### Identificación de genes y marcadores moleculares asociados con la tolerancia a la podredumbre carbonosa causada por *Macrophomina phaseolina* en soja

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires, Área Ciencias Agropecuarias

> **Mariano Hernán Cassina** Ingeniero agrónomo, Universidad de Buenos Aires – 2017

Lugar de trabajo: Facultad de Agronomía, Universidad de Buenos Aires





FAUBA Escuela para Graduados Ing. Agr. Alberto Soriano Facultad de Agronomía – Universidad de Buenos Aires

### **COMITÉ CONSEJERO**

Director de tesis **Eduardo A. Pagano** Ingeniero Agrónomo (Universidad de Buenos Aires) Doctor en Ciencias Biológicas (Universidad de Granada, España)

Co-director de Tesis Javier F. Botto Licenciado en Ciencias Biológicas (Facultad de Ciencias Exactas y Naturales, UBA) Doctor en Biología (UBA)

### JURADO DE TESIS

### JURADO Verónica Lía Doctora de la Universidad de Buenos Aires

### JURADO Sebastián Stenglein Doctor en Ciencias Agropecuarias

### JURADO Mariano Pardo Doctor en Ciencias Biológicas

Fecha de Defensa Tesis: 17 de septiembre de 2024.

## Agradecimientos

Quisiera agradecer a mi director, el Dr. Eduardo A. Pagano, por guiarme a lo largo de todo el trabajo de esta tesis y aconsejarme desde el primer momento. Estoy agradecido con mis compañeros de trabajo, mis amigos y familia, por acompañarme siempre y hacer más fácil el camino.

### Declaración

"Declaro que el material incluido en esta tesis es, a mi mejor saber y entender, original, producto de mi propio trabajo (salvo en la medida en que se identifique explícitamente las contribuciones de otros), y que este material no ha sido presentado, en forma parcial o total, como una tesis en esta u otra institución".

Ing. Agr. Mariano H. Cassina

# Publicaciones derivadas de la tesis

# ÍNDICE GENERAL

Capítulo 1. Introducción general	1
1.1 Estructura de la tesis	1
1.2 Revisión de antecedentes	1
1.2.1 Cultivo de soja en Argentina y sus enfermedades	1
1.2.2 Podredumbre carbonosa causada por M. phaseolina	4
1.2.3 Sistema suelo-planta-atmosfera: factores relacionados a la enfermedad	6
1.2.4 Interacción planta – patógeno y defensa vegetal	9
1.2.5 Genes involucrados	10
1.2.6 Manejo integrado de enfermedades (MIE)	15
1.2.7 Transcriptómica	16
1.3 Objetivo principal	17
1.4 Objetivos específicos	18
1.5 Hipótesis	18
Canítulo 2. Caracterización de conotines con respueste diferencial el etecu	10 dol
natógeno	20
2 1 Introducción	20
2.2 Materiales y métodos	21
2.2.1 Selección de inóculo	21
2.2.1 Ensavo para evaluación de la enfermedad (invernáculo)	26
2 3 Parámetros medidos	30
2.3.1 Ensavo para evaluación de la enfermedad (invernáculo)	31
2 3 1 a) Emergencia de plántulas	31
2.3.1 h) Severidad de raíz y base del tallo de plántulas	32
2 3 1 c) Tasa de infección de la enfermedad	33
2 3 1 d) Área bajo de la curva del progreso de la enfermedad (ABCPE)	34
2 3 1 e) Incidencia en fragmentos de raíz	36
2.3.2 Ensavos complementarios	36
2.3.2 Ensayos comprementarios	36
2.3.2 d) Daño foliar (lesión en cotiledón, hoja cotiledonar y trifoliada)	37
2 3 2 c) Establecimiento de plántulas (longitud y peso fresco de plántula)	37
2.4 Resultados	39
2.4.1 Ensavo para evaluación de la enfermedad (invernáculo)	
2.4.1.a) Emergencia de plántulas	39
2 4 1 b) Severidad de raíz y base del tallo de plántulas	40
2.4.1 c) Tasa de infección de la enfermedad	42
2 4 1 d) Área bajo de la curva del progreso de la enfermedad (ABCPE)	43
2.4.1.e) Incidencia en fragmentos de raíz	45
2.4.2 Ensavos complementarios	48
2 4 2 a) Severidad en semillas	48
242 b) Daño foliar (lesión en cotiledón, hoia cotiledonar y trifoliada)	49
2.4.2 c) Establecimiento de plántulas (longitud y neso fresco de plántula)	50
2.5 Discusión	
2.6 Conclusiones	

Capítulo 3. Análisis transcriptómico de la respuesta de la soja a la podredu	ımbre
carbonosa de un genotipo tolerante	59
3.1 Introducción	59
3.2 Materiales y métodos	61
3.2.1 Ensayo principal in vivo	61
3.2.2 RNA-seq	64
3.3 Resultados	73
3.3.1 Análisis de expresión diferencial entre genotipos contrastantes	73
3.3.2 Análisis de enriquecimiento (GO/KEGG)	75
3.3.3 Selección de genes por tolerancia a la enfermedad	82
3.3.4 Selección de genes candidatos para desarrollar marcadores moleculares	96
3.4 Discusión	97
3.5 Conclusiones	107
Capítulo 4. Identificación de polimorfismos en genes de interés para el desa de marcadores moleculares de tolerancia a la enfermedad	rrollo 110
4.1 Introducción	110
4.2 Materiales y métodos	112
4.3 Resultados	114
4.3.1 Variantes en la secuencia transcriptómica	114
4.4 Discusión	123
4.5 Conclusiones	125
Capítulo 5. Conclusiones generales	127
Bibliografía	
	131

# ÍNDICE DE FIGURAS

Figura 1.1: Vista externa de una planta de soja2
Figura 1.2: Superficie sembrada en verde, representativa de las zonas donde se cultiva soja actualmente en la República Argentina
Figura 1.3: Lote de plantas de soja con podredumbre carbonosa en estadios avanzados del ciclo del cultivo
Figura 1.4: Ciclo epidemiológico de la enfermedad podredumbre carbonosa
Figura 1.5: Área bajo la curva de progreso de la podredumbre carbonosa (PC) de la soja según (a- las temperaturas máximas / b- precipitaciones) en Salta durante la campaña 2006/7
Figura 1.6: (a): Patogenicidad de <i>M. phaseolina</i> en dos cultivares de poroto bajo diferentes concentraciones de NaCl in vitro. Barras indican ± error estándar. (b) Diámetro de colonias (mm) de cuatro aislamientos de <i>M. phaseolina</i> en ágar papa glucosado (APG) ajustado a diferentes concentraciones de NaCl. El diámetro de la colonia se registró después de 72 h de incubación. Líneas verticales indican desvío estándar
Figura1.7: Toxinas actuantes en la PC (podredumbre carbonosa)
Figura 1.8: Modelo de las interacciones entre la sequía y podredumbre carbonosa. Las primeras etapas de estrés por sequía conducen a la producción de especies reactivas de oxígeno (ROS), que implican la reacción de hipersensibilidad y la muerte celular (Laloi et al. 2004); (Guan et al. 2000). La muerte celular del hospedante aumenta el crecimiento del patógeno. Fuente: elaboración propia a partir de los trabajos de (Mauch-Mani y Mauch. 2005), (Pandey, Ramegowda y Senthil-Kumar. 2015) y (Radwan et al. 2014)
Figura 1.9: Porcentaje (%) de plantas enfermas contabilizadas en los diferentes días post inoculación (DPI) de <i>M. phaseolina</i> . Líneas verticales indican desvío estándar de las mediciones
Figura 1.10 : Esquema simplificado de las principales aproximaciones omicas utilizadas en biotecnología17
Figura 2.1: Imágenes indicativas de cada uno de los estadios de semillas descriptos en la escala de Manici et al. 1995, desde el estadio 0 hasta el 5

Figura 2.6: Valores registrados de la escala de Abawi y Pastor Corrales.1990, desde el de menor severidad (1) hasta el valor máximo registrado (7)......33

Figura 2.9: Severidad registrada según la escala Abawi y Pastor Corrales.1990, a los 0, 7, 14 y 21 días desde V1, momento en el cual se inició el estrés hídrico. Las letras indican diferencias significativas según prueba no paramétrica de Kruskal-Wallis. H y S (+) indican la presencia de inóculo y/o sequía. .....40

Figura 2.11: Área bajo la curva del progreso de la enfermedad (ABCPE). La para cada uno de los tratamientos llevados a cabo. Las letras indican diferencias significativas según test Duncan (p<0,05), acorde a lo descripto en (Giovanardi et al. 2018).

Figura 2.14: Severidad según la escala Manici (1995) en semillas. Las letras indican diferencias significativas según test Tukey (valor p<0,05). Las imágenes corresponden a un ejemplo de semilla en placa, y a los diferentes estadios de la escala (0 a 5)......48

Figura 2.15: Vista de las placas incubadas a modo de ejemplo (cotiledón, hoja cotiledonar y folíolo central de la primera hoja trifoliada)......49

Figura 2.20: Box-plots de peso fresco (g) para ambos genotipos tanto en control (c) como con presencia del patógeno (Mp), según test de Tukey (valor p <0,05)......53

Figura 3.1 (Genotipo DM4800) y 3.2 (Genotipo D86: Macetas (repeticiones) representativas de tratamientos aplicados. De izquierda a derecha: H+ S+, H- S+, H+ S-, H- S- (control) en el momento previo a la cosecha (21 días desde V1)......62

Figura 3.8: Esquema de trabajo para análisis bioinformático.	Figura 3.7: Pasos de la secuenciación Illumina67
Figura 3.9: Volcanoplot de la expresión diferencial de genes del genotipo tolerante DT         con respecto al genotipo susceptible DM, en condiciones de infección y sequía simultánea	Figura 3.8: Esquema de trabajo para análisis bioinformático68
Figura 3.10: Categorías de la plataforma GO asociados al conjunto de genes sub- expresados diferencialmente en la comparación DT_HS vs DM_HS.	Figura 3.9: Volcanoplot de la expresión diferencial de genes del genotipo tolerante DT con respecto al genotipo susceptible DM, en condiciones de infección y sequía simultánea
Figura 3.11: Categorías de la plataforma GO asociados al conjunto de genes sobre- expresados diferencialmente en la comparación DT_HS vs DM_HS	Figura 3.10: Categorías de la plataforma GO asociados al conjunto de genes sub- expresados diferencialmente en la comparación DT_HS vs DM_HS.
Figura 3.12: Categorías de la plataforma KEGG asociados al conjunto de genes sub- expresados diferencialmente en la comparación DT_HS vs DM_HS	Figura 3.11: Categorías de la plataforma GO asociados al conjunto de genes sobre- expresados diferencialmente en la comparación DT_HS vs DM_HS
Figura 3.13: Categorías de la plataforma KEGG asociados al conjunto de genes sobre- expresados diferencialmente en la comparación DT_HS vs DM_HS	Figura 3.12: Categorías de la plataforma KEGG asociados al conjunto de genes sub- expresados diferencialmente en la comparación DT_HS vs DM_HS80
Figura 3.14 (a): Esquema resumen del criterio adoptado para la selección de genes de interés que contribuyen a la tolerancia de la enfermedad	Figura 3.13: Categorías de la plataforma KEGG asociados al conjunto de genes sobre- expresados diferencialmente en la comparación DT_HS vs DM_HS81
Figura 3.14 (b): Esquema resumen desde la realización del estudio RNA-seq hasta la identificación de genes asociados a la tolerancia a la enfermedad en el genotipo DT	Figura 3.14 (a): Esquema resumen del criterio adoptado para la selección de genes de interés que contribuyen a la tolerancia de la enfermedad
Figura 3.15: Gráfico de puntos de genes diferencialmente expresados resultantes del enriquecimiento en plataforma GO	Figura 3.14 (b): Esquema resumen desde la realización del estudio RNA-seq hasta la identificación de genes asociados a la tolerancia a la enfermedad en el genotipo DT
Figura 3.16: Gráfico de puntos de genes diferencialmente expresados resultantes del enriquecimiento en plataforma	Figura 3.15: Gráfico de puntos de genes diferencialmente expresados resultantes del enriquecimiento en plataforma GO92
<ul> <li>Figura 3.17: Mapa metabólico correspondiente a las vías de interacción planta – patógeno. Los recuadros en rojo agrupan los genes enriquecidos (sobre-expresados) en el genotipo tolerante en comparación al susceptible</li></ul>	Figura 3.16: Gráfico de puntos de genes diferencialmente expresados resultantes del enriquecimiento en plataforma
Figura 3.18: Mapa metabólico correspondiente a las vías de procesamiento de proteínas en retículo endoplasmático. Los recuadros en rojo agrupan los genes enriquecidos (sobre-expresados) en el genotipo tolerante en comparación al susceptible	Figura 3.17: Mapa metabólico correspondiente a las vías de interacción planta – patógeno. Los recuadros en rojo agrupan los genes enriquecidos (sobre-expresados) en el genotipo tolerante en comparación al susceptible100
Figura 4.1: Proceso secuencial de filtros aplicados para la obtención de un panel de marcadores moleculares frente a la podredumbre carbonosa en soja115 Figura 4.2: Información del impacto y función predichos por SnpEff para las variantes encontradas en la secuencia transcriptómica	Figura 3.18: Mapa metabólico correspondiente a las vías de procesamiento de proteínas en retículo endoplasmático. Los recuadros en rojo agrupan los genes enriquecidos (sobre-expresados) en el genotipo tolerante en comparación al susceptible
Figura 4.2: Información del impacto y función predichos por SnpEff para las variantes encontradas en la secuencia transcriptómica	Figura 4.1: Proceso secuencial de filtros aplicados para la obtención de un panel de marcadores moleculares frente a la podredumbre carbonosa en soja115
Figura 4.3: Hoja de ruta con los criterios establecidos para el filtrado de variantes. El número indica la cantidad de variantes resultantes luego de el filtro especificado en los recuadros principales	Figura 4.2: Información del impacto y función predichos por SnpEff para las variantes encontradas en la secuencia transcriptómica
	Figura 4.3: Hoja de ruta con los criterios establecidos para el filtrado de variantes. El número indica la cantidad de variantes resultantes luego de el filtro especificado en los recuadros principales

# ÍNDICE DE CUADROS

Cuadro 2.1: Inóculos puestos a prueba, su origen geográfico, hospedante y año de recolección (campaña agrícola)23
Cuadro 2.2: Descripción de la escala propuesta por Abawi y Pastor Corrales. 1990
Cuadro 2.3: Cuadro resumen de los tratamientos llevados a cabo
Cuadro 2.4: Severidades promedio registradas a los 7,14 y 21 días desde V1 para todos los genotipos, tanto en capacidad de campo (S-) como en sequía (S+)41
Cuadro 2.5: Valores de R <sup>2</sup> y valor p, para cada uno de los modelos epidemiológicos analizados (Exponencial, monomolecular, Gompertz y logístico), para cada uno de los tratamientos llevados a cabo42
Cuadro 2.6: Tasa de infección (severidad/día) correspondiente a cada uno de los tratamientos llevados a cabo43
Cuadro 2.7: Área bajo la curva del progreso de la enfermedad (ABCPE) calculados a partir de los diferentes momentos de medición. ABCPE (0-1) denota el avance de la enfermedad de V1 hasta la primera medición (7d), ABCPE (1-2) desde la primera medición (7d) hasta la segunda medición (14d) y ABCPE (2-3) desde la segunda medición (14d) hasta la tercera medición (21d). La columna final indica la sumatoria total
Cuadro 2.8: Incidencias promedio registradas en los fragmentos de raíz a los 7,14 y 21 días desde V1, tanto en capacidad de campo (S-) como en sequía (S+)46
Cuadro 3.1: Tabla resumen de cambios de expresión para el enriquecimiento GO76
Cuadro 3.2: Tabla resumen de cambios de expresión para el enriquecimiento KEGG
Cuadro 3.3: Genes diferencialmente expresados en el enriquecimiento GO. Los genes resaltados en amarillo se descartan como contribuyentes de la enfermedad luego del filtrado por las comparaciones de efecto sequía y patógeno. Los 10 genes intervinientes en las vias de unión de proteínas no plegedas y plegado de proteinas son los mismos, con lo cual se los agrupa en el mismo conjunto de celdas

Cuadro 3.4: Tabla resumen de los genes diferencialmente expresados en el análisis de enriquecimiento KEGG para la comparación (DT\_HS vs DM\_HS)......87

Cuadro 3.5: Genes diferencialmente expresados resultantes del análisis Cuadro 3.6: Genes diferencialmente expresados resultantes del análisis KEGG......93 Cuadro 3.7: Genes compartidos en las vías de enriquecimiento tanto de la plataforma GO como KEGG......96 Cuadro 4.1: Cantidad de SNPs e INDELs encontrados en el transcriptoma por Cuadro 4.2: Listado de los 12 SNPs seleccionados para ser utilizados como marcadores moleculares. ChromPos= Numero y posición del cromosoma en donde se encuentra el SNP, GeneName= ID del gen según la notación "Wm82.a4.v1", REF= alelo de la referencia y DM, ALT= alelo alternativo presente en el DT, EFFECT= efecto del SNP predicho por SNPeff, IMPACT= impacto del SNP predicho por SNPeff, log2FC= Expresión diferencial en genotipo DT vs DM en presencia de la enfermedad, padj= valor p ajustado de la expresión diferencial, GO Biological Process Description=

# ABREVIATURAS

ABA	Ácido abscísico
ABCPE	Área bajo la curva del progreso de la enfermedad
APG	Ágar papa glucosado
Avr	Genes de avirulencia del patógeno
CC	Capacidad de campo
CC-BP-MF	Composición celular, proceso biológico, función molecular
CML	Del inglés, calmodulin like protein (calcium)
DM	Genotipo susceptible DM4800
DPI	Del inglés, days after inoculation.
DT / D86	Genotipo tolerante D868286
ERAD	Del inglés, endoplasmic reticulum associated protein degradation
ERF	Del inglés, ethylene respose factors.
ET	Etileno
ETI	Del ingles, effector triggered inmunity
FPKM	Del inglés, Fragments Per Kilobase of transcript per Million mapped reads
GO	Del inglés, gene ontology
GR	Genotipo Toyoshirome
GRP	Del inglés, glucose regulated protein
GWAs	Del inglés, genome wide association
HSP	Del inglés, heat shock proteins
INDELS	Del inglés, insertion and deletions
INT-PP	Interacción planta patógeno
JA	Ácido jasmónico
KEGG	Del inglés, Kyoto Encyclopedia of Genes
MAMPs	Del inglés, metabolism-associated molecular patterns
MIE	Manejo integrado de enfermedades
MJ	Metiljasmonato
MP	Macrophomina pasheolina
NBT	Del inglés, new breeding technologies
PAMPs	Del inglés, pathogen-associated molecular patterns
PC	Podredumbre carbonosa
PPRE	Procesamiento proteínas en retículo endoplasmático
PR	Proteínas relacionadas a la patogénesis
PTI	Del inglés, pattern triggered inmunity
RLK	Del inglés, receptor like protein kinase
ROS	Del inglés, reactive oxygen species
SA	Ácido salicílico
SNP	Del inglés, single nucleotide polimorphisim.
TPR	Del inglés, tetratricopeptide repeat
UPR	Del inglés, unfolded protein response

### RESUMEN

# Identificación de genes y marcadores moleculares asociados con la tolerancia a la podredumbre carbonosa causada por *Macrophomina phaseolina* en soja

La podredumbre carbonosa constituye una de las principales problemáticas actuales a abordar en el cultivo de soja (Glycine max (L.) Merrill) en Argentina. Las estrategias de manejo incluyen métodos culturales, fungicidas y control biológico, pero estos no han sido efectivos o ampliamente adoptados y han proporcionado un control limitado. En este escenario, la resistencia genética puede ser el método más factible y sostenible para manejar la enfermedad. La selección asistida por marcadores acelera los procesos de selección en el marco de un programa de mejoramiento genético. En este sentido, el presente trabajo tuvo como objetivo principal la identificación de genes de interés, para posteriormente proponer un panel de marcadores moleculares de tolerancia al patógeno en el cultivo de soja. Inicialmente se trabajó con genotipos que evidenciaran comportamientos contrastantes frente a la enfermedad. Se realizó una caracterización de la respuesta de dichos genotipos, a partir de la medición de variables epidemiológicas. El genotipo tolerante evidenció menores niveles de severidad y tasa de infección. Posteriormente, se realizó el estudio RNA-seq, el cual mostró que el genotipo tolerante sobreexpresaba genes y vías metabólicas de relevancia frente a la enfermedad, como lo constituye la interacción planta-patógeno y el procesamiento de proteínas a nivel de retículo endoplasmático. Luego, se realizó la búsqueda de variantes en la secuencia transcriptómica que estuvieran asociadas a la respuesta del genotipo tolerante. Finalmente, a partir de dicha información se propuso un panel de 12 SNPs, distribuidos en 9 genes únicos, entre los cuales se destacan los que codifican para proteínas SNAP (involucrados en la respuesta a factores bióticos) y los que están asociados a la respuesta al estrés, tales como los involucrados en el plegamiento de proteínas. Dichos SNPs podrían contribuir en programas de mejoramiento, posterior a un proceso de validación.

**Palabras clave:** podredumbre carbonosa de la soja, genotipos contrastantes, análisis transcriptómico, marcadores moleculares.

### ABSTRACT

# Identification of genes and molecular markers associated with tolerance to charcoal rot caused by *Macrophomina phaseolina* in soybean

Charcoal rot is one of the main current problems to be addressed in soybean crop (Glycine max (L.) Merrill) in Argentina. Management strategies include cultural methods, fungicides and biological control, but these have not been effective or widely adopted and have provided limited control. In this scenario, genetic resistance may be the most feasible and sustainable method to manage the disease. Marker-assisted selection accelerates selection processes within the framework of a breeding program. In this sense, the main objective of this work was the identification of genes of interest, to subsequently propose a panel of molecular markers of tolerance to the pathogen in soybean. Initially, we worked with genotypes that showed contrasting behaviors against the disease. A characterization of the response of these genotypes was carried out, based on the measurement of epidemiological variables. The tolerant genotype showed lower levels of severity and infection rate. Subsequently, the RNA-seq study was carried out, which showed that the tolerant genotype overexpressed genes and metabolic pathways of relevance to the disease, such as the plant-pathogen interaction and protein processing at the endoplasmic reticulum level. Then, we searched for variants in the transcriptomic sequence that were associated with the response of the tolerant genotype. Finally, based on this information, a panel of 12 SNPs was proposed, distributed in 9 unique genes. Proteins involved in the response to biotic factors can be found (such as SNAP proteins) along with genes associated with the response to stress (such as those involved in protein folding). These SNPs could contribute to improvement programs, after a validation process.

**Keywords:** soybean charcoal rot (PC), contrasting genotypes, transcriptomic analysis, molecular markers.

### Capítulo 1. Introducción general

#### 1.1 Estructura de la tesis

Esta tesis se estructura en 5 capítulos a través de los cuales se analizará cuáles son los genes que se activan diferencialmente en genotipos susceptibles y tolerantes al ataque de la podredumbre carbonosa causada por Macrophomina phaseolina en el cultivo de soja, así como la búsqueda de polimorfismos de un solo nucleótido (SNPs, del inglés single nucleotide polimorphism) con el fin de desarrollar marcadores moleculares de tolerancia al patógeno. El presente capítulo constituye una introducción general a los conocimientos sobre el tema y culmina con el planteo de las hipótesis que fueron puestas a prueba durante el desarrollo de este trabajo. El segundo capítulo abordará una caracterización epidemiológica sobre los genotipos empleados y de los aislamientos del patógeno a utilizar. El tercer capítulo detallará el estudio transcriptómico llevado a cabo y el análisis de dicha información, haciendo hincapié en los genes diferencialmente expresados en el genotipo tolerante con respecto al susceptible. El cuarto capítulo consta de la búsqueda de SNPs, su filtrado, y la selección de un panel resultante, propicio para ser utilizado en programas de mejoramiento. Por último, el quinto capítulo (Discusión general) desarrollará una discusión desde una mirada holística de los resultados, y el contraste de estos con las hipótesis planteadas. Por otro lado, en este capítulo se hará una breve reseña acerca de las perspectivas futuras y las conclusiones generales de este trabajo.

#### 1.2 Revisión de antecedentes

#### 1.2.1 Cultivo de soja en Argentina y sus enfermedades

La soja (*Glycine max* (L.) *Merr.*) es una herbácea anual, de ciclo primavero-estival. Su sistema radicular es profuso, pudiendo la raíz principal alcanzar hasta un metro de profundidad En la raíz principal o en las secundarias se encuentran los nódulos, en número variable.

Las flores se encuentran en inflorescencias racimosas axilares en número variable (Figura 1.1). La semilla es rica en proteínas y en aceites. En algunas variedades mejoradas presenta alrededor del 40-42% de proteína y del 20-22% en aceite, respecto a su peso seco (Liener. 1994). La soja en su evolución estableció interacción simbiótica con bacterias del suelo fijadoras de nitrógeno atmosférico (N<sub>2</sub>), denominadas en forma genérica rizobios. Como fruto de esta simbiosis forman nuevas estructuras en las raíces: los nódulos, donde se realiza la fijación biológica de nitrógeno. Las especies más frecuentes en los nódulos pertenecen al género *Bradyrhizobium*.

La soja se cultiva en Argentina desde 1960 y ha sido el cultivo de más rápida adopción y expansión en la historia de la agricultura argentina (Cabrera. 2004). En la actualidad, es el principal producto de la agricultura nacional, por su participación en el valor total de la producción y de las exportaciones (Escande. 2005; Haro. 2017), al igual que en rentabilidad e ingreso económico nacional, con un área sembrada de 16.8 millones de hectáreas en la campaña 2023/24 (Bolsa de Comercio Rosario. 2024).



Figura 1.1: Vista externa de una planta de soja

El aceite y la harina de soja, junto con el grano sin industrializar, representan la mayor fuente de exportación e ingresos para el país. Dentro del territorio argentino, su expansión ha sido notable. Actualmente ocupa diversas regiones del país, desde la región central de la Pampa húmeda (Provincia de Buenos Aires, Córdoba, Santa Fe y Entre Rios) hasta zonas del noroeste (Salta, Tucumán y Santiago del Estero) y noreste argentino (Chaco, Formosa y Corrientes), como se puede ver en los siguientes mapas de zonas productivas representativas de Argentina (Figura 1.2).



Figura 1.2: Superficie sembrada en verde, representativa de las zonas donde se cultiva soja actualmente en la República Argentina (SAGPyA)

Los factores que limitan la productividad de la soja están relacionados, entre otros, con la escasez de recursos tales como la disponibilidad de agua y nutrientes, además de las adversidades de índole biótico, como plagas y enfermedades. Específicamente en cuanto a enfermedades, se estima que a nivel mundial las pérdidas de rendimiento ocasionadas por las enfermedades oscilan entre el 10 y el 15% (Oerke, 2006). Debido a la expansión del cultivo de la soja en el mundo, las enfermedades se han incrementado en número y severidad (Bandara et al. 2020).

En Argentina los daños causados, anualmente, por las enfermedades de soja alcanzan del 8 al 10% del total de la producción (Ivancovich, 2011; Rizzi. 2021). El desarrollo de las

enfermedades de soja en Argentina se ve favorecido, entre otras causas, por las condiciones ambientales, el aumento de la superficie sembrada, el monocultivo, el empleo de germoplasma de escasa variabilidad, y el uso de nuevas técnicas de manejo del cultivo. El panorama sanitario de la soja en Argentina ha cambiado, pasando de ser un cultivo prácticamente libre de enfermedades, en la década del 80, a la situación actual, con dificultades crecientes. Es indispensable tener presente que las enfermedades afectan el normal desarrollo del cultivo perjudicando los procesos de generación del rendimiento, así como la calidad final del grano y provocando grandes pérdidas económicas (Kantolic y Carmona. 2006).

Entre las enfermedades de mayor relevancia por los daños producidos se encuentran las causadas *M. phaseolina, Fusarium tucumaniae, Fusarium virguliforme, Rhizoctonia solani, Colletotrichum truncatum, y Phomopsis sojae,* que tienen importancia económica en otros países productores (Hosseini et al. 2023). El impacto mundial del crecimiento de las enfermedades en soja ha sido muy relevante (Nair. 2023), y en Argentina este fenómeno se profundiza por el incremento en la superficie sembrada (corrimiento de la frontera agrícola) y la incorporación de la siembra directa en las últimas décadas (Carmona et al. 2015). En este sentido, es indispensable la búsqueda de alternativas que ayuden a minimizar los impactos negativos mencionados.

### 1.2.2 Podredumbre carbonosa causada por M. phaseolina

La podredumbre carbonosa (PC) ha surgido como una de las principales preocupaciones en amplias regiones de las zonas productoras de soja de Argentina. El patógeno puede reducir tanto el rendimiento como la calidad de la semilla (Smith y Wyllie. 1999). La abundante producción de microesclerocios negros (Figura 1.3) hace que los tejidos vegetales infectados se ennegrezcan y, por lo tanto, la enfermedad se conoce como podredumbre carbonosa (Sarr et al. 2014). Taxonómicamente, pertenece al filo Ascomycota del reino Fungi, familia *Botryosphaeriaceae*. La PC se distribuye por todo el mundo en los trópicos y subtrópicos (Wyllie. 1988), y se sabe que *M. phaseolina* puede infectar a más de 500 especies de plantas de importancia económica incluyendo maíz, sorgo (Adeyanju et al. 2015) y girasol (Pawlowski et al. 2015). Anualmente, la PC es una preocupación mayor en el NOA debido a las frecuentes condiciones de calor y sequía que tienden a ocurrir durante las etapas importantes del ciclo del cultivo. El hongo sobrevive en el suelo principalmente como microesclerocios que son estimulados por exudados de la raíz. Posteriormente, germinan e infectan los tejidos de la planta hospedante. Aumentos en la temperatura ambiente (en torno al rango 28-35°C) incrementan la severidad de la enfermedad, junto con menores contenidos de humedad del suelo (Smith y Wyllie. 1999). El patógeno puede invadir los tallos a partir de raíces infectadas, obstruyendo los tejidos vasculares en la raíz del cultivo (Kaur et al. 2012). También puede pasar a la semilla (puede transmitirse entre campañas por semilla infectada), lo que causa una germinación reducida y la pudrición de las plántulas.



*Figura 1.3: Lote de plantas de soja con podredumbre carbonosa en estadios avanzados del ciclo del cultivo.* 

La enfermedad puede atacar en cualquier momento del ciclo del cultivo, evidenciando síntomas en estadios tanto vegetativos como reproductivos. Si el ataque se produce en algún estadio vegetativo suele causar tizón de plántula o damping-off, mientras que, si ocurre en post floración, es común que cause intenso amarillamiento foliar, detención del desarrollo de las vainas/semillas y podredumbre carbonosa de la base del tallo/raíces (presencia de microesclerocios negros, interna y/o externamente). Dichas estructuras de resistencia son particularmente evidentes en los lotes de soja en las etapas de crecimiento R5 (comienzo de la semilla), R6 (plena semilla) y R7 (maduración inicial) (Sanches, 2020). En situaciones graves, las plantas enfermas pueden marchitarse y morir prematuramente (Figura 1.4).



Figura 1.4: Ciclo epidemiológico de la enfermedad podredumbre carbonosa (Smith et al. 2014).
(A) M. phaseolina hiberna por microesclerocios en residuos de cultivos. (B) Raíces infectadas después de estar en contacto o en proximidad con microesclerocios. (C) El hongo crece dentro del tallo y la raíz afectando y alterando el sistema vascular. Se producen más microesclerocios.
(D) La presencia abundante de microesclerocios en la parte inferior del tallo y en el tejido de la raíz pivotante da una apariencia similar a la del carbón. Los residuos infectados se convertirán en una fuente potencial de inóculo para el próximo cultivo plantado.

### 1.2.3 Sistema suelo-planta-atmosfera: factores relacionados a la enfermedad

La PC se presenta frecuentemente en las regiones productoras del NEA, NOA y centrosur de Córdoba con elevada intensidad. El estrés hídrico es un factor crítico para la presentación de la enfermedad en el cultivo (Almeida, 2003), por lo que frecuentemente las epifitias se asocian a períodos de altas temperaturas (superiores al rango 28-35°C) coincidentes con bajas precipitaciones (Perez Brandam et al. 2009). Ensayos realizados en la provincia de Salta en la campaña 2006/2007 comprobaron relaciones significativas entre la incidencia de la podredumbre carbonosa y las precipitaciones. Las precipitaciones aparecen como la variable de mayor influencia para ambas campañas agrícolas, señalándose que a mayores precipitaciones menor incidencia de la enfermedad. Al evaluarse el efecto de las precipitaciones y las temperaturas máximas sobre el área bajo de la curva del progreso de la enfermedad (ABPCE), se comprobó que su influencia era altamente significativa (p<0,0001) durante los meses de febrero y marzo en áreas productoras de Salta. En general, a mayores temperaturas durante febrero y marzo mayor es el ABCPE de la PC; por el contrario, a mayores precipitaciones, menor es el ABCPE (Figura 1.5).



*Figura 1.5: Área bajo la curva de progreso de la PC (podredumbre carbonosa) de la soja según (a- las temperaturas máximas / b- precipitaciones) en Salta durante la campaña 2006/7 (Perez Brandam et al. 2009).* 

De hecho, la temperatura optima de crecimiento del patógeno se encuentra en el rango de entre 30 y 35°C. Situaciones hídricas en las cuales el suelo supera el 60% de la capacidad de campo

(CC), se ha visto que la presencia del hongo se ve desfavorecida. Asimismo, la cantidad de microesclerocios generados disminuye con el aumento del contenido hídrico. No obstante, a pesar de ello, la colonización por parte del patógeno no es evitada con altos contenidos hídricos. De hecho, el crecimiento de *M. phaseolina* en tejidos del hospedante a un amplio rango de  $\Psi$ a indica que este organismo ha desarrollado mecanismos de adaptación para sobrevivir a dichas condiciones. La adaptación a rangos de  $\Psi$ a amplios puede ser una estrategia que capacita al saprófito o patógeno latente a sobrevivir en los tejidos del hospedante. El hongo puede ser capaz de continuar el desarrollo vegetativo y crecer bajo condiciones de baja humedad, que, por otra parte, sí afectan al hospedante (Cardona, 2006).

Factores edáficos de relevancia también modulan la posibilidad de ataque por parte de *Macrophomina*, aunque menos relevantes que el estado hídrico del cultivo. Uno de los principales factores edáficos es el efecto salino. Se ha reportado que el patógeno en cuestión puede aumentar sus tasas de crecimiento sin ninguna dificultad cuando las concentraciones totales de sales oscilan entre 0 y 1000 Mm, siendo las concentraciones optimas del 0 al 500 Mm, lo que se refleja a través de la medición de su patogenicidad (Figura 1.6a) y diámetro de colonia in vitro (Figura 1.6b) en diferentes genotipos evaluados (Bañuelos-Balandrán, 2008).



Figura 1.6: (a): Patogenicidad de M. phaseolina en dos cultivares de poroto bajo diferentes concentraciones de NaCl in vitro. Barras indican  $\pm$  error estándar. (b) Diámetro de colonias (mm) de cuatro aislamientos de M. phaseolina en APG (ágar papa glucosado) ajustado a diferentes concentraciones de NaCl. El diámetro de la colonia se registró después de 72 h de incubación. Líneas verticales indican desvío estándar (Bañuelos-Balandrán. 2008)

Como se puede ver en la Figura 1.6b, el diámetro de colonia es decreciente ante el incremento en la salinidad, notándose diferencias en la amplitud de dicha caída, atribuibles a las diferencias entre genotipos. Por otro lado, la Figura 1.6a muestra que la patogenicidad sigue el mismo patrón del diámetro de colonia, pero se pueden distinguir distintos rangos. Entre concentraciones salinas bajas (0-500 mM NaCl) la patogenicidad se mantiene en un rango máximo, cayendo a concentraciones mayores por fuera de dicho rango (>500 mM), siendo la patogenicidad mínima a una concentración en torno a los 1000mM (Bañuelos-Balandrán, 2008).

Además del estado hídrico y la salinidad, la relación C/N (y el pH resultante) del suelo es también importante, dado que, a mayores niveles, el patógeno tiene mayor sustrato para continuar su crecimiento y poder cumplir el ciclo, aumentando la cantidad de inóculo inicial y, por consiguiente, aumentando las posibilidades de infecciones en ciclos posteriores (Li et al. 2023)

### 1.2.4 Interacción planta – patógeno y defensa vegetal

La PC causada por *M. phaseolina* infecta las raíces de las plantas del suelo mediante un mecanismo mediado por toxinas, secretando phaseolinona y botryodiplodin (Figura 1.7).



*Figura 1.7: Toxinas actuantes en la PC (podredumbre carbonosa) causada por M. phaseolina* (Ramezani et al. 2007)

Aunque no se conocen con certeza los mecanismos de la infección por PC a través de toxinas, se estima que la infección del cultivo a través de las raíces se produce por dos mecanismos. El primer mecanismo es la penetración física del tejido y el segundo a través de la secreción de toxinas, permitiendo que el hongo ingrese a través del área necrótica creada por la toxina. La infección de las plantas mediada por toxinas ocurre (1) por secreción de enzimas

hidrolíticas o toxinas que inducen la activación de enzimas hidrolíticas endógenas; y (2) mediante la secreción de toxinas para dividir las células, apuntando al tejido meristemático en las puntas de las raíces y proporcionando un acceso conveniente del hongo al sistema vascular de la planta infectada (Ramezani et al. 2007). Los compuestos fenólicos, como la lignina y los flavonoides son inhibidores efectivos de las enzimas hidrolíticas, que conducen a la resistencia fúngica a la infección enzimática hidrolítica mediada por toxinas (An et al. 2023)

Uno de los principales abordajes actuales para el control de la enfermedad es la existencia de resistencia genética. Hay evidencias que en los genotipos moderadamente resistentes (tales como AG3905) hay mayores concentraciones de fenólicos en semillas, isoflavonoides totales y lignina de la cubierta, tanto en condiciones infectadas y no infectadas, como en condiciones irrigadas o sin irrigación, en comparación con genotipos susceptibles (Bellaloui et al. 2014).

### 1.2.5 Genes, vías metabólicas y mecanismos involucrados

Para mejorar el conocimiento de los cambios moleculares que ocurren cuando la soja es atacada por *M. phaseolina* y la exposición a la sequía, ya sea en conjunto o por separado, es importante centrarse en el sistema de raíces, ya que es la primera región afectada por el patógeno. En estudios genómicos de raíces bajo sequía, los estudios transcriptómicos y proteómicos han revelado la importancia de la producción de especies reactivas de oxígeno (ROS), la señalización celular, la modificación de la pared celular y la respuesta fitohormonal (Vaughn y Nguyen, 2012). Hay muchas familias de genes con cambios en la expresión que se han informado cuando la soja y otros cultivos se colocan bajo condiciones de estrés hídrico y en presencia de patógenos necrotrófos como *M. phaseolina*. La señalización es compleja y las principales interrelaciones están presentas en la siguiente figura (Figura 1.8).



Figura 1.8: Modelo de las interacciones entre la sequía y podredumbre carbonosa. Las primeras etapas de estrés por sequía conducen a la producción de especies de oxígeno reactivo (ROS), que implican la reacción de hipersensibilidad y la muerte celular (Laloi et al. 2004); (Guan et al. 2000). La muerte celular del hospedante aumenta el crecimiento del patógeno. Fuente: elaboración propia a partir de los trabajos de (Mauch-Mani y Mauch. 2005), (Pandey, Ramegowda y Senthil-Kumar. 2015) y (Radwan et al. 2014)

Las primeras etapas de estrés por sequía conducen a la producción de ROS, que implican la reacción de hipersensibilidad y la muerte celular (Laloi et al. 2004; Guan et al. 2000). La muerte de las células del hospedante promueve las condiciones propicias para la infección y el posterior crecimiento del patógeno *M. phaseolina,* que al ser necrotrófo, se nutre de tejido muerto. Dentro de todas las interrelaciones existentes entre ambos estreses: abiótico (sequía) y biótico (patógeno), es importante destacar el rol de ciertas proteínas y segundos mensajeros actuantes, tales como las ROS, el Ca<sup>2+</sup>, las quinasas dependientes de ciclinas (CDPK) y las map-quinasas (MAPK).

