



Universidad Autónoma de Querétaro  
Facultad de Ingeniería  
Doctorado en Ingeniería

Determinación de redes de comunicación molecular en tabaco transgénico que expresa el gen CchGLP.

Opción de titulación  
Tesis

Que como parte de los requisitos para obtener el Grado de  
Doctor en Ingeniería

Presenta:  
Germán Cárdenas Manríquez

Dirigido por:  
Dr. Ramón Gerardo Guevara González

Dr. Ramón Gerardo Guevara González  
Presidente

Dr. Irineo Torres Pacheco  
Secretario

Dr. Aurelio Domínguez González  
Vocal

Dr. Mario M. González Chavira  
Suplente

Dr. Juan Carlos Jaúregui Correa  
Suplente

Dr. Manuel Toledano  
Director de la Facultad

Firma  
Firma  
Firma  
Firma  
Firma  
Firma

Dra. Ma. Guadalupe Flavia Loarca Piña  
Directora de Investigación y Posgrado

Centro Universitario  
Querétaro, Qro.  
Noviembre, 2018.

## RESUMEN

Especies reactivas de oxígeno, peróxido de hidrógeno, redes de comunicación molecular, estrés biótico y abiótico, mecanismos de defensa en plantas.

Las plantas requieren evaluar y responder los estímulos del ambiente adecuadamente, las especies reactivas de oxígeno son moléculas clave en las respuestas al estrés. Previamente en el grupo de trabajo se desarrollaron líneas constitutivamente transfectadas con el gen CchGLP de *Capsicum chinense* que producen altos niveles de peróxido de hidrógeno endógeno. Los niveles de peróxido expresados se correlacionaron con tolerancia a infecciones por geminivirus y estrés por sequía.

El análisis transcriptómico diferencial entre la línea de alta expresión de peróxido (L8) y la línea control de la transfección (L1) mostró cambios en los perfiles relacionados a metabolismo basal, metabolismo especializado y rutas metabólicas relacionadas a estrés.

Los resultados sugieren que los cambios en las concentraciones de peróxido de hidrógeno inducen patrones transcriptómicos, proteómicos y metabolómicos que permiten a la línea transfectada tolerar estrés biótico y abiótico.

Parece ser que la producción constitutiva de niveles específicos de peróxido de hidrógeno dentro de las plantas podría generar una protección contra las condiciones de cultivo adversas.

## SUMMARY

Oxygen reactive species, hydrogen peroxide, molecular communication networks, biotic and abiotic stresses defense mechanisms in plants.

Plants require effective sensing and responses to environmental stimuli, reactive species of oxygen are well-recognized key signal molecules for stress response. Previously in our workgroup were developed several plant lines of constitutively producing germin-like gene from *Capsicum chinense* which produce high levels of endogenous produced peroxide. The levels of peroxide correlate with tolerance to geminivirus infection and drought stress. The differential transcriptomic analysis of the higher H<sub>2</sub>O<sub>2</sub> expressing line (L8) vs transfection control line (L1) showed changes in profiles related to basal metabolism, specialized metabolism and stress-related pathways. Results suggested that changes in peroxide concentrations induce transcriptomic, proteomic and metabolomics patterns that allow tolerance to biotic and abiotic stress.

It seems that constitutive production of specific levels of hydrogen peroxide within the plants might allow plant protection for different adverse conditions in farming.

Al Dr. Ramón por su gran paciencia y enseñanza.

Al Dr. Irineo por ser guía, calma y ejemplo.

Al Dr. Aurelio por su amistad y apoyo incondicional.

Al Dr. Mario por su visión y ayuda.

Al Dr. Juan Carlos por su paciencia y amable escucha.

A mi hermana Vicky por ser la razón y centro de este viaje.

A mi esposa Ana Karen por su amor que no conoce límites y lo ilumina todo.

A mi madre por su increíble e incondicional amor.

A mi padre por mostrarme el camino de la ciencia que tanto amó.

A Marcela por ser mi modelo a seguir, maestra y amiga, además de hermana.

A Julio por enseñarme la humildad y el carisma, por mantener el optimismo.

A ti, que eres pero no estás.

A ti, mi amor moderno, que te adelantaste en el camino.

A mi Hermosa Camila Virginia por hacerme feliz como no sabía que podría ser.

A ti que casi llegas a mi vida y la iluminas.

## **AGRADECIMIENTOS**

Agradezco el apoyo brindado por el Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico brindado mediante la beca de estudios de doctorado con CVU 102046 durante el periodo 01 de agosto del 2014 al 31 de julio del 2018.

Agradezco al director del posgrado Dr. Juan Carlos Jáuregui Correa y al coordinador del programa de doctorado de la Facultad de ingeniería Dr. Irineo Torres Pacheco, por su apoyo y guía, a todo el personal docente administrativo de la Universidad Autónoma de Querétaro.

Agradezco a los Directores y Coordinadores de Investigación y Posgrado de todas las Facultades de la Universidad Autónoma de Querétaro, así como de investigadores, académicos y personal administrativo de la misma.

En particular, a la Dirección de Servicios Escolares y la Dirección de Investigación y Posgrado por el desarrollo del presente formato y en especial al Mtro. Francisco Perusquía Monroy, quien dedicó tiempo a la mejora y corrección del mismo.

## TABLA DE CONTENIDOS

<b>1.</b>	<b>INTRODUCCIÓN.....</b>	<b>10</b>
1.1	ANTECEDENTES.....	10
1.2	DESCRIPCIÓN DEL PROBLEMA. ....	11
1.3	JUSTIFICACIÓN. ....	12
1.4	HIPÓTESIS. ....	13
1.5	OBJETIVOS. ....	13
1.5.1	<i>Objetivo general.</i> .....	13
1.5.2	<i>Objetivos particulares.</i> .....	14
1.6	CONTENIDO DE LA TESIS.....	14
<b>2.</b>	<b>REVISIÓN DE LA LITERATURA.....</b>	<b>15</b>
2.1	MECANISMOS INTRÍNSECOS DE DEFENSA DE LAS PLANTAS.....	15
2.2	REDES DE COMUNICACIÓN Y NIVELES DE MODULACIÓN DE LOS MECANISMOS DE DEFENSA DE LAS PLANTAS.18	18
2.3	VARIEDAD DE TABACO TRANSGÉNICA AL GEN CCHGLP DE CHILE HABANERO.....	19
2.4	HERRAMIENTAS BIOINFORMÁTICAS.....	21
<b>3.</b>	<b>METODOLOGÍA. ....</b>	<b>22</b>
3.2	METODOLOGÍA TRANSCRIPTÓMICA.....	25
3.2.1	<i>Germinación de las semillas.</i> .....	25
3.2.2	<i>Extracción y purificación de ácido ribonucleico (ARN) y síntesis y purificación de ácido desoxiribonucleico complementario (ADNc).</i> .....	26
3.2.3	<i>Hibridación sustractiva en condiciones de supresión ..... <b>¡Error! Marcador no definido.</b></i>	
3.2.4	<i>Purificación de los fragmentos en gel de agarosa y ligación al vector pGEM-T Easy. <b>¡Error! Marcador no definido.</b></i>	
3.2.5	<i>Transformación de células Escherichia coli y selección de colonias de interés. .... <b>¡Error! Marcador no definido.</b></i>	
3.2.6	<i>Extracción de ADN plasmídico y digestión con la enzima Eco R I. ..<b>¡Error! Marcador no definido.</b></i>	
3.2.7	<i>Secuenciación.</i> .....	26
3.2.8	<i>Análisis bioinformático.</i> .....	27

3.3	METODOLOGÍA PROTEÓMICA .....	42
3.3.1	<i>Obtención de las plantas.</i> .....	42
3.3.2	<i>Extracción de proteínas totales por método del ácido tricloro acético y Cuantificación de proteínas.</i> .....	43
3.3.3	<i>Electroforesis en condiciones desnaturalizantes (SDS-PAGE) y purificación de proteínas para isoelectroenfoque.</i> .....	43
3.3.4	<i>Análisis bidimensional de proteínas</i> .....	44
3.3.5	<i>Primera dimensión.</i> .....	44
3.3.6	<i>Equilibrio de las tiras.</i> .....	44
3.3.7	<i>Segunda dimensión (SDS-PAGE).</i> .....	45
3.3.8	<i>Tinción de los geles.</i> .....	45
3.3.9	<i>Análisis por espectrometría de masas.</i> .....	45
3.3.10	<i>Análisis bioinformático.</i> .....	45
3.4	METODOLOGÍA METABOLÓMICA. ....	47
3.4.1	<i>Esterilización, germinación de semillas, colecta de muestras y preparación de extractos.</i> .....	47
3.4.2	<i>Cuantificación de fenoles totales.</i> .....	48
3.4.3	<i>Cuantificación de flavonoids.</i> .....	48
3.4.4	<i>Cuantificación de taninos condensados.</i> .....	48
3.4.5	<i>Cuantificación de antocianinas.</i> .....	49
3.4.6	<i>Determinación de capacidad antioxidante por el ensayo de DPPH.</i> .....	49
3.4.7	<i>Determinación de capacidad antioxidante por el ensayo de ABTS.</i> .....	50
3.4.8	<i>Análisis de compuestos fenólicos en HPLC.</i> .....	50
3.4.9	<i>Análisis metabólico basado en GC-MS.</i> .....	51
3.4.10	<i>Análisis estadístico.</i> .....	51
3.4.11	<i>Análisis bioinformático.</i> .....	52
3.5	INTEGRACIÓN Y ANÁLISIS DE LAS REDES DE COMUNICACIÓN MOLECULAR. ....	53
3.5.1	<i>Gráfico de red de comunicación molecular.</i> .....	55
<b>4.</b>	<b>CONCLUSIONES.</b> .....	<b>55</b>
<b>5.</b>	<b>LITERATURA CITADA.</b> .....	<b>55</b>

6. ANEXOS.....62

## ÍNDICE DE TABLAS

TABLA 1. Proteínas diferenciales detectadas entre las líneas transgénicas de tabaco L1 y L8. Fuente: Cárdenas-Manríquez <i>et, al.</i> , 2016. ....	47
TABLA 2. Lista de rutas metabólicas modificadas por transcritos y proteínas reprimidas.....	53

## ÍNDICE DE FIGURAS

FIGURA 1-1 Ubicación geográfica de los principales centros de investigación enfocados a mecanismos de defensa intrínsecos de las plantas. Fuente: <a href="http://edit.freemap.jp/en/trial_version/edit/world">http://edit.freemap.jp/en/trial_version/edit/world</a> .....	15
FIGURA 3-1 Diagrama de flujo para obtener la información a analizar y generar las redes comunicación molecular. ....	24
FIGURA 3-2 Niveles ómicos de estudio y su interrelación con las bases de datos para determinar diferencias fenotípicas. Fuente: <a href="http://www.encuentros.uma.es/encuentros100/omicas.htm">http://www.encuentros.uma.es/encuentros100/omicas.htm</a> 02/03/2016.....	25
FIGURA 3-3 Diagrama tipo pie de la distribución de la información disponible de secuencias sobre expresadas contra las bases de datos.....	28
FIGURA 3-4. Distribución de especies con las que más coincidencias de mapeo se dieron para los datos usados durante el análisis de secuencias sobre expresadas. ....	30
FIGURA 3-5. Distribución de los tipos de proteínas encontrados al analizar las secuencias sobre expresadas en la línea transgénica (Interpro Scan).....	31
FIGURA 3-6. Histograma de frecuencia para los procesos biológicos en que intervienen las secuencias sobre expresadas.....	32

FIGURA 3-7. Histograma de frecuencia de las 50 secuencias de ARNm sobre expresadas más relevantes en el análisis contra su proceso biológico.....	33
FIGURA 3-8. Histograma de frecuencia de secuencias sobre expresadas contra su localización subcelular. ....	34
FIGURA 3-9. Histograma de frecuencias de transcritos sobre expresados por función a nivel molecular.....	35
FIGURA 3-10. Diagrama tipo pie de la distribución de la información disponible de secuencias reprimidas contra las bases de datos.....	36
FIGURA 3-11. Distribución de especies con las que más coincidencias de mapeo se dieron para los datos usados durante el análisis de secuencias reprimidas.....	37
FIGURA 3-12. Distribución de los tipos de proteínas encontrados al analizar las secuencias reprimidas en la línea transgénica (Interpro Scan). ....	38
FIGURA 3-13. Histograma de frecuencia para los procesos biológicos en que intervienen las secuencias reprimidas. ....	39
FIGURA 3-14. Histograma de frecuencia de las 50 secuencias de ARNm reprimidas más relevantes en el análisis contra su proceso biológico.....	40
FIGURA 3-15. Histograma de frecuencia de secuencias reprimidas contra su localización subcelular. ....	41
FIGURA 3-16. Histograma de frecuencias de transcritos reprimidos por función a nivel molecular. ....	42

## 1. Introducción.

### 1.1 Antecedentes.

Las plantas son organismos inmóviles que continuamente son desafiados por estreses bióticos (debido a otros organismos) y abióticos (debido a factores ambientales) y tienen mecanismos de defensa intrínsecos que les permiten combatir el estrés.

Todas las plantas tienen mecanismos de defensa, que se activan mediante proteínas y algunas de ellas se generan en eventos aleatorios. Por lo tanto, es deseable detectar estos activadores para propagarlos controladamente en cultivos de interés, esto permitirá proponer mejores estrategias de manejo de la fitosanidad, que contribuirán a disminuir el daño por estrés biótico y abiótico en la producción agrícola.

La accesión BG-3821 del chile *Capsicum chinense*, mostró resistencia contra geminivirus en las plantas que lo contenían [Garzón-Tirado *et al.*, 1996]. Se le confrontó exitosamente contra virus [Anaya-López *et al.*, 2002; HéARNdez-Godinez *et al.*, 2002] y se encontró que la proteína CchGLP (proteína parecida a la germina de *Capsicum chinense*) era la responsable principal de la respuesta defensiva.

La proteína CchGLP es una enzima Superóxido-Dismutasa con centro de manganeso (Mn-SOD) y su actividad enzimática se caracterizó como productora de peróxido de hidrógeno (H<sub>2</sub>O<sub>2</sub>). Se ha reportado en diversos artículos que el H<sub>2</sub>O<sub>2</sub> está directamente vinculado con los mecanismos defensivos de las plantas [León-Galván *et al.*, 2012; Jones and Dangl, 2008; Muthamilasaran and Prasad, 2014].

Para estudiar el efecto individual de la proteína CchGLP se construyó un transgén (inserción de información genética exógena) cuyo promotor de expresión es constitutivo, lo que significa que la proteína CchGLP es producida constantemente y en cantidades importantes, aumentando así el contenido de H<sub>2</sub>O<sub>2</sub> en los tejidos de la planta, lo que redundará en un incremento de respuesta defensiva.

El transgén CchGLP se insertó en tabaco (*Nicotiana tabacum*) y se evaluó el incremento en la capacidad defensiva de la planta, encontrando que la presencia de CchGLP aumentó la resistencia de los cultivos donde se insertó. [Ruíz-Nito *et al.*, 2012; Guevara-Olvera *et al.*, 2012].

Se insertó el transgén CchGLP en tabaco (*Nicotiana tabacum*) que es una planta modelo de estudio porque; tiene altas tasas de transformación génica; no requiere cuidados especiales y tiene una tasa de crecimiento acelerado.

Para comprender los mecanismos moleculares de defensa de las plantas el modelo provisto por el tabaco transgénico a CchGLP es una herramienta muy útil, ya que permite conocer los diferentes mecanismos de respuesta defensiva mediada por esta proteína, es decir, las redes de comunicación molecular.

Conociendo estas redes podremos inferir gran parte del resultado de otras aproximaciones que activen la respuesta defensiva de las plantas mediante las mismas enzimas tipo SOD, ya sea elicitando cultivos o insertando genes; o bien, encontrando otros métodos que promuevan la activación de la respuesta defensiva de las plantas.

## 1.2 Descripción del problema.

Las plantas son la fuente principal de alimento del ser humano, su producción y distribución son vitales. El estrés biótico y abiótico causa pérdidas en la producción ocasionando desde malas características organolépticas en el producto hasta pérdidas entre el 70 - 80 % de la cosecha (Pedreschi *et al.*, 2012)

Los métodos actuales de control de la fitosanidad incluyen métodos físicos, químicos y biológicos. Los más comunes son los químicos pues se les considera la opción más barata, no obstante, causan daño ocupacional y ecológico. Los métodos físicos y biológicos no son usados por ser caros y/o poco flexibles.

Las plantas poseen un sistema inmune que puede protegerlas de estrés biótico y abiótico, y conocer este sistema inmune contribuiría a diseñar tratamientos más eficaces para combatir el estrés ambiental y esto se puede lograr conociendo las redes de comunicación molecular activadas por CchGLP.

Determinar las redes y sus mecanismos de activación mediado por enzimas SOD permitirá comprender los mecanismos de defensa de las plantas vinculados a peróxido de hidrógeno, y por tanto, será posible diseñar mejores tratamientos basados en elicitación (exposición por aspersion a sustancias inocuas que activan el sistema defensivo), transformación génica (inserción de genes que expresan proteínas de interés en la defensa) u otros mecanismos.

### 1.3 Justificación.

Las plantas son la principal fuente de alimentos para el ser humano. Debido al incremento de la población previsto que alcanzará 9 a 12 millares de personas para el año 2050, la producción de alimentos se deberá incrementar entre el 34 y 70 %. (Pedreschi *et al.*, 2013).

