Aplicación de nuevas técnicas de mejoramiento y genómica para el desarrollo de marcadores moleculares de tolerancia a sequía en soja

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

Josefina Demicheli Ingeniera Agrónoma - Facultad de Agronomía de la Universidad de Buenos Aires - 2017

Cátedra de Bioquímica de la Facultad de Agronomía, UBA



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



COMITÉ CONSEJERO

Director de Tesis **Eduardo A. Pagano** Ingeniero Agrónomo (Facultad de Agronomía, UBA) Dr. 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) Dr. en Biología (UBA)

JURADO DE TESIS

JURADO Hugo Permingeat Doctor por la Universidad Nacional de Rosario, Argentina

JURADO María Fabiana Drincovich Doctora por la Universidad Nacional de Rosario, Argentina

JURADO Victoria Bonnecarrère

Doctora por la Universidad de la República, Uruguay

Fecha de Defensa Tesis: 30 de noviembre de 2022.

A Rafael, mi mayor logro e inspiración.

Agradecimientos

A Joaquín, por el apoyo, la paciencia y los fines de semanas de mediciones en el invernadero y el laboratorio.

A mis papás, Fernanda y José, por darme la posibilidad de elegir este camino y transitarlo sin preocupaciones.

A mi hermana Juana, por la compañía y la cantidad de lecciones tomadas.

A Eduardo, por la guía y confianza.

A mis compañeras y compañeros de Bioquímica por enseñarme, ayudarme y guiarme en este proceso.

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 lo he presentado, en forma parcial o total, como una tesis en ésta u otra institución.

Ing. Agr. Josefina Demicheli

Publicaciones derivadas de la tesis

Artículos académicos en proceso

 Wide screening assay for rapid drought response charachterization in soybean. Presentación en "Journal of Experimental Botany", estipulada para agosto 2022. Trabajo derivado del Capítulo 4 de esta tesis.

Índice general

3.4.

1. Introducción General 1 1. Importancia económica y social de la sequía en el cultivo de soja 1 2. Importancia agronómica de la sequía en el cultivo de soja 2 3. 3 Respuesta a la sequía 3.1. 6 7 4. Objetivos e Hipótesis 4.1. Objetivo general 7 4.2. 7 4.3. 7 2. Materiales y Métodos Generales 9 Elección de los genotipos y material botánico 1. 9 2. Diseño experimental y tratamientos 10 3. Comparación de los genotipos con respuestas contrastantes a la sequía 13 1. 13 2. Materiales y Métodos 14 2.1. 14 2.2. Metodología Estadística 17 18 3. 3.1. 18 3.2. 20 Fotosíntesis y parámetros asociados 3.3.

Microscopía.....

24

25

	4.	Discusión	27
	5.	Conclusiones	30
4.	Estu	idio Transcriptómico	31
	1.	Introducción	31
	2.	Materiales y Métodos	32
		2.1. Diseño del experimento	32
		2.2. Extracción del RNA y control de calidad	33
		2.3. Secuenciación	33
		2.4. Procesamiento y análisis de los datos	33
	3.	Resultados	35
		3.1. Controles de calidad	35
		3.2. Genes diferencialmente expresados	36
		3.3. Comparaciones entre genotipos contrastantes	48
	4.	Discusión	54
	5.	Conclusiones	58
5.	5. Dise	Conclusiones	58 59
5.	5. Dise 1.	Conclusiones	58 59 59
5.	 5. Dise 1. 2. 	Conclusiones	58 59 59 60
5.	 5. Dise 1. 2. 	Conclusiones	58 59 59 60 60
5.	 5. Dise 1. 2. 	Conclusiones	 58 59 60 60 60
5.	 5. Dise 1. 2. 3. 	Conclusiones	58 59 60 60 60 61
5.	 5. Dise 1. 2. 3. 	Conclusiones	 58 59 59 60 60 60 61
5.	 5. Dise 1. 2. 3. 	Conclusiones	 58 59 59 60 60 60 61 61 62
5.	 5. Dise 1. 2. 3. 4. 	Conclusiones	 58 59 59 60 60 60 61 61 62 67
5.	 5. Dise 1. 2. 3. 4. 5. 	Conclusiones	 58 59 60 60 60 61 61 62 67 69
5.	 5. Dise 1. 2. 3. 4. 5. Con 	Conclusiones	 58 59 60 60 60 61 61 62 67 69 71
5.	 5. Dise 1. 2. 3. 4. 5. Con 1. 	Conclusiones	 58 59 60 60 60 61 61 62 67 69 71 71

3. Perspectivas futuras	. 73
Bibliografía	75
Appendices	81
A. Genes diferencialmente expresados	83
B. Información complementaria de SNPs	111

Índice de figuras

1.1.	Rendimiento de soja (qq/ha) en función de las lluvias acumuladas (mm) para las principales regiones sojeras de la Argenetina. Gráfico tomado de la Bolsa de Cereales de Buenos Aires.	2
2.1.	Imágenes del sustrato utilizado en los experimentos	10
2.2.	Macetas sembradas en el invernáculo de la Facultad de Agronomía en los distintos experi- mentos.	11
3.1.	Escala de marchitamiento utilizada	15
3.2.	Macetas embolsadas y el registro de su peso diario.	15
3.3.	Días que demoró cada genotipo en llegar a un 100 en la escala de marchitamiento. Las barras corresponden al error estándar, letras distintas indican una diferencia con un p< $0,05$. (GS: Genotipo Susceptible, GT: Genotipo Tolerante, PI: PI 416937, W: Williams 82)	18
3.4.	Estado de las plantas de los cuatro genotipos en condición de sequía a 15 días del último riego	19
3.5.	Estado de las plantas del GT bien regadas (izquierda) y sin riego (derecha) por 18 y 19 días.	19
3.6.	Marchitamiento en función del contenido hídrico relativo (RSWC) para los cuatro genoti- pos. El color de los puntos indica la severidad del marchitamiento, siendo gris para valores cercanos a 0, amarillo para valores intermedios (50) y rojo para plantas muertas (100).	20
3.7.	RSWC (contenido relativo de agua) en función de los días desde el último riego para los cuatro genotipos. (Izquierda) Cada panel muestra todas las repeticiones de cada genotipo, el color de los puntos varía según el RSWC. (Derecha) Un gráfico para todos los genotipos, los puntos muestran el promedio, la línea un modelo lineal suavizado y el sombreado los intervalos de confianza para cada línea.	21
3.8.	Transpiración acumulada (L) a través de los días para todas las repeticiones de los cuatro genotipos estudiados. El panel superior muestra las muestras bien regadas, el inferior las sometidas a sequía, el número en la etiqueta indica el número de muestra.	21
3.9.	RSWC (contenido relativo de agua) en función de los días pasados desde el último. Cada panel contiene las repeticiones de un genotipo, el número en la etiqueta corresponde al número de cada muestra.	22
3.10.	Transpiración acumulada por los distintos genotipos bajo condiciones de sequía y control. (GS: Genotipo Susceptible, GT: Genotipo Tolerante, PI: PI 416937, W: Williams 82)	23

3.11.	Índice normalizado de transpiración (NTR) en función del contenido relativo de agua (RSWC) para cada genotipo. Los valores a la derecha de cada gráfico corresponden al $RSWC_C$ (crítico) o X0. Además, se muestra el R^2 ajustado.	23
3.12.	Evolución del RSWC (contenido hídrico relativo) para los genotipos GS y GT (Susceptible y Tolerante).	24
3.13.	Eficiencia del uso del agua (WUE) en función de la irradiancia para los 4 tratamientos 10 días después del último riego. Los asteriscos indican diferencias significativas (p< $0,05$).	25
3.14.	Transpiración, conductancia estomática (g_s), tasa fotosintética y concentración interna de CO_2 (Ci) en función de la irradiancia (PPFD). Las barras representan el error estándar de la media y los asteriscos diferencias significativas (p <0.05). Para la tasa fotosintética se ajustó un modelo hiperbólico significativamente distinto para cada tratamiento.	26
3.15.	Imágenes tomadas de la microscopía electrónica de la epidermis de los tres genotipos, en detalle se muestra un estoma.	27
3.16.	Número de células epidérmicas para los genotipos contrastantes (GS y GT y el de referen- cia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias significativas con un valor $p < 0,05$.	27
3.17.	Densidad estomática (a) e índice estomático (b) para los genotipos contrastantes (GS y GT) y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias con un valor $p < 0.05$ cuando el factor "Genotipo", o la interacción entre ambos factores resultó significativa. Asteriscos indican diferencias significativas entre efectos principales del factor "Cara de la hoja".	28
3.18.	Ancho (a) y largo estomático (b) para los genotipos contrastantes (GS y GT) y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias con un valor $p < 0.05$ cuando el factor "Genotipo", o la interacción entre ambos factores resultó significativa. Asteriscos indican diferencias significativas entre efectos principales del factor "Cara de la hoja".	28
3.19.	Índice de apertura estomático (SAI) para los genotipos contrastantes (GS y GT) y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias con un valor $p < 0,05$ cuando el factor "Genotipo", o la interacción entre ambos factores resultó significativa. Asteriscos indican diferencias significativas entre efectos principales del factor Çara de la hoja".	28
4.1.	Esquema general de la secuenciación.	32
4.2.	Pasos seguidos en el análisis de datos y la metodología utilizada en cada uno	34
4.3.	Gel de agarosa 1,5 % del RNA extraído. El nombre y número de cada muestra se encuentra en cada well.	35
4.4.	Resultado del análisis de clusters de los FPKM para cada tratamiento, utilizando el log10(FPKM El color rojo indica genes que disminuyeron su expresión y el verde genes que la aumen- taron. El rango de colores del verde al rojo indica el nivel de expresión de mayor a menor.	/I +1).
		36
4.5.	Volcano plot. Expresión diferencial de genes para el genotipo tolerante. Nivel de significan- cia expresado como el -log10(padj) en función del nivel de expresión diferencial medido como el log2(Fold Change). Las etiquetas muestran el número de genes regulados hacia arriba (Up) y regulados hacia abajo (Down).	37