Por un lado, los genes implicados en la señalización de ROS y Ca<sup>2+</sup> constituyen una parte importante de la respuesta molecular compartida de las plantas ante situaciones de estrés (Rizhsky et al. 2004; Johnson et al. 2014). Hay evidencia de que los elicitores fúngicos activan

una rama de la red de señalización que se comparte con la señalización ABA en la regulación de los canales de  $Ca^{2+}$  localizado en la membrana plasmática (Klusner et al. 2014).

La señalización de  $Ca^{2+}$  es importante en la expresión de la resistencia a la enfermedad y en movimientos estomáticos controlados por ácido abscísico (ABA) y respuestas a la deshidratación (Blume et al. 2000; Knight et al. 1997).

Por otra parte, análisis de diferentes genes de CDPK realizados en *Triticum aestivum* mostraron que de 12 CDPK que respondían a *Blumeria graminis patovar tritici* (agente causal de mildiú polvoriento en el trigo), 8 también respondieron a los estreses abióticos, lo que corrobora que son un punto importante de *cross-talk* (Li et al. 2008). De forma similar, los genes implicados en la vía de eliminación de las especies reactivas de oxígeno (ROS, del inglés *reactive oxygene especies*) como ascorbato peroxidasa han evidenciado que confieren tolerancia contra diversos estreses abióticos y bióticos (Sarowar et al. 2005; Choi y Hwang, 2012).

Asimismo, las MAPK juegan un rol importante en la señalización de las plantas como respuesta a la sequía. Se ha caracterizado la expresión de las MAPK como un regulador positivo de la resistencia al estrés abiótico. Las proteínas y segundos mensajeros mencionados anteriormente (Ca<sup>2+</sup>, ROS, CDPK, MAPK) modulan la expresión de distintos factores de transcripción, siendo estos grupos numerosos y de distinta índole, que se destacan por desempeñar un papel crucial en la regulación de la respuesta al estrés biótico y abiótico de las plantas. Los de mayor relevancia son aquellos pertenecientes a los grupos de MYB, DREB y bZIP-NAC. Entre los factores MYB, por ejemplo, OsMYB4, cumple una función reguladora importante para diferentes vías de señalización hormonal y tiene un papel en conferir resistencia a estrés biótico y abiótico a las plantas (Mengiste et al. 2003; Iriti et al. 2007; Zhang et al. 2009). Entre los de respuesta a la deshidratación (DREB), las proteínas DREB/CBF son factores de transcripción que inducen un conjunto de genes de respuesta a estrés abiótico. Cumplen una importante función en las vías de señalización independientes de ABA. Los factores de

transcripción de las familias bZIP-NAC promueven la activación de diversos genes, entre los cuales se destacan el rol de las chaperonas, acuaporinas, la osmoprotección y las enzimas superóxido dismutasa y ascorbato peroxidasa. Todas ellas forman parte de una maquinaria altamente efectiva para la detoxificación de las ROS. (Yang et al. 2004; Rabello et al. 2008; Spollen et al. 2008; Gutiérrez- González et al. 2010; Le et al. 2011; Wang et al. 2011).

No obstante, todas estas proteínas, segundos mensajeros y factores de transcripción no son los únicos que intervienen en todo el proceso de respuesta ante el estrés. Como se dijo anteriormente, la infección por parte de *M. phaseolina* genera una reacción de hipersensibilidad, que concluye en la inducción de las vías de defensa vegetal, donde se activan las hormonas ácido jasmónico (JA), salicílico (SA) y etileno (ET). Las hormonas vegetales JA, SA y ET juegan un papel importante en la resistencia a la enfermedad. Estudios recientes muestran que al tratar las plantas con metil jasmonato (MJ) o ET, resistencia parcial a *M. phaseolina* fue observada (Gaige et al. 2010; Marquez et al. 2021).

El siguiente gráfico (Figura 1.9) ilustra el porcentaje (%) de plantas enfermas en función de los días post inoculación (DPI, del inglés *days post inoculation*) de *M. phaseolina* en plántulas de *Medicago truncatula*, a partir de la aplicación de MJ, ET o la combinación de ambos:



Figura 1.9: Porcentaje (%) de plantas enfermas contabilizadas en los diferentes DPI (días post inoculación) de M. phaseolina. Líneas verticales indican desvío estándar de las mediciones.

Como se puede ver en el gráfico, entre el 1° y el 4° DPI, todos los tratamientos concedieron resistencia parcial a *M. phaseolina* en comparación al control. En el 5° DPI, tanto el MJ como el ET obtuvieron resultados similares y superiores al control, pero a su vez fueron superados por la mezcla de ambos (MJ+ET). En el 6° DPI, tanto el MJ como el ET no fueron estadísticamente diferentes al control, en contraste con MJ+ET, que sí continuaban siéndolo. A partir del último DPI analizado, no hubo diferencias entre tratamientos. Es un claro ejemplo de la resistencia parcial que puede conferir el MJ y ET en el control de *M. phaseolina*.

Por otra parte, el SA actúa principalmente a través de la activación de la vía WRKY. Varios genes WRKY están presentes en distintos cultivos agrícolas de interés económico tales como el arroz (*O. sativa* WRKY8 -OsWRKY89-). En el caso del ET actúan factores sensibles al etileno (ERF, del inglés *ethylene respose factors*) como GmERF, que son reguladores importantes de diferentes vías de señalización hormonal y tienen un papel en conferir resistencia a enfermedades. Entonces, a partir del SA (a través de la inducción de la vía WKRY) y el ET-JA (a través de la vía de los ERF), se logra la activación de las proteínas asociadas a la patogénesis (PR) y las defensinas (Radadiya et al. 2021).

No obstante, es importante conocer el rol del ABA para ver su interrelación con estas hormonas. El ABA juega un papel importante en muchos aspectos del desarrollo de la planta, en la regulación de la apertura estomática y en el inicio de la adaptación de respuestas a diversas condiciones ambientales, entre ellas los estreses de diversa índole: adaptación a la sequía, la baja temperatura y la salinidad están reguladas por la actividad combinatoria de las vías de señalización dependientes e independientes de ABA. La evidencia actual sugiere que el ABA afecta la resistencia a la enfermedad principalmente de forma negativa. Las altas concentraciones de ABA reducen fuertemente los niveles de transcripción de JA y/o ET, mientras que en las plantas mutantes de ABA (deficiente) hubo un aumento de la expresión de dichas vías (ET y

JA). El único rol -menor- encontrado actualmente en el cual el ABA actuaría de forma positiva en cuanto a la resistencia vegetal sería a través de la deposición de calosa (Wang et al. 2021).

#### 1.2.6 Manejo integrado de enfermedades (MIE)

Patógenos habitantes del suelo como *M. phaseolina* son de muy difícil manejo. No sólo por ser necrotrófos y estar ampliamente distribuido al ser polífago, sino que adicionalmente puede generar estructuras de resistencia que les asegura viabilidad y persistencia en el suelo. Por todo ello, la eficiencia de alternativas de manejo de estos patógenos tales como control biológico, supresividad y rotación de cultivos, ha sido insuficiente para su control. Asimismo, las estrategias que incluyen métodos culturales, y fungicidas aplicados en semillas, tampoco han sido efectivos o ampliamente adoptados y han proporcionado un control limitado (Mengistu et al. 2015). En este contexto, se hace fundamental una búsqueda constante de nuevas alternativas.

Esta problemática, sumada a la necesidad actual de preservar el ambiente, disminuir los costos de producción y atender a problemáticas de suelo obliga a diseñar nuevas estrategias de manejo de patógenos que permitan complementar o disminuir el uso de los fungicidas. Resulta entonces interesante explorar alternativas de manejo que puedan ser complementarias a las ya conocidas, y que estén enmarcadas dentro de un manejo integrado de enfermedades (MIE). En este escenario, la resistencia genética puede ser el método más factible y sostenible para manejar la PC (Mengistu et al. 2007). La resistencia total a *M. phaseolina* no ha sido informada en ninguna especie vegetal, pero se ha notificado la identificación de resistencia parcial en el cultivo de soja, incluidos cultivares moderadamente resistentes, como DT97-4290, que se utiliza como estándar de control de enfermedades (Paris et al. 2006; Mengistu et al. 2007, 2013; Twizeyimana et al. 2012; Pawlowski et al. 2015), al igual que Munasqa RR (Reznikov et al. 2019). Sin embargo, las investigaciones sobre germoplasma disponible comercialmente y su respuesta general al patógeno no se han realizado ampliamente. En este sentido, la biotecnología

entonces aparece como una herramienta fundamental para ser utilizada, dado que nos permite abordar la problemática de forma holística hacia la mitigación de los efectos de la enfermedad.

#### 1.2.7 Transcriptómica

Como se pudo ver, el esquema de interrelaciones es complejo dado todos los niveles en los cuales actúa la respuesta de la planta ante la infección, como así la respuesta ante el estrés abiótico. En este contexto, cobran importancia las nuevas técnicas de mejoramiento, conocidas como nuevas técnicas de mejoramiento (NBTs del inglés new breeding techniques), que han agilizado el proceso de mejoramiento y selección. Entre ellas, se encuentran técnicas de estudios de asociación del genoma completo (GWAs del inglés genome wide association studies) que permiten asociar grandes cantidades de SNPs a rasgos de interés; o la edición génica mediante CRISPR-CAS9 que permite la introducción de una secuencia deseada a través de una reparación homóloga. Gracias a esta (y otras) técnicas se pueden editar genes de manera precisa, sin embargo, hay que identificar cuáles son aquellos genes que justifican fijar la atención, y cómo. Es por esto que la transcriptómica (Figura 1.10) es una de las herramientas utilizadas para encontrar cuáles son los genes, y las rutas metabólicas, involucrados en la respuesta al patógeno en diferentes condiciones ambientales. Diversos trabajos han abordado a nivel transcriptómico la respuesta vegetal de la soja frente a enfermedades (Marquez et al. 2018), incluyendo las nombradas GWAs (Almeida-Silva, 2021). La realización de estudios transcriptómicos con líneas contrastantes provee información detallada sobre genes diferencialmente expresados que podrían ayudar a la identificación de aquellos genes involucrados en la tolerancia a la enfermedad. Es probable que exista una relación entre los genes expresados diferencialmente y las vías metabólicas explicadas anteriormente.

Las herramientas disponibles para estudios genómicos y transcriptómicos hacen posible el estudio pormenorizado de las bases moleculares que gobiernan los mecanismos de resistencia de

las plantas, así como también, la detección de polimorfismos de secuencia dentro de los genes principales. A partir de este planteo, por ejemplo, sería posible el diseño de un panel de marcadores moleculares SNPs, para asistir el mejoramiento hacia variedades resistentes para ser utilizadas en sistemas de producción más amigables con el ambiente, y, por lo tanto, sustentables. De forma genérica, se puede esquematizar la aplicación agronómica de la biotecnología a través de algunas de las principales ómicas que se utilizan en mejoramiento genético vegetal:



Figura 1.10: Esquema simplificado de algunas de las principales aproximaciones omicas utilizadas en biotecnología.

Los estudios transcriptómicos masivos resultan fundamentales para poder dilucidar los mecanismos que operan detrás de la respuesta vegetal frente a un patógeno. De hecho, este en trabajo de tesis se utilizará una aproximación detranscriptómica masiva, con el objetivo de identificar genes y desarrollar marcadores moleculares SNP con el propósito de aportar nuevas tecnologías que contribuyan a un manejo integrado y sostenible de la enfermedad

### 1.3 Objetivo principal

Identificar cuáles son los genes, y las rutas metabólicas asociadas, que se activan o deprimen diferencialmente en respuesta al patógeno *M. phaseolina* en genotipos de soja (*G. max* L.) con comportamientos contrastantes ante dicho estrés.

### 1.4 Objetivos específicos

I) Identificar, mediante ensayos epidemiológicos, genotipos de soja con respuestas contrastantes frente a la enfermedad, e identificar el momento adecuado para toma de muestra de tejidos radiculares para el estudio RNA-seq.

II) Analizar la expresión diferencial de genes en dichos genotipos contrastantes e identificar los genes indicadores más importantes y las rutas metabólicas asociadas.

III) Identificar polimorfismos de un solo nucleótido (SNPs) en genes seleccionados y proponer marcadores moleculares de tolerancia a la enfermedad.

### 1.5 Hipótesis

I) Existen genotipos de soja que evidencian respuestas contrastantes a la PC. Las sucesivas mediciones epidemiológicas permitirán vislumbrar diferencias significativas, e indicarán, además, el momento propicio para recolectar el tejido a partir de plántulas que se utilizarán en el estudio RNA-seq

II) Es posible detectar genes diferencialmente expresados a través de un análisis transcriptómico y asociar este comportamiento a procesos bioquímicos relacionados.

III) Existe un polimorfismo estructural en los genes mencionados, a partir del cual será posible desarrollar marcadores moleculares.

# Capítulo 2. Caracterización de genotipos con respuesta diferencial al ataque del patógeno

### 2.1 Introducción

La resistencia a las enfermedades están controladas por múltiples genes de herencia cuantitativa, resultando la búsqueda de resistencia genética muy desafiante. En tal escenario, las metodologías que ayudan a dilucidar los mecanismos de resistencia e identificar los genotipos resistentes contribuyen a un mayor éxito en los programas de mejoramiento (St Clair, 2010). Esto cobra aún mayor relevancia en enfermedades tales como la PC, donde no hay ningún genotipo de resistencia total comprobada. Actualmente, hay poca información sobre las regiones genómicas y marcadores moleculares vinculados a la resistencia a la PC en soja (Da Silva et al. 2019).

Por otra parte, el hecho de que la resistencia genética sea de las principales opciones viables para mitigar dicha enfermedad, dada la poca efectividad de otro tipo de medidas, indica que el posible impacto sobre programas de mejoramiento para esta enfermedad en el cultivo de soja puede ser de una relevancia notable. Este impacto seria especialmente importante en campañas argentinas de soja en las cuales hubiera niveles bajos de precipitación, símil sequía, dado que la preponderancia de la enfermedad es mayor.

Lograr una metodología de fenotipado robusta y segura es necesaria para asegurar una correcta estimación de los efectos del patógeno, teniendo en cuenta además el efecto del estrés hídrico particularmente en esta interacción. La PC es una importante enfermedad que reduce el rendimiento en los países productores de soja en todo el mundo. Como ya se ha mencionado, su efecto es más pronunciado en cultivos bajo estrés abiótico del tipo hídrico.
Las condiciones climáticas globales cambiantes (particularmente la aparición de sequías frecuentes o situaciones similares a las sequías), hacen necesaria la caracterización, por ejemplo, condición vía fenotipado incluya hídrica imperante el un que la en proceso de infección de la enfermedad. En este sentido, es esencial la simulación a nivel experimental de las condiciones propicias para la enfermedad, asemejándola a las situaciones productivas reales en las cuales se produce el ataque del patógeno, y el consecuente efecto sobre el cultivo. La caracterización de los genotipos es de vital importancia ya que sienta las bases para el posterior estudio transcriptómico en el cual se puede analizar y dilucidar cuál es el conjunto de genes más relevantes detrás de la respuesta observada.

El objetivo del presente capítulo fue, por un lado, cuantificar diferentes variables epidemiológicas (relacionadas con la enfermedad del cultivo), que denotaron diferentes comportamientos de los genotipos frente al patógeno. Por otra parte, a partir de información recolectada, se buscó identificar 2 genotipos contrastantes frente a la enfermedad. Para ello, se utilizaron inicialmente 4 genotipos de comportamientos disímiles: D86-8286/Toyoshirome como medianamente tolerantes y DM4800/Williams 82 como susceptibles frente a la enfermedad.

#### 2.2 Materiales y Métodos

Los procedimientos detallados comprenderán en un primer estadio, el análisis de diferentes aislados del patógeno y la selección de uno de ellos (2.2.1). Posteriormente, el desarrollo del ensayo en invernáculo (2.2.2) para la evaluación de la enfermedad en diferentes genotipos y la selección de aquellos con comportamiento contrastante. Por último, se detalla la metodología de ensayos complementarios (2.3.2) para el fenotipado de dichos genotipos contrastantes.

#### 2.2.1 Selección de inóculo

Se identificaron genotipos seleccionados que se encuentran documentados bibliográficamente, que difieren en su respuesta a *M. phaseolina*. Contemplando el desempeño

histórico de cada uno de los genotipos en situación de campo y las descripciones bibliográficas, se decidió seleccionar los genotipos D86-8286 -PI518772 (GM:V)- (Mengistu et al. 2011) considerado como resistente o de buen comportamiento, al igual que Toyoshirome -PI 594302 (GM:VII)- (Mengistu et al. 2013), documentado en estudios realizados también en Brasil (Panisson, 2019). Por otra parte, se seleccionaron los genotipos DM4800 (GM:IV) (GDMseeds) y Williams 82 (GM:III) (USDA), considerados como genotipos susceptibles o de mal comportamiento. Asimismo, Williams 82 se utilizó porque se conoce su secuencia genómica completa. A modo de simplificación el genotipo Toyoshirome se nombrará como GR y el D86-8286 como D86.

Dentro del MIE planteado surge como uno de los métodos preferenciales el uso de cultivares resistentes. En la actualidad se han caracterizado algunos genotipos de resistencia moderada, sin embargo, se desconoce el comportamiento de los cultivares comerciales y las líneas avanzadas. El desarrollo de un método de evaluación de la respuesta de genotipos al patógeno constituye una herramienta fundamental para el mejoramiento genético. Se han evaluado numerosos métodos in vitro, en invernáculo y en condiciones de campo (Smith y Carvil, 1997; Mengistu et al. 2015; Reznikov et al. 2019).

Habitualmente se evalúan varios genotipos frente a un solo aislamiento por lo cual la selección del mismo, previamente a la evaluación, debería ser un paso ineludible. Conociendo que el patógeno en cuestión es polífago, y dada su amplitud geográfica, presenta variaciones tanto fisiológicas como morfológicas, además de genéticas. A partir de lo enunciado se puede inferir que los niveles de agresividad de los aislamientos pueden variar, repercutiendo en la interacción con el genotipo de soja (Reznikov et al. 2019).

En el presente ensayo preliminar se trabajó con 3 aislados de distinto origen y agresividad desconocida (cuadro 2.1). Los mismos fueron obtenidos y conservados por la empresa semillera GDM hasta su entrega para los estudios específicos a la cátedra de Bioquímica de la FAUBA, en

el contexto de un convenio. A cada uno de los aislados se los contrastó con un genotipo susceptible (DM 4800) y otro medianamente resistente (DT 974290); (Bellaloui et al. 2008), comúnmente utilizados en el cultivo de soja.

Cuadro 2.1 – Inóculos puestos a prueba, su origen geográfico, hospedante y año de recolección (campaña agrícola)

M. phaseolina	Origen geográfico	Hospedante y año
Inóculo A	Burruyacú, Tucumán	Soja campaña 2017/2018
Inóculo B	Chacabuco, Bs As	Soja campaña 2017/2018
Inóculo C	Rio Cuarto, Córdoba	Soja campaña 2017/2018

El experimento se condujo bajo un diseño experimental completamente aleatorizado, en un arreglo factorial donde los factores fueron el inóculo (A,B y C) y el genotipo (DM 4800 y DT 974290). Se sembraron en cajas de Petri con agar papa glucosado (APG) y se incubaron en oscuridad durante aproximadamente 7 días a 30°C. Se esperó hasta que las placas fueran cubiertas por los microesclerocios del hongo en su totalidad. Cada uno de los 3 aislamientos se sembró en 2 repeticiones (placas de Petri) para cada uno de los genotipos. Se colocaron 5 semillas por placa. Las semillas se desinfectaron previamente con hipoclorito de sodio al 2% durante 2 minutos, se enjuagaron con agua destilada estéril, se dejaron secar sobre papel absorbente y posteriormente se ubicaron en las cajas con las colonias del hongo en estufa a 30°C en ausencia de luz. Se realizó un ANOVA y las comparaciones se probaron con el test Tukey utilizando el sofware estadístico Infostat (Di Rienzo, 2010).

La severidad del daño causado por *M. phaseolina* se evaluó utilizando la escala descripta por (Manici et al. 1995) y reproducida por (Gil-Langarica et al. 2007), la que consta de 6 clases donde 0= semilla sana, 1= semilla con cierto grado de decoloración, 2= tegumento de la semilla invadido por micelio y microesclerocios, 3= tegumento libre pero la semilla está infectada, 4= tegumento y semillas infectados y 5= la semilla no germina y está completamente invadida por

el hongo. Las clases 0, 1 y 2 se consideraron como reacción de baja agresividad de los aislamientos y los valores > 2, como reacción de alta agresividad (Figura 2.1).



Figura 2.1: Imágenes indicativas de cada uno de los estadios de semillas descriptos en la escala de Manici et al. 1995, desde el estadio 0 hasta el 5 (0= semilla sana, 1= semilla con cierto grado de decoloración, 2= tegumento de la semilla invadido por micelio y microesclerocios, 3= tegumento libre pero la semilla está infectada, 4= tegumento y semillas infectados y 5= la semilla no germina y está completamente invadida por el hongo).

Los datos de la reacción in vitro fueron sometidos al análisis de la variancia y las medias comparadas con el estadístico Tukey (p<0.05). Los factores inóculo (p=0,0217) y genotipo (p=0,0475) resultaron estadísticamente significativos a la hora de explicar las diferencias observadas en severidad (valor p<0,05). A la hora de evaluar la agresividad de cada uno de los

inóculos, quedó en evidencia que el inóculo proveniente de Chacabuco, Provincia de Buenos Aires (inóculo B) resultó ser más agresivo que los otros dos (Tucumán –Inóculo A y Córdoba– Inóculo C), no habiendo diferencia significativa entre ellos (Figura 2.2).

Asimismo, hubo efecto del genotipo en las respuestas de agresividad encontradas. Para el genotipo medianamente resistente (DT 97-4290) la agresividad fue menor y diferente frente al genotipo susceptible (DM 4800). El siguiente gráfico muestra la severidad medida según la escala propuesta (Manici *et al.* 1995), para cada uno de los inóculos y genotipos.



Figura 2.2: Severidad medida según la escala propuesta por (Manici et al. 1995). (\*) Indican diferencias significativas con el resto de los genotipos e inóculos.

Se seleccionó el inóculo proveniente de los campos de soja de Chacabuco, dado que es que evidenció mayor agresividad.

## 2.2.2 Ensayo para evaluación de la enfermedad

Se realizó un ensayo en invernáculo de la Facultad de Agronomía. Se realizaron diferentes mediciones, las cuales se explicarán metodológica y estadísticamente a continuación.



Figura 2.3: (a) Vista externa de una placa de Petri colonizada por M. phaseolina, (b) Maceta llena hasta la mitad con sustrato estéril previa a la inoculación, (c) Inoculación de la placa cortada en fragmentos del micelio, (d) Micelio tapado, siembra y posterior cobertura con sustrato estéril, (e y f) emergencia de las 4 plántulas por maceta. (g) Plántulas alcanzando el estadio fenológico V1.

El experimento llevado a cabo en el invernáculo se condujo bajo un diseño experimental completamente aleatorizado (al igual que las mediciones incluidas en dicho ensayo), con arreglo factorial, siendo los factores: genotipo (4 genotipos de comportamiento contrastante), patógeno

(presencia y ausencia de inóculo) y contenido hídrico (sequía y capacidad de campo). Las semillas de los genotipos mencionados fueron sembradas por quintuplicado (5 repeticiones), utilizándose macetas de 31 de capacidad, y se llenaron con sustrato de maceta Growmix Multipro de Terrafertil (turba de musgo, spragnum de fibras medias, perlita, compost de corteza, cal calcita, caldolomita, agentes humectantes, fertilizante). El sustrato se esterilizó en autoclave a 121°C durante 1 hora.

El ensayo se llevó a cabo en invernáculo con condiciones registradas de temperatura (Apéndice). El factor luz no fue un impedimento para el ensayo, dado que la presencia o ausencia de luz no modifica la infección por parte de *M. phaseolina* (Cuevas Moreno, 2017)

Se sembraron 4 semillas por replica biológica. En paralelo, se cultivó el patógeno in vitro, utilizando medio agar papa glucosado (APG), cubriendo completamente la placa. (Figura 2.3). Se realizaron la cantidad de repiques necesarios para contar con la cantidad adecuada de placas con micelio del patógeno. Se contó con la cantidad adecuada posterior al primer repique.

Se promovió la inoculación del patógeno vía la colocación del micelio crecido en APG directamente en el sustrato, utilizando 1 placa por maceta (aproximadamente 5g), con micelios de 7 días de crecimiento (Cassina, 2017). Se llenaron las macetas hasta la mitad con la mezcla del sustrato, para luego colocar el patógeno. Se cubrió nuevamente con el sustrato, formando una capa de 2cm y luego se sembraron las semillas, que fueron finalmente cubiertas con una capa de sustrato. Se registró la variación diaria de la temperatura durante el transcurso del ensayo.

Por otra parte, se aplicó un tratamiento diferencial de riego, a partir de V1 (Fehr y Caviness, 1977). Como primera medida se tomó un peso inicial anterior al primer riego, es decir corresponde al peso inicial de la maceta con el sustrato sin adicionar agua. Para obtener este valor se secó el sustrato en estufa a 80°C hasta constancia de peso. Posteriormente, para cada una de las macetas se regó y al día siguiente (de la noche a la mañana posterior), cuando dejó de drenar, se pesó para ver el peso de referencia o de saturación que nos permita dilucidar el máximo nivel de retención hídrica de las plántulas en la maceta. La diferencia entre ambos pesos nos dará la máxima capacidad de retención de cada una de las macetas, es decir la cantidad de agua almacenada en CC medida en g o ml. Una vez realizadas dichas mediciones, se registró el peso diario para dilucidar si había que proceder al agregado de agua o no. El criterio utilizado corresponde a un rango de CC fijo (Mian et al. 1996). Para los tratamientos con sequía, cuando la humedad del suelo cayera por debajo del 20% de la capacidad de campo, se agregaría agua individualmente a cada maceta para devolver la humedad del suelo a dicho valor de CC fijo (20%). Se definió un 20% de la CC, dado que es una condición altamente estresante, en las cuales las plantas muestran visibles síntomas de estrés hídrico (por ejemplo, orientación o marchitez de las hojas y doblez de los pecíolos) (Mian et al. 1996). Por otra parte, para los tratamientos control o regados se definió un valor del 80% de la CC, que constituye una condición no estresante. A partir de este ensayo se pudieron realizar diferentes mediciones para las plántulas de cada tratamiento (Figura 2.4), las cuales se detallan en materiales y métodos.

# Genotipo D86-8286



# Genotipo DM4800



Genotipo Williams 82



Genotipo GR



Figura 2.4: Plántulas indicativas de cada uno de los tratamientos para cada uno de los genotipos (D86-8286, DM4800, Williams82 y GR). De izquierda a derecha: H+S+, H+S-, H-S+ y H-S-. Las siglas H+ denotan la presencia de inóculo y el S+ presencia de sequía.

#### 2.3 Parámetros medidos

A partir del ensayo general realizado, se realizaron diferentes mediciones. Las siguientes mediciones se listan a continuación:

- Emergencia de plántulas
- Severidad de raíz y base del tallo de plántulas
- Tasa de infección
- Área bajo la curva del progreso de la enfermedad (ABCPE)
- Incidencia en fragmentos de raíz

Para todas las mediciones nombradas, se trabajó con los genotipos y tratamientos descriptos anteriormente. Los genotipos utilizados fueron DM4800, D86, GR y Williams 82. La existencia del estrés hídrico impuesto determina los diferentes tratamientos realizados, junto con la inoculación de cada una de las macetas: H+ S+ (existencia de estrés hídrico junto con la inoculación del patógeno), H+ S- (inoculación del patógeno en capacidad de campo – sin estrés hídrico), H- S+ (existencia de estrés hídrico pero sin inoculación del patógeno) y H- S- (situación del control, sin estrés hídrico ni inoculación del patógeno).

Una vez realizado el ensayo general, y obtenido conclusiones acerca del comportamiento de los genotipos, se utilizaron los genotipos contrastantes D86 y DM4800 para la realización de ensayos complementarios:

- Severidad en semillas:
- Daño foliar (lesión en cotiledón, hoja cotiledonar y trifoliada)
- Establecimiento de plántulas (longitud y peso fresco)

Todo el procedimiento se encuentra representado esquemáticamente en la Figura 2.5.



Figura 2.5: Esquema resumen de la hoja de ruta para la caracterización de genotipos contrastantes frente a la enfermedad (metodología, mediciones y tipos de ensayos llevados a cabo).

# 2.3.1 Ensayo principal

#### 2.3.1.a Emergencia de plántulas

La primera medición efectuada en el ensayo consistió en contabilizar la cantidad de plántulas emergidas sobre el total de las semillas sembradas. Se analizaron al azar 5 repeticiones (macetas) por cada genotipo (en este caso, la medición es previa a la imposición del estrés hídrico). Se consideró que una plántula estaba emergida cuando los bordes de los foliolos de la primera hoja trifoliada no se tocaran. Se registraron dos lecturas correspondientes a los 7 y 14 días desde la siembra. Se realizó un ANOVA y las comparaciones se probaron con el test Tukey al 5% utilizando el software estadístico Infostat. Previamente se comprobaron los supuestos de normalidad y homogeneidad de varianzas.

#### 2.3.1.b Severidad de raíz y base del tallo de plántulas

Asimismo, a partir de los 0,7,14 y 21 días posteriores a V1, la severidad de la enfermedad se midió utilizando la escala de severidad propuesta por (Abawi y Pastor Corrales., 1990). La severidad de *M. phaseolina* fue calculada asignándole valores de 0% al 100% dentro de la escala propuesta. Para cada momento de medición, se seleccionó al azar y se midió la severidad de 1 plántula por maceta. Se consideraron las 5 repeticiones (macetas) por cada tratamiento. El valor final para cada tratamiento resultó del promedio de los valores estimados para cada una de las cinco plántulas. El cuadro 2.2 y la figura 2.6 ilustran los diferentes grados de la escala y los ejemplos de cada uno de los grados adoptados en el ensayo, siendo el máximo valor registrado el 7° de la escala. Estadísticamente, como no se pudo comprobar el supuesto de normalidad, se realizó una prueba no paramétrica del tipo Kruskal-Wallis utilizando el software Infostat.

Cuadro 2.2 – Descripción de la escala propuesta por Abawi y Pastor Corrales, 1990

INFECCIONES EN MACROPHOMINA (Abawi y Pastor Corrales, 1990)

(1) No hay síntomas visibles

(3) Las lesiones se encuentran limitadas a tejidos cotiledonares y/o hay una pequeña decoloración en la raíz (hasta un 10% de lesiones en raíces e hipocótile), pero sin evidenciar necrosis

(5) Las lesiones han progresado desde los cotiledones hasta 2cm de los tejidos corticales o del tallo y/o hay pequeños daños a nivel radical (hasta 25% de lesiones en raíces e hipocótile), pero los tejidos radicales permanecen firmes.

(7) Las lesiones son extensivas al tallo y las ramas. El follaje exhibe necrosis y clorosis y/o el sistema radical sufre daños considerables (hasta 50% de lesiones en raíces e hipocótile). Se pueden observar fructificaciones del hongo.

(9) La mayoría de los tallos, pecíolos y puntos de crecimiento están infectados y/o hay un avanzado deterioro del sistema radical (aproximado 75% de lesiones en raíces e hipocótile). Una considerable cantidad de picnidios y microesclerocios se han producido.

Nota: Los valores (2; 4; 6; 8 y 10) corresponden a estadios visuales intermedios y quedan a criterio del observador.



Figura 2.6: Valores registrados de la escala de Abawi y Pastor Corrales, 1990, desde el de menor severidad (1) hasta el valor máximo registrado (7).

# 2.3.1.c Tasa de infección de la enfermedad

Para analizar la progresividad temporal de la enfermedad, se midió la tasa de infección (r). Dicha variable representa la pendiente que se obtiene al linealizar los datos de progreso de la enfermedad (en este caso la severidad), mediante transformaciones logarítmicas (Gonzalez-Dominguez, 2023).

Dichas regresiones pueden ser llevadas a cabo a partir de diferentes modelos epidemiológicos (exponencial, monomolecular, Gompertz y logístico), con lo cual la regresión debe realizarse para cada uno de los mismos, con el fin de conocer a qué modelo/s ajusta/n cada uno de los tratamientos. Para que un tratamiento ajuste a determinado modelo, se requiere un  $r^2>75\%$  y un p valor<0,005. Se utilizaron los 4 valores de severidad medidos (día 0, 7, 14 y 21). A partir de los datos de severidad, se realizaron las regresiones lineales. Posterior a la selección del modelo, como no se pudo comprobar el supuesto de normalidad, se realizó una prueba no paramétrica del tipo Kruskal-Wallis utilizando el software estadístico Infostat.

# 2.3.1.d Área bajo de la curva del progreso de la enfermedad (ABCPE)

El ABCPE representa el área bajo la curva del progreso de la enfermedad. En los casos en los cuales se dispone información de dichas variables en diferentes estadíos de la enfermedad en el cultivo, puede resolverse matemáticamente de forma relativamente sencilla. Es calculado como una integración trapezoidal (Figura 2.7) de los valores de la enfermedad (incidencia, y más frecuentemente severidad) y se aplica para ello el modelo de (Campbell y Madden, 1990), de acuerdo con la fórmula:

# ABCPE= $\Sigma$ [(x i+xi+1) / 2] (ti+1-ti). n-1, i=1

donde, n es el número de evaluaciones, x es la proporción de la enfermedad y (ti+1-ti) es el intervalo entre dos evaluaciones consecutivas.

El ABCPE se expresa en unidades de % día, dado que denota la progresividad de la enfermedad. Una de las mayores utilidades del cálculo de ABCPE radica en la posibilidad de efectuar comparaciones entre cultivares. A modo de ejemplo, un cultivar más susceptible dará un ABCPE más alto y este análisis de ABCPE es una manera de justificarlo. Como fue explicado anteriormente, esta estrategia sirve para poder comparar entre diferentes epifitias o dentro de una misma, en diferentes situaciones. El ABCPE es una construcción analítica y matemáticas sobre la variable severidad, con lo cual el diseño corresponde al previamente indicado en el ítem severidad. Se realizó un ANOVA y las comparaciones se probaron con el test Duncan al 5% (Giovanardi et al. 2018) utilizando el software estadístico Infostat. Previamente se comprobaron los supuestos de normalidad y homogeneidad de varianzas.





Figura 2.7: Esquemas explicativos de la metodología del área bajo la curva del progreso de la enfermedad (ABCPE). La figura superior muestra la construcción en triángulos / trapecios en base al % enfermedad (en este caso severidad) y la figura inferior la aplicación matemática de dichas superficies.

#### 2.3.1.e Incidencia en fragmentos de raíz

Para cuantificar la incidencia de la PC las evaluaciones se realizaron cada 7 días desde el estadio fenológico V1 siguiendo la técnica del "análisis de fragmentos de raíz". Dicha metodología utilizada para la medición de incidencia fue descripta por Almeida (2003). Se seleccionaron al azar 4 fragmentos de raíz de cada una de las 5 plantas (repeticiones), totalizando 20 fragmentos por tratamiento. Los fragmentos de raíz fueron desinfectados en alcohol al 50% por 30 segundos, y luego transferidos a hipoclorito de sodio (0.5%) durante 1 minuto y enjuagado en agua estéril. Los cuatro fragmentos de raíz de cada planta fueron dispuestos en cajas de Petri conteniendo medio de cultivo APG. Posteriormente, las mismas se incubaron en oscuridad a 28°C durante siete días. Por último, se trasladaron a placas con portaobjetos en cruz rotulados para analizar la presencia del patógeno (por ejemplo, microesclerocios) en microscopio óptico. La incidencia correspondió al porcentaje de fragmentos infectados por M. phaseolina (%). Si bien puede existir dependencia entre fragmentos de una misma raíz, esta aproximación es útil dado que permite una caracterización de la enfermedad, en la cual se pueden dilucidar comportamientos presumiblementes susceptibles como tolerantes. El porcentaje total permite caracterizar la respuesta global. Estadísticamente, como no se pudo comprobar el supuesto de normalidad, se realizó una prueba no paramétrica del tipo Kruskal-Wallis utilizando el software estadístico Infostat.

#### 2.3.2 Ensayos complementarios

A partir de las mediciones llevadas a cabo en el ensayo principal y descriptas en 2.3.1a-e, se realizaron nuevas mediciones y ensayos. Los mismos se denominan como complementarios, y se llevan a cabo utilizando solamente los 2 genotipos contrastantes obtenidos del ensayo principal (genotipos DM4800 y D86-8286).

#### 2.3.2 a- Severidad en semillas

El experimento se condujo bajo un diseño experimental completamente aleatorizado. Se colocó 1 semilla por placa de Petri. Se realizaron 10 repeticiones (placas) por genotipo. La severidad del daño causado por *M. phaseolina* se evaluó utilizando la escala descripta anteriormente (Manici et al. 1995) y reproducida por (Gil-Langarica et al. 2007). Se realizó un ANOVA y las comparaciones se probaron con el test Tukey al 5% utilizando el sofware estadístico Infostat. Previamente se comprobaron los supuestos de normalidad y homogeneidad de varianzas.

#### 2.3.2 b- Daño foliar (lesión en cotiledón, hoja cotiledonar y trifoliada)

Dado que la PC es una enfermedad que afecta principalmente la raíz y base del tallo, las mediciones primordiales son las citadas anteriormente (mediciones 2.3.1a - 2.3.1e). No obstante, a modo de complemento, se procedió a realizar un análisis foliar de las plántulas de ambos genotipos contrastantes.Como primera medida se colectaron los cotiledones de cada una de las 10 plántulas y las hojas cotiledonares correspondientes a los 2 genotipos a los 10 días desde la siembra. Por otra parte, también se colectaron las primeras hojas trifoliadas a los 20 días desde la siembra. Para los ensayos se utilizó el folíolo central de dichas hojas.

Las unidades experimentales son el cotiledón, una hoja cotiledonar, y el folíolo central, respectivamente. Para todos los casos se utilizaron 10 repeticiones (plántulas). Los ensayos se condujeron bajo un diseño completamente aleatorizado. El procedimiento consistió en colocar cada uno de los tejidos en placas de Petri individuales. Las hojas se colocaron con el envés hacia arriba. Se colocó papel secante estéril y se humedeció con agua destilada estéril. Se procedió a una desinfección en campana de los tejidos foliares de la siguiente forma: lavado en hipoclorito de sodio al 2%, y doble lavado con agua destilada. Como última medida se secaron los tejidos. Para evaluar el efecto del patógeno en los tejidos, se colocaron 2 discos de 0,5cm de micelio en cada una de las hojas, simétricamente distribuidos a ambos lados de la nervadura central. Para el

caso de los cotiledones, dado su menor tamaño, se colocó únicamente un disco en el centro del cotiledón. Se incubaron todas las placas en oscuridad y a 30°C. A los 5 días desde la incubación se procedió a medir la longitud de la lesión (en cm) de cada uno de los tejidos foliares. Se realizó un ANOVA y las comparaciones se probaron con el test Tukey al 5% utilizando el software estadístico Infostat. Previamente se comprobaron los supuestos de normalidad y homogeneidad de varianzas.