Para aumentar la producción se requiere disminuir la incidencia de los fitopatógenos que merman la producción y mejoras en el manejo post-cosecha. (Pedreschi *et al.*, 2013). Los virus son importantes patógenos de las plantas responsables del 47% de las epidemias y este valor aumentará ya que constituyen el grupo más grande de enfermedades nuevas y emergentes en las plantas. (Palukaitis *et al.*, 2013).

El combate de patógenos se realiza principalmente por sustancias químicas y frecuentemente se han propuesto métodos de control biológico como alteARNtivas

a los pesticidas para reducir el impacto biológico (Agrios., 2005; Barrera-Pacheco *et al.*, 2008; Buonauro *et al.*, 2009; Anderson *et al.*, 2010; Mejía-Teniente *et al.*, 2010; Kim *et al.*, 2012; Foley *et al.*, 2013; Almario *et al.*, 2014).

Se sabe que se puede inducir la expresión del sistema defensivo de las plantas mediante el uso de elicitores, moléculas inocuas a los vertebrados y el medio ambiente (Mejía-Teniente *et al.*, 2014). Sin embargo, se desconoce el total de los mecanismos que inducen, silencian y modulan las respuestas del sistema defensivo debido a la enorme cantidad de datos generados en los diferentes modelos de plantas de interés. (Jones y Dangl, 2006). Al ser muchos los eventos celulares que se entrecruzan con los mecanismos de defensa es importante reconocer la influencia de los sistemas de defensa intrínseca de las plantas en cada uno de estos niveles, esto es, un estudio de redes moleculares (Muthamilarasan y Prasad., 2013). Las redes de comunicación molecular aportan información sobre el resultado de un estímulo específico en la planta en un estado único, son útiles para analizar la integración de grandes cantidades de información biológica con alta exactitud, en menos tiempo del que solía hacerse en análisis individual (Nana *et al.*, 2014). Analizar las redes de comunicación molecular permitirá ofrecer una nueva opción en fitosanidad con mejores características económicas, de impacto a la salud y al ambiente.

#### 1.4 Hipótesis.

El fenotipo de las plantas de tabaco transgénicas al gen CchGLP, permite determinar las redes de comunicación molecular de defensa activadas por enzimas productoras de peróxido de hidrógeno, mediante el análisis de la información de diferentes niveles ómicos..

#### 1.5 Objetivos.

##### 1.5.1 Objetivo general.

Determinar redes de comunicación de los mecanismos moleculares de defensa en tabaco transgénico que expresa el gen CchGLP.

### 1.5.2 Objetivos particulares.

1. Generar perfiles ómicos diferenciales en planta azigota.
2. Generar perfiles ómicos diferenciales en planta transgénica.
3. Definir las redes de comunicación molecular en planta azigota.
4. Definir las redes de comunicación molecular en planta transgénica.
5. Comparar las redes de comunicación molecular en planta azigota y planta transgénica de tabaco.

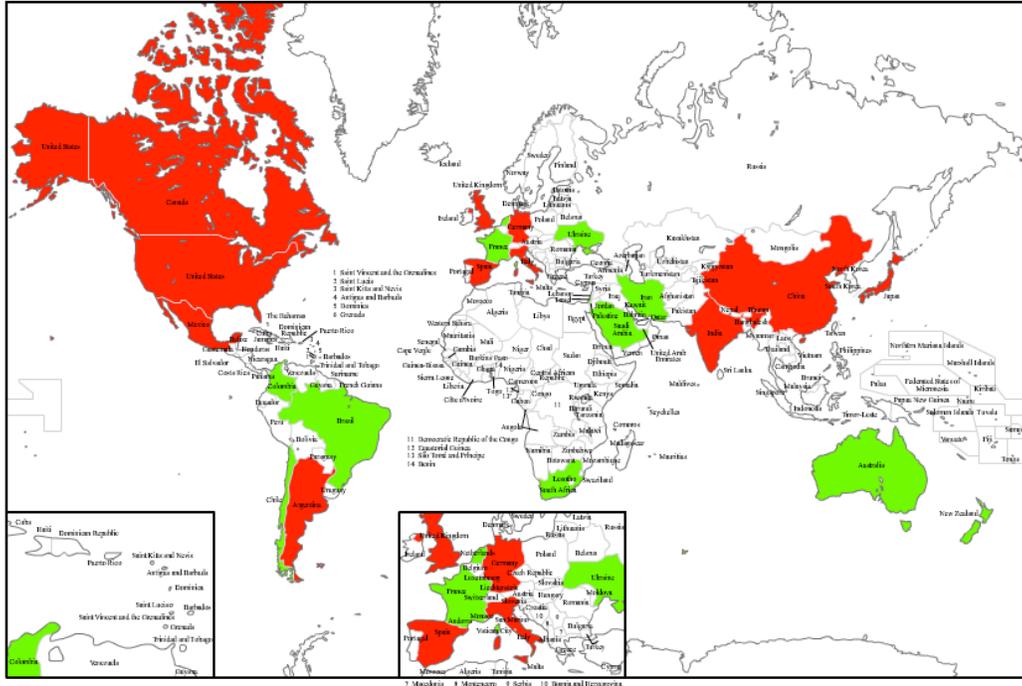
### 1.6 Contenido de la tesis.

El primer capítulo contiene la información necesaria para comprender el estado actual de la producción agrícola y sus problemas presentes y futuros; la necesidad de producir alimentos nutritivos para la creciente población que impulsa el desarrollo del presente trabajo; los avances científicos y conocimientos previos que se han realizado en el campo de desarrollo de la presente tesis; se enuncia el problema existente y la mejora que aportará el presente estudio en este campo. En este capítulo también se postula la hipótesis de trabajo, el objetivo general y los objetivos particulares y se describe brevemente el contenido de esta tesis. Se presenta un mapamundi que muestra las tendencias globales actuales en el campo de investigación de la presente tesis.

El segundo capítulo presenta de forma cronológica las más recientes investigaciones y la información técnica necesaria para comprender los las bases de los sistemas de defensa intrínsecos de las plantas y su activación.

El tercer capítulo detalla los procedimientos metodológicos, los resultados y la discusión sobre su valor o relevancia.

El cuarto capítulo resume las conclusiones derivadas de la discusión de resultados.



**FIGURA 1-1** Ubicación geográfica de los principales centros de investigación enfocados a mecanismos de defensa intrínsecos de las plantas. Fuente: [http://edit.freemap.jp/en/trial\\_version/edit/world](http://edit.freemap.jp/en/trial_version/edit/world).

En color rojo se denotan los países en los que se encuentran universidades y grupos de investigación en general que se han dedicado por más de 10 años a desarrollar investigación reconocida en el área de cultivo y relacionado a mecanismos de defensa de las plantas. En color verde se colorean los países donde se desarrolla investigación universitaria o general reconocida desde hace menos de 10 años. Este mapa muestra las áreas con mayor interés en el desarrollo de mejoras en las tecnologías de cultivo existentes y que podrían ser susceptibles de requerir los conocimientos de personal capacitado en ésta área del conocimiento humano.

## 2. Revisión de la literatura.

### 2.1 Mecanismos intrínsecos de defensa de las plantas.

Los avances en las investigaciones sobre los mecanismos de señalización durante la defensa, la regulación del estoma y las interacciones entre un gen que confiere resistencia y el efector que lo activa, así como la producción de toxinas específicas al huésped, permiten comprender la relación planta-patógeno y el resultado en el balance de esa relación. La cuestión de cómo las plantas han desarrollado la capacidad de distinguir entre simbioses, como los rizobios y las micorrizas, de sus muchos enemigos debe estar relacionada también con los mecanismos de defensa. (Anderson *et al.*,2010). La mayoría de los microorganismos asociados a las plantas son los patógenos que afectan el crecimiento y reproducción, y éstas suelen responder a la infección con la activación de su sistema inmune innato (Jones y Dangl., 2006; Mejía-Teniente *et al.*,2010). En términos generales, la inmunidad se refiere al estado fisiológico que tiene suficientes defensas biológicas para evitar la infección, la enfermedad o la invasión biológica no deseada. Las plantas reconocen y responden a moléculas comunes de diferentes tipos de microorganismos, incluyendo factores de avirulencia no patógenos y patógenos que son reconocidos directamente o a través de sus efectos en el hospedero. (Mejía-Teniente *et al.*,2010).

Existen 2 grandes ramas en los mecanismos de defensa de las plantas, la defensa activada por patógenos PTI (inmunidad evocada por PAMPs) o respuesta inmune primaria que es conservada, no específica al patógeno, de respuesta lenta y sostenida, provocada por el reconocimiento de estructuras invariables de superficie microbiana llamados patrones moleculares asociados a patógenos (PAMP's / Ma) a través de las proteínas receptoras llamadas receptores de reconocimiento (PRR's) importante para orientar la inmunidad a infecciones microbianas en todas las especies vegetales. Los PAMP's son normalmente los componentes esenciales de toda clase de agentes patógenos, como la flagelina bacteriana y la quitina de los hongos. Las plantas también responden a las moléculas endógenas liberadas por la invasión de un patógeno, como fragmentos de la pared celular o fragmentos cuticulares llamados patrones moleculares asociados a peligro o daño (DAMP's). La segunda es debida a señales

específicas del patógeno sobre el tejido o ETI (inmunidad evocada por efector) es especializada, específica al patógeno, de respuesta fuerte y rápida, activada mediante receptores intracelulares a moléculas de virulencia del patógeno, llamadas efectoras y a menudo implica una forma de muerte celular localizada denominada la respuesta de hipersensibilidad (HR). Ambas utilizan muchas vías de transducción de señales de diversos niveles, desde las involucradas en vías de crecimiento y diferenciación celular (como MAPK), así como control de los flujos iónicos, respuestas vinculadas a especies reactivas de oxígeno, mecanismos de reforzamiento de la pared celular, reorganización del citoesqueleto, traslocación de organelos, modificación del tráfico de vesículas, cambios en la citoarquitectura celular, modificación de sistemas endomembranales llegando incluso a la respuesta hipersensitiva o muerte súbita de la célula vegetal (Jones y Dangl., 2006; Mejía-Teniente *et al.*, 2010; Muthamilarasan y Prasad, 2013).

El estrés es usualmente definido como un factor externo que ejerce una alteración en la homeostasis de la planta iniciando un conjunto de respuestas fisiológicas o bioquímicas que definen a un estado en particular del organismo diferente al que se observa bajo un rango de condiciones óptimas. Por lo tanto estrés biótico se define como el estrés que se produce como resultado de los daños causados a las plantas de otros organismos vivos, y el abiótico el ocasionado por factores ambientales. El estrés se inicia cuando las plantas reconocen el estrés a nivel celular, activando las vías de transducción de señales que transmiten información dentro de cada célula y a lo largo de la planta, lo que da lugar a cambios en la expresión de muchas redes génicas. El concepto de estrés está estrechamente asociado con tolerancia al estrés, es decir, la capacidad de la planta para hacer frente con éxito a cierto entorno desfavorable y permanecer bajo un estado de estrés sin sufrir cambios significativos en su fenotipo. Los enfoques modernos de producción agrícola incluyen la aplicación de mecanismos de envío de señales de estrés (muchos de los cuales dependen de la generación de especies reactivas de oxígeno), como una estrategia para aumentar la tolerancia de las plantas contra diversos factores ambientales. (Mejía-Teniente. *et al.*, 2010).

## 2.2 Redes de comunicación y niveles de modulación de los mecanismos de defensa de las plantas.

Recientemente se ha realizado un cambio en el que se usa la ontología génica para evaluar datos de redes moleculares y viceversa, usar los datos de redes moleculares para construir y evaluar ontologías genéticas. Existe evidencia que una gran parte de la ontología genética puede ser reconstruida a partir de topologías de redes moleculares. En diversas áreas de la investigación el análisis de la ontología juega un rol importante en la unificación del conocimiento y jerarquización de las relaciones entre seres vivos. Esta información ayuda a la descripción de los procesos biológicos, función molecular y componentes celulares. La ontología genética se ha vuelto la mayor fuente de información en anotación de genes y proteínas entre varias especies, proveyendo herramientas para la valoración sistemática de conjuntos de genes obtenidos por los nuevos análisis de secuencias. La ontología génica ha crecido en tamaño y complejidad, pero no deja de ser un campo difícil, y adicionalmente el rápido desarrollo de nuevas tecnologías de adquisición de datos biológicos ha generado una acumulación de información que excede nuestra habilidad para interpretarlos. Para superar esos problemas se han desarrollado diversas herramientas computacionales que realizan el análisis para la anotación de genes y proteínas. Se ha realizado un gran esfuerzo en evaluar la precisión de estas predicciones. La disponibilidad de variada información de nivel génico, interacciones proteínas y genes, perfiles de expresión de ARNm y vías metabólicas, ha creado nuevas oportunidades de predicción, comparadas contra los viejos métodos basados solamente en similitud de secuencias o de función. Se ha demostrado que la integración de datos biológicos complementarios mejora significativamente la precisión de la predicción en la anotación de función de genes desconocidos. Esta información ha sido usada en la predicción de interacciones huésped-hospedero durante infecciones, interacciones proteína-proteína y detección de asociaciones gen-fenotipo. (Gligorijevic *et al.*, 2014).

Un número creciente de experimentos ha sido diseñado para detectar las interacciones moleculares intracelulares e intercelulares. Basados en estas interacciones se generan las redes moleculares que se usan en muchas aplicaciones típicas como el descubrimiento de nuevos genes relacionados a infecciones y la identificación de blancos para nuevos fármacos y complejos moleculares. Debido a que los bancos de datos están incompletos y el alto número de falsos positivos en las interacciones, las interacciones proteicas se integran a redes de análisis para construir una red molecular estable. Obtener una red molecular estable depende de los métodos bioinformáticos usados, los resultados obtenidos de una red y su confiabilidad puede variar dependiendo de estos métodos, donde los más integrales han mostrado ser los que mejor predicen los resultados. Las interacciones moleculares tales como proteína-ADN, proteína-ARN, ADN-ADN, ARN-ARN y proteína-proteína facilitan varias funciones de los organismos, como la transcripción, interacción entre promotores y elementos distales, y la regulación de la expresión génica. Las redes moleculares se construyen basadas en éstas interacciones para elucidar sus roles subyacentes en la biología. Entre los problemas de confiabilidad está la falta de uso de varios métodos de análisis sobre las múltiples interacciones para comparar la información, la falta de anotaciones completas, las diferentes bases matemáticas y probabilísticas de los métodos, la discriminación de falsos positivos obtenidos. (Nana *et al.*, 2014).

### 2.3 Variedad de tabaco transgénica al gen CchGLP de chile habanero.

Un gen que codifica para una proteína relacionada con las germinas (CchGLP) fué clonado de la línea BG-3821 resistente a geminivirus de *Capsicum chinense* Jacq. Se caracterizó y la actividad enzimática de la proteína expresada fue analizada. (Anaya-López *et al.*, 2003, León-Galván *et al.*, 2011).

La Germina y las proteínas relacionadas a la Germina (GPLs) han sido implicadas en el refuerzo de resistencia de la pared celular de plantas, por lo que proveen resistencia a factores de estrés biótico y abiótico. Una característica clave de las

subfamilias relacionadas a las GLPs incluyendo germinas, es la presencia de cajas de germina y segmentos conservados de la superfamilia de cupinas (proteínas con estructura secundaria tipo bucle o hélice beta). Esta estructura contiene tres residuos de histidina y un glutamato altamente conservados y participan del enlace de un ión metálico. La proteína consta de 203 aminoácidos y su peso molecular es de 20.5 kDa, similar a la de otras proteínas relacionadas a la germina. Contiene un péptido de señal en los primeros 18 aminoácidos en el extremo amino-terminal, así como una N-glicosilación al nivel de los residuos 44-47. La secuencia de ADNc de CchGLP contiene una sección de control de la transcripción de 24 pares de bases, un segmento de poli Adeninas, y un marco de lectura abierto que codifica una proteína de 203 aminoácidos. Una secuencia putativamente importante llamada KGD (o a veces KGE) que se ha detectado en más de 50% de GPLs, pero no en germinas, y fue detectada en la proteína recombinante del gen CchGLP. El número de aminoácidos en la proteína obtenida, así como las características de la secuencia fueron típicas las reportadas para GLPs en la literatura. La glicosilación en germinas parece ser esencial en las interacciones proteína-proteína, pero no para la actividad de la enzima. Las GPLs se han clasificado en subfamilias según sus actividades oxalato oxidasa (OXO) o superóxido dismutasa (SOD). Las actividades OXO y SOD generan la producción de peróxido de hidrógeno. La proteína recombinante muestra actividad manganeso superóxido dismutasa (Mn-SOD). Se ha reportado que la actividad Mn-SOD ocurre extracelularmente, en la mitocondria y peroxisomas, y se les ha vinculado con los mecanismos de defensa contra estrés biótico en las plantas. El análisis Molecular mostró que CchGLP está presente en una sola copia en el genoma de *C. chinense Jacq* y fue inducida en las plantas con aplicaciones de etileno (Et) y ácido salicílico (SA) pero no por ácido jasmónico (JA) en la ausencia de agentes patógenos. CchGLP se consideró una proteína relacionada a la patogénesis (PR) de la familia PR-16. Las proteínas PR se definen como proteínas inducidas por el huésped específicamente en situaciones patológicas. No se acumulan localmente en las hojas infectadas, pero son

sistémicamente inducidas, asociado con el desarrollo de resistencia sistémica adquirida (SAR) contra una nueva infección por hongos, bacterias y virus. (León-Galván *et al.*, 2011).