4.6.	Las 10 categorías funcionales más significativas enriquecidas para cada sub-ontología de GO (BP= procesos biológicos, CC= componente celular, MF= función molecular) para la comparación del GT en sequía con respecto a su control. En rojo (izquierda) aquellas que resultaron subrepresentadas y en verde (derecha) aquellas que resultaron sobrerepresentadas. El número dentro de la barra indica la cantidad de genes pertenecientes a esa categoría que resultaron enriquecidos. Los asteriscos indican enriquecimientos significativos (padj<0,05).	38
4.7.	Diagrama de los componentes del ribosoma (KEGG gmx03010). El color corresponde al nivel de expresión para el genotipo tolerante bajo condiciones de sequía con respecto a su control.	40
4.8.	Volcano plot. Expresión diferencial de genes para el GS. Se muestra el nivel de signifi- cancia como el -log10(padj) en función del nivel de expresión diferencial medido como el log2(Fold Change). Las etiquetas muestran el número de genes regulados hacia arriba (Up) y regulados hacia abajo (Down).	41
4.9.	Las 10 categorías funcionales más significativas enriquecidas para cada sub-ontología de GO (BP= procesos biológicos, CC= componente celular, MF= función molecular) para la comparación del GS en sequía con respecto a su control. En rojo (izquierda) aquellas que resultaron subrepresentadas y en verde (derecha) aquellas que resultaron sobrerepresentadas. El número dentro de la barra indica la cantidad de genes pertenecientes a esa categoría que resultaron enriquecidos. Los asteriscos indican enriquecimientos significativos (padj<0,05).	42
4.10.	DAG (Directed acyclic graph) con los términos correspondientes a componentes celula- res de GO que disminuyeron su expresión significativamente en el GS frente al estrés por sequía. El color sombreado de los términos corresponde a su nivel de expresión	44
4.11.	Principales reacciones involucradas en la fijación de C en organismos fotosintéticos (<i>KEGG gmx 00710</i>). El color sombreado de los genes corresponde a su nivel de expresión. \ldots	45
4.12.	DAG (Directed acyclic graph) con los términos correspondientes a funciones moleculares de GO que disminuyeron su expresión significativamente en el GS frente al estrés por se- quía. El color sombreado de los términos corresponde a su nivel de expresión	46
4.13.	Principales reacciones involucradas en el metabolismo del N (<i>KEGG gmx 00910</i>). El color del sombreado de los genes corresponde a su nivel de expresión.	47
4.14.	Transducción de señales mediadas por fitohormonas (KEGG gmx 04075).El color sombrea- do de los genes corresponde a su nivel de expresión para el GS frente a la sequía	49
4.15.	Transducción de señales mediadas por fitohormonas (KEGG gmx_04075). El color sombreado de los genes corresponde a su nivel de expresión para el GT frente a la sequía	50
4.16.	Análisis preliminar de la expresión diferencial de factores de transcripción para los genoti- pos contrastantes.	52
4.17.	Familias de factores de transcripción que cambiaron significativamente su expresión en el GS bajo condiciones de sequía.	52
4.18.	Familias de factores de transcripción que cambiaron significativamente su expresión en el GT bajo condiciones de sequía.	53
4.19.	Gráfico de dispersión de las categorías enriquecidas resultantes del análisis de AgriGO y REViGO para el GT en situación sequía con respecto a su control. Los ejes representan espacios semánticos que agrupan las categorías similares.	55

4.20.	Gráfico de dispersión de las categorías enriquecidas resultantes del análisis de AgriGO y REViGO para el GS en situación sequía con respecto a su control. Los ejes representan espacios semánticos que agrupan las categorías similares.	56
5.1.	a : Diagrama de Venn mostrando los SNPs compartidos y exclusivos para cada genotipo. (GS= Genotipo Sensible, GT= Genotipo Tolerante). b : Dendrograma realizado a partir de la clusterización de los resultados del SoySNP6K chip. (GS= Genotipo Sensible, GT= Ge- notipo Tolerante, W=Williams)	61
5.2.	Información del impacto y función predichos por SNPeff para las variantes encontradas en la secuencia transcriptómica.	63
5.3.	Pipeline utilizado para el filtrado de variantes. El número indica la cantidad de variantes resultantes luego de el filtro especificado en la derecha.	64

Índice de cuadros

3.	1. Escala de marchitamiento utilizada	14
5.	 Información del SNP en la secuencia genómica que es polimórfico entre líneas y está en un gen diferencialmente expresado en el GT frente al estrés por sequía. (CHROMPOS= Núme- ro de cromosoma y posición donde cae el SNP, REF= alelo de referencia presente en GS y Williams, ALT= alelo alternativo presente en GT, log2FoldChange= nivel de expresión di- ferencial en el GT en respuesta a la sequía, padj= valor p ajustado de la exresión diferencial, locusName= identificación del locus polimórfico, chr= cromosoma donde se ubica, start= posición de inicio del gen, end= posición de finalización del gen, strand= hebra del ADN, descr= Descripción del gen). 	62
5.	2. Cantidad de SNPs e INDELs encontrados en el transcriptoma por cromosoma	62
5.	3. Listado de los 25 SNPs seleccionados para ser utilizados como marcadores moleculares. CHROMPOS= Número y posición del cromosoma en donde se encuentra el SNP, Gene- Name= ID del gen según la notación "Wm82.a4.v1", REF= alelo de la referencia y GS, ALT= alelo alternativo presente en el GT, EFFECT= efecto del SNP predicho por SNPeff, IMPACT= impacto del SNP predicho por SNPeff, log2FoldChange= Expresión diferencial en el GT en respuesta a la sequía, padj= valor p ajustado de la expresión diferencial, PFAM Names= Nombre del gen según la base Pfam, GO Biological Process Description= des- cripción 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	66
А	.1. Genes con expresión diferencial significativa para el GS (genotipo susceptible) en respues- ta a la sequía. GeneID= ID del gen en la notación Wm82.a4.v1, padj= valor p ajustado de la expresión diferencial, log2FoldChange= expresión diferencial en logaritmo, PFAM Description= Descripción del gen según la base de datos PFAM.	83
А	.2. Genes con expresión diferencial significativa para el GT (genotipo tolerante) en respues- ta a la sequía. GeneID= ID del gen en la notación Wm82.a4.v1, padj= valor p ajustado de la expresión diferencial, log2FoldChange= expresión diferencial en logaritmo, PFAM Description= Descripción del gen según la base de datos PFAM.	91
В	 Tabla con información de secuencias para el pedido de sondas KASP específicas de cada SNP. La secuencia comprende las regiones flanqueantes del SNP (50 pares de bases <i>ups-tream</i> y <i>downstream</i>), con el alelo de referencia y el alternativo entre corchetes, tal como 	

B.2. SNPs e INDELs alternativos a utilizarse como marcadores moleculares. CHROMPOS= Cromosoma y posición donde se encuentra, REF= alelo o variante de referencia presente en el GS, ALT= alelo o variante alternativa presente en el GT, EFFECT= efecto predicho por SNPeff, IMPACT= impacto predicho por SNPeff, GO Molecular Function Description= Descripción de la función moleular cumplida por el gen anotada en Gene Ontology. 113

Abreviaturas

Abreviatura	Completo
ABA	Ácido abscisico
A _{max}	Fotosíntesis a saturación lumínica
ANOVA	Análisis de la varianza
CC	Capacidad de campo
Ci	Carbono intracelular
DTR	Índice de transpiración diario, del inglés Daily Transpiration Ratio
eQTL	del inglés, expression QTL
EUR	Eficiencia de uso de los recursos
FBN	Fijación biológica de nitrógeno
FPKM	del inglés, fragments per kilobase of transcript sequence per millions base pairs sequenced.
FTSW	Fracción transpirable del agua del suelo, del inglés Fraction of Transpirable Soil Water
g _s	Conductancia estomática
GS	Genotipo susceptible
GT	Genotipo tolerante
GWAS	del inglés, Genome Wide Association Studies
HSP	del inglés, Heat shock proteins
IC	Índice de Cosecha
INDELs	Inserciones o Deleciones
LEA	del inglés, Late embryogensis abundant proteins
NGS	Secuenciación de nueva generación, del inglés Next Generation Sequencing
NTR	Índice de transpiración normalizado, del inglés Normalized Transpiration Ratio
PAR	Radiación Fotosintéticamente Activa, del inglés Photosynthetic Active Radiation
QTL	del inglés, Qauntitative Trait Locus
RILs	Líneas endocriadas recombinantes, del inglés Recombinant Inbred Lines
ROS	Especies reactivas de oxígeno, del inglés Reactive Oxygen Species
RSWC	Contenido hídrico relativo, del inglés Relative Soil Water Content
RSWCc	Contenido hídrico relativo crítico, del inglés critic Relative Soil Water Content
RTO	Rendimiento
SAI	Índice de apertura estomático, del inglés Stomatal Aperture Index
SNP	Polimorfismo de un solo nucleótido, del inglés Single Nucleotide Polimorphism
TA	Transpiración Acumulada
VPD	Déficit de presión de vapor
WUE	Eficiencia del uso del agua, del inglés Water Use Efficiency

Resumen

Aplicación de nuevas técnicas de mejoramiento y genómica para el desarrollo de marcadores moleculares de tolerancia a sequía en soja

El desarrollo de materiales de soja tolerantes a sequía es primordial dado que este cultivo explora diversos ambientes a lo largo del país y, además, se estima que debido al cambio climático los eventos de seguía se tornarán más frecuentes e intensos. La selección asistida por marcadores moleculares permite acelerar y hacer más eficientes los procesos de selección en programas de mejoramiento. Este trabajo tuvo como objetivo el desarrollo de marcadores moleculares de tolerancia a sequía en soja. Se partió de genotipos con comportamientos contrastantes a la sequía seleccionados del programa de mejoramiento de una semillera nacional. Primero, se caracterizó la respuesta a la sequía de ambos genotipos, y se encontró que el tolerante presentó un menor marchitamiento y una mayor eficiencia del uso del agua producto de una conservación del agua del suelo a través de un cierre parcial estomático a contenidos hídricos más altos. Luego, se realizó una RNAseq que develó que en el genotipo tolerante aumentó la expresión de genes y vías metabólicas propias de defensa al estrés, como la acumulación de solutos compatibles, procesos de señalización y vía del ABA. Contrariamente, el genotipo susceptible presentó disminuida la expresión de genes involucrados en el metabolismo primario y procesos claves para la homeostasis celular, como la fotosíntesis, degradación de almidón, glucólisis y expansión celular. Finalmente, se buscaron variantes en la secuencia genómica y transcriptómica que estuvieran asociadas a la respuesta del genotipo tolerante. Se construyó un panel de 26 SNPs, distribuidos en 16 genes que cambiaron su expresión en el genotipo tolerante frente a la seguía, entre los cuales hubo inhibidores de proteasas, enzimas de síntesis de solutos compatibles, activadores de la rubisco o con actividad oxidorreductasa. Previo a una validación, estos SNPs podrían utilizarse como marcadores moleculares para seleccionar materiales con una mayor tolerancia a la sequía.

