ZIP transcription factor 7 identical to	STMJN02	F1E22.18
STMIU09	expressed protein contains Pfam profile PF04520: Protein of unknown function DUF584	STMEP15	fringe-related protein similarity to predicted proteins + similar to hypothetical
STMGR44	Putative non-specific lipid transfer protein StnsLTP	STMJO58	Putative delta TIP
STMGQ20	Putative non-specific lipid transfer protein StnsLTP	STMFB15	Pepper esterase
STMHS17	Protein phosphatase 2C	STMIV08	Putative delta TIP
STMEF70	Putative ankyrin protein	STMFB96	Pepper esterase
STMGE78	probable glycine rich RNA binding protein - potato {Solanum tuberosum;}	STMIN19	fringe-related protein similarity to predicted proteins + similar to hypothetical
STMHQ67	Cysteine protease inhibitor 9 precursor (PKIX) (pT1)	STMFA06	hypothetical protein
STMGO76	poly (A) polymerase family protein similar to SP Q9BWT3 Poly(A) polymerase gamma	STMCX38	L-asparaginase
STMiy82	Jasmonic acid 2	STMDG54	Putative short-chain type alcohol dehydrogenase
STMHQ63	EbiP4168 (Fragment)	STMEU17	AUX1-like protein
STMHO27	expressed protein	STMIQ88	Putative short-chain type alcohol dehydrogenase
STMGA34	Ethylene-responsive late embryogenesis-like protein	STMER46	amidohydrolase family protein similar to SP P32375 Allantoinase
STMiy72	UDP-glucuronosyl/UDP-glucosyl transferase family protein contains Pfam profile:	STMGU05	Protein kinase domain putative
STMHZ25	Ci21B protein	STMCP35	Putative allantoinase (Fragment)
STMEW74	Ci21A protein	STMDH21	putative receptor kinase {Oryza sativa (japonica cultivar-group)}
STMHO79	Protein phosphatase 2C	STMEJ65	Methionine synthase (EC 2.1.1.14)
STMGO56	zinc finger (C3HC4-type RING finger) family protein contains Pfam profile: PF00097	STMCQ50	myrcene/ocimene synthase putative similar to GI:9957293; contains Pfam profile:
STMGU47	Putative Kunitz-type tuber invertase inhibitor	STMEG13	AUX1-like protein
STMHT83	hypothetical protein {Solanum demissum;}	STMCP54	Receptor serine/threonine kinase
STMEG23	Induced stolen tip protein TUB8 (Fragment)	STMJM74	expressed protein
STMGB34	17.6 kD class I small heat shock protein (Type I small heat shock protein 17.6 kDa isoform)	STMIN17	Hypothetical protein AT4g15070
STMHZ63	expressed protein	STMCG17	Putative proline-rich protein (Fragment)
STMHO80	Hsp20.0 protein	STMJA22	Receptor serine/threonine kinase
STMHN79	TATA-binding protein-associated factor TAFII55 family protein contains Pfam profile:	STMCP60	Putative allantoinase (Fragment)
STMGP14	expressed protein similar to Hypothetical 72.2 kDa protein in RPS27A-GPM1 intergenic	STMFB43	amidohydrolase family protein similar to SP P32375 Allantoinase
STMJO47	Early light inducible protein	STMEW76	glutaredoxin family protein contains INTERPRO Domain IPR002109 Glutaredoxin
STMHS26	hypothetical protein	STMEH13	Glucan endo-1 3-beta-glucosidase basic isoform 2 precursor ((1->3)-beta-glucan
STMEI68	Mitochondrial small heat shock protein	STMCM83	endonuclease/exonuclease/phosphatase family protein contains Pfam profile PF03372:
STMGZ28	Nonspecific lipid-transfer protein (Fragment)	STMCX35	24K germin like protein precursor
STMIM42	glycine-rich protein	STMEN96	S-adenosyl-methionine-sterol-C-methyltransferase (EC 2.1.1.41)
STMIR06	dehydrodolichyl diphosphate synthase putative / DEDOL-PP synthase putative similar	STMEA15	LEXYL1 protein

Sobre expresados en T07 y T10		Sub expresados en T07 y T10	
STMHL10	null	STMDV08	null
STMHU70	putative CCAAT-binding transcription factor	STMEP30	Ammonium transporter 1 member 1 (LeAMT1;1)
STMGQ82	Chloroplast small heat shock protein class 1	STMGF63	Pectin methyl esterase (EC 3.1.1.11)
STMIC42	Early light inducible protein	STMEB78	Glucan endo-1 3-beta-glucosidase basic isoform 2 precursor ((1->3)-beta-glucan
		STMIZ76	Patatin-like protein 1
		STMIO46	S-adenosyl-L-methionine:carboxyl methyltransferase family protein similar to

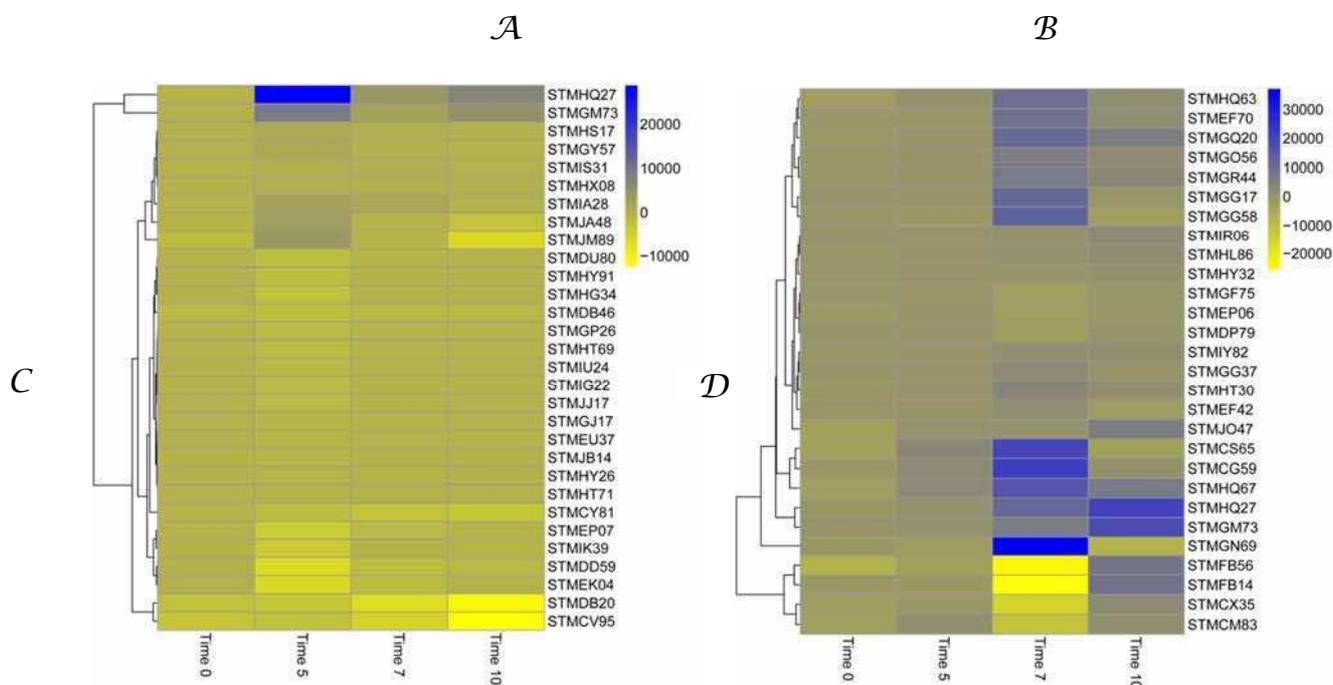
Para visualizar el comportamiento de los genes identificados en la Tabla 24, sus perfiles temporales de logFC se graficaron en la Figura 22. Nuevamente perfiles contrastantes pueden observarse en T07, comportamiento que ya se insinúa en T05.



**Figura 22. Perfiles de logFC para los genes listados en la Tabla 24.**

Las estrategias de análisis presentadas muestran que hay un cambio en la expresión diferencial de genes frente al estrés que es dependiente de la especie. La dinámica que adquieren esos genes se pudo apreciar en los gráficos de perfiles de expresión presentados. Seguidamente se muestra otro recurso grafico que refleja el tipo de respuesta más temprano de CPH en

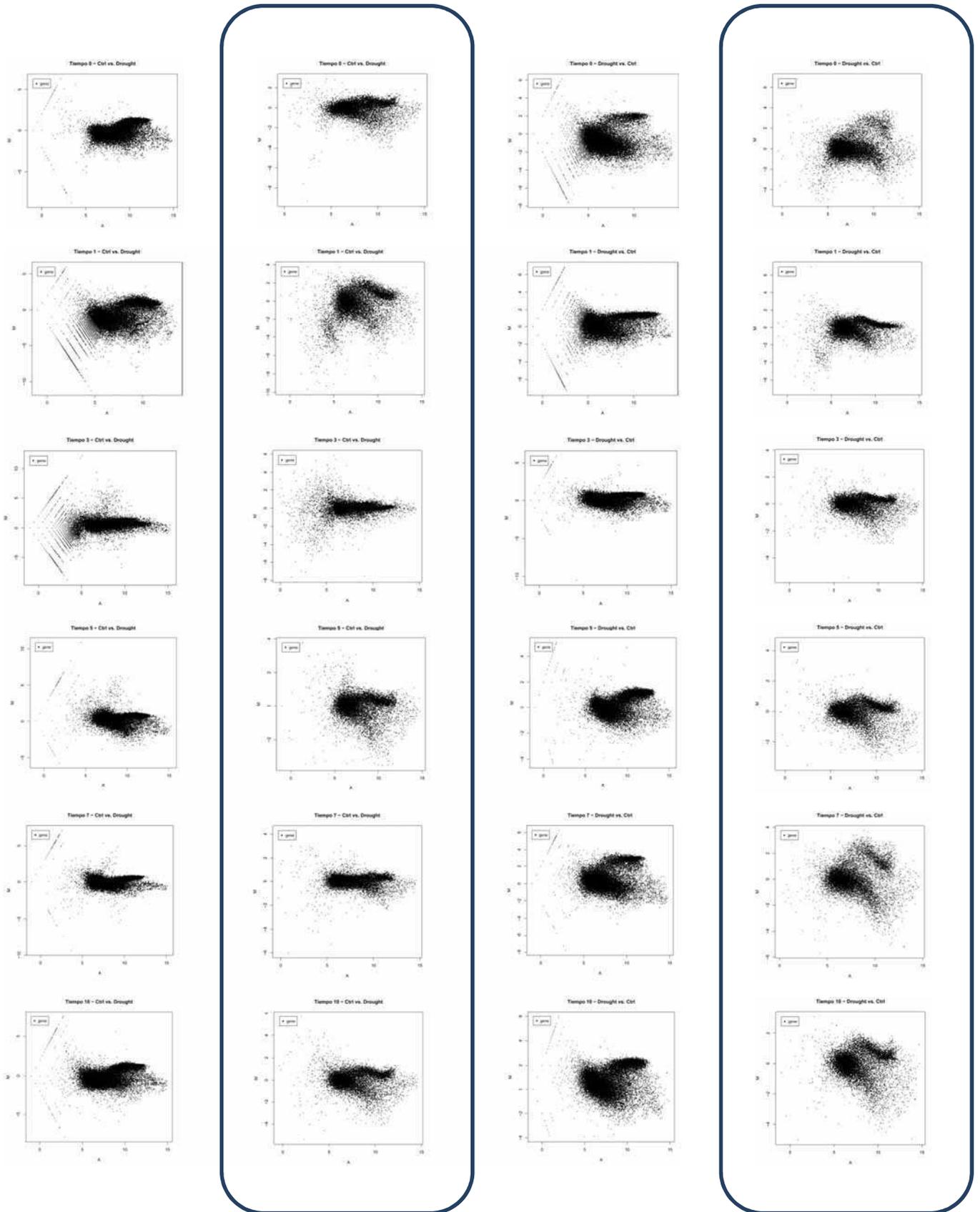
relación a VNT para hacer frente al estrés (Figura 23). En CPH el tiempo más cambiante muestra ser el día 05 del estrés y en VNT los cambios más importantes se dan en T07 y T10.



**Figura 23. Expresión relativa de los genes diferenciales más extremos bajo condiciones de estrés hídrico en los genotipos CPH (A) y VNT (B) en los tiempos T00, T05, T07 y T10 del estrés.** En la escala cromática, el color azul hace referencia a la máxima expresión de genes inducidos bajo estrés y el amarillo refiere a la máxima represión de genes bajo condiciones de sequía. El color gris denota expresiones no diferenciales. C y D corresponden al agrupamiento jerárquico de los genes en estudio.

### ***Solanum pascoense***

Tras aplicar la misma rutina de análisis utilizada para las especies CPH y VNT, los datos expresión de *S. pascoense* arrojaron inconsistencias estadísticas que llevaron a tomar la decisión de no ser considerados para estudios posteriores, por la información equivocada que se pudiera inferir del análisis de los mismos. Como hemos visto, es muy útil en la práctica realizar gráficos que permitan examinar los resultados del experimento de micromatrices. En la Figura 18 se muestran los gráficos tipo MA para la especie PSC, donde se ve el efecto del proceso de normalización dentro de cada arreglo por el método *Print-tip-lowess*, a fin de detectar problemas con el comportamiento de los datos del experimento.



**Figura 24. Gráficos MA del ensayo de micromatrices con la especie PSC.** Incluye hibridaciones con la combinación “control – Cy3” vs “estresado-Cy5” y su dye-swap en 6 momentos de estrés hídrico: T00, T01, T03, T05, T07 y T10. Las columnas resaltadas en azul muestran el efecto de la normalización aplicada sobre los datos.

Estas salidas revelan que en T00, una de las repeticiones (PSC) no se comporta como lo esperado, mostrándose como uno de los tiempos con mayor diferenciación cuando se aplica el análisis de componentes principales (Di Rienzo, J. comunicación personal, 2010). Teniendo en cuenta que en el tiempo 00 hay sólo dos micromatrices, aun implementando el procedimiento bayesiano para aumentar la fiabilidad estadística, los p-valores resultan ser muy altos, en comparación con los otros experimentos llevados a cabo con las especies CPH y VNT y por ende, no se puede explicar que el comportamiento diferencial de los genes sea en respuesta al estrés o producto del azar. Tal es así, que aun implementando esta corrección para controlar falsos positivos, no es posible identificar genes expresados diferencialmente en forma significativa.

Sobre la base de los otros 5 tiempos (10 micromatrices), los resultados de la especie PSC siguen mostrando falta de coherencia. En cada tiempo son pocos los datos obtenidos por gen para estimar correctamente su expresión.

La falta de repeticiones en cada tiempo dificulta sortear los problemas encontrados en algunas de las repeticiones y es muy difícil tener una buena estimación de lo que está ocurriendo en el perfil temporal. Por otra parte no es posible asumir un modelo simplificado de la evolución en el tiempo que permita un modelado de la evolución temporal.

### **Análisis bioinformático con el programa Limma**

El análisis de los datos de expresión de las especies CPH y VNT se realizó utilizando el conjunto de herramientas del programa Limma (Smyth, 2005) siguiendo una rutina diseñada especialmente para este tipo de micromatrices y diseño experimental (ver Material complementario en CD adjunto). Este análisis arrojó diferencias significativas en los perfiles transcripcionales de las especies estudiadas. El análisis comparativo de los perfiles de expresión de estas dos especies silvestres de *Solanum* reveló que en CPH se ven afectados por el tratamiento de sequía un total de 235 genes mientras que en VNT la diferenciación involucra 4133 genes. Ambas especies comparten 151 genes de expresión diferencial por la sequía (Tabla 25).

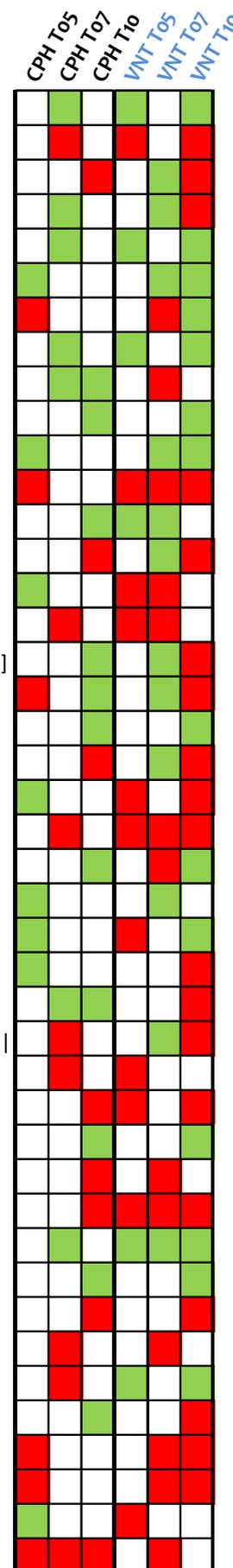
**Tabla 25. Genes diferencialmente expresados en paralelo en CPH y VNT frente al tratamiento de sequía en los tiempos T05, T07 y T10.** Se incluye el nombre del clon asignado a la micromatriz y su anotación según el Índice de Genes de *Solanum tuberosum* de TIGR (GISt). A la derecha de la tabla y con colores se muestra la inducción (rojo) o represión (verde) de cada gen y tiempo analizado. Casillas en blanco: sin expresión diferencial.

Noml Clon	Anotación TIGR StGI (Índice de genes de <i>S. tuberosum</i> - TIGR )	CPH T <sub>0</sub>	CPH T <sub>5</sub>	CPH T <sub>7</sub>	VNT T <sub>05</sub>	VNT T <sub>07</sub>	VNT T <sub>10</sub>
STMCF50	Lipase-like protein						
STMCH16	Ovule/fiber cell elongation protein Ghfe1						
STMCK16	GDSL-motif lipase/hydrolase family protein similar to family II lipases EXL3 GI:						
STMCM56	expressed protein						
STMCN22	Putative 7-transmembrane G-protein-coupled receptor						
STMCN43	Glutathione peroxidase (EC 1.11.1.9)						
STMCO76	Type-2 metallothionein						
STMCP56	Polyphenol oxidase B chloroplast precursor (PPO) (Catechol oxidase) (Fragment)						
STMQC37	oxidoreductase 2OG-Fe(II) oxygenase family protein similar to flavonol synthase						
STMQC50	myrcene/ocimene synthase putative similar to GI:9957293						
STMQC64	RAB11B						
STMCR58	Bactinecin 11						
STMCS04	BAG domain-containing protein contains Pfam:PF02179 BAG domain						
STMCS77	unknown protein {Arabidopsis thaliana						
STMCV58	SRE1a						
STMCV86	expressed protein						
STMCV95	GDSL-motif lipase/hydrolase family protein similar to family II lipase EXL3						
STMCX15	BAG domain-containing protein contains Pfam:PF02179 BAG domain						
STMCY65	1-deoxyxylulose 5-phosphate synthase						
STMCY81	Dehydration-responsive element binding protein 3						
STMZ36	Hypothetical protein PPS3						
STMZ41	Alpha-amylase precursor						
STMDB20	GDSL-motif lipase/hydrolase family protein similar to family II lipase EXL3						
STMDB38	SRE1a						
STMDB46	Propolyphenol oxidase precursor (Fragment)						
STMDD06	null						
STMDD41	protodermal factor 1 (PDF1) identical to protodermal factor 1 [Arabidopsis thaliana]						
STMDD59	CCR4-NOT transcription complex protein putative similar to SWISS-PROT:Q9UFF9 CCR4-NOT						
STMDF56	probable arabinogalactan protein precursor - tomato {Lycopersicon esculentum						
STM DG54	Putative short-chain type alcohol dehydrogenase						
STM DH21	putative receptor kinase {Oryza sativa (japonica cultivar-group)}						
STM DH35	RING-H2 finger protein RHC1a [imported] - Arabidopsis thaliana						
STM DH85	aspartyl-tRNA synthetase						
STM DH95	null						
STM DI20	unknown protein {Arabidopsis thaliana						
STM DP28	expressed protein						
STM DP89	hypothetical protein F28A21.170 - Arabidopsis thaliana {Arabidopsis thaliana						
STM DS45	ribosomal protein L14 putative						
STM DT08	pectinesterase family protein contains Pfam profile: PF01095 pectinesterase						

Nombre del Clon

Anotación TIGR StGI (Índice de genes de *S. tuberosum* - TIGR)

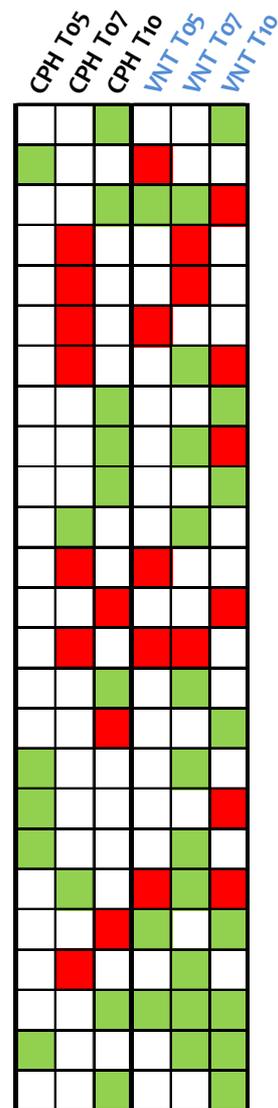
STMDV05	60S acidic ribosomal protein P0
STMDW37	expressed protein
STMDZ61	Fe-superoxide dismutase (EC 1.15.1.1)
STMEA30	Putative gamma TIP
STMEA73	MAR-binding protein
STMEC75	Isochorismate synthase (EC 5.4.99.6)
STMED56	Hypothetical protein
STMEF15	expressed protein { <i>Arabidopsis thaliana</i> }
STMEG89	Peroxidase
STMEH63	phytochrome-interacting factor 4 (PIF4) / basic helix-loop-helix protein 9 (bHLH9) /
STMEH75	Similar to <i>Dictyostelium discoideum</i> (Slime mold). prespore-specific protein
STMEI68	Mitochondrial small heat shock protein
STMEI79	Probable xyloglucan endotransglucosylase/hydrolase 1 precursor (LeXTH1)
STMEI91	Fe-superoxide dismutase (EC 1.15.1.1)
STMEK04	Class I chitinase (EC 3.2.1.14 EC 3.2.1.14)
STMEN18	ML domain putative
STMEN59	protodermal factor 1 (PDF1) identical to protodermal factor 1 [ <i>Arabidopsis thaliana</i> ]
STMEP27	null
STMEP30	Ammonium transporter 1 member 1 (LeAMT1
STMES65	expressed protein
STMES89	calmodulin-related protein 2 touch-induced
STMET16	expressed protein
STMEU54	Putative ABC transporter
STMEV24	Tuber-specific and sucrose-responsive element binding factor
STMEV47	ZPT2-13
STMEW40	Predicted protein
STMEY77	Beta-1 3 glucanase precursor (EC 3.2.1.39 EC 3.2.1.39)
STMFA20	DNAJ heat shock N-terminal domain-containing protein low similarity to SP O75190
STMFB70	expressed protein
STMGC64	acireductone dioxygenase (ARD/ARD') family protein similar to iron-deficiency induced
STMGD38	exostosin family protein contains Pfam profile: PF03016 Exostosin family
STMGD68	Putative late blight resistance protein
STMGD72	Class II small heat shock protein Le-HSP17.6
STMGE12	exostosin family protein contains Pfam profile: PF03016 Exostosin family
STMGE39	EEF53 protein (Fragment)
STMGF18	Phosphoenolpyruvate carboxylase kinase 2
STMGG17	M355
STMGG40	Plastidic ATP/ADP-transporter
STMGG43	Patatin
STMGM73	Hypothetical protein B1027A11.138
STMGM76	leucine-rich repeat family protein / extensin family protein similar to extensin-like
STMGP26	Avr9/Cf-9 rapidly elicited protein 284
STMGS26	Metallothionein-like protein type 2 B



Nombre del Clon

Anotación TIGR StGI (Índice de genes de *S. tuberosum* - TIGR )

STMIS68	putative receptor-like protein kinase
STMIU24	zinc finger (C3HC4-type RING finger) family protein contains Pfam domain PF00097:
STMIW53	Aspartic proteinase 2
STMJY47	digalactosyldiacylglycerol synthase
STMIZ30	Phosphatidylinositol-4-phosphate 5-Kinase putative
STMIZ51	ribosomal protein L33 family protein ribosomal protein L33 Rickettsia prowazekii PIR:
STMJC21	Uncharacterised protein family
STMJD58	subtilase family protein similar to subtilase GI:9957714 from [ <i>Oryza sativa</i> ]
STMJF44	unknown { <i>Prunus armeniaca</i>
STMJF86	UDP-glucuronosyl/UDP-glucosyl transferase family protein contains Pfam profile:
STMJG06	Putative thaumatin-like protein
STMJG33	hypothetical protein
STMJG42	hydroxyproline-rich glycoprotein family protein contains proline-rich extensin domains
STMJG44	expressed protein
STMJH02	Putative anthocyanidine rhamnosyl-transferase
STMJJ13	Putative lipid transfer protein
STMJJ17	Xyloglucan endotransglucosylase-hydrolase XTH3
STMJJ23	Putative ripening-related protein
STMJK74	Pectinesterase 3 precursor (Pectin methylesterase 3) (PE 3) (EC 3.1.1.11 EC 3.1.1.11)
STMJL93	expressed protein
STMJN56	Glycoprotein
STMJN87	Farnesyl-protein transferase beta subunit (EC 2.5.1.-)
STMJP48	Serine protease
STMJP70	Transcription regulatory protein SWI1 (SWI/SNF complex component SWI1)
STMJP82	Hypothetical protein PPS3



Por su parte del total de los genes diferenciales, la especie CPH expresa 57 genes exclusivos y VNT por su parte 3967. Los genes específicos en CPH están todos reprimidos en T05 (17%), mayoritariamente reprimidos en T07 y en proporciones similares de inducción y represión en T10. En la Tabla 26 se listan los genes diferenciales que sólo se expresan en CPH, el tiempo en el que se expresan y el tipo de expresión que presentan (inducidos y reprimidos). En el material complementario se encuentran los datos de VNT.