La cepa BG-3821 de *C. Chinense* ha demostrado ser un excelente fenotipo para estudiar las interacciones planta-patógeno. La expresión de la proteína recombinante no parece estar vinculada al número de copias del gen. Esta variabilidad en la expresión del transgén se atribuye a muchos factores incluyendo las diferencias en la posición dentro del cromosoma, secuencias repetidas y número de copias. La producción de especies reactivas de oxígeno (ROS) fue mayor en la línea transgénica que expresa el gen CchGLP comparada contra la silvestre. El incremento en la producción de peróxido de hidrógeno se puede deber al promotor constitutivo de la expresión del transgén CchGLP. Sin embargo se considera que otros genes pueden estar envueltos en la respuesta diferencial de especies reactivas de oxígeno (ROS). El hecho que no se percibió el aumento de expresión de CAT-1, proteína involucrada en el metabolismo de peróxido de hidrógeno, en las plantas transgénicas que contenían el gen CchGLP sugiere que este transgén activa otros genes que inducen un incremento en la producción de ROS. Sin embargo la tolerancia a los ataques de los geminivirus lograda con el transgén CchGLP no fue completa, ya que se presentó sintomatología ligera, lo que sugiere que se necesitan otras respuestas vinculadas a la defensa para lograr la inmunidad total (Guevara-Olvera *et al.*, 2012).

#### 2.4 Herramientas bioinformáticas.

Recientemente se usa la ontología génica para evaluar datos de redes moleculares y viceversa. Existe evidencia que una gran parte de la ontología genética puede ser reconstruida a partir de topologías de redes moleculares. La ontología juega un rol importante en la unificación del conocimiento y jerarquización de las relaciones entre seres vivos. Esta información ayuda a la descripción de los procesos biológicos, función molecular y componentes celulares. La ontología genética se ha vuelto la mayor fuente de información en

anotación de genes y proteínas entre varias especies, proveyendo herramientas para la valoración sistemática de conjuntos de genes obtenidos por los nuevos análisis de secuencias. El rápido desarrollo de nuevas tecnologías de adquisición de datos biológicos ha generado una acumulación de información que excede nuestra habilidad para interpretarlos. Para superar esos problemas se han desarrollado diferentes herramientas computacionales que realizan el análisis para la anotación de genes y proteínas. Se han realizado diversos esfuerzos en evaluar la precisión de estas predicciones. La disponibilidad de información de nivel génico, interacciones proteínas y genes, perfiles de expresión de ARNm y vías metabólicas ha creado nuevas oportunidades de predicción, comparadas contra los viejos métodos basados solamente en similitud de secuencias o de función. Se ha demostrado que la integración de datos biológicos complementarios mejora significativamente la precisión de la predicción en la anotación de función de genes desconocidos. Ésta información ha sido usada en la predicción de interacciones huésped-hospedero durante infecciones, interacciones proteína-proteína y detección de asociaciones gen-fenotipo. (Gligorijevic *et al.*, 2014).

Obtener una red molecular estable depende de los métodos bioinformáticos usados, los resultados obtenidos de una red y su confiabilidad puede variar dependiendo de éstos métodos donde los más integrales han mostrado ser los que mejor predicen los resultados. Entre los problemas de confiabilidad está la falta de uso de varios métodos de análisis sobre las múltiples interacciones para comparar la información, la falta de anotaciones completas, las diferentes bases matemáticas y probabilísticas de los métodos, la discriminación de falsos positivos obtenidos. (Nana *et al.*, 2014).

### **3. Metodología.**

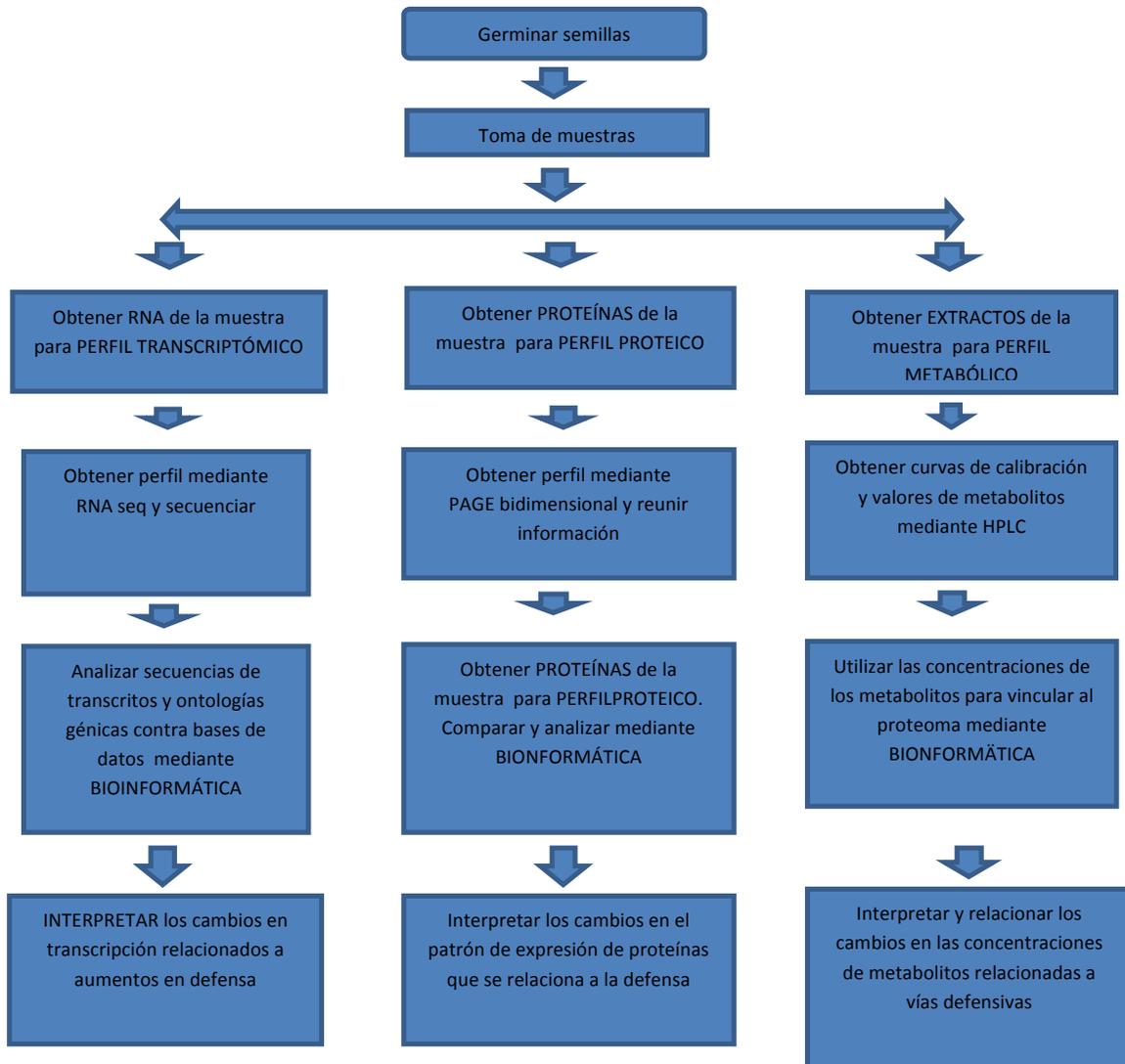
Se deberá trabajar 3 líneas simultáneas de muestras biológicas para los 3 niveles de información biológica a obtener. De cada una de ellas se generará un perfil biológico que sera analizado bioinformáticamente.

El análisis del transcriptoma permite determinar cuáles genes están transcribiéndose en mayor o en menor cantidad ante la sobreexpresión de peróxido de hidrógeno endógena provocada por la inserción del transgén CchGLP. Estas diferencias de expresión son las responsables de los cambios de función molecular y comportamiento celular que diferencia a la células que recibieron el transgén y lo expresan de aquellas que no (FIGURA 3-2)

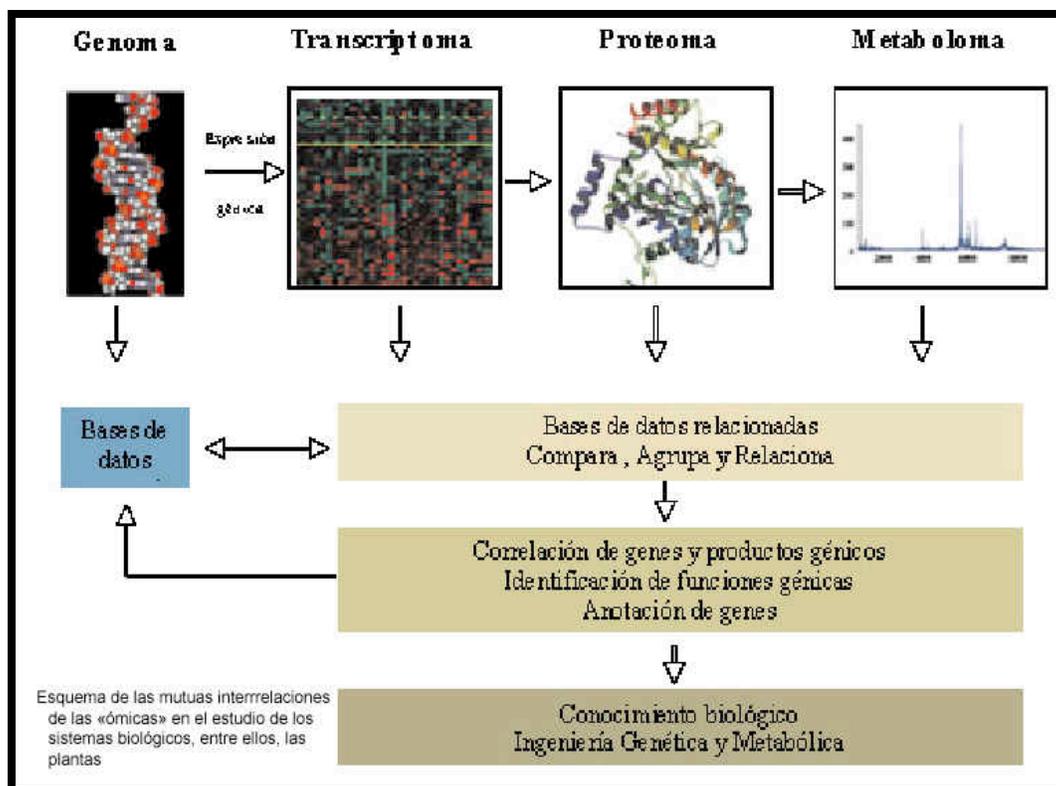
El análisis del proteoma permite determinar el cambio en la expresión de proteínas provocado por la sobreexpresión de peróxido de hidrógeno inducido por la inserción y expresión constitutiva del transgén CchGLP. Al cambiar los niveles de proteínas en la célula cambian también las actividades en las rutas metabólicas en las que participan debido a cambios internos o externos en la célula (FIGURA 3-2).

El análisis del metaboloma es parcial, ya que permitirá determinar el cambio relativo en los metabolitos analizados. Este cambio permite determinar el cambio en los productos de las rutas metabólicas en respuesta a la sobreexpresión de peróxido de hidrógeno y aumento de la respuesta defensiva generados por la inserción del transgén CchGLP (FIGURA 3-2).

### 3.1 Esquema general de trabajo.



**FIGURA 3-1** Diagrama de flujo para obtener la información a analizar y generar las redes comunicación molecular.



**FIGURA 3-2 Niveles ómicos de estudio y su interrelación con las bases de datos para determinar diferencias fenotípicas. Fuente:**

<http://www.encuentros.uma.es/encuentros100/omicas.htm> 02/03/2016.

## 3.2 Metodología transcriptómica.

### 3.2.1 Germinación de las semillas.

Se utilizaron semillas de dos líneas diferentes de plantas transgénicas al gen CchGLP de tabaco (*Nicotiana tabacum* Xanthi nc) la línea 1 (L1) no lo expresa y la línea 8 (L8) tiene altos niveles de expresión. Las semillas de la línea 1 se usaron como control. Se incubaron, para esterilizarlas superficialmente durante 20 minutos con hipoclorito de sodio al 13%, después se enjuagaron con agua estéril. Posteriormente, se colocaron directamente en medio MS adicionado con el marcador de restricción kanamicina (100 mg/mL) contenido en frascos previamente esterilizados, este proceso se llevó a cabo en condiciones asépticas en una campana de flujo laminar, para evitar posibles contaminaciones. Los

frascos con las semillas crecieron a 22°C bajo un fotoperiodo de 16 h de luz y 8 de oscuridad (Guevara-Olvera *et al.*, 2012).

### 3.2.2 Extracción y purificación de ácido ribonucleico (ARN) y síntesis y purificación de ácido desoxiribonucleico complementario (ADNc)

Se tomaron muestras de las hojas verdaderas de las plantas azigotas y transgénicas, se les realizó una extracción y purificación de ARN total con un híbrido de los protocolos de TRIzol® Reagent (Invitrogen, Carlsbad, CA, USA) y SV Total ARN Isolation System (Promega, Madison, WI, USA). Mediante el ARN se sintetizó ADNc de cadena sencilla de acuerdo con las instrucciones del manual PCR-Select™ ADNc subtraction kit (Clontech, Palo Alto, CA, USA). Se purificó el ADNc mediante el método descrito en el manual PCR-Select™ ADNc subtraction kit (Clontech, Palo Alto, CA, USA), para obtener un ADNc que aumentó la pureza mejorando los análisis posteriores.

### 3.2.3 RNA seq.

Revela la presencia y cantidad de RNA en una muestra biológica en una muestra en el momento específico usando tecnología de secuenciación de próxima generación. Facilita el monitoreo del cambiante transcriptoma al permitir analizar las diferencias entre transcritos derivados de cortes específicos (alternative splicing), modificaciones post-transcripcionales, fusiones génicas, mutaciones puntuales y polimorfismos de uno varios nucleótidos, cambios en la expresión génica a través del tiempo, diferencias en la expresión entre 2 o más grupos.

Permite el análisis entre diferentes subgrupos de RNA como rRNA, mRNA, miRNA, tRNA y perfiles ribosomales.

Se envió la muestra a Langebio para secuenciación en illumina Hiseq 2500.

### 3.2.4 Secuenciación.

El ADN plasmídico se mandó a secuenciar a Langebio del CINVESTAV Irapuato, Obteniendo las secuencias que posteriormente se analizaron mediante el programa BLAST (Altschul y col., 1990), análisis tipo BLASTn y BLASTx, para

comparar los genes con las bases de datos del National Center for Biotechnology Information (NCBI).

### 3.2.5 Análisis bioinformático.

Para analizar conjuntos de datos obtenidos por secuenciación de próxima generación para transcritos (ARN-seq) y/o librerías por hibridación sustractiva se instaló una partición de Linux fedora versión 22 y se descargaron e instalaron los siguientes programas desde la terminal: cufflinks, bowtie, tophat

Se realizaron mapeos contra el genoma reportado en la base de datos de NCBI mediante blast-n con corte estadístico del e-value de  $1 \times 10^{-6}$  para asegurar la confiabilidad de los resultados. Al no tener disponible aún un genoma de tabaco completo se comparó primero en marco de lectura libre y después contra el genoma de *A. thaliana* que arrojó resultados filogenéticamente próximos a tabaco.

Se realizó un mapeo con bowtie ([bowtie-bio.sourceforge.net/](http://bowtie-bio.sourceforge.net/)) y TopHat (<https://ccb.jhu.edu/>). Bowtie es una herramienta que analiza secuencias cortas (desde 50 pares de bases) y las alinea contra genomas relativamente grandes (mamíferos por ejemplo). TopHat es una herramienta que analiza las posibles secuencias de corte postranscripcional (splice) entre los exones (secuencias de los genomas que se transcriben en una secuencia de ARN funcional) para alinear aún mejor contra las secuencias reportadas en las bases de datos y generar un ensamble más completo del transcriptoma. También se les hizo el corte estadístico con e-value de  $1 \times 10^{-6}$ .

Después se realizó un análisis con Cufflinks ([cufflinks.cbc.umd.edu/](http://cufflinks.cbc.umd.edu/)) que permite ensamblar un transcriptoma completo con la mejor aproximación y después permite realizar normalización de los valores de expresión de los 2 grupos analizados, o bien, permite obtener un valor diferencial comparativo.