PALABRAS CLAVE: estrés abiótico, genotipos contrastantes, estudio transcriptómico, genotipado.

Abstract

Application of new breeding techniques and genomics to obtain molecular markers for drought tolerant soybean.

Developing drought tolerant soybeans should be an outmost priority, since this crop explores a wide range of environments across the country, and experts predict that drought episodes are going to become more frequent and intense due to climate change. Marker assisted selection enables an acceleration and increased efficiency of breeding programs. The main objective of the current study was to develop molecular markers for selecting drought tolerant soybeans, using the variation present in two contrasting genotypes from the breeding program of a national seed company. Firstly, a characterization of the drought response was performed, which revealed that the tolerant genotype presented a slow wilting phenotype and a higher water use efficiency due to a soil water conservation strategy product of an early partial stomatal closure in response to drying soil. Later, an RNAseq was conducted which explained the phenotypes observed previously. Whilst the tolerant genotype presented up regulated genes and metabolic pathways involving a drought tolerant response, such as accumulation of compatible solutes, signaling processes and ABA mediated pathway; the sensitive genotype down regulated genes encoding for key proteins in cell homeostasis. These were involved in photosynthetic structural components, starch degradation, glycolysis, or cell wall expansion. Finally, a DNA microarray and transcriptomic variant calling was performed to detect SNPs, and those associated to the drought response of the tolerant genotype were selected. A panel of 26 molecular markers distributed across 16 different genes with differential expression in the tolerant genotype was constructed. These genes were protease inhibitors, enzymes involved in compatible solute synthesis, rubisco activators or with oxidoreductase activity. Prior validation, these SNPs might be used as molecular markers for drought tolerant soybeans selection in breeding programs.

KEYWORDS: abiotic stress, contrasting genotypes, transcriptomic analysis, genotyping.

Capítulo 1

Introducción General

1. Importancia económica y social de la sequía en el cultivo de soja

La Argentina es uno de los principales productores de soja a nivel mundial, se encuentra en tercer lugar detrás de Brasil y USA, quienes en conjunto agrupan el 80 % de la producción total de esta proteoleaginosa (Ybran y Lacelli, 2016; Purcell y col., 2017; Alam Khan, 2018). En Argentina es el cultivo más expandido, abarcando en la campaña 2021/22 la suma de casi 16 millones de hectáreas, un número ampliamente superior a las 8 millones de hectáreas destinadas a maíz o las 7 millones destinadas a trigo (Bolsa de Comercio de Rosario, 2022). La producción de soja tiene una singular importancia económica para el país, ya que es uno de los productos que provee mayores ingresos de divisas; no solo a través de la exportación de granos, sino también por el desarrollo de la industria de subproductos (especialmente harina y aceite) que provee un alto valor agregado a la cadena de producción.

La sequía es una de las principales causas de pérdidas en el rendimiento de los cultivos de grano en el mundo, y su impacto depende del momento, duración e intensidad de ocurrencia (Hua y col., 2018; Jenks, Hasegawa y Jain, 2007; Chamarthi y col., 2021). En Argentina, la devastadora sequía ocurrida en la campaña 17/18 fue considerada "la peor en 44 años" afectándose 62 millones de hectáreas (Ministerio de Agroin-dustria, 2018) y llevando a la declaración de una emergencia agropecuaria (La Nación, 2018b; La Nación, 2018c). Este desastre climático causó pérdidas millonarias a los productores, teniendo también repercusión en otros sectores agropecuarios como la ganadería y la lechería, al disminuir las reservas forrajeras y aumentar el precio de los granos destinados a la alimentación animal (Mira, 2018).

En el cultivo de soja, la sequía es considerado el estrés abiótico de mayor importancia debido a las catastróficas pérdidas que produce. Por ejemplo, en la campaña 17/18 en Santa Fe, una de las principales regiones sojeras, el 75 % del área de soja de segunda resultó afectada por la sequía. Tal fue el impacto del estrés que la cosecha se vio imposibilitada y se tuvieron que destinar aquellos lotes a la alimentación animal directa o el confeccionado de rollos para reserva (Hernández, 2018).

Este problema, además, es de difícil solución. Los eventos de sequía son difíciles de predecir y, más aún, difíciles de evitar o subsanar sus pérdidas. Teóricamente, la solución definitiva sería el riego, pero en la práctica esta opción cae por su propio peso. Es técnicamente imposible aplicar riego a las millones de hectáreas sembradas, y por sobre todas las cosas, completamente errado desde el punto de vista socio-ambiental. A través de distintas prácticas agronómicas, como la elección de la variedad, densidad y fecha de siembra (entre otros), se puede tratar de minimizar las pérdidas producidas por un evento de sequía. Sin embargo, la sequía puede afectar sin previo aviso a este, y cualquier otro, cultivo a lo largo de todo el ciclo del mismo.

Sumado a esto, el panorama es desalentador ya que a causa del cambio climático, distintos modelos estadísticos predicen que las sequías podrían eventualmente exacerbarse, y las precipitaciones ser cada vez más erráticas y dispares (Fritsche-Neto y Borém, 2012; Purcell y col., 2017). La demanda de alimento,



Figura 1.1: Rendimiento de soja (qq/ha) en función de las lluvias acumuladas (mm) para las principales regiones sojeras de la Argenetina. Gráfico tomado de la Bolsa de Cereales de Buenos Aires.

sin embargo, aumenta diariamente junto con el crecimiento demográfico, el cual ha llegado a una tasa del 1 % anual en nuestro país. Es por esto que se hace evidente la necesidad de incrementar el rendimiento de los cultivos, no solo en áreas que ocasionalmente son afectadas por sequías, sino en zonas marginales con climas áridos y semiáridos y suelos con regímenes ústicos y arídicos, las cuales abundan en el país.

2. Importancia agronómica de la sequía en el cultivo de soja

El nivel de precipitaciones impacta directamente en el rendimiento de la soja en Argentina, alcanzándose mayores rendimientos en los años más lluviosos. Este nivel de precipitaciones impacta de distinta manera a las principales regiones sojeras del país, siendo algunas más dependientes y otras más estables ante el ingreso de agua (Ibarguren, 2021). Por ejemplo, Buenos Aires presenta un rendimiento más estable ante el nivel de precipitaciones, con una media de 28,5 qq/ha. Santa Fe y Córdoba son las provincias con mayor potencial de rendimiento, pudiendo superar los 37 y 34 qq/ha respectivamente, pero este rendimiento es sumamente dependiente de las lluvias acumuladas, pudiendo disminuir casi un 40 % en años secos. Por otro lado, La Pampa presenta un potencial de rendimiento mucho menor dado que su nivel de precipitaciones es menor también. Por último, Entre Ríos tiene el menor rendimiento medio de entre todas las regiones, y este también es altamente variable según el nivel de lluvias acumuladas. Es evidente entonces que el rendimiento del cultivo está directamente influenciado por el nivel de precipitaciones alcanzado.

Cuando hablamos del impacto de la escasez de agua en los cultivos es importante distinguir la diferencia entre sequía y estrés hídrico. La sequía es un evento meteorlógico provocado por escasas precipitaciones. Sin embargo, el estrés hídrico no es necesariamente provocado por un déficit hídrico. Por ejemplo, el agua puede estar en cantidades suficientes, pero su calidad ser limitante, resultando en un estrés hídrico al verse imposibilitada la absorción por las raíces. Las bajas temperaturas también pueden provocar un estrés hídrico al salir el agua de las células y formar cristales de hielo en los espacios intercelulares. Asimismo, hasta plantas que se encuentran con una buena disposición hídrica pueden sufrir eventos de déficit a lo largo del día, disminuyendo la turgencia en horas del mediodía (Buchanan, Gruissem y Jones, 2015).

La sequía es el estrés abiótico que causa mayores pérdidas en el cultivo de soja. Esto se debe a que esta puede aparecer en cualquier momento el ciclo, afectando desde la germinación hasta la maduración de los granos (Chen y col., 2016). En zonas donde el agua es considerada un factor limitante para la producción, el rendimiento puede definirse como:

RTO = TxEUAxIC

(RTO: Rendimiento, T: Transpiración, EUA: Eficiencia en el uso del agua, IC: Índice de Cosecha)

Históricamente se han logrado grandes incrementos en el rendimiento a través de un aumento en el índice de cosecha, pero hoy en día, para los cultivos de grano usualmente producidos, este estaría alcanzando el máximo. Es por esto entonces, que el desafío para los programas de mejoramiento reside en aumentar los rindes a través de una mayor producción de biomasa combinada con una baja utilización de agua. Es decir, aumentar la eficiencia del uso del agua (Fritsche-Neto y Borém, 2012), lo que se conoce en inglés como *"more crop per drop"*.