#### 2.3.2-c Establecimiento de plántulas (longitud y peso fresco de plántula)

Se midió la longitud y el peso fresco de plántula a través de un ensayo llevado a cabo en condiciones registradas utilizando bandejas plásticas. Las mismas fueron llenadas con sustrato GrowMix Multipro de Terrafertil. Se sembraron semillas de cada uno de los genotipos evaluados (DM4800 y D86-8286) en las bandejas. Se aplicaron los tratamientos detallados en la cuadro 2.3:

Tratamiento llevado a cabo	Sigla del tratamiento
Inoculación al sustrato con el patógeno en el genotipo DM4800	DM-MP
Control del genotipo DM4800 (sin inocular)	DM-C
Inoculación al sustrato con el patógeno en el genotipo D86	DT-MP
Control del genotipo D86 (sin inocular)	DT-C

Cuadro 2.3- Cuadro resumen de los tratamientos llevados a cabo

Se regaron todas las bandejas, procurando mantener una buena condición hídrica a todo momento. En los tratamientos con inoculación, se colocó en el sustrato una placa de Petri cubierta con micelio de *M. phaseolina* previamente crecido en cámara de cultivo durante 7 días. Posteriormente, se cubrió con una fina capa de sustrato y se sembraron las semillas. A diferencia del ensayo general, aquí el patógeno estaba prácticamente muy cercano o en íntimo contacto con las semillas a germinar. Para los controles, únicamente se procedió a la siembra de las semillas, dado que son sin inoculación. Las bandejas fueron colocadas debajo de iluminación artificial permanente (24hs de luz) a temperatura ambiente (en torno a 20°C). Se realizaron 4 repeticiones por tratamiento, siendo la unidad experimental una semilla (plántula) por bandeja. Los ensayos se condujeron bajo un diseño completamente aleatorizado. A los 7 días desde la siembra, se procedió a evaluar la longitud (cm) y el peso fresco de las plántulas (g). Para evaluar el peso se extrajeron las plántulas del sustrato y se lavaron las raíces para quitar el exceso del mismo. Por último, se secaron y se pesaron en la balanza. Se realizó un ANOVA y las comparaciones se probaron con el test Tukey al 5% utilizando el software estadístico Infostat. Previamente se comprobaron los supuestos de normalidad y homogeneidad de varianzas.

# 2.4 Resultados

#### 2.4. 1Ensayo de evaluación de la enfermedad

# 2.4.1.a Emergencia de plántulas

Es posible observar que inicialmente (es decir a los 7 días desde la siembra), hubo diferencias en la emergencia. Por ejemplo, el genotipo GR (Toyoshirome) evidenció una menor emergencia comparada a los otros 3 genotipos. No obstante, en la segunda medición correspondiente a los 14 días desde la siembra, no hubo diferencia en la emergencia para ninguno de los tratamientos. Es decir, que, si bien hubo una distinta velocidad de emergencia, el porcentaje final de todos los tratamientos no evidenció diferencias significativas según test de comparaciones múltiples Tukey (valor p > 0,05). (Figura 2.8).



Figura 2.8– Emergencia medida a los 7 y 14 días desde la siembra, para todos los tratamientos llevados a cabo. H+ y H- indican la presencia o ausencia de inóculo. Las barras indican el desvío estándar o error. No hubo diferencia significativa a los 14 días desde la siembra. La imagen representa la vista general de las semillas en plena emergencia.

#### 2.4.1.b Severidad de raíz y base del tallo de plántulas

Como se puede apreciar en la Figura 2.9 la misma resultó significativamente diferente para los tratamientos evaluados.

Se pueden observar como tendencia el incremento de la severidad en la sucesión temporal de mediciones. Para el momento 0, correspondiente al inicio del tratamiento de sequía, la severidad fue nula para todos los tratamientos. Para los momentos 7 y 14 días ya se registró severidad, siendo las diferencias no estadísticamente diferentes, pero si lo fueron a los 21 días de medición, con lo cual se analizará particularmente este momento de medición, conocida como severidad final.



Figura 2.9: Severidad registrada según la escala Abawi y Pastor Corrales (1990), a los 0, 7, 14 y 21 días desde V1, momento en el cual se inició el estrés hídrico. Las letras indican diferencias significativas según prueba no paramétrica de Kruskal-Wallis (en este caso, correspondiente a los 21 días). H y S (+) indican la presencia de inóculo y/o sequía. Líneas punteadas indican tratamientos de sequía y continuas, situación de capacidad de campo.

El factor contenido hídrico es muy importante a la hora de analizar los resultados, dado que la sequía favorece el avance de *M. phaseolina* (Cuadro 2.4). El genotipo D86 resultó ser el que evidenció una menor severidad de la enfermedad, tanto en presencia de sequía como en capacidad de campo (S=38% y S=36% respectivamente). Esto indica que el genotipo tendría un buen comportamiento frente a *Macrophomina*. En cuanto al genotipo GR (Toyoshirome), el mismo se comportó similar a D86 únicamente en situación de capacidad de campo (S=42%), mientras que el mismo genotipo en condiciones de estrés hídrico empeoró su desempeño, arrojando valores de severidad mayores (S=52%), siendo un comportamiento regular en la situación más favorable para el avance del patógeno (estrés hídrico). El genotipo Williams 82 tuvo comportamientos disímiles en condiciones hídricas contrastantes. Mientras que para la situación de capacidad de campo logró un aceptable desempeño (S=48%), para la situación de estrés hídrico tuvo valores de severidad altos (S=62%), lo que nos indica que el genotipo es susceptible a *Macrophomina*. No obstante, el genotipo de peor desempeño fue DM4800, dado que su desempeño fue malo tanto en situaciones de capacidad de campo como de estrés hídrico (S= 58% para ambos casos), no habiendo diferencias con el comportamiento de Williams 82.

	Severidad			
	Sevenaud			
	7d	14d	21d	
Williams H+ S+	38%	50%	62%	
Williams H+ S-	30%	40%	48%	
GR H+ S+	34%	40%	52%	
GR H+ S-	28%	38%	42%	
D86 H+ S+	30%	34%	38%	
D86 H+ S-	24%	30%	36%	
DM 4800 H+ S+	40%	52%	58%	
DM 4800 H+ S-	40%	50%	58%	

Cuadro 2.4– Severidades promedio registradas a los 7,14 y 21 días desde V1 para todos los genotipos, tanto en capacidad de campo (S-) como en sequía (S+).

#### 2.4.1. c -Tasa de infección de la enfermedad

Se probaron todos los tratamientos para cada uno de los modelos epidemiológicos existentes (Cuadro 2.5) con el objetivo de obtener el modelo de mejor ajuste.

Cuadro 2.5- Valores de  $R^2$  y valor p, para cada uno de los modelos epidemiológicos analizados (Exponencial, monomolecular, Gompertz y logístico), para cada uno de los tratamientos llevados a cabo.

	Expo	onencial	Monomolecular		Gompertz		Logístico	
Tratamiento	R <sup>2</sup>	Valor p	$R^2$	Valor p	$R^2$	Valor p	R <sup>2</sup>	Valor p
Williams H+ S+	0,85	0,18	0,94	0,03	0,96	0,01	0,85	0,07
Williams H+ S-	0,66	0,19	0,97	0,01	0,96	0,01	0,84	0,08
GR H+ S+	0,83	0,08	0,92	0,04	0,94	0,03	0,83	0,08
GR H+ S-	0,65	0,19	0,99	0,0003	0,95	0,02	0,82	0,09
D86 H+ S+	0,63	0,2	0,97	0,01	0,89	0,06	0,77	0,12
D86 H+ S-	0,66	0,19	0,97	0,01	0,94	0,02	0,82	0,09
DM 4800 H+ S+	0,65	0,19	0,99	0,0034	0,96	0,02	0,83	0,09
DM 4800 H+ S-	0,65	0,19	0,97	0,01	0,95	0,02	0,83	0,09

Todos los tratamientos ajustan al modelo monomolecular, con lo cual se ha seleccionado dicho modelo para el análisis de la variable tasa de infección. Una vez que se ha tomado esto en consideración, es posible realizar una prueba de comparaciones múltiples para evaluar diferencias entre tratamientos. En la siguiente Figura 2.10 se muestra la tasa de infección (severidad/día) para cada uno de los tratamientos:



Figura 2.10: Tasa de infección (r) expresado en severidad/día para cada uno de los tratamientos llevados a cabo. Las letras denotan diferencias significativas según prueba no paramétrica de Kruskal – Wallis.

Tratamiento	Tasa de infección (severidad/día)
Williams H+ S+	2,95
Williams H+ S-	2,28
GR H+ S+	2,48
GR H+ S-	2
D86 H+ S+	1,81
D86 H+ S-	1,71
DM 4800 H+ S+	2,76
DM 4800 H+ S-	2,76

*Cuadro 2.6: Tasa de infección (severidad/día) correspondiente a cada uno de los tratamientos llevados a cabo.* 

Como se puede apreciar en el cuadro 2.6, el genotipo D86 tanto en sequía como a capacidad de campo mostró ser el de menor tasa de infección (1,71 y 1,81 severidad/día, respectivamente). Asimismo, no hubo diferencias significativas con el genotipo GR en condición de capacidad de campo (2 severidad/día). Todos los demás tratamientos resultaron tener tasas estadísticamente mayores. En rangos intermedios se encontró tanto Williams 82 en capacidad de campo (2,28 severidad/día) como GR en condición de sequía (2,48 severidad/día). El genotipo DM4800 resultó tener una mal desempeño tanto en sequía como en capacidad de campo (2,76 severidad/día para ambos casos), no habiendo diferencias con GR y Williams 82 en sequía. El de mayor tasa fue el genotipo Williams 82 en sequía, alcanzando casi los 3 severidad/día (2,95 severidad/día), no habiendo diferencias con DM4800 en ambas condiciones.

# 2.4.1.d - Área bajo de la curva del progreso de la enfermedad (ABCPE)

Tomando como punto de partida la Figura 2.7 (Progresión de la severidad de la enfermedad), es posible calcular el valor total de ABCPE para cada uno de los tratamientos, a partir del cálculo del área de cada una de las secciones. Los valores de cada sección y el valor del ABCPE total están proporcionados a continuación en la cuadro 2.7:

Cuadro 2.7: Área bajo la curva del progreso de la enfermedad (ABCPE) calculados a partir de los diferentes momentos de medición. ABCPE (0-1) denota el avance de la enfermedad de VI hasta la primera medición (7d), ABCPE (1-2) desde la primera medición (7d) hasta la segunda medición (14d) y ABCPE (2-3) desde la segunda medición (14d) hasta la tercera medición (21d). La columna final indica la sumatoria total. Las letras indican diferencias significativas para la sumatoria de ABCPE.

Tratamiento	ABCPE (0-1)	ABCPE (1-2)	ABCPE (2-3)	∑ABCPE	
Williams H+ S+	133	308	392	833	В
Williams H+ S-	105	245	308	658	AB
GR H+ S+	119	259	322	700	AB
GR H+ S-	98	231	280	609	AB
D86 H+ S+	105	224	252	581	Α
D86 H+ S-	84	189	231	504	Α
DM 4800 H+ S+	140	322	385	847	В
DM 4800 H+ S-	140	315	378	833	В



Figura 2.11: ABCPE (área bajo la curva del progreso de la enfermedad) para cada uno de los tratamientos llevados a cabo. Las letras indican diferencias significativas según test Duncan (p < 0.05), acorde a lo descripto en (Giovanardi et al. 2018).

El menor valor registrado de ABCPE corresponde a D86 en capacidad de campo (ABCPEcc=504). El mismo genotipo en sequía mostró un valor de ABCPE numéricamente superior (ABCPEs=581), no habiendo diferencias significativas entre sí (Figura 2.11). El genotipo GR (Toyoshirome), tampoco mostró diferencias significativas con el genotipo D86, aunque los valores de ABCPE tanto en capacidad de campo como en sequía fueron numéricamente mayores (ABCPEcc=609 y ABCPEs=700). El genotipo Williams 82 mostró diferencias según se tratara de la situación a capacidad de campo o en sequía (ABCPEcc=658 y ABCPEs=833). El genotipo DM 4800 fue el de peor comportamiento tanto en capacidad de campo como en sequía (ABCPEcc=833 y ABCPEs=847), junto con el genotipo Williams 82 en sequía, mencionado anteriormente.

## 2.4.1.e - Incidencia en fragmentos raíz

La incidencia (%) medida en fragmentos de raíz mostró diferencias significativas entre tratamientos en las 3 mediciones, siendo las más notorias en la medición correspondiente a los 21 días desde V1 (última medición). Los resultados y valores se observan en la figura 2.12 y el cuadro 2.8.



Figura 2.12: Incidencia en fragmentos de raíces a los 7, 14 y 21 días desde V1. Las letras indican diferencias significativas según la prueba no parámetrica de Kruskal-Wallis. H y S (+) indican la presencia de inóculo y/o sequía. Líneas punteadas indican tratamientos de sequía y continuas, situación de capacidad de campo.

	Incidencia (%)			
Tratamientos	7 días	14 días	21 días	
Williams H+ S+	30%	35%	50%	
Williams H+ S-	35%	40%	40%	
GR H+ S+	15%	25%	35%	
GR H+ S-	10%	15%	15%	
D86 H+ S+	5%	15%	15%	
D86 H+ S-	5%	5%	5%	
DM 4800 H+ S+	40%	45%	65%	
DM 4800 H+ S-	40%	40%	40%	

Cuadro 2.8: Incidencias promedio registradas en los fragmentos de raíz a los 7,14 y 21 días desde V1, tanto en capacidad de campo (S-) como en sequía (S+).

La incidencia mayor fue para DM4800 en la situación de sequía (Is=65%). El desempeño del mismo genotipo en situación hídrica favorable fue superior (Icc=40%), aunque estadísticamente no fueron diferentes, y ambas situadas en mal desempeño dentro del ranking. Analizando la sucesión de mediciones, es interesante destacar que en ambas situaciones la incidencia inicial (medida a los 7 días de V1) fue alta (I=40% para ambos casos), pero se evidenciaron diferencias con la profundización del déficit hídrico (I=65% en la situación de sequía versus 40% en la situación control). Esta descripción es concluyente dado que implicaría un mal comportamiento del genotipo per se. Por otro lado, el genotipo Williams 82 también evidenció un comportamiento disímil según la condición hídrica, a pesar de que estadísticamente no hubo diferencias entre dichos tratamientos. En la situación regada siempre se manifestó una incidencia superior numéricamente desde el inicio (Icc=35% vs Is=30%) y en la medición fuertemente en la situación de sequía, sobrepasando el registro observado para la situación de capacidad de campo (Is=50% vs Icc=40%), siguiendo la misma lógica de mayores niveles de enfermedad en la situación de estrés hídrico.

El comportamiento de GR (Toyoshirome) fue regular, coincidente con las mediciones de severidad y ABCPE. No hubo diferencias significativas en la situación regada y control en la primera medición (Is=15% e Icc= 10%, respectivamente). Esta tendencia se mantuvo e incluso se pronunció en la última medición correspondiente a los 21 días desde V1, más que duplicando la situación de sequía a la de control (Is=35% vs Icc=15%), aunque estadísticamente no implicó una diferencia significativa. Si bien la incidencia en la situación control es baja, esto no representa un buen comportamiento dado que la incidencia en la situación estresante fue más del doble. El genotipo D86 siempre resultó ser el de mejor desempeño (tanto en severidad, ABCPE como incidencia) lo que implicaría que es un genotipo tolerante a la enfermedad. Los valores iniciales de incidencia fueron muy bajos (I=5% en ambos casos). Esta baja incidencia se mantuvo en la situación control (continuó siendo un 5% en las otras 2 mediciones, a los 14 y 21 días desde V1), y se incrementó en la situación de sequía (Is=15%), pero aun así siendo el de mejor comportamiento en condición estresante comparado con el resto de los genotipos, en especial con los genotipos de mal comportamiento (DM4800: Is=65% y Williams: Is= 50%). La vista general del procedimiento se puede apreciar en la Figura 2.13.



Figura 2.13: (a) Vista de plántulas de un tratamiento. (b) Placa de Petri conteniendo 4 fragmentos de cada repetición (5 plántulas). (c) Fragmentos de raíz con presencia de microesclerocios en vista a través de la lupa, considerados positivos para el cálculo de incidencia.

#### 2.4.2 Ensayos complementarios

#### 2.4.2 a- Severidad en semillas

El genotipo de buen comportamiento (D86) tuvo una severidad sensiblemente menor (S=2,2) en comparación al genotipo susceptible (DM4800), cuya S=3,5 (Figura 2.14). Esto refuerza lo observado en la severidad de raíz y base del tallo (según la escala Abawi y Pastor Corrales 1990).



Figura 2.14: Severidad según la escala Manici (1995) en semillas. Las letras indican diferencias significativas según test Tukey (valor p<0,05). Foto a la derecha del gráfico ejemplifica placa de Petri con micelio crecido de M.phaseolina. Las imágenes inferiores corresponden a un ejemplo de semilla en placa, y a los diferentes estadios de la escala de Manici (0 a 5) (0= semilla sana, 1= semilla con cierto grado de decoloración, 2= tegumento de la semilla invadido por micelio y microesclerocios, 3= tegumento libre pero la semilla está infectada, 4= tegumento y semillas infectados y 5= la semilla no germina y está completamente invadida por el hongo).

#### 2.4.2.b- Daño foliar (lesión en cotiledón, hoja cotiledonar y trifoliada)

Los resultaron no arrojaron diferencias a nivel de cotiledón. Esto puede deberse a que es el tejido de menor tamaño (Figura 2.15), y por ende la capacidad de expansión de la enfermedad se encuentra más limitada. Consecuentemente, la lesión también se acota para ambos genotipos. Con respecto al primer par de hojas (simples o cotiledonares), ya se observaron diferencias significativas entre genotipos (Figura 2.16). La longitud de la lesión es sensiblemente menor (1,96cm) en el genotipo de mejor comportamiento (D86), con respecto al genotipo susceptible (2,89cm), correspondiente a DM4800. El incremento de la lesión en el genotipo susceptible es de un 47%. La misma tendencia se puede observar en las hojas trifoliadas. Incluso tienden profundizarse los comportamientos observados en las hojas cotiledonares. La brecha aumenta a un 54%. La longitud de la lesión es de 3,22cm en el genotipo DM4800, mientras que en D86 es de 2,08cm. La hipótesis principal de la profundización de las tendencias radica en la mayor área foliar y la nutrición disponible para la infección del patógeno en hojas trifoliadas.



*Figura 2.15: Vista de las placas incubadas a modo de ejemplo (cotiledón, hoja cotiledonar y folíolo central de la primera hoja trifoliada).* 



Figura 2.16: Longitud de la lesión (cm) para distintos tejidos foliares (cotiledón, hoja cotiledonar o simple y hoja trifoliada). Las letras indican diferencias significativas según test de comparaciones Tukey (valor p < 0.05).

# 2.4.2.c- Establecimiento de plántulas (longitud y peso fresco de plántula)

Los resultados fueron graficados en diagramas de caja y bigotes (Figura 2.17 y 2.20). Con respecto al primer gráfico (Figura 2.17), se observa que el genotipo DM evidenció una mayor longitud de plántula en la situación control con respecto a la situación inoculada (DM-Mp). Es decir, que la infección por parte del patógeno repercutió en el crecimiento de la plántula a través de una menor longitud. Mientras que la longitud en DM-c fue del 13,38cm, para el genotipo en DM-Mp la misma resultó ser de 9,75cm. Es decir, se evidenció una caída del 27,18% de la longitud de plántula.

Cabe destacar la presencia de un valor atípico en el caso del genotipo DM-Mp cuyo desempeño fue muy superior a la media del tratamiento, con lo cual se infiere que la longitud de plántula podría haber sido incluso mucho menor aún, en el caso de desestimar dicho dato. Como

contracara, en el genotipo de mejor comportamiento (DT), no se observaron diferencias significativas en DT-Mp (15cm) con respecto a DT-c (16cm). Es decir que no se observó un decrecimiento en la longitud de plántula en presencia del patógeno en términos de diferencias significativas (menor al 10% numéricamente). La vista general de los tratamientos se puede observar en las Figuras 2.18 y 2.19.



Figura 2.17: Box-plots de longitud de plántula (cm) para ambos genotipos tanto en control (c) como con presencia del patógeno (Mp), según test de Tukey (valor p < 0,05). Hubo diferencias significativas para el caso de longitud de plántula en el genotipo susceptible DM.



Figura 2.18 – La imagen superior hace referencia a DM-C (control), mientras que la inferior hace referencia a DM-MP (inoculado). Imagen tomada en el laboratorio de Bioquímica de la FAUBA, donde se realizó el ensayo in vitro con luz artificial.



Figura 2.19 – La imagen superior hace referencia a DT-C (control), mientras que la inferior hace referencia a DT-MP (inoculado). Imagen tomada en el laboratorio de Bioquímica de la FAUBA, donde se realizó el ensayo in vitro con luz artificial.

En cuanto a la variable peso fresco total (g) de las plántulas (Figura 2.20), no se observaron diferencias significativas para ninguno de los genotipos. Numéricamente los pesos frescos totales fueron similares al control, e incluso levemente superiores.



Figura 2.20: Box-plots de peso fresco (g) para ambos genotipos tanto en control (c) como con presencia del patógeno (Mp), según test de Tukey (valor p < 0.05).

En el caso del genotipo DM, sería correcto señalar que mientras la longitud de plántula disminuyó, el peso fresco total se mantuvo constante o incluso ligeramente superior, aunque no estadísticamente.



Figura 2.21 – Mediciones de peso fresco y aspecto general de las plántulas. Las imágenes superiores corresponden al genotipo DM (control –a- e inoculado-b-) y las imágenes inferiores al genotipo DT (control –c- e inoculado-d-).

No obstante, es indispensable destacar nuevamente en este punto que la presencia del valor atípico podría estar llevando a postular la hipótesis nombrada anteriormente, cuando en realidad la caída en la longitud de plántula hubiera sido mayor, y con ello, un deterioro del peso fresco (g).

En el caso del genotipo tolerante DT, se evidenció claramente su condición de genotipo de buen comportamiento dado que no sólo la disminución en el crecimiento en la longitud de la plántula no fue significativa, sino que el peso fresco total ha sido similar a la situación control (sin diferencias significativas). Es decir, que frente a la presencia del patógeno no sólo pudo mantener el crecimiento aéreo (vía longitud de plántula) sino que además hubo un sostenimiento del peso fresco (y presumiblemente, del aporte radical, siendo la longitud de plántula similar).

# 2.5 Discusión

La opción más útil y ambientalmente amigable, además de económicamente viable (Deshmukh, 2021) para combatir la PC es desarrollar cultivares de soja que sean innatamente resistentes o tolerantes a la enfermedad (Coser, 2017). Se han realizado múltiples estudios para identificar cultivares de soja tolerantes con éxito variable. Dichos esfuerzos se han observado en otros cultivos de interés agrícola tales como frutilla (Pickel et al. 2020), poroto (Kaur et al. 2023) y maní (Badana et al. 2021). Los cultivares actualmente tolerantes exhiben una respuesta cuantitativa de múltiples genes, que se caracteriza por la reducción de la severidad de la enfermedad al limitar la colonización por el patógeno. El método más utilizado para evaluar la resistencia en los cultivares de prueba ha sido cuantificar la cantidad de *M. phaseolina* presente en genotipos de soja en ensayos a campo. Esto incluye estudios realizados por Smith y Carvil, (1997), Smith y Wyllie (1999), (Mengistu et al. 2007) y (Upadhyay et al. 2022). Asimismo, la evaluación a nivel fenotipíca y el desarrollo de un modelo, resulta de vital importancia para este tipo de adversidades por diversas razones. No solo permite eventualmente la detección temprana de enfermedades sino que también es una herramienta eficaz para la caracterización de genotipos

de mayor tolerancia frente a dichos estreses.

En experimentos previos realizados con la variedad de soja DM4800 también se ha observado comportamiento susceptible frente a la enfermedad. Este genotipo se utilizó como referencia de susceptibilidad en diferentes ensayos, como los realizados por (Scandiani et al. 2012) para el análisis de agresividad de inóculos.

Si bien no existen abundantes referencias bibliográficas respecto al genotipo D86, se ha reportado su aporte a la tolerancia, junto con otros genotipos conocidos como DT 974290. Dicho genotipo se suele utilizar como control de tolerancia frente a la enfermedad. En ese sentido, las mediciones llevadas a cabo apuntan a corroborar dicho comportamiento, tanto en el genotipo susceptible como el tolerante.

En la última medición ejecutada (correspondiente a los 21 días desde V1) se observa, como en la medición intermedia (correspondiente a los 14 días desde V1), un incremento de la enfermedad por ambas vías, tanto severidad como incidencia. Es interesante destacar que en este momento se observa con mayor claridad, como el genotipo intermedio (GR) tiende a asemejarse a los desempeños de los genotipos de mal comportamiento (DM4800 y Williams 82). Esta descripción es particularmente evidente en el tratamiento de GR en la situación de sequía: valores de severidad de 52% e incidencia del 35%, mientras que el rango de los genotipos susceptibles oscila entre el 50 y el 65%. Otro aspecto para destacar es la total divergencia entre genotipos contrastantes: estadísticamente hay diferencias significativas en ambos tratamientos (S+ y S-), para ambas mediciones (severidad e incidencia). Este análisis refuerza lo ya observado para ABCPE y tasa de infección. El desempeño del genotipo tolerante D86 en condiciones de sequía fue sensiblemente superior (menor nivel de ABCPEs=581 y tasa infección=1,81), comparado al genotipo susceptible DM4800 (ABCPEs= 847 y tasa infección= 2,76). Este análisis es de relevancia dado que la enfermedad suele ocurrir – o profundizar su avance – en condiciones de estrés hídrico, como las descriptas.

Los ensayos complementarios en bandejas proveyeron más información respecto al contraste de los genotipos y reforzaron lo observado previamente. Se puede hipotetizar que, si la longitud de plántula fue menor y el peso fresco total similar en el genotipo susceptible DM, ha habido un incremento en la proporción aportada por la biomasa radical en la situación tratada con respecto al control. En ese sentido, el menor crecimiento aéreo evidenciado en el crecimiento de la plántula es parcialmente compensada por un crecimiento radical más robusto. Por las imágenes provistas (Figura 2.21), el mayor aporte estaría dado por una mayor profundización de la raíz principal. Es decir que la plántula prioriza el anclaje frente al avance de la enfermedad, en claro detrimento del crecimiento de la plántula. Este atributo sería indeseable pensando en la futura productividad de la planta. Asimismo, caso se desestimara la presencia del valor atípico, el promedio de la longitud de plántula para DM-Mp sería menor al observado (8,17cm en lugar del 9,75cm actual). Es decir, que la caída en el crecimiento en longitud de plántula se profundizaría aún más. En cuanto al peso fresco, se observaron similitudes para tratado y control. Sin embargo, si se desestimara la presencia del valor atípico, se evidenciarían diferencias, provocando disminuciones en el peso fresco para DM-Mp (0,86g en lugar de 0,98g actual). Como contracara, las variables morfológicas de longitud y peso fresco confirmaron el buen desempeño del genotipo DT.

#### **2.6** Conclusiones

Como primera medida, se pudo corroborar que la técnica de aplicación directa de micelio crecido en Placa de Petri al sustrato fue efectiva, logrando una infección adecuada, es decir, que se reprodujo el patosistema de interés, permitiendo la evaluación de los comportamientos diferentes y la medición de las variables propuestas. Hubo interacción entre genotipo, estado hídrico e inóculo. Es decir que las respuestas observadas, tanto en ABCPE, severidad como tasa de infección se debieron a la diferente combinación de los factores nombrados anteriormente. Es importante destacar, a partir de los cuadros de información procesada, que al genotipo DM4800
tiene un mal comportamiento tanto en sequía como en regado, mientras que Williams 82 tiene un mal comportamiento en sequía y regular en capacidad de campo, comparado al genotipo DM4800.

Se pudieron dilucidar comportamientos diferentes entre los genotipos seleccionados. El genotipo superior o de mejor comportamiento fue el D86, mientras que GR (Toyoshirome) tuvo un comportamiento regular o intermedio en general. Tanto Williams 82 como DM4800 tuvieron desempeños pobres, siendo este último genotipo el de peor comportamiento.

A partir de este ensayo preliminar ya se encuentran seleccionados y caracterizados los genotipos que se utilizaron en el próximo ensayo definitivo en el cual se extraerán muestras de los tejidos vegetales para proceder a la realización del RNA-seq (análisis transcriptómico).

# Capítulo 3. Perfil transcriptómico de la tolerancia de la soja al ataque del patógeno

# 3.1 Introducción

El transcriptoma consiste en el conjunto completo de transcriptos de una célula, e incluye al ARN mensajero, ARN de transferencia, ARN ribosomal y otros ARNs no codificantes (Jacquier, 2009; Srivastava et al. 2018). La cantidad de cada tipo de ARN varía entre tejidos y condiciones fisiológicas, lo que torna esencial su estudio para poder entender funcionalmente al genoma (considerado como la secuencia de nucleótidos que constituye el ADN de un individuo o especie). Asimismo, de ese modo conocer, por ejemplo, cómo una planta responde frente a un agente externo como puede ser el ataque de un patógeno.

Se conoce como RNA-seq a una tecnología que se encuentra dentro de las nuevas tecnologías de secuenciación masiva, denominadas en inglés Next Generation Sequencing (NGS), que permiten obtener secuencias de diversas longitudes, acorde a la tecnología utilizada (por ejemplo, Illumina, ion torrent, etc). Las tecnologías NGS convierten el ARN en millones de lecturas cortas a través de plataformas modernas de secuenciación, basadas en "secuenciación por síntesis", que utilizan una ADN polimerasa (ej.: Roche 454, Illumina, Helicos, Pacific Biosciences) o una ADN ligasa (ej.: Life Technologies SOLiD, Complete Genomics) para llevar a cabo la reacción. Comparado con los microarreglos que se basan en información previa, el RNA-seq permite obtener casi todos los transcriptos expresados en un tejido y en una determinada situación, y por lo tanto logra detectar nuevos transcriptos. Los experimentos con RNA-seq han demostrado que tanto las secuencias transcriptas como los niveles de expresión pueden variar significativamente tanto con el tipo de célula, como con el estado de desarrollo y según varíe la condición biótica y abiótica (Hartmann et al. 2022). Este aspecto resultó central al

momento de seleccionar la tecnología adecuada, entre las disponibles. Existen diversos trabajos realizados a nivel transcriptómico en el estudio de *M. phaseolina*, en especies tales como sorgo (Bandara, 2019) y sésamo (Dutta, 2020), que permitieron dilucidar los mecanismos que operan en genotipos tolerantes así como identificar SNPs de utilidad.

La secuenciación con plataforma Illumina es fundamental para la ejecución de los objetivos del presente capítulo. El proceso comienza con la preparación de una biblioteca. La muestra de ADN (para secuenciación genómica) o ADNc (para secuenciación de ARN) se corta, generalmente por método físico, enzimático o químico, en fragmentos cortos predeterminados para un tamaño específico, y luego los adaptadores de secuenciación se ligan a ambos extremos de cada fragmento. Los fragmentos se cargan en una celda de flujo. La celda de flujo tiene oligonucleótidos unidos a la superficie de la misma (cebadores), y sus secuencias son complementarias a los adaptadores de tal manera que los fragmentos de ADN o ARN por secuenciar quedan unidos por complementariedad de bases a la celda de flujo.Un primer paso de extensión amplifica los fragmentos. Como consecuencia, se crea un fragmento complementario que queda unido covalentemente a la celda de flujo. Tras un lavado, se remueven los fragmentos originales, que solamente estaban unidos por complementariedad de bases, y quedan sólo los unidos covalentemente. En un segundo paso de extensión, los fragmentos quedan unidos en la celda de flujo por ambos extremos, formando estructura de puentes. Esta extensión se repite varias veces generando agrupamientos o clusters (lugares de la celda de flujo donde se agrupan numerosas secuencias iguales). Tras el agregado de la polimerasa, los cebadores y los nucleótidos marcados fluorescentemente, comienza la secuenciación propiamente dicha. Los cebadores se van a unir por complementariedad de bases a los extremos de los miles de fragmentos que quedaron unidos covalentemente a la celda de flujo, dejando un extremo 3' libre donde la polimerasa va a comenzar la extensión. Cada uno de los nucleótidos que se usan para esta síntesis están marcados con un fluoróforo distinto, que al ser irradiados emiten una

fluorescencia característica como un color distinto. La fluorescencia que emite el agrupamiento en su conjunto (al ser miles de copias de ese fragmento) a medida que se va sintetizando es lo suficientemente fuerte como para que la detecte una cámara. Una vez adicionado el nucleótido fluorescente a la cadena en crecimiento, se debe remover químicamente la señal fluorescente para poder adicionar el nucleótido sucesivo. De esta manera, a medida que se sintetiza el fragmento se puede determinar, según la fluorescencia emitida, la base que se adicionó en la síntesis. Por ello se llama secuenciación por síntesis.

Una vez que sea realiza la secuenciación, se obtiene lo que se conoce como información cruda a partir de la cual se realiza el control de la calidad de la información y el posterior análisis bioinformático. Las lecturas sin procesar del formato fastq se procesan primero. En este paso, se obtienen lecturas limpias al eliminar las lecturas que contienen el adaptador y las lecturas de baja calidad de los datos sin procesar. Al mismo tiempo, el contenido de Q20, Q30 y GC. Se efectuan los cálculos son las lecturas limpias de alta calidad, y los análisis posteriores se basan en ello. Es una herramienta propicia para ejectutar los objetivos propuestos en el presente capítulo.

#### 3.2 Materiales y métodos

#### 3.2.1 Ensayo principal in vivo

El ensayo principal, del cual se obtuvieron las muestras destinadas a la secuenciación, se realizó bajo invernáculo en la Facultad de Agronomía, bajo un diseño experimental completamente aleatorizado, con arreglo factorial, siendo los factores: genotipo (D86 y DM4800), patógeno (presencia y ausencia de inóculo) y contenido hídrico (sequía y CC). Se sembraron 3 semillas de cada uno de dos genotipos de interés en macetas con un sustrato Growmix Multipro de Terrafertil, por quintuplicado. Se raleó a una planta por maceta alcanzado el estado de VC.

La inoculación del patógeno se realizó de la misma forma que en el ensayo preliminar. Las plantas se mantuvieron sin déficit hídrico hasta alcanzar el estado de V1, y a partir de este momento se aplicó el tratamiento diferencial de riego (misma metodología que la utilizada en el ensayo preliminar). Por otra parte, el análisis preliminar permitió dilucidar el momento más propicio para efectuar la cosecha de las plántulas, siendo este tiempo coincidente con la medición de mayor severidad de la enfermedad (en torno a las 3 semanas -21 días- desde V1). Se cosecharon las plántulas de ambos genotipos contrastantes D86 y DM4800, correspondientes a todos los tratamientos, que surgen de la combinación de factor hídrico (presencia o ausencia de sequía) e inóculo (presencia o ausencia), resultando en (H-S-), (H-S+), (H+S-) y (H+S+).

El material recolectado fue guardado en sobres de aluminio individuales y congelado en N líquido inmediatamente. Posteriormente, se guardaron los sobres en freezer a -80°C hasta la extracción del ARN. Las raíces fueron lavadas con agua destilada y secadas con papel absorbente para eliminar los excesos de sustrato. El tejido seleccionado para la secuenciación fueron justamente las raíces, ya que es la zona principal de afectación por parte del patógeno (Babu, 2011; Bressano et al. 2010; Chowdhury y Basu, 2017; Kendig, et al. 2000). se prosiguió a cumplir los pasos necesarios para la realización del RNA-seq, y el análisis de los datos resultantes de dicho estudio.



Figura 3.1 (Genotipo DM4800) y 3.2 (Genotipo D86: Macetas (repeticiones) representativas de tratamientos aplicados. De izquierda a derecha: H+ S+, H- S+, H+ S-, H- S- (control) en el momento previo a la cosecha (21 días desde V1)



Figura 3.3: Macetas (repeticiones) representativas de tratamientos aplicados al genotipo DM 4800. (A) H+ S+, (B) H- S+, (C) H+ S-, (D) H- S- (control) en el momento previo a la cosecha (21 días desde V1)



Figura 3.4: Macetas (repeticiones) representativas de tratamientos aplicados al genotipo D86. (A) H+ S+, (B) H- S+, (C) H+ S-, (D) H- S- (control) en el momento previo a la cosecha (21 días desde V1)

# 3.2.2 RNA-seq

# Extracción del ARN, secuenciación y construcción de la biblioteca

A continuación se detallan todos los materiales y métodos utilizados, desde la extracción del ARN de las plantas cosechadas , la secuenciación propiamente dicha y los análisis de datos llevados a cabo:

El objetivo del estudio RNA-seq es identificar qué genes y rutas metabólicas están *"up- regulated"* o *"down-regulated"* en respuesta a la enfermedad en el genotipo tolerante. Para cumplir con dichas premisas, el procedimiento fue el siguiente:

- Extracción del ARN
- Análisis transcriptómico mediante secuenciación de ARN (RNA-seq)
- Análisis de los resultados de la secuenciación

El siguiente esquema resume la modalidad de trabajo y los pasos a seguir (Figura 3.5):



Figura 3.5: Pasos llevados a cabo para poder analizar los genes up y down regulated: Primero consta de la preparación del ARN, luego la preparación de las bibliotecas de ADNc, la secuenciación de éste y por último propiamente el análisis bioinformático a llevar a cabo

La extracción del ARN se llevó a cabo con el kit Qiagen (RNeasy Plant Mini Kit) y se controló que cumpla con las condiciones de masa, pureza e integridad requeridas para la secuenciación.

Se siguió el protocolo RNeasy Plant Mini Kit (RNeasy mini hand book), introduciendo las modificaciones: (I) Agregado de  $\beta$ -mercaptoetanol al buffer RLT (10µ1  $\beta$ -Me/1ml de buffer) y (II) Modificación de la concentración del buffer RPE, agregando 4 veces el volumen de etanol (96-100%).

En primer lugar, se pesaron 100mg (0,1g) de tejido radical de cada una de las muestras. Las mismas luego se malaxaron en un mortero y pilón, con N líquido, procurando que en ningún momento se descongele la muestra en cuestión. Una vez lograda una consistencia polvorienta, semejante a una pasta, se colocaron las muestras en tubos de 2ml con ayuda de una espátula. Se agregaron 450 µl de buffer RLT en el tubo que contenía el polvo. Se dio un ligero vórtex sobre las muestras.

Una vez introducido el buffer, se incubaron los tubos en estufa a 56°C por 1 a 3 minutos, con el objetivo de facilitar la acción del mismo. Posteriormente, se pasó el lisado al tubo QIAshredder provisto por el fabricante. Se centrifugaron las muestras a máxima velocidad (11.000 rpm) por 2 minutos y se trasvasó el sobrenadante a un tubo nuevo de 2µl. Se cortó la punta del tip para facilitar el trasvasamiento. En este paso fue importante considerar qué volumen de sobrenadante se está trasvasando.

El siguiente paso constó del agregado de etanol al 96%. El volumen fue del 50% en referencia a la cantidad de sobrenadante trasvasado anteriormente. Automáticamente se resuspendió el volumen. Una vez concluido el paso anterior, se trasvasó la totalidad del volumen a la columna MiniSpin. Se centrifugaron las muestras a máxima velocidad por 15 segundos, y se procedió al descarte del flow-through, es decir lo que no queda retenido en la columna, procurando seguir utilizando el mismo tubo de colección que está recubriendo la columna

MiniSpin. El siguiente paso constó de 3 agregados de buffer, siendo el buffer RW1 la primera ocasión (700  $\mu$ l) y el RPE en la 2° y 3° (500 $\mu$ l en cada ocasión). Cabe aclarar que entre cada una de las ocasiones se centrifugó a máxima velocidad (por 15 segundos, salvo en la 3° ocasión por 2 minutos), descartando el flow-through y manteniendo el mismo tubo de colección. Luego se colocó la columna mini Spin en un tubo de 2ml, centrifugando por 1 minuto a máxima velocidad. El objetivo de este paso fue el descarte del exceso de buffer, si lo hubiera. Por último, se produjo la elución. Primero se colocó la columna miniSpin en un tubo de 1,5ml. Se procedió al agregado de 30 $\mu$ l de agua libre de ARNsas, y se realizó la ultima centrifugación a máxima velocidad por 1 minuto. Una vez finalizado, se disupuso del ARN de cada una de las muestras para su cuantificación.

Se cuantificó la concentración de ARN total mediante el fluorómetro Qubit (Invitrogen Corporation, USA), siguiendo las indicaciones del fabricante.