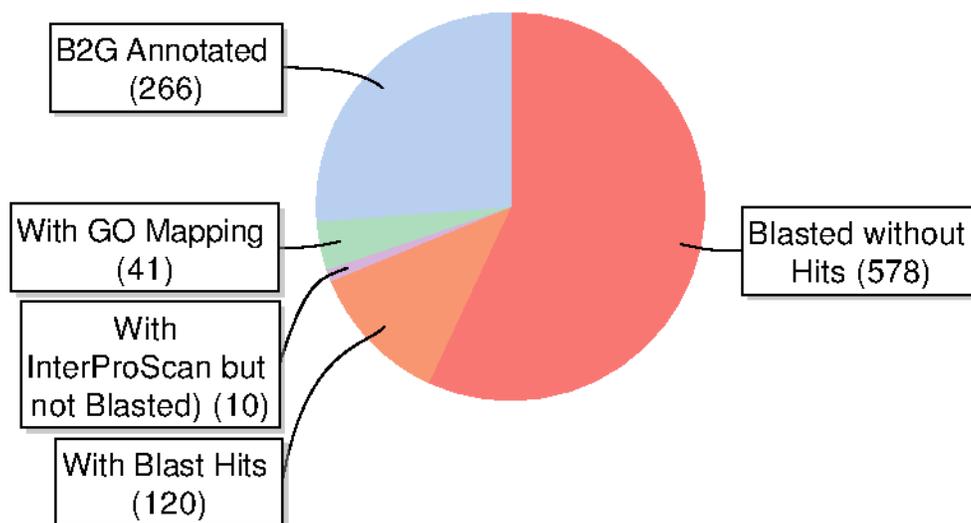
Para obtener una tabla diferencial normalizada (valores relativos de expresión de los transcritos que aumentan en un condición contra otra) se usó EdgeR con E

value < 0.1 (<http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>) en formato GFF.

Con las tablas de expresión de los transcritos que aumentaron su abundancia relativa en la línea experimental L8 (upregulated) y los que disminuyeron su abundancia relativa en la misma línea (downregulated) se realizó un análisis mediante Blast2GO.

Blast2GO es un programa que genera las Ontologías Génicas de las secuencias suministradas (la ubicación del gen, la proteína que produce, los procesos biológicos que controla, su localización en la célula y su función molecular) y permite analizar las posibilidades de expresión proteica de las secuencias propuestas para los marcos de lectura encontrados.

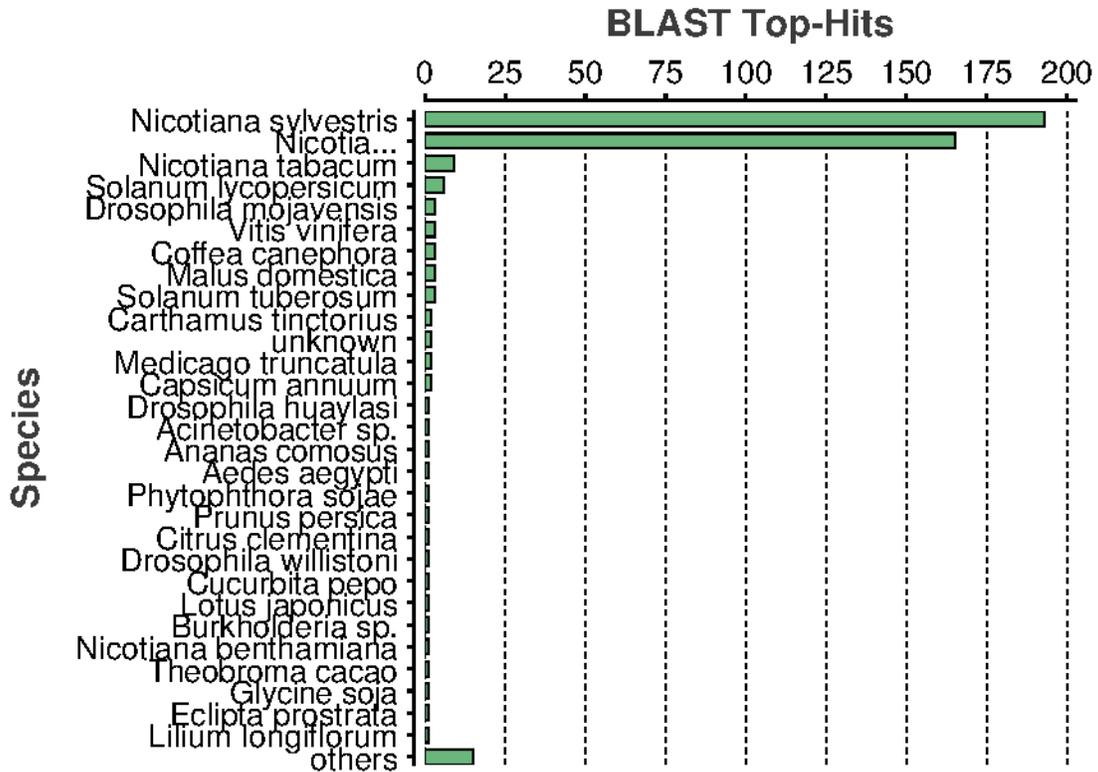
De tal manera se obtuvieron los siguientes resultados:



**FIGURA 3-3 Diagrama tipo pie de la distribución de la información disponible de secuencias sobre expresadas contra las bases de datos.**

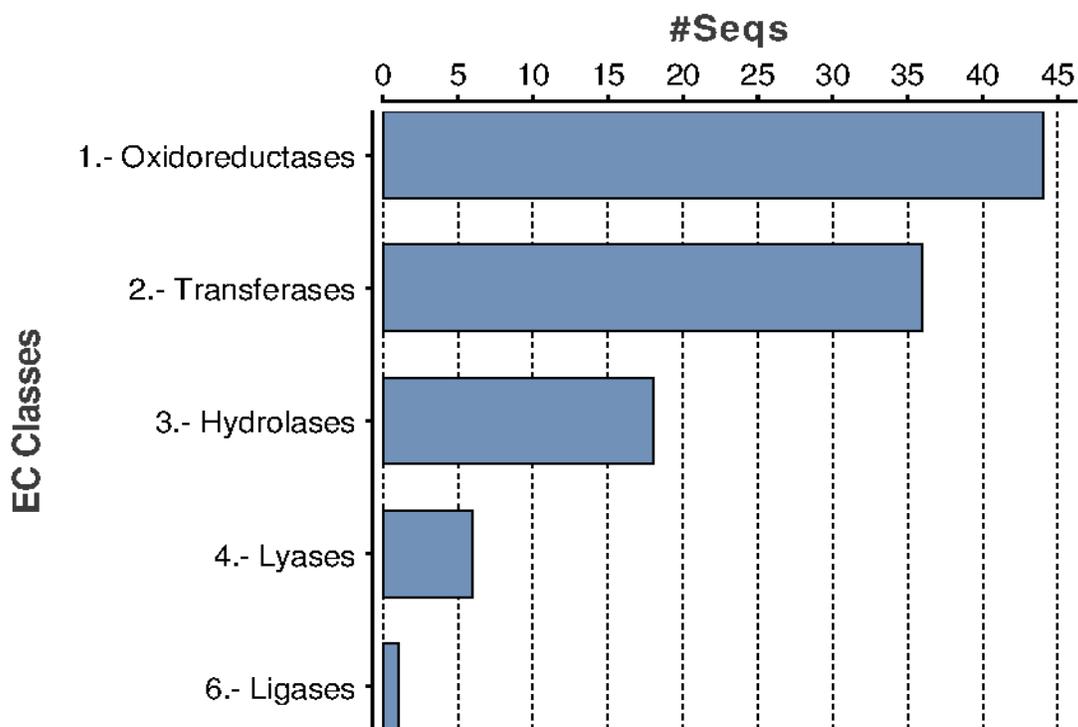
La FIGURA 3-3 muestra que hay 1015 secuencias de ARNm sobre expresadas en la línea 8 (que expresa el transgén CchGLP) lo que significa que ante la producción constitutiva de peróxido de hidrógeno hay aumento en 1015

mensajeros. Analizando en sentido de las manecillas del reloj, de estas secuencias 578 no tienen homólogos (blast hits) reportados en las bases de datos, es decir, son transcritos aún no reportados o nuevos, un número de 120 presentaron homología (with blast hits) pero no tiene reportada información acerca de su ontología génica (información precisa del gen y de su producto génico), esto implicaría que la secuencia de ARNm está reportada pero no se sabe aún para que proteína codifica y cuál es su función en la célula, un número de 10 secuencias coinciden con la secuencia de marco de lectura abierto para proteínas para las cuales se conoce su función, pero no se tiene la información del gen que la codifica (interpro scan but not blasted), la cantidad de 41 secuencias si tienen la información del gen que codifican y la proteína traducida, pero no se saben con certeza su los procesos biológicos en los que intervienen, su localización celular y/o su función a nivel molecular (with GO mapping). De tal manera que las 266 secuencia restantes son las que contienen información completa para realizar análisis comparativos (B2G annotated) y serán las más significativas durante el análisis.



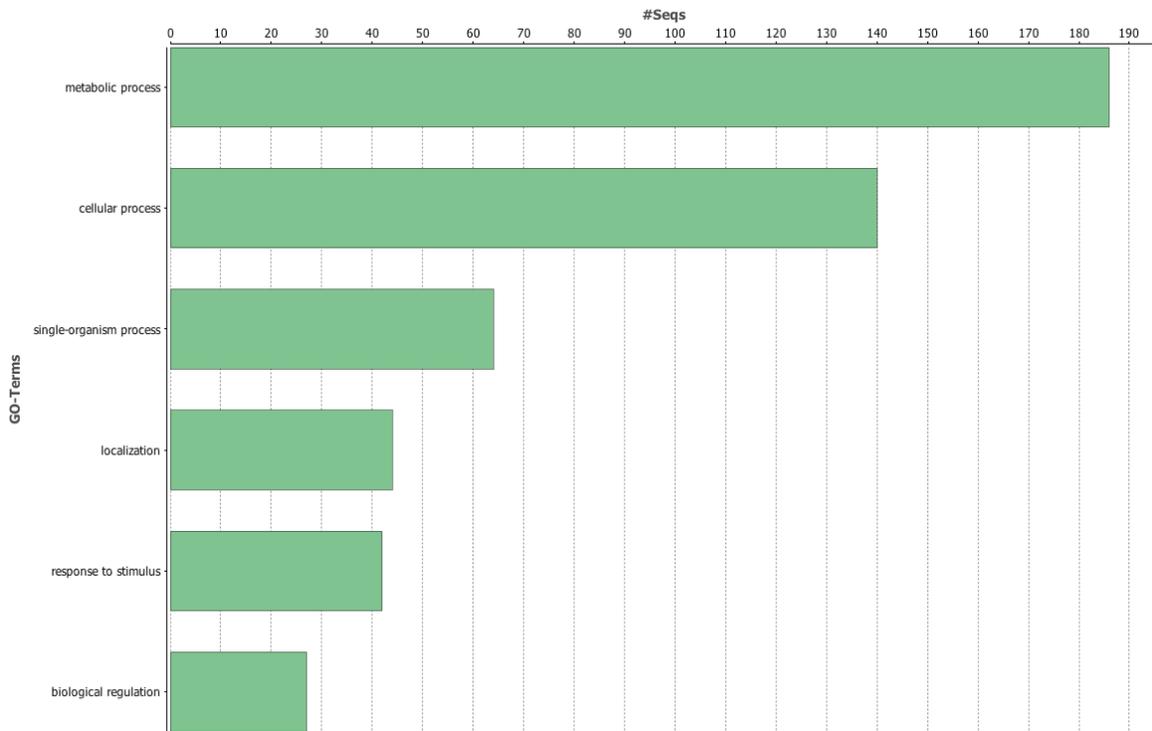
**FIGURA 3-4.** Distribución de especies con las que más coincidencias de mapeo se dieron para los datos usados durante el análisis de secuencias sobre expresadas.

La FIGURA 3-4 muestra que de las secuencias que sobre expresaron y están vinculadas a los mecanismos de defensa activados por peróxido de hidrógeno, 195 secuencias coincidieron con las reportadas para *N. Silvestris* y 165 coincidieron para *N. Benthamiana*, esto implica que la mayoría de secuencias reconocidas coinciden con secuencias del género de tabaco, las que no coinciden con especies de tabaco posiblemente sean secuencias relacionadas a la defensa de la plantas activada por peróxido de hidrógeno pero que aún no han sido reportadas para especies del género del tabaco. Esto aumenta la solidez de los resultados pues la mayoría de secuencias sí tienen reporte en la planta modelo que se usó.



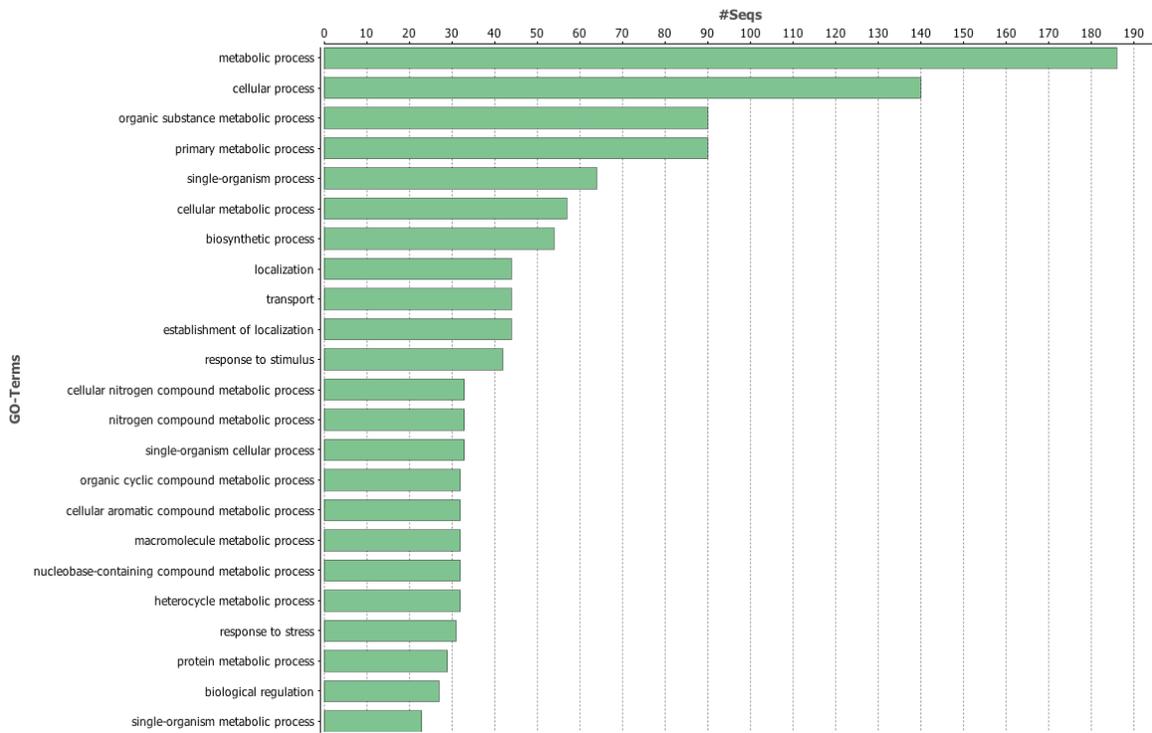
**FIGURA 3-5. Distribución de los tipos de proteínas encontrados al analizar las secuencias sobre expresadas en la línea transgénica (Interpro Scan).**

La FIGURA 3-5 muestra que de las 125 secuencias de ARNm que generaron una secuencia de proteína reportada y caracterizada, 44 (35.2%) son del grupo de las oxido-reductasas, esto muestra que si hay un aumento sustancial de peróxido de hidrógeno en la línea transgénica y que se activan mecanismos que recapturan este radical pues su exceso mataría a la planta, se sospecha un balance que aumenta la concentración intracelular del peróxido lo suficiente para promover la defensa a estrés activada por esta molécula sin dañar a la célula o tejidos.



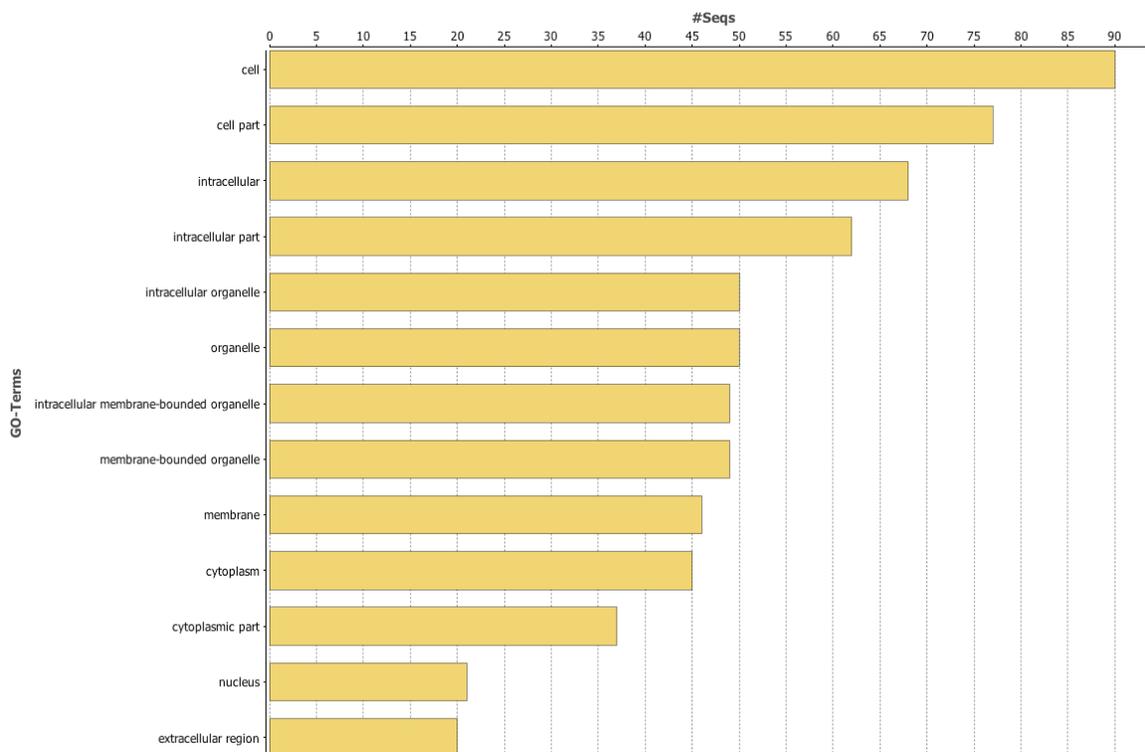
**FIGURA 3-6. Histograma de frecuencia para los procesos biológicos en que intervienen las secuencias sobre expresadas.**

El histograma figura 3-6 demuestra que el mayor cambio de procesos biológicos ocurre a nivel de procesos metabólicos, lo cual es de esperarse, pues la planta al estar sometida constitutivamente a estrés debe aumentar el gasto de energía para realizar sus funciones vitales y sobrepasar el estrés. El quinto grupo representa la respuesta a estímulos, como el estrés provocado por la producción constante de peróxido de hidrógeno, éste subgrupo será analizado a detalle para saber cuáles son las secuencias clave que responden al estímulo dañino proveyendo defensa a la planta.