Meteorológicamente, la sequía indica un período prolongado libre de lluvias, situación que se observa a menudo en los ambientes agrícolas. Cuando esto ocurre, en la planta se genera un desbalance hídrico, donde las salidas de agua por transpiración superan las entradas por absorción desde la raíz. A causa de este desbalance hídrico, las plantas exponen una serie de respuestas que varían en complejidad y escala temporal. En primer lugar, en cuestión de minutos, se produce un reciclado de algunas proteínas y movimiento estomático. En segundo lugar, al transcurrir las horas, se producen *heat shock proteins*, movimiento foliar, marchitamiento y ajuste osmótico en respuesta al ABA (ácido abscisíco, Buchanan, Gruissem y Jones, 2015). Al cabo de uno o dos días, se produce una rustificación celular, se inducen los llamados genes de mantenimiento y se acelera la iniciación floral, floración o cuaje de granos. Finalmente, si la sequía se prolonga por varios días o semanas, se afecta el desarrollo y crecimiento del canopeo y se acelera la senescencia foliar. Consecuentemente, las reducciones del rendimiento en un corto plazo pueden deberse a una reducida conductancia estomática que genera una disminución en la tasa fotosintética. Sin embargo, a largo plazo, las reducciones en el rendimiento parecerían estar más asociadas a la menor área foliar producto de una disminución en la tasa de crecimiento y una acelerada senescencia (Fritsche-Neto y Borém, 2012).

Otra manera de definir el rendimieno es según sus componentes ecofisiológicos, siendo estos:

RTO = PARabsxEURxIC

(RTO: Rendimiento, PARabs: Radiación fotosintéticamente activa absorbida, EUR: Eficiencia de uso de los recursos, IC: Índice de Cosecha).

La sequía afecta todos y cada uno de estos parámetros. La radiación absorbida se ve limitada a causa de una reducción en el área foliar, producto de una disminución en el crecimiento vegetativo y división celular. Dichos procesos requieren un aumento de la turgencia, por lo que se ven disminuidos ante un desbalance hídrico, donde la elongación celular puede cesarse por una interrupción del flujo de agua del xilema hacia las células vecinas (Farooq y col., 2009). Otro mecanismo por el cual se disminuye la radiación absorbida, en el corto plazo, es el movimiento foliar. La pérdida de turgencia lleva al marchitamiento, y en caso de prolongarse, se intensifica a través de la senescencia foliar. El aumento en la resistencia estomática; sobre todo en soja, al ser una planta isohídrica con alta sensibilidad estomática (Jenks, Hasegawa y Jain, 2007; Satorre y col., 2003), resulta en una disminución de la tasa fotosintética y producción de especies reactivas de oxígeno (ROS). Además, aumenta la temperatura foliar que provoca subas en los costos de mantenimiento, la respiración y fotorrespiración; disminuyendo así la eficiencia en el uso de los recursos (Fritsche-Neto y Borém, 2012). El índice de cosecha también se ve afectado, soliendo ser menor en los cultivos de secano que en aquellos que crecen en condiciones óptimas, ya que este se relaciona con la proporción de transpiración que ocurre post-antesis, y la contribución relativa del agua almacenada en el suelo es menor en el período de interés (Satorre y col., 2003).

3. Respuesta a la sequía

Las plantas cuentan con tres grandes mecanismos para lidiar con la sequía: el escape, la evitación y la tolerancia (Chen y col., 2016; Manavalan y col., 2009; Sanchez y col., 2012; Zhao, Aleem y Sharmin, 2018).

El primero consiste en completar el ciclo de vida cuando el agua abunda, antes de un evento de sequía, teniendo inevitablemente un ciclo de menor duración. El segundo mecanismo consiste en evitar eventos de sequía al mantener un mejor estado hídrico, ya sea por una mayor y más eficiente absorción de agua desde las raíces, o por una menor transpiración o pérdida de agua por las partes aéreas. El último mecanismo, la tolerancia, le permite a la planta seguir metabolizando normalmente aún con bajos potenciales agua o mantener la turgencia a través del ajuste osmótico.

La respuesta a la sequía es sumamente compleja ya que está afectada por factores tanto del estrés, como de la planta en sí. Es decir, las características particulares del estrés como la duración, la severidad; y de la planta, como el estado fenológico, la aclimatación por estreses previos y el genotipo, condicionan la respuesta (Carrera y col., 2021). Además, cada respuesta esta sustentada por una variedad de genes y rutas metabólicas. Por lo tanto, no hay un solo mecanismo, o gen, posible que explique la tolerancia al estrés, sino que hay que contemplarlos a todos en conjunto (Buchanan, Gruissem y Jones, 2015; Prince y col., 2015).

Una alta WUE, producto del mantenimiento de una transpiración óptima, es una de las estrategias utilizadas para el incremento del rendimiento en soja. La fenología, la sensibilidad fotoperiódica, la plasticidad, el mantenimiento del área foliar, la tolerancia a altas temperaturas, el ajuste osmótico, el vigor en etapas tempranas, la longitud y densidad radical, son distintos atributos que están asociados positivamente con un mantenimiento de la transpiración durante la sequía. Sin embargo, cuales de estos rasgos serán incorporados en un plan de mejoramiento dependerá del ambiente objetivo para el cuál se desea mejorar.

Desde el punto de vista fisiológico y bioquímico, estos rasgos pueden ser:

Rasgos asociados a la raíz - La morfología y la plasticidad radical son atributos involucrados en la evitación de un evento de sequía. Una raíz principal más profunda acompañada de raíces laterales fibrosas pero moderadas, ayudan a las plantas a extraer agua de un mayor volumen de suelo de manera eficiente (Fried, Narayanan y Fallen, 2019; Manavalan y col., 2009). La elongación de la raíz principal en estadios tempranos, aún en condiciones no estresantes, es un atributo a tener en cuenta ya que indica la capacidad de enraizamiento de un genotipo. Se han encontrado correlaciones positivas entre la plasticidad radical y el buen comportamiento ante la sequía. Ante la falta de agua, las plantas resistentes aumentan la relación raíz:tallo, la longitud y peso seco da la raíz y el número de raíces laterales (Zhao, Aleem y Sharmin, 2018).

Aunque los rasgos asociados a la raíz son promisorios para la selección y desarrollo de resistencia a la sequía, es de suma dificultad seleccionar en base a mediciones fenotípicas de la raíz. La clave sería dilucidar los mecanismos fisiológicos y las regulaciones genéticas que gobiernan la plasticidad radicular y su adaptación a la sequía. La identificación de los genes y las rutas metabólicas encargadas de este aspecto ayudarían a la realización de selección asistida por marcadores moleculares o edición génica.

- Mantenimiento de la fijación biológica de nitrógeno La fijación biológica de nitrógeno (FBN) en leguminosas es sumamente sensible a la disponibilidad hídrica. El estrés hídrico afecta la FBN independientemente de su efecto sobre la tasa fotosintética, se ha encontrado que con disminuciones de tan solo un 5 % se produce una baja del 70 % en la FBN (Manavalan y col., 2009). Esta inhibición en la fijación frente a la sequía está asociada a una menor disponibilidad de oxígeno, un menor flujo de carbono hacia los nódulos, una disminuida síntesis de sacarosa en los nódulos y un aumento de ureídos y aminoácidos libres. Se ha demostrado que aquellos genotipos capaces de mantener una alta turgencia en sus tejidos ante situaciones de sequía (conocidos como de marchitamiento lento o *slow wilting*), y por consiguiente, sostener la función de sus hojas y nódulos, son aquellos que menos reducen su rendimiento ante el estrés (Prince y col., 2015; Sinclair y col., 2010).
- **Conductancia estomática** A medida que el reservorio de agua en el suelo va disminuyendo la planta atraviesa distintas etapas. Una primera en donde el agua aún no es limitante y la transpiración no depende del contenido de agua en el suelo, sino de la demanda atmosférica; y luego, una segunda etapa en donde la transpiración está limitada por la disponibilidad hídrica. Durante esta transición entre fases suele haber una disminución en la conductancia estomática (g_s) que influye en la cantidad de agua que pierde la planta ante el desbalance entre la demanda y la disponibilidad hídrica. El contenido hídrico al cuál ocurre esta transición varía entre especies y genotipos (Hufstetler y col., 2007; Manavalan y col., 2009). En un trabajo realizado por Ray y Sinclair 1997 con distintos genotipos

de maíz se encontraron variaciones significativas en el contenido de agua del suelo (expresado como FTSW, del inglés *Fraction of Transpirable Soil Water*) al cual el uso del agua por parte de la planta comenzaba a declinar. En soja, se ha demostrado que genotipos que exhiben tolerancia a la sequía tienen una mayor y más rápida disminución de la g_s , en comparación con los susceptibles (Zhao, Aleem y Sharmin, 2018). Esta rápida disminución de la pérdida de agua por parte de la planta permite preservar el agua del suelo y usarla de manera más eficiente. Este mecanismo puede evidenciarse macroscópicamente como un marchitamiento lento, y se ha descripto como un atributo de tolerancia a la sequía que no compromete el rendimiento (Sinclair y col., 2010; Kunert y Vorster, 2020).