**Tabla 26. Genes únicos diferencialmente expresados en la especie CPH frente al tratamiento de sequía en los tiempos T05, T07 y T10.** Se incluye el nombre del clon asignado a la micromatriz y su anotación según el Índice de Genes de Solanum tuberosum de TIGR (GISt). A la derecha de la tabla y con colores se muestra la inducción (rojo) o represión (verde) de cada gen y tiempo analizado

Nombre del Clon	Anotación TIGR StGI (Índice de genes de <i>S. tuberosum</i> - TIGR)	CPH T05	CPH T07	CPH T10
STMDJ76	DNA-binding protein 3	Green	White	White
STMEK70	BURP domain-containing protein similarity to SP Q08298 Dehydration-responsive protein	Green	White	White
STMEP07	WIZZ	Green	White	White
STMEU37	CaCBF1B	Green	White	White
STMGD50	expressed protein contains similarity to GPI-anchored protein	Green	White	White
STMGJ17	expressed protein	Green	White	White
STMGL65	myb family transcription factor contains Pfam profile: PF00249 Myb DNA binding domain	Green	White	White
STMHT71	MA3 domain-containing protein low similarity to programmed cell death 4 protein	Green	White	White
STMJB12	expressed protein	Green	White	White
STMJB14	Xyloglucan endotransglucosylase-hydrolase XTH3	Green	White	White
STMBB19	tomato HSP90	White	Red	White
STMBB31	tomato HSP90	White	Red	White
STMCK87	Malonyl CoA:anthocyanin 5-O-glucoside-6'''-O-malonyltransferase	White	Green	White
STMDM93	Similar to Dictyostelium discoideum (Slime mold). Non-receptor tyrosine kinase spore lysis	White	Red	White
STMEH57	expressed protein	White	Green	White
STMGT74	high mobility group protein alpha (HMGalpha)/ HMG protein alpha nearly identical to	White	Red	White
STMGU33	multi-copper oxidase type I family protein contains Pfam profile: PF00394 Multicopper	White	Green	White
STMHH88	MATH domain putative	White	Red	White
STMHP89	Calcium/calmodulin-dependent protein kinase homolog CaM kinase homolog MCK1 protein	White	Red	White
STMHW31	expressed protein	White	Red	White
STMHW35	calcineurin-like phosphoesterase family protein contains Pfam profile: PF00149	White	Red	White
STMHW51	Copper chaperone	White	Red	White
STMIB22	putative zinc finger protein	White	Red	White
STMID62	F24B9.7 protein	White	Red	White
STMIR80	Storekeeper protein	White	Green	White
STMJB92	integral membrane Yip1 family protein contains Pfam domain PF04893: Yip1 domain	White	Green	White
STMJL31	histone H3.2 protein	White	Red	White
STMCB40	F20B17.23 {Arabidopsis thaliana;}	White	White	Red
STM CX59	3-beta hydroxysteroid dehydrogenase/isomerase family protein contains Pfam profile	White	White	Green
STM CX70	Flavonol 3-O-glucosyltransferase (UDP-glucose flavonoid 3-O-glucosyltransferase)	White	White	Red
STM DU36	proton-dependent oligopeptide transport (POT) family protein contains Pfam profile:	White	White	Green
STM EF22	Xyloglucan-specific fungal endoglucanase inhibitor protein	White	White	Green
STM EH59	Alcohol NADP+ oxidoreductase	White	White	Green
STM EJ39	Putative metallophosphatase	White	White	Green
STM EY41	ribosomal protein S14p/S29e	White	White	Red
STM GD48	Patatin T5 precursor (Potato tuber protein)	White	White	Green
STM GF92	CMP-sialic acid transporter-like protein	White	White	Green
STM GG89	Phosphoenolpyruvate carboxylase kinase 2	White	White	Red

Nombre del Clon	Anotación TIGR StGI (Índice de genes de <i>S. tuberosum</i> - TIGR)	CPH To5	CPH To7	CPH T10
STMGR18	expressed protein			Green
STMGV53	galactosyl transferase GMA12/MNN10 family protein very low similarity to alpha-1			Green
STMGY54	expressed protein predicted protein <i>Caenorhabditis elegans</i> gb:Z70780		Red	Red
STMHI53	expressed protein similar to Symplekin (SP:Q92797) { <i>Homo sapiens</i> }			Green
STMHK02	short-chain dehydrogenase/reductase (SDR) family protein contains INTERPRO family			Green
STMHK83	Similar to <i>Dictyostelium discoideum</i> (Slime mold). Homeobox-containing protein			Green
STMHL62	Patatin precursor			Green
STMHR14	expressed protein		Red	Red
STMIN51	expressed protein		Red	Red
STMJF29	Ribosomal protein S21e		Red	Red
STMJF94	Chalcone--flavonone isomerase B (Chalcone isomerase B)		Red	Red
STMJJ47	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein contains		Red	Red
STMJM09	Ribosomal L39 protein		Red	Red
STMJM89	vacuolar protein VAC7 - yeast ( <i>Saccharomyces cerevisiae</i> ) { <i>Saccharomyces cerevisiae</i> ;}}			Green
STMJN59	Helicase conserved C-terminal domain putative		Red	Red
STMJN78	expressed protein		Red	Red
STMJO27	dynein light chain type 1 family protein similar to SP O02414 Dynein light chain LC6		Red	Red
STMJO65	CycD3;2 protein			Green

Los datos de expresión en cada tiempo así como las referencias de clones y sus anotaciones están disponibles en su totalidad para las dos especies en Tablas suplementarias en el siguiente vínculo: <http://m4m.ugr.es/Flor/Tesis.html> o en el CD adjunto a este trabajo como Material Complementario.

Para realizar un análisis global de la expresión diferencial de los genes obtenidos tras el análisis bioinformático, se utilizó el Programa MapMan (Thimm *et al*, 2004), una herramienta desarrollada en el grupo del Dr. Marc Stitt para el estudio de micromatrices, donde en cada tiempo del tratamiento analizado y con cada especie, se generó una agrupación de los genes dentro de las principales rutas metabólicas de plantas o procesos afectados. Esta agrupación arbitraria seguida por el Programa, se muestra en diagramas y Tablas.

### **Análisis de micromatrices a través de rutas metabólicas**

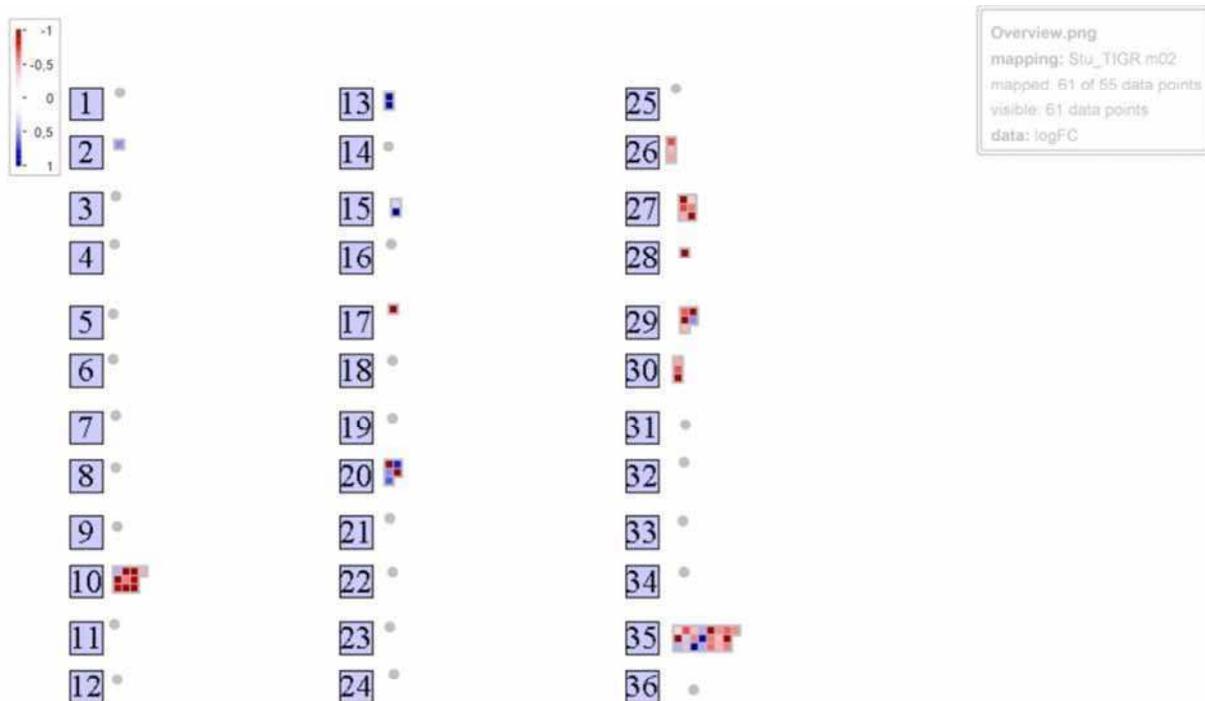
#### ***Solanum cardiophyllum***

Del total de genes expresados diferencialmente en respuesta al estrés hídrico en la especie CPH, un 23% se observa en el T05, de los cuales el 75% están reprimidos. En líneas generales en este tiempo, los procesos más afectados tienen relación con el metabolismo de la pared celular, con proteínas de respuesta al estrés biótico y abiótico, con factores de transcripción avocados al procesamiento del ARN, con metabolismo de aminoácidos y señalización (Figuras 25, 26 y 27).

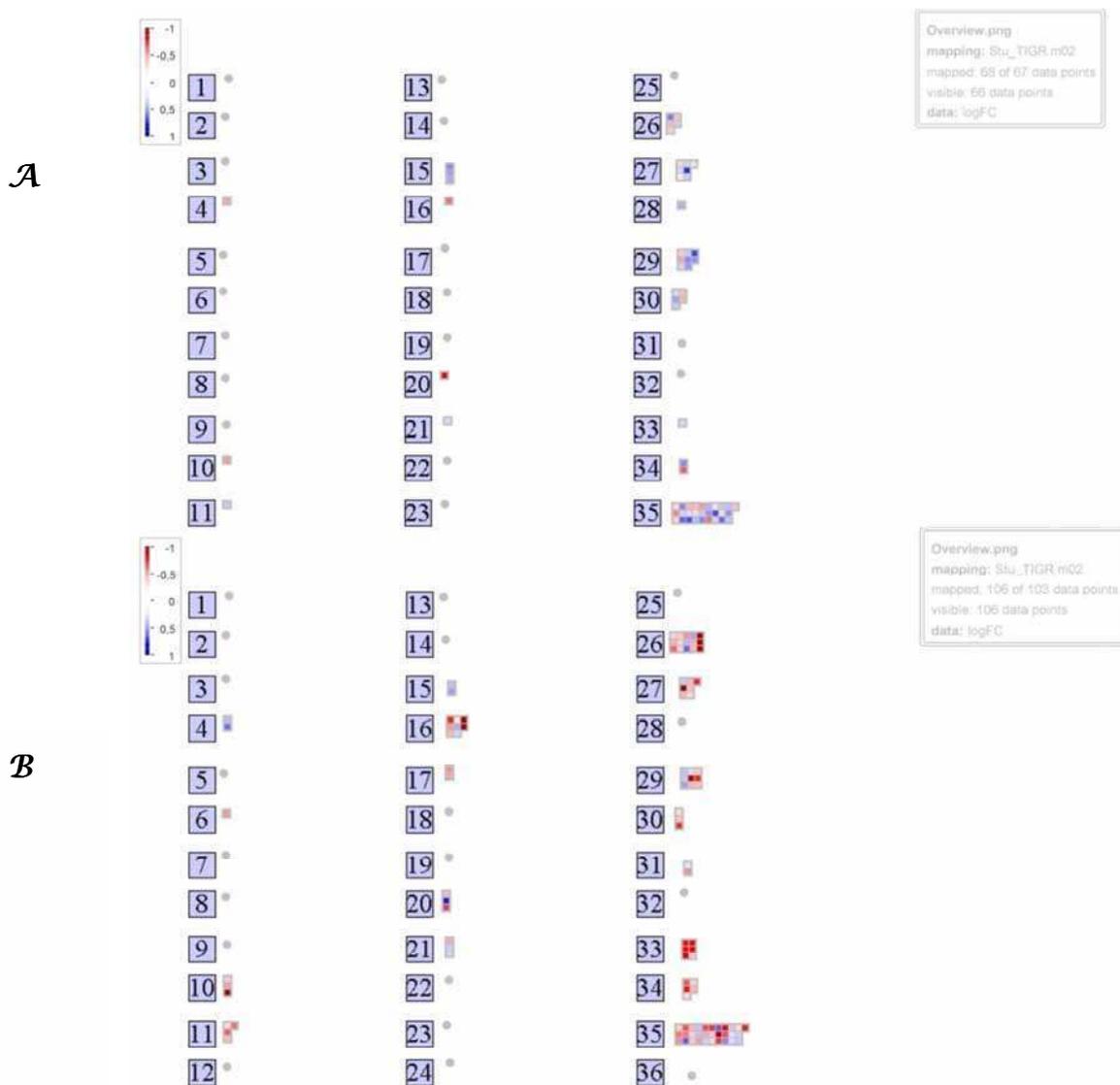
En el tiempo 07 del estrés, la especie CPH muestra la activación y represión de un número mayor de genes que comprenden el 28% del total de expresión diferencial durante la sequía. Hay una mayor proporción de genes inducidos. Los principales procesos afectados involucran la regulación del metabolismo de proteínas, la regulación de la transcripción del ARN, las vías de señalización así como la glicólisis se ve afectada en este tiempo.

Finalmente, el día 10 del estrés revela ser el momento de mayor expresión diferencial de todo el tratamiento, donde el 44% del total de genes se muestran

afectados por el estrés y donde particularmente el 70% de ellos se han sub-expresado respecto del control irrigado. Aquí observamos cambios en genes relacionados con procesos redox celulares, en los de glicólisis y de gluconeogénesis, los genes relacionados con la pared celular y el metabolismo secundario. Esta última categoría es la que posee mayor número de genes afectados diferencialmente expresados. Hay genes relacionados con el desarrollo que inician su regulación en este tiempo. También es importante la acción de los factores de transcripción, el metabolismo proteico y los mecanismos de transporte celular.



**Figura 25. Visión general del metabolismo puesto en marcha en la especie CPH en el tiempo T05 del tratamiento de sequía.** Del 1 al 36 se muestran los grupos de clasificación utilizados por el programa MapMan para el análisis de las principales vías metabólicas afectadas. **1:** fotosíntesis, **2:** metabolismo principal de carbohidratos, **3:** metabolismo secundario de Hidratos de carbono, **4:** Glucólisis, **5:** Fermentación, **6:** gluconeogénesis, **7:** OPP, **8:** TCA, **9:** transporte de electrones mitocondrial/síntesis de ATP, **10:** Pared celular, **11:** metabolismo de lípidos, **12:** metabolismo de N, **13:** metabolismo de aminoácidos, **14:** asimilación de S, **15:** manejo de metales, **16:** metabolismo secundario, **17:** metabolismo hormonal, **18:** metabolismo de cofactores y vitaminas, **19:** síntesis de tetrapirroles, **20:** estrés, **21:** redox, **22:** metabolismo de poliaminas, **23:** metabolismo de nucleótidos, **24:** biodegradación de xenobióticos, **25:** metabolismo C1, **26:** misc, **27:** ARN, **28:** ADN, **29:** proteína, **30:** señalización, **31:** célula, **33:** desarrollo, **34:** transporte, **35:** sin asignación.



**Figura 26. Visión general del metabolismo de la especie CPH en el tiempo T07 y T10 y sus cambios durante el tratamiento de sequía (A y B respectivamente).** Del 1 al 36 se muestran los grupos de clasificación utilizados por el programa MapMan para el análisis de las principales vías metabólicas afectadas. 1: fotosíntesis, 2: metabolismo principal de carbohidratos, 3: metabolismo secundario de Hidratos de carbono, 4: Glucólisis, 5: Fermentación, 6: gluconeogénesis, 7: OPP, 8: TCA, 9: transporte de electrones mitocondrial/síntesis de ATP, 10: Pared celular, 11: metabolismo de lípidos, 12: metabolismo de N, 13: metabolismo de aminoácidos, 14: asimilación de S, 15: manejo de metales, 16: metabolismo secundario, 17: metabolismo hormonal, 18: metabolismo de cofactores y vitaminas, 19: síntesis de tetrapirroles, 20: estrés, 21: redox, 22: metabolismo de poliaminas, 23: metabolismo de nucleótidos, 24: biodegradación de xenobióticos, 25: metabolismo C1, 26: misc, 27: ARN, 28: ADN, 29: proteína, 30: señalización, 31: célula, 33: desarrollo, 34: transporte, 35: sin asignación. En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad basada en logFC.

## **Metabolismo de hidratos de carbono**

Durante todo el proceso de estrés, sólo vemos afectado los genes intervinientes en el metabolismo de hidratos de carbono en el tiempo T05. Se activan  $\alpha$ -amilasas, que intervienen en la degradación del almidón, posicionándose entre de los genes de expresión más contrastantes. Con el avance del estrés no se ven afectados otros genes de este metabolismo.

## **Genes asociados a la Pared celular**

Durante todo el tratamiento de estrés, los genes relacionados con la pared celular se mostraron en general reprimidos. La expresión diferencial de estos genes se produce en todos los tiempos analizados. En T05 se reprimen los genes que participan en la degradación y modificación celular como los Xiloglucanos Endotransglicosilasas-Hidrolasa y proteínas de respuesta a la deshidratación del tipo RD22 (Pectato Liasas), al igual que se reprime la Pectín Esterasa. Sólo se induce en este tiempo las proteínas tipo Extensinas.

En T07 sólo interviene una Pectín Esterasa, que se sub-expresa en las plantas estresadas. En T10 se reprimen Xiloglucanos Endotransglicosilasas-Hidrolasas y Pectato Liasas que intervienen en la degradación de pared.

## **Glucólisis**

La sequía no afecta sustancialmente la expresión de genes asociados a la glucólisis en la especie CPH. En el tiempo T07, se reprimen Glucosa-6-fosfato Isomerasa que cataliza la síntesis de fructosa-6-fosfato a partir de glucose-6-fosfato. En T10 se induce la Fosfoenolpiruvato carboxilasa-kinasa 2.

## **Gluconeogénesis**

Durante el día 10 del estrés, se promueve la represión de la enzima Piruvato Fosfato Dikinasa, que cataliza la conversión de piruvato en fosfoenolpiruvato en las células del mesófilo.

## **Metabolismo de aminoácidos**

Sólo se observa la activación de genes relacionados con el metabolismo de aminoácidos en el tiempo T05, en donde se activan específicamente Acetilornitina Aminotransferasas.

## **Transporte de iones metálicos**

Los genes afectados al almacenamiento y quelación de metales en respuesta al estrés se indujeron en todos los tiempos analizados. Se afectaron las Metalotioninas, Ferritinas y Chaperonas de Cobre.

## **Metabolismo de hormonas**

En T05 se reprimen el Factor 5 de Respuesta al Etileno (del inglés Ethylene response factor 5). En T07 no hay genes del metabolismo hormonal que estén diferencialmente expresados. En T10 se reprimen genes involucrados en el metabolismo de brasinoesteroides, como la proteína de elongación del óvulo/fibra celular (del inglés Ovule/fiber cell elongation protein Ghfe1) y del metabolismo del etileno como 2-Oxoglutarato-dependiente Dioxigenasa.

## **Metabolismo secundario**

En T07 se reprime la enzima Malonil CoA:Antocianina 5-O-glucosido-6"-O-Maloniltransferasa, que interviene en el metabolismo de antocianinas. En T10,

los genes involucrados en el metabolismo secundario aumentan marcadamente, induciéndose genes afectados al metabolismo de flavonoides, en particular Dihidroflavonoles (Flavonol 3-O-Glucosiltransferasa) y Chalconas (Chalcona--flavonone Isomerasa B). Se reprimen genes afectados al metabolismo de fenoles (Propolifenol Oxidasa), de terpenoides (Monoterpeno Sintasa), fenilpropanoides (Alcohol NADP+ Oxidoreductasa) y antocianinas (Leucoantocianidin Dioxigenasa, UDP-Glucoronosil y UDP-Glucosil Transferasa).

### **Proteínas de respuesta al estrés**

En T05 se inducen y reprimen genes de este tipo de proteínas. Particularmente se reprimen genes de enzimas de respuesta al estrés biótico como la Quitinasa y se inducen inhibidores de Cisteín Proteasas. Se inducen genes de proteínas de respuesta al estrés abiótico, en particular las relacionadas con altas temperaturas, como pequeñas proteínas de choque térmico (del inglés Small Heat Shock Protein – Small HSP) y se reprimen genes de proteínas de respuesta a la deshidratación como por ejemplo proteínas con dominio APETALA2/EREBPs.

En T10 se inducen nuevamente pequeñas proteínas de choque térmico y se reprimen otras de mayor peso molecular HSP70kD.

### **Procesos Rédox Celulares**

En T10 intervienen por primera vez genes con actividad rédox. Particularmente se inducen Dismutasas como Fe-superoxido Dismutasas y se reprimen las 2-Oxoglutarato-dependente Dioxigenasa.

### **Regulación de la transcripción**

La regulación de la transcripción sufre una dinámica particular a lo largo del estrés. Se reprimen esencialmente los genes en el T05, se inducen

mayoritariamente en T07 y en T10 se activan y reprimen en proporciones similares.

En T05 se reprimen genes de factores de transcripción como proteínas con dominios de dedos de zinc C2H2, MYB, y WRKY.

En T07 se inducen los de factores de transcripción del tipo GRAS, de proteínas de alta movilidad, de ensamblaje de cromatina y de unión a ADN. También se reprimen genes de proteínas de unión a ADN (proteínas de unión-MAR), factores de transcripción con dominio de dedos de zinc C2H2 y proteínas tipo GeBP, denominadas store-keeper, relacionadas con la síntesis de Patatina.

En T10 siguen reprimidos los genes de factores de transcripción que observáramos en los tiempos anteriores y se suman los de proteínas con dominios bHLH. En este período se inducen Helicasas que intervienen en el procesamiento del ARN.

### **Metabolismo de proteínas**

El metabolismo de las proteínas se ve reprimido en T05, dinámica que se va modificando conforme avanza el estrés. En el día 05, se reprimen genes de proteínas de síntesis específicas de cloroplastos, de degradación y modificación post-transcripcional. Como parte del proceso de degradación se induce en este tiempo un inhibidor de Cisteín Proteasa. En T07, la mayoría de los genes diferencialmente expresados están inducidos, representados por enzimas intervinientes en la degradación asociada a Ubiquitina E2, de modificación post-traducciona, con función secretoria como el factor de ribosilación del ADP, así como proteínas ribosomales que intervienen en la síntesis proteica. Se reprimen particularmente enzimas que intervienen en la activación de aminoácidos como Aspartil-ARNt Sintetasa. En T10 se inducen proteínas involucradas en la síntesis proteica a la vez que se reprimen aquellas afectadas a la degradación de proteínas (Subtilasas, Aspartato Proteasas y proteínas de degradación dependiente de Ubiquitina).