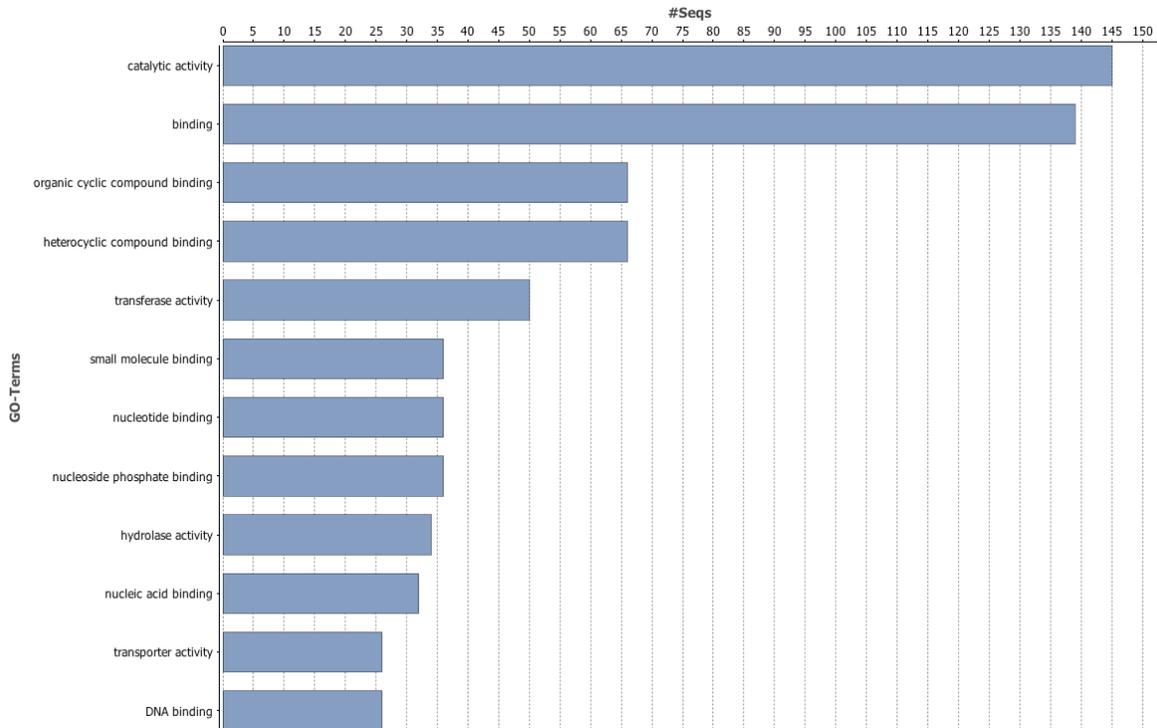
**FIGURA 3-7. Histograma de frecuencia de las 50 secuencias de ARNm sobre expresadas más relevantes en el análisis contra su proceso biológico.**

Al analizar las 50 secuencias con mayor información y que estadísticamente fueron más significativas según el filtrado de los mapeos ( $e$  value =  $1 \times 10^{-6}$ ) por su función biológica, se encontraron 31 transcritos sobre expresados cuya función específica es la respuesta a estrés, serán una de las metas de los análisis de subgrupos a realizar en los experimentos posteriores para determinar los 31 transcritos más relevantes según su proceso biológico.



**FIGURA 3-8. Histograma de frecuencia de secuencias sobre expresadas contra su localización subcelular.**

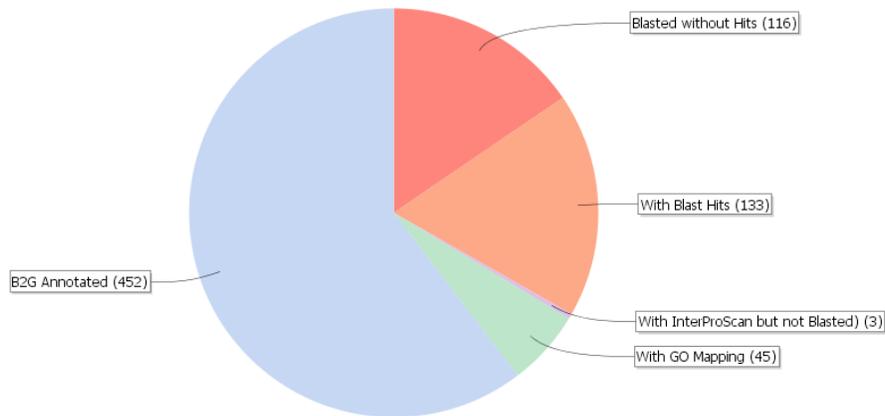
La FIGURA 3-8 muestra que más del 50% de las secuencias que se sobre expresan están localizadas en el interior celular. Se ha reportado que las especies reactivas de oxígeno activan proteínas del peroxisoma (Akpinar *et al.*, 2012), sin embargo encontramos una distribución homogénea de blancos de los transcritos analizados ya que la muestra fue de tejido completo. Una de las posibilidades en futuros estudios será segmentar organelos por ultracentrifugación y gradientes de sales como cloruro de Cesio para separar los componentes celulares y analizarlos por separado para determinar mediante la localización subcelular los tipos de enzimas que se activan al aumentar la concentración de peróxido de hidrógeno y promover la defensa.



**FIGURA 3-9. Histograma de frecuencias de transcritos sobre expresados por función a nivel molecular.**

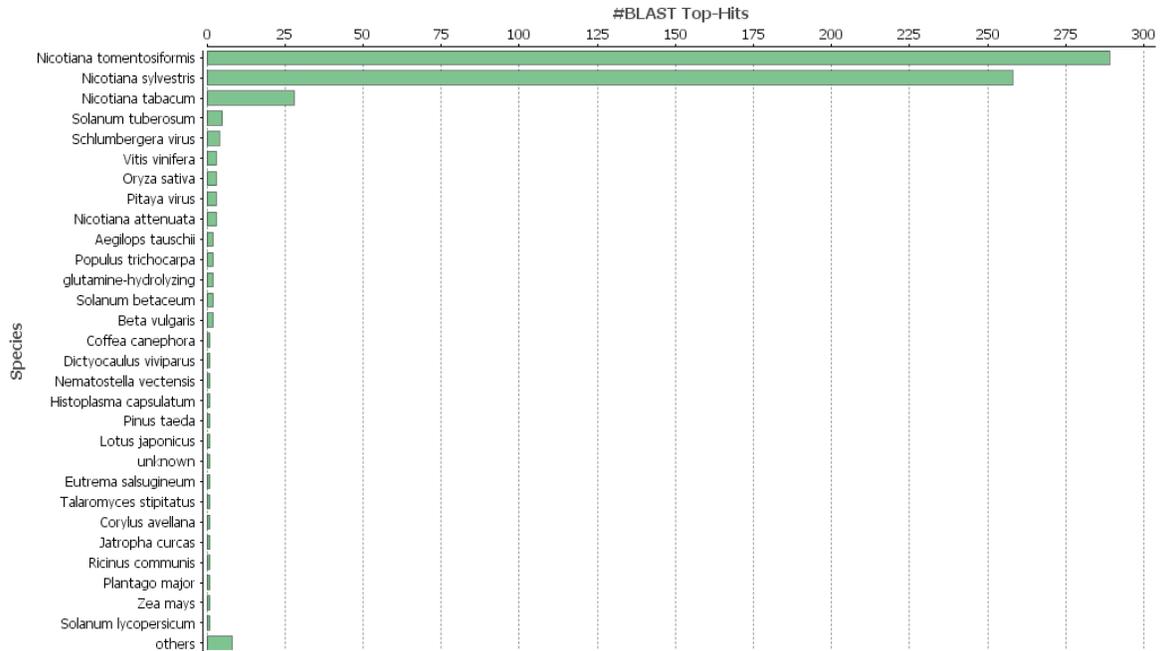
La FIGURA 3-9 muestra que una de las principales funciones moleculares de los transcritos sobre expresados es la actividad catalítica (relacionada a cambios en el metabolismo basal, metabolismo secundario y respuesta a estímulos). Muchos de los grupos restantes que se agrupan en el histograma muestran unión a ADN, nucleótidos o nucleósidos, lo que implica un cambio relevante en el uso de la información genética que posee el organismo, se sospecha que en respuesta al estrés promoviendo mecanismos de adaptación y defensa que sólo se despliegan ante activadores específicos. Además, se piensa que estos cambios no se darían si este activador (en este caso, peróxido de hidrógeno) no estuviese presente en las concentraciones necesarias.

Para los transcritos que se reprimieron los resultados fueron los siguientes:



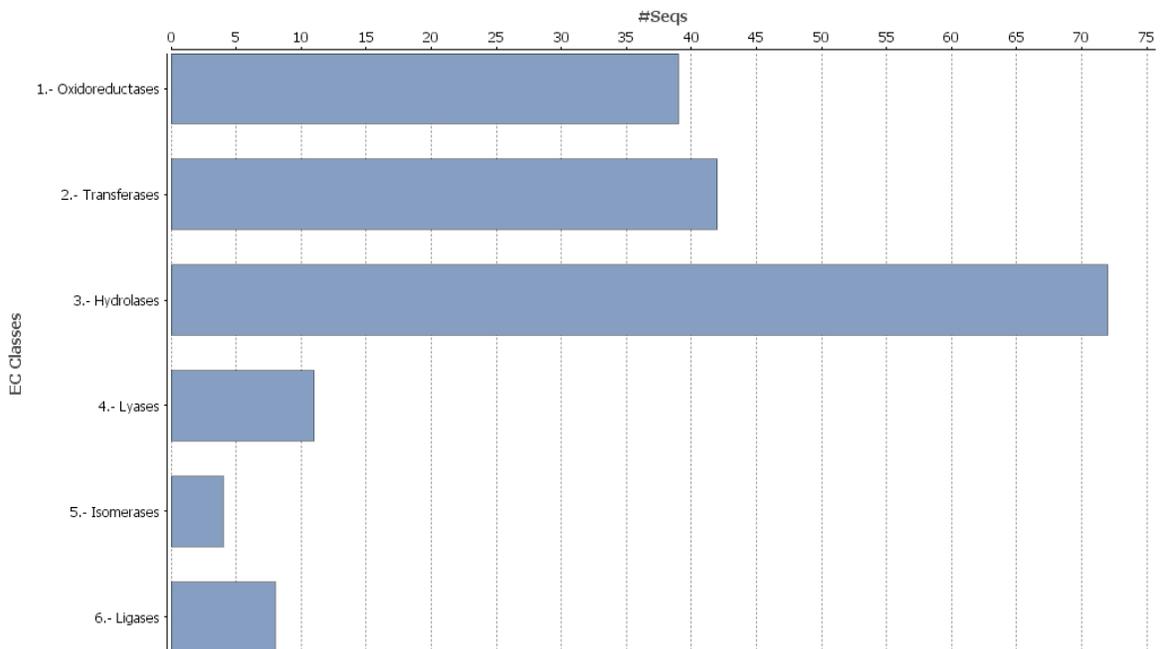
**FIGURA 3-10. Diagrama tipo pie de la distribución de la información disponible de secuencias reprimidas contra las bases de datos.**

La FIGURA 3-10 muestra que 749 transcritos disminuyeron su expresión relativa a causa de la expresión constitutiva de peróxido de hidrógeno debida al transgén CchGLP. Al analizarlos comenzando en sentido de las manecillas del reloj, 116 secuencias no tienen reporte en la base de datos o son nuevas (blasted without hits), 133 secuencias de ARNm sí están reportadas pero sólo tienen información del transcrito y no de la proteína que codifican (with blast hits), 3 secuencias sí tienen información de una proteína que putativamente codifican pero no del transcrito (with interpro Scan but not blasted), 45 secuencias tienen información del gen y la proteína pero no de sus acciones sobre la célula (with GO mapping). Y 452 tienen información completa del transcrito, la proteína y las funciones celulares (B2G annotated). La búsqueda principal de subgrupos se enfocará a este último.



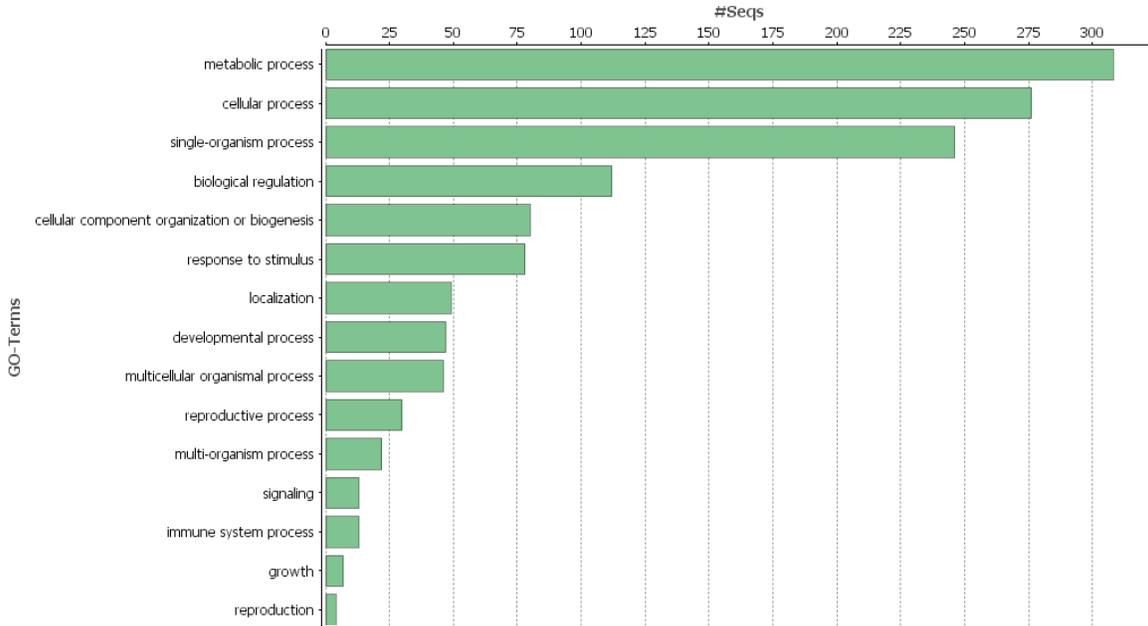
**FIGURA 3-11. Distribución de especies con las que más coincidencias de mapeo se dieron para los datos usados durante el análisis de secuencias reprimidas.**

El histograma presentado en la FIGURA 3-11 demuestra que la mayoría de transcritos reprimidos si están reportados para especies del género de tabaco, lo que demuestra que los resultados son estadísticamente significativos por ser este nuestro modelo de estudio. Los transcritos que se alejan filogenéticamente del tabaco como género y de las solanáceas en general pueden ser genes muy conservados entre las plantas que se reprimen durante eventos defensivos vinculados a la exposición ante peróxido de hidrógeno o aún no han sido reportados para tabaco y solanáceas en general.