Luego de la cuantificación de ARN para cada una de las muestras, se procedió a la preparación de un gel doble de agarosa de 1,5% (Apéndice – Figura 3.6).

Las muestras fueron enviadas a Sacramento (California) para su secuenciación, cumpliendo ampliamente los requisitos (concentración >20 ng/µl, volumen >50µl). Allí la empresa Novogene realizó un estudio de control de calidad, donde se aseguró la calidad de las muestras: su concentración (µg/µl), volumen (µl), masa (µg) y valor de pureza e integridad a través del RIN, medido con el equipo Agilent Bioanalyzer.

La degradación y contaminación del ARN se controló en geles de agarosa al 1%. La pureza del ARN se verificó utilizando el espectrofotómetro NanoPhotometer® (IMPLEN, CA, EE.UU.). La integridad y cuantificación del ARN se evaluaron utilizando el kit de ensayo ARN Nano 6000 del sistema Bioanalyzer 2100 (Agilent Technologies, CA, EE. UU.).

Para la secuenciación, en primer lugar, se fragmentó químicamente el ARN para luego sintetizar el ADNc doble hebra. Luego se realizó la reparación y adenilación de los extremos.

Por último, se ligó los adaptadores y se procedió con el enriquecimiento de los fragmentos de la biblioteca de acuerdo con el protocolo descripto por en TruSeq® ARN Sample Preparation Guide Illumina. La secuenciación de ARN constó de las siguientes etapas:

- Análisis de calidad del ARN (pureza, integridad y concentración)
- Construcción de la librería y control de su calidad
- Secuenciación propiamente dicha

Los controles de calidad y la construcción de la biblioteca se realizaron a continuación de la secuenciación por Novogene en Sacramento, California. Los pasos de la secuenciación se pueden observar en (Apéndice – Figura 3.7). Se utilizó una cantidad total de 1µg de ARN por muestra como material de entrada para las preparaciones de muestra de ARN (concentración>20ng/µL, volumen> 15 µL y RIN(integridad ARN) >7). Las bibliotecas de secuenciación se generaron usando NEBNext®UltraTMRNA Library Prep Kit para Illumina® (NEB, EEUU). Siguiendo las recomendaciones del fabricante, se agregaron códigos numéricos para atribuir secuencias a cada muestra. Brevemente, el ARNm se purificó del ARN total usando partículas magnéticas unidas a oligo poli-T. Posteriormente, la fragmentación de ARNm se llevó a cabo utilizando cationes divalentes a temperatura elevada en el buffer (NEBNext First Strand Synthesis Reaction buffer (5X)). La primera cadena de ADNc se sintetizó usando cebadores (primers) hexámeros aleatorios y transcriptasa inversa M- MuLV (ARNasa H-).

La síntesis de ADNc de la segunda cadena se realizó posteriormente usando ARNasa H y ADN polimerasa I. A posteriori se ligaron los adaptadores (NEBNext adaptador). Los fragmentos de la biblioteca se purificaron con el sistema AMPure XP (Beckman Coulter, Beverly, EEUU) para seleccionar fragmentos de ADNc de preferentemente 150 ~ 200 pb de longitud. Luego, se utilizaron 3µl de enzima USER (NEB, EEUU) para remover los restos uracilo. Asimismo, dicha enzima fue utilizada junto con el ADNc a una temperatura de 37°C durante 15 minutos (seguidos de 5 minutos a 95°C), previo a la realización de la PCR. Posteriormente, se realizó la PCR utilizando la polimerasa Phusion High-Fidelity ADN, primers universales de PCR e Index (X) Primer. Finalmente, se purificaron los productos de PCR (sistema AMPure XP) y se evaluó la calidad de la biblioteca en el sistema Agilent Bioanalyzer 2100.

El análisis bioinformático constó de diversas etapas, siendo el punto de partida la información cruda (raw data) obtenida a partir de la secuenciación. La figura 3.8 muestra el flujo de trabajo para el análisis de los datos obtenidos de la secuenciación.



Figura 3.8: Esquema de trabajo para análisis bioinformático

# Control de la calidad de la información (data QC)

a) <u>Tasa de error</u>: La tasa de error de la secuenciación y la calidad de la base varían según los secuenciadores, los residuos de reactivos y los diferentes tipos de muestra. Para la tecnología RNA-seq, se puede presentar la distribución de la tasa de error de secuenciación. La tasa de error aumenta con las lecturas de secuenciación. Es común observar este patrón en la plataforma de secuenciación de alto rendimiento Illumina (Erlich Y, Mitra PP et al.2008; Jiang L, Schlesinger F et al, 2011). Las primeras seis bases tienen una tasa de error relativamente alta debido a la unión incompleta de hexámeros aleatorios utilizados en el cebado de la síntesis de ADNc. En general, una tasa de error base única debe ser inferior al 1%.

b) <u>Distribución de contenido de GC</u>: La distribución del contenido de GC es para detectar la separación potencial de AT / GC, que afecta la cuantificación posterior de la expresión génica. En vista de la fragmentación aleatoria y la ley biológica del contenido de G/C - A/T, G-C por un lado, y A-T por el otro, deben ser respectivamente iguales, y el contenido debe ser estable durante todo el proceso de secuenciación para la biblioteca. Se permite una gran variación del error de secuenciación en las primeras 6-7 bases considerando el uso de cebador aleatorio en la construcción de bibliotecas.

c) <u>Filtrado de la información</u>: Las lecturas secuenciadas / lecturas sin formato a menudo contienen lecturas de baja calidad o lecturas con adaptadores, lo que afecta la calidad del análisis posterior. Para evitar esto, fue necesario filtrar las lecturas sin procesar y obtener las lecturas limpias. Los parámetros establecidos consistieron en la eliminación de lecturas que contenían fragmentos de los adaptadores, más del 10% de bases indeterminadas o un Q-score (parámetro de calidad que hace referencia al 50% de las bases de lecturas) bajo.

# Mapeado al genoma de referencia

Se utilizó un algoritmo para mapear secuencias conocido como HISAT2 (Wen, 2017).

Se seleccionó para mapear las lecturas secuenciadas filtradas al genoma de referencia (disponible en Phytozome correspondiente a la versión 4 de la variedad Williams 82). En general, HISAT2 admite genomas de cualquier tamaño, incluidos los de más de 4 mil millones de bases y la mayoría de los parámetros están configurados de forma predeterminada, o mejor precisión que cualquier otro método. El algoritmo HISAT2 se puede dividir en tres partes. En principio, se lee por completo a un solo exón del genoma.

Posteriormente, las lecturas se segmentan y luego se asignan a los exones adyacentes. Por último, las lecturas se segmentan y luego se asignan a tres o más exones. Se calcula la cantidad de lecturas asignadas totales (y su porcentaje de lecturas limpias), incluida la cantidad de lecturas asignadas múltiples (y su porcentaje de lecturas limpias), y la cantidad de lecturas asignadas de forma única (y su porcentaje de lecturas limpias). Las lecturas/fragmentos mapeados totales deben ser mayor al 70% y las lecturas o fragmentos mapeados múltiples no deben ser mayor al 10%.

Las regiones mapeadas se pueden clasificar como exones, intrones o regiones intergénicas. Las lecturas mapeadas de exones deberían ser el tipo de lectura más abundante cuando el genoma de referencia está bien anotado. Las lecturas de intrones pueden derivarse de la contaminación previa al ARNm o la retención de intrones a partir de empalmes alternativos. Las lecturas asignadas a regiones intergénicas se atribuyen principalmente a la anotación débil del genoma de referencia. A partir de dicho análisis se obtuvo la distribución de las lecturas de secuenciación en la región genómica.

#### Cuantificación del nivel de expresión génica

El análisis del nivel de expresión génica es la tarea central en el experimento de RNAseq. El nivel de expresión génica se calcula por el número de lecturas asignadas. La abundancia de transcripción refleja el nivel de expresión génica directamente. En los experimentos de RNAseq, el nivel de expresión génica se estima por la abundancia de lecturas (recuento de secuenciación) que se mapean en genoma o exón. FeatureCounts v1.5.0-p3 se utilizó para contabilizar los números de lecturas mapeadas a cada gen (Liao et al. 2013).

Los recuentos de lectura son proporcionales al nivel de expresión génica, la longitud del gen y la profundidad de secuenciación. La cantidad esperada de fragmentos por kilobase de secuencia de transcripción por millones de pares de bases secuenciados (FPKM, del inglés *Fragments Per Kilobase of transcript per Million mapped reads*) es el método más común para estimar los niveles de expresión génica, que tiene en cuenta los efectos de la profundidad de secuenciación y la longitud del gen en el conteo de fragmentos (Trapnell y Cole. 2010).

# Análisis de correlación

Las réplicas biológicas son necesarias para cualquier experimento biológico, incluida la tecnología RNA-seq. Se realizaron 3 repeticiones, considerando la restricción presupuestaria y la cantidad de tratamientos a evaluar. La correlación de los niveles de expresión génica entre muestras desempeña un papel importante para verificar la fiabilidad y la selección de la muestra, que no solo puede demostrar la repetibilidad del experimento, sino también estimar el análisis diferencial de la expresión génica. Se realizó el análisis de correlación entre las réplicas biológicas, en donde se calculó la correlación de Pearson entre los niveles de expresión de todos los genes. Es necesario altos niveles de correlación ( $r^2$ ) entre muestras de un mismo tratamiento.

# Análisis de expresión diferencial

El recuento de lecturas obtenido del análisis de expresión génica se utilizó para el análisis de expresión diferencial. Al tratarse de muestras con réplicas biológicas, se realizó un análisis de expresión diferencial de dos condiciones / grupos utilizando el paquete DESeq2 R (Anders et al. 2010). Este programa proporciona rutinas estadísticas para determinar la expresión diferencial en datos digitales de expresión génica utilizando un modelo basado en la distribución binomial negativa. Los valores p resultantes se ajustaron utilizando el enfoque de Benjamini y Hochberg

para controlar la Tasa de Descubrimiento Falso. Los genes con un valor p ajustado <0,05 encontrado por DESeq2 se asignaron como expresados diferencialmente.

# Análisis de enriquecimiento

A través del análisis de enriquecimiento de los genes expresados diferenciales, se puede descubrir qué funciones biológicas o vías están significativamente asociadas con los genes expresados diferenciales. Se utilizó el software clusterProfiler (Yu G, 2012) para el análisis de enriquecimiento, que incluye enriquecimiento GO y KEGG. Dicho software compara perfiles funcionales de varias condiciones a mismo nivel (diferentes tratamientos), valiéndose de la admisión de diversas ontologías y rutas, anotaciones genéticas de múltiples especies, incluidas las anotaciones emergentes. La biblioteca clusterProfiler tiene muchas características únicas, incluida una interfaz ordenada que pueden manipular el resultado del enriquecimiento y apoyar directamente la visualización del resultado del enriquecimiento (Wu, 2021).

<u>GO</u> (del inglés *Gene Ontology*) (http://www.geneontology.org/), es un sistema de clasificación bioinformática para unificar la presentación de propiedades genéticas en todas las especies. Incluye tres ramas principales: componente celular, función molecular y proceso biológico. Los términos GO con padj <0.05 son de enriquecimiento significativo.

<u>KEGG</u> (del inglés *Kyoto Enciclopedia of genes and and genomes*) (http://www.kegg.jp/) . Las interacciones de múltiples genes pueden estar involucradas en ciertas funciones biológicas. KEGG es una colección de bases de datos curadas manualmente que contienen recursos de información genómica, de vías biológicas y de enfermedades (Kanehisa, 2008). El análisis de enriquecimiento identifica rutas metabólicas significativamente enriquecidas o rutas de transducción de señales asociadas con genes expresados diferencialmente, en comparación con todo el fondo del genoma. Los términos KEGG con padj <0.05 son enriquecimiento significativo.

A partir de toda la metodología explicada, se identificaron qué genes se inducen como

respuesta a la enfermedad aumentando su expresión, y qué genes disminuyen su expresión, junto con las vías metabólicas asociadas a ellos, en el genotipo tolerante DT con respecto al susceptible DM.

#### 3.3 Resultados

# 3.3.1 Análisis de expresión diferencial entre genotipos contrastantes

Se está analizando el comportamiento de ambos genotipos en presencia de sequía y del patógeno. El aspecto general de las plantas para cada uno de los tratamientos para dicho momento de pre-cosecha se puede observar en las Figuras 3.1, 3.2, 3.3 y 3.4.

Tal como se expresó anteriormente, el recuento de lecturas obtenido del análisis de expresión génica se utiliza para el análisis de expresión diferencial. El diseño experimental contó de ambos genotipos (DM y DT) en los 4 tratamientos descriptos del capítulo II para cada uno (condiciones combinadas de factor hídrico e inóculo). Al tratarse de muestras con repeticiones biológicas (3 repeticiones por tratamiento), se realizó un análisis de expresión diferencial de dos condiciones / grupos utilizando el paquete DESeq2 R (Anders et al. 2010).

A partir del uso de dicho paquete se pudo obtener el listado de todos los genes que se expresan diferencialmente en la comparación mencionada (DT\_HS vs DM\_HS). El análisis permitió dilucidar un total de 686 genes que cambiaron su expresión, siendo 373 los genes que aumentaron su expresión en DT\_HS con respecto a DM\_HS, y 313 disminuyeron su expresión.

Contemplando la totalidad de los genes de dicha lista, se construyó un gráfico de volcán (Figura 3.9), donde se puede ver claramente la distribución de todos los 686 genes (en rojo, aquellos que aumentaron su expresión y en verde los que disminuyeron su expresión). Dicho gráfico muestra en simultáneo la significancia de cada uno de los genes (a través del logaritmo negativo base 10 del valor p ajustado), con la magnitud del cambio de expresión de cada uno de los genes en el genotipo DT con respecto a DM (expresado a través del logaritmo base 2 del fold

change). Cobra relevancia el umbral de significancia, representado con el  $-\log_{10}(\text{padj}) = 1,3$ , que marca los genes de respuesta significativa y los no significativos. Este valor surge de considerar el valor p ajustado de 0,05 que es ampliamente utilizado como valor de significancia en los análisis estadísticos.



Figura 3.9 – Volcanoplot de la expresión diferencial de genes del genotipo tolerante DT, con respecto al genotipo susceptible DM, en condiciones de infección y sequía simultánea.

Analizando la figura 3.9, se puede observar el conjunto de puntos superpuestos en azul, que representan todo el conjunto de genes que no resultaron significativos, y por ende no están incluidos en los 686 genes totales. Asimismo, se puede observar en rojo los 373 genes que aumentaron su expresión en el genotipo tolerante DT con respecto al genotipo DM. Por otra parte, los puntos verdes representan los 313 genes que disminuyeron su expresión en DT con respecto a DM. El listado completo puede observarse en el <u>Apéndice.</u>

El análisis de expresión diferencial cobra una gran relevancia dado que permite filtrar el comportamiento de la totalidad de genes presentes, aportando un panorama más simplificado acerca de qué genes cambian su expresión (aumentando o disminuyendo) en la comparación

analizada. No obstante, el paso posterior corresponde al análisis de enriquecimiento, que permite extraer mayores conclusiones a partir de vincular la expresión de los genes a categorías funcionales y vías metabólicas.

# 3.3.2 Análisis de enriquecimiento (GO/KEGG)

El análisis de enriquecimiento funcional se utiliza para determinar si una clase particular de genes o proteínas está sobre-representada dentro de un conjunto mayor de genes. Es decir, sirve para evaluar si dentro de los genes diferencialmente expresados hay una clase particular que se destaque significativamente, ya sea por su disminución o aumento en la expresión. Se realizó un análisis de 2 bases de datos distintas (GO y KEGG), donde se representan las categorías enriquecidas (sobreexpresadas y subexpresadas):

# Análisis de Enriquecimiento utilizando la herramienta Gene Ontology (GO)

A través de gene ontology (GO), utilizando los diferentes genes sobre (+) y subexpresados (-), se los puede asociar a diferentes categorías GO, agrupados en 3 ramas principales: función molecular (MF), componente celular (CC) y procesos biológico (BP). La siguiente tabla (Cuadro 3.1) resume las categorías GO que resultaron significativas, es decir con un padj <0,05, junto con el número de genes involucrados en dichas categorías:

CATEGORIAS ENRIQUECIDAS EN GENES SOBREEXPRESADOS (DT_HS vs DM_HS)							
Categoría	GO - ID	padj	Cantidad genes				
BP	GO:0006457	Plegado proteínas	4,68E-07	10			
MF	GO:0051082	Unión proteínas no plegadas	5,95E-06	10			
MF	GO:0045735	0.00637	5				

*Cuadro 3.1 – Tabla resumen de cambios de expresión para el enriquecimiento GO* 

CATEGORIAS ENRIQUECIDAS EN GENES SUBEXPRESADOS (DT_HS vs DM_HS)						
Categoría GO - ID Descripción padj Cantidad de gen						
CC	GO:0071944	Periferia celular	0.02794	5		
CC	GO:0005618	Pared celular	0.02794	3		
CC	GO:0048046	Apoplasto	0.02794	3		
MF	GO:0043531	Unión al ADP	7,97E-08	13		

Como se puede apreciar, en el genotipo tolerante DT, considerando la rama de MF, hay una sobre-expresión de un grupo de 10 genes asociados al término GO de unión de proteínas no plegadas. Asimismo, dentro de esta interpretación se suma el grupo de 5 genes representados dentro del la categoría de actividad asociada a reserva de nutrientes. Como contrapartida, se puede ver un grupo de 13 genes sub-expresados asociados a la categoría unión al ADP. Cuando se analiza el CC, se puede vislumbrar únicamente una sub-expresión génica en el genotipo tolerante DT compuesta por pocos genes (3-5), asociados a todas categorías GO muy similares entre si, como lo son los de pared celular, apoplasto y periferia celular. Por último, en la rama BP, podemos observar un grupo de 10 genes sobreexpresados en el genotipo tolerante, asociados a la categoría GO de plegado de proteínas. Los genes resultantes del análisis totales y únicos son 33: los genes de periferia celular, pared celular y apoplasto son los mismos (5), lo mismo ocurre con los genes asociados al plegado de proteínas y la unión de proteínas no plegadas (10), sumado a la actividad de reserva de nutrientes (5) y la unión al ADP (13).

El siguiente gráfico (Figura 3.10) muestran todas las categorías resultantes de los genes sub- expresados en el genotipo tolerante con respecto al susceptible:



DT\_HSvsDM\_HS\_down

Figura 3.10: Categorías de la plataforma GO asociados al conjunto de genes sub- expresados diferencialmente en la comparación DT\_HS vs DM\_HS.

El gráfico tiene muchas variables involucradas. Para cada uno de las categorías GO, se pueden destacar puntos en diferentes posiciones. El tamaño de los puntos indica el conteo de genes, es decir el número de genes involucrados dentro de dicho término. En simultáneo, cada uno de los puntos tiene una gama de colores diferente. Dichos colores representan la significancia de cada una de las categorías. Para que una categoría sea significativa, el padj debe ser < 0,05. Los puntos que se asemejen al color rojizo, en base a la paleta de colores de la leyenda del gráfico, representan categorías significativas.

El gráfico resume los términos ejemplificados en el cuadro 3.1. Se pueden examinar las categorías significativas de periferia celular, apoplasto, unión al ADP, pared celular y otros términos asociados a componente celular, tales como región extracelular. Los tamaños de los puntos correlacionan con el número de genes de cada una de las categorías descriptos

anteriormente en el cuadro 3.1. El siguiente gráfico (Figura 3.11) muestra todas las categorías resultantes de los genes sobre- expresados en el genotipo tolerante con respecto al susceptible:



Figura 3.11: Categorías de la plataforma GO asociados al conjunto de genes sobre- expresados diferencialmente en la comparación DT\_HS vs DM\_HS.

Es posible visualizar las categorías estadísticamente significativas (asociadas a genes sobre-expresados) en la comparación. La unión de proteínas no plegadas se destaca como una de las categorías significativas, representada por un conjunto de 10 genes. Asimismo, el plegado de proteínas también resulta como una de las categorías significativas, también representado por 10 genes. La última categoría significativa es la actividad de reserva de nutrientes, representada por 5 genes.

Como una primera aproximación que se puede realizar a partir del análisis GO, es evidente que hay un conjunto de genes sobreexpresados en el genotipo tolerante DT involucrados en diferentes procesos asociados a proteínas, como se observa en las categorías de unión de proteínas no plegadas y el plegado de proteínas. Es posible que dichos mecanismos contribuyan en mayor o en menor medida a la tolerancia observada en el genotipo DT.

# Análisis de Enriquecimiento utilizando la herramienta Kyoto Encyclopedia of Genes and Genomes (KEGG)

Como se mencionó anteriormente, KEGG (Enciclopedia de Genes y Genomas de Kioto, http://www.kegg.jp/) es una colección de bases de datos curadas manualmente que contienen recursos de información genómica, de vías biológicas y de enfermedades (Kanehisa, 2008). El análisis de enriquecimiento identifica rutas metabólicas significativamente enriquecidas o rutas de transducción de señales asociadas con genes expresados diferencialmente, en comparación con todo el fondo del genoma. La siguiente tabla (Cuadro 3.2) resume las categorías KEGG estadísticamente significativas (es decir, con padj <0.05):

Cuadro 3.2 – Tabla resumen de cambios de expresión para el enriquecimiento KEGG

CATEGORIAS ENRIQUECIDAS GENES SOBREEXPRESADOS (DT_HS vs DM_HS)					
KEGG - IDDescripciónpadjCantidad de g					
gmx04141	Procesamiento proteínas en retículo	2,19E-10	34		
gmx04626	Interacción planta-patógeno	0.041	9		

La siguiente figura 3.12 muestra las categorías de enriquecimiento KEGG de la comparacion DT\_HS vs DM\_HS para el caso de sub-expresión génica:



Figura 3.12: Categorías de la plataforma KEGG asociados al conjunto de genes sub expresados diferencialmente en la comparación DT\_HS vs DM\_HS.

Analizando el cuadro 3.2 y la figura 3.12 se puede interpretar que no hubo categorías KEGG asociadas a disminuciones de expresión en el genotipo tolerante DT con respecto al genotipo DM. Dentro de las 20 categorías de mayor enriquecimiento evidenciados en la figura 3.12, se puede señalar que ninguna resultó significativa (valor padj < 0,05), para el caso de categorías KEGG asociadas a sub-expresión de genes en el genotipo tolerante DT.

La siguiente figura 3.13 muestra las categorías de enriquecimiento KEGG de la comparacion DT\_HS vs DM\_HS para el caso de sobre-expresión génica:



*Figura 3.13: Categorías de la plataforma KEGG asociados al conjunto de genes sobreexpresados diferencialmente en la comparación DT\_HS vs DM\_HS.* 

A partir de la visualización del cuadro 3.2, y figura 3.13, solamente se observan 2 categorías KEGG que resultaron significativas (valor padj < 0,05) en el análisis, asociadas a sobre-expresión génica en el genotipo tolerante DT con respecto al genotipo susceptible DM.

Por un lado, hay una sobre-expresión de un grupo de 34 genes asociados al procesamiento de proteínas a nivel de retículo endoplasmático. Dentro de esta interpretación se suma el grupo de 9 genes representados dentro del término interacción planta – patógeno. Estas vías metabólicas sobre-representadas sugieren mecanismos responsables de la tolerancia observada en el genotipo DT frente a la enfermedad.

## 3.3.3 Selección de genes por tolerancia a la enfermedad

A través de los análisis de enriquecimiento de las bases de datos correspondientes a GO y KEGG se pudieron identificar las vías metabólicas predominantes en cada una de las comparaciones y la cantidad de genes representativos de cada una de dichas vías citadas.

Es necesario a partir de ahora identificar aquellos DEGs que estén presentes en la comparación entre genotipos (DT\_HS vs DM\_HS), los cuales representarán un indicio muy importante del origen del mejor desempeño del genotipo DT frente a la enfermedad. El criterio adoptado para la identificación de genes se explica en la figura 3.14(a).



Figura 3.14 (a): Esquema resumen del criterio adoptado para la selección de genes de interés que contribuyen a la tolerancia de la enfermedad.

Los criterios adoptados se detallan a continuación, de modo secuencial del (1) al (3):

- (1) Identificar los genes sub y sobre-expresados en la comparación DT\_HS vs DM\_HS
- (2) Considerar aquellos genes identificados en el efecto únicamente por sequía del genotipo tolerante una vez que ya está infectado (DT\_HS vs DT\_H), tanto aquellos sub como sobre-expresados. Aquellos genes que sean identificados podrían atribuirse al efecto sequía, con lo cual los mismos no estarían asociados a la tolerancia de la enfermedad por parte del genotipo DT (Figura 3.14)
- (3) No obstante, se corrobora la presencia de los genes identificados en el ítem 2, en la situación de efecto patógeno en el genotipo tolerante una vez que ya hay sequía (DT\_HS vs DT\_S). Aquellos

genes considerados en efecto sequía (2) que se encuentren presentes en efecto patógeno (3) no han de descartarse producto que los mismos estarían asociados a la tolerancia de la enfermedad. Por el contrario aquellos genes identificados en (2) que no estén presentes en (3) se descartaran producto de que su expresión es respuesta únicamente al estrés abiótico de la sequía (Figura 3.14).

#### Selección de genes identificados en el análisis (GO)

Se cumplió con los pasos descriptos anteriormente. Se analizaron los 33 genes presentes en el análisis GO (destacados anteriormente) correspondientes a la comparación de interés (DT\_HS vs DM\_HS). De los 33 genes, se observó que 3 de los mismos, correspondientes a CC, e involucrados en la periferia celular y el apoplasto, también disminuían su expresión cuando se observaba su comportamiento en la comparación donde se analizaba el efecto sequía en dicho genotipo (DT\_HS vs DT\_H).

Dado que dichos genes no estaban presentes en la comparación que evalúa el efecto patógeno (DT\_HS vs DT\_S), se procede al descarte de los mismos (destacados en color amarillo), producto de que no contribuirían a la tolerancia de la enfermedad. Dichos genes están asociados a actividad hidrolasa y trans-glicolasa. La tabla (Cuadro 3.3) cita los 33 genes resultantes del enriquecimiento GO, resaltando los 3 genes que se descartan.

Cuadro 3.3 – Genes diferencialmente expresados en el enriquecimiento GO. Los genes resaltados en amarillo se descartan como contribuyentes de la enfermedad luego del filtrado por las comparaciones de efecto sequía y patógeno. Los 10 genes intervinientes en las vias de unión de proteínas no plegadas y plegado de proteinas pertenecen a la misma familia, con lo cual se ha tomado la decisión de agruparlos en el mismo conjunto de celdas dentro del cuadro.

	Glyma.02G305600	Molecular chaperone [HSP90 family], HEAT SHOCK PROTEIN 90, Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase		
	Glyma.03G057500	Molecular chaperone [DnaJ superfamily]		
	Glyma.03G218300	Molecular chaperone [DnaJ superfamily]		
	Glyma.06G289000	Molecular chaperone [DnaJ superfamily]		
UP: union de proteínas no plegadas	Glyma.08G332900	Molecular chaperone [HSP90 family], HEAT SHOCK PROTEIN 90, Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase		
plegado de proteínas (n=10)	Glyma.09G131500	Molecular chaperone [HSP90 family], HEAT SHOCK PROTEIN 90, Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase		
	Glyma.11G172401	Molecular chaperone [HSP90 family], HEAT SHOCK PROTEIN 90, Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase		
	Glyma.12G117900	Molecular chaperone [DnaJ superfamily]		
	Glyma.15G077700	Molecular chaperone [DnaJ superfamily]		
	Glyma.16G178800	Molecular chaperone [HSP90 family], HEAT SHOCK PROTEIN 90,Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase		
	Glyma.05G169200	PF00190, cupin		
UP: actividad de nutrients de reserva (n=5)	Glyma.11G153587	PF00190, cupin		
	Glyma.19G086200	PF00190, cupin		
	Glyma.19G086400	PF00190, cupin		
	Glyma.19G086500	PF00190, cupin		

	Glyma.01G032900	Leucine rich repeat proteins, some proteins contain F-box, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.01G033200	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.01G046900	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.03G034500	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.03G043000	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
DOWN: Unión al ADP (n=13)	Glyma.03G048200	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.06G311100	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.07G077700	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.08G318000	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.13G188300	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.16G214100	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.19G135600	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
	Glyma.U032405	Apoptotic ATPase, LEUCINE-RICH REPEAT-CONTAINING PROTEIN		
		EXOCYST COMPLEX COMPONENT		
		SEC6-RELATED EXOCYST		
	Glyma.03G026400	COMPLEX COMPONENT 3, Exocyst		
	0171111100 0020100	complex component Sec6		
		Glycosyl hydrolases family 16,		
	Glyma.07G174800	Xyloglucan endo-transglycosylase		
DOWN: Periferia		[XE1] C-terminus		
cellular, apoplasto	Glyma.15G154700	SPORULATION PROTEIN RELATED		
(n=5)		YCOSYL HYDROLASE-RELATED,		
	Glyma.15G169100	Glycosyl hydrolases family 16, Vyloglycon ando transglycosylase [VET]		
		C-terminus		
		Glycosyl hydrolases family 16.		
	Glyma.19G184900	Xyloglucan endo-transglycosylase		
		[XET] C-terminus		

#### Selección de genes identificados en enriquecimiento (KEGG)

Se cumplieron con los pasos descriptos anteriormente. Se analizaron los 43 genes presentes en el enriquecimiento KEGG, correspondientes a la comparación de interés (DT\_HS vs DM\_HS). De los 43 genes que evidenciaron aumento de expresión, hay que destacar que genes únicos son 39, dado que 4 de ellos están presentes en ambas categorías de análisis KEGG. Es decir que están presentes en procesamiento de proteínas a nivel de retículo endoplasmático y la interacción planta – patógeno. Dichos genes son proteínas del choque de calor del tipo 90 (HSP90, del inglés *heat shock proteins*), comportándose una de ellas como la chaperona involucrada en el reconocimiento de proteínas (GRP94, del inglés *glucose regulated protein*) en el caso de procesamiento de proteínas a nivel de retículo endoplasmático. Estos genes se encuentran resaltados en color naranja (Cuadro 3.4).

De los 39 genes únicos, se observó que todos los genes correspondientes a procesamiento de proteínas en retículo endoplasmatico también aumentaban su expresión en efecto sequia (DT\_HS vs DT\_H). El mismo patrón se observa para 4 de los 9 genes que están asociados al término de interacción planta – patógeno.

Los 5 genes restantes de dicha vía no están presentes en efecto sequia, con lo cual contribuirían a la tolerancia a la enfermedad es el caso de los genes que codifican para proteínas de unión al calcio (CML, del inglés *calmodulin like protein*), la proteína asociada a la patogénesis (PR), y la isoforma con dominio receptor del tipo quinasa 3 (CERK 1). Dichos genes están representados en blanco (Cuadro 3.4). Los genes descriptos anteriormente que están representados en simultáneo en DT\_HS vs DM\_HS y en efecto sequía (DT\_HS vs DT\_H) deben analizarse en la comparación de efecto patógeno (DT\_HS vs DT\_S) para dilucidar su contribución o no a la tolerancia a la enfermedad. En este sentido, se observó que la mayoría de dichos genes también cumplían un rol de sobre-expresión en la comparación de efecto patógeno (DT\_HS vs DT\_S). Dichos genes entonces son contemplados como

contribuyentes a la tolerancia del genotipo DT a la enfermedad de la PC. En el cuadro 3.4 dichos genes están representados en color verde.

Por el contrario, 5 genes de sobre-expresión en la vía de procesamiento de proteínas a nivel de retículo endosplasmático únicamente se sobre-expresaban en el efecto sequía (y por ende no en efecto patógeno), con lo cual los mismos se han descartado como genes que aporten a la tolerancia a la enfermedad. Los mismos se pueden observar en color amarillo en el cuadro 3.4. La siguiente tabla cita los 43 genes resultantes del enriquecimiento KEGG (39 genes únicos), resaltando todas las situaciones descriptas anteriormente

	Glyma.03G218300	Hsp40	DNAJ, dnaJ protein ERDJ3B, DnaJ homolog subfamily B member 11
	Glyma.02G205600	Hsp70	Heat shock 70 kDa protein-like, heat shock 70kDa protein 1/2/6/8
	Glyma.08G332900	Hsp90	uncharacterized protein LOC100819568, molecular chaperone HtpG
	Glyma.09G131500	Hsp90	molecular chaperone HtpG, heat shock protein 83
	Glyma.16G178800	Hsp90	molecular chaperone HtpG, heat shock protein 83
	Glyma.01G103200	NEF	hsp70-interacting protein, hsp70 nucleotide exchange factor fes1
	Glyma.03G070400	NEF	hsp70-interacting protein, hsp70 nucleotide exchange factor fes1
Procesamiento proteínas en	Glyma.02G305600	GRP94	heat shock protein 90-6, mitochondrial, heat shock protein 90kDa beta
retículo endoplasmático (n=34)	Glyma.02G252800	sHSF'	18.2 kDa class I heat shock protein, HSP20 family protein
(1-51)	Glyma.04G208000	sHSF'	17.5 kDa class I heat shock protein, HSP20 family protein
	Glyma.04G229800	sHSF'	small heat shock protein, chloroplastic, HSP20 family protein
	Glyma.06G054700	sHSF'	HSP20 family protein, class II heat shock protein
	Glyma.06G134900	aller,	uncharacterized protein LOC100798019,
		51151	HSP20 family protein
	Glyma.06G157800	sHSF'	18.5 kDa class I heat shock protein,

Cuadro 3.4 – Tabla resumen de los genes diferencialmente expresados en el análisis de enriquecimiento KEGG para la comparación (DT\_HS vs DM\_HS).

			HSP20 family protein
	C1 07C042600	sHSF'	HSP20-like chaperone protein,
	Giyilia.07G045000		HSP20 family protein
	Cluma 07C200000	ause,	17.5 kDa class I heat shock protein,
	Gryma.07G200000	SUPL	HSP20 family protein
	Cluma 07C200200	ause,	18.5 kDa class I heat shock protein,
	Giyina.07G200200	SU2L	HSP20 family protein
	CI 07C200200	ause,	putative class I heat shock protein,
	Glyma.07G200300	51151	HSP20 family protein
	Clama 07C200700	aller?	17.5 kDa class I heat shock protein,
	Glyma.0/G200/00	SHSF	HSP20 family protein
	Classe 08C0(0000		17.3 kDa class I heat shock protein,
	Glyma.08G069000	SHSF	HSP20 family protein
	G1 00 G212000	LICE	uncharacterized protein LOC100787482,
	Glyma.08G212000	sHSF	HSP20 family protein
	G1 00G210000	LICE	uncharacterized protein LOC100791734,
Procesamiento	Glyma.08G318900	sHSF'	HSP20 family protein
proteínas en	Glyma.11G246600	sHSF'	small heat shock protein, chloroplastic-like,
endoplasmático			HSP20 family protein
(n=34)	Glyma.13G071400	sHSF'	17.6 kDa class I heat shock protein,
			HSP20 family protein
	Glyma.13G176000	sHSF'	17.6 kDa class I heat shock protein,
			HSP20 family protein
	Glyma 13G176200	sHSF'	17.5 kDa class I heat shock protein,
			HSP20 family protein
	Glyma 14G063800	sHSF'	17.3 kDa class I heat shock protein,
			HSP20 family protein
	Glyma.14G099900	sHSF'	alpha-crystallin-Hsps_p23-like superfamily
			protein, HSP20 family protein
	Glyma.14G100000	sHSF'	alpha-crystallin domain-containing protein, HSP20 family protein
	Classic 14C214800		uncharacterized protein LOC100802108,
	Glyma.14G214800	SHSF	HSP20 family protein
	Glyma.15G264600	JUCEY	16.9 kDa class I heat shock protein 1,
		sHSF'	HSP20 family protein
	Glyma.16G012000	sHSF'	17.4 kDa class III heat shock protein,
			HSP20 family protein

		sHSF'	low molecular weight heat shock protein Hsp22.3
	Glyma.19G011400		precursor,
			HSP20 family protein
	Glyma.20G015900	sHSF'	putative class II heat shock, HSP20 family protein
	Glyma 14G156300	СМІ	calcium-binding EF-hand family protein,
	Olyma.140130300	CIVIL	calcium-binding protein CML
	Glyma.18G260700	CML	probable calcium-binding protein CML45 isoform X1, calmodulin
Internación	Glyma.13G035100	CML	probable calcium-binding protein CML44, calcium-binding protein CML
Planta-	Glyma.15G062500	PR	pathogenesis-related protein 1
patógeno (n=9)	Glyma.02G270800	CERK1	lysM domain receptor-like kinase 3 isoform X2 (RJ1)
	Glyma.08G332900	HSP90	uncharacterized protein LOC100819568, molecular chaperone HtpG
	Glyma.16G178800	HSP90	molecular chaperone HtpG, heat shock protein 83
	Glyma.09G131500	HSP90	molecular chaperone HtpG, heat shock protein 83
	Glyma.02G305600	HSP90	heat shock protein 90-6, mitochondrial, heat shock protein 90kDa beta

De esta manera, considerando 43 genes iniciales, siendo 39 genes únicos, y descartando 5 genes por el criterio descripto, se obtienen 34 genes únicos de interés en el enriquecimiento KEGG. Como se puede observar, los genes descartados corresponden a proteínas de choque de calor (HSP) siendo las mismas del tipo 40 y 70 respectivamente. Asimismo, se descartaron otros 3 genes que codifican para proteínas de choque de calor de tipo 20, siendo una de ellas, del tipo cloroplastídica. Una vez filtrado los genes que no son atribuibles a la tolerancia a la enfermedad, se pudieron identificar aquellos genes que serían responsables de la tolerancia a la PC, tanto a través del análisis de la plataforma GO, como KEGG.

El siguiente esquema (Figura 3.14 b) sintetiza el diagrama de trabajo con el cual se identificó el conjunto de genes que serían responsables de la tolerancia del genotipo DT a la PC en el cultivo de soja.



Figura 3.14 (b): Esquema resumen desde la realización del estudio RNA-seq hasta la identificación de genes asociados a la tolerancia a la enfermedad en el genotipo DT.

Partiendo del estudio RNA-seq, se analizaron en un principio los genes diferencialmente expresados (DEGs). Posteriormente se obtuvieron aquellos genes de interés (sobre- y sub-expresados) asociados a componentes celulares, funciones molecular y procesos biológicos, a través del análisis GO. En paralelo, también se llevó a cabo el análisis a través de la plataforma KEGG, que nos permitió asociar el comportamiento de dichos genes a vías y rutas metabólicas particulares. Luego, se procedió al filtrado de genes según el criterio desarrollado anteriormente, consiguiendo identificar aquellos genes a los cuales se le puede atribuir alguna implicancia en la tolerancia del genotipo DT a la enfermedad. Identificados los 60 genes únicos, se analizaron cuantitativamente. La siguientes tablas (Cuadro 3.5 y 3.6), muestran dichos genes descriptos, considerando la relación de cambio (representado por el fold change). El cálculo efectuado considera el logaritmo del cambio de expresión (log2 fold

change). Asimismo, se incorporaron los valores de padj que indican la significancia de cada uno de los genes (padj < 0,05). Por último, se efectúa el cálculo del logaritmo negativo del valor p ajustado (-log10 padj). Tanto para el análisis GO como KEGG, se pueden construir gráficos de puntos que relacionen el nivel de significancia de cada uno de los genes con el cambio de expresión (ya sea sobre o sub – expresión). El objetivo de dichas gráficas consiste en visualizar aquellos genes que muestran cambios de expresión de mayor y menor magnitud.

# Análisis de los genes resultantes de la plataforma GO

Cuadro 3.5 – Genes	diferencialmente	expresados	resultantes del	análisis GO.