## **Señalización celular**

Esencialmente se sub-expresan genes relacionados con la señalización celular en los tiempo T05 y T10. En T07 se inducen genes asociados al proceso de señalización mediado por calcio y el fosfatidilinositol y se reprimen genes que regulan la señalización a través de proteínas G. En T05 se reprimen principalmente genes que intervienen en la señalización de calcio (calmodulinas y proteínas de unión al calcio) y de receptores de kinasas. En T10 se encuentran reprimidos los genes de receptores de Kinasas y los factores de interacción lumínica como los Fitocromos.

## **Estructura y división celular**

En el T07 se observan cambios a nivel de organización y división celular. Se reprimen genes de Ciclinas que intervienen en la división celular, induciéndose por su parte Dineínas, que son proteínas del motor molecular en las células que convierten la energía química contenida en el ATP en energía mecánica del movimiento.

## **Transporte celular**

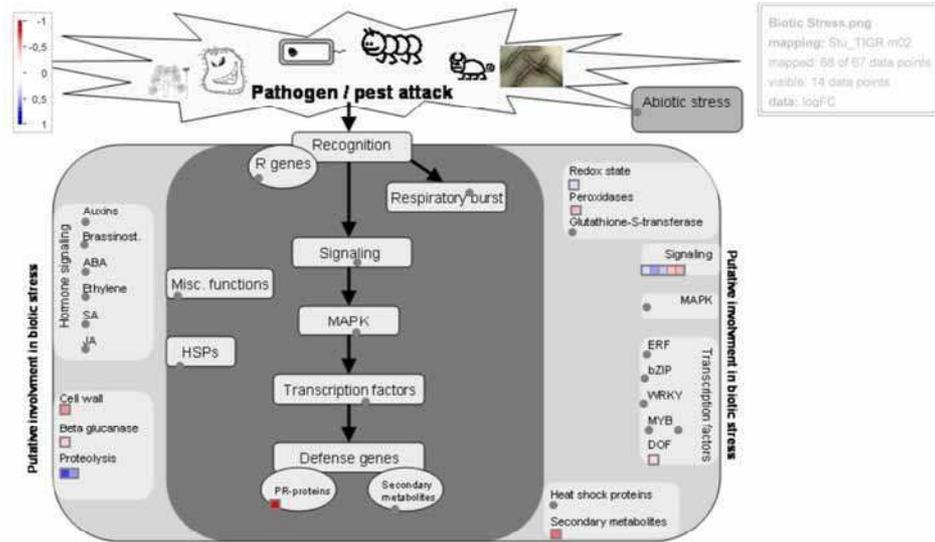
Empiezan a ser evidentes cambios en la expresión de genes de transporte en T07, donde se inducen genes de transportadores de metabolitos de membrana como el Transportador Plastídico ATP/ADP. En este tiempo, también se reprimen los de Acuaporinas.

En el T10, hay una represión generalizada de genes de transportadores de amonio, péptidos y oligopéptidos.

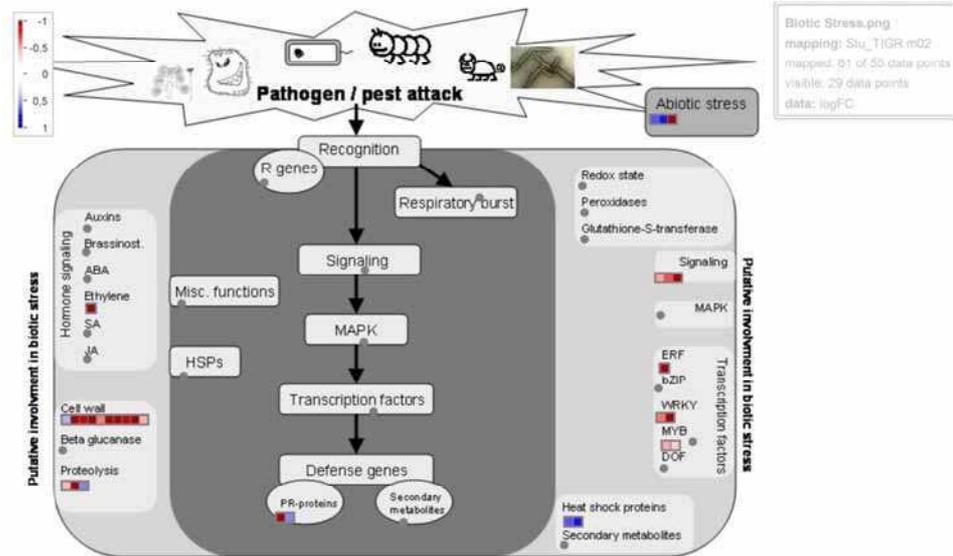
## **Desarrollo y almacenamiento**

En T10 se reprimen proteínas de almacenamiento como Patatinas y de desarrollo como Transcetolasas.

A



B



C

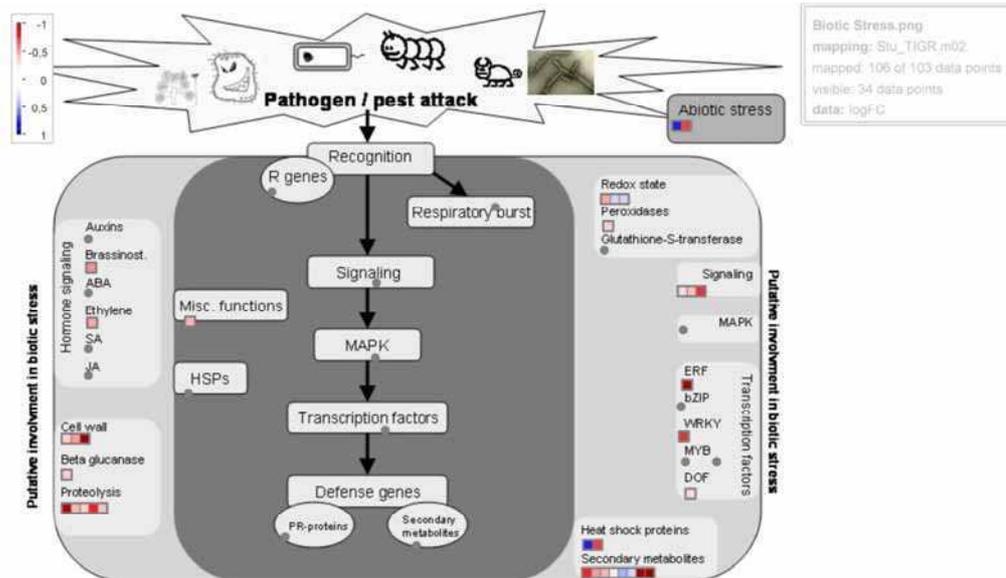


Figura 27. Esquema general de la respuesta de CPH al estrés hídrico y los cambios en la expresión de genes durante los tiempos T05, T07 y T10 (A, B y C respectivamente). En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad, basada en logFC.

## **Solanum venturii**

En la especie VNT se expresan diferencialmente 4133 genes en respuesta al estrés hídrico. En el tiempo T05 se ven afectados el 31% del total de estos genes. Mayoritariamente se encuentran sub-expresados (67%). Se inducen las enzimas asociadas al fotosistema 2 (PS2) como por ejemplo Proteínas represora ultravioleta-B (del inglés *Ultraviolet-B-repressible protein*) proteínas de membrana, proteína W del centro del fotosistema II (PsbW), polipéptidos del fotosistema II. Se inducen los transportadores de electrones como: Ferredoxina I y en paralelo, se reprimen la Ferredoxina NADP reductasa. Las enzimas intervinientes en el proceso de fotorespiración tales como Serina Hidroxi-metil-Transferasa y Glicina Hidroxi-metil-Transferasa están reprimidas.

Las enzimas del ciclo del Calvin, las Fosfoglicerato Kinasas y las Transcetolasa están reprimidas, mientras las enzimas Gliceraldehido3-Fosfato Deshidrogenasa y Fructuosa Bifosfato Aldolasas están inducidas.

Los metabolismos más afectados en *S. venturii* pueden verse en los diagramas generados por el Programa MapMan para cada tiempo de análisis en las Figuras 28, 29 y 30. En particular con la especie VNT, donde la respuesta a la sequía implica una compleja maquinaria de expresión de genes, la utilización de herramientas de visualización desarrolladas específicamente para sortear estos problemas nos ofreció una visión global y coordinada de los metabolismos implicados y que describiremos en detalle a continuación.

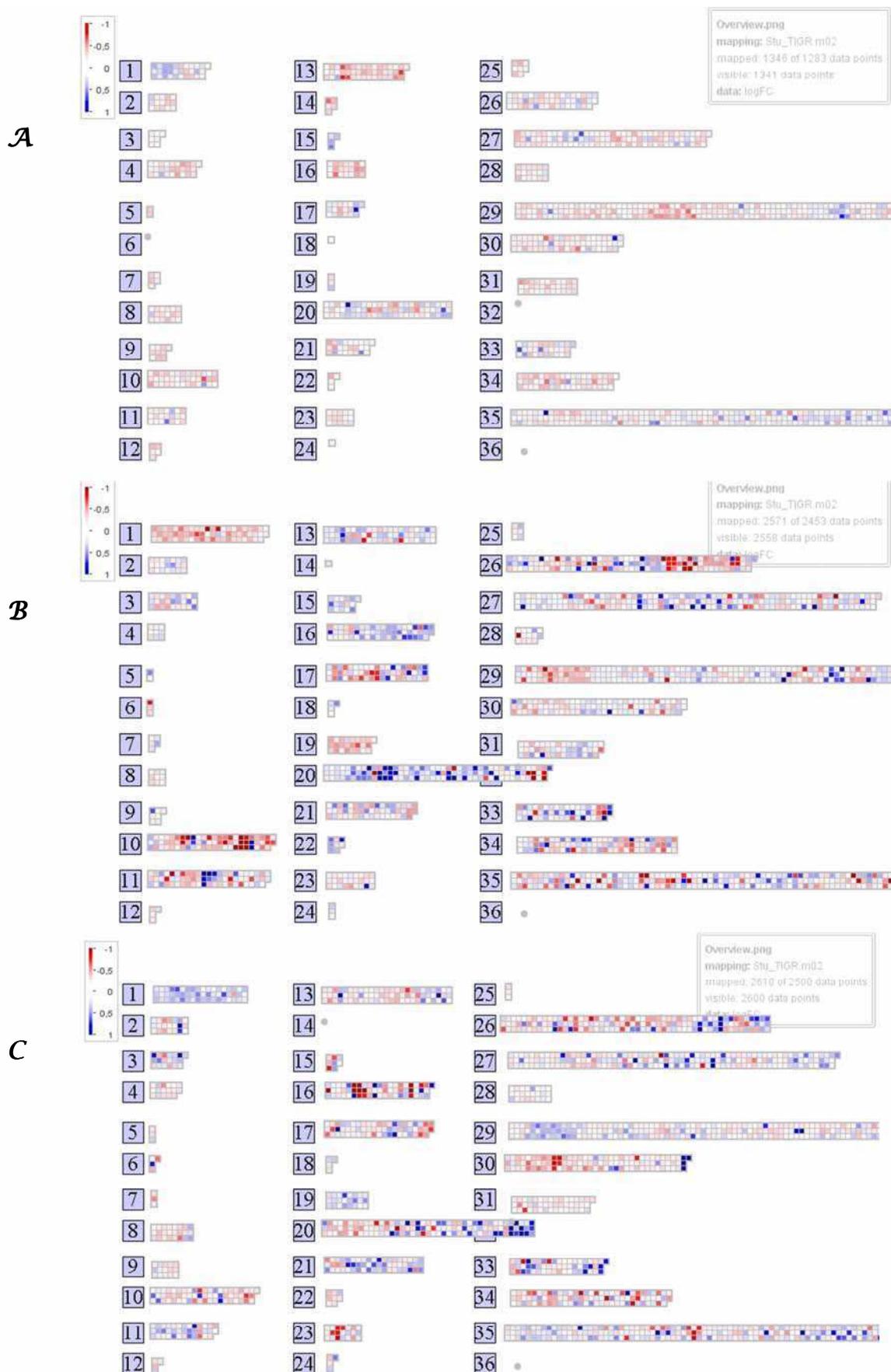
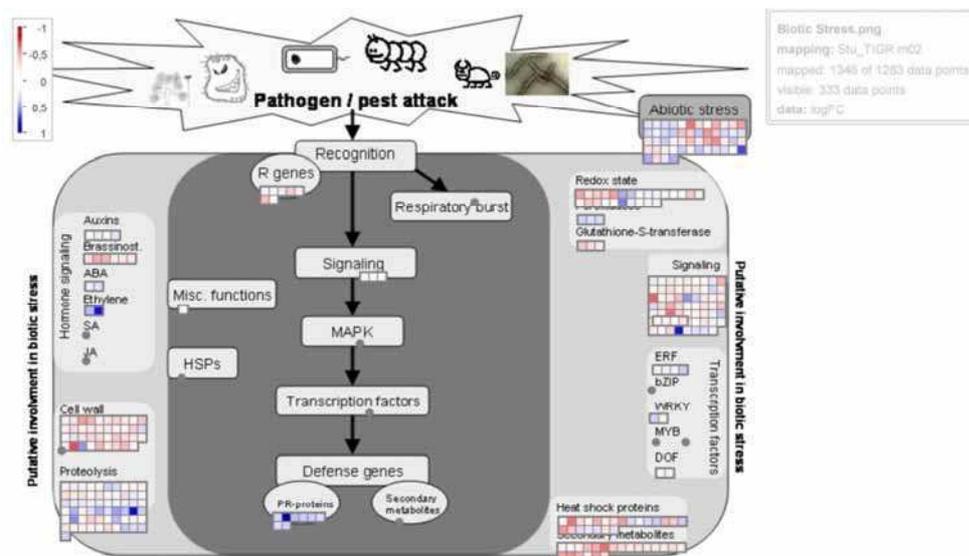
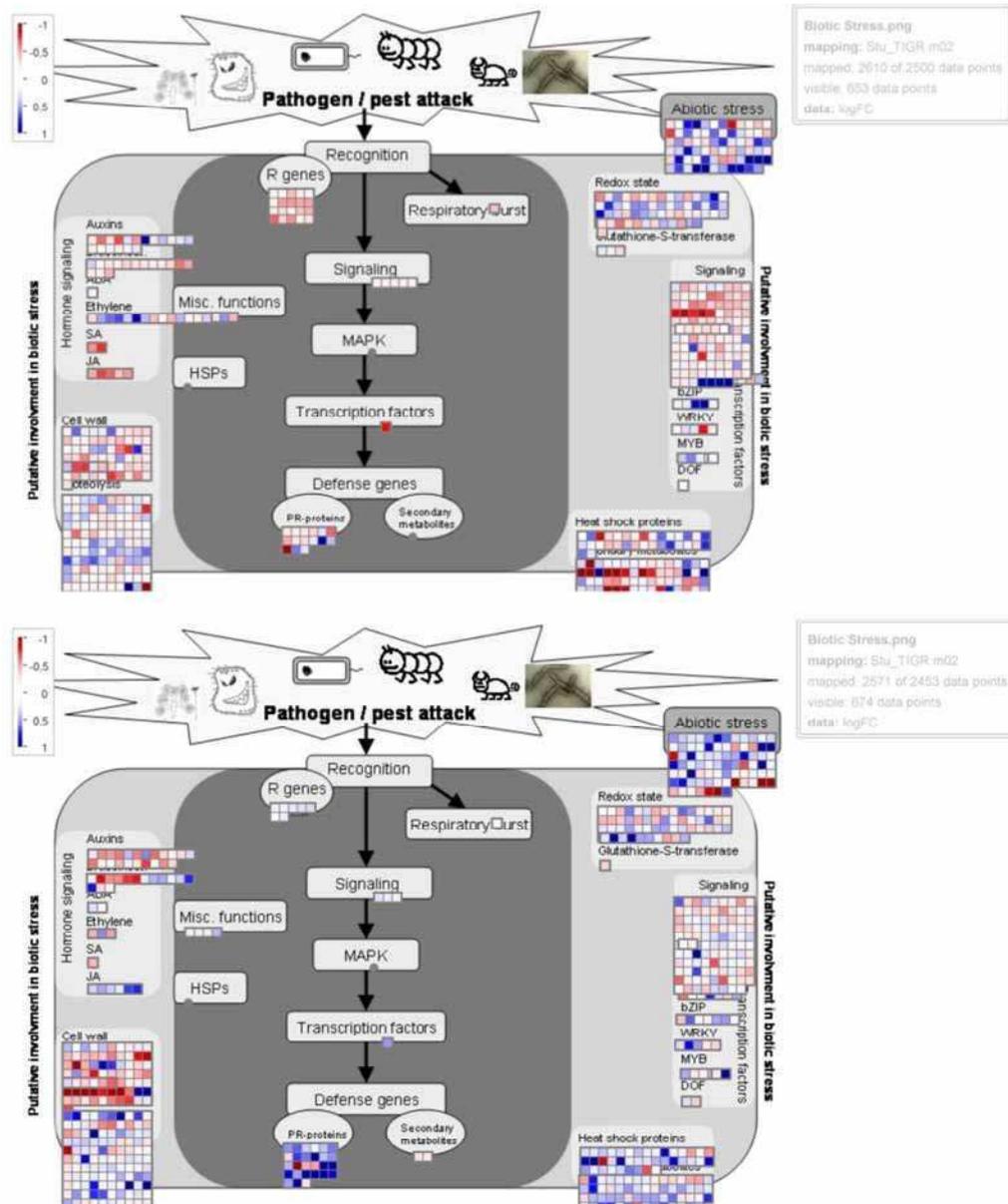


Figura 28. Visión general del metabolismo de la especie VNT y sus cambios durante los tiempos T05, T07 y T10 del tratamiento de sequía (**A**, **B** y **C** respectivamente). Del 1 al 36 se muestran los grupos de clasificación utilizados por el programa MapMan para el análisis de

las principales vías metabólicas afectadas. 1: fotosíntesis, 2: metabolismo principal de carbohidratos, 3: metabolismo secundario de Hidratos de carbono, 4: Glucólisis, 5: Fermentación, 6: gluconeogénesis, 7: OPP, 8: TCA, 9: transporte de electrones mitocondrial/síntesis de ATP, 10: Pared celular, 11: metabolismo de lípidos, 12: metabolismo de N, 13: metabolismo de aminoácidos, 14: asimilación de S, 15: manejo de metales, 16: metabolismo secundario, 17: metabolismo hormonal, 18: metabolismo de cofactores y vitaminas, 19: síntesis de tetrapirroles, 20: estrés, 21: redox, 22: metabolismo de poliaminas, 23: metabolismo de nucleótidos, 24: biodegradación de xenobióticos, 25: metabolismo C1, 26: misc, 27: ARN, 28: ADN, 29: proteína, 30: señalización, 31: célula, 33: desarrollo, 34: transporte, 35: sin asignación. En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad basada en logFC.



**Figura 29. Cambios en la expresión de la especie VNT sometida a sequía en T05.** En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad, basada en logFC.



**Figura 30. Cambios en la expresión de la especie VNT sometida a sequía en T07 y T10 (A y B respectivamente).** En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad, basada en logFC.

Hay una marcada expresión diferencial de genes en respuesta a la sequía en la mayor parte de los metabolismos analizados. Los procesos más afectados a lo largo del estrés hídrico en la especie VNT, están relacionados con la fotosíntesis, la pared celular, el metabolismo de lípidos, el metabolismo de aminoácidos, el metabolismo secundario, la acción de las hormonas, las proteínas de respuesta al estrés, la regulación de los procesos redox celulares, el procesamiento del ARN, el metabolismo de proteínas y la señalización.

A fin de explorar los datos que resultaron más relevantes para explicar la respuesta de esta especie frente al estrés hídrico, se describen seguidamente los cambios que se observaron en los principales procesos afectados.

## **Fotosíntesis**

Durante el período de estrés, la fotosíntesis es uno de los procesos que se ha visto más afectado por la disminución del agua disponible. Se observa una expresión diferencial de genes, con una dinámica muy particular ya que en T05, hay una proporción similar de los genes que se inducen y reprimen, mientras que en T07 todos los genes relacionados con la fotosíntesis se reprimen y concluyen en el tiempo 10 con una inducción generalizada de estos genes.

Particularmente en T05 se inducen las Subunidades de Polipéptidos del fotosistema II (Polipéptido de 10 kDa del fotosistema II, Proteína Ultravioleta-B-Reprimible, Proteínas W del centro de reacción del fotosistema II, Núcleo de proteínas complejas del fotosistema II) y I (Subunidades IV, IX, III del centro de reacción del fotosistema I), el transporte de electrones por medio de la Ferredoxina I del cloroplasto. En del Ciclo de Calvin se induce las Gliceraldehido-3-Fosfato Deshidrogenasas y Aldolasas como la Fructosa-bifosfato Aldolasa, excepto una de ellas (putativa Fructosa Bifosfato Aldolasa similar a Aldolasa Plastídica NPALDP1). En este período se reprimen en el transporte de electrones, la Ferredoxina-NADP Reductasa. En el Ciclo de Calvin, se reprimen Fosfoglicerato Kinasa y las Transcetolasas. Al transitar el día 07 del estrés, todos los genes afectados a este proceso se ven reprimidos, desde los que intervienen en el fotosistema I y II, así como en el transporte de electrones, las enzimas del ciclo de Calvin. En T07 también aparecen afectados genes involucrados en la fotorespiración, situación que no se había observado en T05, como por ejemplo las Fosfoglicolato Fosfatasas y las Proteínas de clivaje de Glicinas y Serina-Hidroximetiltransferasa que se reprimen. También se reprime la Ribulosa-3-Fosfato-Epimerasa y enzimas de interacción con Rubisco (Ribulosa 1,5-Bisfosfato Carboxilasa-Oxigenasa). De los pocos genes activados podemos mencionar a Citocromo C6.

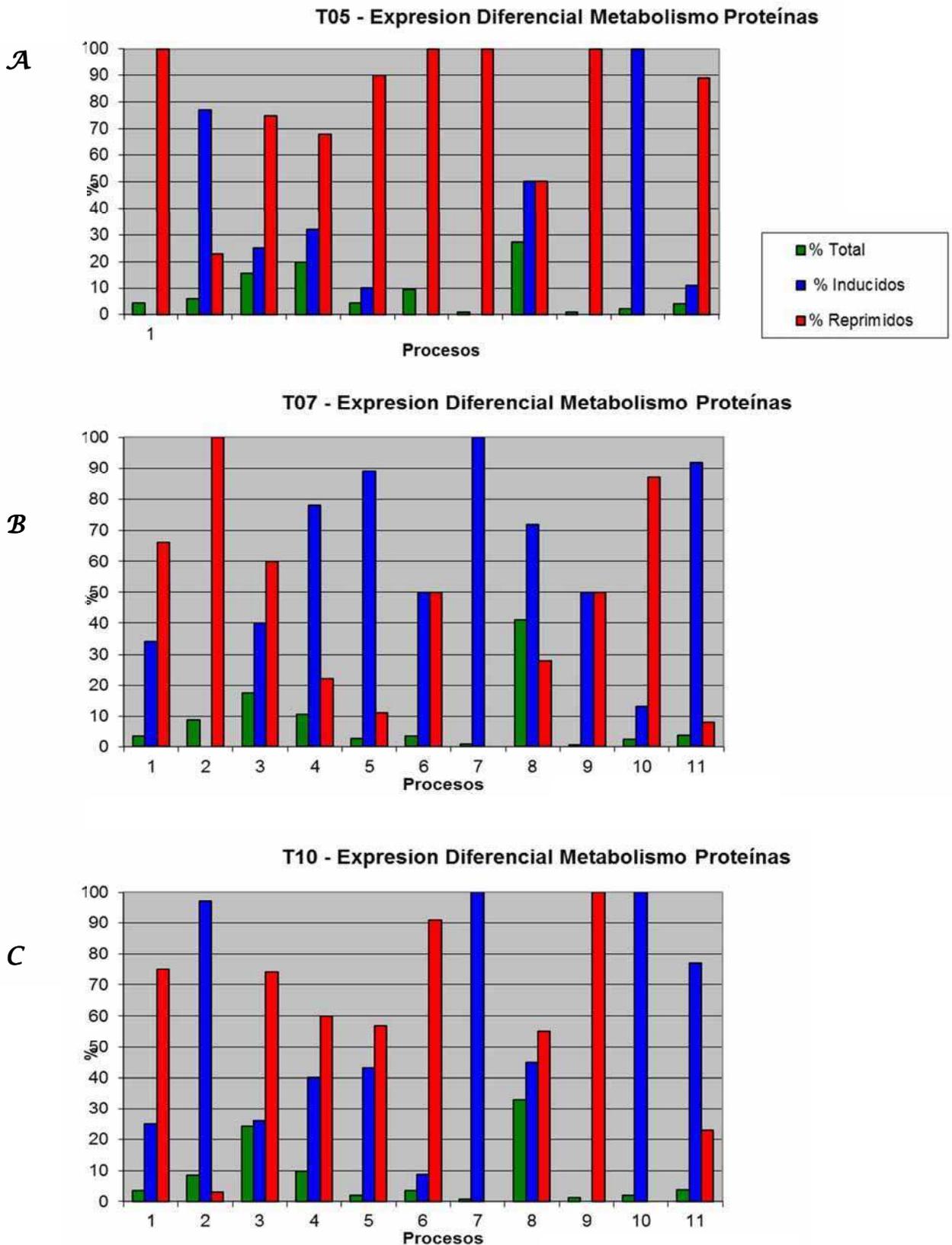
En T10 la situación es completamente opuesta a lo sucedido tres días antes, donde ahora los genes afectados a la fotosíntesis se encuentran todos inducidos como las subunidades de los fotosistemas I Y II, el Citocromo b6, las ATP Sintetas, el Citocromo C6, transportadores de electrones como Plastocianinas y Ferredoxinas. En relación a la fotorespiración, hay inducción de las proteínas de clivaje de Glicinas. En el ciclo de Calvin, vuelven a inducirse las Gliceraldehido-3-Fosfato Deshidrogenasas de T05, Aldolasa, Fructuosa-1,6-Bifosfatasa, Ribulosa-Fosfato-3-epimerasa y la Rubisco.