- Conductancia epidérmica La conductancia epidérmica de la hoja es la resultante de la combinación entre la conductancia estomática y cuticular. Cuando los estomas están abiertos, el agua que se pierde a través de la cutícula es despreciable. Sin embargo, en condiciones intensas de sequía cuando los estomas están completamente cerrados, el agua que se pierde a través de la cutícula cobra importancia, pudiendo ser su conductancia incluso mayor que la estomática. Es decir, cuando los estomas están cerrados, la conductancia epidérmica termina determinando la pérdida de agua de la hoja que puede llegar a niveles críticos dañándolas. Seleccionar a favor de una baja conductancia epidérmica podría ayudar a generar genotipos que puedan aumentar su supervivencia ante períodos de sequía. Sin embargo, en estudios anteriores se ha encontrado una correlación negativa entre la conductancia epidérmica y la WUE (Hufstetler y col., 2007), uno de los principales rasgos que se buscan potenciar.
- Pubescencia foliar La pubescencia foliar aumenta la reflectancia de la hoja, logrando mantener menores temperaturas ante altas irradiancias. Además, una mayor pubescencia aumenta la resistencia de la capa límite hasta un 50%. Al tener una menor temperatura y mayor resistencia de la capa límite, la planta tiene una menor transpiración y una menor saturación lumínica al penetrar una menor radiación al canopeo. Sin embargo, no es recomendable la selección a favor de este carácter ya que está gobernado por los alelos *Pd1* y *Pd2* que han sido asociados con reducciones en el rendimiento, retrasos en la maduración y aumentos en la altura de la planta (Miransari, 2016).
- **Eficiencia del uso del agua** La WUE puede ser considerada a distintos niveles, tanto de una hoja en particular como de la planta entera. En el primer caso, la WUE esta definida por el ratio entre la fotosíntesis a saturación lumínica ($A_{máx}$) y la conductancia estomática (g_s) para el vapor de agua. En este enfoque, la WUE de la hoja podría maximizarse mediante un cierre parcial de estomas que garantice una concentración de carbono intercelular suficiente para saturar la fotosíntesis, pero al mismo tiempo minimizar la pérdida de agua a través de los estomas. Si consideramos la WUE de la planta entera esta está definida por la relación entre la acumulación de materia seca y la transpiración. En soja se ha demostrado que ante déficits hídricos moderados aumenta la WUE a ambos niveles, siendo un rasgo manipulable, y por ende, objetivo de estudio del mejoramiento (Liu y col., 2005). Sin embargo, debe tenerse en cuenta que este atributo es desable siempre y cuando no traiga aparejado pérdidas en el rendimiento; por lo que podría buscarse su incremento a través de una mayor producción de biomasa y no una reducción en el uso del agua (Fried, Narayanan y Fallen, 2019).
- **Ajuste osmótico** El ajuste osmótico es la activa acumulación de solutos en respuesta a una condición de agua limitante, que permite mantener una alta turgencia celular aún cuando el potencial agua de los tejidos está en detrimento. Los solutos acumulados pueden ser prolina, sacarosa, carbohidratos solubles, glicina beatina, entre otros (Zhao, Aleem y Sharmin, 2018; Buchanan, Gruissem y Jones, 2015). Este mecanismo permite mantener la g_s y la tasa fotosintética, minimizar la senescencia y muerte foliar, reducir el aborto floral, potenciar el crecimiento radicular y la extracción de agua del suelo.
- Acumulación de proteínas Existen distintos tipos de proteínas que protegen a las células contra el estrés osmótico e hídrico. Estas actúan como proteínas altamente hidrofílicas que retienen agua o como chaperonas que protegen a macromoléculas y membranas de la desnaturalización. Estas proteínas se acumulan activamente durante procesos de estrés abiótico, indicando así su importancia en la defensa de la célula. Las LEA (late embryogenesis abundant proteins), son un grupo de proteínas que, en un principio, se pensaba que solamente se producían durante la embriogénesis; sin embargo, se ha descubierto que se producen también en respuesta al frío, estrés osmótico, y ABA exógeno. Su función no es conocida con precisión, pero dada su activa acumulación bajo condiciones de estrés, se cree que las mismas están involucradas en la protección y defensa de la célula confiriendo tolerancia (Aroca, 2012; Cutler y col., 2010). Las condiciones de sequía también inducen la acumulación de HSPs (*heat shock proteins*), que cumplen la función de chaperonas moleculares. Estas cumplen un rol importante

en el plegado, ensamblaje y estabilización de otras proteínas, estabilización de membranas, replegado de proteínas defectuosas y removilización de proteínas no funcionales o degradadas bajo condiciones de estrés (Joshi y col., 2016; Sanità de Toppi y Pawlik-skowrońska, 2003; Lopes-Caitar y col., 2013).

ABA - El ABA es una fitohormona que, además de estar involucrada en la germinación y la maduración, juega un rol importante en la respuesta a la sequía (Cutler y col., 2010). Distintos tipos de estreses abióticos, como la salinidad y la sequía, inducen la producción de ABA en las raíces, que luego es transportado a las hojas induciendo el cierre estomático. Esta disminución en la g_s produce consecuentemente una disminución en la fotosíntesis (Chaves, Flexas y Pinheiro, 2009). La producción de ABA también cumple un rol esencial en la acumulación de metabolitos y proteínas que ayudan en la resistencia a la sequía. Por ejemplo, actúa promoviendo la degradación de almidón transitorio en respuesta al estrés osmótico a través de la actividad de la α -amilasa y β -amilasa, produciendo hexosas que son solutos compatibles acumulados activamente para mantener la turgencia y mantener la inegridad de la membrana celular (Cardoso y col., 2020). Además, su acumulación induce la expresión de distintos genes involucrados en la respuesta y tolerancia a la deshidratación.

3.1. Mejoramiento genético para tolerancia a sequía

El mejoramiento, o *breeding*, para la obtención de plantas de soja tolerantes a sequía es generalmente dejado de lado en los planes de mejoramiento de las grandes empresas debido a su alta complejidad. El fenotipado a gran escala de rasgos asociados a la tolerancia a la sequía, o una elevada WUE, suele ser un cuello de botella (Jumrani y Bhatia, 2019). Esto se debe a que la sequía tiene distintos efectos dependiendo de la duración, la severidad y el momento del ciclo en el cuál ocurra. Además, no existe una única cualidad que otorgue tolerancia ante este estrés, sino que hay varios atributos que deben considerarse siempre teniendo en cuenta el ambiente objetivo.

La tolerancia a la sequía está gobernada por múltiples genes que interaccionan entre sí, por lo que un enfoque en donde se contemplen numerosos genes en conjunto podría tener ventajas por sobre aquellos en donde se seleccionan genotipos a partir de unos pocos genes (Tripathi y col., 2016; Alam Khan, 2018). En Argentina se ha desarrollado el evento transgénico HB4, un factor de transcripción de girasol que confiere tolerancia a la sequía, y se ha aprobado su inserción en trigo y soja (Gabriela González y col., 2019; Ribichich y col., 2020). Por más que su utilización ha sido aprobada por el Ministerio de Agricultura, hoy en día su participación en el mercado sigue siendo muy baja. Su bajo *market share* se debe, según la misma obtentora del evento, principalmente a dos causas: la mala percepción pública de los transgénicos y la incapacidad de un solo gen o genotipo de brindar tolerancia a estreses abióticos complejos. La sequía, salinidad o inundación no son estreses üniversales ´´, es decir no afectan siempre en el mismo momento, con la misma intensidad o duración; por lo que se requieren tantos *traits* y genotipos como escenarios posibles. Por más que hay áreas del país en donde el evento HB4 podría representar una ventaja en años secos, la tecnología representa un costo en años en donde no hay sequía, por lo que el productor es reticente a incorporarla (Chan, Trucco y Otegui, 2020).

En este sentido, los marcadores moleculares funcionarían como una herramienta clave, ya que permiten el análisis de múltiples genes en simultáneo sobre el germoplasma propio de una semillera. Esta tecnología le permitiría al *breeder* identificar los individuos más tolerantes dentro de su programa de mejoramiento, en donde la selección ya es direccionada hacia un ambiente objetivo particular. Además, los marcadores moleculares pueden contribuir a la identificación de loci asociados al estrés hídrico mediante el uso de técnicas cuantitativas para la identificación de QTLs (del inglés, *Quantitative Trait Loci*). Una vez identificar los genes causales asociados a la tolerancia al estrés hídrico. Esta estrategia es costosa y de alto riesgo por lo que otra alternativa posible son las aproximaciones genómicas y transcriptómicas para la identificación de genes marcadores al estrés hídrico.

Uno de los principales desafíos para el mejoramiento reside en encontrar cuáles son los genes causantes de la tolerancia. Un primer paso hacia la resolución de este interrogante es el estudio del transcriptoma de genotipos contrastantes (Prince y col., 2015; Hua y col., 2018). Este enfoque permite conocer en detalle los genes, y las rutas metabólicas asociadas, que responden frente al estrés. Asimismo, la búsqueda de polimorfismos en la secuencia transcriptómica para el desarrollo de marcadores moleculares tendría como

ventaja la utilización únicamente de las regiones exónicas, que tienen mayor impacto que las regiones no codificantes. Además, se podrían priorizar aquellos polimorfismos presentes en genes claves involucrados en la respuesta, conocidos como eQTLs (del inglés, *expression QTLs*) maximizando el impacto (Bernardo y Woodbury, 2020). El análisis de numerosos marcadores moleculares en simultáneo hoy en día es posible y sencillo gracias a las plataformas de genotipado como la SNP Line (LGC Biosearch Technologies), que permiten el análisis de un gran número de muestras por día. En síntesis, la selección de genotipos tolerantes a la sequía a través del análisis de un panel de marcadores moleculares parecería ser la alternativa más apropiada y económica frente a aquellas que se basan en la manipulación de uno o pocos genes para aplicar en planes de mejoramiento.

4. Objetivos e Hipótesis

4.1. Objetivo general

Identificar los genes clave que participan en la tolerancia a la sequía en el cultivo de soja y analizar sus polimorfismos con vistas al diseño de marcadores moleculares para asistir al mejoramiento.