	Genes	Log2FC	padj	-log10padj
	Glyma.02G305600	1,807	0,00090941	3,04
	Glyma.03G057500	5,298	6,14E-05	4,21
	Glyma.03G218300	1,366	0,043524	1,36
	Glyma.06G289000	0,259	0,0052537	2,28
UP	Glyma.08G332900	2,029	0,032918	1,48
Unión de proteínas	Glyma.09G131500	3,251	0,0033731	2,47
no plegadas, plegado	Glyma.11G172401	3,560	0,00064006	3,19
de proteínas (n=10)	Glyma.12G117900	0,354	2,82E-05	4,55
	Glyma.15G077700	3,423	0,00083888	3,08
	Glyma.16G178800	0,488	3,52E-04	3,45
	Glyma.05G169200	5,722	5,36E-06	5,27
LIP: activided de putrientes	Glyma.11G153587	3,163	0,0029075	2,54
de reserva (n-5)	Glyma.19G086200	0,277	0,0066965	2,17
de reserva (II–5)	Glyma.19G086400	3,025	0,031631	1,50
	Glyma.19G086500	0,635	1,09E-14	13,96
	Glyma.01G032900	-2,092	0,0087134	2,06
	Glyma.01G033200	-6,659	5,47E-12	11,26
	Glyma.01G046900	-0,434	4,57E-03	2,34
	Glyma.03G034500	-2,630	0,046173	1,34
	Glyma.03G043000	-0,239	0,027242	1,56
DOWN: unión al ADP	Glyma.03G048200	-3,059	0,035776	1,45
(n=13)	Glyma.06G311100	-0,655	4,32E-11	10,36
	Glyma.07G077700	-4,254	9,48E-12	11,02
	Glyma.08G318000	-0,249	0,0043054	2,37
	Glyma.13G188300	-0,219	0,019187	1,72
	Glyma.16G214100	-4,317	2,36E-08	7,63
	Glyma.19G135600	-0,163	0,023522	1,63
	Glyma.U032405	-1,794	0,014977	1,82
DOWN: periferia celular,	Glyma.03G026400	-4,177	0,00022185	3,65
apoplasto (n=2)	Glyma.15G169100	-2,999	4,79E-02	1,32

Como se puede observar en el gráfico de puntos GO (Figura 3.15), son 30 los genes resultantes del análisis, los cuales se distribuyen en las categorías funcionales de ADP-B (unión al ADP), CP + A (funciones de periferia celular y apoplasto), NRA (actividad de nutrientes de reserva) y UPB + PF (unión de proteínas no plegadas y plegado de proteínas). Como se observa en el cuadro 3.5 y se visualiza en el gráfico, a partir del eje de log2FoldChange (y la recta perpendicular indicando el no cambio de expresión= 0), las únicas categorías con genes sub-expresados en el genotipo tolerante con respecto al susceptible son ADP-B y CP + A. Como contracara, los genes de las categorías NRA y UPB + PF se sobreexpresaron en el genotipo tolerante.



Figura 3.15 – Gráfico de puntos de genes diferencialmente pertenecientes a las categorías GO enriquecidas.

Analizando cuantitativamente, se puede observar que los rangos de expresión representados a partir del log2FoldChange varían entre -7 y 7.
Asimismo, todos los las diferencias de expresión de los genes resultaron estadísticamente significativos, representado con el  $-\log_{10}(\text{padj}) = 1,3$  y valores superiores.

Observando los cambios de expresión, se puede observar que la mayoría de los genes se sub-expresaron entre 0,1 y 3,5 veces para el caso de ADP-B y CP+A, y se sobre-expresaron en la misma magnitud para los casos de las categorías NRA y UPB + PF. Como se observa en el cuadro 3.5, en el genotipo tolerante hay aumentos de expresión en genes asociados al plegado de proteínas y la unión de proteínas no plegadas, por ejemplo a través genes del tipo chaperonas moleculares y proteínas del choque de calor, indispensables en las respuestas ante el estrés generado a nivel proteico. Como contracara, hay una disminución en la expresión de genes asociados a la unión del ADP, representado en su mayoría por genes del tipo ATP-asa.

Por otra parte, se activan respuestas del tipo de reservorio de nutrientes, asociados a la superfamilia del tipo cupin (en este caso PF00190 cupin). Esta familia agrupa proteínas de almacenamiento, siendo las principales fuentes de N para las plantas en desarrollo.

### Análisis de los genes de la plataforma KEGG

	Genes	log2Fold Change	padj	Fórmula -log10padj
	Glyma.08G332900	2,029	0,032918	1,48
	Glyma.09G131500	3,251	0,003373	2,47
	Glyma.16G178800	0,488	3,52E-04	3,45
	Glyma.01G103200	2,002	0,039514	1,40
	Glyma.03G070400	2,545	0,0025329	2,60
UP: Procesamiento	Glyma.02G305600	1,807	0,0009094	3,04
proteinas en	Glyma.02G252800	3,713	0,0001446	3,84
endoplásmático	Glyma.04G208000	2,583	0,025826	1,59
(n=25+4 en común)	Glyma.04G229800	5,661	4,94E-09	8,31
	Glyma.06G054700	3,953	4,84E-02	1,32
	Glyma.06G134900	5,27	2,66E-07	6,57
	Glyma.06G157800	3,457	0,0014188	2,85
	Glyma.07G043600	2,707	0,033221	1,48

Cuadro 3.6 – Genes diferencialmente expresados correspondientes a las categorías KEGG sobre-representadas

	Glyma.07G200000	4,532	4,91E-03	2,31
	Glyma.07G200200	2,677	0,015084	1,82
	Glyma.07G200300	4,714	4,48E-04	3,35
	Glyma.08G212000	3,224	0,0016484	2,78
	Glyma.08G318900	6,004	1,10E-08	7,96
	Glyma.13G071400	3,994	9,58E-03	2,02
	Glyma.13G176000	4,153	2,33E-02	1,63
	Glyma.13G176200	3,857	2,51E-02	1,60
	Glyma.14G063800	2,923	0,0079328	2,10
	Glyma.14G099900	4,688	5,86E-03	2,23
	Glyma.14G100000	3,009	0,040425	1,39
	Glyma.14G214800	3,451	7,98E-03	2,10
	Glyma.15G264600	3,789	0,00021	3,68
	Glyma.16G012000	3,958	4,44E-02	1,35
	Glyma.19G011400	2,893	0,00066	3,18
	Glyma.20G015900	3,745	3,33E-02	1,48
	Glyma.14G156300	3,043	0,016059	1,79
Interacción	Glyma.18G260700	1,517	0,0069004	2,16
Planta-patogeno $(n-5+4)$ en	Glyma.13G035100	2,919	0,038992	1,41
$\frac{1-5+7}{\text{común}}$	Glyma.15G062500	1,656	0,0305	1,52
	Glyma.02G270800	4,740	1,40E-04	3,85

Como se puede contemplar en el gráfico de puntos KEGG (Figura 3.16), son 34 los genes resultantes del análisis. De la totalidad de los genes, 25 genes pertenecen a las vías PPRE (procesamiento proteico en retículo endoplasmático) y 5 a la vía INT-PP (interacción planta– patógeno). Los 4 genes restantes del análisis, son compartidos por ambas vías (PPRE y INT-PP), y se encuentran resaltados en el cuadro 3.6 en color celeste.

Como se observa en el cuadro 3.6 y se visualiza en el gráfico de la figura 3.16, a partir del eje de log2FoldChange, todos los genes resultantes del análisis evidenciaron un aumento de expresión en el genotipo tolerante, con respecto al genotipo susceptible:



Figura 3.16 – Gráfico de puntos de genes diferencialmente expresados resultantes del análisis en la plataforma KEGG

Todos los genes se sobre-expresaron, siendo el log2FoldChange de dichos genes en el rango de 0,48 a 6. Como se dilucida en la cuadro 3.6, en el genotipo tolerante hay aumentos de expresión en genes asociados al procesamiento de proteínas en el retículo, principalmente representado por familias del tipo proteínas de choque de calor (heat shock proteins), siendo las HSP 20 las más abundantes y de mayor aumento de la expresión. No obstante, los genes que codifican para HSP40, HSP70 y HSP90 también evidencian aumentos de expresión en el genotipo tolerante.

Asimismo, genes involucrados en la interacción planta – patógeno también evidenciaron aumentos de expresión, principalmente los receptores del tipo quinasa con dominio LysM, junto con proteínas asociadas a la patogénesis y proteínas de unión al calcio. Estos genes mencionados en ambas vías (procesamiento proteínas en retículo endoplasmático e interacción planta-patógeno) cumplen roles específicos dentro de dicho mapa metabólico.

### 3.3.4 Selección de genes candidatos para desarrollar marcadores moleculares

A partir del análisis realizado en las plataformas GO y KEGG, se pudo dilucidar la expresión diferencial de 64 genes en cuanto al análisis de enriquecimiento. Del conjunto de dichos genes, se pudieron identificar 4 que resultaron de particular relevancia dado que se detectaron en ambas metodologías en simultáneo. Dichos genes se describen en la tabla a continuación (Cuadro 3.7).

Cuadro 3.7 – Genes compartidos en las vías de enriquecimiento tanto de la plataforma GO como KEGG

Gen ID	Enriquecimiento GO	Enriquecimiento KEGG		
Cluma 02C205600 1	(+) Unión de proteínas	(+) Procesamiento de proteínas		
(HSD 00 / GDD04)	no plegadas / plegado de	en retículo endoplasmático /		
(113F 907 0KF 94)	proteínas	Interacción planta-patógeno		
Glyma 08G332000 1	(+) Unión de proteínas	(+) Procesamiento de proteínas		
(HSP 00 / HSP 00)	no plegadas / plegado de	en retículo endoplasmático /		
(1151 307 1151 30)	proteínas	Interacción planta-patógeno		
Glyma 00G131500 1	(+) Unión de proteínas	(+) Procesamiento de proteínas		
(HSP 00 / HSP 00)	no plegadas / plegado de	en retículo endoplasmático /		
(1131 90 / 1131 90)	proteínas	Interacción planta-patógeno		
Clyma 16G178800 1	(+) Unión de proteínas	(+) Procesamiento de proteínas		
(HSD 00 / HSD 00)	no plegadas / plegado de	en retículo endoplasmático /		
(1151 90 / 1151 90)	proteínas	Interacción planta-patógeno		

Dichos genes tienen una particularidad: no sólo se expresan en ambas plataformas, sino que dentro del enriquecimiento KEGG se expresan tanto en la vía de procesamiento de proteínas en retículo endoplasmático como en la de interacción planta-patógeno. Las implicancias a nivel de procesamiento proteico también se pueden dilucidar en los términos destacados en GO, como la unión de las proteínas no plegadas y el plegado de proteínas. El criterio de la selección de dichos 4 genes justamente radica en su importancia dentro del análisis llevado a cabo: su relevancia en ambas plataformas (GO y KEGG), junto con las funciones dentro de las vías destacadas anteriormente.

# 3.4 Discusión

No hay que perder de vista que, a nivel productivo, la PC aumenta sus niveles de incidencia y severidad en situaciones de estrés hídrico, como se ha mencionado anteriormente. Campañas con escasa precipitación como la de 2018 o 2023 en Argentina son propensas a experimentar mayores niveles de la enfermedad. En este sentido, en esta comparación se está poniendo a prueba justamente el desempeño genotípico de ambos (DM y DT), otorgándole las condiciones propicias al patógeno para infectar. Un gen que se exprese diferencialmente en esta situación en el genotipo tolerante DT podría indicar las posibles vías y procesos metabólicos sobre las cuales el genotipo actúa para frenar el avance de la enfermedad. Se pudo observar un mecanismo en el genotipo tolerante de mayor procesamiento proteico ante el estrés, representado por el plegado de proteínas y la unión de las proteínas que no se encuentran plegadas, junto con una mayor actividad de los reservorios de nutrientes. Estos mecanismos descriptos otorgarían una ventaja al genotipo DT para la tolerancia a la PC, en comparación con el genotipo susceptible.

### Biología de Sistemas

Para analizar la evolución de dichos distintos procesos metabólicos, se utilizó la metodología KEGG, lo que permitió arribar a las siguientes discusiones:

### Interacción planta – patógeno

El sistema inmune de las plantas está constituido por una inmunidad innata que actúa de dos formas fundamentales. La primera está basada en un reconocimiento de PAMP (del inglés: *pathogen-associated molecular pattern*), mediante receptores que reconocen estos patrones y que se encuentran en la superficie de las células vegetales, y se denomina PTI (del inglés: *PAMP-triggered immunity*). La resistencia basal es la primera línea de defensa y se induce cuando los receptores de la membrana plasmática identifican ciertos compuestos de los microorganismos, los MAMP. Los MAMP son componentes vitales de muchos

microorganismos, como así también de los patógenos, y se presentan de manera constitutiva, o sea, siempre están presentes durante la vida de los microorganismos. Se ha identificado a varios de ellos en patógenos de plantas, como la flagelina, los lipopolisacáridos, los factores de elongación y los  $\alpha$ -glucanos, entre otros, y dependiendo de qué patógeno se trate (Hussan et al. 2020)

Como se comentó anteriormente, el reconocimiento de los patógenos está mediado por un grupo de proteínas receptoras. Las proteínas símil receptor y los receptores con actividad de quinasa (RLK, en inglés *receptor like protein kinase*) están localizados en la membrana plasmática y contienen dominios con repeticiones ricas en leucina. Los receptores con actividad de quinasa contienen un dominio citosólico con actividad de quinasa que media la transducción. La señalización por receptores con actividad de quinasa se inicia cuando el receptor se autofosforila y se forma un complejo con el ligando que se internaliza y luego se degrada.

Las proteínas con localización citoplasmática forman otras dos clases de receptores. Además del dominio rico en leucina, estas proteínas contienen un dominio de unión de nucleótidos. En el genotipo tolerante se observa una sobre-expresión del gen que codifica para uno de estos receptores RLK (lysM domain receptor-like kinase 3 isoform X2 (RJ1) -Glyma.02G270800.1).

En el mapa metabólico (Figura 3.17) se lo puede observar como receptor CERK1. Estos receptores juegan un papel clave en el reconomiciento del patógeno y la inducción de otros genes de respuestas de defensa. En *Arabidopsis thaliana*, se identificó una proteína similar al receptor LysM (LysM RLK1) necesaria para la señalización de quitina (Wan et al. 2008). La quitina, un polímero de N-acetil-D-glucosamina, se encuentra en las paredes celulares de los hongos, pero no en las plantas. Las células vegetales pueden percibir fragmentos de quitina (quitooligosacáridos) que conducen a la inducción de genes y respuestas de defensa. Asimismo,

se demostró que mutaciones en este gen bloquearon la inducción de genes responsivos al ataque fúngico, provocando aumento de la susceptibilidad a las enfermedades.

Dentro del mapa metabólico, también se puede observar sobre-expresión en el genotipo tolerante de los genes que codifican para proteínas de unión al calcio, conocidas como CML (Glyma.13G035100.1, Glyma.14G156300.1 y Glyma.18G260700.1). El calcio juega un rol importante en la interacción planta-patógeno. La percepción de patógenos se traduce en un Ca<sup>2+</sup> citosólico elevado (mediado por la membrana plasmática y los canales intracelulares) como un paso temprano en una cascada de señalización (Cheval et al. 2013). El aumento de Ca<sup>2+</sup> que ocurre en respuesta a la percepción de PAMP podría ocurrir debido a eventos de fosforilación o un aumento en nucleótidos cíclicos. Aguas abajo del aumento inicial de Ca<sup>2+</sup>, la cascada de señalización incluye la activación de calmodulina y proteínas del tipo quinasa, y generación de NO<sup>-</sup> y ROS. Algunos de estos eventos aguas abajo amplifican la señal de Ca<sup>2+</sup> mediante la activación adicional de los transportadores de Ca<sup>2+</sup> (Ma y Berkowitz, 2007).

Como se ha podido observar, a través de la PTI, por ejemplo, se producen respuestas de defensa tales como la acumulación de ROS, que conducen a respuestas múltiples tales como el refuerzo de la pared celular, el cierre estomático y las respuestas de hipersensibilidad, cuyo objetivo es limitar el avance de la enfermedad.



Figura 3.17 – Mapa metabólico correspondiente a las vías de interacción planta – patógeno. Los recuadros en rojo agrupan los genes enriquecidos (sobre-expresados) en el genotipo tolerante en comparación al susceptible. Mapa extraído de Kegg (genome.jp)

Sin embargo, hay patógenos exitosos que producen efectores que inhiben la PTI, los cuales son reconocidos por las plantas mediante receptores adicionales, que desencadenan respuestas efectoras contra los mismos. Esta respuesta constituye la segunda forma de actuación de la inmunidad innata y se denomina ETI (del inglés: *effector-triggered immunity*) (Jones y Dangl, 2006; He et al. 2007; Boller y He, 2009). Este segundo nivel de reconocimiento es altamente específico y se establece cuando los productos de los genes de resistencia de la planta interactúan con moléculas específicas del patógeno.

Estas moléculas se denominan efectores y son productos de los genes de avirulencia del patógeno: los genes *Avr*, que no son los que específicamente producen la enfermedad; éstos se

liberan en el interior de la célula vegetal durante la infección. Las respuestas primaria y secundaria crean resistencia y se basan en la activación rápida de cascadas de transducción de señales que promueven cambios intracelulares; los más importantes son los que desencadenan las respuestas secundarias, o sea, las inducidas por la presencia del patógeno. Las proteínas de resistencia registran a estos efectores, o también identifican ciertas proteínas de la planta (proteínas blanco) que han sido alteradas por estos efectores, y desencadenan la respuesta de defensa secundaria.

Esta resistencia inducida en general se activa junto con un proceso de muerte celular local en los puntos de infección, que forma lesiones necróticas. Esta reacción se conoce como respuesta hipersensible, que asociada a otras respuestas localizadas de defensa bloquea la propagación posterior del patógeno a las células circundantes.

Observando el mapa metabólico, se observa una sobre-expresión de las proteínas de choque de calor, HSP (específicamente HSP90, representado por varios genes), que contribuyen a los procesos mencionados anteriormente, principalmente al desencadenante de la respuesta de hipersensibilidad. Las HSP funcionan como chaperonas moleculares responsables del plegamiento, ensamblaje, translocación y degradación de proteínas en condiciones de estrés y en muchos procesos celulares normales. Las HSP desempeñan un papel indispensable como chaperonas moleculares en el control de calidad de los receptores residentes en la membrana plasmática y las proteínas R intracelulares frente a posibles invasores microbianos.

El último eslabón sobre-expresado dentro del mapa metabólico corresponde a los genes que codifican para proteínas asociadas a la patogénesis (PR, en este caso la PR-1, correspondiente a Glyma.15G062500.1). Este mecanismo es parte de la inducción de genes asociados a la defensa vegetal. Las proteínas PR constituyen un grupo heterogéneo de proteínas, constitutivas o inducidas, muy estables a pH bajos, relativamente resistentes a la acción de proteasas, tóxicas para los patógenos, las que se acumulan predominantemente en los espacios extracelulares. Se inducen diferencialmente dependiendo de la especie vegetal y del tipo de infección que ha sufrido la planta.

Estas PR cumplen un rol central, por ejemplo en la acumulación de fitoalexinas. Las fitoalexinas son compuestos antimicrobianos que se acumulan en algunas plantas en altas concentraciones, después de infecciones bacterianas o fúngicas y ayudan a limitar la dispersión del patógeno. Tienen varias características, entre ellas el hecho de que se sintetizan muy rápido, en pocas horas después del ataque microbiano.

A modo de resumen, se puede observar los diferentes modos en los cuales los genes sobre-expresados en el mapa de interacción planta-patógeno contribuyen a la tolerancia de la enfermedad en el genotipo tolerante DT: desde el rol de los receptores CERK1, la acción de las proteínas del tipo CML, la expresión de HSP – en su rol de control de calidad de los receptores de reconocimiento y la acumulación de PR1, junto con la producción de fitoalexinas.

### Procesamiento de proteínas en retículo endoplasmático (PPRE)

La respuesta al estrés del retículo endoplásmico es un mecanismo altamente conservado que resulta de la acumulación de proteínas desplegadas (no plegadas) o mal plegadas (Simoni et al. 2022) en el retículo endoplasmático (Figura 3.18). La respuesta juega un papel importante al permitir que las plantas perciban y respondan a condiciones ambientales adversas, como estrés por calor, estrés por sal e infección por patógenos, como constituye en este caso por la infección de *M. phaseolina*.

Dado que el retículo endoplasmático es un microambiente bien controlado para la síntesis y el plegamiento adecuados de proteínas, es altamente susceptible a las condiciones de estrés (Korner et al. 2015). La acumulación de proteínas no plegadas o mal plegadas activa una vía de señalización, llamada respuesta de proteína desplegada/no plegada (UPR, del inglés *unfolded protein response*), que actúa para aliviar el estrés del retículo, procurando de esta forma evitar los procesos que conducen a la muerte celular (Deng et al. 2013).

Las plantas mantienen un equilibrio entre la demanda de plegamiento de proteínas y la capacidad de plegado. Cuando ese equilibrio se ve alterado por condiciones, como el estrés ambiental o biótico, las proteínas desplegadas y mal plegadas se acumulan en el retículo, causando "estrés de retículo endoplasmático", una condición que puede dañar las células.

Para hacer frente a este problema, las células están equipadas con un sistema de control de calidad del retículo endoplasmático. La acumulación de proteínas desplegadas o mal plegadas activa la UPR que regula al alza la producción de factores que promueven el plegamiento de proteínas y/o eliminan las proteínas desplegadas/mal plegadas de retículo endoplasmático a través de la degradación asociada a él (ERAD, del inglés *endoplasmic reticulum associated protein degradation*).

# Respuesta de proteinas no plegadas (UPR)

El plegamiento de proteínas implica todos los procesos mediante los cuales las proteínas nacientes alcanzan su conformación nativa y funcional. Las proteínas nacientes destinadas a la secreción portan péptidos con señal N-terminales que interactúan con partículas de reconocimiento de señales. La unión de dichas partículas detiene la traducción, y el complejo de ARNm, ribosoma y polipéptido naciente que contiene las partícula se dirige a la membrana del retículo endoplasmático, donde se disocia. A continuación, se reanuda la traducción y el polipéptido en crecimiento se transloca al lumen del retículo endoplasmático a través del complejo de translocón Sec61.

El plegamiento de proteínas comienza inmediatamente después de la entrada al lumen del retículo endoplasmático y es ayudado por varias enzimas y chaperonas moleculares. El papel principal de las chaperonas es unirse a las proteínas nacientes y prevenir su agregación. En un estado desplegado/no plegado, las regiones hidrófobas de la proteína pueden exponerse a la superficie, haciendo que las proteínas sean vulnerables a la agregación.

Hay dos vías principales para el plegamiento de proteínas, vías dependientes de N-

glucanos e independientes de N-glucanos. La vía independiente de N-glicanos involucra la proteína de unión del lumen (BiP) o la proteína regulada por glucosa de 94 kDa (GRP94). En el mapa metabólico se puede observar la sobre-expresión de un gen que codifica para GRP94 (Glyma.02G305600.1, heat shock protein 90kDa beta, heat shock protein 90-6, mitochondrial). BiP es un acompañante abundante en el retículo endoplasmático, donde se asocia con el translocón Sec61 e interactúa con polipéptidos recién sintetizados. Su interacción es ayudada por una co-chaperona, ERdj3, una proteína de ADN-J (Park y Seo, 2015), de una manera dependiente de nucleótidos. BiP tiene un dominio de unión a nucleótidos N-terminal y un dominio de unión a sustrato C-terminal, que puede unirse a parches hidrófobos de proteínas de estado no nativo, como las que forman cadenas  $\beta$  enterradas en el núcleo de la proteína. ERdj3 se une directamente a proteínas nacientes y recluta BiP. En el mapa metabólico (figura 3.18) se observa una sobre-expresión de HSP40 que tiene función Erdj3, que a pesar de haberse catalogado exclusivamente para efecto sequía anteriormente, cumple un rol en la vía metabólica (Glyma.03G218300.1, DNAJ, dnaJ protein ERDJ3B, DnaJ homolog subfamily B member 11).

Esto induce la actividad ATPasa de BiP y la hidrólisis del ATP unido a ADP. En la forma unida a ADP, BiP se une a sustratos con alta afinidad. El intercambio de ADP con ATP libera el sustrato, que luego puede progresar más a través del proceso de plegado. Se ha observado que, por ejemplo, el silenciamiento de genes HSP40 provocó mayor susceptibilidad a *Soybean mosaic virus* en el caso del cultivo de soja (Liu y Whitham, 2013). La otra vía de plegamiento es dependiente de N-glicanos. La mayoría de las proteínas secretadas y de membrana están glucosiladas co o postraduccionalmente en residuos de asparagina dependientes del contexto.

El plegamiento adecuado de proteínas, en particular de proteínas grandes, puede ser un desafío. Incluso en condiciones óptimas, todavía puede ocurrir un plegado incorrecto de

proteínas. En condiciones de estrés o en etapas de desarrollo específicas en las que hay una gran secreción de proteínas, las demandas de plegamiento de proteínas en el retículo endoplasmático pueden superar su capacidad de plegamiento. La acumulación resultante de proteínas desplegadas / mal plegadas produce estrés en el retículo, que puede ser dañino para las células y para todo el organismo. Para hacer frente a esto, el retículo endoplasmático tiene un sofisticado sistema de control que puede detectar y retener proteínas mal plegadas para rondas adicionales de plegado y eliminar proteínas mal plegadas terminalmente a través de ERAD o degradación autofágica. Las glicoproteínas liberadas que no se pliegan correctamente son reconocidas y reglucosiladas (monoglucosiladas) por difosfato de uridina (UDP)-glucosa: glicoproteína glucosiltransferasa (UGGT), tal como se puede observar en el centro del mapa metabólico.



Figura 3.18 – Mapa metabólico correspondiente a las vías de procesamiento de proteínas en retículo endoplasmático. Los recuadros en rojo agrupan los genes enriquecidos (sobreexpresados) en el genotipo tolerante en comparación al susceptible. Mapa extraído de Kegg (genome.jp)

# ERAD

Las glicoproteínas se extraen de la maquinaria de plegado mediante la escisión de los residuos  $\alpha$ -1,2-manosilo por las  $\alpha$ -manosidasas. Las ubiquitina-proteína ligasas del sistema ERAD reconocen glicoproteínas mal plegadas a través de una señal bipartita (parches proteicos hidrófobos expuestos y una manosa unida a  $\alpha$ -1,6 en la cadena C del oligosacárido) y las dirigen para su eliminación en el citosol.

La degradación de proteínas mal plegadas por el sistema ERAD es clave para restablecer el equilibrio homeostático en respuesta al estrés del retículo endoplasmático. Todas las células eucariotas poseen un sistema de control de calidad para identificar y eliminar proteínas desplegadas o mal plegadas en la vía secretora. Las proteínas mal plegadas que no alcanzan su estado nativo son degradadas por el sistema ERAD. ERAD consta de cuatro pasos: reconocimiento, ubiquitinación, dislocación y degradación de las proteínas para ser eliminadas. El reconocimiento de proteínas mal plegadas para su degradación por parte del sistema ERAD es un desafío considerando los diferentes sustratos de proteínas con los que debe enfrentarse el sistema de control de calidad. El sistema de reconocimiento involucra ubiquitinas E3-ligasas equipadas con adaptadores, que dotan a las ligasas de notables capacidades de reconocimiento (Howell, 2013). La ubiquitinización de dichas proteínas concluye con su envío al protesoma. En el mapa metabólico se puede observar todos los grupos de genes sobre-expresados en el genotipo tolerante, desde las sHSP, HSP 70, HSP 90 (listados en el cuadro 3.4) y los NEF, que contribuyen al proceso de degradación de proteínas mal plegadas.

A modo de resumen, se puede ver los diferentes modos en los cuales los genes sobreexpresados en el mapa de PPRE contribuyen a la tolerancia de la enfermedad en el genotipo tolerante DT: desde el rol de GRP94, la acción de HSP40 como co- chaperona (ERdj3), la expresión de HSP (HSP 70, 90 y sHSF), junto con las NEF como actores clave dentro de la degradación de proteínas mal plegadas vía el sistema ERAD.

# **3.5** Conclusiones

El estudio RNA-seq permitió dilucidar el conjunto de DEGs expresados en el genotipo tolerante DT respecto al susceptible DM, en condiciones en las cuales el patógeno ataca a nivel productivo (proceso de infección con situación de sequía en simultáneo). A partir de dichos DEGs, se llevó a cabo el análisis tanto en la plataforma GO como KEGG para dilucidar cuáles genes y en qué medida están relacionados a funciones biológicas específicas o vías metabólicas. Posteriormente, se dilucidó a través del contraste con otras comparaciones de tratamientos, cuáles de dichos genes resultantes del análisis eran exclusivamente atribuibles a una respuesta ante la sequía, siendo la estrategia el descarte en principio de los mismos.

Una vez que se realizó el enriquecimiento, se pudo obtener el conjunto de genes asociados a las categorías específicas GO y las vías metabólicas y funciones específicas para el caso de KEGG. A partir de estos análisis se pudo concluir el rol preponderante en la tolerancia en la enfermedad. Desde la plataforma GO, se verificó la importancia del aumento de expresión de genes asociados a la unión de proteínas no plegadas en el genotipo tolerante, junto con el aumento de expresión de la actividad de reservorios de nutrientes.

Asimismo, la vía KEGG permitió identificar la importancia de las vías de interacción planta-patógeno, a partir de una sobre-expresión de proteínas de unión al calcio, los receptores CERK1, las HSP90 (en su rol y aporte a la reacción de hipersensibilidad) y las PR, asociadas a patogénesis, obviamente que contribuyen a la tolerancia a la enfermedad. Este panorama se completa con el análisis del procesamiento de proteínas en el retículo endoplasmático, donde se observa sobre-expresión de la respuesta UPR (vía GRP94 y HSP40) y la actividad de ERAD, representado por un conjunto de genes sobre-expresados en el genotipo tolerante (HSP 70-90, sHSF y NEF).

A partir de analizar los genes resultantes de los DEGs y el enriquecimiento funcional,

junto con la biología de sistemas desarrollada a partir de los mapas metabólicos KEGG, se pudo dilucidar el conjunto de 4 genes candidatos para el desarrollo de marcadores moleculares. Dichos genes no solamente se destacaron por su expresión en ambas metodologías de enriquecimiento, sino que se pudo dilucidar sus roles preponderantes dentro de las vías de procesamiento de proteínas en retículo endoplasmatico y la interacción planta – patógeno.

# Capítulo 4. Identificación de polimorfismos en genes de interés para el desarrollo de marcadores moleculares de tolerancia a la enfermedad

### 1. Introducción

Con el avance de la investigación en el área de la biotecnología y la biología molecular, se ha recurrido a herramientas basadas en el material genético, como lo es el caso de los marcadores moleculares. En este sentido, la selección asistida por marcadores moleculares es una herramienta moderna que ha cobrado un rol protagónico en el mejoramiento genético en los últimos años. Asimismo, esta herramienta ofrece diversas ventajas, desde su mayor eficiencia hasta su capacidad de mejorar los programas de mejoramiento de los diferentes cultivos. De hecho, uno de los principales fines aplicativos implica el desarrollo de nuevos cultivares. Un ejemplo concreto lo constituye el mejoramiento genético asistido por marcadores moleculares para intentar controlar o mitigar los efectos de diversas adversidades bióticas. La tolerancia a las enfermedades, por ejemplo, suele estar gobernada por un conjunto de genes. Esfuerzos orientados a mejorar genéticamente las variedades existentes para el aumento de la tolerancia resulta la principal vía de acción frente a las enfermedades. Este aspecto se ha resaltado aún más en las últimas décadas ante la insuficiencia de los manejos agronómicos actuales, como pueden ser las medidas culturales (manejo de fechas de siembra, rotaciones) y químicas (ante la sostenida capacidad de resistencia y adaptación de los agentes patógenos).

En soja se desarrollaron e implementaron numerosas técnicas de marcado molecular. Entre los marcadores moleculares más utilizados se puede mencionar los marcadores microsatélites y los SNPs. Los marcadores microsatélites se usan comúnmente para estimar el grado de parentesco entre genotipos, y debido a su abundancia, sensibilidad y reproducibilidad

111

resultan de preferencia para estudiar la diversidad genética y para la identificación varietal (Song et al. 2010). Los SNPs son un tipo de marcador que consiste en un polimorfismo de nucleótido único, es decir que hay una variante genómica en la posición de una base única en el ADN.

La búsqueda de SNPs es una estrategia útil a la hora de pensar un programa de mejoramiento para mitigar estreses que afecten el ciclo del cultivo. Esta herramienta es ampliamente utilizada en diversos cultivos, tales como en durazno (Cirilli et al. 2017), algodón (Sun et al. 2019), arvejas (Leonforte et al. 2013) y soja (Shi et al. 2015) entre otros. Las plataformas que se conocen como de NGS también han cobrado una relevancia significativa, principalmente por su capacidad de secuenciación y genotipado de alto rendimiento. Las tecnologías de NGS y marcadores moleculares se utilizan cada vez más en la actualidad, a tal punto de favorecer el mejoramiento, por ejemplo, a través de la identificación de polimorfismos en genes de interés. En este sentido, la detección de variaciones a través de SNPs y las inserciones y deleciones (INDELs, del inglés *insertions and deletions*) es una de las principales vías para explicar comportamientos genéticos que permitan incluirse en programas de mejoramiento.

Los marcadores moleculares son segmentos de ADN con una ubicación específica en un cromosoma y constituyen puntos de referencia en el genoma, no son influenciados por el ambiente y brindan información sobre las diferencias, a nivel genotípico, entre individuos (Amiteye, 2021). En el caso de *M. phaseolina*, desarrollar marcadores específicos a la resistencia cobra mayor relevancia dado la complejidad del manejo agronómico para mitigar los efectos de la enfermedad y la inexistencia actual de resistencia genética (Poudel, 2023). A partir de lo explicado anteriormente, es posible pensar en la búsqueda de SNPs en el transcriptoma, dada su fortaleza en comparación a la búsqueda en el genoma. El análisis vía transcriptoma se realiza analizando solamente las regiones codificantes de los genes, lo cual representa una ventaja dado que el impacto esperado es mayor que el que puede encontrarse en regiones no codificantes.

Trabajos publicados demuestran la utilidad del análisis transcriptómico para la búsqueda e identificación de polimorfismos en el genoma, haciendo foco en el combate de adversidades, como en soja (Yadav et al. 2015) sorgo (Chopra et al. 2015) y girasol (Livaja et al. 2013). Asimismo, los polimorfismos encontrados podrán ser analizados y priorizados en función a diferentes criterios. En este sentido, la identificación de variantes en genes relevantes o que desempeñen una función central frente a la PC de la soja es un criterio a tener en cuenta. Esta información puede ser combinada con otros criterios indispensables como resulta los genes expresados diferencialmente y el posterior enriquecimiento funcional vía las plataformas GO y KEGG (analizados en el capítulo 3).

En este capítulo se abordará el último objetivo, que consiste en la búsqueda de poliformismos en la secuencia transcriptómica para el desarrollo de marcadores moleculares. Los polimorfismos buscados serán priorizados o filtrados según los genes diferencialmente expresados en los genotipos frente a la enfermedad. Una vez aplicados los criterios seleccionados, se propondrá un panel de marcadores moleculares que puedan eventualmente utilizarse en un programa de mejoramiento.

### 2. Materiales y métodos

### 2.1. Búsqueda de mutaciones en la secuencia transcriptómica

La búsqueda de polimorfismos en la secuencia transcriptómica constituye una herramienta fundamental para la propuesta de un panel de marcadores moleculares. Se utilizaron diferentes herramientas bioinformáticas. Para la búsqueda de mutaciones en la secuencia transcriptómica, ya sean SNPs o INDELs se utilizó el software GATK (del ingles, *Genome Analysis ToolKit*, McKenna y col., 2010). Para el filtrado inicial se utilizaron los parámetros estándar del programa y se definieron otros parámetros (cluster:3, WindowSize:35; QD < 2.0; FS > 30.0; DP<10.) Una vez identificadas las mutaciones se anotaron utilizando el software SnpEff (Cingolani et al. 2012), incluyendo la función e impacto del cambio aparecido. La función de la

mutación puede ocurrir en diversos sentidos, y con diferentes implicancias. Las mismas pueden ser con sentido (cambio de un aminoácido por otro), sin sentido (introducción de un codón stop) o silenciosa (sin cambio de fenotipo). Asimismo, el impacto predicho por el software también puede arrojar diferentes resultados. El impacto puede clasificarse en alto (si SnpEff calcula que tendrá efectos deletéreos), moderado o bajo. Por otra parte, también puede clasificarse como modificador. Es importante destacar que el software SnpEff puede asignarle mas de un efecto a una misma mutación. Para seleccionar las mutaciones de interés e impacto se aplicaron sucesivos filtros:

(i) En primer lugar, se filtraron aquellos SNPs que cayeron en *contigs* o *scaffolds*, (es decir no se pudieron mapear a regiones cromosómicas concretas). Los *contigs* son secuencias continuas ensambladas en base a fragmentos de secuencias, mientras que los *scaffolds* son secuencias genómicas reconstruidas a través del ensamblado de *contigs*. Luego, se seleccionaron aquellas asociadas al genotipo tolerante DT, tomando aquellas variantes donde el genotipo sensible DM (tanto en control como en presencia de la enfermedad) presentó la misma variante que la referencia (disponible en Phytozome - versión 4 de la variedad Williams 82), pero el genotipo DT presentó la alternativa (Figura 4.1).

(ii) A continuación, se seleccionaron las mutaciones que el software SnpEff indicó como de impacto alto o moderado.

(iii) Por último, se seleccionaron aquellas que cayeran en genes diferencialmente expresados frente al patógeno en el genotipo DT. Solo se admitieron aquellos SNPs en donde todas las repeticiones biológicas de cada tratamiento presentaran la misma variante (aquí quedaron descartadas variantes que implicaran heterocigosis y también diferencias entre repeticiones biológicas). Este criterio intentó lograr un enfoque lo más restrictivo posible, con el objetivo final de lograr un número de marcadores moleculares que sea posible de aplicar en un programa de mejoramiento de soja frente a la PC.



Figura 4.1: Proceso secuencial de filtros aplicados para la obtención de un panel de marcadores moleculares frente a la PC en soja.

# 3. Resultados

# 3.1 Variantes en la secuencia transcriptómica

La búsqueda de mutaciones en la secuencia transcriptómica arrojó un total (Cuadro 4.1) de 55780 SNPs y 22523 INDELs (inserciones o deleciones). A través del programa SnpEff se calculó el impacto, función y región donde cayeron las mutaciones.

	1	1
	SNPs	INDELs
Gm1	2160	949
Gm2	2457	1243
Gm3	3795	1003
Gm4	2159	995
Gm5	2663	1021
Gm6	3352	1401
Gm7	3679	1260
Gm8	2557	1457
Gm9	3652	1246
Gm10	2547	1195
Gm11	2012	1114
Gm12	1605	878
Gm13	4469	1559
Gm14	1866	887
Gm15	3448	1126
Gm16	3227	960
Gm17	1833	1131
Gm18	4788	1224
Gm19	1313	928
Gm20	2198	946
	55780	22523

Cuadro 4.1: Cantidad de SNPs e INDELs encontrados en el transcriptoma por cromosoma.

Al analizar específicamente los SNPs e INDELs, podemos observar que la mayoría de ellos cayeron en la categoría descripta como "MODIFIER" (Figura 4.2). Esta categoría implica que probablemente los cambios que estén introduciendo no tengan impacto relevante o que el efecto que provoca no sea de fácil interpretación (esto sucede, por ejemplo, al caer en regiones intergénicas). Luego, hubo gran cantidad de variaciones, tanto SNPs como INDELs, de impacto bajo ("LOW") que indican mutaciones no deletéreas que no provocan cambio en la funcionalidad de la proteína, es decir que no afectan el rol que cumple la presencia activa de dicha proteína. Distinguiendo dentro de las variantes, el impacto ("LOW") fue más apreciable para el caso de los SNPs (en ambos genotipos), con una frecuencia mucho mayor. Para el caso de los INDELs, la frecuencia ("LOW") es menos apreciable. Asimismo, el análisis de los impactos de índole moderado ("MODERATE") resulta de gran interés dado que este tipo de impacto implica posibles cambios en la funcionalidad de la proteína resultante, aunque no necesariamente ese cambio es absoluto ni totalmente disruptivo. Por último, se encuentra el grupo de impacto alto ("HIGH"), que representa una baja frecuencia sobre el total de variantes

analizadas para ambos genotipos. Esta descripción era esperable, dado que son aquellas variantes que implican un impacto muy relevante sobre la proteína resultante, ya sea cambiando por completo la funcionalidad, eventualmente provocando degradación e incluso pérdida de funcionalidad. Por último, analizando específicamente la función de los SNPs, es posible destacar que la mayoría se encontraban en regiones donde provocaron ya sea cambios no silenciosos (o con sentido) y del tipo silenciosos. Por otra parte, las variantes atribuibles a pérdidas de sentido fueron pocos relevantes en ambos genotipos.