### **Metabolismo de proteínas**

De todos los procesos en donde existe expresión diferencial, es en el metabolismo de proteínas donde se observa una mayor cantidad de genes involucrados (Figura 31, A, B, C).

En T05 los genes de proteínas de activación de aminoácidos se encuentran reprimidos. Esto mismo se mantiene en T07 y T10 en alta proporción, pero se observa inducción de otros genes relacionados a la activación de aminoácidos conforme avanza el estrés.

Los genes involucrados en la síntesis de proteínas (factores de iniciación y elongación) y en las modificaciones postraduccionales muestran el mismo comportamiento general que las proteínas de activación de aminoácidos.



**Figura 31. Expresión diferencial relacionada con el Metabolismo de Proteínas en la especie VNT en T05, T07 y T10 (A, B y C respectivamente). En verde se muestra el % del total de genes diferencialmente expresados, en azul los genes inducidos (en %) y en rojo los reprimidos. Procesos: 1) Activación de aminoácidos, 2) Proteínas de organelas, 3) Regulación**

post-Traduccional, 4) Proteínas ribosomales, 5) Síntesis-Iniciación, 6) Síntesis-Elongación, 7) Síntesis-Release, 8) Degradación, 9) Glicosilación, 10) Ensamblaje, 11) Plegamiento

Los genes de proteínas cloroplásticas se muestran mayoritariamente inducidos en T05 y T10, mostrando en T07 una tendencia inversa.

Algunos de los genes que muestran mayores modificaciones de expresión de cada uno de estos procesos son:

- Activación de aminoácidos: Asparaginil-ARNt Sintetasa (reprimidos), Isoleucina-ARNt Sintetasa (reprimidos), Glicina-ARNt Ligasa (reprimido e inducido), Histidil- ARNt Sintetasa (inducidos).
- Síntesis de proteínas de cloroplasto: Proteínas Ribosomal L16. Ribosomal protein L1, L10, L4, L24, etc (inducidas en T05 y T10 y reprimidas en T07).
- Modificaciones post-traduccionales: las Proteín Kinasas, Fosfatasa 2A y 2B, Proteínas del complejo T y Caseín Kinasas.

Por su parte los inhibidores de la degradación tales como inhibidores Cisteín Proteasas se encuentran arrestados en todos los tiempos en respuesta a la sequía.

Los genes que intervienen en el plegamiento de proteínas están reprimidos en T05 y T07, entre ellos Proteínas GrpE y varios tipos de Chaperoninas. En T10 se observa la inducción de una gran proporción de los genes afectados a este proceso.

Las proteínas de ensamblaje, así como los factores de ligación están inducidos en T05 y T10 y marcadamente reprimidos en T07.

En T05 y T10 los genes que intervienen en la glicosilación están reprimidos, en T07 se encuentran inducidos y reprimidos.

## **Pared celular**

En los tres tiempos la expresión diferencial de los genes relacionados con la pared celular que intervienen en la síntesis, degradación y modificación, están mayoritariamente reprimidos (90, 80 y 75% de los genes observados en los tiempos T05, T07 y T10 respectivamente). A medida que avanza el tratamiento de estrés hídrico, se observa el aumento de genes inducidos como  $\alpha$ -

Expansina, Xiloglicano Endo-Transglicosilasa y los relacionados con precursores de la pared celular como UDP-Glc-4-Epimerasa y con proteínas estructurales ricas en Hidroxiprolina y Arabinogalactanas.

### **Metabolismo de Hidratos de carbono**

Los procesos bioquímicos de formación, ruptura y conversión de los hidratos de carbono (relacionados con la sacarosa y el almidón) se vieron afectados por el estrés. Los genes involucrados en la síntesis y degradación de hidratos de carbono no son en proporción elevados respecto del total de genes diferencialmente expresados, pero sí son clave en el aporte de energía a través de su metabolismo.

En T05, hay una represión general de los genes que intervienen en la síntesis como en la degradación de hidratos de carbono, mostrando la inducción específica de genes de Fructosa-1,6-Bisfosfatasa, enzima que interviene en la síntesis de sacarosa a partir de glicerol.

En T07, comienzan a inducirse más genes relacionados con el metabolismo del almidón como Isoamilasas, que intervienen en la desramificación del almidón. Asimismo se activan en este tiempo genes involucrados en la degradación de la sacarosa como Sacarosa Sintasa e Invertasa Neutra y del almidón, como  $\beta$  y  $\alpha$ -Amilasas y 4- $\alpha$ -Glucanotransferasas. En T10 el 70% de los genes se muestran reprimidos y los inducidos siguen la tendencia presentada en T07.

Esta dinámica de genes relacionados con el metabolismo de hidratos de carbono a lo largo del tratamiento de sequía puede observarse en la Figura 32 (A, B y C).

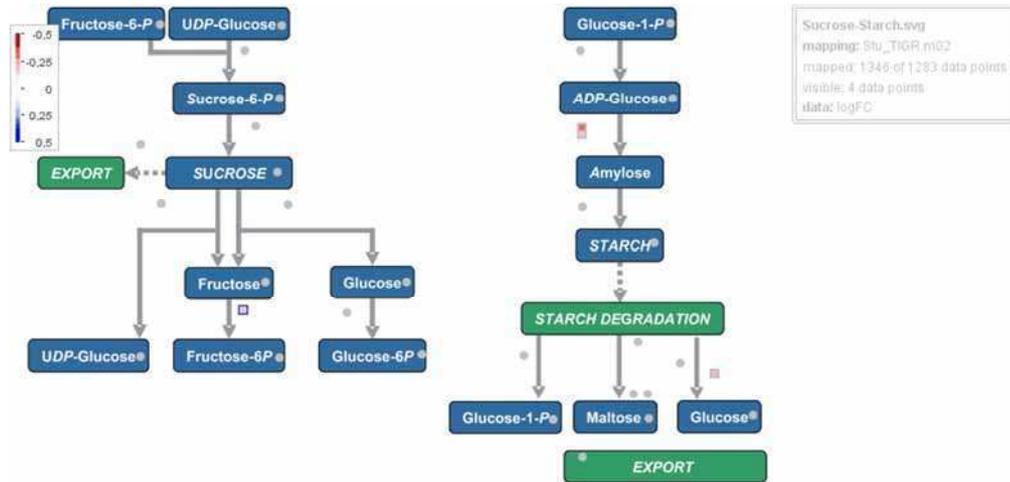
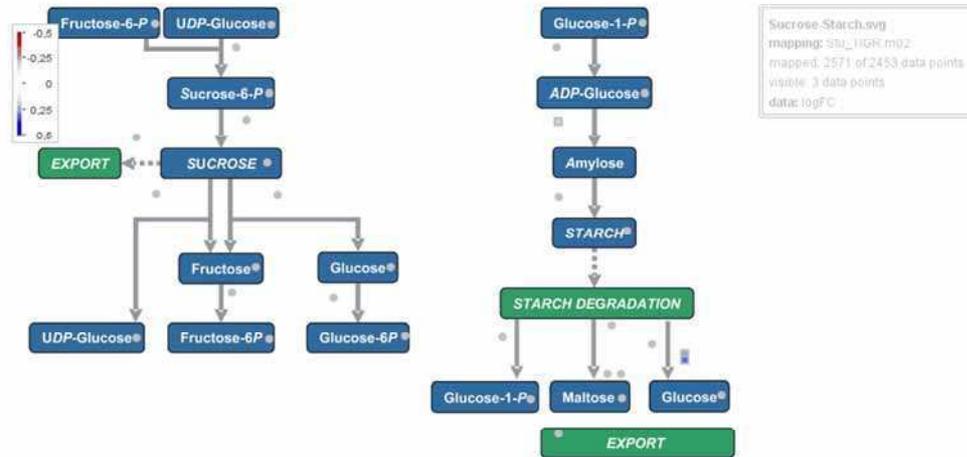
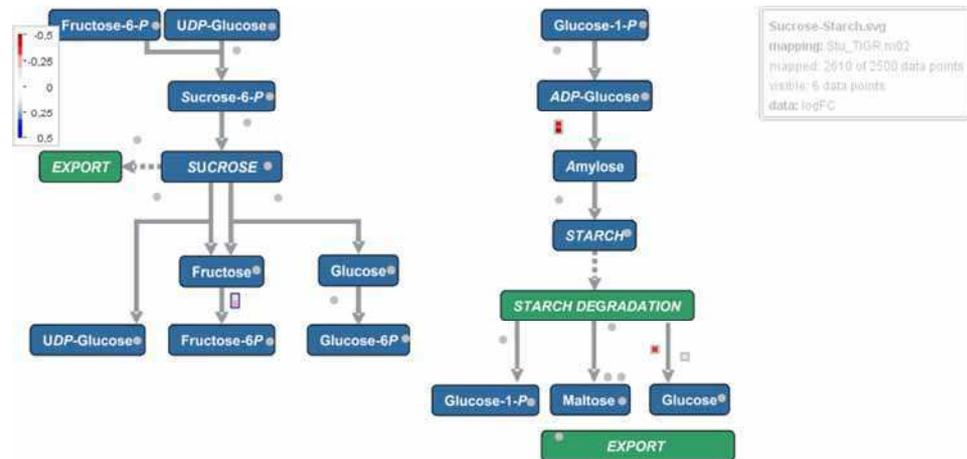
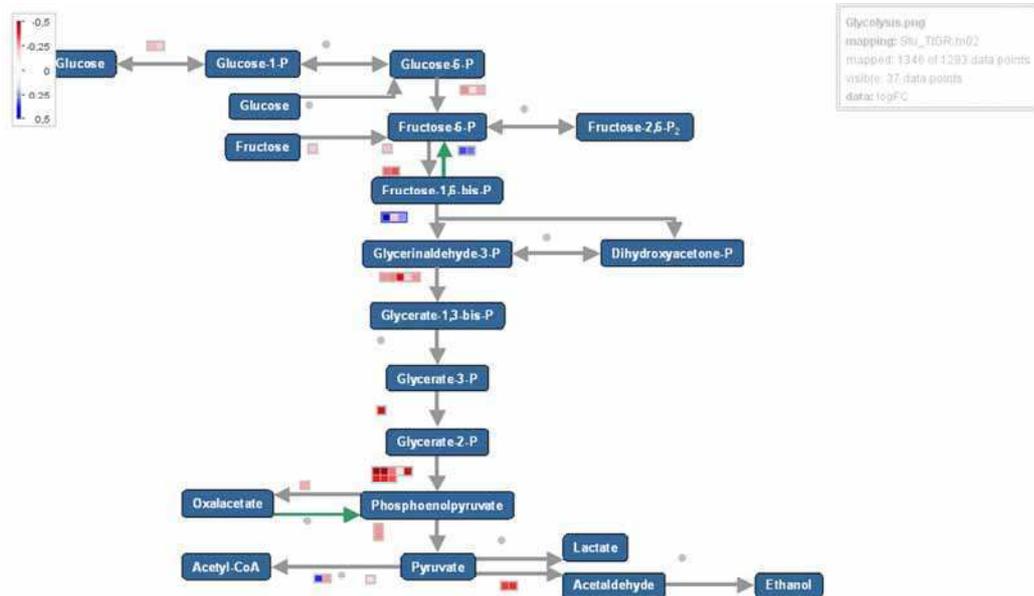
**A****B****C**

Figura 32. Expresión diferencial de genes relacionados con el metabolismo de hidratos de carbono, en particular de azúcares y almidón en T05, T07 y T10 (**A**, **B** y **C** respectivamente). En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad, basada en logFC.

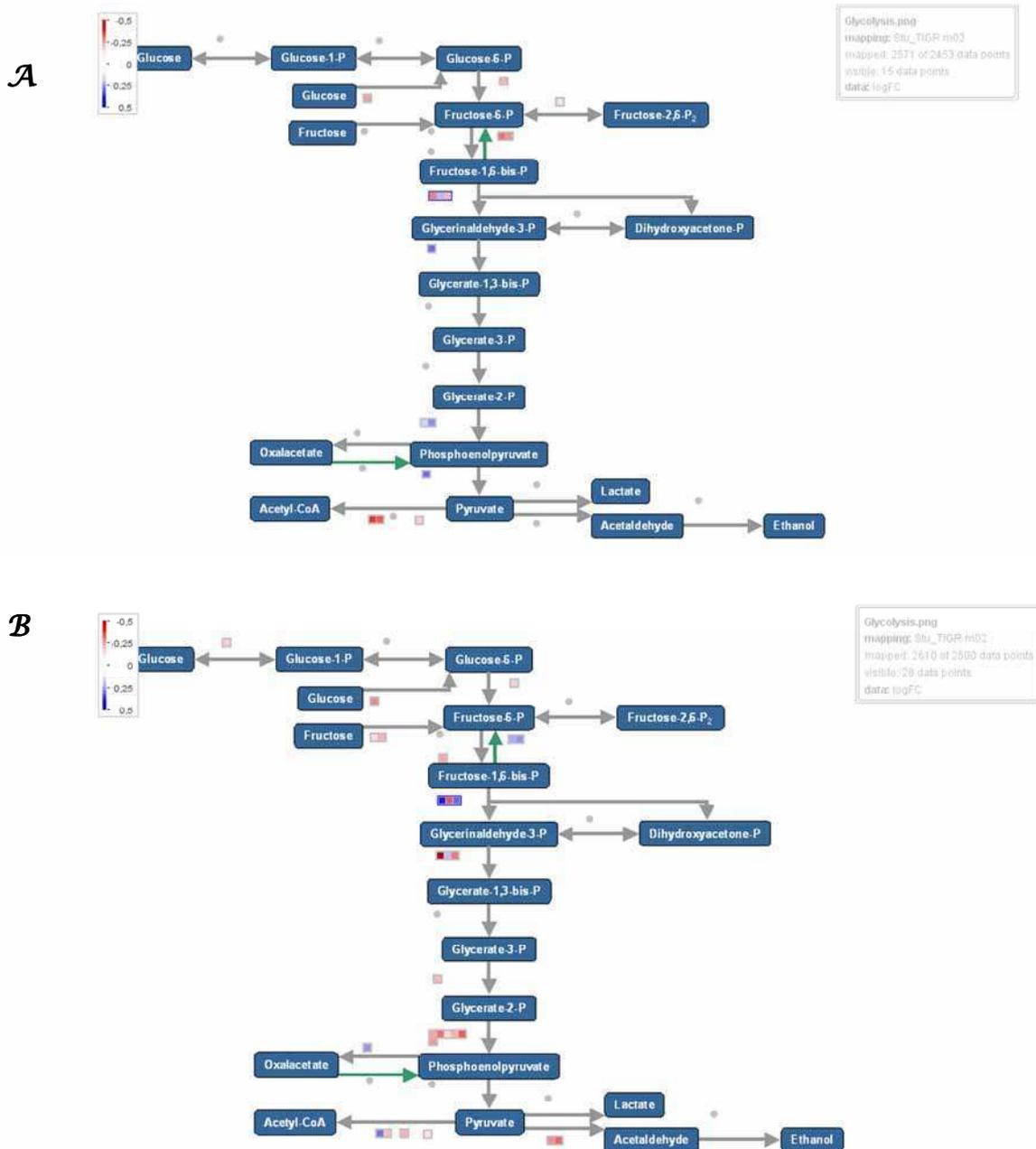
## Glucólisis

La vía metabólica encargada de oxidar la glucosa con la finalidad de obtener energía para la célula se vio afectada por el estrés hídrico.

En la glucólisis se ven comprometidos más genes en el tiempo 05 (Figura 33) observándose sub-expresión de genes tales como UDP-glucosa Pirofosforilasa, Fosfoglicerato Mutasa, Enolasa, Piruvato Kinasa, Fosfoenolpiruvato Carboxilasa, Glucosa-6-Fosfato Isomerasa, Fosfofructoquinasa y Glicerinaldehído 3-Fosfato Deshidrogenasa. Se induce la enzima Fructosa-Bisfosfato Aldolasa. En T07 la cantidad de genes involucrados es menor (Figura 34A), observándose inducción y represión de genes en partes iguales. Entre los inducidos se encuentran Enolasas, Piruvato Kinasas, Aldolasas y Glicerinaldehído-3-Fosfato Deshidrogenasa, manteniéndose reprimidas otras Aldolasas y Glucosa-6-Fosfato Isomerasa. En T10 persiste la tendencia de una represión mayoritaria de los genes que intervinieron en los otros tiempos (Figura 34B), induciéndose en este en particular Fosfoenolpiruvato Carboxilasa Kinasa2.



**Figura 33. Expresión diferencial de genes relacionados con el metabolismo de hidratos de carbono, en particular el proceso de glucólisis en T05. En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad, basada en logFC.**



**Figura 34. Expresión diferencial de genes relacionados con el metabolismo de hidratos de carbono, en particular el proceso de glucólisis en T07 y T10 (A y B respectivamente).** En rojo se muestran los genes reprimidos y en azul los inducidos en una escala cromática de intensidad, basada en logFC.

### Metabolismo de hormonas

Las hormonas participaron de la regulación de la respuesta al estrés. Los genes que intervienen en el metabolismo de las diferentes fitohormonas fueron incrementándose desde el tiempo 05 hasta el final del tratamiento. Inicialmente se produjo la activación del metabolismo del Ácido Abscísico como elementos de respuesta al Ácido Abscísico unidos al factor 2, de auxinas con Aldo/Keto

Oxidoreductasa, citoquininas con histidina con proteínas de fosfotransferencia), etileno (factor de unión a elemento de respuesta al etileno) y giberelinas (proteínas de respuesta a giberelinas). Por otro lado, los genes del metabolismo de brasinoesteroides presentes en mayor proporción en relación al total diferencialmente expresados, se presentan reprimidos (proteínas de elongación del óvulo / fibra celular, proteína de interacción con el receptor TGF-beta). En T07, el metabolismo de las auxinas y brasinoesteroides son los más sensibles, involucrando más genes en su metabolismo. Gran proporción de los genes del metabolismo de las auxinas están reprimidos, induciéndose específicamente Aldo/keto Oxidoreductasas y proteínas de respuesta a auxinas. Por su parte en el metabolismo de brasinoesteroides hay mayor proporción de genes inducidos. Se activa en este tiempo de forma diferencial el metabolismo del jasmonato, a través de S-adenosil-L-metionina: ácido salicílico carboxilo-metiltransferasa, lipoxigenasas (LOX1 y LOX2). Al final del tratamiento de sequía, los genes afectados al metabolismo hormonal se han incrementado marcadamente respecto del tiempo 05, mostrando genes inducidos y reprimidos en las distintas vías. Las vías que se observan más implicadas comprenden a las auxinas, brasinoesteroides y al etileno. En el metabolismo de brasinoesteroides, casi la totalidad de los genes están reprimidos, al igual que en el metabolismo del jasmonato y el ácido salicílico.

### **Transporte de iones metálicos**

Los genes asociados al transporte de iones metálicos que intervienen en la adquisición y quelación de éstos, cumplen un rol importante en la detoxificación celular bajo condiciones de estrés oxidativo. En todos los tiempos vemos una inducción generalizada de Metalotioninas, que pueden quelar zinc y otros metales y con sus residuos de cisteína pueden capturar especies reactivas de oxígeno como superóxido. En T05 y T07 se inducen también Hierro-reductasas y particularmente en T07 y T10, la Ferritina I.

Las Chaperonas de Cobre están particularmente reprimidas en T07 así como las Hierro-Reductasas en T10.

## **Proteínas LEA**

Genes que codifican para proteínas LEA, se inducen en todos los tiempos, aumentando paulatinamente su proporción a medida que avanza la sequía, algunas responden al etileno y otras inducen tuberización.

## **Proteínas de respuesta al estrés**

Varias proteínas asociadas al estrés biótico se vieron afectadas por la sequía, aumentando la proporción de genes diferenciales en el tiempo, como Cisteína proteasas, Endoquitinasa, Quitinasas, proteínas con motivos LRR y proteínas generales relacionadas con la patogénesis, cuya expresión varía a lo largo del estrés.

El número de genes que codifican proteínas de respuesta al estrés abiótico aumentó de T05 a T07 y se mantuvo similar en T10. Éstas proteínas están mayoritariamente inducidas en todos los tiempos con una tendencia creciente desde T05 a T10. Están reprimidos en todos los tiempos genes que codifican para proteínas que responde a la temperatura como Proteínas de Choque Térmico tipo Hsp90, Hsp70, Chaperonas y Proteínas pequeñas mitocondriales de Choque Térmico como HSP20. De las proteínas reportadas en respuesta al estrés hídrico y salino están reprimidos los genes que codifican para Proteínas Anquirinas, relacionadas con respuestas tempranas a la deshidratación del tipo 3 (del inglés early-responsive to dehydration stress protein ERD3) y Oligosacaril Transferasas.

Se inducen principalmente Osmotinas, Proteínas pequeñas de Choque Térmico HSP20, Proteínas asociadas a la respuesta temprana a la deshidratación del tipo ERD15 y RD22 y un complejo de proteínas inducidas por la luz.

## **Señalización**

En la señalización mediada por luz intervienen en proporciones similares genes que se inducen y reprimen en respuesta al estrés hídrico en todos los tiempos analizados como Fitocromo A y B2, Criptocromo 1B, Kinasas, Fotoliasas y Proteínas de inducción temprana por la luz.

Se reprimen en todos los tiempos genes de receptores de Kinasas como Serín/ Treonín Proteín-Kinasa y Lectín-Proteín Kinasa.

Los genes que gobiernan la señalización del calcio están diferencialmente expresados en todos los tiempos del estrés. Estos involucran genes con alta similitud a Calreticulinas (reprimidas en T05 y T10), a Calmodulinas (inducidas en todos los tiempos) y Proteínas de unión a calcio (inducidas en T05 y T10).

La señalización mediante proteínas G muestra sus genes principalmente arrestados en T05 y T10, con mayores niveles de inducción en T07.

En la señalización mediada por MAPKinasas se observó un comportamiento similar al de las Proteínas G.

### **Procesos Rédox celulares**

El número de genes diferencialmente expresados aumentan con el tratamiento de estrés hídrico. La Disulfuro isomerasa se reprime en T05 y T10, las Tiorredoxinas H inducidas en T05 y Tiorredoxina Peroxidasa inducidas en T07 y T10. Los genes que codifican para las Glutación Reductasas están principalmente reprimidos en todos los tiempos. Los genes con alta similitud con Glutarredoxinas están inducidos en T05 y T10.