4.2. Objetivos específicos

- Caracterizar la respuesta a la sequía de genotipos de soja con comportamientos contrastantes ante este estrés.
- Identificar los genes, y rutas metabólicas asociadas, con expresión diferencial ante el estrés por sequía en los genotipos contrastantes.
- Analizar el polimorfismo de los principales genes de respuesta a la sequía y utilizarlo en el diseño de marcadores moleculares.

4.3. Hipótesis

- El genotipo tolerante se diferenciará del sensible en diversos parámetros fisiológicos y morfológicos que le permitirán utilizar el agua de una manera más eficiente minimizando el estrés provocado por la sequía.
- II. El genotipo tolerante presentará enriquecidas rutas metabólicas propias de la tolerancia al estrés.
- Genes presentes en estas rutas metabólicas presentarán un polimorfismo estructural que permitirá su utilización como potenciales marcadores moleculares.

Capítulo 2

Materiales y Métodos Generales

1. Elección de los genotipos y material botánico

En mi trabajo final de grado he identificado dos genotipos con comportamientos contrastantes a la sequía (GT, genotipo tolerante; GS, genotipo susceptible). Estos genotipos se seleccionaron a partir de la evaluación del germoplasma de una semillera nacional realizada anteriormente por el grupo de trabajo, en donde se sometieron 630 genotipos a estrés por sequía y se evaluó su respuesta. Los genotipos seleccionados pertenecen al grupo de madurez IV y difirieron en su respuesta ante dicho estrés a través de diferencias en características morfológicas (peso y longitud de tallo y raíces), fenológicas (días a VC y VE de la escala de Fehr y Caviness) y a nivel molecular (distintas expresiones relativas de tiorredoxinas, (Demicheli, 2017). Sin embargo, surge como interrogante si estos genotipos pueden ser catalogados y tomados como ejemplos de "tolerancia y susceptibilidad a la sequía". Es decir, el hecho de que sean contrastantes no los hace comparables ante genotipos de conocido comportamiento ante la sequía, como el PI 416937 o Williams 82.

Es por esto por lo que, además de los genotipos ya estudiados (el genotipo susceptible y el genotipo tolerante) siempre que se pudo se incluyeron también dentro del análisis los genotipos PI 416937 y Williams 82. El genotipo PI 416937 es *slow wilting* gracias a un sistema que le permite conservar el agua en el suelo (Ries y col., 2012; Song y col., 2016; Manavalan y col., 2009; Earl, 2002; Mian y col., 1996; King, Purcell y Brye, 2009; Shin y col., 2015). Es decir, tiene mecanismos que le permiten conservar el agua en el suelo incluso cuando abunda, dejándola disponible para momentos del ciclo en la que pueda faltar. Debido a esto, el contenido hídrico relativo crítico (RSWC_C) de este genotipo (contenido de agua en el suelo al cual el genotipo disminuye su transpiración) suele ser más alto que el de otros genotipos (Hufstetler y col., 2007). Además, ante una alta demanda atmosférica, es decir, a valores de VPD superiores a 2 kPa mantiene su transpiración, mientras que los demás genotipos siguen respondiendo a la demanda. Esta limitación en la transpiración se debe a una menor conductancia hidráulica en las hojas (Sinclair, Zwieniecki y Holbrook, 2008; Sinclair, 2017). El flujo de agua desde el xilema de la hoja hasta las células oclusivas es interrumpido por una baja conductancia hidráulica ante una situación de alto VPD, por lo que estas pierden turgencia aumentando la resistencia estomática. Esto permite que ante condiciones de VPD > 2kPa (situación común a campo al mediodía) este genotipo conserve el agua del suelo. Este rasgo puede ser útil en el breeding de genotipos especialmente pensados para ambientes con alta demanda atmosférica y bajas precipitaciones. Por el contrario, el genotipo Williams 82 es considerado sensible a la sequía y es utilizado como genotipo de referencia en diferentes estudios dado que se encuentra secuenciado en su totalidad en las distintas bases de datos (Ha y col., 2015; Hua y col., 2018; Song y col., 2016).

En todos los experimentos se trabajó con las hojas de las plantas tratadas. Se eligieron estas por sobre las raíces ya que son los órganos en los cuáles se va dar una gran parte de la respuesta, como el cierre estomático, la respuesta a fitohormonas, la fotosíntesis y el comienzo de la senescencia. Además, son mucho más fáciles de manipular y seleccionar que las raíces. Al trabajar con un sustrato arenoso el lavado de las raíces resulta muy dificultoso ya que muchas partículas quedan adheridas y resultan en un problema en las posteriores aplicaciones biomoleculares. Uno de los problemas más comunes es la baja pureza y



Figura 2.1: Imágenes del sustrato utilizado en los experimentos

concentración del RNA extraído por el tapado de las columnas de extracción.

2. Diseño experimental y tratamientos

En todos los experimentos se utilizó un diseño completamente aleatorizado (DCA) multifactorial, siendo los dos factores genotipo y condición hídrica. Los niveles para el factor condición hídrica fueron siempre dos: control y sequía; pero para el factor genotipo variaron según el experimento y se detallarán en cada capítulo. Las repeticiones y las mediciones específicas realizadas en cada experimento también se detallarán en cada capítulo.

Generalmente, los genotipos fueron sembrados por quintuplicado en macetas de 3L (17 cm de alto x 15 de diámetro) durante el verano 2018/19 en un invernáculo de la Facultad de Agronomía, UBA. El sustrato utilizado fue una mezcla de growmix (sustrato profesional marca terrafertil, 33 %) y arena entrefina (marca ciudad floral, 67 %). Ambos sustratos fueron vertidos proporcionalmente en un contenedor y mezclados hasta homogeneizar. Una vez listo el sustrato se rellenaron las macetas y llevaron a saturación permitiéndo-les drenar durante la noche. Al día siguiente, una vez que habían perdido el exceso de agua se sembraron 3 semillas por maceta. Previo a la siembra se calculó la capacidad de retención hídrica, resultado de la diferencia entre el peso del sustrato a capacidad de campo y el peso seco. El peso a capacidad de campo fue el registrado a la mañana siguiente de que se dejaron drenar *overnight*, mientras que el peso seco se registró luego de secar el sustrato hasta constancia de peso en estufa a 80 °C.

Una vez alcanzado el estado de VE se raleó y se dejó una sola planta por maceta. Hasta el estado de V2 las plantas se mantuvieron bien regadas y se controlaron las plagas mediante la aplicación de un insecticida/ acaricida (Grhesatión 25 ME). El día previo al comienzo del tratamiento de sequía se regaron las macetas hasta saturación y se dejaron drenar durante la noche.



Figura 2.2: Macetas sembradas en el invernáculo de la Facultad de Agronomía en los distintos experimentos.

Capítulo 3

Comparación de los genotipos con respuestas contrastantes a la sequía

1. Introducción

La obtención de plantas tolerantes a la sequía suele ser dejada de lado en planes de mejoramiento dada la complejidad de este rasgo (conocido en inglés como *trait*), el fenotipado a gran escala suele ser una limitante (Jumrani y Bhatia, 2019). Es por esto que la obtención de marcadores moleculares que permitan detectar tempranamente genotipos con una performance superior, o al menos descartar aquellos con excesiva susceptibilidad, sería una gran herramienta para la selección de lineas tolerantes. La comparación de genotipos contrastantes es un punto de partida que aporta información valiosa sobre los posibles mecanismos detrás de la tolerancia. Una exhaustiva caracterización de los genotipos es la base para luego poder rastrear los genes responsables y evitar caer en el camino inverso; en donde se salta del gen de interés a la planta entera sin considerar las posibles interacciones y artefactos (Passioura, 2020).

La respuesta a la sequía es compleja y abarca distintos mecanismos que pueden involucrar desde la raíz hasta las hojas de la planta. Por eso, a la hora de evaluar el comportamiento de un genotipo ante este estrés es importante medir más de un parámetro, ya que muchos en conjunto pueden ser los causantes de la respuesta observada. El marchitamiento, o *wilting*, se da cuando la planta comienza a perder turgencia y es el primer síntoma visible del estrés hídrico. Aquellas plantas que presentan un fenotipo de marchitamiento lento, conocido como *slow wilting*, lo hacen a través de conservar el agua del suelo al disminuir la pérdida de agua a través de la transpiración, dando como resultado una mayor eficiencia en el uso del agua (WUE), turgencia y retención hídrica (Kunert y Vorster, 2020); o a través de una mayor capacidad de captación de agua. Esta respuesta puede lograrse a través de varios mecanismos, como puede ser un mayor volumen de suelo explorado a causa de un sistema radical más profundo o denso, menor conductancia estomática (g_s), acumulación de minerales o solutos para reducir el potencial osmótico, o el mantenimiento de la transpiración por sobre determinado DPV (déficit de presión de vapor). Además, estos mecanismos parecerían aumentar la tolerancia a la sequía sin comprometer rendimiento (Pathan y col., 2014; Sinclair y col., 2010; Kunert y Vorster, 2020). Dado que el fenotipo de marchitamiento lento es el primer síntoma visible del estrés, y trae aparejados todos estos mecanismos detrás, es un rasgo valuable y fácil de medir.

Como se dijo anteriormente, el fenotipo de marchitamiento lento puede ser producto de una menor pérdida de agua. Para evaluar si esto es así, es importante estudiar la dinámica de transpiración ante contenidos hídricos restrictivos. Una manera muy utilizada es la propuesta por Ray y Sinclair (Ray y Sinclair, 1997), que permite identificar el contenido hídrico a partir del cual un genotipo en particular cierra los estomas para conservar el agua del suelo. De esta manera, se puede comparar este valor (llamado RSWC_c) entre genotipos e identificar cuál tiene una mayor sensibilidad estomática. Esta metodología fue utilizada en diversos trabajos a lo largo del tiempo en soja, girasol, maní y maíz, entre otros (Hufstetler y col., 2007; Earl, 2003; Casadebaig, Debaeke y Lecoeur, 2008; Bhatnagar-Mathur y col., 2009; Sinclair y col., 2010; King y Purcell, 2017).