Figura 4.2: Información del impacto y función predichos por SnpEff para las variantes encontradas en la secuencia transcriptómica.

Teniendo en cuenta que el objetivo final es la propuesta de un posible panel de marcadores moleculares, se procedió a la selección de variantes siguiendo diversos criterios de

forma secuencial (4 criterios sucesivos), y resumidos en la Figura 4.3:

 i) En primer medida, con el fin de analizar únicamente aquellas variaciones presentes en regiones conocidas dentro de los cromosomas, se descartaron aquellas variantes que se situaran en contigs o scaffolds. A partir de este primer filtro, de las 78303 variantes iniciales se obtienen 77387 variantes totales, distribuidas en 20 cromosomas.

ii) Consecuentemente, el segundo criterio para el filtrado consistió en trabajar con aquellas variantes en las cuales el genotipo tolerante DT presentó la alternativa respecto al genotipo susceptible DM. Este criterio resulta altamente restrictivo, dado que dicha condición debe cumplirse para todas las repeticiones biológicas. Cabe recalcar, de este modo, que la sola presencia de un alelo en el genotipo DT igual al genotipo DM sería descartado como potencial marcador. Es decir, el polimorfismo que puede ser útil para asociar un marcador molecular es el que ocurre entre ambos genotipos contrastantes. A partir del cumplimiento del criterio, de las 77387 variantes previas se obtuvieron 2340 totales.

iii) El tercer criterio establecido está relacionado a características atribuibles a dichas variantes a través de la utilización del software SnpEff. Dicho software permite analizar el tipo de impacto que tiene cada una de dichas variables. Para seguir restringiendo el panel de variantes candidatas, se retuvieran aquellas que presentaran un impacto alto ("HIGH") o moderado ("MODERATE").
Como resultado de la aplicación de dicho filtro, de las 2340 variantes anteriores se obtuvieron 1040.

iv) Para finalizar con el filtrado, se estableció como criterio la vinculación de dichas variantes con el comportamiento génico observado en el RNA-seq. Específicamente, los SNPs a seleccionar debieron estar relacionados a genes diferencialmente expresados en el genotipo tolerante DT (de forma sobre-expresado), frente al genotipo susceptible DM, en condiciones de presencia de la enfermedad. Como resultado, de los 1040 SNPs previos, se obtuvieron 12 SNPs resultantes, que a su vez representaban 9 genes (es decir, en algunos casos hubo más de 1 SNP para el mismo gen).



Figura 4.3: Hoja de ruta con los criterios establecidos para el filtrado de variantes. El número indica la cantidad de variantes resultantes luego de el filtro especificado en los recuadros principales.

Se obtuvo como consecuencia de los sucesivos filtros aplicados, un panel de 12 marcadores moleculares SNP, que a su vez representan 9 genes distintos (Cuadro 4.2). Esto implica que algunos genes presentaron más de 1 SNP de interés. Dentro de esta caracterización se encuentra tanto el gen Glyma.18G022500 (1 gen, 3 SNPs) como el gen Glyma.20G110900 (1 gen, 2 SNPs). Al intentar agrupar los 9 genes, es posible vislumbrar diferentes categorías o grupos funcionales:

(i) SNPs asociados a genes de tolerancia contra factores bióticos (1 gen, 3 SNPs): Dentro de este grupo se encuentra el gen Glyma.18G022500 (1 gen, 3 SNPs). Este gen cumple un rol preponderante dado que codifica para una proteína SNAP (proteínas solubles de unión a NSF).

Este grupo de proteínas se caracterizan por la presencia de un dominio de repetición de tetratricopéptido (TPR, del inglés *tetratricopeptide repeat*) que es compartido por un gran número de proteínas en diversas especies (D'Andrea y Regan, 2003). Se ha descubierto que las proteínas vegetales que contienen TPR son esenciales para las respuestas a hormonas como el etileno, la citoquinina, la giberelina, el salicilato y la auxina en *Arabidopsis*. Además, se han asignado a las proteínas TPR muchas funciones celulares como el plegamiento de proteínas, el crecimiento y la integridad de las raíces, el transporte de proteínas mitocondriales y las interacciones proteína-proteína. Recientes trabajos publicados demuestran el rol de dicho gen en el control de estreses bióticos tales como el nematodo del nudo en el cultivo de soja (Wei et al. 2022). Resulta indispensable destacar los puntos en contacto entre la adversidad descripta y la PC causada por *M. phaseolina*: ambas son adversidades bióticas de difícil control, tienen como hospedante el cultivo de soja y afectan principalmente el tejido de la base del tallo y sistema radical.

(ii) SNPs asociados a regulación negativa actividad ATPasa (1 gen, 2 SNPs): Dentro de este grupo se encuentra 1 gen representado por 2 SNPs, el Glyma.20G110900. Este gen cumple un rol regulando negativamente la actividad de la ATPasa. Implica entonces, que dichos genes están relacionados con la preservación del ATP dentro de la célula, optimizando el balance energético en la planta. Si bien es posible teorizar que dicho atributo sería potencialmente beneficioso ante situaciones de estrés, el rol de dichos genes no está del todo dilucidado.

(iii) SNPs de respuesta al estrés (5 genes, 5 SNPs): Dentro de este grupo se encuentran 5 genes, totalizando cada uno de ellos 1 sólo SNP. Estos genes se han agrupado dentro de la misma categoría ya que comparten vías de acción relevantes en la mitigación de estreses de diverso índole, incluyendo hídrico, salino y oxidativo, entre otros. Los genes involucrados incluyen al gen Glyma.01G022800 (involucrado en la unión de proteínas y plegamiento de las mismas) y Glyma.07G200200 (comparte cierto rol análogo en el plegamiento de proteínas, además de

detoxificación de ROS, respuesta a estrés hídrico, salino, térmico y de hipoxia celular). Asimismo, dentro de esta categoría se incluye el gen Glyma.15G252000, involucrado en el metabolismo de detoxificación de ROS, glutatión, proceso de catabolismo de toxinas y respuesta celular a la falta de agua. Por último dentro de esta categoría se encuentran los genes Glyma.19G086500 y Glyma.17G205900, involucrados en funciones asociadas a la reserva de nutrientes, uniones entre iones (principalmente manganeso) y respuestas generales frente al estrés.

(iv) SNPs sin función clara identificable (2 genes, 2 SNPs): Esta categoría podríamos definirla como de función desconocida o miscelánea, incluyendo 2 genes principales que a su vez representan cada uno 1 sólo SNP. Los genes involucrados son el Glyma.04G004500 y Glyma.15G199700.

Cuadro 4.2: Listado de los 12 SNPs seleccionados para ser utilizados como marcadores moleculares. ChromPos= Numero y posición del cromosoma en donde se encuentra el SNP, GeneName= ID del gen según la notación "Wm82.a4.v1", REF= alelo de la referencia y DM, ALT= alelo alternativo presente en el DT, EFFECT= efecto del SNP predicho por SNPeff, IMPACT= impacto del SNP predicho por SNPeff, log2FC= Expresión diferencial en genotipo DT vs DM en presencia de la enfermedad, padj= valor p ajustado de la expresión diferencial, GO Biological Process Description= descripción del proceso biológico llevado a cabo por el gen/ proteína según la base de datos Gene Ontology, GO Molecular Function Description= descripción de la función molecular cumplida por el gen/ proteína según la base de datos Gene Ontology.

Chrom	Pos	Gene name (Glyma)	REF	ALT	EFFECT	IMPACT	log2FC	padj	GO Biological Process Descriptions	GO Molecular Process Descriptions
Gm01	2312171	01G022800	G	А	missense_variant	Moderate	2,76	5,25E-21	biological process.	protein binding.
Gm04	387419	04G004500	G	А	missense_variant	Moderate	1,71	0,03	NA	NA
Gm07	37206032	07G200200	Т	A	missense_variant	Moderate	2,67	0,01	response to reactive oxygen species; protein folding; oxidative stress; heat; salt stress; hydrogen peroxide; protein complex oligomerization; cellular response to hypoxia.	Protein binding; protein self-association; unfolded protein binding.
Gm15	24677076	15G199700	G	С	missense_variant	Moderate	1,92	0,01	NA	NA
Gm15	50006550	15G252000	А	Т	stop_gained	High	4,82	2,51E-15	glutathione metabolic process; response to oxidative stress; toxin catabolic process; cellular water deprivation; cadmium ion; oxidation- reduction process.	glutathione transferase activity; peroxidase activity; glutathione binding.
Gm17	33451256	17G205900	А	Т	missense_variant	Moderate	4,62	1,64E-1	cellular manganese ion homeostasis.	NA

Gm18	1641829	18G022500	С	G	missense_variant	Moderate	1,8	4,33E-5	intracellular protein transport; SNARE complex disassembly.	soluble NSF attachment protein activity; syntaxin binding.
Gm18	1643572	18G022500	G	Т	missense_variant	Moderate	1,8	4,33E-5	intracellular protein transport; SNARE complex disassembly.	soluble NSF attachment protein activity; syntaxin binding.
Gm18	1643578	18G022500	С	А	missense_variant	Moderate	1,8	4,33E-5	intracellular protein transport; SNARE complex disassembly.	soluble NSF attachment protein activity; syntaxin binding.
Gm19	31215238	19G086500	С	G	missense_variant	Moderate	0,63	1,09E-14	response to stress	manganese ion binding; nutrient reservoir activity.
Gm20	35276668	20G110900	А	Т	missense_variant	Moderate	1,52	0,01	negative regulation of	ATPase inhibitor
									ATPase activity.	activity.
Gm20	35276679	20G110900	C	G	missense_variant	Moderate	1,52	0,01	negative regulation of	ATPase inhibitor
									ATPase activity.	activity.

### 4. Discusión

La ruta elegida para el filtrado de los SNPs permite obtener como resultado final un panel acotado de marcadores moleculares, teniendo en cuenta la vasta cantidad de variantes iniciales. Es importante destacar que, el panel descripto en el cuadro 4.2 surge de los diferentes criterios adoptados y explicados anteriormente. Esto significa que el panel resultante propuesto no es el único aplicable a un plan de mejoramiento de soja frente a la PC. La resultante de 12 SNPs comprendiendo 9 genes implica un panel diverso pero no inmanejable, que puede significar una óptima alternativa. Otros trabajos llevados a cabo también en el cultivo de soja destacan paneles de amplitudes similares como propuestas de mejoramiento, como es el caso de la mancha roja de la hoja (Lukanda, 2023), donde se identificaron 19 marcadores moleculares de interés. Asimismo, también se han analizado abordajes similares para otras enfermedades tales como el mal de los almácigos ("damping off"), en el cual se obtuvieron 7 SNPs de interés (Lin, 2020). Cabe destacar que, si se cambiara cualquiera de los criterios establecidos en la aplicación de los filtros, el panel resultante sería diferente, e igualmente válido. Ampliar, acotar o modificar el panel propuesto es posible siempre y cuando se encuentren especificados los criterios establecidos. La búsqueda de polimorfismos vía SNPs es una herramienta fundamental, y dada la realización previa del RNA-seq, nos permite aproximar dichos resultados a los genes diferencialmente expresados y enriquecidos. Es interesante como parte de la discusión, analizar si los genes resultantes de la búsqueda de SNPs forman parte de las diferentes vías y categorías enriquecidas tanto en la plataforma GO como KEGG.

Analizando los genes, 2 (dos) de un total de 9 resultaron sobre-expresados en el enriquecimiento funcional. Uno de ellos respecto a la plataforma GO (Glyma.19G086500) y el otro de ellos respecto de KEGG (Glyma.07G200200). El gen presente en la plataforma KEGG está asociado a la vía de procesamiento de proteínas dentro del retículo (gmx041041), específicamente dentro del término gmx:100499727), el cual está asociado a HSP (proteínas de

choque de calor). Dichas proteínas cumplen un rol específico dentro del mapa metabólico, principalmente dentro del mecanismo de degradación de proteínas mal plegadas a través del sistema ERAD (Manghwar, 2022). Por otra parte, el gen presente en la plataforma GO está asociado al incremento de la capacidad de almacenamiento de nutrientes de reserva, además de la contribución a la unión con el manganeso y la respuesta genérica ante el estrés. Los demás genes (7), no estuvieron presentes en los enriquecimientos correspondientes del RNA-seq, pero esto no implica que no tengan relevancia en el aumento de la respuesta de tolerancia de la soja frente a la enfermedad. Como se destacó previamente en este capítulo, muchos de ellos cumplen una función de respuesta frente a los diversos estreses desencadenados en la planta, como respuesta general (Glyma.01G022800, Glyma.15G252000 y Glyma.17G205900), o bien como respuesta más específica a factores bióticos (Glyma.18G022500). Por último, el resto de ellos desempeñaba un rol o bien en funciones generales a nivel energético (Glyma.20G110900) o de función desconocida sin función atribuible (Glyma.04G004500 una clara 0 y Glyma.15G199700).

El panel propuesto en este capítulo puede utilizarse siempre y cuando se valide previamente en poblaciones segregantes. Dicho paso queda fuera del alcance y los objetivos de la presente tesis por diversos motivos, siendo el principal la limitante temporal (extensión en el tiempo de los pasos que requiere la validación). Estos pasos implican, como primer medida, el cruzamiento de los genotipos comentados (susceptible DM y tolerante DT). Asimismo, debería simularse la condición estresante (inoculación, infección y recrear el ciclo de la enfermedad), para finalmente evaluar si el desempeño observado tiene relación con los SNPs seleccionados. Las plantas que evidencien un mejor comportamiento (entendido como una mejor tolerancia frente al patógeno) sería esperable que tuvieran los alelos asociados al genotipo DT dentro de los SNPs.

# 5. Conclusiones

Se establecieron criterios específicos para poder obtener un panel sólido de marcadores moleculares que constituya una propuesta plausible de ser incluida en un programa de mejoramiento de soja. En particular, se eligieron variantes donde el genotipo susceptible DM (tanto en control como en presencia de la enfermedad) presentó la misma variante que la referencia, pero el genotipo DT la alternativa. Seguidamente, se filtró según el tipo de efecto observado en el software SNPeff (efecto moderado o alto). Por último, se vincularon las variantes elegidas hasta dicho momento con los resultados arrojados previamente en el estudio RNA-seq. Este criterio permitió directamente obtener como producto final un panel de SNPs, con efectos posibles (ej.: que se genere un posible cambio que pudiera afectar un marco de lectura o una traducción) y genes identificables, además de función conocida dentro de las diferentes rutas metabólicas existentes dentro de la planta. El panel actual propuesto es manejable, guarda relación con los estudios realizados previamente, es económicamente viable (Yang, 2020) y razonable de ser aplicado en los planes de mejoramiento.

# **Capítulo 5: Conclusiones Generales**

El cultivo de la soja es de los más relevantes, no sólo por la extensión dedicada al mismo, sino también por su aporte al PBI nacional. El cultivo se desarrolla tanto en regiones pampeanas como extra-pampeanas y es la principal fuente de divisas para el país. Uno de los principales desafíos en la actualidad son el conjunto de adversidades tanto bióticas como abióticas que padece el cultivo. Las enfermedades, dada su complejidad y por los daños que ocasionan, han sido objeto de diversos estudios, siendo la finalidad encontrar alternativas de control a las ya existentes. En ese contexto, la PC de la soja es una de los principales adversidades a atender, dado que suele estar acompañado de eventos de sequía, cada vez más usuales y frecuentes en la Argentina. La generación de variedades o materiales de soja con buen desempeño o tolerancia frente a la enfermedad es la alternativa más relevante.

Se trabajó inicialmente con diversos genotipos de soja de comportamiento presumiblemente diferente e inóculos del patógeno. Como primer medida, se analizaron los inóculos para trabajar con aquel que fuera más agresivo frente a la soja. Asimismo, se diseñaron ensayos en invernáculo con el objetivo de poder caracterizar el comportamiento de cada uno de esos genotipos frente a la enfermedad. Se planearon protocolos pertinentes tanto de inoculación (Cassina, 2017) como de simulación de condiciones hídricas estresantes (Mian et al. 1996). A partir de dicho estudio, se efectuaron mediciones de índole epidemiológica: severidad, tasa de infección y ABCPE, entre otras. Los análisis efectuados permitieron distinguir el desempeño de cada uno de los genotipos, además de definir el momento de cosecha propicio para los tejidos, que serían necesarios para estudios posteriores. De esta forma se dio cumplimiento al primero de los objetivos.

Asimismo, de dichos genotipos, 2 evidenciaron un comportamiento contrastante (DM-susceptible) y (DT-tolerante). Una vez seleccionados dichos genotipos contrastantes, se realizaron ensayos complementarios diferentes, tanto a nivel de plántula (longitud y peso fresco), hojas (lesiones en cotiledón, hojas unifoliadas y trifoliadas) y semillas (severidad). Se pudo verificar y aceptar la hipótesis inicial que enunciaba que existen genotipos de soja que evidencian respuestas contrastantes a la PC y que las sucesivas mediciones epidemiológicas permitirían vislumbrar diferencias significativas, e indicarían, además, el momento propicio para recolectar el tejido a partir de plántulas que se utilizarán en el estudio RNA-seq.

La selección de genotipos contrastantes resultó el insumo fundamental para a partir de allí entender la respuesta a nivel génico. El RNA-seq permitió una caracterización de los genes y vías metabólicas responsables de la respuesta diferencial observada entre ambos genotipos. Principalmente, se observó una subexpresión en el genotipo tolerante de genes asociados al metabolismo basal o funciones primarias (como las de apoplasto, periferia celular y pared celular). Como contracara, se evidenció una sobreexpresión de genes asociados a optimización de los niveles de reserva, unión de proteínas, y unión de proteínas no plegadas. Asimismo, se vislumbró la participación de genes en las vías metabólicas de interacción planta-patógeno y de procesamiento de proteínas a nivel de retículo, que desempeñan funciones clave para la mitigación del estrés producido. De esta forma se dio cumplimiento al segundo de los objetivos y se aceptó la segunda hipótesis que enunciaba que era posible detectar genes diferencialmente expresados a través de un análisis transcriptómico y asociar este comportamiento a procesos bioquímicos relacionados.

Por último, se realizó la búsqueda de polimorfismos en la secuencia transcriptómica que estuvieran asociados a la respuesta del genotipo tolerante frente a la enfermedad. Se utilizaron herramientas bioinformáticas para ejecutar el variant calling y la predicción de efectos de dichas
variantes vía SnpEff. Se establecieron diferentes criterios que permitieron obtener un panel de SNPs que pueden formar parte de un plan de mejoramiento. Se llegó a un panel de 12 SNPs, involucrando 9 genes únicos. Todos evidencian un efecto moderado o alto según el efecto predicho por SnpEff. Dichos genes están relacionados específicamente a la resistencia a factores bióticos y la mitigación de estrés, como lo constituye por ejemplo la mitigación de ROS y la codificación de proteínas SNAP, responsables del plegamiento proteico y la respuesta hormonal. De esta forma se dio cumplimiento al tercero y último de los objetivos planteados. Se aceptó la última hipótesis que enunciaba la existencia de un polimorfismo estructural en los genes mencionados, a partir del cual será posible desarrollar marcadores moleculares.

El abordaje de la podredumbre carbonosa en soja constituye un gran desafío, entre diversos factores, por su presencia en condiciones de estrés hídrico (a diferencia de las principales enfermedades de soja) y la inexistencia actual de cultivares con resistencia genética. En este sentido, haber desarrollado un modelo de evaluación a nivel fenotípico cobra mucha relevancia. Por un lado, permite realizar un testeo rápido y fidedigno a ser aplicado en planes de mejoramiento. Por otro lado, constituye una herramienta eficaz para caracterizar la respuesta a la enfermedad de genotipos o variedades que resulten de interés en futuros estudios. La evaluación de la enfermedad aplicando las condiciones estresantes asociadas (como la sequía), junto con una metodología eficaz de inoculación, suma robustez al testeo mencionado.

Una de las principales contribuciones de la presente tesis radica en la generación de información novedosa obtenida a partir del transcriptoma de genotipos contrastantes de soja respecto a la respuesta ante *M. phaseolina*, que permitió no solo la identificación de genes clave, sino también conocer el rol que cumplen parte de ellos dentro de vías metabólicas de relevancia para el cultivo (base bioquímica y de genómica funcional detrás). Conocer los mecanismos que operan detrás del comportamiento de genotipos de buen desempeño frente al patógeno es fundamental ya que permite además agregar información valiosa al conocimiento del

patosistema de estudio.

Los programas de soja actuales dependen, en parte, del uso de marcadores moleculares para mejorar las eficiencias de selección, y para asegurar la transferencia precisa de genes entre variedades. En este sentido, panel de SNPs propuesto en la presente tesis es de singular importancia y de un aporte indispensable, dado que constituye un panel sólido y acotado, a partir del cual los mejoradores podrían diseñar estrategias a ser implementadas en futuros programas.

No obstante, pensando a futuro, es posible seguir trabajando con los resultados hasta aquí obtenidos. Principalmente, vía la validación de los marcadores moleculares propuestos (por ejemplo, a través de poblaciones segregantes). Si bien es una alternativa que requiere esfuerzo y tiempo, también existen otras alternativas más ágiles para cumplir dicho objetivo, como por ejemplo validar los marcadores utilizando una población comercial (donde se conocen de antemano los comportamientos diferenciales de cada uno de los genotipos frente a la enfermedad).

A modo de conclusión, es importante reforzar que el aporte generado hasta aquí puede ser de utilidad para mejorar el abordaje de una de las principales problemáticas actuales referidas al manejo del cultivo de soja en Argentina. Asimismo, sirve como punto de partida para futuros trabajos, como por ejemplo, la validación de los marcadores moleculares aquí propuestos y/o el desarrollo de nuevos marcadores, que puedan llegar a sumarse a los ya propuestos. Por otra parte, existen otras posibilidades de estudio a partir de la profundización del estudio en alguna/s de las vías metabólicas aquí presentadas como de relevancia, o la redefinición del panel de marcadores a partir de la modificación de los criterios de filtro aplicados a la búsqueda de variante

## **BIBLIOGRAFÍA**

- Abawi, G. S., y Pastor-Corrales, M. A. (1990). Root rots of beans in Latin America and Africa: diagnosis, research methodologies, and management strategies. Cali, Colombia: Centro Internacional de Agricultura Tropical (CIAT).
- Adeyanju, A., Little, C., Yu, J. M., y Tesso, T. (2015). Genome-wide association study on resistance to stalk rot diseases in grain sorghum. G3, 5, 1165–1175.
- Almeida, Á. M., Abdelnoor, R. V., Arias, C. A. A., Carvalho, V. P., Jacoud Filho, D. S., Marin, S. R., Carvalho, C. G. (2003). Genotypic diversity among Brazilian isolates of *Macrophomina phaseolina* revealed by RAPD. Fitopatologia Brasileira, 28, 279-285.
- Almeida, A. M., Ferreira Saravia, O., Farias, J. R. B., Almeida, C., y Torres, E. (2001). Survival of pathogens on soybean debris under no-tillage and conventional tillage systems. Pesquisa Agropecuaria Brasilera, 36, 1231-1238.
- Almeida-Silva, F., y Venancio, T. M. (2021). Integration of genome-wide association studies and gene coexpression networks unveils promising soybean resistance genes against five common fungal pathogens. Scientific Reports, 11(1), 24453.
- Amiteye, S. (2021). Basic concepts and methodologies of DNA marker systems in plant molecular breeding. Heliyon, 7(10).
- An, J., Kim, S. H., Bahk, S., Le Anh Pham, M., Park, J., Ramadany, Z., ... Chung, W. S. (2023). Quercetin induces pathogen resistance through the increase of salicylic acid biosynthesis in Arabidopsis. Plant Signaling y Behavior, 18(1), 2270835.
- Anders, S., y Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology.
- Babu, B. S., Pandravada, S. R., Rao, R. P., Anitha, K., Chakrabarty, S. K., y Varaprasad, K. S. (2011). Global sources of pepper genetic resources against arthropods, nematodes and pathogens. Crop Protection, 30(4), 389-400.
- Badana, M. K., Kelaiya, D. S., y Jadav, A. H. (2021). Screening of groundnut varieties and different types of host species against dry root rot disease.
- Bandara, A. Y., Weerasooriya, D. K., Bradley, C. A., Allen, T. W., y Esker, P. D. (2020). Dissecting the economic impact of soybean diseases in the United States over two decades. PLoS ONE, 15(4), e0231141.
- Bañuelos-Balandrán, J. J., y Mayek-Pérez, N. (2008). Evaluación no Destructiva de la Patogenicidad de *Macrophomina phaseolina* (Tassi) Goid. en Frijol (*Phaseolus vulgaris* L.).

- Bellaloui, N., Mengistu, A., y Paris, R. L. (2008). Soybean seed composition in cultivars differing in resistance to charcoal rot (*Macrophomina phaseolina*). The Journal of Agricultural Science, 146(6), 667-675.
- Bellaloui, N., Mengistu, A., Zobiole, L., Abbas, H., y Kassem, M. (2014). Soybean Seed Phenolics, Sugars, and Minerals Are Altered by Charcoal Rot Infection in MG III Soybean Cultivars. Food and Nutrition Sciences, 5, 1843-1859.
- Boller, T., y He, S. Y. (2009). Effectors in microbial pathogens pattern recognition receptors in plants and innate immunity in plants: An arms race between. Science, 324, 742-44.
- Bolsa de Comercio de Rosario, 2024. https://www.bcr.com.ar/es/mercados/gea/estimacionesnacionales-de-produccion/estimaciones
- Blume, B., Nürnberger, T., Nass, N., y Scheel, D. (2000). Receptor-mediated increase in cytoplasmic free calcium required for activation of pathogen defense in parsley. Plant Cell, 12, 1425-1440.
- Bressano, M., Curetti, M., Giachero, L., Gil, S. V., Cabello, M., March, G., ... Luna, C. M. (2010). Mycorrhizal fungi symbiosis as a strategy against oxidative stress in soybean plants. Journal of Plant Physiology, 167(18), 1622-1626.
- Campbell, C. L., y Madden, L. V. (1990). Introduction to Plant Disease Epidemiology. New York: John Wiley y Sons.
- Cardona, R. (2006). Distribución vertical de esclerocios de *Macrophomina phaseolina* en un suelo infestado naturalmente en el estado Portuguesa. Revista Facultad Agronomía [online], 23(3).
- Carmona M.A, Gally M.E, Grijalba P.E, Sautua F.J. 2015. Evolución de las enfermedades de la soja en la Argentina: pasado, y presente. Aportes de la FAUBA al manejo integrado. Agronomía y Ambiente. Rev. Facultad de Agronomía UBA 35(1): 37-52.
- Cassina, M. (2017). Utilización de fosfito de manganeso y potasio como tratamiento químico alternativo de semillas de soja tanto in vivo como in vitro para el control de tizón de plántulas: causadas por diferentes especies de *Pythium*. [Tesis de Grado, Universidad de Buenos Aires. Facultad de Agronomía].
- Cheval, C., Aldon, D., Galaud, J. P., y Ranty, B. (2013). Calcium/calmodulin-mediated regulation of plant immunity. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1833(7), 1766-1771.
- Choi, H. W., y Hwang, B. K. (2012). Molecular and cellular control of cell death and defense signaling in pepper. Planta, 241, 1-27.
- Chopra, R., Burow, G., Hayes, C., et al. (2015). Transcriptome profiling and validation of genebased single nucleotide polymorphisms (SNPs) in sorghum genotypes with contrasting

responses to cold stress. BMC Genomics, 16, 1040.

- Chowdhury, S., Basu, A., y Kundu, S. (2017). Biotrophy-necrotrophy switch in pathogen evoke differential response in resistant and susceptible sesame involving multiple signaling pathways at different phases. Scientific Reports, 7(1), 1-17.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... y Ruden, D. M. (2012).
  A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly, 6(2), 80-92.
- Cirilli, M., Rossini, L., Geuna, F., et al. (2017). Genetic dissection of Sharka disease tolerance in peach (P. persica L. Batsch). BMC Plant Biology, 17, 192.
- Coser, S. M., et al. (2017). Genetic architecture of charcoal rot (*Macrophomina phaseolina*) resistance in soybean revealed using a diverse panel. Frontiers in Plant Science, 8, 1626.
- Cuevas Moreno, J. A. (2017). Análisis de los perfiles de expresión de genes involucrados en la patogenicidad del hongo *Macrophomina phaseolina Tassi* (Goid.) en distintos hospederos. [Master's thesis, Instituto Politécnico Nacional, Centro de Biotecnología Genómica, México].
- D'Andrea, L. D., y Regan, L. (2003). TPR proteins: The versatile helix. Trends in Biochemical Sciences, 28(12), 655-662.
- Da Silva, M. P., Klepadlo, M., Gbur, E. E., Pereira, A., Mason, R. E., Rupe, J. C., ... y Chen, P. (2019). QTL mapping of charcoal rot resistance in PI 567562A soybean accession. Crop Science, 59(2), 474-479.
- Deng, Y., Srivastava, R., y Howell, S. H. (2013). Endoplasmic reticulum (ER) stress response and its physiological roles in plants. International Journal of Molecular Sciences, 14(4), 8188-212.
- Deshmukh, R., y Tiwari, S. (2021). Molecular interaction of charcoal rot pathogenesis in soybean: A complex interaction. Plant Cell Reports, 40(10), 1799-1812.
- Di Rienzo, J., Balzarini, M., Gonzalez, L., Casanoves, F., Tablada, M., y Walter Robledo, C. (2010). Infostat: software para análisis estadístico.
- Dutta, D., Awon, V. K., y Gangopadhyay, G. (2020). Transcriptomic dataset of cultivated (Sesamum indicum), wild (S. mulayanum), and interspecific hybrid sesame in response to induced *Macrophomina phaseolina* infection. Data in Brief, 33, 106448.
- Erlich, Y., Mitra, P. P., de la Bastide, M., et al. (2008). Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. Nature Methods, 5(8), 679-682.

Escande, A.; 2005. Simposios. Enfermedades de soja, girasol, maíz y trigo en sistemas sustentables

(siembra directa). Manejo de la sanidad del girasol en sistemas sustentables (siembra directa) - En: XIII Congreso Latinoamericano de Fitopatología. III Taller de la Asociación Argentina de Fitopatólogos. Año 2005. Córdoba, Argentina.. - Páginas/s: 639

- Fehr, W. R., y Caviness, C. E. (1977). Stages of soybean development. Special Report, (80). Ames, Iowa: Iowa State University of Science and Technology, Cooperative Extension Service.
- Gaige, A. R., Ayella, A., y Shuai, B. (2010). Methyl jasmonate and ethylene induce partial resistance in *Medicago Truncatula* against the charcoal rot pathogen *Macrophomina phaseolina*. Physiological and molecular plant pathology, 74(5–6), 412–418.
- Gill-Langarica, H., Maldonado-Moreno, N., Quintero, V., y Mayek-Pérez, N. (2007). Reacción de Germoplasma Mejorado de Soya [Glycine max (L.) Merr.] a *Macrophomina phaseolina* (Tassi) Goidanich y Déficit Hídrico. Revista Mexicana de Fitopatología, 26, 105-113.
- Giovanardi, D., Biondi, E., Ignjatov, M., Jevtić, R., y Stefani, E. (2018). Impact of bacterial spot outbreaks on the phytosanitary quality of tomato and pepper seeds. Plant Pathology, 67(5), 1168-1176.
- Gene Ontology. (n.d.). Retrieved from http://www.geneontology.org/
- González-Domínguez, E., Caffi, T., Rossi, V., Salotti, I., y Fedele, G. (2023). Plant disease models and forecasting: changes in principles and applications over the last 50 years. Phytopathology®, 113(4), 678-693.
- Guan, L. M., Zhao, J., y Scandalios, J. G. (2000). Cis-elements and trans-factors that regulate expression of the maize Cat1 antioxidant gene in response to ABA and osmotic stress: H2O2 is the likely intermediary signaling.
- Gutierrez-Gonzalez, J. J., Guttikonda, S. K., Tran, L. S., Aldrich, D. L., Zhong, R., Yu, O., Nguyen,
  H. T., y Sleper, D. A. (2010). Differential expression of isoflavone biosynthetic genes in soybean during water deficits. Plant Cell Physiology, 51, 936–948.
- Haro Sly, M. (2017). The Argentine portion of the soybean commodity chain. Palgrave Communications, 3, palcomms201795.
- Hartmann, A., Berkowitz, O., Whelan, J., et al. (2022). Cross-species transcriptomic analyses reveal common and opposite responses in Arabidopsis, rice, and barley following oxidative stress and hormone treatment. BMC Plant Biology, 22(1), 62.
- He, P., Shan, L., y Sheen, J. (2007). Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant-microbe interactions. Cellular Microbiology, 9(6), 1385-1396.
- Hosseini, B., Voegele, R. T., y Link, T. I. (2023). Diagnosis of soybean diseases caused by fungal and oomycete pathogens: Existing methods and new developments. Journal of Fungi, 9(5), 587.

- Howell, S. H. (2013). Endoplasmic reticulum stress responses in plants. Annual Review of Plant Biology, 64, 477-499.
- Hussan, R. H., Dubery, I. A., y Piater, L. A. (2020). Identification of MAMP-responsive plasma membrane-associated proteins in Arabidopsis thaliana following challenge with different LPS chemotypes from *Xanthomonas campestris*. Pathogens, 9(10), 787.
- Iriti, M., Vannini, C., Carravieri, S., Genga, A., Bracale, M., y Faoro, F. (2007). The rice Osmyb4 gene modulates the resistance to tomato mosaic virus in transformed tomato plants improving the quality traits of fruits. In Abs XIII Int Congress on Molecular Plant-Microbe Interactions (p. 282). Sorrento.
- Ivancovich, A. (2011). Enfermedades de Soja: Diagnóstico y manejo [Soybean diseases: Diagnosis and management]. Inta Pergamino.
- Jacquier, A. (2009). The complex eukaryotic transcriptome: Unexpected pervasive transcription and novel small RNAs. Nature Reviews Genetics, 10(12), 833-844.
- Jiang, L., Schlesinger, F., Davis, C. A., et al. (2011). Synthetic spike-in standards for RNA-seq experiments. Genome Research, 21(9), 1543-1551.
- Johnson, S. M., Lim, F. L., Finkler, A., Fromm, H., Slabas, A. R., y Knight, M. R. (2014). Transcriptomic analysis of Sorghum bicolor responding to combined heat and drought stress. BMC Genomics, 15(1), 1-19.
- Jones, J. D., y Dangl, J. L. (2006). The plant immune system. Nature, 444, 323-329.
- Kanehisa, M., y Goto, S. (2008). KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research, 28(1), 27-30.
- Kantolic, A. G., y Carmona, M. A. (2006). Bases ecofisiológicas de la generación de rendimiento: relación con el efecto de las enfermedades foliares y el uso de fungicidas en el cultivo de soja. Buenos Aires: FAUBA.
- Kaur, S., Dhillon, G. S., Brar, S. K., Vallad, G. E., Chand, R., y Chauhan, V. B. (2012). Emerging phytopathogen *Macrophomina phaseolina*: biology, economic importance and current diagnostic trends. Critical Reviews in Microbiology, 38, 136–151.
- Kaur, S., Kumari, N., Parmar, P., y Sharma, V. (2023). Simultaneous exposure of contrasting *Vigna radiata* cultivars to *Macrophomina phaseolina* infection and drought stress: morphophysiological and biochemical implications. Plant Biosystems An International Journal Dealing with all Aspects of Plant Biology, 157(5), 1100-1113.
- KEGG: Enciclopedia de Genes y Genomas de Kioto. (s. f.). Recuperado de http://www.kegg.jp/

Kendig, S. R., Rupe, J. C., y Scott, H. D. (2000). Effect of irrigation and soil water stress on

densities of *Macrophomina phaseolina* in soil and roots of two soybean cultivars. Plant Disease, 84(8), 895-900.

- Kørner, C. J., Du, X., Vollmer, M. E., y Pajerowska-Mukhtar, K. M. (2015). Endoplasmic reticulum stress signaling in plant immunity--At the crossroad of life and death. International Journal of Molecular Sciences, 16(11), 26582-98.
- Klusner, B., Young, J. J., Murata, Y., Allen, G. J., Mori, I. C., Hugovieux, V., y Schroeder, J. I. (2002). Convergence of calcium signaling pathways of pathogenic elicitors and abscisic acid in Arabidopsis guard cells. Plant Physiology, 130, 2152-2163.
- Knight, H., Trewavas, A. J., y Knight, M. R. (1997). Calcium signalling in *Arabidopsis thaliana* responding to drought and salinity. The Plant Journal, 12, 1067-1078.
- Laloi, C., Apel, K., y Danon, A. (2004). Reactive oxygen signaling: The latest news. Current Opinion in Plant Biology, 7, 323-328.
- Le, D. T., Nishiyama, R., Watanabe, Y., Mochida, K., Yamaguchi-Shinozaki, K., Shinozaki, K., y Tran, L. S. (2011). Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. DNA Research, 18, 263–276.
- Leonforte, A., Sudheesh, S., Cogan, N. O., et al. (2013). SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (*Pisum sativum* L.). BMC Plant Biology, 13, 161.
- Li, A. L., Zhu, Y. F., Tan, X. M., Wang, X., Wei, B., Guo, H. Z., ... y Mao, L. (2008). Evolutionary and functional study of the CDPK gene family in wheat (*Triticum aestivum* L.). Plant Molecular Biology, 66, 429-443.
- Li, X., et al. (2023). Acidification suppresses the natural capacity of soil microbiome to fight pathogenic Fusarium infections. Nature Communications, 14(1), 5090.
- Liao, Y., Smyth, G. K., y Shi, W. (2013). featureCounts: An efficient general-purpose program for assigning sequence reads to genomic features. Bioinformatics.
- Liener, I. E. (1994). Implications of antinutritional components in soybean foods. Critical Reviews in Food Science and Nutrition, 34(1), 31-67.
- Lin, F., Wani, S. H., Collins, P. J., Wen, Z., Li, W., Zhang, N., ... y Wang, D. (2020). QTL mapping and GWAS for identification of loci conferring partial resistance to *Pythium sylvaticum* in soybean (*Glycine max* (L.) *Merr*). Molecular Breeding, 40, 1-11.
- Liu, J. Z., y Whitham, S. A. (2013). Overexpression of a soybean nuclear localized type-III DnaJ domain-containing HSP40 reveals its roles in cell death and disease resistance. The Plant Journal, 74, 110-121.