Las Catalasas están reprimidas en T05 y T10. En T10 se inducen las Dismutas Superóxido Dismutasas y Fe-Superóxido Dismutasas.

## Validaciones mediante qPCR

Para validar experimentalmente los datos de los perfiles de expresión genética, se llevaron a cabo reacciones de PCR en tiempo real con 6 transcriptos selectos. Los transcriptos fueron elegidos sobre la base de señales de hibridación más intensas y valores de fold change más extremos. Se focalizó la validación en dos momentos del tratamiento de estrés, en donde los cambios de expresión según la especie fueron más marcados. Para CPH se seleccionaron los tiempos T05 y T07 y para VNT, T07 y T10 dado que la validación se realizó a partir de experimentos independientes (realizados en Argentina) y, por lo tanto, con ARNs diferentes de los utilizados para las hibridaciones de las micromatrices.

Previo a la realización de las amplificaciones en tiempo real se realizaron diluciones seriadas al décimo del templado de ARN de CPH y VNT utilizado como sustrato de las reacciones de PCR. Las reacciones de amplificación se realizaron por triplicado tanto con los transcriptos seleccionados así como con el gen utilizado como referencia, el factor de elongación 1 alfa (EF-1 $\alpha$ , elongation factor 1-  $\alpha$ ). Los controles negativos que se incluyeron fueron reacciones sin templado de ADNc para cada par de oligonucleótidos ensayados, incluyendo el factor de elongación 1-  $\alpha$ .

La cuantificación de los cambios relativos en la expresión génica se realizó utilizando el método  $2^{-\Delta\Delta CT}$  según lo descrito por Livak y Schmittgen (2001).

El análisis estadístico de los perfiles de expresión obtenidos en los ensayos por qPCR fue realizado con el programa fgStatistics (Di Rienzo, 2009).

Se encontraron patrones consistentes entre ambas metodologías. Es decir, los genes que se mostraron inducidos o reprimidos por estudios con micromatrices también lo hicieron en los ensayos de expresión de qPCR. Sin embargo, y como era de esperar por los antecedentes en la literatura, las diferencias cuantitativas absolutas entre las micromatrices y los datos de PCR en tiempo real sugieren que las relaciones de expresión (control/estrés) observadas en los experimentos con micromatrices no deben interpretarse como el exacto número

de veces en que se induce un gen. Aún así, los datos de micromatrices demostraron ser lo suficientemente consistentes con los validados por qPCR como para sustanciar los cambios más extremos en los patrones de expresión génica descritos en esta tesis.

Analizando cada transcripto en particular podemos observar, por ejemplo, el perfil de expresión del fragmento de la enzima Ferritina I (STMGY57) en la especie CPH mediante la metodología de micromatrices en la Figura 35. Allí se muestra una marcada inducción de la expresión de Ferritina I en las plantas sometidas a estrés (línea roja en la figura), respecto del control (línea verde). La validación por qPCR en los tiempos T05 y T07, dónde se puede monitorear este cambio, confirmó esta tendencia. El punto de expresión más extremo, T05, fue estadísticamente significativo con un p-valor de 0,0291.

Respecto de la proteína con motivo GDSSL-Lipasa/Hidrolasa (STMCK16), la misma se encuentra reprimida por efecto del estrés en la especie CPH. En la Figura 36 puede observarse el perfil de expresión generado en el ensayo de micromatrices. Mediante qPCR, se validó esta directriz en la expresión, que arrojó un p-valor de 0.0431 en T07. Es más, considerando un valor de significancia de 0.10, también es estadísticamente significativa la diferencia de expresión en T05.

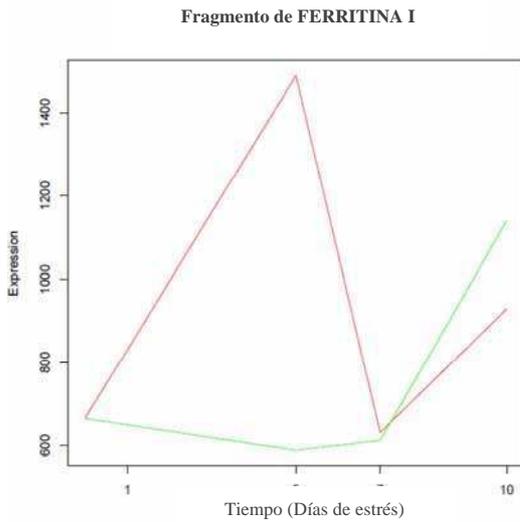
En relación al perfil de expresión diferencial en la especie VNT de la Proteína de Transferencia de Lípidos 1 (STMGN69) utilizando micromatrices, se observa un pico de inducción en el tiempo T07 (Figura 37). La validación por real time, mostró diferencias en los Cts entre el control y el material estresado, reflejando una tendencia conservada con las micromatrices. Estadísticamente estas diferencias fueron marginalmente significativas.

La Proteína de Transferencia de Lípidos No Específica de *Solanum tuberosum* ó StnsLPT (del inglés *Solanum tuberosum non-specif Lipid Trans Protein*) (STMGQ20) muestra en su perfil de transcripción una sobreexpresión en las plantas estresadas hídricamente en los tiempos T07 y T10, que se insinúa desde el tiempo T05 (Figura 38). La validación por real time PCR corrobora

estos resultados, con p-valores significativos de 0,0783 y 0,0592 para cada tiempo respectivamente (considerando p-valores significativos de 0,1).

La Dehidrina TAS14 (STMHQ27), presenta un perfil transcriptómico similar a la STnsLPT (Figura 39) y su expresión fue validada mediante qPCR con p-valores de 0,01 y 0,0084 en los tiempos T07 y T10, respectivamente.

Por su parte también se comprobó la expresión del gen dedos de zinc (C3HC4-tipo dedo RING) (STMDB72) (Figura 40), que no presentaba cambios en la expresión bajo condiciones de estrés en ninguna de las especies silvestres de papa tanto mediante micromatrices como mediante qPCR.



**Figura 35.** Gráfico del perfil de expresión del fragmento de Ferritina I mediante micromatrices. Líneas de color rojo corresponden a la expresión de plantas estresadas y de color verde a las plantas control. Se muestra la intensidad de expresión en el tiempo.

Genotipo\*Target= CPH:**Ferritin**

Technical replicates sampled at random

Resampling cycles=5000

Average CP's in Target and Reference

Treatment	Samples	Target	Reference
5:C	4	33.22	22.17
5:S	4	30.11	22.19
7:C	4	33.85	21.86
7:S	4	33.15	21.37

Target efficiency: Ef. Target1; Reference efficiency: Ef. Ref.

Ratio DCP.Target(Control-Sample)/DCP.Reference(Control-Sample)

Ctrl\Samp	S5:C	S5:S	S7:C	S7:S
C5:C		<b>9.91</b>	0.53	0.52
C5:S	0.10		0.05	0.05
C7:C	1.89	18.73		0.99
C7:S	1.91	18.94	1.01	

P-value matrix

Resampling cycles: 5000

Ctrl\Samp	S5:C	S5:S	S7:C
C5:S	<b>0.0291</b>		
C7:C	0.2749	0.0074	
C7:S	0.2893	0.0159	0.4962

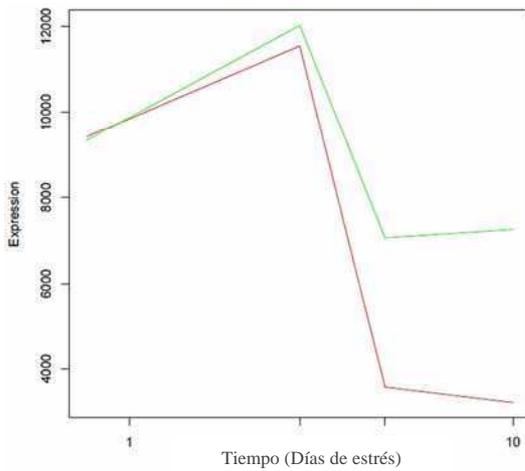
Summary

Resampling cycles: 5000

Ctrl	Sample	Ratio	SE	p-value
5:C	5:S	<b>9.91</b>	4.90	<b>0.0291</b>
5:C	7:C	0.53	0.29	0.2749
5:C	7:S	0.52	0.25	0.2893
5:S	5:C	0.10	0.18	0.0291
5:S	7:C	0.05	0.06	0.0074
5:S	7:S	0.05	0.05	0.0159
7:C	5:C	1.89	1.70	0.2749
7:C	5:S	18.73	10.38	0.0074
7:C	7:S	0.99	0.64	0.4962
7:S	5:C	1.91	6.38	0.2893
7:S	5:S	18.94	40.18	0.0159
7:S	7:C	1.01	2.52	0.4962

**fgstatistic**

Proteína con motivo GDSL – lipasa/hidrolasa



**Figura 36.** Gráfico del perfil de expresión de la proteína con motivo GDSL-Lipasa/Hidrolasa mediante micromatrices. Líneas de color rojo corresponden a la expresión de plantas estresadas y de color verde a las plantas control. Se muestra la intensidad de expresión en el tiempo.

Genotipo\*Target= CPH:GDSL

Technical replicates sampled at random

Resampling cycles=5000

Average CP's in Target and Reference

Treatment	Samples	Target	Reference
5:C	4	25.46	22.17
5:S	4	27.52	22.19
7:C	4	26.10	21.86
7:S	3	29.21	21.37

Target efficiency: Ef. Target1; Reference efficiency: Ef. Ref.

Ratio DCP.Target(Control-Sample)/DCP.Reference(Control-Sample)

Ctrl\Samp	S5:C	S5:S	S7:C	S7:S
C5:C		0.27	0.56	0.06
C5:S	3.71		2.06	0.21
C7:C	1.80	0.49		0.10
C7:S	17.69	4.77	<b>9.82</b>	

P-value matrix

Resampling cycles: 5000

Ctrl\Samp	S5:C	S5:S	S7:C
C5:S	0.0950		
C7:C	0.2367	0.1205	
C7:S	0.0398	0.1050	<b>0.0431</b>

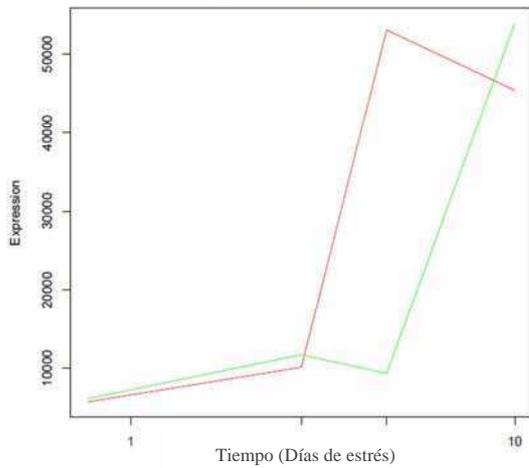
Summary

Resampling cycles: 5000

Ctrl	Sample	Ratio	SE	p-value
5:C	5:S	0.27	0.16	0.0950
5:C	7:C	0.56	0.44	0.2367
5:C	7:S	0.06	0.27	0.0398
5:S	5:C	3.71	5.26	0.0950
5:S	7:C	2.06	1.20	0.1205
5:S	7:S	0.21	0.72	0.1050
7:C	5:C	1.80	3.02	0.2367
7:C	5:S	0.49	0.26	0.1205
7:C	7:S	0.10	0.42	0.0431
7:S	5:C	17.69	3.50	0.0398
7:S	5:S	4.77	0.32	0.1050
7:S	7:C	9.82	1.05	0.0431

fgstatistics

Proteínas de Transferencia de Lípidos



**Figura 37.** Gráfico del perfil de expresión de la Proteína de Transferencia Lípidos 1 mediante micromatrices. Líneas de color rojo corresponden a la expresión de plantas estresadas y de color verde a las plantas control. Se muestra la intensidad de expresión en el tiempo.

```

Genotipo*Target= VNT:LIP TRANS

Technical replicates sampled at random

Resampling cycles=5000

Average CP's in Target and Reference

Treatment      Samples      Target      Reference
7:C             4            27.18       23.30
7:S             4            25.25       23.19
10:C            4            27.41       23.79
10:S            4            25.18       23.16

Target efficiency: Ef. Target1; Reference efficiency: Ef. Ref.

Ratio DCP.Target(Control-Sample)/DCP.Reference(Control-Sample)

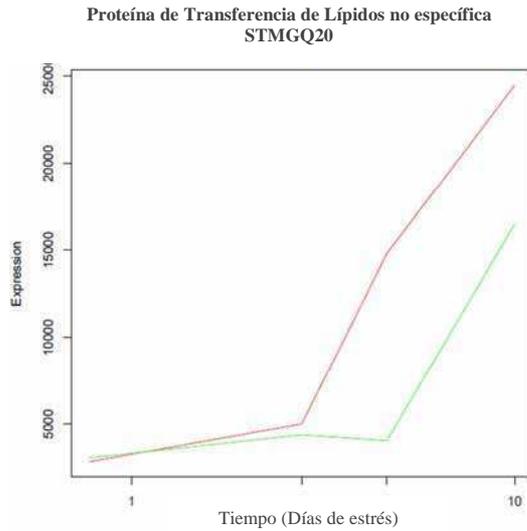
Ctrl\Samp      S7:C      S7:S      S10:C      S10:S
C7:C           3.22      1.15      3.32
C7:S           0.31      0.36      1.03
C10:C          0.87      2.72      2.81
C10:S          0.30      0.97      0.35

P-value matrix
Resampling cycles: 5000

Ctrl\Samp      S7:C      S7:S      S10:C
C7:S           0.1562
C10:C          0.4311    0.1298
C10:S          0.1374    0.4859    0.1020

Summary
Resampling cycles: 5000

Ctrl   Sample   Ratio   SE     p-value
7:C    7:S      3.22    10.23  0.1562
7:C    10:C     1.15    2.11   0.4311
7:C    10:S     3.32    3.23   0.1374
7:S    7:C      0.31    0.22   0.1562
7:S    10:C     0.36    0.43   0.1298
7:S    10:S     1.03    0.66   0.4859
10:C   7:C      0.87    0.65   0.4311
10:C   7:S      2.72    3.54   0.1298
10:C   10:S     2.81    1.09   0.1020
10:S   7:C      0.30    0.37   0.1374
10:S   7:S      0.97    2.54   0.4859
10:S   10:C     0.35    0.46   0.1020
    
```



**Figura 38.** Gráfico del perfil de expresión de la Proteína de Transferencia de Lípidos no específica mediante micromatrices. Líneas de color rojo corresponden a la expresión de plantas estresadas y de color verde a las plantas control. Se muestra la intensidad de expresión en el tiempo.

```

Genotipo*Target= VNT:STNLP
Technical replicates sampled at random
Resampling cycles=5000
Average CP's in Target and Reference
Treatment      Samples      Target      Reference
7:C            4            33.60       23.30
7:S            4            31.61       23.19
10:C           4            33.44       23.79
10:S           4            30.38       23.16

Target efficiency: Ef. Target1; Reference efficiency: Ef. Ref.
Ratio DCP.Target(Control-Sample)/DCP.Reference(Control-Sample)

Ctrl\Samp      S7:C      S7:S      S10:C      S10:S
C7:C           3.35      1.48      7.30
C7:S           0.30      0.44      2.18
C10:C          0.68      2.27      4.94
C10:S          0.14      0.46      0.20

P-value matrix
Resampling cycles: 5000

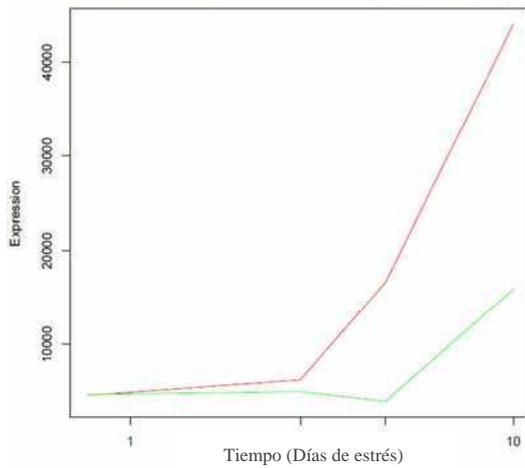
Ctrl\Samp      S7:C      S7:S      S10:C
C7:S           0.0592
C10:C          0.3465   0.2305
C10:S          0.0132   0.1312   0.0783

Summary
Resampling cycles: 5000

Ctrl      Sample      Ratio      SE      p-value
7:C       7:S         3.35       3.89    0.0592
7:C       10:C        1.48       2.47    0.3465
7:C       10:S        7.30       4.35    0.0132
7:S       7:C         0.30       0.04    0.0592
7:S       10:C        0.44       0.28    0.2305
7:S       10:S        2.18       0.55    0.1312
10:C      7:C         0.68       1.20    0.3465
10:C      7:S        2.27      10.43    0.2305
10:C      10:S        4.94      11.18    0.0783
10:S      7:C         0.14       0.04    0.0132
10:S      7:S         0.46       0.40    0.1312
10:S      10:C        0.20       0.25    0.0783
fgstatistics

```

Dehidrina TAS14



**Figura 39.** Gráfico del perfil de expresión de la Proteína inducida por estrés ambiental y ácido abscísico, Dehidrina TAS14, mediante micromatrices. Líneas de color rojo corresponden a la expresión de plantas estresadas y de color verde a las plantas control. Se muestra la intensidad de expresión en el tiempo.

```

Genotipo*Target= VNT:TAS14

Technical replicates sampled at random

Resampling cycles=5000

Average CP's in Target and Reference

Treatment      Samples      Target      Reference
7:C            4            28.01       23.30
7:S            4            18.24       23.19
10:C           4            29.14       23.79
10:S           4            22.68       23.16

Target efficiency: Ef. Target1; Reference efficiency: Ef. Ref.

Ratio DCP.Target(Control-Sample)/DCP.Reference(Control-Sample)

Ctrl\Samp      S7:C         S7:S         S10:C        S10:S
C7:C           3.4E-03      291.68       0.68         21.15
C7:S           1.46         425.84       2.3E-03      0.07
C10:C          0.05         13.79        0.03         30.88
C10:S          0.05         13.79        0.03         30.88

P-value matrix
Resampling cycles: 5000

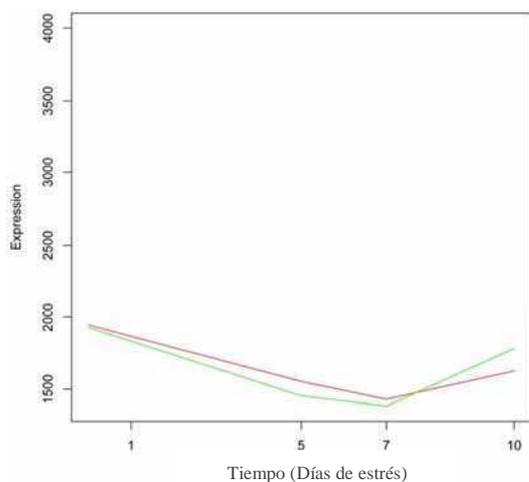
Ctrl\Samp      S7:C         S7:S         S10:C
C7:S           0.0100
C10:C          0.2802      0.0083
C10:S          0.0054      0.0594      0.0084

Summary
Resampling cycles: 5000

Ctrl  Sample  Ratio      SE      p-value
7:C   7:S      291.68     123.24  0.0100
7:C   10:C     0.68       0.60    0.2802
7:C   10:S     21.15      18.41   0.0054
7:S   7:C      3.4E-03    0.01    0.0100
7:S   10:C     2.3E-03    4.9E-03 0.0083
7:S   10:S     0.07       0.19    0.0594
10:C  7:C      1.46       2.77    0.2802
10:C  7:S     425.84     424.36  0.0083
10:C  10:S     30.88      54.50   0.0084
10:S  7:C      0.05       0.06    0.0054
10:S  7:S     13.79      11.93   0.0594
10:S  10:C     0.03       0.04    0.0084
    
```

fgstatistics

Proteína con motivo de dedos de zinc tipo C3HC4 RING



**Figura 40.** Gráfico del perfil de expresión de la proteína con motivo dedos de zinc tipo C3HC4 RING mediante micromatrices. Líneas de color rojo corresponden a la expresión de plantas estresadas y de color verde a las plantas control. Se muestra la intensidad de expresión en el tiempo.

```

Genotipo*Target= VNT:C3HC4

Technical replicates sampled at random

Resampling cycles=5000

Average CP's in Target and Reference

Treatment      Samples      Target      Reference
7:C             4            34.53       23.30
7:S             4            33.33       23.19
10:C            4            34.72       23.79
10:S            4            34.24       23.16

Target efficiency: Ef. Target1; Reference efficiency: Ef. Ref.

Ratio DCP.Target(Control-Sample)/DCP.Reference(Control-Sample)

Ctrl\Samp      S7:C      S7:S      S10:C      S10:S
C7:C           2.15      1.17      1.13
C7:S           0.47      0.54      0.53
C10:C          0.86      1.84      0.97
C10:S          0.88      1.90      1.03

P-value matrix
Resampling cycles: 5000

Ctrl\Samp      S7:C      S7:S      S10:C
C7:S           0.0765
C10:C          0.4172   0.1955
C10:S          0.4129   0.1273   0.4824

Summary
Resampling cycles: 5000

Ctrl  Sample      Ratio      SE      p-value
7:C   7:S           2.15      1.37    0.0765
7:C   10:C          1.17      1.24    0.4172
7:C   10:S          1.13      0.52    0.4129
7:S   7:C           0.47      0.31    0.0765
7:S   10:C          0.54      0.60    0.1955
7:S   10:S          0.53      0.27    0.1273
10:C  7:C           0.86      0.89    0.4172
10:C  7:S           1.84      1.98    0.1955
10:C  10:S          0.97      0.60    0.4824
10:S  7:C           0.88      0.79    0.4129
10:S  7:S           1.90      1.73    0.1273
10:S  10:C          1.03      1.74    0.4824
    
```

fgstatistics

## **6- DISCUSIÓN**

## 6. DISCUSIÓN

La papa cultivada (*Solanum tuberosum*), recurso decisivo para la seguridad alimentaria de cientos de millones de personas del mundo en desarrollo es muy susceptible a la sequía en comparación con otros cultivos (Weisz, 1994).

La respuesta al estrés por sequía en *S. tuberosum* es dependiente del cultivar, reflejando una variabilidad genética que puede ser aprovechada para descubrir mecanismos sólidos de resistencia al estrés hídrico (Levy, 1983, Jefferies, 1993, Heuer y Nadler, 1998). Levy identificó diferentes respuestas osmóticas a la sequía en los cultivares de *S. tuberosum* (Levy, 1992). Esto demuestra que hay diversidad de mecanismos de adaptación y aclimatación presente en el germoplasma de *Solanum*.

Las especies silvestres de papa representan un vasto recurso genético de caracteres de interés agronómico que contribuyan a desarrollar variedades de papa más robustas para otras regiones (Contreras, 2008). La disponibilidad de germoplasma es imprescindible para el mejoramiento genético ya que se requiere de variabilidad para aquellos caracteres que limitan la producción del cultivo de papa (Ross, 1986). La diversidad genética de papa, conservada *ex situ* en los bancos de germoplasma así como en otras colecciones, constituye un importante recurso fitogenético con el que se cuenta para afrontar el impacto del cambio climático y de numerosos estreses bióticos a los que se enfrenta el cultivo (Coleman, 2008, Ross, 1986; Jansky y Peloquin, 2005; Arvin y Donnelly, 2008; Levy y Veilleux, 2007, Contreras, 2008; Bamberg y del Río, 2010, Palta, 2010). El banco del Instituto Vasco de Investigación y Desarrollo Agrario de Neiker-Tecnalia, está nutrido por el material proveniente de las principales colecciones mundiales de germoplasma de papa como Dutch – German Potato Collection (BGRC), IPK Genbank Groß Lüsewitz (GLKS) y Commonwealth Potato Collection (CPC).

Bradshaw (2007) menciona que las futuras iniciativas de mejoramiento de la resistencia de las papas cultivadas frente a estreses abióticos y bióticos deben considerar una mayor utilización de especies silvestres en el mejoramiento de la papa, dada la amplia gama de hábitats en los que han evolucionado. Los mejoradores deberían aprovechar la diversidad que existe en las principales

colecciones de germoplasma de papa en el mundo para hacer frente a estos desafíos, utilizando las herramientas que permitan encontrar este potencial genético en el germoplasma disponible.