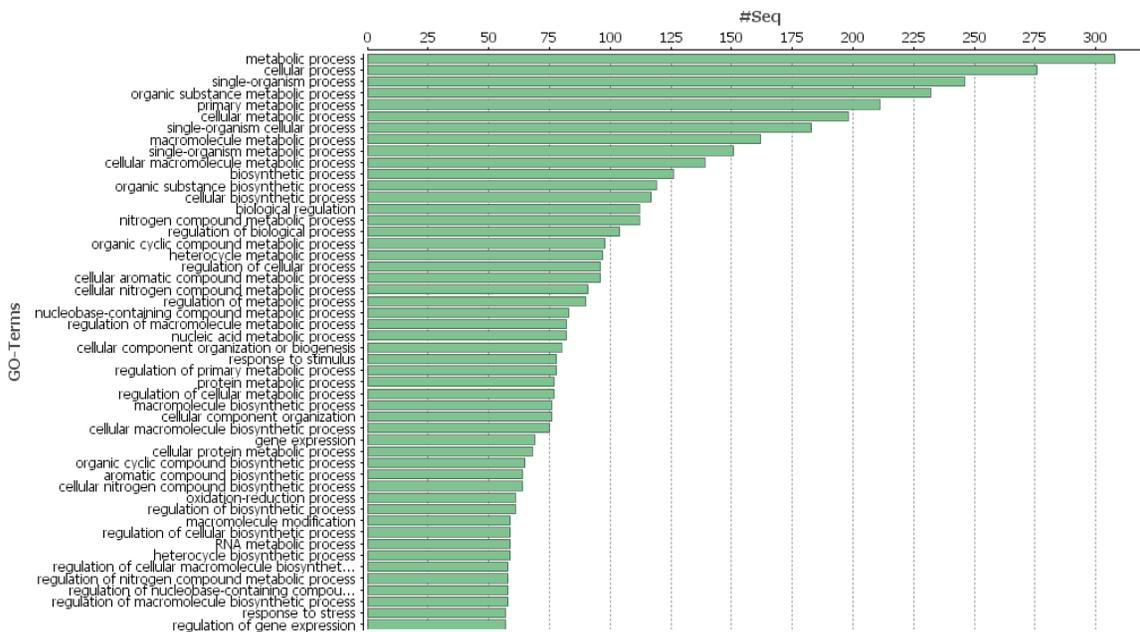
**FIGURA 3-12. Distribución de los tipos de proteínas encontrados al analizar las secuencias reprimidas en la línea transgénica (Interpro Scan).**

El histograma presentado en la FIGURA 3-12 muestra que el grupo de transcritos que disminuye más es el de las hidrolasas, concuerda con la sospecha que el catabolismo se ve restringido a vías metabólicas basales, la degradación por recambio de componentes disminuye ante la presión metabólica de utilizar preferentemente la energía para mecanismos defensivos disminuyendo algunas vías catabólicas no esenciales. En el primer grupo se ve también represión de oxido-reductasas, la sospecha es que disminuye la oxido-reducción en algunos procesos anabólicos no esenciales para evitar fugas energéticas innecesarias, sin embargo, éste sub-grupo será analizado para comprobar que no esté relacionado a oxido-reducción de especies reactivas de oxígeno que se contraponen a lo esperado.



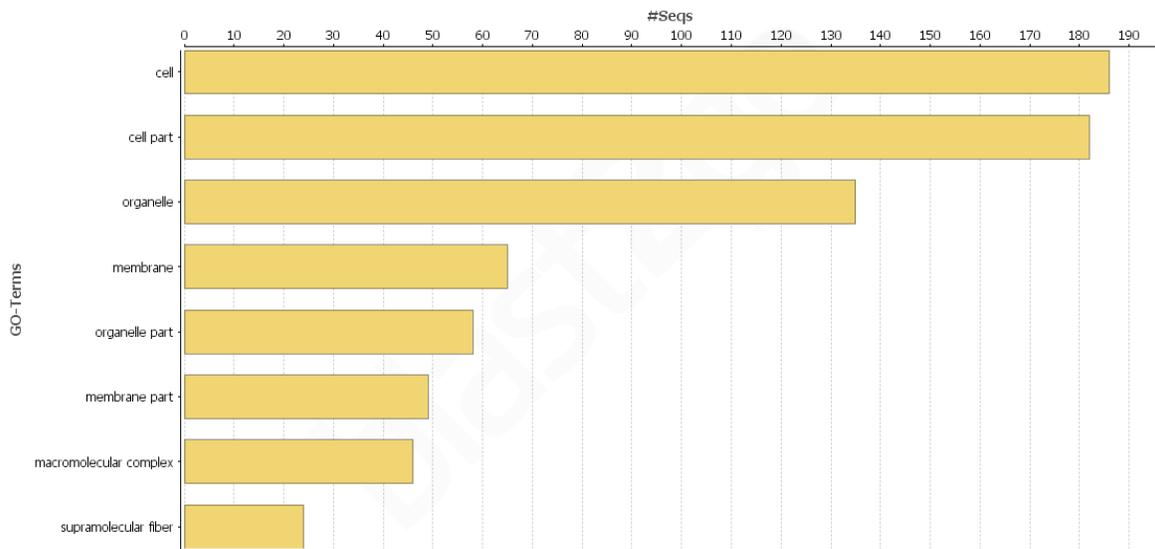
**FIGURA 3-13. Histograma de frecuencia para los procesos biológicos en que intervienen las secuencias reprimidas.**

El histograma de la FIGURA 3-13 muestra concordancia con su antagónico de la FIGURA 3-6, algunos procesos metabólicos se reprimen y se sospecha que es debido al entrecruce de vías metabólicas basales no esenciales que se apagan mientras en los transcritos sobre expresados las rutas metabólicas vinculadas a respuesta a estímulos como el estrés se encienden. La barra correspondiente a la respuesta a estímulos muestra más de 75 transcritos que se reprimen y esto podría deberse a que se reprime la transcripción de elementos clave en las vías de irritabilidad (entrada del estímulo) que generan respuestas laxas o bien se reprime la transcripción de elementos importantes en vías de excitabilidad (respuesta ante el estímulo) que generan respuestas lentas. Esto coincidiría con la sospecha que la célula cambia su configuración de transcripción hacia un estado de transcripción de máxima producción de energía que se use en los mecanismos de defensa más rápidos y contundentes para sobrevivir al estrés. Y estos eventos estarían desencadenados por el aumento del peróxido de hidrógeno.



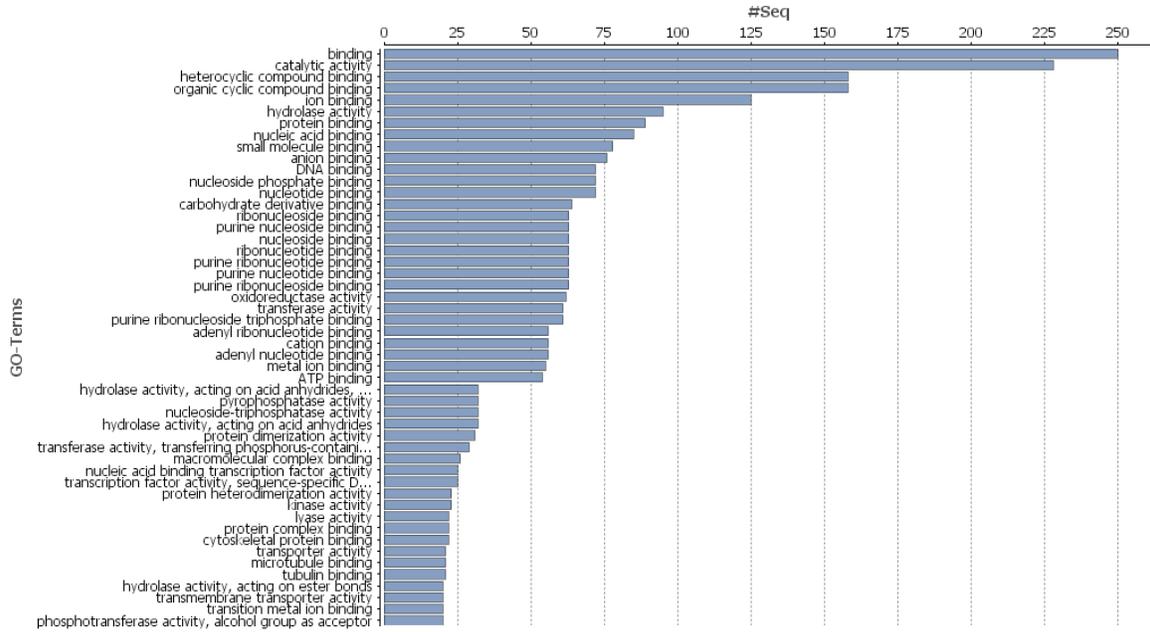
**FIGURA 3-14. Histograma de frecuencia de las 50 secuencias de ARNm reprimidas más relevantes en el análisis contra su proceso biológico.**

La FIGURA 3-14 muestra un histograma donde hay muchos grupos de procesos biológicos con transcritos reprimidos, la mayoría están relacionados con eventos metabólicos de naturaleza muy diversa, los grupos que más dominan intervienen en procesos biosintéticos, lo que podría implicar que muchas rutas anabólicas no esenciales sean reprimidas para aumentar el estatus energético y dirigirlo hacia la defensa celular y tisular. El histograma muestra en penúltima posición el grupo de respuesta a estrés lo que significa que son pocos los transcritos de defensa a estrés que se reprimen, esto concuerda con la idea de reforzar los mecanismos de defensa ante el aumento de peróxido de hidrógeno.



**FIGURA 3-15. Histograma de frecuencia de secuencias reprimidas contra su localización subcelular.**

La FIGURA 3-15 muestra que más del 50% de las secuencias que se sobre expresan están localizadas en el interior celular. Al igual que su contraparte, la figura 9, será necesario en posteriores estudios analizar estructuras específicas mediante métodos de separación de organelos para determinar la represión de funciones características de algunos partes de la célula que arrojen más información sobre el estado metabólico.



**FIGURA 3-16. Histograma de frecuencias de transcritos reprimidos por función a nivel molecular.**

En la primera mitad del histograma representado en la FIGURA 3-16 muestra que la función molecular más coincidente es la de unión a ADN o elementos del mismo, lo que implicaría que muchos de los elementos que se reprimen pueden ser mecanismos de bloqueo de secuencias de ADN que codifican para elementos defensivos, que sólo ante un estímulo específico se reprimen permitiendo así que se transcriban transcritos usados en condiciones muy específicas. Algunos de estos subgrupos serán analizados a fondo en secciones posteriores de este trabajo.

### 3.3 Metodología proteómica

#### 3.3.1 Obtención de las plantas.

Las semillas se esterilizaron con hipoclorito de sodio al 2%, Tween 20 ® al 0.1% por cinco minutos, se enjuagaron con agua destilada estéril. Se colocaron 10 semillas por frasco en medio de cultivo Murashige-Skoog (1962). Estas actividades se realizaron en una campana de flujo laminar. Los frascos se mantuvieron a temperatura ambiente con un fotoperiodo de 16 horas de luz y 8 de

oscuridad. Una vez germinadas, las plantas se mantuvieron en medio de cultivo durante 6 semanas, después se transfirieron a suelo. La aclimatación se llevó a cabo en macetas con sustrato peat-moss ® previamente esterilizado y consistió en cubrir las macetas con bolsas estériles y perforarlas cada tres días. Las plantas se dejaron crecer 4 meses.

### 3.3.2 Extracción de proteínas totales por método del ácido tricloro acético y Cuantificación de proteínas.

Se pesaron 5 g de tejido fresco y se molieron con nitrógeno líquido, se mezclaron con 15 ml de solución de extracción (ver anexo), la mezcla se agitó por 15 minutos en hielo, después se centrifugó a 10000 g durante 10 minutos a 4°C, se recuperó el sobrenadante, posteriormente se agregaron 3 volúmenes de solución de precipitación (ver anexo) y se dejó precipitar por 4 horas. Se centrifugó, lavó, secó y re suspendió en 1 ml de solución de rehidratación (ver anexo). Una vez obtenido el extracto de proteínas, se realizó la cuantificación para determinar la concentración de los extractos. Se utilizó el método de Bradford y albúmina de suero bovino (BSA) como proteína estándar a una concentración de 10, 15, 25, 50, 100 y 150 µg/ml. Se tomaron 100 µl de extracto y se agregó 1 ml de reactivo de Bradford (ver anexo), se mezcló y se dejó reposar en oscuridad por 5 minutos, se leyó la absorbancia a 595 nm (Bradford, 1976)

### 3.3.3 Electroforesis en condiciones desnaturalizantes (SDS-PAGE) y purificación de proteínas para isoelectroenfoque.

Se realizaron electroforesis SDS-PAGE con el fin de analizar la calidad y pureza de los extractos proteicos. Esta metodología permite separar por peso molecular, las proteínas de las hojas de tabaco. Se analizaron 10 µg de proteína por muestra, se mezclaron con buffer de muestra (Ver anexo) en un volumen final de 20 µl. La electroforesis fueron realizadas en mini-geles al 12% con una cámara vertical Mini-PROTEAN® a temperatura ambiente y 15 mA constantes por gel de 0.75 mm por dos horas o hasta que el frente de corrida llegó al tope. Los geles se tiñeron con azul de coomasie (Laemmli, 1970). La presencia de sales, lípidos, ácidos fenólicos y nucleicos, pueden interferir con la separación en el IEF y en la visualización de

los ensayos bidimensionales, por esto, las proteínas fueron tratadas con el ReadyPrep™ 2-D Cleanup Kit (Bio-Rad Laboratories, Inc. California, USA), utilizando el protocolo propuesto por la casa comercial.

#### 3.3.4 Análisis bidimensional de proteínas

La electroforesis bidimensional se basa en una separación de las proteínas en una dimensión con base a su punto isoeléctrico (isoelectroenfoque) y en una segunda dimensión con base a su peso molecular (SDS-PAGE). La metodología realizada se basa en la propuesta por Weiss y Görg (2007). Una vez determinada la concentración de proteínas se realizó el isoelectroenfoque.

#### 3.3.5 Primera dimension.

Se utilizó el equipo Protean® i12™ (Bio-Rad) y tiras de poliacrilamida de 7 cm con un gradiente de pH inmovilizado (IPG) de 3-10 (Bio-Rad). Se colocaron 125 µl de la muestra de proteína limpia por carril de la bandeja de rehidratación, en seguida se colocaron las tiras IPG en contacto con la muestra evitando la formación de burbujas y enseguida se agregaron 2 ml de aceite mineral (Bio-Rad, USA) para evitar la evaporación de la muestra, así como la cristalización de la urea. Se utilizó el método Running Preset (Bio-Rad) para correr el IEF, este consistió en una rehidratación activa a 50 V por 16 h a una temperatura de 20°C, inmediatamente se realizó el paso de acondicionamiento a 250 V por 15 minutos, posteriormente se incrementó el voltaje de manera gradual hasta 4000 V como máximo por 1 h. El enfoque final consistió en mantener el voltaje al máximo y completar un total de 15000 V. Todo esto se realizó evitando que el amperaje excediera los 50 µA/tira y a una temperatura de 20°C.

#### 3.3.6 Equilibrio de las tiras.

Se equilibraron las tiras de IPG adicionando 2 ml de buffer de equilibrio I (ver anexo 5) por 15 minutos en agitación para reducir los grupos sulfhídricos. Después se decantó el buffer de equilibrio I y se agregaron 2 ml de buffer de equilibrio II para alquilar los grupos sulfhídricos, se agitó durante 15 minutos y posteriormente se colocó en el gel de poliacrilamida para correr la segunda dimensión.

### 3.3.7 Segunda dimensión (SDS-PAGE).

Después del equilibrio, las tiras IPG se colocaron en geles SDS-PAGE al 12%, en los cuales las proteínas se separaron por su PM de forma perpendicular a la primera dimensión. Una vez colocada la tira se agregó solución de sellado (ver Anexo 5) La electroforesis fue realizada en minigeles con una cámara vertical Mini-PROTEAN® a temperatura ambiente a 15 mA constantes por gel de 0.75mm por dos horas o hasta que el frente de corrida llegó al tope. Los geles se tiñeron con azul de coomassie (Laemmli, 1970).

### 3.3.8 Tinción de los geles.

Tinción con azul de Coomasie: Los geles se sumergieron durante 30 minutos bajo agitación suave en solución de fijado (metanol 50%, ácido acético 10%), una vez retirada la solución de fijado se sumergieron durante toda la noche bajo agitación en solución de azul brillante de Coomassie R-250 al 0.025% (ver Anexo 6) y se destiñeron mediante varios lavados con solución de desteñido hasta observar los geles libres de tinción de fondo.

Tinción con SYPRO® Ruby: Los geles se sumergieron durante 1 hora en solución de fijado (ver Anexo 6) bajo agitación, después se lavaron 3 veces con agua Milli-Q durante 10 minutos bajo agitación. Una vez lavados, se sumergieron en Sypro® Ruby Protein Gel Stain (Molecular Probes™) y se agitaron durante toda la noche, protegidos de la luz. Una vez teñidos los geles, se transfirieron a un contenedor con solución de lavado y se agitaron durante 30 minutos.

### 3.3.9 Análisis por espectrometría de masas.

Una vez teñidos los geles fueron enviados a LABANMA para la secuenciación de los péptidos. Usando el programa MASCOT.

### 3.3.10 Análisis bioinformático.

La identificación de los péptidos se hizo a través del programa MASCOT (<http://www.matrixscience.com>); se comprobó realizando un alineamiento comparativo mediante el programa Blast-P(<https://www.ncbi.nlm.nih.gov/blast/>) los

resultados de este alineamiento se usaron para generar una tabla de valores normalizados mediante el programa FASTA36 ajustando el punto de corte estadístico con e value menor a 0.01, con esto se generó una matriz de expresión diferencial.