- Livaja, M., Wang, Y., Wieckhorst, S., et al. (2013). BSTA: A targeted approach combines bulked segregant analysis with next-generation sequencing and de novo transcriptome assembly for SNP discovery in sunflower. BMC Genomics, 14, 628.
- Lukanda, M. M., Dramadri, I. O., Adjei, E. A., Badji, A., Arusei, P., Gitonga, H. W., ... y Tusiime, G. (2023). Genome-wide association analysis for eesistance to *Coniothyrium glycines* causing red leaf blotch disease in soybean. Genes, 14(6), 1271.
- Ma, W., y Berkowitz, G. A. (2007). The grateful dead: calcium and cell death in plant innate immunity. Cell Microbiol, 9(11), 2571-2585.
- Manghwar, H., y Li, J. (2022). Endoplasmic reticulum stress and unfolded protein response signaling in plants. International Journal of Molecular Sciences, 23(2), 828.
- Manici, L. M., Caputo, F., y Cerato, C. (1995). Temperature responses of isolates of *Macrophomina phaseolina* from different climatic regions of sunflower production in Italy. Plant Disease.
- Márquez, N., Giachero, M. L., Declerck, S., y Ducasse, D. A. (2021). *Macrophomina phaseolina*: General characteristics of pathogenicity and methods of control. Frontiers in Plant Science, 12, 634397.
- Márquez, N., Giachero, M. L., Gallou, A., Debat, H. J., Cranenbrouck, S., Di Rienzo, J. A., ... y Declerck, S. (2018). Transcriptional changes in mycorrhizal and nonmycorrhizal soybean plants upon infection with the fungal pathogen *Macrophomina phaseolina*. Molecular Plant-Microbe Interactions, 31(8), 842-855.
- Mauch-Mani, B., y Mauch, F. (2005). The role of abscisic acid in plant-pathogen interactions. Current Opinion in Plant Biology, 8(4), 409-414.
- McKenna, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20(9), 1297-1303.
- Mengiste, T., Chen, X., Salmeron, J., y Dietrich, R. (2003). The BOTRYTIS SUSCEPTIBLE1 gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in Arabidopsis. The Plant Cell, 15(11), 2551-2565.
- Mengistu, A., Bond, J., Nelson, R. L., Rupe, J., Shannon, G., Arelli, P. R., y Wrather, A. (2011). Identification of soybean accessions resistant to *Cercospora sojina* by field screening, molecular markers, and phenotyping. Crop Science, 51, 1101-1109.
- Mengistu, A., Bond, J., Nelson, R. L., Rupe, J. C., Shannon, G., Arelli, P. R., y Wrather, A. (2013). Identification of soybean accessions resistant to *Macrophomina phaseolina* by field screening and laboratory validation. Plant Health Progress.
- Mengistu, A., Ray, J. D., Smith, J. R., y Paris, R. L. (2007). Charcoal rot disease assessment of soybean genotypes using a colony-forming unit index. Crop Science, 47, 2453.

- Mengistu, A., Wrather, A., y Rupe, J. C. (2015). Charcoal rot. In G. L. Hartman, J. C. Rupe, E. F. Sikora, L. L. Domier, J. A. Davis, y K. L. Steffey (Eds.), Compendium of Soybean Diseases and Pests. St. Paul, MN: American Phytopathological Society Press.
- Mian, M. A. R., et al. (1996). Molecular markers associated with water use efficiency and leaf ash in soybean. Crop Science, 36(5), 1252-1257.
- Nair, R. M., Boddepalli, V. N., Yan, M. R., Kumar, V., Gill, B., Pan, R. S., ... y Somta, P. (2023). Global status of vegetable soybean. Plants, 12(3), 609.
- Oerke, E. C. (2006). Crop losses to pests. The Journal of agricultural science, 144(1), 31-43.
- Pandey, P., Ramegowda, V., y Senthil-Kumar, M. (2015). Shared and unique responses of plants to multiple individual stresses and stress combinations: physiological and molecular mechanisms. Frontiers in Plant Science, 6.
- Panisson, F. T. (2019). Comparação de métodos de fenotipagem em soja e variabilidade genética de *Macrophomina phaseolina*.
- Paris, R. L., Mengistu, A., Tyler, J. M., y Smith, J. R. (2006). Registration of soybean germplasm line DT97-4290 with moderate resistance to charcoal rot. Crop Science, 46(5), 2324.
- Park, C. J., y Seo, Y. S. (2015). Heat Shock Proteins: A Review of the Molecular Chaperones for Plant Immunity. Plant Pathology Journal, 31(4), 323-333.
- Pawlowski, M. L., Hill, C. B., y Hartman, G. L. (2015). Resistance to Charcoal Rot Identified in Ancestral Soybean Germplasm. Crop Science, 55(3), 1230–1235.
- Pérez Brandán, C., Díaz, C., Carmona, M., y March, G. (2009). Influencia de las precipitaciones y las temperaturas sobre la intensidad de la podredumbre carbonosa de la soja (*Macrophomina phaseolina*) en Salta (Argentina). In Soja Actualización 2009 EEA Marcos Juárez (pp. 37-40). INTA.
- Pickel, B., Dai, N., Maymon, M., Elazar, M., Tanami, Z., Frenkel, O., ... y Freeman, S. (2020). Development of a reliable screening technique for determining tolerance to *Macrophomina phaseolina* in strawberry. European Journal of Plant Pathology, 157, 707-718.
- Poudel, B. y Vaghefi, N. (2023). Taxonomy of *Macrophomina*—traditional to molecular approaches in *Macrophomina Phaseolina* (pp. 3-8). Academic Press.
- Rabello, A. R., Guimaraes, C. M., Rangel, P. H., da Silva, F. R., Seixas, D., de Souza, E., ... y Mehta, A. (2008). Identification of drought-responsive genes in roots of upland rice (*Oryza* sativa L). BMC Genomics, 9(485).
- Radadiya, N., Mangukia, N., Antala, V., Desai, H., Chaudhari, H., Dholaria, T. L., ... y Mahatma, M. K. (2021). Transcriptome analysis of sesame-*Macrophomina phaseolina* interactions revealing the distinct genetic components for early defense responses. Physiology and

Molecular Biology of Plants, 27(8), 1675–1693.

- Radwan, O., Rouhana, L. V., Hartman, G. L., y Korban, S. S. (2014). Genetic mechanisms of hostpathogen interactions for charcoal rot in soybean. Plant Molecular Biology Reporter, 32(3), 617–629.
- Ramezani, M., Shier, W. T., Abbas, H. K., Tonos, J. L., Baird, R. E., et al. (2007). Soybean charcoal rot disease fungus *Macrophomina phaseolina* in Mississippi produces the phytotoxin (-)-botryodiplodin but no detectable phaseolinone. Journal of Natural Products, 70, 128-129.
- Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., y Mittler, R. (2004). When defense pathways collide: The response of Arabidopsis to a combination of drought and heat stress. Plant Physiology, 134(4), 1683-1696.
- Rizzi, M. (2021). Evaluación de las mermas de rendimiento del cultivo de Soja (Glycine max L.) causadas por enfermedades de raíz y tallo en el partido de Pergamino, Buenos Aires.
- Reznikov, S., et al. (2019). Soybean-*Macrophomina phaseolina*-specific interactions and identification of a novel source of resistance. Phytopathology, 109(1), 63-73.
- Sanches, A. C. N. (2020). Qualidade de sementes de plantas de soja infectadas com *Macrophomina phaseolina* (Tassi) *Goid*. em estádios fenológicos reprodutivos.
- Sarowar, S., Kim, Y. J., Kim, E. N., Kim, K. D., Hwang, B. K., Islam, R., y Shin, J. S. (2005). Overexpression of a pepper basic pathogenesis-related protein 1 gene in tobacco plants enhances resistance to heavy metal and pathogen stresses. Plant Cell Reports, 24, 216-224.
- Sarr, M., Ndiaye, M., Groenewald, J., y Crous, P. (2014). Genetic diversity in *Macrophomina phaseolina*, the causal agent of charcoal rot. Phytopathologia Mediterranea, 53, 250–268.
- Scandiani, M. M., Carmona, M. A., Luque, A. G., Matos, K. D. S., Lenzi, L., Formento, Á. N., ... y Sautua, F. (2012). Aislamiento, identificación y daños asociados al síndrome de la muerte súbita en el cultivo de soja en Argentina. Tropical Plant Pathology, 37, 358-362.
- Shi, Z., Liu, S., Noe, J., Arelli, P., Meksem, K., y Li, Z. (2015). SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC genomics, 16, 1-12."
- Simoni, E. B., et al. (2022). Cell death signaling from endoplasmic reticulum stress: Plant-specific and conserved features. Frontiers in Plant Science, 13, 835738.
- Smith, D., Chilvers, M., Dorrance, A., Hughes, T., Mueller, D., Niblack, T., y Wise, K. (2014). Charcoal rot management in the north central region. University of Wisconsin Extension Bulletin A4037.

Smith, G. S., y Carvill, O. N. (1997). Field screening of commercial and experimental soybean

cultivars for their reaction to Macrophomina phaseolina. Plant Disease, 81, 363-368.

- Smith, G., y Wyllie, T. (1999). Charcoal rot. In Compendium of Soybean Disease, 4th Edn. St. Paul (Minnesota, MN): American Phytopathological Society.
- Song, Q., Jia, G., Zhu, Y., Grant, D., Nelson, R. T., Hwang, E. Y., ... y Cregan, P. B. (2010). Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR\_1.0) in soybean. Crop Science, 50(5), 1950-1960.
- Spollen, W. G., Tao, W., Valliyodan, B., Chen, K., Hejlek, L. G., Kim, J. J., ... y Nguyen, H. T. (2008). Spatial distribution of transcript changes in the maize primary root elongation zone at low water potential. BMC Plant Biology, 8(32).
- Srivastava, A., Gupta, A. K., Sarkar, S., Lal, R. K., Yadav, A., Gupta, P., y Chanotiya, C. S. (2018). Genetic and chemotypic variability in basil (*Ocimum basilicum* L.) germplasm towards future exploitation. Industrial Crops and Products, 112, 815-820.
- St. Clair, D. A. (2010). Quantitative disease resistance and quantitative resistance loci in breeding. Annual Review of Phytopathology, 48, 247-268.
- Sun, H., Meng, M., Yan, Z., Lin, Z., Nie, X., y Yang, X. (2019). Genome-wide association mapping of stress-tolerance traits in cotton. The Crop Journal, 7(1).
- Trapnell, C., et al. (2010). Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology, 28(5), 511–515.
- Twizeyimana, M., Hill, C. B., Pawlowski, M., Paul, C., y Hartman, G. L. (2012). A cut-stem inoculation technique to evaluate soybean for resistance to *M. phaseolina*. Plant Disease, 96(8), 1210–1215.
- Upadhyay, Shikha, et al. (2022). Biotechnological interventions to combat against charcoal rot and Rhizoctonia root rot diseases of soybean (*Glycine max* (L.) *Merrill*). In Current Topics in Agricultural Sciences, 6, 1–18.
- Vaughn, L. M., y Nguyen, H. T. (2012). The effects of moisture extremes on plant roots and their connections with other abiotic stresses. In M. Crespi (Ed.), Root Genomics and Soil Interactions.
- Wan, J., Zhang, X. C., Neece, D., et al. (2008). A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. The Plant Cell, 20(2), 471–481.
- Wang, D., Pan, Y., Zhao, X., Zhu, L., Fu, B., y Li, Z. (2011). Genome-wide temporal-spatial gene expression profiling of drought responsiveness in rice. BMC Genomics, 12, 149.
- Wang, Y., Li, X., Fan, B., Zhu, C., y Chen, Z. (2021). Regulation and function of defense-related callose deposition in plants. International Journal of Molecular Sciences, 22(5), 2393.

- Wen, G. (2017, September). A simple process of RNA-sequence analyses by Hisat2, Htseq and DESeq2. In Proceedings of the 2017 International Conference on Biomedical Engineering and Bioinformatics (pp. 11–15).
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., ... y Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation, 2(3).
- Wyllie, T. D. (1988). Charcoal rot of soybean: Current status. In T. D. Wyllie y D. H. Scott (Eds.), Soybean Diseases of the North Central Region (pp. 106–113). St. Paul, MN: American Phytopathological Society.
- Yadav, C. B., Bhareti, P., Muthamilarasan, M., Mukherjee, M., Khan, Y., Rathi, P., y Prasad, M. (2015). Genome-wide SNP identification and characterization in two soybean cultivars with contrasting mungbean yellow mosaic India virus disease resistance traits. *PloS one*, 10(4).
- Yang, L., Zheng, B., Mao, C., Qi, X., Liu, F., y Wu, P. (2004). Analysis of transcripts that are differentially expressed in three sectors of the rice root system under water deficit. Molecular Genetics and Genomics, 272, 433–442.
- Yang, T., Miller, M., Forgacs, D., Derr, J., y Stothard, P. (2020). Development of SNP-based genomic tools for the Canadian bison industry: Parentage verification and subspecies composition. Frontiers in Genetics, 11, 585999.
- Yu, G., Wang, L., Han, Y., y He, Q. (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS: A Journal of Integrative Biology, 16(5), 284–287.
- Zhang, Y., Cao, G., Qu, L. J., y Gu, H. (2009). Characterization of Arabidopsis MYB transcription factor gene AtMYB17 and its possible regulation by LEAFY and AGL15. Journal of Genetics and Genomics, 36(2), 99–107.

## APÉNDICE

Temperaturas diarias del ensayo preliminar (entre el 8 marzo y 26 abril)



Figura 3.6: Gel de agarosa para cada una de las muestras (1.5%)

DM		12	37	8	22	26	15	30	34	-9	17	24
	-	-	-		-	-	-	-	-		-	-
					•		-	5			1000	
DT												
				-	-	1			1			
					-			-		-	-	-

## Figura 3.7: Pasos secuenciación Illumina



Gene_ID	log2FoldChange	pvalue	padj
Glyma.01G016700.1.Wm82.a4.v1	6,8381	5,18E-24	5,25E-21
Glyma.01G021000.1.Wm82.a4.v1	2,0671	0.00036245	0.021907
Glyma.01G022800.2.Wm82.a4.v1	2,7621	2,66E-02	0.00039185
Glyma.01G032300.1.Wm82.a4.v1	2,2068	2,18E-01	0.0024774
Glyma.01G035500.1.Wm82.a4.v1	0,3079	0.00012936	0.010377
Glyma.01G048700.1.Wm82.a4.v1	2,8658	8,55E-02	0.0075748
Glyma.01G091400.1.Wm82.a4.v1	0,6798	4,37E-19	1,21E-15
Glyma.01G103200.1.Wm82.a4.v1	2,0018	0.00081491	0.039514
Glyma.01G152200.1.Wm82.a4.v1	3,0795	0.00021732	0.015084
Glyma.01G155400.1.Wm82.a4.v1	0,3016	8,11E-07	3,74E-04
Glyma.01G167400.1.Wm82.a4.v1	3,5082	3,92E-02	0.00055955
Glyma.01G211200.3.Wm82.a4.v1	2,8829	4,88E-03	8,88E-01
Glyma.01G212000.1.Wm82.a4.v1	1,4142	0.00030136	0.018969
Glyma.01G225100.1.Wm82.a4.v1	4,0819	6,28E-03	0.00011094
Glyma.01G238600.2.Wm82.a4.v1	4,9406	5,95E-06	2,38E-03
Glyma.01G241500.1.Wm82.a4.v1	1,7316	0.00073296	0.036233
Glyma.02G012200.1.Wm82.a4.v1	3,5133	8,93E-01	0.0078225
Glyma.02G013402.1.Wm82.a4.v1	2,5736	0.00062769	0.032072
Glyma.02G017100.1.Wm82.a4.v1	0,2385	0.0003512	0.02144
Glyma.02G018600.1.Wm82.a4.v1	2,0363	0.00014629	0.01158
Glyma.02G022100.1.Wm82.a4.v1	3,3215	0.0001904	0.013717
Glyma.02G038400.2.Wm82.a4.v1	1,9592	0.00032455	0.020137
Glyma.02G086100.10.Wm82.a4.v1	3,3125	0.00023798	0.016114
Glyma.02G093500.1.Wm82.a4.v1	1,3608	0.00028488	0.018234
Glyma.02G108100.1.Wm82.a4.v1	1,3462	0.00030783	0.019296
Glyma.02G127300.1.Wm82.a4.v1	1,6391	0.00027871	0.018007
Glyma.02G144700.2.Wm82.a4.v1	2,2126	4,40E-01	0.0044432
Glyma.02G196301.1.Wm82.a4.v1	2,9361	0.00052886	0.028109
Glyma.02G205600.1.Wm82.a4.v1	3,0723	0.00063017	0.032145
Glyma.02G218200.1.Wm82.a4.v1	2,1874	2,30E-01	0.0025666
Glyma.02G230000.1.Wm82.a4.v1	2,1196	6,79E-02	0.00090725
Glyma.02G242000.1.Wm82.a4.v1	2,1646	6,45E-01	0.0061165
Glyma.02G252800.1.Wm82.a4.v1	3,7129	8,37E-03	0.00014462
Glyma.02G259700.1.Wm82.a4.v1	3,1233	4,66E-01	0.0046153
Glyma.02G261700.1.Wm82.a4.v1	2,2615	0.00040464	0.023522
Glyma.02G268200.2.Wm82.a4.v1	1,7115	6,40E-01	0.0061165
Glyma.02G270800.1.Wm82.a4.v1	4,7398	2,95E-07	1,40E-04

Genes sobre-expresados (373)

Glyma.02G305600.1.Wm82.a4.v1	1,8068	6,88E-02	0.00090941
Glyma.03G035900.1.Wm82.a4.v1	0,1172	0.00019174	0.013749
Glyma.03G041400.1.Wm82.a4.v1	6,3055	5,59E-12	5,00E-09
Glyma.03G053602.1.Wm82.a4.v1	5,1354	6,28E-06	2,48E-03
Glyma.03G055300.2.Wm82.a4.v1	3,6328	4,57E-01	0.0045505
Glyma.03G057500.1.Wm82.a4.v1	5,2976	1,21E-07	6,14E-05
Glyma.03G066601.1.Wm82.a4.v1	3,3366	0.0002165	0.015084
Glyma.03G067100.1.Wm82.a4.v1	4,2781	1.58e-15	1,33E-08
Glyma.03G068900.5.Wm82.a4.v1	6,1311	1,33E-10	1,01E-07
Glyma.03G070400.1.Wm82.a4.v1	2,5445	2,26E-02	0.0025329
Glyma.03G084025.1.Wm82.a4.v1	3,7033	2,01E-01	0.0023183
Glyma.03G086200.1.Wm82.a4.v1	3,2658	1,93E-05	5,58E-02
Glyma.03G126000.1.Wm82.a4.v1	2,0521	1,69E-04	5,02E-02
Glyma.03G146100.1.Wm82.a4.v1	3,2295	4,58E-03	8,56E-01
Glyma.03G150500.1.Wm82.a4.v1	2,6839	2,81E-06	1,14E-03
Glyma.03G167400.9.Wm82.a4.v1	3,0156	0.0006357	0.032373
Glyma.03G190100.1.Wm82.a4.v1	3,1737	0.00012154	0.0098011
Glyma.03G207800.1.Wm82.a4.v1	3,5456	1,47E-01	0.0017513
Glyma.03G213900.1.Wm82.a4.v1	2,9144	5,54E-02	0.00076878
Glyma.03G218300.1.Wm82.a4.v1	1,3658	0.00093628	0.043524
Glyma.03G246800.1.Wm82.a4.v1	2,9145	0.00016524	0.012497
Glyma.03G247200.1.Wm82.a4.v1	2,1891	4,69E-03	8,58E-01
Glyma.03G249100.1.Wm82.a4.v1	1,8679	2,32E-01	0.0025666
Glyma.04G004500.1.Wm82.a4.v1	1,7196	0.00063836	0.0324
Glyma.04G007200.1.Wm82.a4.v1	3,1803	0.00043509	0.024864
Glyma.04G037100.1.Wm82.a4.v1	1,5103	0.00026056	0.017146
Glyma.04G044800.1.Wm82.a4.v1	0,2477	6,28E-01	0.0060231
Glyma.04G052000.5.Wm82.a4.v1	3,2166	3,31E-01	0.0035324
Glyma.04G054400.1.Wm82.a4.v1	3,9781	2,13E-03	4,45E-01
Glyma.04G069700.4.Wm82.a4.v1	1,6547	1,01E-02	0.00016775
Glyma.04G105900.1.Wm82.a4.v1	0,7808	1,81E-22	1,10E-18
Glyma.04G113900.1.Wm82.a4.v1	3,2832	0.00027642	0.017956
Glyma.04G178500.1.Wm82.a4.v1	2,6148	0.00093556	0.043524
Glyma.04G208000.1.Wm82.a4.v1	2,5829	0.00046211	0.025826
Glyma.04G220900.2.Wm82.a4.v1	1,3788	0.00068273	0.034139
Glyma.04G229800.1.Wm82.a4.v1	5,6608	5,36E-12	4,94E-09
Glyma.04G230500.1.Wm82.a4.v1	2,1698	0.0003986	0.023414
Glyma.04G233800.3.Wm82.a4.v1	3,7479	3,34E-03	6,63E-01
Glyma.05G022200.1.Wm82.a4.v1	4,1068	1,31E-06	5,61E-04
Glyma.05G056200.1.Wm82.a4.v1	4,0662	3,59E-03	0.0005174
Glyma.05G056300.1.Wm82.a4.v1	4,2611	8,74E-03	0.0001484

Glyma.05G064700.1.Wm82.a4.v1	2,1341	0.00049281	0.02685
Glyma.05G065550.1.Wm82.a4.v1	6,3926	7,70E-13	7,81E-10
Glyma.05G169200.1.Wm82.a4.v1	5,7217	8,81E-09	5,36E-06
Glyma.05G173100.2.Wm82.a4.v1	1,4763	0.00056643	0.02964
Glyma.05G199300.1.Wm82.a4.v1	3,3272	8,07E-02	0.0010391
Glyma.05G211700.1.Wm82.a4.v1	1,5068	0.00045083	0.025395
Glyma.05G221800.1.Wm82.a4.v1	3,1402	2,30E-03	4,72E-01
Glyma.05G222400.1.Wm82.a4.v1	3,9006	1,48E-02	0.00023024
Glyma.05G237200.1.Wm82.a4.v1	1,7561	0.0010447	0.047404
Glyma.06G031000.1.Wm82.a4.v1	2,9225	0.00019449	0.01388
Glyma.06G031100.1.Wm82.a4.v1	4,4348	1,18E-03	2,73E-01
Glyma.06G054700.1.Wm82.a4.v1	3,9525	1,61E-05	4,84E-02
Glyma.06G062000.2.Wm82.a4.v1	3,7605	6,24E-03	0.00011094
Glyma.06G134900.4.Wm82.a4.v1	5,2701	3,68E-11	2,66E-07
Glyma.06G135900.1.Wm82.a4.v1	2,3783	0.00025405	0.016901
Glyma.06G145000.2.Wm82.a4.v1	2,0318	9,08E-02	0.0079316
Glyma.06G157800.1.Wm82.a4.v1	3,4573	1,15E-01	0.0014188
Glyma.06G172551.1.Wm82.a4.v1	0,3424	2,89E-01	0.0031242
Glyma.06G195000.1.Wm82.a4.v1	2,3687	0.00037857	0.022544
Glyma.06G205200.1.Wm82.a4.v1	3,8236	8,73E-04	0.0001484
Glyma.06G241100.1.Wm82.a4.v1	3,8273	2,73E-03	5,46E-01
Glyma.06G244302.1.Wm82.a4.v1	4,1153	5,60E-03	0.00010071
Glyma.06G259100.3.Wm82.a4.v1	3,0214	8,75E-01	0.0077083
Glyma.06G265420.1.Wm82.a4.v1	2,2246	0.00039503	0.023275
Glyma.06G289000.1.Wm82.a4.v1	0,2585	5,36E-01	0.0052537
Glyma.07G014300.1.Wm82.a4.v1	2,5098	1,11E-02	0.00018278
Glyma.07G024000.1.Wm82.a4.v1	1,3135	0.00079608	0.038848
Glyma.07G030800.1.Wm82.a4.v1	3,6199	6,58E-02	0.00088938
Glyma.07G043600.1.Wm82.a4.v1	2,7069	0.00065672	0.033221
Glyma.07G052632.1.Wm82.a4.v1	3,7544	4,55E-02	0.00064096
Glyma.07G055300.1.Wm82.a4.v1	1,3253	0.00049642	0.026999
Glyma.07G061500.1.Wm82.a4.v1	3,1748	2,57E-03	5,23E-01
Glyma.07G074400.1.Wm82.a4.v1	2,4471	0.00051274	0.027639
Glyma.07G087400.1.Wm82.a4.v1	2,4395	4,27E-01	0.0043428
Glyma.07G091600.1.Wm82.a4.v1	1,7707	0.00049032	0.026811
Glyma.07G096000.1.Wm82.a4.v1	3,6106	2,20E-01	0.0024838
Glyma.07G110000.1.Wm82.a4.v1	2,6921	1,53E-01	0.0018168
Glyma.07G140700.1.Wm82.a4.v1	3,0291	4,14E-01	0.0042555
Glyma.07G156901.1.Wm82.a4.v1	6,6346	2,39E-14	3,30E-11
Glyma.07G157000.1.Wm82.a4.v1	7,4901	4,70E-21	2,04E-18
Glyma.07G199900.3.Wm82.a4.v1	1,1806	0.00030066	0.018964

Glyma.07G200000.1.Wm82.a4.v1	4,5317	1,39E-05	4,91E-03
Glyma.07G200200.1.Wm82.a4.v1	2,6773	0.00021716	0.015084
Glyma.07G200300.1.Wm82.a4.v1	4,7142	1,03E-06	4,48E-04
Glyma.07G200700.1.Wm82.a4.v1	3,1283	9,52E-01	0.0081293
Glyma.08G028300.1.Wm82.a4.v1	3,0353	1,77E-01	0.0020795
Glyma.08G056600.1.Wm82.a4.v1	2,6025	0.00055288	0.029085
Glyma.08G059900.1.Wm82.a4.v1	1,9525	0.00037654	0.022544
Glyma.08G062300.1.Wm82.a4.v1	1,6994	0.00095061	0.043989
Glyma.08G068900.1.Wm82.a4.v1	3,3237	0.00023728	0.016102
Glyma.08G069000.1.Wm82.a4.v1	2,9118	9,15E-01	0.0079446
Glyma.08G084600.1.Wm82.a4.v1	2,0017	0.0010685	0.048055
Glyma.08G106700.1.Wm82.a4.v1	5,1589	5,68E-09	3,52E-07
Glyma.08G145300.1.Wm82.a4.v1	1,7783	0.0007488	0.036896
Glyma.08G175200.1.Wm82.a4.v1	2,2937	0.00021159	0.014891
Glyma.08G176600.1.Wm82.a4.v1	3,1575	0.00025957	0.017118
Glyma.08G212000.3.Wm82.a4.v1	3,2235	1,37E-01	0.0016484
Glyma.08G273600.1.Wm82.a4.v1	1,9631	0.00031991	0.01993
Glyma.08G285900.1.Wm82.a4.v1	2,3512	0.00012299	0.0098919
Glyma.08G318900.1.Wm82.a4.v1	6,0039	1,48E-11	1,10E-08
Glyma.08G324700.1.Wm82.a4.v1	2,6846	1,11E-03	0.00018278
Glyma.08G332900.1.Wm82.a4.v1	2,0288	0.00064965	0.032918
Glyma.09G029900.1.Wm82.a4.v1	2,6111	0.00090479	0.04245
Glyma.09G040300.1.Wm82.a4.v1	4,0705	5,89E-05	1,52E-01
Glyma.09G060700.1.Wm82.a4.v1	1,8708	0.00047853	0.026355
Glyma.09G065900.1.Wm82.a4.v1	3,4926	1,58E-03	3,44E-01
Glyma.09G093050.1.Wm82.a4.v1	1,6429	1,42E-01	0.0017013
Glyma.09G110700.1.Wm82.a4.v1	2,7874	7,13E-01	0.0065901
Glyma.09G131500.1.Wm82.a4.v1	3,2514	3,14E-01	0.0033731
Glyma.09G138000.1.Wm82.a4.v1	2,7288	0.00027737	0.01798
Glyma.09G161500.1.Wm82.a4.v1	3,1417	4,59E-03	8,56E-01
Glyma.09G171600.1.Wm82.a4.v1	3,5071	6,56E-01	0.0061774
Glyma.09G192800.1.Wm82.a4.v1	1,3809	6,85E-02	0.00090941
Glyma.09G230600.1.Wm82.a4.v1	3,1221	0.00056178	0.029447
Glyma.09G233800.2.Wm82.a4.v1	0,000231	0.00024283	0.016333
Glyma.09G253200.1.Wm82.a4.v1	2,3087	0.00010709	0.0089442
Glyma.10G012900.1.Wm82.a4.v1	2,7119	1,52E-03	3,34E-01
Glyma.10G029400.1.Wm82.a4.v1	2,1547	0.00036167	0.021903
Glyma.10G029600.4.Wm82.a4.v1	4,9666	1,65E-06	6,89E-04
Glyma.10G040700.1.Wm82.a4.v1	1,6341	0.00051244	0.027639
Glyma.10G051100.1.Wm82.a4.v1	2,5645	0.00026669	0.017436
Glyma.10G051900.1.Wm82.a4.v1	1,3636	0.0010807	0.048174

Glyma.10G113000.1.Wm82.a4.v1	4,4512	1,45E-04	3,26E-02
Glyma.10G121524.1.Wm82.a4.v1	3,0799	1,53E-03	3,34E-01
Glyma.10G141300.1.Wm82.a4.v1	3,2893	2,82E-02	0.003063
Glyma.10G161500.1.Wm82.a4.v1	1,9564	1,30E-04	2,97E-02
Glyma.10G193200.1.Wm82.a4.v1	2,0982	2,03E-01	0.0023342
Glyma.10G204900.2.Wm82.a4.v1	2,8691	0.00083619	0.040225
Glyma.10G237800.1.Wm82.a4.v1	3,4817	2,87E-02	0.00041886
Glyma.10G241700.1.Wm82.a4.v1	3,3116	9,42E-01	0.0080667
Glyma.11G030500.1.Wm82.a4.v1	2,9414	1,38E-02	0.00021842
Glyma.11G043600.1.Wm82.a4.v1	3,0928	9,21E-01	0.0079551
Glyma.11G077900.1.Wm82.a4.v1	0,2499	0.0003331	0.0205
Glyma.11G078000.1.Wm82.a4.v1	3,2274	3,99E-03	7,82E-02
Glyma.11G095000.1.Wm82.a4.v1	2,8811	0.00052541	0.028024
Glyma.11G099400.2.Wm82.a4.v1	3,9123	5,61E-02	0.00077547
Glyma.11G124500.2.Wm82.a4.v1	0,2071	0.00050233	0.027223
Glyma.11G140500.1.Wm82.a4.v1	2,9946	2,56E-01	0.0028077
Glyma.11G153587.1.Wm82.a4.v1	3,1629	2,67E-01	0.0029075
Glyma.11G172401.3.Wm82.a4.v1	3,5602	4,53E-02	0.00064006
Glyma.11G179100.1.Wm82.a4.v1	4,0256	9,50E-04	2,33E-01
Glyma.11G186500.1.Wm82.a4.v1	3,2866	0.00017983	0.013335
Glyma.11G207300.1.Wm82.a4.v1	2,6488	0.00089713	0.04222
Glyma.11G223000.1.Wm82.a4.v1	2,4826	0.00021906	0.015171
Glyma.11G246600.1.Wm82.a4.v1	4,0802	4,28E-03	8,14E-01
Glyma.11G247600.1.Wm82.a4.v1	1,8927	0.00092102	0.043078
Glyma.11G250500.1.Wm82.a4.v1	2,2843	6,67E-01	0.0062239
Glyma.11G252800.1.Wm82.a4.v1	2,9072	0.00048549	0.026595
Glyma.12G039100.1.Wm82.a4.v1	2,7424	0.00011554	0.0094172
Glyma.12G046800.1.Wm82.a4.v1	3,0102	2,28E-02	0.00034311
Glyma.12G049300.1.Wm82.a4.v1	2,0099	0.00051565	0.027698
Glyma.12G078100.1.Wm82.a4.v1	2,2984	2,37E-01	0.0026077
Glyma.12G103500.1.Wm82.a4.v1	5,4918	1,14E-07	5,87E-05
Glyma.12G117900.1.Wm82.a4.v1	0,3537	1,22E-03	2.82e-05
Glyma.12G163900.1.Wm82.a4.v1	1,7756	0.00085825	0.041026
Glyma.12G188500.1.Wm82.a4.v1	2,4847	0.00036397	0.021955
Glyma.12G220600.2.Wm82.a4.v1	1,0486	8,08E-01	0.0072282
Glyma.13G000200.1.Wm82.a4.v1	5,9432	5,71E-10	3,94E-07
Glyma.13G001200.1.Wm82.a4.v1	0,2455	0.00067724	0.033976
Glyma.13G003500.1.Wm82.a4.v1	3,1988	0.00035228	0.021463
Glyma.13G005833.1.Wm82.a4.v1	4,2735	2,28E-03	4,71E-01
Glyma.13G033300.3.Wm82.a4.v1	3,8745	1,84E-05	6,27E-03
Glyma.13G033700.1.Wm82.a4.v1	3,1101	0.00060687	0.031271

Glyma.13G033800.1.Wm82.a4.v1	3,4619	1,98E-01	0.002303
Glyma.13G035100.1.Wm82.a4.v1	2,9185	0.00080159	0.038992
Glyma.13G071400.1.Wm82.a4.v1	3,9936	2,84E-05	9,58E-03
Glyma.13G086800.1.Wm82.a4.v1	3,2387	0.00015439	0.011883
Glyma.13G098600.1.Wm82.a4.v1	7,0143	4,38E-17	8,33E-14
Glyma.13G105700.1.Wm82.a4.v1	0,2241	0.0010775	0.048147
Glyma.13G116600.1.Wm82.a4.v1	2,8751	3,96E-01	0.00408
Glyma.13G166850.1.Wm82.a4.v1	3,2879	7,65E-01	0.0069196
Glyma.13G176000.1.Wm82.a4.v1	4,1528	9,51E-04	2,33E-01
Glyma.13G176100.1.Wm82.a4.v1	2,0872	0.00026123	0.017153
Glyma.13G176200.1.Wm82.a4.v1	3,8568	1,06E-03	2,51E-01
Glyma.13G176300.1.Wm82.a4.v1	2,6385	0.00025481	0.016914
Glyma.13G176500.1.Wm82.a4.v1	3,1024	4,24E-02	0.0043267
Glyma.13G188400.1.Wm82.a4.v1	4,4372	2,68E-03	5,40E-02
Glyma.13G206700.1.Wm82.a4.v1	2,8604	5,52E-03	9,98E-01
Glyma.13G215500.3.Wm82.a4.v1	1,6744	0.00015308	0.011812
Glyma.13G221200.1.Wm82.a4.v1	2,8378	0.0010294	0.046918
Glyma.13G240100.1.Wm82.a4.v1	2,2608	0.00011043	0.0091354
Glyma.13G250900.2.Wm82.a4.v1	3,1629	0.00047199	0.026167
Glyma.13G351050.1.Wm82.a4.v1	0,3021	0.00058874	0.0305
Glyma.13G359500.3.Wm82.a4.v1	3,2219	0.00015824	0.012118
Glyma.14G003800.1.Wm82.a4.v1	2,1571	0.00067923	0.03402
Glyma.14G060700.1.Wm82.a4.v1	2,7881	0.00044576	0.025289
Glyma.14G063800.1.Wm82.a4.v1	2,9228	9,11E-01	0.0079328
Glyma.14G098100.3.Wm82.a4.v1	1,5041	0.00017907	0.013311
Glyma.14G099900.1.Wm82.a4.v1	4,6878	1,70E-05	5,86E-03
Glyma.14G100000.1.Wm82.a4.v1	3,0085	0.00084435	0.040425
Glyma.14G118550.1.Wm82.a4.v1	6,4945	4,65E-13	4,87E-10
Glyma.14G118600.1.Wm82.a4.v1	4,5638	9,55E-08	5,00E-05
Glyma.14G120800.1.Wm82.a4.v1	3,5676	7,09E-02	0.0065735
Glyma.14G136300.1.Wm82.a4.v1	4,0551	1,28E-07	6,40E-05
Glyma.14G136351.1.Wm82.a4.v1	4,8652	5,36E-14	6,52E-11
Glyma.14G156300.1.Wm82.a4.v1	3,0434	0.00023612	0.016059
Glyma.14G186400.1.Wm82.a4.v1	0,000271	0.00070249	0.034954
Glyma.14G214800.1.Wm82.a4.v1	3,4514	4,12E-03	7,98E-01
Glyma.15G004200.1.Wm82.a4.v1	2,1831	7,57E-02	0.0068925
Glyma.15G057400.1.Wm82.a4.v1	3,1107	0.00016018	0.012174
Glyma.15G059900.1.Wm82.a4.v1	1,7203	0.0011279	0.049984
Glyma.15G062500.1.Wm82.a4.v1	1,6562	0.00058727	0.0305
Glyma.15G073200.1.Wm82.a4.v1	2,2334	4,54E-01	0.004536
Glyma.15G077700.1.Wm82.a4.v1	3,4232	6,10E-03	0.00083888

Glyma.15G099600.2.Wm82.a4.v1	4,2074	3,03E-04	8,21E-02
Glyma.15G112600.1.Wm82.a4.v1	0,2851	8,16E-08	4,43E-05
Glyma.15G128800.1.Wm82.a4.v1	2,8308	7,45E-02	0.00096797
Glyma.15G154500.1.Wm82.a4.v1	3,6147	9,53E-03	0.0001601
Glyma.15G160950.1.Wm82.a4.v1	3,3739	8,66E-01	0.0076528
Glyma.15G161700.1.Wm82.a4.v1	3,0925	8,23E-01	0.0073341
Glyma.15G167300.1.Wm82.a4.v1	3,3401	1,72E-07	8,44E-05
Glyma.15G170000.4.Wm82.a4.v1	4,7543	3,55E-05	1,17E-02
Glyma.15G192800.1.Wm82.a4.v1	3,7051	2,14E-02	0.00032462
Glyma.15G196500.1.Wm82.a4.v1	4,7245	1,23E-05	4,41E-03
Glyma.15G199700.1.Wm82.a4.v1	1,9245	0.00016154	0.012247
Glyma.15G240066.1.Wm82.a4.v1	2,5862	0.0003003	0.018964
Glyma.15G241400.1.Wm82.a4.v1	4,4897	1,72E-03	3,71E-01
Glyma.15G248300.1.Wm82.a4.v1	2,8397	7,20E-02	0.00094003
Glyma.15G249800.1.Wm82.a4.v1	2,1258	0.00082794	0.039954
Glyma.15G250300.1.Wm82.a4.v1	6,0649	8,65E-11	6,74E-08
Glyma.15G252000.1.Wm82.a4.v1	4,8208	9,90E-19	2,51E-15
Glyma.15G252700.1.Wm82.a4.v1	3,0211	0.0003767	0.022544
Glyma.15G264600.1.Wm82.a4.v1	3,7889	1,31E-02	0.00021005
Glyma.16G012000.1.Wm82.a4.v1	3,9578	2,09E-03	4,44E-01
Glyma.16G037600.1.Wm82.a4.v1	2,6489	0.00076719	0.037559
Glyma.16G042000.1.Wm82.a4.v1	2,4807	9,32E-01	0.0080072
Glyma.16G093750.2.Wm82.a4.v1	3,3767	0.0001866	0.013539
Glyma.16G096200.1.Wm82.a4.v1	4,4917	1,91E-03	4,08E-01
Glyma.16G105700.1.Wm82.a4.v1	2,5836	0.00039503	0.023275
Glyma.16G125000.1.Wm82.a4.v1	3,1463	0.00040114	0.023498
Glyma.16G127960.1.Wm82.a4.v1	4,7541	1,98E-04	5,69E-02
Glyma.16G128020.1.Wm82.a4.v1	3,3371	0.00012102	0.0097849
Glyma.16G128140.1.Wm82.a4.v1	2,3741	1,00E-01	0.0012617
Glyma.16G133600.1.Wm82.a4.v1	2,3393	0.00016601	0.012524
Glyma.16G158100.1.Wm82.a4.v1	4,8906	3,23E-09	2,09E-06
Glyma.16G173000.1.Wm82.a4.v1	3,0951	1,33E-01	0.001605
Glyma.16G176600.2.Wm82.a4.v1	1,9573	0.000515	0.027698
Glyma.16G178800.1.Wm82.a4.v1	0,4877	7,52E-07	3,52E-04
Glyma.16G178900.1.Wm82.a4.v1	1,9517	9,97E-04	2,43E-01
Glyma.16G193600.1.Wm82.a4.v1	3,6399	9,91E-03	0.00016546
Glyma.16G195000.1.Wm82.a4.v1	1,8238	0.00041766	0.024082
Glyma.16G212300.1.Wm82.a4.v1	3,4457	5,30E-01	0.0052225
Glyma.16G215100.1.Wm82.a4.v1	4,1039	3,17E-20	1,20E-16
Glyma.17G006900.1.Wm82.a4.v1	3,6437	1,05E-01	0.0013049
Glyma.17G036100.1.Wm82.a4.v1	2,1204	0.00041999	0.024137