Entre la papa y sus congéneres silvestres se han reportado más de 200 especies (Hawkes, 1990, Clausen, 2003, Hijmans *et al*, 2002).

Las papas silvestres crecen en 16 países, desde el suroeste de los Estados Unidos hasta la costa central de Chile y regiones adyacentes de Argentina. Hay una importante riqueza de especies en el norte de Argentina, en el centro Bolivia, el centro de Ecuador, el centro y norte de México, y en el sur y centro-norte de Perú (Hijmans *et al* y Spooner, 2001) y las diferencias en su sitio de origen genera una gran variabilidad en la capacidad de adaptarse o aclimatarse a la sequía, al frío o condiciones salinas extremas (Correll, 1962). Estos datos muestran que el comportamiento diferencial de las especies de *Solanum* justifica ampliamente el utilizarlas en la búsqueda de resistencia a la sequía.

Con este objeto de evaluar germoplasma disponible de especies silvestres de *Solanum* que tuberizan, se estudiaron papas originarias de distintas regiones de América como: *S. berthaultii*, *S. gandarillasii*, *S. leptophyes*, *S. doddsii*, *S. infundibuliforme* (Bolivia), *S. sparsipilum* (Bolivia y Argentina), *S. bulbocastanum*, *S. polytrichon*, *S. papita*, *S. cardiophyllum* (México), *S. sanctae-rosae*, *S. marinasense* (Perú), *S. pascoense* (Perú y Ecuador), *S. venturii* (Argentina) (Tabla 3), bajo condiciones controladas en uno de los invernaderos de la Granja Modelo Arkaute de Neiker-Tecnalia, a fin de seleccionar nuevas especies con capacidad de adaptación a la sequía y para identificar material base para el mejoramiento genético.

La sub-colección de genotipos de *Solanum* utilizada en este estudio, no había sido previamente analizada por su comportamiento frente al estrés hídrico. Este material biológico era de interés en las investigaciones científicas del grupo y había sido evaluado frente a diferentes patógenos de papa como *Phytophthora infestans*, *Globodera pallida* y los virus de la papa Y<sup>N</sup>, del enrollado de la hoja, M y el virus S (Ruiz de Galarreta, 1998).

El presente estudio se focalizó en estudiar la respuesta temprana a la sequía implementada por especies silvestres de papa, para lo cual se sometió el material al estrés por supresión del riego por un período de 10 días. Una estrategia similar había sido implementada por TIGR para la evaluación de

solanáceas frente a estrés hídrico

([ftp://ftp.tigr.org/pub/data/s\\_tuberosum/SGED//055\\_TIGR\\_drought1/055\\_Study\\_Description.pdf](ftp://ftp.tigr.org/pub/data/s_tuberosum/SGED//055_TIGR_drought1/055_Study_Description.pdf)). Los ensayos de sequía con el germoplasma silvestre de *Solanum* se realizaron en el Instituto Vasco de Investigación y Desarrollo Agrario, Neiker-Tecnalia, de Vitoria-Gasteiz, España, en el marco del Proyecto Europeo UE (QLK5-CT-2002-01849) “*Developing a physical and functional map of potato: creating new sources for molecular markers to breed cultivars with multiple resistances and quality traits (APOPHYS)*”.

Dentro de las especies silvestres de papa, encontramos una gran diversidad morfológica y fenotípica, que dista mucho de la papa cultivada que conocemos y que se caracteriza principalmente por hojas y folíolos más pequeños, tallos más numerosos y pequeños tubérculos (Rousselle *et al*, 1999, Spooner y Salas, 2006).

Las especies silvestres de *Solanum* evaluadas mostraron una respuesta diferente frente al tratamiento de sequía. La gran plasticidad frente a la sequía mostrada por el material evaluado puede observarse en las Figuras 7 (A, B y C) y Figura 8. Las papas silvestres mostraron valores de RWC que oscilaron entre el 60% y el 99%, prevaleciendo en el germoplasma evaluado, las especies con RWC asociadas a un buen comportamiento frente a la sequía.

De todo el material evaluado, se seleccionaron en particular para estudios transcripcionales posteriores tres especies de *Solanum* que presentaron un RWC alto, medio y bajo al finalizar el tratamiento de sequía (Tabla 17, Figura 8), a fin de tener información sobre estrategias contrastantes implementadas por las especies silvestres frente al estrés. Los genotipos seleccionados en base a este criterio fueron: *S. pascoense*, *S. cardiophyllum* y *S. venturii*, con un RWC de 99, 87 y 68% respectivamente. Estos materiales en particular provienen de zonas geográficas muy distintas, lo que hace suponer que las diferencias en la respuesta al estrés muestran la adaptación al medioambiente de origen a través de la expresión de genes clave para hacer frente en forma exitosa a la sequía (Coleman, 2008).

Las condiciones de cultivo que limitan el crecimiento de las plantas, restringen también la productividad (Boyer, 1982). Los períodos de sequía son una característica inevitable y recurrente de la agricultura mundial (Bazzaz y

Sombroek, 1996, Reynolds, 2010, Ahuja *et al*, 2010). Una de las estrategias para contribuir a solucionar la problemática mundial que la sequía ocasiona en la seguridad alimentaria de la población, se basa en la búsqueda de nuevos cultivares que presenten una mayor eficiencia en el uso de agua y nutrientes (Hawkes, 1978, Blum, 2005).

El estrés hídrico ha sido estudiado a nivel fisiológico y por métodos moleculares (Ingram y Bartels, 1996, Liu *et al* 2006, Vos y Haverkort, 2007, Coleman, 2008, Vinod *et al*, 2006, Legay, 2011). Los efectos de la sequía se han correlacionado con la expresión de ciertos genes mediante el uso de estudios de gen único y datos de micromatrices (Alscher *et al*, 1997, Watkinson *et al*, 2006, Evers *et al*, 2010, Sakuma *et al*, 2006). Se han descrito genes que responden a la sequía a nivel transcripcional (Bohnert *et al*, 1995, Ingram y Bartels, 1996; Bray, 1997; Shinozaki y Yamaguchi-Shinozaki, 1997, 1999, 2000, Avni, *et al*, 2008, Watkinson *et al*, 2006). Los genes inducibles por el estrés se han utilizado para mejorar la tolerancia al estrés de plantas mediante transferencia de genes (Holmberg y Bülow, 1998; Bajaj *et al.*, 1999, Xiaolan, 2009, Abogadallah *et al*, 2011).

Se sabe muy poco acerca de los eventos iniciales relacionados con la percepción de las señales de estrés (Urao *et al.* 1999, Ueguchi *et al*, 2001; Wohlbach *et al*, 2008). Entre las vías de señalización, se han distinguido las dependientes de ácido abscísico (ABA) y las independientes del ABA (Shinozaki y Yamaguchi Shinozaki, 1997, 2007). El ABA se ha visto implicado en las respuestas de estrés por sequía donde su actividad está modulada por factores de transcripción como bZIPs (Uno *et al*, 2000, Kang *et al*, 2002) y MYB (Abe *et al*, 2003). La represión de los factores de transcripción es un factor importante de la tolerancia al estrés, como las proteínas de dedos de zinc tipo Cys2/His, que han sido descritos en condiciones de sequía, frío y estrés salino en *Arabidopsis* (Sakamoto *et al*, 2004). *Thellungiella halophila*, pariente de *Arabidopsis* tolerante a la salinidad, presenta mayores niveles constitutivos de expresión de los genes de respuesta a sequía y sal de *Arabidopsis* y, además contiene genes que son miembros parálogos de las familias de genes de *Arabidopsis* de respuesta a estrés (Inan *et al*, 2004 y Taji *et al*, 2004). Estas diferencias en la resistencia al estrés entre especies relacionadas, estarían

asociadas con una mayor presencia de dichos genes parálogos y/o un mayor grado de preparación mediante la expresión constitutiva de genes de defensa. La papa es muy susceptible a la sequía en comparación con otros cultivos (Weisz, 1994). El estrés hídrico es reconocido mundialmente como la mayor limitante a nivel mundial de la producción de papa (Schapendonk *et al* 1989, Van Loon, 1981). La falta de agua reduce el número y tamaño de hojas, la altura de los brotes (Deblonde y Ledent 2001) y la tasa fotosintética (Coleman, 2008). Todas las etapas de desarrollo son susceptibles al estrés hídrico, desde la emergencia hasta la iniciación del tubérculo, del llenado y el crecimiento de mismos. La exposición disminuye el desarrollo del tubérculo, aún en casos de cortos períodos de sequía, dependiendo de la oferta de hidratos de carbono desde el follaje que se vea afectada en estas condiciones de estrés por la reducción en la tasa fotosintética (Jefferies y MacKerron, 1989).

Es importante considerar que las variaciones en el contenido de agua en las plantas son claves para limitar los efectos negativos del estrés hídrico en su metabolismo. Cuando la tasa de transpiración es alta, la humedad del suelo baja y si el contenido de agua de la planta decrece un 10%, el crecimiento se vuelve más lento, pudiendo eventualmente cesar. Si decrece un 20%, la planta se marchita y los principales procesos se vuelven más lentos o se detienen y si llega a decrecer un 40%, la planta usualmente muere (Vos y Haverkort, 2007, Coleman, 2008). Estos materiales buscan cubrir los rangos de RWC asociados a diferentes estrategias de las plantas frente al estrés.

En los ensayos de estrés se incluyó a modo de referencia la variedad Kennebec de *S. tuberosum*, para monitorear en las papas cultivadas el estatus hídrico de las mismas durante el tratamiento y poder compararlas con las observadas en las especies silvestres. Como se observa en la tabla 17, Kennebec se comportó entre el grupo de las papas evaluadas más sensibles al estrés.

Es difícil consensuar un criterio que defina cuándo una planta es tolerante a la sequía. El concepto de tolerancia se redefine constantemente. Desde el punto de vista agronómico, podemos definir la tolerancia de un cultivo al estrés hídrico como la capacidad generar un rendimiento económicamente rentable aún con condiciones limitantes de agua.

Con las especies PSC, CPH y VNT se realizaron estudios transcripcionales monitoreando el estrés hídrico en distintos momentos del tratamiento utilizando la técnica de micromatrices de ADNc.

El procesamiento y análisis de las micromatrices se realizó durante los años 2006 y 2007 en la Universidad de Salamanca y Granada de España. Posteriormente se realizaron las validaciones mediante qPCR y ensayos biológicos adicionales en INTA Castelar y Balcarce respectivamente.

Las micromatrices de ADN constituyen una de las tecnologías disponibles más utilizadas para el análisis funcional de los genomas a gran escala. Existen dos tecnologías básicas para la producción de micromatrices de ADN, una de ellas utiliza como sustrato moléculas de ADNc y la otra oligonucleótidos sintéticos, que fueron descritas en detalle en la Introducción. A partir de cualquiera de ellas, se genera una matriz de expresiones génicas cuyo tratamiento estadístico es común en cualquier plataforma y de allí el carácter general de los instrumentos de análisis. Las micromatrices utilizadas en el presente trabajo son de ADNc. Al momento de realizar los estudios transcripcionales, este tipo de micromatrices eran las únicas disponibles comercialmente. Los trabajos publicados disponibles en la actualidad sobre transcriptómica en papa han utilizado ambos tipos de micromatrices (Watkinson *et al*, 2006, Schafleitner *et al*, 2007, Kloosterman *et al*, 2008, Vasquez-Robinet *et al*, 2008 Evers *et al*, 2010)

Las mayores limitaciones en la utilización de estas tecnologías giran en torno del análisis de datos (sin dejar de considerar el diseño del experimento y el número de réplicas). Aún cuando en los últimos años se han hecho muchas contribuciones en este campo, no existen soluciones ampliamente aceptadas y experimentos comparativos dentro y entre plataformas revelan la variabilidad inherente a estas tecnologías tanto a nivel de réplicas técnicas como biológicas, haciéndose cada vez más evidente la dificultad del análisis e interpretación de sus resultados (Michiels *et al*, 2005).

En el marco del Proyecto APOPHYS se inició el estudio de sequía con la especie PSC. Mediante la selección del Proyecto denominado "*Comparative Solanum wild species transcription profiles under controlled water stress conditions*", financiado por el Servicio de Perfiles de Expresión del Instituto

TIGR y la NSP (*National Science Foundation*) fue posible incluir las otras dos especies: CPH y VNT para tener una visión amplia de la respuesta a la sequía de especies silvestres de *Solanum* de diverso origen geográfico.

Con PSC, la primera de las especies ensayadas, se diseñó un experimento que cubrió los seis momentos del tratamiento de sequía T00, T01, T03, T05, T07 y T10 (corresponden a días de iniciado el estrés), utilizando un total de 12 micromatrices, que contemplaban en el diseño la implementación de *dye-swap* (Tabla 8). Éstas fueron procesadas en el laboratorio del Dr. Angel Dominguez (Instituto de Microbiología Bioquímica de la Universidad de Salamanca) y escaneadas en un Centro de la misma Universidad, resultando micromatrices de buena calidad, con buena señal y sin artefactos (Figura 11).

CPH y VNT, por su parte fueron procesadas por el Servicio de Expresión de TIGR en paralelo a la realización de los ensayos con la especie PSC. Dado que se contaba con mayor disponibilidad de micromatrices para realizar el estudio, se hicieron unas consideraciones en el diseño del experimento a fin de hacerlo más balanceado que implicaron la inclusión de un mayor número de réplicas técnicas y biológicas, acotando el número de tiempos evaluados a cuatro T00, T05, T07 y T10 (Tablas 9 y 10).

El ARN obtenido de las especies PCS, VNT y CPH (de las plantas control y sometidas a sequía), fue de excelente calidad (Figura 10), acorde para estudios de este tipo. Las extracciones de ARN de las tres especies se realizaron en forma conjunta. Para los estudios transcripcionales con CPH y VNT, el ARN fue enviado para su procesamiento al Servicio de Expresión de TIGR (USA). Previo a la utilización del ARN en las hibridaciones de las micromatrices (Salamanca como TIGR), la calidad y la integridad del mismo fue verificada como excelente para estos estudios.

De las hibridaciones se obtuvieron imágenes escaneadas que se procesaron para convertirlos en valores de expresión con el Programa GenePix Pro 5.0.0.49 (Axon).

Las micromatrices de la especie PSC (Salamanca) fueron pre-procesadas en el laboratorio hasta la obtención de los datos de expresión (archivos GPR). Una vez obtenidas las matrices de expresión en ambos canales, se inició el procesamiento de las micromatrices mediante un análisis estadístico y bioinformático de los datos.

Se analizaron 1.427.712 datos de expresión correspondientes a la respuesta al estrés hídrico de 15.264 ADNcs depositados por duplicado sobre los 44 soportes de cristal de papa, que monitoreaban 10 días de evolución del estrés hídrico en las especies CPH, VNT y PSC (Figura 11 y Material Complementario).

El análisis exploratorio inicial mediante componentes principales de los datos de expresión de las micromatrices de la especie PSC, reveló inconsistencias estadísticas en los mismos que llevaron a tomar la decisión de no ser consideradas en el análisis comparativo, por la información equivocada que se pudiera inferir del análisis de los mismos. El análisis de componentes principales mostró que en el tiempo 00, una de las repeticiones no se comportaba como lo esperado, demostrando una variabilidad no correlacionada con el inicio del tratamiento de sequía (Figura 24). Considerando la eliminación de esa repetición, el análisis estadístico revela que los p-valores son marginalmente significativos, aún implementando un procedimiento bayesiano para aumentar la fiabilidad estadística. Sobre la base de los otros 5 tiempos (10 micromatrices), los resultados de la especie PSC siguen mostrando falta de coherencia. En cada tiempo son pocos los datos obtenidos por gen para estimar correctamente su expresión.

La falta de repeticiones en cada tiempo dificulta sortear los problemas encontrados en algunas de las réplicas y es muy difícil tener una buena estimación de lo que está ocurriendo en el perfil temporal. Por otra parte no es posible asumir un modelo simplificado de la evolución en el tiempo que permita un modelado de la evolución temporal.

Los procesos de normalización aplicados a la matriz de datos de expresión, no llegaron a corregir esta situación posiblemente atribuible a errores sistemáticos vinculados con varios factores inherentes al material biológico, a la técnica de micromatrices y al tratamiento de la imagen, como por ejemplo diferencias en la eficiencia de la incorporación de los tintes en el proceso de marcación, diferencias en la cantidad de ARNm, diferencias en los parámetros de escaneado, diferencias entre los grupos de agujas, efectos espaciales etc. Muchos trabajos han tratado este tema (Murphy 2002, Sasik *et al*, 2004, Lee y Saeed, 2007).

Dado que el procesamiento del material biológico (ARN) utilizado en los ensayos de micromatrices fue tratado de igual manera y su calidad e integridad controlada mediante distintas técnicas, las variaciones deberían ser atribuibles a problemas en las hibridaciones o en la adquisición de datos, estimando, dada la calidad de las hibridaciones que haya habido algún problema en el paso de conversión de la imagen en valores de expresión.

Por su parte, la normalización de las micromatrices de CPH y VNT no mostró problemas y logró un ajuste adecuado de los datos de expresión para los análisis posteriores. La literatura discute mucho sobre la necesidad de realizar la normalización de los datos de expresión y sobre las estrategias para llevarlo a cabo, no llegando a un consenso sobre una técnica en particular, sino que deben evaluarse las diferentes opciones a fin de seleccionar la que mejor se adapte a las micromatrices utilizadas (Quackenbush, 2002; Smyth y Speed 2003). En este caso, tras probar diferentes alternativas de normalización de los datos (vs<sub>n</sub>, loewss, etc) y consultar con los expertos, (Beissbarth, T., comunicación personal, 2006), se aplicó una normalización local, denominada *print-tip-lowess*. Por su parte, para uniformar la distribución de las intensidades de los canales rojo y verde de todos las micromatrices se normalizaron con el método Aquantile (Yang *et al*, 2002, Smyth y Speed, 2003). El efecto de la normalización aplicada es muy claro y se observa en la las Figuras 12, 13 y 14.

Se realizó un estudio estadístico basado en componentes principales con el objeto de analizar los datos de expresión diferencial frente a sequía de las micromatrices de las especies CPH y VNT desde una perspectiva multivariada. Éste análisis mostró que al relacionar la ubicación de los genes en el espacio de los componentes principales, con su expresión en los distintos tiempos de evaluación, el tiempo T00 en ambas especies tiene un peso casi nulo en la conformación de los componentes principales 1, 2 y 3 (Figura 15 A-B y 16 A-B). Estos resultados muestran el diseño coherente del experimento dado que no debería haber cambios en la expresión génica entre las plantas estresadas hídricamente y sus controles al inicio del experimento.

A medida que avanza el estrés, los tiempos se van proyectando en el espacio de los componentes principales de manera diferencial, revelando una

respuesta al estrés pero con una dinámica de respuesta al estrés dependiente de la especie (Figuras 15 A-B y 16 A-B).

Al estudiar el comportamiento de los genes diferencialmente expresados surgido del ordenamiento espacial mediante componentes principales se vio un comportamiento en grupos de los mismos. Basados en los perfiles de expresión, se pudieron identificar en CPH, al menos 5 conglomerados que resumen las principales tendencias del comportamiento de los transcritos frente al estrés (Figura 18, 19). Por otro lado, la dinámica de expresión de los transcritos diferenciales de VNT, mostró expresiones contrastantes y agrupadas, que sugieren una regulación común sincrónica (como se ve en las Figuras 21 y 22) confirmando la diversidad genética de respuestas en el germoplasma de especies silvestres de papa.

El análisis bioinformático de los datos de expresión de las especies CPH y VNT arrojó diferencias significativas en los perfiles transcripcionales de las especies estudiadas. El análisis comparativo de los perfiles de expresión de estas dos especies silvestres de *Solanum* reveló que en CPH se ven afectados por el tratamiento de sequía un total de 235 genes durante los cuatros tiempos monitoreados del tratamiento de sequía, mientras que en VNT la diferenciación involucra 4133 genes. Ambas especies presentan una expresión común de 151 genes principalmente centrada en de expresión diferencial por la sequía (Tabla 25). Las respuestas al estrés dependientes de especie incluyen 57 genes en CPH (Tabla 26) y 3967 genes en VNT (Material Complementario).

Los genes seleccionados según los patrones de expresión diferencial más extremos frente al tratamiento de sequía, fueron validados por qPCR. Si bien son pocos los genes validados al momento de presentar este trabajo, los resultados son muy prometedores y validan los análisis realizados con los datos de expresión diferencial de las micromatrices (Figuras 35, 36, 37, 38 y 39).

Dado que CPH no presenta diferencias fenotípicas muy marcadas respecto del no tratado y se expresan bajo condiciones de sequía sólo 235 genes, estos se presentan como interesantes genes candidatos para ser evaluados como nuevos genes de resistencia a la sequía. En el futuro se validarán los mismos en los otros genotipos que mostraron altos índices de RWC, de los cuales ya se está preparando material.

Los genes expresados diferencialmente en cada especie se analizaron con el Programa MapMan (Thimm *et al*, 2004). Este programa permite organizar la información de matrices de datos expresión génica según la ontología establecida por MapMan, en grupos funcionales y procesos. Esta herramienta desarrollada inicialmente para trabajar con datos de expresión de *Arabidopsis* se fue actualizando (Usadel *et al*, 2005) y se logró adaptar para su utilización en solanáceas (Urbanczyk-Wochniak *et al*, 2006, Rotter *et al*, 2007), permitiendo trabajar con micromatrices 10 K de papa, cuyos clones están pre-asignados en los mapas. Para este trabajo se utilizó la versión 3.5.1 (19.11.2010) del programa. Los datos con los trabaja el programa están estructurados en tres bases de datos: los mapas, las rutas metabólicas y los datos de expresión.

En este contexto, el enfoque integrado de análisis de expresión presentado en este trabajo mediante la herramienta bioinformática Mapman, representa un recurso muy valioso que facilita la comprensión de perfiles de expresión, permitiendo obtener una visión global de la respuesta de las especies silvestres de papa frente al estrés hídrico.

Los mapas y las rutas metabólicas se descargaron del sitio <http://mapman.gabipd.org/web/guest/mapmanstore>, en particular accediendo al vínculo denominado “Stu\_TIGR download”.

Este vínculo, presenta la clasificación de los contigs provisional (TC) de papa obtenidos de la base de datos del Índice de Genes de la Papa (StGI) disponibles en Dana-Farber Cancer Institute (DFCI) (<http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=potato>).

De esta manera puede utilizarse esta herramienta para el análisis particular de micromatrices de papa. En el desarrollo de las bases que utiliza el MapMan, se puso especial atención en el mapeo de genes que no podrían ser anotados en base a su similitud con los genes de *Arabidopsis*, dado que posiblemente representarían genes únicos de la papa. A 97 de estos genes se los ha clasificado en “BINs” o clases funcionales, luego de su anotación manual.

El programa consta de dos módulos separados, el módulo SCAVENGER que recoge y clasifica los parámetros medidos en un conjunto de categorías jerárquicas funcionales (BINs, subBINs...enzima individual) y el módulo

IMAGEANNOTATOR que importa las clasificaciones y las utiliza para organizar y mostrar datos en diagramas.

### **Análisis de los transcriptomas de *S. cardiophyllum* y *S. venturii* bajo estrés hídrico**

Los ambientes originarios de las especies CPH y VNT son muy distintos. *S. cardiophyllum* fue reportada como originaria de México, de los estados de Hidalgo, México, Michoacán, Morelos, Querétaro y Puebla, desarrollándose junto a vegetaciones secas arbustivas, en el borde de cultivos y en campos viejos (Hawkes, 1990). Spooner la detalla más claramente en la monografía de Spooner (2004). Por otro lado, *S. venturii* fue descrita por Hawkes y Hjerting en 1960 en el norte de Argentina, en la provincia de Catamarca y reportada por Spooner y Clausen (1993) en una colección de germoplasma en las provincias de Catamarca, La Rioja, Jujuy, Salta y Tucumán, particularmente en las laderas húmedas, entre bosques de *Alnus* y *Polylepis*, o junto a campos cultivados, rutas y terrazas de ríos y a una altitud de entre 1900-3000 msnm.

Watkinson *et al*, 2006 proponen que la adaptación y aclimatación a la sequía implica la participación coordinada de las redes de regulación y, en concreto, de los genes asociadas a la señalización, la transcripción de genes y rutas metabólicas involucradas en la repartición de recursos, defensa de estrés y protección.

En esta tesis, el análisis comparativo de los transcriptomas de CPH y VNT sometidos a déficit hídrico de sequía puso de manifiesto diferencias en la velocidad de respuesta al estrés.