En este capítulo se abordará el primer objetivo específico planteado en el Capítulo 1 poniendo a prueba la primera hipótesis, que establece que el genotipo tolerante se diferenciará del sensible en diversos parámetros fisiológicos y morfológicos que le permitirán utilizar el agua de una manera más eficiente minimizando el estrés provocado por la sequía. Para comprobarlo, se evaluó la respuesta de cada genotipo frente al estrés a través de distintos parámetros (marchitamiento, transpiración, fotosíntesis y características estomáticas).

2. Materiales y Métodos

2.1. Condiciones de crecimiento y mediciones

2.1.1. Wilting

Las condiciones de crecimiento hasta el inicio del tratamiento fueron las detalladas en el Capítulo 2. Pasados 26 días después de la siembra (específicamente, el 6/11/18) se alcanzó el estado de V2 y se comenzó con el tratamiento que consistió en cortar el riego a las plantas destinadas a sequía. Por el contrario, las plantas control se mantuvieron bien regadas manteniendo siempre una disponibilidad hídrica de entre un 90 y 80 % de capacidad de campo (CC). El contenido hídrico relativo (RSWC, del inglés *relative soil water content*) de cada maceta se registró diariamente entre las 15 y las 17 hs.

El wilting diario de cada planta en sequía se clasificó según la escala propuesta por King (King, Purcell y Brye, 2009) que consiste en 6 valores entre 0 y 100 (Cuadro 3.1). Se asignó un valor de la escala a través de la observación individual de cada planta diariamente entre las 15 y 17 hs.

Valor	Descripción
0	Plantas sin marchitamiento.
20	Plantas con un leve marchitamiento y plegamiento foliar en la cima del canopeo.
40	Plantas con severo plegamiento foliar en la cima del canopeo, moderado marchitamiento de hojas a lo largo del canopeo y cierta pérdida de turgencia en los pecíolos.
60	Plantas con marchitamiento severo en hojas a lo largo del canopeo y avanzada perdida de turgencia en pecíolos.
80	Plantas con marchitamiento severo en pecíolos y hojas muertas a lo largo del canopeo.
100	Muerte de la planta.

Cuadro 3.1: Escala de marchitamiento utilizada.

2.1.2. Transpiración

Hasta el inicio del tratamiento las plantas fueron crecidas según las condiciones detalladas en el segundo capítulo (véase Sección 2, Capítulo 2). Una vez alcanzado el estado de V1 de la escala de Fehr y Caviness (Fehr y Caviness, 1977) las macetas fueron regadas hasta saturar y se dejaron drenar durante la noche. Para contabilizar la transpiración de cada planta se cubrió cada maceta con una bolsa plástica y se insertó entre el tallo y la bolsa un tubo para facilitar el riego (Figura 3.2). De esta manera se logró controlar la evaporación, siendo la diferencia de peso obtenida entre días sucesivos producto únicamente de lo transpirado por la planta. Las macetas se pesaron diariamente alrededor de las 15 hs.

Luego del embolsado se registró el peso de cada maceta y se calculó el peso inicial. Este peso inicial, o peso a CC, estaba compuesto por: el peso del sustrato a CC, el peso de la maceta, el peso de la bolsa, el peso del tubo plástico insertado para el riego y el peso promedio de 60 plantas de los mismos genotipos en V4 (de esta manera se buscó tener en cuenta el peso de la planta a lo largo del experimento). A partir de este momento, las plantas destinadas a control se mantuvieron entre un 80 y 90% de CC para lograr una buena



Figura 3.1: Escala de marchitamiento utilizada





Figura 3.2: Macetas embolsadas y el registro de su peso diario.

disponibilidad hídrica, pero evitando condiciones de anaerobiosis. Asimismo, se cortó el suministro hídrico a las plantas destinadas a sequía.

Para minimizar las fluctuaciones en la transpiración a lo largo de los días debido a factores ambientales, la transpiración diaria de las plantas estresadas se relativizó a la transpiración diaria de las plantas bien regadas. La relativización consistió en dividir la transpiración diaria individual de cada planta estresada por el promedio de la transpiración diaria de las plantas bien regadas del mismo genotipo, obteniéndose así el índice de transpiración diario (DTR, del inglés *Daily Transpiration Ratio*). Luego, se realizó una segunda normalización para mitigar las variaciones entre plantas. Para esto, se relativizó el DTR diario de cada planta estresada al promedio del DTR de esa misma planta al inicio del ciclo, cuando aún tenía una buena disponibilidad hídrica. Para lograr esto, se dividió el DTR diario de cada planta estresada por el promedio del DTR de esa misma planta para los primeros 4 días del experimento, previo a que las condiciones hídricas fueran limitantes. Es decir, antes de que comiencen las diferencias entre las plantas bien regadas y las estresadas. A través de la segunda normalización se construye el índice de transpiración normalizado (NTR, del inglés *Normalized Transpiration Ratio*; Ray y Sinclair, 1997). Las plantas estresadas fueron descartadas una vez que su NTR alcanzó un valor de 0,15; este punto se tomó como el peso final de la maceta.

2.1.3. Fotosíntesis y parámetros asociados

Las condiciones de crecimiento generales fueron las detalladas en el Capítulo 2. En este experimento no se utilizaron los genotipos de referencia ya que en base a los resultados obtenidos en las mediciones de marchitamiento y transpiración se comprobó que la respuesta de los genotipos contrastantes era comparable a la de estos de referencia. Al inicio del experimento se llenaron las macetas con 3,2 kg de sustrato y se sembraron cuatro repeticiones biológicas de los genotipos de interés (GT y GS) usando tres semillas por maceta para asegurar una buena emergencia. Las plantas se mantuvieron bien regadas, en el estadío de VC se raleó a una planta por maceta, se aplicó fertilizante y un insecticida para el control de plagas. Una vez alcanzado el estadío V2 se comenzó con el riego diferencial. En primer lugar, se regó a saturación y se dejó drenar durante la noche. A la mañana siguiente se registró el peso a capacidad de campo, y a partir de ese momento, se cortó el suministro hídrico a aquellas plantas destinadas a sequía mientras que las demás, a modo de control, se mantuvieron entre un 80 y 90 % de CC. Se registró diariamente el peso de cada maceta para la determinación de su contenido hídrico.

Pasados 3 y 10 días desde el último riego se realizaron mediciones de fotosíntesis con el equipo LiCor-6400. Las mediciones se realizaron al mediodía (de 11 a 13:30 hs) cuando las condiciones de irradiancia eran máximas. Se midió sobre el folíolo central de la segunda hoja trifoliada expandida. Los bordes de la hoja inmediatamente superior, por lo general, aún se tocaban. Se realizaron mediciones de la tasa fotosintética y tanspiratoria, conductancia estomática (g_s) y eficiencia del uso del agua (WUE) instantánea, estimada como la relación entre la asimilación de carbono y la transpiración. Finalizadas las mediciones se cosecharon las hojas sobre las cuales se había trabajado y se congelaron en N líquido inmediatamente. Posteriormente, se guardaron a -80°C hasta su utilización.

2.1.4. Microscopía

Este experimento se realizó en la Estación Experimental del Zaidín ubicada en Granada, España. Se utilizaron los genotipos de interés y el control Williams 82, dejando de lado al PI 416937 en base a los resultados obtenidos para las mediciones de marchitamiento y tanspiración. Los tres genotipos fueron germinados en papel de filtro humedecido por 5 días y luego se transfirieron a macetas que se mantuvieron bien irrigadas durante hasta la toma de las muestras. Las macetas se introdujeron en una cámara de crecimiento con un fotoperíodo largo (16 horas de luz/ 8 horas de oscuridad), con una temperatura de 21°C, 60 % de humedad relativa y 500 mumol fotones m⁻² s⁻¹ de intesidad lumínica. Estas condiciones se mantuvieron hasta que las plantas alcanzaron el estadío de V1, momento en el cual se tomaron las muestras para su análisis bajo microscopio.

Para la estimación de las variables estomáticas y epidérmicas se tomaron 10 muestras de cada genotipo, 5 de
la cara adaxial y 5 de la cara abaxial. Las imágenes se tomaron con un microscopio electrónico de barrido de alta resolución (HRSEM AURIGA, Carl Zeiss SMT) y se procesaron con el software ImageJ (Schneider, Rasband y Eliceiri, 2012). Para el cálculo del índice de apertura estomático se midieron 6 estomas dentro de cada imagen.

2.2. Metodología Estadística

2.2.1. Marchitamiento

Para los datos que corresponden a los días que tardó cada genotipo en alcanzar un 100 en la escala de marchitamiento (Figura 3.3) se realizó un ANOVA y las comparaciones múltiples se probaron con el test de Tukey utilizando el software R (R Core Team, 2020). Previamente se comprobó que los datos cumplieran con los supuestos de distribución normal y homogeneidad de varianzas. Dos de las cinco repeticiones del genotipo Williams fueron desestimadas dado que hubo un error en la asignación de las semillas.

2.2.2. Transpiración

Para evaluar si las diferencias encontradas en la transpiración acumulada al día 24 del tratamiento se construyó un modelo lineal multifactorial. La transpiración acumulada se tomó como la variable respuesta, y el genotipo (GS, GT, PI 416937, Williams 82) y la condición hídrica (control, sequía) se tomaron como las variables predictoras. Se comprobaron los supuestos de normalidad y homogeneidad de varianzas.

Para el análisis del NTR se ajustó a los datos una regresión lineal segmentada a través de los mínimos cuadrados utilizando el software GraphPad (*GraphPad Prism Software*). Una vez ajustado el modelo se compararon los valores estimados para los parámetros X0 y la pendiente de los distintos sets de datos. Esta comparación se realizó para identificar si cada parámetro difería entre los sets de datos o si un valor común ajustaba mejor a todos los datos en conjunto. La comparación se realizó haciendo una suma de cuadrados extra y una prueba con el estadístico F, seleccionando como diferencias significativas aquellas con un valor p<0,05.