Glyma.17G042100.1.Wm82.a4.v1	1,8038	7,57E-01	0.0068925
Glyma.17G045900.1.Wm82.a4.v1	2,3221	2,13E-01	0.0024233
Glyma.17G077500.1.Wm82.a4.v1	6,7171	2,36E-14	3,30E-11
Glyma.17G089400.1.Wm82.a4.v1	0,2163	0.00067587	0.033963
Glyma.17G117150.1.Wm82.a4.v1	4,3324	4,16E-04	8,00E-01
Glyma.17G172800.2.Wm82.a4.v1	3,6845	3,69E-01	0.0038569
Glyma.17G177800.1.Wm82.a4.v1	4,3847	4,04E-03	7,87E-01
Glyma.17G179000.1.Wm82.a4.v1	6,4327	2,33E-13	2,53E-10
Glyma.17G182166.1.Wm82.a4.v1	5,3824	2,38E-09	1,57E-06
Glyma.17G195000.1.Wm82.a4.v1	3,2623	0.0003135	0.019571
Glyma.17G203200.1.Wm82.a4.v1	3,1744	0.00018548	0.013505
Glyma.17G205900.7.Wm82.a4.v1	4,6263	6,46E-04	1,64E-01
Glyma.17G206001.1.Wm82.a4.v1	7,7016	1,53E-23	1,17E-19
Glyma.17G212500.1.Wm82.a4.v1	1,2464	0.00056111	0.029447
Glyma.17G217700.1.Wm82.a4.v1	3,1893	0.00037892	0.022544
Glyma.17G235366.1.Wm82.a4.v1	5,7545	4,12E-09	2,61E-06
Glyma.17G235432.1.Wm82.a4.v1	5,5173	7,97E-08	4,41E-05
Glyma.17G235500.1.Wm82.a4.v1	5,6535	3,23E-08	1,89E-05
Glyma.17G247800.1.Wm82.a4.v1	2,7135	0.00081375	0.039514
Glyma.18G009700.1.Wm82.a4.v1	2,3049	0.00048193	0.026447
Glyma.18G022400.1.Wm82.a4.v1	1,7152	9,14E-04	2,28E-01
Glyma.18G022500.1.Wm82.a4.v1	1,8097	7,69E-08	4,33E-05
Glyma.18G022600.3.Wm82.a4.v1	2,5871	9,18E-01	0.0079475
Glyma.18G022700.3.Wm82.a4.v1	0,000172	3,86E-01	0.0040091
Glyma.18G028233.1.Wm82.a4.v1	2,0894	0.00022885	0.015741
Glyma.18G034100.1.Wm82.a4.v1	1,6664	1,24E-01	0.0015085
Glyma.18G056800.1.Wm82.a4.v1	2,4525	0.0001832	0.013505
Glyma.18G068250.1.Wm82.a4.v1	4,4662	1,43E-04	4,39E-02
Glyma.18G070000.1.Wm82.a4.v1	6,2294	1,15E-11	9,95E-09
Glyma.18G087000.3.Wm82.a4.v1	2,2133	0.00044781	0.025353
Glyma.18G087100.1.Wm82.a4.v1	4,8822	1,57E-06	6,63E-04
Glyma.18G107000.1.Wm82.a4.v1	0,1828	0.00021097	0.014891
Glyma.18G109000.1.Wm82.a4.v1	3,4767	0.00011407	0.0093479
Glyma.18G112500.1.Wm82.a4.v1	5,4433	2,55E-07	1,23E-04
Glyma.18G127200.1.Wm82.a4.v1	4,1254	1,90E-04	5,56E-02
Glyma.18G144200.1.Wm82.a4.v1	0,3556	3,85E-01	0.004004
Glyma.18G170900.1.Wm82.a4.v1	1,9613	0.00049215	0.02685
Glyma.18G204000.1.Wm82.a4.v1	2,4693	0.00024225	0.01633
Glyma.18G219200.1.Wm82.a4.v1	3,6241	4,65E-03	8,56E-01
Glyma.18G225500.1.Wm82.a4.v1	1,2465	0.0006029	0.031119
Glyma.18G229700.1.Wm82.a4.v1	2,5046	2,60E-03	5,27E-01

Glyma.18G231800.1.Wm82.a4.v1	2,2854	0.00029455	0.018761
Glyma.18G247500.1.Wm82.a4.v1	1,5199	0.00076533	0.037529
Glyma.18G247650.1.Wm82.a4.v1	4,3734	4,83E-04	1,28E-01
Glyma.18G253100.1.Wm82.a4.v1	4,7276	2,30E-04	6,41E-02
Glyma.18G259700.1.Wm82.a4.v1	0,2169	0.00017299	0.012922
Glyma.18G260700.1.Wm82.a4.v1	1,5168	7,60E-01	0.0069004
Glyma.18G271167.1.Wm82.a4.v1	4,6571	4,74E-04	1,26E-01
Glyma.18G293400.3.Wm82.a4.v1	3,1966	6,88E-01	0.0063924
Glyma.19G001600.1.Wm82.a4.v1	2,3802	9,29E-01	0.0080024
Glyma.19G011400.1.Wm82.a4.v1	2,8933	4,72E-02	0.00066083
Glyma.19G050300.3.Wm82.a4.v1	1,3451	0.00014868	0.01168
Glyma.19G066350.1.Wm82.a4.v1	3,3203	0.0002331	0.015925
Glyma.19G085100.1.Wm82.a4.v1	2,6265	5,52E-01	0.0053974
Glyma.19G086200.1.Wm82.a4.v1	0,2765	7,27E-01	0.0066965
Glyma.19G086400.1.Wm82.a4.v1	3,0249	0.0006149	0.031631
Glyma.19G086500.1.Wm82.a4.v1	0,6345	5,02E-18	1,09E-14
Glyma.19G098801.1.Wm82.a4.v1	3,6368	2,56E-02	0.00037964
Glyma.19G113400.4.Wm82.a4.v1	1,8811	0.0010913	0.048575
Glyma.19G177450.1.Wm82.a4.v1	1,9008	0.00093472	0.043524
Glyma.19G190400.1.Wm82.a4.v1	3,4405	1,04E-01	0.0013023
Glyma.19G191700.1.Wm82.a4.v1	2,9005	0.00089314	0.042163
Glyma.19G213000.1.Wm82.a4.v1	4,3501	4.63e-07	8,56E-01
Glyma.20G005300.1.Wm82.a4.v1	1,5171	0.00045107	0.025395
Glyma.20G013600.1.Wm82.a4.v1	3,9327	2,14E-03	4,45E-01
Glyma.20G015900.1.Wm82.a4.v1	3,7453	1,50E-03	3,33E-01
Glyma.20G017200.3.Wm82.a4.v1	1,5065	0.00104	0.04726
Glyma.20G018100.3.Wm82.a4.v1	2,8733	6,42E-02	0.00087155
Glyma.20G018500.1.Wm82.a4.v1	2,5603	1,26E-01	0.0015334
Glyma.20G019200.1.Wm82.a4.v1	1,5867	4,27E-03	8,14E-01
Glyma.20G028300.2.Wm82.a4.v1	4,3416	5,99E-04	1,53E-01
Glyma.20G028350.1.Wm82.a4.v1	3,4755	0.00011451	0.0093585
Glyma.20G032400.2.Wm82.a4.v1	0,3217	0.00020811	0.014783
Glyma.20G038302.1.Wm82.a4.v1	4,7031	2,68E-04	7,40E-02
Glyma.20G099200.1.Wm82.a4.v1	1,9444	0.00035513	0.021594
Glyma.20G105200.1.Wm82.a4.v1	3,1051	8,92E-02	0.0078225
Glyma.20G110900.1.Wm82.a4.v1	1,5208	0.00014753	0.01162
Glyma.20G116800.1.Wm82.a4.v1	3,2393	6,48E-01	0.0061165
Glyma.20G128800.1.Wm82.a4.v1	1,2142	0.00066165	0.033359
Glyma.20G143700.2.Wm82.a4.v1	2,8339	2,31E-01	0.0025666
Glyma.20G156800.1.Wm82.a4.v1	2,6505	3,52E-01	0.0037278
Glyma.20G210800.1.Wm82.a4.v1	1,9106	0.00045317	0.025419

Glyma.20G231900.1.Wm82.a4.v1	1,9496	1,45E-01	0.0017309
Glyma.20G236900.1.Wm82.a4.v1	1,7491	0.00019402	0.013879
Glyma.20G238550.1.Wm82.a4.v1	3,0492	0.00052158	0.027869
Glyma.20G241000.1.Wm82.a4.v1	3,0533	0.00046411	0.02589
Glyma.U002150.1.Wm82.a4.v1	4,7605	9,03E-07	4,10E-04
Glyma.U031132.1.Wm82.a4.v1	3,0748	0.00069505	0.034672
Glyma.U032205.1.Wm82.a4.v1	6,5716	4,01E-15	6,41E-13
Glyma.U044201.1.Wm82.a4.v1	4,7267	8,22E-06	3,08E-03

Genes subexpresados (313)

Gene_ID	log2FoldChange	pvalue	padj
Glyma.01G010300.2.Wm82.a4.v1	-2,0315	0.00037631	0.022544
Glyma.01G019300.1.Wm82.a4.v1	-2,3541	0.00047485	0.026248
Glyma.01G019400.2.Wm82.a4.v1	-3,0492	0.00075905	0.03728
Glyma.01G019400.4.Wm82.a4.v1	-3,7204	6,24E-02	0.00085426
Glyma.01G032900.1.Wm82.a4.v1	-2,0924	0.00010404	0.0087134
Glyma.01G033200.1.Wm82.a4.v1	-6,6588	3,24E-15	5,47E-12
Glyma.01G046900.1.Wm82.a4.v1	-0,4343	4,57E-03	8,56E-01
Glyma.01G082000.1.Wm82.a4.v1	-2,8964	0.0010585	0.047938
Glyma.01G092300.1.Wm82.a4.v1	-1,5345	0.00063692	0.032381
Glyma.01G141100.1.Wm82.a4.v1	-2,9412	0.0010659	0.048009
Glyma.01G150900.1.Wm82.a4.v1	-1,9682	0.0010121	0.04627
Glyma.01G151100.1.Wm82.a4.v1	-2,2416	0.00055297	0.029085
Glyma.01G154400.1.Wm82.a4.v1	-2,2847	0.00017039	0.012762
Glyma.01G156200.1.Wm82.a4.v1	-3,3664	0.00018201	0.013464
Glyma.01G186800.1.Wm82.a4.v1	-2,9653	0.00061999	0.031763
Glyma.01G229600.1.Wm82.a4.v1	-3,9074	1,20E-01	0.0014684
Glyma.01G234300.9.Wm82.a4.v1	-0,3106	0.00060119	0.031084
Glyma.02G001700.1.Wm82.a4.v1	-2,0817	0.00084301	0.040425
Glyma.02G005500.2.Wm82.a4.v1	-3,0832	0.00044586	0.025289
Glyma.02G013300.1.Wm82.a4.v1	-2,5312	9,94E-01	0.0084659
Glyma.02G024600.1.Wm82.a4.v1	-2,5631	0.00038047	0.022592
Glyma.02G068100.1.Wm82.a4.v1	-2,6392	0.00020918	0.014824
Glyma.02G082500.1.Wm82.a4.v1	-2,8905	0.00078554	0.038396
Glyma.02G089000.1.Wm82.a4.v1	-3,0682	1,16E-12	1,10E-09
Glyma.02G093251.1.Wm82.a4.v1	-0,4142	1,93E-02	0.00029536
Glyma.02G111100.1.Wm82.a4.v1	-0,3074	4,30E-01	0.0043574
Glyma.02G111200.1.Wm82.a4.v1	-4,7863	1,07E-05	3,92E-03
Glyma.02G197100.1.Wm82.a4.v1	-0,4138	1,16E-03	2,71E-01
Glyma.02G205102.1.Wm82.a4.v1	-6,1595	3,78E-16	6,76E-13

Glyma.02G253000.1.Wm82.a4.v1	-2,5568	0.00022728	0.015669
Glyma.02G262100.1.Wm82.a4.v1	-1,9499	0.00090433	0.04245
Glyma.03G000800.1.Wm82.a4.v1	-1,2845	0.00041228	0.02392
Glyma.03G026400.1.Wm82.a4.v1	-4,1772	1,42E-02	0.00022185
Glyma.03G029900.1.Wm82.a4.v1	-3,5159	6,00E-05	1,88E-02
Glyma.03G031200.3.Wm82.a4.v1	-1,8214	7,88E-01	0.0071051
Glyma.03G031400.1.Wm82.a4.v1	-6,4736	1,63E-13	1,91E-10
Glyma.03G034500.2.Wm82.a4.v1	-2,6303	0.0010069	0.046173
Glyma.03G036000.1.Wm82.a4.v1	-1,7015	6,70E-02	0.00090186
Glyma.03G043000.1.Wm82.a4.v1	-0,2387	0.00050358	0.027242
Glyma.03G048200.1.Wm82.a4.v1	-3,0592	0.00072136	0.035776
Glyma.03G050100.2.Wm82.a4.v1	-3,3692	0.00019025	0.013717
Glyma.03G072100.1.Wm82.a4.v1	-2,4038	2,02E-01	0.0023259
Glyma.03G077900.1.Wm82.a4.v1	-1,6365	0.0010119	0.04627
Glyma.03G083600.1.Wm82.a4.v1	-5,9282	3,76E-10	2,66E-07
Glyma.03G100800.2.Wm82.a4.v1	-2,9541	0.00088427	0.041875
Glyma.03G129400.1.Wm82.a4.v1	-2,3923	0.00013232	0.010558
Glyma.03G180000.1.Wm82.a4.v1	-2,5657	0.00052871	0.028109
Glyma.03G196900.3.Wm82.a4.v1	-3,1575	0.00047609	0.026269
Glyma.03G213000.1.Wm82.a4.v1	-1,2303	0.00033521	0.020588
Glyma.03G234800.1.Wm82.a4.v1	-2,2158	0.00023102	0.015818
Glyma.03G248500.1.Wm82.a4.v1	-1,3628	0.00017042	0.012762
Glyma.03G264902.1.Wm82.a4.v1	-2,1882	4,37E-02	0.00062093
Glyma.04G022620.1.Wm82.a4.v1	-0,000425	1,04E-03	2,49E-01
Glyma.04G022660.1.Wm82.a4.v1	-0,000745	1,60E-21	8,11E-18
Glyma.04G028600.1.Wm82.a4.v1	-2,6151	0.0010311	0.046926
Glyma.04G039600.1.Wm82.a4.v1	-3,4526	7,51E-01	0.0068796
Glyma.04G064500.1.Wm82.a4.v1	-2,6643	0.0010729	0.048147
Glyma.04G066500.1.Wm82.a4.v1	-3,3119	6,65E-01	0.0062234
Glyma.04G073400.3.Wm82.a4.v1	-3,1903	0.00040276	0.023522
Glyma.04G078500.2.Wm82.a4.v1	-6,8929	1,02E-30	1,56E-26
Glyma.04G085800.1.Wm82.a4.v1	-3,0032	0.00023099	0.015818
Glyma.04G088500.1.Wm82.a4.v1	-2,9812	0.00017488	0.013031
Glyma.04G098800.1.Wm82.a4.v1	-2,9805	0.0010288	0.046918
Glyma.04G104200.1.Wm82.a4.v1	-1,3982	0.00086659	0.041233
Glyma.04G106300.1.Wm82.a4.v1	-3,4724	0.00010359	0.0087002
Glyma.04G114500.1.Wm82.a4.v1	-4,0458	3,48E-02	0.00050314
Glyma.04G114600.1.Wm82.a4.v1	-3,0128	0.00045035	0.025395
Glyma.04G115300.1.Wm82.a4.v1	-5,6745	1,53E-08	9,13E-06
Glyma.04G140132.1.Wm82.a4.v1	-4,5237	2,23E-04	6,26E-02
Glyma.04G143500.2.Wm82.a4.v1	-4,6831	6,98E-06	2,69E-04

Glyma.04G155500.6.Wm82.a4.v1	-4,5636	4,21E-05	1,36E-02
Glyma.04G159300.1.Wm82.a4.v1	-2,6519	0.00083551	0.040225
Glyma.04G173500.1.Wm82.a4.v1	-0,2993	0.00097572	0.045014
Glyma.04G214400.3.Wm82.a4.v1	-1,8721	0.0009434	0.043788
Glyma.04G231700.3.Wm82.a4.v1	-2,8066	0.00043063	0.024655
Glyma.05G035800.1.Wm82.a4.v1	-3,6795	3,61E-01	0.0037883
Glyma.05G037200.1.Wm82.a4.v1	-0,2998	0.00096897	0.04477
Glyma.05G037800.1.Wm82.a4.v1	-3,1925	0.0002515	0.016768
Glyma.05G042600.1.Wm82.a4.v1	-3,7818	1,05E-01	0.0013049
Glyma.05G084900.2.Wm82.a4.v1	-3,4325	4,51E-01	0.0045278
Glyma.05G095600.5.Wm82.a4.v1	-3,6074	5,54E-01	0.0054018
Glyma.05G130300.2.Wm82.a4.v1	-1,4312	8,07E-01	0.0072282
Glyma.05G155600.1.Wm82.a4.v1	-1,4141	0.00066704	0.033575
Glyma.05G165700.2.Wm82.a4.v1	-7,9228	2,30E-31	6.98e-31
Glyma.05G198900.1.Wm82.a4.v1	-3,4583	0.00011723	0.0095042
Glyma.05G204400.1.Wm82.a4.v1	-0,1541	6,46E-01	0.0061165
Glyma.05G248700.1.Wm82.a4.v1	-2,4409	1,84E-01	0.0021505
Glyma.06G013200.1.Wm82.a4.v1	-3,3741	3,58E-01	0.0037644
Glyma.06G022500.1.Wm82.a4.v1	-3,6652	2,57E-01	0.0028077
Glyma.06G041800.1.Wm82.a4.v1	-2,2036	0.0010739	0.048147
Glyma.06G059500.1.Wm82.a4.v1	-2,4689	2,07E-01	0.0023687
Glyma.06G064700.1.Wm82.a4.v1	-0,3562	7,42E-01	0.0068122
Glyma.06G065000.1.Wm82.a4.v1	-2,0507	0.00025666	0.016963
Glyma.06G087400.1.Wm82.a4.v1	-2,5119	5,97E-01	0.005765
Glyma.06G100300.1.Wm82.a4.v1	-0,1989	0.00073824	0.036435
Glyma.06G111600.1.Wm82.a4.v1	-2,7845	1,64E-01	0.0019386
Glyma.06G115800.2.Wm82.a4.v1	-3,0354	0.00010942	0.0090894
Glyma.06G145100.1.Wm82.a4.v1	-1,4534	0.00039893	0.023414
Glyma.06G164000.1.Wm82.a4.v1	-3,8041	1,96E-01	0.0022814
Glyma.06G164200.1.Wm82.a4.v1	-3,1565	0.00026436	0.017321
Glyma.06G180300.1.Wm82.a4.v1	-3,0738	0.00057291	0.029876
Glyma.06G195600.1.Wm82.a4.v1	-0,2608	0.00082707	0.039954
Glyma.06G198100.2.Wm82.a4.v1	-3,2705	0.00029843	0.018902
Glyma.06G205800.1.Wm82.a4.v1	-3,3658	9,24E-07	4,13E-04
Glyma.06G207900.3.Wm82.a4.v1	-3,8957	9,25E-02	0.0011764
Glyma.06G212601.1.Wm82.a4.v1	-6,5557	3,15E-14	4,16E-12
Glyma.06G212700.1.Wm82.a4.v1	-3,3849	4,62E-01	0.0045893
Glyma.06G212800.1.Wm82.a4.v1	-7,1049	1,16E-18	2,72E-16
Glyma.06G213101.1.Wm82.a4.v1	-7,0047	3,23E-17	6,54E-14
Glyma.06G226700.1.Wm82.a4.v1	-3,8749	2,75E-02	0.00040395
Glyma.06G233300.1.Wm82.a4.v1	-3,8071	7,19E-02	0.00094003

Glyma.06G237100.1.Wm82.a4.v1	-4,2671	7,97E-03	0.00013851
Glyma.06G254800.1.Wm82.a4.v1	-3,8786	1,44E-04	3,26E-02
Glyma.06G264000.1.Wm82.a4.v1	-3,5612	8,79E-02	0.0011222
Glyma.06G265000.2.Wm82.a4.v1	-2,3039	0.00032576	0.02017
Glyma.06G265100.1.Wm82.a4.v1	-1,8889	0.00015212	0.011812
Glyma.06G279000.1.Wm82.a4.v1	-4,8372	1,16E-04	3,60E-02
Glyma.06G311100.1.Wm82.a4.v1	-0,6549	3,41E-14	4,32E-11
Glyma.07G000600.1.Wm82.a4.v1	-3,8649	1,50E-03	3,33E-01
Glyma.07G017800.1.Wm82.a4.v1	-3,3845	0.00016695	0.012564
Glyma.07G027900.1.Wm82.a4.v1	-3,0836	0.00042316	0.024273
Glyma.07G055800.1.Wm82.a4.v1	-3,4188	0.0001035	0.0087002
Glyma.07G073700.3.Wm82.a4.v1	-1,6758	0.0010596	0.047938
Glyma.07G077700.2.Wm82.a4.v1	-4,2538	6,24E-15	9,48E-12
Glyma.07G080001.1.Wm82.a4.v1	-3,2698	0.0002836	0.018229
Glyma.07G082100.1.Wm82.a4.v1	-3,7433	2,32E-01	0.0025666
Glyma.07G083800.3.Wm82.a4.v1	-4,1455	2,20E-02	0.00033256
Glyma.07G091800.1.Wm82.a4.v1	-0,1102	0.000244	0.016375
Glyma.07G094000.1.Wm82.a4.v1	-5,6618	9,11E-14	8,94E-10
Glyma.07G094051.1.Wm82.a4.v1	-3,9221	1,25E-02	0.00020073
Glyma.07G111300.1.Wm82.a4.v1	-3,2063	2,92E-02	0.00042433
Glyma.07G133900.1.Wm82.a4.v1	-2,6146	0.00018568	0.013505
Glyma.07G174800.1.Wm82.a4.v1	-1,8689	1,23E-02	0.00019967
Glyma.07G203100.1.Wm82.a4.v1	-4,0457	3,55E-03	7,01E-01
Glyma.07G236100.1.Wm82.a4.v1	-3,2653	0.00028762	0.01837
Glyma.07G249700.4.Wm82.a4.v1	-3,1664	0.00045222	0.025413
Glyma.07G253500.1.Wm82.a4.v1	-3,7037	3,32E-02	0.0035324
Glyma.08G014200.1.Wm82.a4.v1	-4,0261	2,35E-02	0.00035184
Glyma.08G017600.1.Wm82.a4.v1	-3,2555	0.00011088	0.0091354
Glyma.08G097300.1.Wm82.a4.v1	-3,0047	0.00015166	0.011812
Glyma.08G119600.1.Wm82.a4.v1	-0,2628	0.00071952	0.035743
Glyma.08G126100.1.Wm82.a4.v1	-1,8522	3,80E-03	0.0005445
Glyma.08G149700.1.Wm82.a4.v1	-0,3946	7,13E-02	0.00093899
Glyma.08G177700.1.Wm82.a4.v1	-3,3739	0.00013043	0.010435
Glyma.08G201400.1.Wm82.a4.v1	-3,0758	0.00026835	0.01747
Glyma.08G216200.1.Wm82.a4.v1	-4,2427	1,07E-05	3,92E-03
Glyma.08G220000.1.Wm82.a4.v1	-1,3595	0.00072473	0.035885
Glyma.08G221300.1.Wm82.a4.v1	-2,6646	9,38E-02	0.0011887
Glyma.08G261450.1.Wm82.a4.v1	-4,4291	2,13E-03	4,45E-01
Glyma.08G261500.4.Wm82.a4.v1	-7,2105	1,36E-19	4,31E-16
Glyma.08G273000.1.Wm82.a4.v1	-5,0511	1,09E-05	3,96E-03
Glyma.08G318000.1.Wm82.a4.v1	-0,2486	4,21E-01	0.0043054

Glyma.08G328600.2.Wm82.a4.v1	-3,4094	0.00015264	0.011812
Glyma.08G338500.1.Wm82.a4.v1	-3,2847	0.00028244	0.018192
Glyma.09G015900.1.Wm82.a4.v1	-3,1904	0.00041823	0.024082
Glyma.09G019800.1.Wm82.a4.v1	-1,0582	0.00018789	0.0136
Glyma.09G034700.2.Wm82.a4.v1	-3,3964	2,26E-01	0.0025329
Glyma.09G036700.1.Wm82.a4.v1	-2,9093	0.0003854	0.02284
Glyma.09G038500.1.Wm82.a4.v1	-2,3699	0.00015028	0.011745
Glyma.09G043400.2.Wm82.a4.v1	-2,5258	0.00015272	0.011812
Glyma.09G051100.1.Wm82.a4.v1	-1,8776	0.00098441	0.045277
Glyma.09G075100.4.Wm82.a4.v1	-4,8858	5,29E-05	1,68E-02
Glyma.09G107600.1.Wm82.a4.v1	-3,9547	6,41E-02	0.00087155
Glyma.09G109700.1.Wm82.a4.v1	-0,2963	0.00083759	0.040228
Glyma.09G140050.1.Wm82.a4.v1	-3,9719	1,19E-02	0.00019399
Glyma.09G189500.1.Wm82.a4.v1	-4,2924	1,41E-05	4,93E-03
Glyma.09G203200.1.Wm82.a4.v1	-1,8091	1,04E-03	2,49E-01
Glyma.09G222300.1.Wm82.a4.v1	-2,3905	0.00010818	0.0090105
Glyma.10G003900.1.Wm82.a4.v1	-1,7199	0.00065975	0.033319
Glyma.10G059500.1.Wm82.a4.v1	-1,9239	8,69E-03	0.001115
Glyma.10G122800.1.Wm82.a4.v1	-3,6518	3,91E-01	0.0040479
Glyma.10G158300.2.Wm82.a4.v1	-5,4314	1,87E-11	1,54E-08
Glyma.10G158350.1.Wm82.a4.v1	-2,9215	0.00075816	0.03728
Glyma.10G164400.2.Wm82.a4.v1	-1,8423	0.00034273	0.021007
Glyma.10G187700.1.Wm82.a4.v1	-2,8404	0.00089588	0.04222
Glyma.10G238550.1.Wm82.a4.v1	-4,7416	1,70E-04	5,02E-02
Glyma.10G269200.1.Wm82.a4.v1	-0,000236	0.00010239	0.0086471
Glyma.10G278400.1.Wm82.a4.v1	-1,4363	0.00086833	0.041233
Glyma.10G282200.1.Wm82.a4.v1	-2,7602	3,32E-02	0.0035324
Glyma.11G012400.1.Wm82.a4.v1	-3,0122	0.00082469	0.039924
Glyma.11G053200.1.Wm82.a4.v1	-2,2689	0.00086937	0.041233
Glyma.11G080800.1.Wm82.a4.v1	-1,6401	0.00032296	0.020079
Glyma.11G088000.1.Wm82.a4.v1	-2,8128	0.00048005	0.026392
Glyma.11G093400.1.Wm82.a4.v1	-2,0174	0.0011005	0.048916
Glyma.11G117400.1.Wm82.a4.v1	-1,0859	0.00018497	0.013505
Glyma.11G137151.1.Wm82.a4.v1	-4,7514	9,72E-07	4,28E-04
Glyma.11G170202.1.Wm82.a4.v1	-5,8864	1,42E-09	9,57E-07
Glyma.11G189600.1.Wm82.a4.v1	-0,2833	0.00031318	0.019571
Glyma.11G196500.3.Wm82.a4.v1	-0,3095	0.00015963	0.012174
Glyma.11G204500.1.Wm82.a4.v1	-3,9817	7,91E-03	0.00013819
Glyma.11G253100.6.Wm82.a4.v1	-3,6894	3,11E-01	0.003356
Glyma.12G002500.2.Wm82.a4.v1	-1,9794	0.00018441	0.013505
Glyma.12G002900.2.Wm82.a4.v1	-2,8066	0.0004701	0.026128

Glyma.12G007200.1.Wm82.a4.v1	-2,7848	6,23E-01	0.005997
Glyma.12G012600.1.Wm82.a4.v1	-1,8676	0.00025585	0.016946
Glyma.12G015300.1.Wm82.a4.v1	-1,6653	0.00011073	0.0091354
Glyma.12G032000.1.Wm82.a4.v1	-3,4801	9,81E-02	0.0012373
Glyma.12G078700.1.Wm82.a4.v1	-1,6618	0.00052018	0.027842
Glyma.12G080300.1.Wm82.a4.v1	-1,6761	0.00027896	0.018007
Glyma.12G090250.1.Wm82.a4.v1	-3,1174	0.00056897	0.029722
Glyma.12G108200.3.Wm82.a4.v1	-3,2797	0.0001438	0.011415
Glyma.12G142300.1.Wm82.a4.v1	-3,0423	0.00061988	0.031763
Glyma.12G164000.1.Wm82.a4.v1	-4,1994	1,85E-03	0.00028535
Glyma.12G187600.1.Wm82.a4.v1	-3,2178	0.00036658	0.022069
Glyma.12G238600.1.Wm82.a4.v1	-3,8392	2,37E-02	0.00035288
Glyma.13G001100.1.Wm82.a4.v1	-3,2942	0.0002476	0.01658
Glyma.13G004200.3.Wm82.a4.v1	-3,0009	0.00095039	0.043989
Glyma.13G027700.1.Wm82.a4.v1	-3,0095	0.00069568	0.034672
Glyma.13G034000.1.Wm82.a4.v1	-6,8457	1,42E-19	4,31E-16
Glyma.13G043800.1.Wm82.a4.v1	-0,1542	0.0008001	0.038982
Glyma.13G053600.2.Wm82.a4.v1	-2,2427	0.00011325	0.0093055
Glyma.13G110250.1.Wm82.a4.v1	-3,6795	4,41E-01	0.0044443
Glyma.13G133900.1.Wm82.a4.v1	-1,0875	0.0010768	0.048147
Glyma.13G136500.7.Wm82.a4.v1	-2,7316	0.00045551	0.025504
Glyma.13G183700.1.Wm82.a4.v1	-2,3849	0.00014911	0.011684
Glyma.13G188300.1.Wm82.a4.v1	-0,2187	0.00030545	0.019187
Glyma.13G201100.1.Wm82.a4.v1	-1,6757	0.00016003	0.012174
Glyma.13G207800.1.Wm82.a4.v1	-0,1704	0.00091524	0.042874
Glyma.13G251400.1.Wm82.a4.v1	-2,8258	0.00011697	0.0095042
Glyma.13G258951.1.Wm82.a4.v1	-3,5686	3,54E-01	0.0037388
Glyma.13G260200.1.Wm82.a4.v1	-6,6853	1,86E-13	2,10E-11
Glyma.13G275600.3.Wm82.a4.v1	-1,8089	0.0010637	0.04799
Glyma.13G296900.2.Wm82.a4.v1	-3,5296	8,04E-01	0.0072282
Glyma.13G323800.1.Wm82.a4.v1	-0,00032	0.00028474	0.018234
Glyma.13G334500.1.Wm82.a4.v1	-2,8317	0.00014352	0.011415
Glyma.14G041100.1.Wm82.a4.v1	-3,1727	0.00043816	0.024992
Glyma.14G057400.1.Wm82.a4.v1	-2,0062	0.00050163	0.027223
Glyma.14G058500.1.Wm82.a4.v1	-0,2098	0.00051874	0.027814
Glyma.14G065100.3.Wm82.a4.v1	-2,6671	0.00029585	0.018778
Glyma.14G081800.1.Wm82.a4.v1	-1,7155	0.00029498	0.018761
Glyma.14G082400.2.Wm82.a4.v1	-1,7432	0.00058889	0.0305
Glyma.14G120900.1.Wm82.a4.v1	-3,9005	7,55E-03	0.00097623
Glyma.14G131651.1.Wm82.a4.v1	-4,5515	9,09E-04	2,28E-01
Glyma.14G212600.2.Wm82.a4.v1	-4,7152	4,38E-05	1,40E-02

Glyma.15G021300.1.Wm82.a4.v1	-4,8982	4,10E-05	1,34E-02
Glyma.15G023800.1.Wm82.a4.v1	-1,8893	0.00034875	0.021333
Glyma.15G026000.1.Wm82.a4.v1	-2,8405	0.00088581	0.041882
Glyma.15G034500.2.Wm82.a4.v1	-3,4058	1,39E-02	0.0002185
Glyma.15G044200.1.Wm82.a4.v1	-1,2815	0.00037892	0.022544
Glyma.15G044400.1.Wm82.a4.v1	-3,2933	0.0001979	0.01409
Glyma.15G070300.1.Wm82.a4.v1	-2,4069	6,84E-04	0.00012024
Glyma.15G106000.1.Wm82.a4.v1	-4,0689	5,76E-03	0.00010306
Glyma.15G115300.1.Wm82.a4.v1	-3,9247	9,44E-03	0.00015952
Glyma.15G139801.1.Wm82.a4.v1	-3,4309	5,68E-01	0.0055037
Glyma.15G154700.1.Wm82.a4.v1	-2,7144	1,74E-01	0.0020467
Glyma.15G169100.1.Wm82.a4.v1	-2,9993	1,58E-04	4,79E-02
Glyma.15G173300.2.Wm82.a4.v1	-1,6773	0.00021503	0.015063
Glyma.15G211200.4.Wm82.a4.v1	-3,5881	5,65E-01	0.0054916
Glyma.15G213000.3.Wm82.a4.v1	-4,2511	8,43E-03	0.00014478
Glyma.15G226200.1.Wm82.a4.v1	-3,4061	0.00014665	0.01158
Glyma.15G228500.1.Wm82.a4.v1	-0,3991	5,27E-04	1,37E-01
Glyma.15G231500.2.Wm82.a4.v1	-2,4274	0.00026823	0.01747
Glyma.15G244400.1.Wm82.a4.v1	-4,6104	4,95E-04	1,30E-01
Glyma.15G247800.1.Wm82.a4.v1	-0,3125	0.00041823	0.024082
Glyma.15G263266.1.Wm82.a4.v1	-3,0796	0.00046837	0.02608
Glyma.15G263800.1.Wm82.a4.v1	-3,1253	0.00055281	0.029085
Glyma.16G040100.1.Wm82.a4.v1	-1,4777	8,35E-01	0.0074197
Glyma.16G098100.1.Wm82.a4.v1	-4,8545	1,81E-06	7,44E-04
Glyma.16G109600.1.Wm82.a4.v1	-5,5101	5,26E-08	3,02E-05
Glyma.16G128200.1.Wm82.a4.v1	-2,6625	6,52E-06	2,54E-03
Glyma.16G149600.1.Wm82.a4.v1	-4,5398	3,91E-04	1,05E-01
Glyma.16G169700.1.Wm82.a4.v1	-2,0533	0.0011272	0.049984
Glyma.16G214100.1.Wm82.a4.v1	-4,3166	2,95E-11	2,36E-08
Glyma.16G214500.1.Wm82.a4.v1	-3,8278	1,12E-02	0.00018278
Glyma.17G110100.1.Wm82.a4.v1	-3,1097	0.00022488	0.015538
Glyma.17G126500.1.Wm82.a4.v1	-3,1681	6,60E-01	0.0061945
Glyma.17G158900.1.Wm82.a4.v1	-2,1856	0.0010049	0.04615
Glyma.17G162500.1.Wm82.a4.v1	-1,8408	0.00018493	0.013505
Glyma.17G162633.1.Wm82.a4.v1	-2,2284	0.00054178	0.028696
Glyma.17G171400.2.Wm82.a4.v1	-2,7642	0.00053408	0.028337
Glyma.17G189751.1.Wm82.a4.v1	-4,1599	1,68E-02	0.00026126
Glyma.17G191800.1.Wm82.a4.v1	-3,4794	0.00010104	0.0085565
Glyma.17G207600.1.Wm82.a4.v1	-3,4787	9,99E-01	0.0084815
Glyma.17G209900.1.Wm82.a4.v1	-0,4703	2,89E-04	7,90E-02
Glyma.17G210100.3.Wm82.a4.v1	-5,5081	9,00E-08	4,80E-05

Glyma.17G210200.1.Wm82.a4.v1	-3,9889	5,20E-02	0.00072535
Glyma.17G220900.1.Wm82.a4.v1	-3,0765	0.00025023	0.01672
Glyma.17G245000.1.Wm82.a4.v1	-3,1735	0.00044085	0.025099
Glyma.17G253200.3.Wm82.a4.v1	-0,2844	0.00097737	0.045021
Glyma.17G257200.1.Wm82.a4.v1	-2,0686	0.00019153	0.013749
Glyma.17G262300.1.Wm82.a4.v1	-2,9498	0.0010639	0.04799
Glyma.18G018600.1.Wm82.a4.v1	-1,9535	0.00054718	0.028931
Glyma.18G043400.1.Wm82.a4.v1	-3,1563	0.00047253	0.026167
Glyma.18G043600.1.Wm82.a4.v1	-1,6043	0.00041637	0.024082
Glyma.18G055400.1.Wm82.a4.v1	-3,0051	0.00093319	0.043524
Glyma.18G071900.6.Wm82.a4.v1	-2,6534	1,10E-03	2,59E-01
Glyma.18G137100.1.Wm82.a4.v1	-1,0988	0.00040428	0.023522
Glyma.18G138700.3.Wm82.a4.v1	-0,2685	0.00038905	0.023011
Glyma.18G139000.2.Wm82.a4.v1	-3,3543	0.00018563	0.013505
Glyma.18G191600.1.Wm82.a4.v1	-2,2577	0.00057556	0.029962
Glyma.18G203100.1.Wm82.a4.v1	-3,0257	0.00086694	0.041233
Glyma.18G209000.11.Wm82.a4.v1	-0,4715	2,05E-04	5,83E-02
Glyma.18G220500.1.Wm82.a4.v1	-1,5815	0.00024133	0.016304
Glyma.18G228500.1.Wm82.a4.v1	-1,7163	6,46E-01	0.0061165
Glyma.18G244900.1.Wm82.a4.v1	-1,4855	0.00062059	0.031763
Glyma.18G272700.1.Wm82.a4.v1	-1,8817	0.00015507	0.011905
Glyma.18G283100.1.Wm82.a4.v1	-1,0027	0.00021584	0.015084
Glyma.19G010767.1.Wm82.a4.v1	-2,4402	0.00035996	0.021843
Glyma.19G012300.11.Wm82.a4.v1	-2,8677	0.0010785	0.048147
Glyma.19G032500.1.Wm82.a4.v1	-1,2313	0.0008633	0.041202
Glyma.19G069300.1.Wm82.a4.v1	-3,9456	1,89E-03	0.00029082
Glyma.19G085600.1.Wm82.a4.v1	-3,8707	1,15E-02	0.0014188
Glyma.19G135600.1.Wm82.a4.v1	-0,1633	0.00040397	0.023522
Glyma.19G184900.1.Wm82.a4.v1	-4,2491	1,33E-02	0.0002114
Glyma.19G207300.2.Wm82.a4.v1	-1,5311	5,31E-02	0.0052225
Glyma.19G261600.1.Wm82.a4.v1	-2,9205	0.00032694	0.020202
Glyma.20G094300.1.Wm82.a4.v1	-3,3069	0.00023569	0.016059
Glyma.20G094400.4.Wm82.a4.v1	-4,9406	3,30E-06	1,10E-02
Glyma.20G102851.1.Wm82.a4.v1	-5,1852	7,94E-06	3,02E-03
Glyma.20G106000.1.Wm82.a4.v1	-3,9432	6,80E-02	0.00090725
Glyma.20G204000.1.Wm82.a4.v1	-2,3557	0.00021134	0.014891
Glyma.U032405.1.Wm82.a4.v1	-1,7939	0.00021331	0.014977