Los resultados de la matriz de expresión diferencial se muestran en la siguiente tabla:

Número de Adhesión	Proteína Identificada	Proceso Biológico en que participa	Frecuencia de Ocurrencia	Función Molecular
O81905	proteína cinasa parecida a receptor serina-treonina	expansión celular y diferenciación	3.4	Crecimiento y diferenciación celular
Q40567	calretulina de tabaco	doblamiento de proteínas	3.2	Metabolismo proteico
P11605	Nitrato reductasa (NADH 1)	asimilación de nitrato	3.1	Metabolismo basal
O22467	proteína de unión a histonas MSI1	proliferación celular	3.1	Crecimiento y diferenciación celular
Q9XIV8	peroxidasa N1	producción y recaptura de especies reactivas de oxígeno	2.93	proteínas relacionadas al estrés
Q1W1Q6	Gliceraldehido-3-fosfato deshidrogenasa A	generación de NADPH	2.8	Metabolismo basal
Q9XEN8	Anexina	ciclo celular	2.8	Crecimiento y diferenciación celular
B9RMA5	Peptidil-prolil cis-trans isomerasa	metabolismo proteico	2.7	Metabolismo proteico
Q7DMA9	Peptidil-prolil-cis-trans isomerasa	división celular/control de citocinas y auxinas	2.6	Crecimiento y diferenciación celular
Q9FM96	Glucosidasa 2 subunidad beta	respuesta a estrés (percepción bacterial de elongacion TU) patrones moleculares asociados a patógenos (PAMPs)	2.3	proteínas relacionadas al estrés
Q9FPL3	fosforibosil-amino-imidazol-carboxamida-formil-transferasa/IMP ciclohidrolasa	biosíntesis de purinas	2.3	Metabolismo basal
Q9FXT9	subunidad regulatoria 7 de la proteasa 26 S	degradación de proteínas	2.2	Metabolismo proteico
B3F8G8	1-hidroxi-2-metil-2-(E)-butenil 4-difosfato sintasa	procesamiento biosintético de terpenoides	2.14	Metabolismo basal
O65719	70 kDa proteína de choque por calor	respuesta a estrés (choque por calor, frío, infección viral)	2.1	proteínas relacionadas al estrés
Q2QMV8	supuesta Adenina fosforibosil transferasa 1	procesamiento metabólico de nucleósidos	2.1	Metabolismo basal
P10896	activador de ribulosa 1,5 bifosfato carboxilasa oxigenasa	metabolismo central	2.01	Crecimiento y diferenciación celular
B8AZJ5	fosfoglicerato cinasa	metabolismo central	2	Metabolismo basal
Q9LW96	Inositol-3-fosfate sintasa	metabolismo de Polioles	1.9	Metabolismo basal

QBLNYS	ascorbato peroxidasa del estroma	Respuesta a estrés oxidativo	1.89	proteínas relacionadas al estrés
Q14TB1	proteína de choque térmico de 90 kDa	respuesta a estrés (virus)	1.87	proteínas relacionadas al estrés

**TABLA 3-1** Proteínas diferenciales detectadas entre las líneas transgénicas de tabaco L1 y L8. Fuente: Cárdenas-Manríquez *et al.*, 2016.

En la tabla 1 se muestran las 20 proteínas cuya expresión relativa fue la mayor entre las 58 proteínas diferenciales encontradas, en sombreado se muestran 5 con altos valores de expresión y que participan en funciones vinculadas a la respuesta a estrés. De hecho, 22 de las 58 proteínas diferenciales están vinculadas con respuesta a estrés. Esto implica que entre el 25 y 30 % de las proteínas diferenciales encontradas tiene actividad defensiva, lo que se corresponde con los resultados predichos en los análisis transcriptómicos.

### 3.4 Metodología Metabólica.

#### 3.4.1 Esterilización, germinación de semillas, colecta de muestras y preparación de extractos.

Las semillas de las líneas 1, 8, 25 y 26 de tabaco transgénico *N. tabacum cv. Xanthi* y la línea silvestre *xnc*, se esterilizaron por inmersión en solución de NaClO al 13% durante 15 minutos, tras lo cual se realizaron tres lavados con agua destilada estéril. Las semillas se colocaron en frascos de vidrio con 25 ml de medio MS adicionado con fitagel como gelificante, kanamicina como agente de selección y ácido ascórbico como antioxidante. Los frascos se mantuvieron a temperatura ambiente con un fotoperiodo aproximado de 16 h de luz y 8 h de oscuridad, durante 6 semanas hasta la obtención de las plántulas (Ruiz, 2009).

Se colectaron hojas de cada una de las líneas de tabaco a los 2, 4 y 6 meses, se congelaron y molieron inmersas en nitrógeno líquido. Muestras de tejido de cada una de las líneas de tabaco se liofilizaron durante 72 horas. Las muestras se almacenaron a -80°C hasta su análisis.

Muestras de tejido de cada línea de tabaco de 200 mg se homogenizaron en 2 mL de metanol, y se mantuvieron en agitación durante 24 h en oscuridad. Se

centrifugaron las muestras a 5000 rpm por 10 min a 4 °C, tras lo cual se separó la fase orgánica para su análisis. (Atanassova *et al.*, 2011).

#### 3.4.2 Cuantificación de fenoles totales

La cuantificación de fenoles totales se hizo por el método de Folin-Ciocalteu en una microplaca de 96 pozos de acuerdo al método propuesto por Attard (2013). A una alícuota de 40 µL de extracto metanólico de cada una de las líneas de tabaco se le adicionaron 460 µL de agua destilada, 250 µL del reactivo de Folin-Ciocalteu, y tras 5 min se agregaron 1250 µL de Na<sub>2</sub>CO<sub>3</sub> al 2 %. La mezcla se agitó en vórtex por 10 segundos y se incubó a temperatura ambiente por 60 min, tras lo cual se midió la absorbancia de la muestra a 765 nm en un espectrofotómetro multiskan para microplacas. La concentración de fenoles totales se determinó con base en una curva estándar realizada con ácido gálico, expresando los resultados como µg equivalentes de ácido gálico por gramo de tejido (µg EAG/g).

#### 3.4.3 Cuantificación de flavonoides.

La cuantificación de flavonoides totales se realizó con base en el método de aminoetil difenilborato propuesto por Ooman y col. (2005). Una alícuota 50 µL del extracto metanólico se mezcló con 180 µL de agua destilada y 20 µL de 2-aminoetil difenilborato al 1 % en una microplaca de 96 pozos. Se midió la absorbancia de la mezcla a 404 nm en un espectrómetro multiskan para microplacas. El contenido de flavonoides se obtuvo de acuerdo a una curva estándar de rutina, expresando los resultados como µg equivalentes de rutina por gramo de tejido (µg ER/g).

#### 3.4.4 Cuantificación de taninos condensados.

Para determinar la cantidad de taninos condensados se utilizó el ensayo de la vainillina descrito por Hagerman (2002) modificado. A una alícuota de 50 µL del extracto metanólico de cada muestra, se adicionaron 200 µL de una solución de vainillina 1 % con HCl 8 % en una proporción 1:1, en tantos de 50 µL. Se dejó reposar la mezcla a temperatura ambiente por 20 min, se midió la absorbancia de

cada reacción a 500 nm en un espectrófotometro. Se utilizó como línea base de absorbancia una solución de HCl al 4 %. El contenido de taninos se obtuvo por comparación con una curva estándar de catequina, siendo expresados los resultados como  $\mu\text{g}$  equivalentes de catequina por gramo de tejido ( $\mu\text{g EC/g}$ ).

#### 3.4.5 Cuantificación de antocianinas.

Para la determinación de antocianinas se empleó el método de pH diferencial descrito por Sutharut y Sudarat (2012). Se prepararon dos diluciones de la muestra a partir de una alícuota de 200  $\mu\text{L}$  de extracto metanólico y 2 mL de buffer. Para la primera dilución se utilizó buffer de cloruro de potasio (0.025 M, pH 1.0), y para la segunda buffer de acetato de sodio (0.4 M, pH 4.5). Se midió la absorbancia de ambas diluciones de la muestra a 520 y 700 nm tras 20 min. El contenido de antocianinas totales se determinó de acuerdo con las formulas:

$$A = (A_{520\text{nm}} - A_{700\text{nm}})_{\text{pH } 1.0} - (A_{520\text{nm}} - A_{700\text{nm}})_{\text{pH } 4.5} \quad \text{Ec. 1}$$

$$\text{Antocianinas (mg/L)} = A \cdot \text{MW} \cdot \text{DF} \cdot 1000 / \epsilon \cdot 1 \quad \text{Ec. 2}$$

Donde A es la absorbancia, MW es el peso molecular, DF es el factor de dilución, y  $\epsilon$  es la absorptividad molar. Se utilizó como antocianina de referencia la cianidina-3-glucosido (MW= 449.2 g Lmol<sup>-1</sup>,  $\epsilon=26\ 900$  Lmol<sup>-1</sup>cm<sup>-1</sup>).

El resultado de antocianinas presentes en la muestra se expresa como  $\mu\text{g}$  equivalentes de cianidina-3-glucosido ( $\mu\text{g ECG/g}$ ).

#### 3.4.6 Determinación de capacidad antioxidante por el ensayo de DPPH.

Para el ensayo de capacidad antioxidante se usó el método de DPPH (2,2-difenil-1-picrilhidrazil) propuesto por Ooman y col. (2005). A una alícuota de 20  $\mu\text{L}$  de extracto se adicionaron 200  $\mu\text{L}$  de una solución de DPPH 150  $\mu\text{M}$  en metanol al 80%. La placa se mantuvo en oscuridad a temperatura ambiente, y se midió la absorbancia a 520 nm a los 0, 5, 15, 30, 60, 75, 90 y 120 min posterior a la preparación de la reacción. Como referencia se realizó una curva estándar con

Trolox, expresando el resultado como  $\mu\text{M}$  equivalentes de Trolox por g de tejido ( $\mu\text{M ET/g}$ ). Para determinar la actividad anti-radical (ARA) se utilizó la ecuación:

$$\text{ARA} = (1 - A_{\text{muestra}} / A_{\text{control}}) \cdot 100 \quad \text{Ec. 3}$$

#### 3.4.7 Determinación de capacidad antioxidante por el ensayo de ABTS.

La capacidad antioxidante se evaluó simultáneamente mediante el ensayo de ABTS con modificaciones según el método de Gupta y col., 2009. Para el reactivo se prepara una solución stock de ABTS 7 mM y una solución de persulfato de potasio 2.6 mM, las cuales se mezclaron en las mismas proporciones. La solución se dejó en reposo durante 12 h en oscuridad a temperatura ambiente, tras lo cual se diluyó con metanol hasta obtener una absorbancia entre 0.8 y 1.0 unidades a 734 nm en el espectrofotómetro. Esta solución se usó como solución de trabajo.

Para el ensayo, una alícuota de 20  $\mu\text{L}$  del extracto se mezcló con 230  $\mu\text{L}$  de la solución ABTS en una microplaca y se midió su absorbancia a 734 nm a 0 y 6 min tras la preparación de la reacción. Como referencia se utiliza una curva estándar con Trolox, tal que el resultado se expresa como  $\mu\text{M}$  equivalentes de Trolox por g de tejido ( $\mu\text{M ET/g}$ ). La actividad antiradical (ARA) se determinó de acuerdo a la ecuación:

$$\text{ARA} = (1 - A_{\text{muestra}} / A_{\text{control}}) \cdot 100 \quad \text{Ec. 4}$$

#### 3.4.8 Análisis de compuestos fenólicos en HPLC.

Se realizó una extracción a partir de una muestra de 50 mg de tejido liofilizado en 10 mL de metanol por incubación en el sonicador por 40 min a temperatura ambiente. La muestra se centrifugó a 3000 G por 20 min a 4 °C; posteriormente se filtró con acrodiscos Whatman de 25 mm con tamaño de poro de 0.45  $\mu\text{m}$ .

Para la identificación y cuantificación de compuestos fenólicos se utilizó un sistema Waters 600 HPLC-DAD, equipado con muestreo automático Waters 717 plus y un detector UV/Vis Waters 2487 (Milford, MA, USA). La separación se realizó en una columna capilar de fase reversa Phenomenex ODS-C18 (250 x 4.6

mm, 5  $\mu$ m). La fase móvil utilizada consiste en dos solventes: (A) ácido fórmico-agua 1:99 v/v y (B) acetonitrilo; a una proporción A/B inicial de 98/2, 68/32 a los 30 min, 45/55 a los 48 min, 5/95 a 53 min, 98/2 a 57 min. El tiempo total de corrida fué de 67 min. Las mediciones de absorbancia se realizaron a 260-320 nm. Se inyectó una muestra de 10  $\mu$ L a un flujo constante de 1  $\mu$ L/min.

Los compuestos se determinaron contra curvas estándares preparadas previamente. Los compuestos utilizados como estándares fueron en ácido clorogénico, epicatequina, hesperidina, ácido cumárico, ácido cafeico, epigallocatequín galato, rutina, ácido sinápico, ácido elágico, vainillina, ácido rosmarínico, quercetina, y resveratrol.

#### 3.4.9 Análisis metabolómico basado en GC-MS.

Se realizó una extracción de metabolitos para el perfil a partir de 100 mg de tejido de cada línea en 10 mL de metanol mediante incubación en equipo sonicador durante 40 min a temperatura ambiente. Se centrifuga a 3000 G por 20 min a 4 °C, y se evaporar el solvente de la fase metanólica colectada bajo corriente de gas nitrógeno.

El análisis se realizó en un cromatógrafo de gases Agilent 7890<sup>a</sup> acoplado a un detector de masas cuadrupolo Agilent 5975C con una columna capilar HP-5MS (30 m x 0.25 mm, 0.25 $\mu$ m). El análisis se realizó en modo sin fraccionamiento a 250 °C. La programación de temperatura del sistema fué: temperatura inicial de 100 °C incrementando a 310 °C en rampa de 40 °C/min, y sosteniendo constante durante 7.5 min. Se usó helio como gas acarreador a un flujo de 1 mL/min.

Los metabolitos fueron identificados de acuerdo al tiempo de retención del ion objetivo y la razón cualificada del ion para cada compuesto con apoyo del software ChemStation. Solo se consideraron compuestos con una probabilidad de identidad superior a un 50 %.

#### 3.4.10 Análisis estadístico.

Los datos obtenidos de la concentración de compuestos bioactivos se sometieron a un análisis de varianza (ANOVA) simple para determinar la diferencia entre tratamientos. Así mismo se realizó una prueba de diferencia mínima significativa de Fisher (DSM) para determinar la diferencia entre las medias de cada línea de tabaco para cada tipo de compuestos.

### 3.4.11 Análisis bioinformático.

Se usará KEGG ([www.genome.jp/](http://www.genome.jp/)) para análisis de rutas metabólicas y pasos clave de cada una de ellas. Para el análisis de datos del perfil metabólico se utilizó el paquete estadístico “gplots” en el software R, para la elaboración de heatmaps. Así mismo, para analizar el panorama de vías metabólicas relacionadas a los metabolitos identificados y su impacto se recurrirá a la herramienta de libre acceso MetaboAnalyst 2.0 y KEEG en la web.

Los resultados de la integración metabólica son:

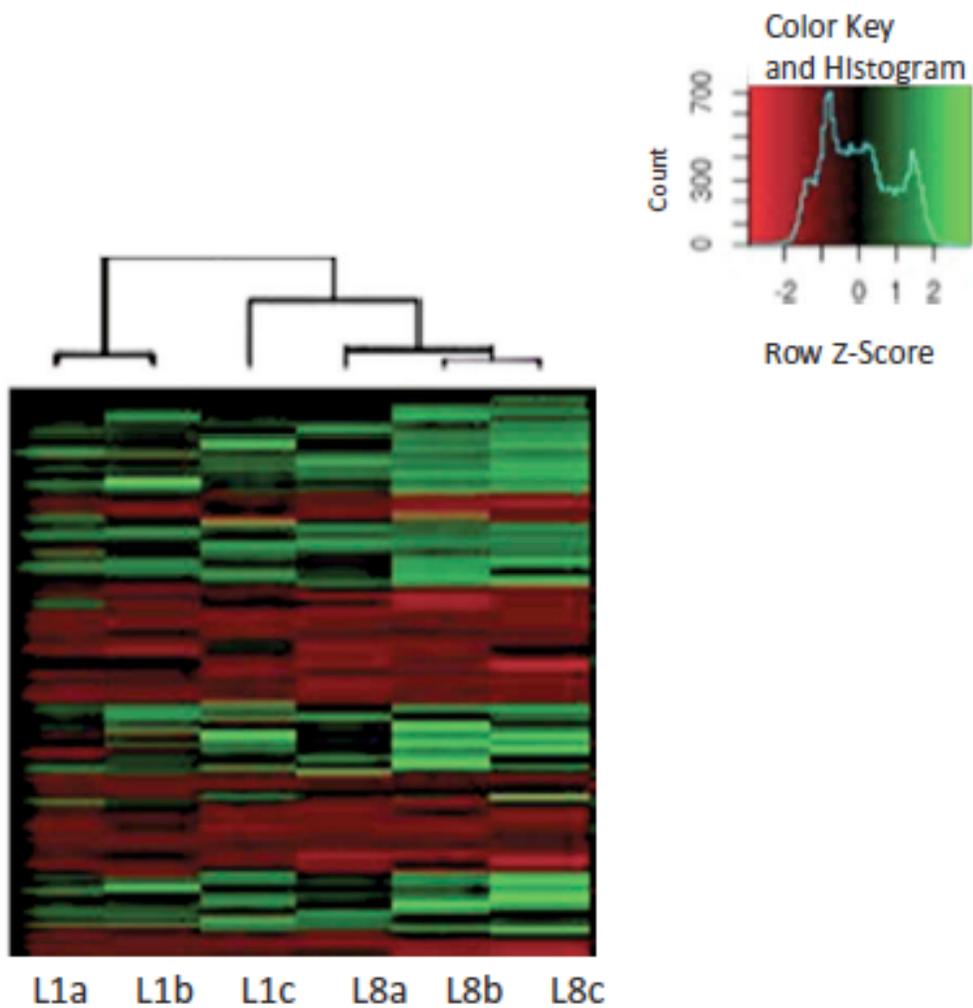
Vía metabólica modificada	No. de secuencias	No. de enzimas	Vía metabólica modificada	No. de secuencias	No. de enzimas
metabolismo de purinas	9	6	metabolismo de amino-azúcares y nucleótido-azúcares	2	2
metabolismo de azufre	9	5	metabolismo de metano	2	1
biosíntesis de antibióticos	8	5	metabolismo de glutatión	2	1
fosforilación oxidativa	6	4	metabolismo de fenilalanina	2	2
metabolismo de glioxilato y dicarboxilato	5	4	biosíntesis de pantotenato y CoA	1	1
metabolismo de ácido alfa linolénico	5	2	metabolismo de seleno-compuestos	1	1
metabolismo de almidón y sacarosa	5	4	glucolisis y gluconeogénesis	1	1
metabolismo de fármacos por citocromo p450	4	3	síntesis de cutina suberina y ceras	1	1
biosíntesis de fenilpropanoides	4	1	metabolismo de alanina aspartato y glutamato	1	1
metabolismo de beta-alanina	3	3	fijación de carbono por organismos fotosintéticos	1	1
metabolismo de tiamina	3	1	metabolismo de propanoato	1	1
biosíntesis de alcaloides de isoquinolina	3	2	biosíntesis de alcaloides tropanos piperidínicos y piridínicos	1	1
metabolismo de cisteína y metionina	3	1	metabolismo de glicerolípidos	1	1
metabolismo de xenobióticos por citocromo p450	3	2	metabolismo de nicotinato y nicotinamida	1	2
metabolismo de ácido cianoamino	3	2	metabolismo de histidina	1	1
metabolismo de pirimidina	3	1	interconversiones pentosa-glucuronato	1	1
metabolismo de glicina serina y treonina	3	2	biosíntesis de monobactam	1	1

metabolismo de tirosina	2	2	vía de señalización del receptor de la célula T	1	1
reserva de un carbono por folato	2	1	degradación de aminobenzoato	1	1
metabolismo de ácido linoléico	2	1	biosíntesis de isoflavonoides	1	1
metabolismo de porfirinas y clorofila	2	2	biosíntesis de isoterpeno	1	1