El análisis de la expresión diferencial frente a la sequía de las especies CPH y VNT mostró que las mismas implementaron distintas estrategias para hacer frente al estrés. Los genes expresados diferencialmente no solo variaron en el número sino también en sus perfiles temporales de expresión (Figura 18, 19, 21, 22, 23 y Material Complementario). Una cifra altamente significativa de los genes diferencialmente expresados como respuesta al tratamiento tendieron a agruparse mostrando comportamientos particulares en ambas especies.

En CPH, los perfiles de expresión permitieron identificar al menos 5 conglomerados que resumen las principales tendencias del comportamiento de los transcritos frente al estrés (ver Figura 18 y 19). Por otro lado, la dinámica de expresión de los transcritos diferenciales de VNT, mostró expresiones contrastantes y agrupadas, que sugieren una regulación común sincrónica (como se ve en las Figuras 21 y 22) confirmando la diversidad genética de respuestas en el germoplasma de especies silvestres de papa frente a la sequía. Estas diferencias de expresión entre papa silvestres sugieren nuevas vías para explotar la variabilidad genética del germoplasma americano para el mejoramiento del cultivo.

La expresión diferencial de genes es notablemente diferente en número y en cinética durante el período de estrés ensayado. Mientras CPH muestra muy poca expresión diferencial de genes concentrados en los metabolismos asociados a pared celular, transcripción, metabolismo de proteínas, señalización y genes de respuesta a estrés biótico y abiótico (Figura 25 y 26 A-B y 27 A-B-C) VNT muestra en todos los tiempos ensayados una expresión diferencial masiva en casi todo los grupos clasificados por MapMan (Figuras 28, 29 y 30).

De la comparación de los perfiles de expresión entre CPH y VNT a lo largo de tratamiento de sequía, surgen 150 genes comunes a ambas especies. Entre los genes de expresión única en CPH encontramos, HSP90, Entensinas, Patatinas, HistonaH3.2. Por su parte en VNT se expresan de forma única Citocromo P450 (EC 1.14.-.-), Glutamina Sintetasa (EC 6.3.1.2 EC 6.3.1.2), proteínas Ankirinas, etc.

A partir del análisis del MapMan con la especie VNT surgen grupos de clasificación y genes expresados diferencialmente únicos. Los grupos con más genes asociados corresponden a:

- Fotosíntesis
- Regulación rédox.
- Metabolismo de aminoácidos
- Metabolismo de hidratos de carbono

Una de las primeras respuestas frente al estrés hídrico es el cierre total o parcial de los estomas. Ekanayake y de Jong (1992) mostraron que especies

de *Solanum* cultivadas y silvestres presentaban un control de la pérdida de agua mediado por los estomas muy variable.

La falta de agua reduce el número y tamaño de hojas, la altura de los brotes (Deblonde y Ledent 2001) y la tasa fotosintética (Coleman, 2008). Todas las etapas de desarrollo son susceptibles al estrés hídrico, desde la emergencia hasta la iniciación del tubérculo, del llenado y el crecimiento de mismos. La exposición disminuye el desarrollo del tubérculo, aún en casos de cortos períodos de sequía, dependiendo de la oferta de hidratos de carbono desde el follaje que se vea afectada en estas condiciones de estrés por la reducción en la tasa fotosintética (Jefferies y MacKerron, 1989).

Las diferencias de adaptación podrían ser explicadas por diferencias de reacción de cada especie en términos de percepción de estrés, transducción de señales y programas apropiados de expresión génica y vías de transducción de señales (Bartels y Sunkar, 2005). Coleman (2008) observó respuestas fisiológicas superiores a nivel celular y de planta entera en la especie silvestre *S. gandarillasii* en relación con *S. tuberosum* cv Kennebec.

Los cambios en la expresión de genes bajo condiciones de sequía promueven una serie de alteraciones bioquímicas y fisiológicas. La fotosíntesis, una de las vías primarias biosintéticas está fuertemente afectada por el estrés hídrico, dado que se limita la disponibilidad de CO<sub>2</sub> para la fotosíntesis y por lo tanto favorece la fotorrespiración, por el cierre de los estomas. Análisis globales de la expresión génica mostraron una represión sustancial de los genes afectados a la fotosíntesis, bajo condiciones de sequía (Harb *et al*, 2010). La disminución de la actividad fotosintética es acompañada por una menor abundancia de genes relacionados con la fotosíntesis (Rizhsky *et al*, 2002). En estudios recientes de análisis transcriptómicos de diferentes especies de papas se ha observado que los genotipos susceptibles muestran una mayor expresión de genes diferenciales expresados que indicarían la necesidad de adaptación del aparato fotosintético en condiciones de deshidratación (Watkinson *et al*, 2006, Schafleitner *et al*, 2007, Vasquez-Robinet *et al*, 2008 Evers *et al*, 2010). En VNT el grupo de clasificación asociados a la fotosíntesis (Grupo 1) es uno de los que cuenta con mayor cantidad de genes expresados diferencialmente.

En CPH los genes de fotosíntesis no se vieron afectados. La activación de genes asociados a la fotorrespiración, así como la gran reducción

de la fotosíntesis sugiere un mayor impacto de la sequía en la especie VNT en relación con CPH.

Asociado a los cambios de expresión en genes de fotosíntesis se encuentran los del metabolismo del carbono, estos genes responden fuertemente al estrés por sequía, regulando genes de enzimas de degradación como lo observado para las  $\alpha$ -amilasas para ambas especies. La expresión de genes de hidratos de carbono, en particular de degradación de sacarosa, también se mostraron afectadas por la sequía (Sacarosa Sintasa e Invertasa Neutra), coincidiendo con lo encontrado en la bibliografía (Schafleitner *et al*, 2007).

Los genes de la biosíntesis de aminoácidos se vieron fuertemente afectados en la especie VNT. Estos genes están relacionados con el metabolismo de prolina, al igual que el metabolismo de lípidos (Schafleitner *et al*, 2007).

Otro de los grupos que está fuertemente representado es el de genes relacionados con detoxificación celular de especies reactivas de oxígeno. La sequía, al igual que muchos estreses aumentan la producción de especies reactivas de oxígeno (EROS) que afectan negativamente las estructuras celulares y el metabolismo (Alscher, 1997). Si bien la producción de estas especies de oxígeno cumple una función señal, las plantas deben protegerse del exceso de ROS para mantener la homeostasis metabólica. Es esperable que bajo condiciones de estrés, los genes asociados con los antioxidantes sean inducidos por la sequía (Bueno *et al*, 1998, Smirnov, 1998). Un ejemplo de antioxidantes, lo constituyen las metalotioninas, que se vieron inducidas en ambas especies. Aun así, la especie CPH, mostró una anticipada producción de estos antioxidantes desde T05 y la mantuvo hasta la finalización del tratamiento. Por su parte, VNT, muestra una expresión diferencial sólo en T07. El patrón de expresión de las metalotioninas podría correlacionarse con la adquisición de tolerancia a la sequía que lleva a que VNT requiera de una aclimatación antes de producir estos antioxidantes. Este tipo de respuesta fue observado en estudios transcripcionales similares con papas andígenas en condiciones de campo (Watkinson *et al*, 2006). Las metalotioninas han sido reportadas en la homeostasis de metales, en la desintoxicación de EROS y en el mantenimiento del equilibrio redox celular (Theocharis *et al*, 2003; Ghoshal, *et al* 1999).

A pesar de todo lo antedicho, si bien resultó muy útil para interpretar la respuesta al estrés hídrico de dos especies silvestres de papa, las agrupaciones arbitrarias asignadas a los genes dentro de las rutas metabólicas establecidas por este tipo de programas bioinformáticos, puede hacernos perder información valiosa. Un ejemplo en donde el programa MapMan no pudo agrupar en ninguno de los Bins o grupos funcionales que estructuran su base un gen de interés de respuesta a la sequía, fue con la Dehidrina TAS14, proteína inducida por ácido abscísico (STMHQ27) y que fuera validada por qPCR en este trabajo por presentar una expresión diferencial extrema frente al estrés.

### ***Estrategias de adaptación al estrés hídrico en especies silvestres de papa***

La respuesta al estrés por sequía en *S. tuberosum* es dependiente del cultivar, reflejando una variabilidad genética que puede ser aprovechada para descubrir mecanismos sólidos de resistencia al estrés hídrico (Levy, 1983, Jefferies, 1993, Heuer y Nadler, 1998). Levy identificó diferentes respuestas osmóticas a la sequía en los cultivares de *S. tuberosum* (Levy, 1992). Por un lado están las consideradas respuestas de aclimatación de corto plazo, donde producen un ajuste osmótico de corta duración (alta concentración de solutos y menor potencial osmótico) y por otro una constitutiva y continua elevada concentración de solutos y bajo potencial osmótico reflejo de un genotipo adaptado para hacer frente a estreses hídricos transitorios.

Blum (2005) describe diferentes estrategias de adaptación de las plantas a los diferentes estreses a los que se deben adaptar. El escape a la deshidratación es definido como la capacidad de mantener un estado hídrico alto bajo condiciones de estrés. Las plantas no parecen estar afectadas porque están poco expuestas a la deshidratación. Las estrategias que implementan las plantas en estas condiciones aumentando la capacidad de captura del agua del suelo, limitando las pérdidas de agua y reteniendo la hidratación celular aún con un menor potencial de agua en la planta.

La tolerancia a la deshidratación por su parte está definida como la capacidad de conservar las funciones celulares en un estado de deshidratación

Lo observado en el presente estudio apoya la idea de que ambas especies estudiadas a nivel transcriptómico poseen diferentes estrategias de adaptación al estrés hídrico impartido: *S. cardiophyllum* mostraría una estrategia de escape, mientras que *S. venturii* se adaptaría a través de la tolerancia al estrés hídrico.

## **7- CONCLUSIONES**

## CONCLUSIONES

- ✓ Las especies silvestres de *Solanum* evaluadas mostraron respuestas diferentes frente al tratamiento de sequía. Estos materiales en particular provienen de zonas geográficas muy distintas, lo que hace suponer que las diferencias en la respuesta al estrés reflejan la adaptación al ambiente de origen a través de estrategias específicas de expresión diferencial de genes para hacer frente en forma exitosa a la sequía. La información que se obtuvo sobre estas estrategias adaptativas podría ser utilizada en programas de mejoramiento de la papa cultivada (*S. tuberosum*) para aprovechar la diversidad genética y obtener variedades más resistentes a sequía.
  
- ✓ Los estudios de perfiles transcriptómicos realizados en micromatrices resultaron estadísticamente satisfactorios para dos especies silvestres de papa (*S. cardiophyllum* y *S. venturii*), ambas con comportamientos contrastantes bajo el tratamiento de estrés hídrico ensayado, mostrándose *S. cardiophyllum* como especie más resistente que *S. venturii*.
  
- ✓ Los análisis bioinformáticos realizados permitieron describir en forma tanto detallada como en forma más sistémica las respuestas al estrés en término de vías metabólicas y fisiológicas, aunque podrían afinarse construyendo una ontología más detallada y que clasifique mejor los genes diferencialmente expresados no clasificados por programas como el MapMan pero encontrados como diferencialmente expresados en estos estudios.
  
- ✓ Entre las vías metabólicas y procesos fisiológicos diferencialmente más afectados en la especie más susceptible a la sequía, *S. venturii*, se encontraron: la fotosíntesis, el metabolismo de aminoácidos, el metabolismo de hidratos de carbono y los mecanismos de detoxificación de especies reactivas de oxígeno

- ✓ Estos estudios sugieren que las dos especies silvestres de papa: *S. cardiophyllum* y *S. venturii* muestran diferente respuesta a la exposición a estrés hídrico. El análisis transcriptómico de *S. cardiophyllum* mostró variación en la expresión de pocos genes entre las plantas control y las sometidas a sequía, mientras que en el análisis de *S. venturii* se encontraron 4133 genes diferencialmente expresados que denotan diferentes mecanismos involucrados para lograr la respuesta adaptativa.
  
- ✓ Dado que *S. cardiophyllum* varió en su expresión sólo 235 genes, de los cuales 57 genes fueron únicos, éstos son interesantes genes candidatos para ser evaluados como nuevos genes de resistencia a la sequía.
  
- ✓ *S. cardiophyllum* respondería implementando una estrategia de escape, ya que tiene la capacidad de mantener el estado hídrico alto, bajo condiciones de sequía y parece no estar mayormente afectada en sus patrones transcripcionales.  
*S. venturii* mostró una estrategia de tolerancia, ya que sus patrones de transcripción activados más tardíamente a lo largo del estrés hídrico, lograrían conservar las funciones celulares aún mostrando síntomas de deshidratación.

### **Perspectivas:**

- ✓ Los presentes análisis transcripcionales deberán ser complementados con estudios bioquímicos y metabólicos que describan en los diferentes metabolismos afectados, lo que ocurre a cada una de estas especies durante el estrés hídrico. Esto posibilitará tener una visión global y completa de los diferentes mecanismos que gobiernan las diferentes estrategias observadas en estas dos especies silvestres de papa.

## **8- BIBLIOGRAFÍA**

## 8. BIBLIOGRAFÍA

- Abe H., Urao T., Ito T., Seki M., Shinozaki K., Yamaguchi-Shinozaki K. (2003) *Arabidopsis AtMyc2 (bHLH) and AtMyb2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell* 15: 63–78.
- Abogadallah G.M., Nada R.M., Malinowski and Quick P. (2010) *Overexpression of HARDY, an AP2/ERF gene from Arabidopsis, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic Trifolium alexandrinum L. PLANTA. Volume 233, Number 6: 1265-1276.*
- Abogadallah, G. M., Nada, R. M., Malinowski, R. y Quick, P. (2011) *Overexpression of HARDY, an AP2/ERF gene from Arabidopsis, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic Trifolium alexandrinum L. PLANTA, Vol. 233, Núm. 6, 1265-1276.*
- Agrios, GN. (1997) *Plant Pathology. London Academic. 4th ed. Págs 25-37.*
- Ahuja, I. Ric, C.H. de Vos Atle Bones M., Hall, R.D. (2010) *Plant molecular stress responses face climate change. Trends in Plant Science, Vol. 15, No. 12: 664-674.*
- Alonso, F. (1996) *El Cultivo de la papa. Ediciones Mundi-Prensa.*
- Alscher R.G., Donahue J.L., Cramer C.L. (1997) *Reactive oxygen species and antioxidants: relationships in green cell. Physiol. Plantarum* 100: 224–233.
- Altschul S.F, Gish W, Miller W, Myers EW and Lipman DJ. (1990) *J Mol Biol* 215: 403-440.
- ARGENPAPA: <http://www.argenpapa.com.ar>
- Arumuganathan K, Earle E D. (1991) *Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter. Publisher: Springer Netherlands. Volume: 9, Issue: 3: 208-218.*

- Arvin M. J.; D. J. Donnelly. (2008). *Screening Potato Cultivars and Wild Species to Abiotic Stresses Using an Electrolyte Leakage Bioassay*. *J. Agric. Sci. Technol.* Vol. 10: 33-42
- Attwood, T.K. y D.J. Parry-Smith. 2002. *Introducción a la Bioinformática*. Prentice Hall.
- Avni, H., Füsün, Ö., Selçuk, F., Tufan, M., Teixeira, J. A. y Yücel, M. (2008) *Revealing Response of Plants to Biotic and Abiotic Stresses with Microarray Technology*. *Genes, Genomes and Genomics*, 2(1), 14-48.
- Bajaj, S., Targolli, J., Liu, L.F., Ho, T.H.D. y Wu, R. (1999) *Transgenic approaches to increase dehydration-stress tolerance in plants*. *Mol Breed* 5: 493–503
- Bamberg, J. y del Rio, A. (2010) *La diversidad genética y los desafíos frente al cambio climático*. Charla Magistral. XXIV Congreso ALAP 2010, Cusco, Perú. <http://www.papaslatinas.org/alap2010.html>
- Bartels, D. y Sunkar, R. (2005) *Drought and salt tolerance in plants*. *Crit Rev Plant Sci* 24: 23–58.
- Bazzaz, F. y Sombroek. (1996) Cap. 12. *Global change impacts on agriculture, forestry and soils: The programme of the global change and terrestrial ecosystems core project of IGBP*. En: "Global climate change and agricultural production. Direct and indirect effects of changing hydrological, pedological and plant physiological processes". Editado por FAO, Roma, Italia. ISBN 0-471-95763-1
- Benjamini, Y., Hochberg, Y. (1995) *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. *Journal of the Royal Statistical Society, B*, 57: 289-300.
- Blum, A. (2005) *Drought resistance, water-use efficiency, and yield potential: Are they compatible, dissonant, or mutually exclusive?* *Aust J Agric Res* 56: 1159–1168
- Bohnert H.J, Nelson D.E., Jensen R.G (1995) *Adaptation to environmental stresses*. *Plant Cell* 7: 1099-1111.
- Bonierbale M.W., Plaisted, R.L. y Tanksley, S.D. (1988) *RFLP Maps Based on a Common Set of Clones Reveal Modes of Chromosomal Evolution in Potato and Tomato*. *Genetics*, 120(4):1095–1103.

- Bowden, R.L, y Rouse, D. I. (1991) *Effects of Verticillium dahliae on gas exchange of potato. Phytopathology* 81, 293–301.
- Boyer J. S. (1982) *Plant Productivity and Environment. Science*, Vol. 218 (4571): 443-448
- Bradshaw J. E., Pande B., Bryan G. J., Hackett, C.A., McLean, K., Stewart, H.E. y Waugh, R. (2004), "Intervalmapping of quantitative trait loci for resistance to late blight [*Phytophthora infestans* (Mont.) de Bary], height and maturity in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*)," *Genetics*, vol. 168, no. 2, pp. 983–995.
- Bradshaw, J. E. (2007) *Cap. Potato-Breeding Strategy. En: D. Vreugdenhil (Ed.) Potato Biology and Biotechnology Advances and Perspectives. Elsevier, Oxford. Pág. 157-177.*
- Bray E.A. (2002) *Classification of Genes Differentially Expressed during Water-deficit Stress in Arabidopsis thaliana: an Analysis using Microarray and Differential Expression Data. Ann Bot*, 89 (7): 803-811.
- Bray, E. A., Bailey-Serres, J., y Weretilnyk, E. (2000) *Responses to abiotic stresses. En: Biochemistry and Molecular Biology of Plants (B. B. Buchanan, W. Gruissem, y R. L. Jones, eds), Pág: 1158-1203. American Society of Plant Physiologists, Rockville, Md..*
- Bray, E.A. (1997) *Plant responses to water deficit. Trends in Plant Science*, Vol. 2, Núm. 2: 48-54.
- Brazma, A. (2001). *Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet.* 29, pág.365-371.
- Breyne, P., Dreesen, R., Cannoot, B., Rombaut, D., Vandepoele, K., Rombauts, S., Vandeerhaeghen, R., Inze, D. and Zabeau. (2003) *Mol Gen Genomics* 269: 173-79.
- Brown, P. y Botstein, D. (1999) *Exploring the new world of the genome with DNA microarrays. Nature Genetic.* 21(Suppl.), 33-37.
- Bryan G. J, Hein I. (2008) *Genomic resources and tools for gene function analysis in potato. International journal of plant genomics* 2008 p. 216 513.

- Bueno P., Piqueras A., Kurepa J., Savoure A., Verbruggen N., Van Montagu M. e Inze D. (1998). *Expression of antioxidant enzymes in response to abscisic acid and high osmoticum in tobacco BY-2 cell cultures. Plant Sci.* 138. P 27–34.
- Burton, W. G. (1989) *The potato - 3rd ed. John Wiley & Sons Inc. New York.* 742 pp.
- Centro Internacional de la papa (CIP) (1999). *Principales Enfermedades, Nematodos a Insectos de la Papa.*
- Chaves, M.M., Pereira, J.S., Maroco, J., Rodriguez, M.L., Ricardo, C.P.P., Osorio, M.L., Carvalho, I., Faria, T. y Piñeiro, C. (2002) *How Plants Cope with Water Stress in the Field. Photosynthesis and Growth. Annals of Botany* 89: 907-916.
- Clausen, A. (2003) *Conservación de germoplasma de papa. Actualidad Papera - Año 2 Enero-Febrero-Marzo N° 5.*
- Coker, J. y Davies, E. (2003) *Selection of candidate housekeeping controls in tomato plants using EST data. BioTechniques*, 35(4), 740-748.
- Coleman, W. K. (2008) *Evaluation of wild Solanum species for drought resistance 1 .Solanum gandarillasii Cardenas. Env and Experimental Botany*, 62, 221-230.
- Commonwealth Potato Collection (CPC)  
[http://germinate.scri.ac.uk/germinate\\_cpc/app/index.pl](http://germinate.scri.ac.uk/germinate_cpc/app/index.pl)
- Contreras, A. (2008) *Uso de especies silvestres y cultivadas en el Mejoramiento de la papa. Agro Sur* 36 (3) 115-129.
- Cook D., Fowler S., Fiehn O., Thomashow M.F. (2004) *A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. Proc. Natl. Acad. Sci. U.S.A.* 101, 15243–15248.
- Correll, D. S., (1962) *The potato and its wild relatives. Contributions from the Texas Research Foundation, Botanical Studies* 44: 1- 606.
- Corvo Dolcet, S. (2005) *“Cultivo de la papa en la Argentina”.* Secretaría de agricultura, ganadería, pesca y alimentos. Subsecretaría de agricultura, ganadería y alimentación. Dirección de agricultura

- David L., Veilleux, R. E. (2007) *Adaptation of potato to high temperatures and salinity-a review. Invited review. american journal of potato research. Vol. 84, Núm. 6, 487-506.*
- Deblonde, P.M.K. y Ledent, J.F. (2001) *Effects of moderate drought on green leaf number, stem height, leaf length and tuber yield of potato cultivars, Eur. J.Agron. 14, Págs 31–41.*
- Di Rienzo, J.A. (2009) *fgStatistics. Statistical software for the analysis of experiments of functional genomics. RDNDA. Argentina. 756587. <http://sites.google.com/site/fgStatistics/>*
- Dong, F. Song, J. Naess, S.K. Helgeson J. P. Gebhardt, C. and Jiang, J. (2000) *Development and applications of a set of chromosome-specific cytogenetic DNA markers in potato,"Theoretical and Applied Genetics, Vol. 101, N°7, pág, 1001–1007.*
- Durbin R., Eddy S., Krogh A. y Mitchinson G. (1998) *Biological sequence analysis. Probabilistic models of proteins and nucleic acids. Cambridge: Cambridge University Press. 356 pág.*
- Dutch – German Potato Collection (BGRC).
- Ekanayake, I.J. y De Jong, J.P. (1992) *Stomatal response of some cultivated and wild tuber-bearing potatoes in warm tropics as influenced by water deficits. Ann Bot 70:53–60.*
- Evans, K., Parkinson, K.J. y Trudgill, D.L. (1975) *Effects of Potato Cyst-Nematodes On Potato Plants. Nematologica, Vol.21, Núm. 3, 1975, pp. 273-280 (8)*
- Evers, D., Lefevre, I., Legay, S., Lamoureux, D., Hausman, J.-F., Rosales, R.O.G., Marca, L.R.T., Hoffmann, L., Bonierbale, M. y Schafleitner, R. (2010) *Identification of drought-responsive compounds in potato through a combined transcriptomic and targeted metabolite approach. Journal of Experimental Botany 61; 2327-2343.*
- F.A.O. Food and Agriculture Organization of The United Nations:<[www.fao.org/](http://www.fao.org/)>
- Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). *Crosstalk between abiotic and biotic stress responses: a current view from the points of*

- convergence in the stress signaling networks. *Curr Opin Plant Biol* 9:436–442.
- Gebhardt C. (2007) Molecular markers, maps, and population genetics, in *Potato Biology and Biotechnology: Advances and Perspectives*, D. Vreugdenhil, Elsevier, Amsterdam, The Netherlands Ed., chapter 5, pp. 77–89.
  - Gebhardt, C., Ritter, E. y Salamini, F. (2001) RFLP map of the potato. In: RL Philipps, IK Vasil (eds), *DNA-based Markers in Plants*, Vol 6. Huwer Academic Publishers, Dordrecht, 319-336.
  - Gebhardt, C., Ritter, E., Barone, A., Debener, T., Schachtschabel, U., Kaufman, H., Thompson, R.D., Bonierbale, M.W., Ganal, M.W., Tanksley, S.D. y Salamini, F. (1991) RFLP maps of potato and their alignment with the homeologous tomato genome. *Theor Appl Genet* 83:49–57.
  - Gebhardt, C., Ritter, E., Debener, T., Schachtschabel, U., Walkemeir, B., Ghoshal K., Majumder S., Li Z., Bray T.M. y Jacob S.T. (1999). Transcriptional induction of metallothionein-I and -II genes in the livers of Cu,Zn-superoxide dismutase knockout mice. *Biochem. Biophys. Res. Commun.* 264. Págs 735–742.
  - González, L., González-Vilar, M. (2003). Determination of Relative Water Content. *Handbook of Plant Ecophysiology Techniques*. Chapter 14. Pp. 207-212.
  - Greilhuber, J., Borsch, T., Müller, K., Worberg, A., Porembski, S., y Barthlott, W. (2006) Smallest angiosperm genomes found in *Lentibulariaceae*, with chromosomes of bacterial size. *Plant Biology*, 8: 770-777.
  - Hackett C. A, Pande B. y Bryan G. J. (2003) “Constructing linkage maps in autotetraploid species using simulated annealing. *Theoretical and Applied Genetics*. Vol. 106, N° 6, pp. 1107–1115.
  - Hand D., Mannila H., y Smyth P. (2001) *Principles of Data Mining*. MIT Press, Cambridge, MA, 2001, 546 pp.
  - Harb, A., Krishnan, A., Ambavaram, M. M. R. y Pereira, A. (2010) *Molecular and Physiological Analysis of Drought Stress in Arabidopsis*

- Reveals Early Responses Leading to Acclimation in Plant Growth. Plant Physiology, Vol. 154, 1254–1271*
- Haverkort A.J. and C. Grashoff, 2004, In: D.K.L. MacKerron and A.J. Haverkort (eds), *Decision Support*
  - Haverkort A.J., D.K.L. MacKerron (1995). *Potato Ecology and Modelling of Crops under Conditions Limiting Growth. Kluwer, Dordrecht.*
  - Haverkort, A.J., Rouse D.I. y Turkensteen, L.J. (1990) *The influence of Verticillium dahliae and drought on potato crop growth. Effects on gas exchange and stomatal behaviour of individual leaves and crop canopies*
  - Haverkort, A.J., Verhagen, A. (2008) *Climate change and its repercussions for the potato supply chain. Potato Research, Vol 51, Núm 3-4, p. 223-237.*
  - Hawkes, J. G. (1978) *Biosystematics of the potato. Cap. 2. En: The potato Crop. Editor: Harris, P. M. pp. 15-69.*
  - Hawkes, J.G. (1990) *The potato: evolution, biodiversity and genetic resources. Belhaven Press, London.*
  - Hazen, S. P., Wu, Y. y Kreps J. A. (2003) *Gene expression profiling of plant responses to abiotic stresses. Funct. Integr. Genomics 3:105-111.*
  - Heuer, B., Nadler, A. (1998) *Physiological response of potato plants to soil salinity and water deficit. Plant Sci. 137. Págs 43–51.*
  - Hijmans, R.J. y Spooner, D. M. (2001) *Geographic distribution of wild potato species. Am. J. Bot., Vol. 88 N. 11, 2101-2112.*
  - Hijmans, R.J., Spooner, D.M., Salas, A.R., Guarino, L. y De la Cruz, J. (2002). *Atlas of wild potatoes. Systematic and Ecogeographic Studies in Crop Gene pools 8. IPGRI, Rome.*
  - Holmberg, N. and Bülow. (1998) *Improving stress tolerance in plants by gene transfer. Trends in Plant Science 3, 61-66.*
  - Hooker, W.J. (1980) *Compendio de enfermedades de la papa. Imp / Ed.: Lima, Perú :Centro internacional de la papa, 1980. Descripción:166 pp.*
  - Inan G., Zhang Q., Li P., Wang Z., Cao Z., Zhang H., Zhang C., Quist T.M., Goodwin S.M., Zhu J., Shi H., Damsz B., Charbaji T., Gong Q., Ma S., Fredricksen M., Galbraith D.W., Jenks M.A., Rhodes D., Hasegawa

- P.M., Bohnert H.J., Joly R.J., Bressan R.A y, Zhu J.K. (2004) Salt cress. A halophyte and cryophyte *Arabidopsis* relative model system and its applicability to molecular genetic growth and development of extremophiles. *Plant Physiol.* 135 Págs 1718–1737.
- Ingram, J. y Bartels, D. (1996) The molecular basis of dehydration tolerance
  - IPK Genbank Groß Lüsewitz (GLKS). Institut für Pflanzen genetik und Kulturpflanzen forschung Gatersleben (Instituto de Genética Vegetal y de Investigación de Cultivos Vegetales) (IPK), Genbank, Außenstelle GroßLüsewit. <http://www.ipk-gatersleben.de/Internet/Forschung/Genbank>
  - Jansky, S.H., Peloquin, S.J. (2005). Advantages of wild diploid *Solanum* species over cultivated diploid relatives in potato breeding programs. *Genet. Resour. Crop Evol.* 56: Págs 669–674.
  - Jefferies R. A., Heilbronn T. D. (1991) Water stress as a constraint on growth in the potato crop. 1. Model development. *Agricultural and Forest Meteorology.* Volume 53, Issue 3: 185-196.
  - Jefferies, R.A. (1993) Responses of potato genotypes to drought. I. Expansion of individual leaves and osmotic adjustment. *Ann. Appl. Biol.* 122. Págs 93–104.
  - Jeffries, R.A., Heilbron, T.D. y Mackerron, D.K.L. (1989) Estimating tuber dry matter concentration from accumulated time and soil moisture. *Potato Res.* 32, 411-417.
  - Kang J., Choi H., Im M. y Kim S.Y (2002) *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *Plant Cell* 14. Págs 343–357.
  - Kloosterman, B., De Koeyer, D., Griffiths, R., Flinn, B., Stenernagel, B., Scholz, U., Sonnewald, S., Sonnewald, U., Bryan, G.J., Prat, S., Banfalvi, Z., Hammond, J.P. Geigenberger, P., Nielsen, K., Vissier, R.G. y Bachan, C.W. (2008) Genes driving potato tuber initiation and growth: identification based on transcriptional changes using POCl array. *Funct. Integr. Genomics* 8: 329-340.
  - Lee, N.H. y Saeed, A. (2007) Microarrays: an Overview. *Methods in Molecular Biology*, 353:265-300.

- Legay, S., Lefevre, I., Lamoureux, D., Barreda, C., Luz, RT, Gutierrez, R., Quiroz, R., Hoffmann, L., Hausman, J.F., Bonierbale, M., Evers, D. y Schafleitner, R.( 2011) Carbohydrate metabolism and cell protection mechanisms differentiate drought tolerance and sensitivity in advanced potato clones (*Solanum tuberosum* L.). *Funct Integr Genomics*. 11(2):275-91.
- Levitt J. (1980) *Responses of plants to environmental stresses*, Vol. I y II, Academic Press.
- Levy, D. (1983) *Varietal differences in the response of potatoes to repeated short periods of water stress in hot climates. 2. Tuber yield and dry matter accumulation and other tuber properties*. *Potato Res.* 26, 313–321
- Levy, D. (1992) *Osmotic potential of potatoes subjected to a single cycle of water deficit*. *Potato Res.* 35: 17-24.
- Levy, D. y Veilleux R.E. (2007) *Adaptation of Potato to High Temperatures and Salinity-A Review*. *American Journal of Potato Research* 84:437-456
- Linneo, C. (1753) *Species Plantarum*:  
<http://www.botanicus.org/title.aspx?bibid=b12069590#>.
- Liu, F., Shahnazari, A., Andersen, M. N., Jacobsen, S.-Erik, y Jensen, C. R. (2006). *Physiological responses of potato ( Solanumtuberosum L .) to partial root-zone drying : ABA signalling , leaf gas exchange , and water use efficiency*. *Access*, 57(14), 3727-3735. doi: 10.1093/jxb/erl131.
- Livak K.J., Schmittgen T.D. (2001): *Analysis of relative gene expression data usingn real-time quantitative PCR and the 2-  $\Delta\Delta$ Ct method*. *Methods*, 25: 402-408.
- Lopez C.G., Banowitz G.M., Peterson C.J., Kronstad W.E. (2003). *Dehydrin expressionand drought tolerance in seven wheat cultivars*. *Crop Sci.* 43: 577-582.
- Meehl, G.A.(2007) *Global Climate Projections*. In *Climate Change (2007): The Physical Science Basis. Contribution of WorkingGroup I to the Fourth Assessment Report of the IntergovernmentalPanel on Climate*

- Change* (Solomon, S. et al., eds), pág. 749–845. Cambridge University Press.
- Michiels, S. (2005) *Prediction of cancer outcome with microarrays: a multiple random validation strategy*. *Lancet*, 365, 488-492.
  - Moreno, L.P. (2009) *Respuesta de las plantas al estrés por déficit hídrico: Una revisión* *Plant responses to water deficit stress. A review*. *Agronomía Colombiana* 27 (2) 179-191.
  - Murphy D. (2002). *Gene expression studies using microarrays: principles, problems, and prospects*. *Advan in Physiol Edu* 26: Págs 256-270.
  - Nicot, N., Jean-Francois, H., Hoffmann, L. y Evers, D. (2005) *Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress*. *J. Exp. Bot.*, 56: 2907-2914.
  - Öktem, H. A., Eyidoğan, F., Selçuk, F., Mehmet Tufan Öz; Teixeira da Silva, J.A, Yücel M. (2008). *Revealing Response of Plants to Biotic and Abiotic Stresses with Microarray Technology*. *Genes, Genomes and Genomics*, 2(1) Págs 14-48.
  - Palta, J. (2010) *Understanding and Managing the Impact of Abiotic Stresses*. XXIV Charla Magistral. XXIV Congreso ALAP 2010, Cusco, Perú. <http://www.papaslatinas.org/alap2010.html>
  - Plaisted, R. (1982). «Potato». En W. Fehr y H. Hadley. *Hybridization of Crop Plants*. Nueva York: American Society of Agronomy, Crop Science Society of America. pp. 483-494.
  - Pradeep, K.A., Parinita, A., Reddy, M. y Sopory Sudhir, K. (2006) *Role of DREB transcription factors in abiotic and biotic stress tolerance in plants*. *Plant Cell Reports* 2006, 25(12):1263.
  - Quackenbush J, Liang F, Holt I, Perteau G, Upton J (2000) *The TIGR gene indices: reconstruction and representation of expressed gene sequences*. *Nucleic acids research* 28 (1) pág 141-5.
  - Quackenbush J. (2002). *Microarray data normalization and Transformation*. *Nature Genetics Supplement*. Volume 32. Suppl: 496-501.

- Renne, P. , Dressen U., Hebbeker U., Hille D., Flu"gge U.I., Westhoff P., Weber A.P.M. (2003). *The Arabidopsis mutant dct is deficient in the plastidicglutamate/malate translocator DiT2*. *Plant J.* 35 Págs 316-331.
- Reynolds, M.P. (2010) *Cap. 1. Crop and Climate. En: Climate change and crop production / editor, Matthew P. Reynolds, ISBN 978-1-84593-633-4.*
- Rizhsky L., Liang H. y Mittler R. (2002). *The combined effect of drought stress and heat shock on gene expression in tobacco*. *Plant Physiol* 130, Págs 1143-1151.
- Romero – Zaliz; R. (2005) *Reconocimiento de Perfiles de Regulación Genética mediante Algoritmos Evolutivos Multiobjetivo*. Universidad de Granada. ISBN 978-84-338-3531-4.
- Ross, H. (1986) *Supplements to J. of Plant Breeding* 13 (eds. Röbbelen, W.H.a.G.) (Verlag Paul Parey, Berlin, FRG, Berlin and Hamburg, Germany).
- Rotter A, Usadel B, Baebler S, Stitt M, Gruden K (2007) *Adaptation of the MapMan ontology to biotic stress responses: application in solanaceous species*. *Plant methods* 3: 10.
- Rousselle, P., Robert, Y.y Crosnier, J. C. (1999) *La Patata: producción, mejora, plagas y enfermedades, utilización INRA*. Ediciones Mundi-Prensa, 607pp.
- Ruiz de Galaterra, J.I., Carrasco, A., Salazar, A., Barrena, I., Iturrutxa, E., Marquez, R, Legorboru, F.J. y Ritter, E., (1998). *Wild Solanum species as resistance sources against different pathogens of potato*. *Potato Research* 41(1998): 57 – 68.
- Sakamoto H., Maruyama K., Sakuma Y., Meshi T., Iwabuchi M., Shinozaki K. y Yamaguchi-Shinozaki K. (2004) *Arabidopsis Cys2/His-typezinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions*. *Plant Physiol.* 136 Págs 2734–2746.
- Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M. y Shinozaki, K. (2006). *Functional Analysis of an Arabidopsis Transcription Factor , DREB2A , Involved in Drought-Responsive Gene Expression*. *Gene Expression*, 18, 1292-1309.

- Salaman, R.N. (1926) *Potato Varieties*. Cambridge University Press, Cambridge, UK. 378pp.
- Sambrook J., Fritsch EF, Maniatis T. (1989) *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*.
- Šášik R, Woelk C H y Corbeil J (2004). Review. *Microarray truths and consequences. Journal of Molecular Endocrinology*, 33, 1–9
- Schafleitner, R., Gutierrez Rosales, R.O., Gaudin, A., Alvarado Aliaga, C.A., Nomberto Martinez, G., Tincopa Marca, L.R., Bolivar, L.A., Mendiburu Delgado, F., Simon, R. y Bonierbale, M. (2007) *Capturing candidate drought tolerance traits in two native Andean potato clones by transcription profiling of field grown plants under water stress. Plant Physiology and Biochemistry* 45, 673–690.
- Schapendonk, A.H.C.M., Spitters, C.J.T. y Groot, P.J. (1989) *Effects of water stress on photosynthesis and chlorophyll fluorescence of five potato cultivars. Potato Research* 32, 17- 32.
- Scott, C.P., VanWye, J., McDonald, M.D., Crawford, D.L. (2009) *Technical Analysis of cDNA Microarrays. PLoS ONE* 4(2): e4486. doi:10.1371/journal.pone.0004486
- Seki M., Narusaka M., Abe H., Kasuga M., Yamaguchi-Shinozaki K., Carninci P., Hayashizaki Y., Shinozaki K. (2001). *Monitoring the expression patterns of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell* 13: 61–72.
- Shinozaki K, Yamaguchi-Shinozaki K y Seki M. (2003). *Regulatory network of gene expression in the drought and cold stress responses. Current Opinion in Plant Biology* 6, 410–417.
- Shinozaki K, Yamaguchi-Shinozaki K. (1997) *Gene Expression and signal transduction in water-stress response. Plant Physiol.* 115: 327-334.
- 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. Plant Biol.* 3, 217-223
- Shinozaki, K. y Yamaguchi-Shinozaki, K. (1999) *Molecular responses to drought stress. In Molecular Responses to Cold, Drought, Heat and Salt*

- Stress in Higher Plants* (Shinozaki, K. and Yamaguchi-Shinozaki, K., eds). Austin, TX: RG Landes: 11-28.
- Shinozaki, K. y Yamaguchi-Shinozaki, K. (2007) Gene networks involved in drought stress response and tolerance. *J. Exp. Bot.* (2007) 58 (2): 221-227.
  - Smirnoff, N. (1998) Plant resistance to environmental stress. *Curr. Opin. Biotechnol.* 9 Págs 214–219.
  - Smyth, G. K. (2005). *Limma: linear models for microarray data*. In: *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds.), Springer, New York, Págs 397–420.
  - Smyth, G. K., Speed, T. P. (2003). Normalization of cDNA microarray data. In: *METHODS: Selecting Candidate Genes from DNA Array Screens: Application to Neuroscience*, D. Carter (ed.). *Methods Volume 31, Issue 4*, Págs 265-273.
  - Speed, T. (2002) *Statistical Methods For Identifying Differentially Expressed Genes in Replicated cDNA Microarray Experiments*, *Statistica Sinica*, 12: 111–139.
  - Spitters, C.J.T.y Schapendonk, A.H.C.M. (1990). Evaluation of breeding strategies for drought tolerance in potato by means of crop growth simulation. *Plant and Soil* 123, p 193-203.
  - Spooner DM, Salas A (2006): *Structure, Biosystematics, and genetic resources*. In *Handbook of potato production, improvement, and post-harvest management*. Edited by Gopal J, Khurana SMP. New York: The Haworth Press; Págs 1-39.
  - Spooner, D.M. (2004) *Thirty wild potato descriptions from North and Central America for NSF plant biotic inventories monograph project*. *Systematic Botany Monografía*. 68:1-209.
  - Spooner, D.M. y Clausen, A.M. (1993) *Wild potato (Solanum sect. Petota) germplasm collecting expedition to Argentina in 1990, and status of Argentinian potato germplasm resources*. *Potato Res.* 36:3-12.
  - Spooner, D.M. y Salas, A. (2006) *Structure, biosystematics, and genetic resources*. *Handbook of potato production, improvement, and post-*

- harvest management (ed. by J. Gopal and S.M Paul Khurana), pp. 1–39. Haworth Press, Inc., Binghamton, NY.
- Steyn, J.M., Du Plessis, H.F. y Nortje, P.F. (1992) Die invloed van verskillende waterregimes op Up-todate aartappels I. Vegetatiewe ontwikkeling, fotosintese tempo en stomatadiffusieweerstand. *S. Afr. J. Plant Soil.* 9(3), 113-117.
  - Struik, P.C., Geertsema, J. y Custers, C.H.M.G. (1989) Effects of shoot, root and stolon temperature on the development of the potato (*Solanum tuberosum* L.) plant. III. Development of tubers. *Potato Res.* 32, 151.
  - Taji T., Seki M., Satou M., Sakurai T., Kobayashi M., Ishiyama K., Narusaka Y., Narusaka M., Zhu J.K. y Shinozaki K. (2004) Comparative genomics in salt tolerance between *Arabidopsis* and *Arabidopsis*-related halophyte salt cress using *Arabidopsis* microarray. *Plant Physiol.* 135: 1697-1709.
  - Tanksley, S.D., Ganai, M.W., Prince, J.P., de Vicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giovannoni, J.J., Grandillo, S., Martin, G.B., Messeguer, R., Miller, J.C., Miller, L., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W. y Young, N.D. (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141-1160.
  - The Gene Ontology Consortium (2000) *Nature Genet.* 25: 25-29
  - The TIGR Solanaceae Gene Expression Database (SGED). [http://www.tigr.org/tdb/potato/SGED\\_datasearch.shtml](http://www.tigr.org/tdb/potato/SGED_datasearch.shtml)
  - Theocharis, S.E., Margeli, A.P. y Koutselini, A. (2003). Metallothionein: a multifunctional protein from toxicity to cancer. *Int. J. Biol. Marker.* 18: 162–169.
  - Thimm, O., Blaesing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y. y Stitt, M. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37(6):914-39.
  - TIGR: The Institute for Genomic Research: [www.tigr.org](http://www.tigr.org).

- Ueguchi, C., Sato, S., Kato, T., y Tabata, S. (2001) *The AHK4 gene involved in the cytokinin-signaling pathway as a direct receptor molecule in Arabidopsis thaliana. Plant Cell Physiol. 42, 751–755*
- Uhrig, H. y Salamini, F. (1989) *RFLP analysis and linkage mapping in Solanum tuberosum. Theor Appl Genet 78:65–75*
- Uno Y., Furihata T., Abe H., Yoshida R., Shinozaki K. y Yamaguchi-Shinozaki K. (2000) *Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. Proc. Natl. Acad. Sci. U.S.A. 97: 11632-11637.*
- Urao, T., Yakubov, B., Satoh, R., Yamaguchi-Shinozaki, K., Seki, B., Hirayama, T., y Shinozaki, K. (1999) *A transmembrane hybrid-type histidine kinase in Arabidopsis functions as an osmosensor. Plant Cell 11, 1743–1754*
- Urbanczyk-Wochniak E, Usadel B, Thimm O, Nunes-Nesi A, Carrari F, Davy M, Bläsing O, Kowalczyk M, Weicht D, Polinceusz A, Meyer S, Stitt M, Fernie A R (2006) *Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: effects of genetic and environmental alterations in energy metabolism in the leaf. Plant molecular biology 60 (5) p. 773-92.*
- Usadel, B., Nagel, A., Thimm, O., Redestig, H., Blaesing, O.E., Palacios-Rojas, N., Selbig, J., Hannemann, J., Piques, M.C., Steinhauser, D., Scheible, W.R., Gibon, Y., Morcuende, R., Weicht, D., Meyer, S. y Stitt, M. (2005) *Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. Plant Physiol. 138: 1195-204.*
- Valliyodan, B. y Nguyen, H.T. (2006) *Understanding regulatory networks and engineering for enhanced drought tolerance in plants. Current Opinion in Plant Biology 9: 1-7.*
- Van Loon, C.D., (1981) *The effect of water stress on potato growth, development and yield. Am. Potato J. 38, 51–69.*

- Vasquez-Robinet, C., Mane, S.P., Ulanov, A., Watkinson, J.I. y Stromberg, V.K. (2008) *Physiological and molecular adaptations to drought in andean potato genotypes*. *J. Exp. Bot.*, 59: 2109-2123.
- Vinod, M. S., Sharma, N., Manjunath, K., Kanbar, A., Prakash, N. B. y Shashidhar, H. E. (2006) *Candidate genes for drought tolerance and improved productivity in rice (Oryza sativa L)*. *Journal of biosciences*, 31(1): 69-74.
- Vos, J. & Haverkort, A.J. (2007) *Water availability and potato crop performance*. En: *Potato Biology and Biotechnology- Advances and Perspectives* (edited by D. Vreugdenhil, J. Bradshaw, C. Gebhardt, F. Govers, D.L. MacKerron, M.A. Taylor & H.A. Ross. Oxford: Elsevier. Pp. 333-351.
- Vos, J. y Groenwold, J. (1989) *Genetic differences in water-use efficiency, stomatal conductance and carbon isotope fractionation in potato*. *Potato Research*, Vol. 32, Núm. 2: 113-121.
- Watkinson, J.I., Sioson, A.A., Vasquez-Robinet, C., Shukla, M., Kumar, D., Elis, M., Heath, L.S., Ramakrishnan, N., Chevone, B., Watson, L.T., Van Zyl, L., Egertsdotter, U., Sederoff, R.R. y Grene, R. (2003) *Photosynthetic acclimation is reflected in specific patterns of gene expression in drought-stressed loblolly pine*. *Plant Physiol.* 133: 1702–1716.
- Weisz, R., Kaminski, J. y Smilowitz, S. (1994) *Water-deficit effects on potato leafgrowth and transpiration—utilizing fraction extractable soil-water for comparison with other crops*. *Am. Potato J*, 71: 829–840.
- Wohlbach, D.J., Quirino, B.F., Sussman, M.R. (2008) *Analysis of the Arabidopsis histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation*. *Plant Cell* 20, 1101-1117.
- Xiaolan, W., Sachie, K., Yukihiro, I. y Kinya, T. (2009) *Accumulation of raffinose in rice seedlings overexpressing OsWRKY11 in relation to desiccation tolerance*. *Plant Biotechnology*, Vol. 26, No. 4: 431-434

- Yamaguchi-Shinozaki K, Shinozaki K. (1994). A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* 6:251–64.
- Yamaguchi-Shinozaki K. y Shinozaki, K. (2005) Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science* 10:88-94.
- Yang, Y. H., and Thorne, N. P. (2003). Normalization for two-color cDNA microarray data. In: D. R. Goldstein (ed.), *Science and Statistics: A Festschrift for Terry Speed, IMS Lecture Notes - Monograph Series, Volume 40*, pp. 403-418.
- Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., y Speed, T. P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Research* 30(4):e15.
- Yordanov, I., Velikova, V. y Tsonev, T. (2003). Plant responses to drought and stress tolerance. *Bulg. J. Plant Physiol.*, 187–206.
- Yoshioka, K. y Shinozaki, K. (2009) *Signal Crosstalk in Plant Stress Responses*. Yoshioka y Shinozaki (Editores). Wiley-Blackwell. ISBN: 978-0-8138-1963-1. pp 224.
- Yu J. et al. (2002); A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *indica*) *Science* 296, pág 79.
- Zhu, J. K (2001) Plant salt tolerance. *Trends in Plant Science*, Vol. 6, (2): 66-71.
- Zhu, J.K. (2002). Salt and drought stress signal transduction in plants. *Annu. Rev. Plant Biol.* 53: 247–273.