**TABLA 3-2. Lista de rutas metabólicas modificadas por transcritos y proteínas reprimidas**

### 3.5 Integración y análisis de las redes de comunicación molecular.

Una de las figuras integrativas más usadas actualmente en análisis ómicos son los heatmaps. Ya que permiten comparar niveles de expresión en un número diferente de muestras.



**Figura 3-17**

La figura 3-17 muestra el heatmap correspondiente al comparativo de genes en que aumentó/disminuyó la expresión de los transcritos comparativamente entre la línea control L1 y la línea de alta expresión del CchGLP L8.

Las letras a, b, c, representan un replicado diferente que se sometió a comparación.

Mediante paquetes como Mixomics (exclusivo para R), pheNetic, IIS, SAMnetweb, se pueden hacer análisis probabilístico de grupos y subgrupos de los resultados de una secuenciación masiva de datos, sin embargo, para realizar corridas de

prueba se requieren 32 Gb en ram, comunicación mayor a 100mbps y una suscripción a un grupo de intercambio de información que solicita un apoyo mínimo de \$100 U.S.D pagado a través de tarjeta.

Esta será la parte más relevante del estudio en cuanto pueda realizarse.

#### 3.5.1 Gráfico de red de comunicación molecular.

Será la representación central de la intercomunicación entre los 3 niveles ómicos, un gráfico de red sería la mejor representación para este caso en específico.

### 4. Conclusiones.

Se puede concluir que el incremento sostenido en los niveles de peróxido de hidrógeno ocasionado por la expresión constitutiva del transgen CchGLP comparado en la línea L8 vs la línea L1 **aumenta la expresión de genes relacionados a respuestas favorables ante condiciones de cultivo adversas:**

Malas condiciones térmicas (heat shock protein hsf30 isoforma 2x, cognate 70KDa, 83-like);

Sequía (suberisation-associated anionic peroxidase-like protein, dehydration-responsive protein rd22, lignin-forming anionic peroxidase like);

Regulación y destino metabólico ante el ambiente (Bag-family molecular chaperone regulator 6, auxin-binding protein abp19a-like, high-affinity nitrate transporter-like protein, copper transporter protein Cch isoform x2, Em protein h5);

Defensa ante stress biótico y abiótico directo (peroxidase 15-like, stress response protein nst1-like, major allergen pru ar 1-like, snaking-2, endochitinase mcht-2, defensin-like protein 1, abscisic acid and environmental stress-inducible protein tas 14-like, kirola-like).

Mientras que el aumento del nivel de peróxido de hidrógeno comparado entre plantas de la línea L8 y L1 **disminuye la expresión de transcritos no-favorables ante condiciones de cultivo adversas:**

Respuesta osmótica (transcription factor pur-alpha 1 isoform x2, caspase like protein pimp 1, cellulose synthase-like protein d5)

Regulación y destino metabólico (possible WRKY transcription factor 26 isoform 1x, late embryogenesis abundant protein lea5-like, calmodulin-binding-like protein, tumor-related protein)

inhibidores de respuestas a estrés (T-ascorbate peroxidase peroxisomal-like protein, cellulose synthase-like protein d5)

seguramente los transcritos que aumentan / disminuyen su expresión se relacionan con disminución del metabolismo basal, aumento del metabolismo secundario específico a respuestas defensivas, modificaciones a mediano y largo plazo del estado transcripcional y posiblemente cambio del estado epigenético de las semillas generadas durante esta fase.

**Esto demuestra que el aumento de peróxido de hidrógeno endógeno puede ser usado como una herramienta de cambio ómico estable en cultivos que enfrentan condiciones adversas de manera exitosa.**

## **5. Literatura citada.**

Agrios G., 2005. Plant pathology. 5th ed. Elsevier Academic Press. ISBN 0-12-044565-4. 30 corporate drive, suite 400, Burlington, MA 01803, USA 525 B street, suite 1900, San Diego, California 92101-4495, USA 84 Teobald's Road, London WC1X8RR, UK.

Almario J. Muller D. Défago G. Moënne-Loccoz Y (2014). Rhizosphere ecology and phytoprotection in soils naturally suppressive to Thielaviopsis black root rot of tobacco. Environmental Microbiology, doi:10.1111/1462-2920.12459.

Anaya-López J. L. Torres-Pacheco I. González-Chavira M. Garzón-Tiznado J. A. Pons-Hernández J. L. (2003) resistance to geminivirus mixed infections in mexican wild peppers. *Hortscience* 38(2):251-255

Anderson J. P. Gleason C. A. Foley R. C. Thrall P. H. Burdon J. B. Singh K. B. (2010). Plants versus pathogens: an evolutionary arms race. *Funct Plant Biol.* 2010 May 20; 37(6):499–512. doi:10.1071/FP09304.

Barrera-Pacheco A. Joaquín-Ramos A. J. Torres-Pacheco I. González-Chavira M. M. Pérez-Pérez M. C. I. Guevara-Olvera L. Guevara-González R. G. (2008) Analysis of transcriptional expression induced in *Capsicum chinense* BG-3821 under conditions of biotic and abiotic stress. *Agroscience* 42: 95-106.

Aureliano Bombarely, Hen G. Rosli, Julia Vrebalov, Peter Moffett, Lukas A. Mueller, and Gregory B. Martin. (2012). A Draft Genome Sequence of *Nicotiana benthamiana* to Enhance Molecular Plant-Microbe Biology Research. *MPMI* Vol. 25, No. 12, pp. 1523–1530. <http://dx.doi.org/10.1094/MPMI-06-12-0148-TA>

Buonaurio R. Iriti M. Romano G. (2009) induced resistance to plant diseases caused by oomycetes and fungi. *Petria*, 19(3): 130-148.

Kieron D Edwards, Aureliano Bombarely, Geraint W Story, Fraser Allen, Lukas A Mueller, Steve A Coates, Louise Jones. (2010). TobEA: an atlas of tobacco gene expression from seed to senescence. *BMC Genomics* 2010, 11:142 <http://www.biomedcentral.com/1471-2164/11/142>

Foley R. C. Cynthia A. Gleason C. A. Anderson J. P. Hamann T. Singh K. B. (2013). Genetic and Genomic Analysis of *Rhizoctonia solani* Interactions with *Arabidopsis*; Evidence of Resistance Mediated through NADPH Oxidases. *PLoS ONE* 8(2): e56814. doi:10.1371/journal.pone.0056814

García-Gutiérrez, Cipriano; Rodríguez-Meza, Guadalupe Durga. (2012). Problemática y riesgo ambiental por el uso de plaguicidas en sinaloa. *Ra Ximhai*, vol. 8, núm. 3b, pp. 1-10 <http://www.redalyc.org/articulo.oa?id=46125177005>

Lina García-Mier, Ramón G. Guevara-González, Víctor M. Mondragón-Olguín, Beatriz del Rocío Verduzco-Cuellar and Irineo Torres-Pacheco. (2013). Agriculture and Bioactives: Achieving Both Crop Yield and Phytochemicals. *Int. J. Mol. Sci.* 14, 4203-4222; doi:10.3390/ijms14024203 ISSN 1422-0067

Vladimir Gligorijevic, Vuk Janjic and Natasa Przulj. (2014). Integration of molecular network data reconstructs Gene Ontology. *BIOINFORMATICS* Vol. 30 ECCB, pages i594–i600 doi:10.1093/bioinformatics/btu470

Goodin M. M. Zaitlin D. Naidu R. A. Lommel S. A. (2008) *Nicotiana benthamiana*: Its History and Future as a Model for Plant–Pathogen Interactions. *Molecular Plant–Microbe Interactions* Vol. 21, No. 8, 2008, pp. 1015–1026. doi:10.1094/MPMI -21-8-1015.

Guevara-Olvera L. Ruíz-Nito M. L. Rangel-Cano R. M. Torres-Pacheco I. Rivera-Bustamante R. F. Muñoz-Sánchez C. I. González-Chavira M. M. Cruz-HeARNdez A. Guevara-González R. G. (2012) Expression of a germin-like protein gene (CchGLP) from a geminivirus-resistant pepper (*Capsicum chinense* Jacq.) enhances tolerance to geminivirus infection in transgenic tobacco. *Physiological and Molecular Plant Pathology* 78:45-50.

Jones D.G. y Dangl J. L. (2006). The plant immune system. *Nature*, 444: 323-329.

Junping Gao, Genhong Wang, Sanyuan Ma, Xiaodong Xie, Xiangwei Wu, Xingtang Zhang, Yuqian Wu, Ping Zhao, Qingyou Xia. (2014). CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol Biol*. DOI 10.1007/s11103-014-0263-0

Kim S. W. Jung J. H. Lamsal K. Kim Y. S. Min J. S. Lee Y. S (2012). Antifungal Effects of Silver Nanoparticles (AgNPs) against Various plant Pathogenic Fungi. *Mycobiology* 40(1):53-58.

León-Galván F. Joaquín-Ramos A. J. Torres-Pacheco I. Barba de la Rosa A. P. Guevara-Olvera L. González-Chavira M. M. Ocampo-Velazquez R. V. Rico-García

E. Guevara-González R. G. (2012). A Germin-Like Protein Gene (*CchGLP*) of *Capsicum chinense* Jacq. Is Induced during Incompatible Interactions and Displays Mn-Superoxide Dismutase Activity. *InteARNtional jouARNI of molecular sciences*. 12, 7301-7313; doi:10.3390/ijms12117301.

Martínez-Valenzuela C. Gómez-Arroyo S. (2007). Riesgo genotóxico por exposición a plaguicidas en trabajadores agrícolas. *Rev. Int. Contam. Ambient.* 23 (4) 185-200.

Mejía-Teniente L. Torres-Pacheco I. González-Chavira M. M. Ocampo-Velázquez R. V. Herrera-Ruíz G. Chapa-Oliver A. M. Guevara-González R. G. (2010). Use of elicitors as an approach for sustainable agricultura. *African jouARNI of biotechnology* Vol. 9 (54) pp. 9155-9162.

Laura Mejía-Teniente, Flor de Dalia Durán-Flores, Angela María Chapa-Oliver, Irineo Torres-Pacheco, Andrés Cruz-Hernández, Mario M. González-Chavira, Rosalía V. Ocampo-Velázquez and Ramón G. Guevara-González. (2013). Oxidative and Molecular Responses in *Capsicum annum* L. after Hydrogen Peroxide, Salicylic Acid and Chitosan Foliar Applications. *Int. J. Mol. Sci.* 14, 10178-10196; doi:10.3390/ijms140510178 ISSN 1422-0067

Muthamilarasan M. and Prasad M . (2013) Plant innate immunity: An updated insight into defense mechanism. *J. Biosci.* 38 433–449] DOI 10.1007/s12038-013-9302-2

Nana Jin, Deng Wu, Yonghui Gong, Xiaoman Bi, Hong Jiang, Kongning Li, and Qianghu Wang. (2014). Integration Strategy Is a Key Step in Network-Based Analysis and Dramatically Affects Network Topological Properties and Inferring Outcomes. *Biomed research inteARNtional*. Article ID 296349, 13 pgs. <http://dx.doi.org/10.1155/2014/296349>

Peter Palaukaitis, Simon C. Groen and John P. Carr. (2013). The Rumsfeld paradox: some of the things we know that we don't know about plant virus

infection. *Current Opinion in Plant Biology* 2013, 16:513–519.  
<http://dx.doi.org/10.1016/j.pbi.2013.06.004>

Romina Pedreschi, Susan Lurie, Maarten Hertog, Bart Nicolai, Jurriaan Mes and Ernst Woltering. (2013). Post-harvest proteomics and food security. *Proteomics* 13, 1772–1783 DOI 10.1002/pmic.201200387

Ruíz-Nito María de la Luz (2009). Expresión transgénica de CchGLP (Capsicum chinense CchGLP) en *Nicotiana tabacum xanthi* nc para evaluar la resistencia a geminivirus. Tesis presentada para obtener el título de maestro en ciencias en ingeniería bioquímica. Instituto tecnológico de Celaya, Guanajuato.

Kai Sun, Joana P Goncalves, Chris Larminie and Nataša Pržulj. (2014). Predicting disease associations via biological network analysis. *BMC Bioinformatics*, 15:304.  
<http://www.biomedcentral.com/1471-2105/15/304>

Listado de revistas utilizadas

Revista	Impacto	No. de veces utilizada
African jouARNI of biotechnology		6
Agroscience	0,374	1
BIOINFORMATICS	5,486	6
Biomed research inteARNtional		7
BMC BIOINFORMATICS	2,751	1
BMC Genomics	4,073	3
Current Opinion in Plant Biology	7,711	1
Environmental Microbiology	5,843	1

Functional Plant Biology	2,929	2
Hortscience	0,477	1
InteARNtional jouARNI of molecular sciences	2,598	1
InteARNtional jouARNI of molecular sciences	2,598	2
InteARNtional jouARNI of molecular sciences	2,598	4
JouARNI of Biosciences	1,648	2
Molecular Plant-Microbe Interactions	4,431	2
Molecular Plant-Microbe Interactions	4,431	2
Mycobiology	2,031	1
Nature	36,28	3
Petria		1
Physiological and Molecular Plant Pathology	1,377	11
Plant Molecular Biology	4,15	2
PLoS ONE	4,092	1
Proteomics	4,505	2
Ra Ximhai		1
Internacional de Contaminación Ambiental	0,36	1

- Los años de publicación de las revistas consultadas abarcan del 2008 a la fecha. Se consultaron publicaciones anteriores cuando la relevancia de la información lo indicaba.

## 5.6 ANEXOS.

### **Anexo 1. Soluciones para la extracción fenólica de proteínas.**

1. Buffer extracción: 500 mM de Tris-HCl, 50 mM EDTA, 700 mM sucrosa, 100 mM KCl y ajustar pH 8 con HCl.

Antes de la extracción agregar  $\beta$ -mercaptoetanol al 2% y 1 mM de fenilmetilsulfonil fluorido (PMSF)

2. Fenol: Tris-HCl saturado, pH6.6/7.9

3. Solución de precipitación: 0.1mM de acetato de amonio en metanol frío. Se almacena a -20°C.

4. Buffer de enfoque isoeléctrico: 9M urea, 4% CHAPS, 0.5% Triton X-100, 20 mM DTT, 1.2% Pharmalytes pH 3 a 10.

### **Anexo 2. Soluciones para la extracción de proteínas por ácido tricloro acético.**

1. Buffer de extracción: tris 40 mM, EDTA 10 mM, PVPP 1%.

2. Acetona

### **Anexo 3. Soluciones para el equilibrio de las tiras de IPG**

1. Buffer de equilibrio I: Urea 6M, SDS 2%, tris-HCl 0.05M pH8.8, glicerol 20%, DTT 2%.

2. Buffer de equilibrio II: Urea 6M, SDS 2%, tris-HCl 0.05M pH8.8, glicerol 20%, iodoacetamida 2.5%.