2.2.3. Fotosíntesis y parámetros asociados

Los resultados fueron analizados a través del software GraphPad (*GraphPad Prism Software*). Para las mediciones de fotosíntesis se ajustó un modelo hiperbólico y se realizó una suma de cuadrados extra con una prueba F para comprobar la conveniencia de ajustar un modelo distinto para cada set de datos. Para las mediciones de conductancia, transpiración, WUE y C_i se ajustaron modelos lineales multifactoriales, siendo los factores tratamiento e irradiancia, y se analizaron a través de análisis de la varianza con medidas repetidas en el tiempo. Las comparaciones se realizaron utilizando el test de Tukey corregido por comparaciones múltiples a través de la prueba estadística de hipótesis.

2.2.4. Microscopía

Para el análisis estadístico se ajustó un modelo multifactorial en donde las distintas variables respuesta dependían de los factores genotipo (GT, GS y WIlliams), cara de la hoja (adaxial o abaxial) y la interacción entre ambos. Luego de la comprobación de los supuestos de los distintos modelos se realizó un ANOVA y en caso de ser significativa la interacción entre factores, o el factor con más de dos niveles (genotipo) se realizó una prueba de comparaciones múltiples (HSD Tukey) con el paquete agricolae de R (Felipe de Mendiburu y Muhammad Yaseen, 2020).

3. Resultados

3.1. Marchitamiento

En primer lugar, se contabilizó la cantidad de días sin riego que tardaron los distintos genotipos en alcanzar un 100 en la escala de marchitamiento, y se pudo observar que estos respondieron de manera distinta (Figura 3.3). Luego de 15 días sin agua las plantas de los genotipos PI 416937, GS y Williams 82 habían muerto en su totalidad, presentando hojas, pecíolos y tallos completamente secos (Figura 3.4). Sin embargo, para ese mismo día el genotipo GT aún no presentaba síntomas de marchitamiento.



Figura 3.3: Días que demoró cada genotipo en llegar a un 100 en la escala de marchitamiento. Las barras corresponden al error estándar, letras distintas indican una diferencia con un p<0,05. (GS: Genotipo Susceptible, GT: Genotipo Tolerante, PI: PI 416937, W: Williams 82)

Recién al cabo de 18 días sin riego el GT comenzó a evidenciar síntomas de marchitamiento con una leve pérdida de turgencia y plegamiento foliar. Al día siguiente, pasados 19 días desde el último riego, el 40 % sus plantas se vieron severamente afectadas (Figura 3.5). El 60 % restante de las plantas del GT continuaron con moderados síntomas de marchitamiento y murieron dentro de la semana siguiente.

En segundo lugar, se relacionó el marchitamiento con el contenido de agua relativo para los cuatro genotipos. En la figura 3.6 cada panel muestra la relación para un genotipo distinto y el color de los puntos indica la severidad del marchitamiento; gris para plantas turgentes, amarillo para plantas con severo plegamiento foliar en la cima del canopeo, moderado marchitamiento de hojas a lo largo del canopeo y cierta pérdida de turgencia en los pecíolos; y por último, rojo para plantas completamente secas. Se graficaron, además, tres líneas verticales a contenidos hídricos de: 0,25; 0,15 y 0,05. Es interesante notar como, para RSWC entre 0,25 y 0,15, la mayoría de los genotipos ya presentaban varias plantas con síntomas, pero el GT solo una. Dos de las plantas del genotipo PI 416937 transitaron toda la escala de marchitamiento entre estos dos contenidos. Para contenidos hídricos aún menores (cercanos a 0,1) las plantas correspondientes a todos los genotipos ya habían alcanzado un 100 en la escala, exceptuando a la mitad de las plantas correspondientes al GT.

Aunque a un mismo contenido relativo de agua los genotipos de interés, GT y GS, obtuvieron valores de marchitamiento similares se pudo observar que el GT siempre demoró más días en llegar a ese contenido relativo de agua. Es decir, el GT pudo conservar de manera más eficiente el agua del suelo y se distinguió de los demás genotipos de manera notable (Figura 3.7). La figura 3.7 muestra en el panel izquierdo la disminución progresiva del RSWC diario para todas las repeticiones de los cuatro genotipos, el color de los puntos representados varía según el nivel hídrico, siendo azul para valores elevados (buena disponibilidad hídrica), verde para valores intermedios (contenidos hídricos que comienzan a ser limitantes) y rojo para valores bajos (a partir de los cuales las planas ya comienzan a evidenciar pérdidas de turgencia notables). Se puede observar como el GT, a un mismo día, siempre presentó valores mayores para el RSWC. Esta diferencia se ve marcadamente en el panel derecho en donde se graficó el promedio para cada uno de los



Figura 3.4: Estado de las plantas de los cuatro genotipos en condición de sequía a 15 días del último riego



Figura 3.5: Estado de las plantas del GT bien regadas (izquierda) y sin riego (derecha) por 18 y 19 días.



Figura 3.6: Marchitamiento en función del contenido hídrico relativo (RSWC) para los cuatro genotipos. El color de los puntos indica la severidad del marchitamiento, siendo gris para valores cercanos a 0, amarillo para valores intermedios (50) y rojo para plantas muertas (100).

cuatro genotipos (puntos), se ajustó un modelo lineal (línea) y se calcularon los intervalos de confianza (área sombreada alrededor de cada línea). En este panel se puede ver claramente como el GT se diferenció de los demás genotipos, presentando siempre contenidos hídricos relativos superiores.

3.2. Transpiración

La primera exploración de los datos que corresponden a la transpiración acumulada (TA) para los cuatro genotipos en las distintas condiciones hídricas se muestran en la figura 3.8, en donde se representó la TA medida en litros (L) para cada una de las seis repeticiones de cada tratamiento. Todos los genotipos comenzaron el experimento con seis repeticiones biológicas, pero se puede observar en el gráfico como las muestras 4 y 24; ambas pertenecientes al GS, murieron a los pocos días de comenzar. La variabilidad entre las repeticiones biológicas fue máxima para el genotipo PI 416937 en situación control donde la muestra 21 duplicó el volumen de agua transpirado por las demás plantas, no solo del mismo genotipo, sino de las otras 47 plantas presentes en el experimento. Se decidió eliminar esta muestra dado que no guardó relación con la transpiración de las demás muestras del grupo biológico, ni con ninguna otra muestra dentro del experimento. Además, se decidió eliminar la muestra número 36 del genotipo Williams 82 dado que tenía manchas foliares y murió pasadas dos semanas. Durante ese tiempo, esta muestra transpiró marcadamente más que todas las demás plantas de este genotipo debido a la infección fúngica. Todos los análisis y gráficos sobre transpiración mostrados a continuación exceptúan la muestra 21 y la 36.

El RSWC se presenta como la contrapartida de la transpiración dado que, al estar las macetas embolsadas, toda el agua disponible debe haber sido absorbido por las raíces, movilizado a través del xilema y transpirado por las hojas. Se puede observar que estos resultados guardaron relación con la TA (Figura 3.9). El GS tuvo una amplia variabilidad habiendo una separación en dos grupos distintos conformados por las muestras 34, 32 y 22 por un lado; y las muestras 28 y 25 por otro. A diferencia del segundo, el primer grupo transpiró más y consecuentemente, alcanzó valores de RSWC más bajos. Esta relación inversa en donde las plantas que más transpiraron fueron las que menores contenidos hídricos alcanzaron se mantuvo para los demás genotipos, como también se mantuvo la variabilidad interna entre las muestras de cada genotipo. El PI 416937 fue el que menos variabilidad presentó.

La evaluación de la TA demostró que los genotipos tuvieron diferencias intrínsecas en el consumo de agua



Figura 3.7: RSWC (contenido relativo de agua) en función de los días desde el último riego para los cuatro genotipos. (Izquierda) Cada panel muestra todas las repeticiones de cada genotipo, el color de los puntos varía según el RSWC. (Derecha) Un gráfico para todos los genotipos, los puntos muestran el promedio, la línea un modelo lineal suavizado y el sombreado los intervalos de confianza para cada línea.



Figura 3.8: Transpiración acumulada (L) a través de los días para todas las repeticiones de los cuatro genotipos estudiados. El panel superior muestra las muestras bien regadas, el inferior las sometidas a sequía, el número en la etiqueta indica el número de muestra.



Figura 3.9: RSWC (contenido relativo de agua) en función de los días pasados desde el último. Cada panel contiene las repeticiones de un genotipo, el número en la etiqueta corresponde al número de cada muestra.

total a lo largo del período evaluado. En la figura 3.10a se puede ver el consumo de agua diario acumulado para los cuatro genotipos, el color indica la cantidad acumulada siendo verde para valores menores de 0,5 L; azul para valores cercanos a 1,5 L y gris para los intermedios. Como era de esperarse, todos los genotipos transpiraron menos bajo condiciones de sequía que en control. Si evaluamos la evolución de la transpiración acumulada vemos que el genotipo PI 416937 fue el que alcanzó valores mayores en ambas situaciones, control y sequía. Le siguió en volumen transpirado en situación control el GS, que también se ve azulado en el *heatmap*. En condiciones de sequía puede verse como los genotipos PI 416937 y Williams 82 alcanzaron valores más elevados que los GT y GS (se ven mas azulados en el gráfico).

Si se evalúa la transpiración acumulada en un punto, como ser el día 24 desde el inicio del tratamiento se confirma que todos los genotipos disminuyeron su transpiración ante un evento de sequía (Figura 3.10b). Aunque las diferencias entre genotipos no resultaron significativas, se puede observar una clara tendencia, en donde el genotipo PI 416937 fue el que más transpiró, tanto en sequía como bajo riego, alcanzando valores de 0,89 L y 0,64 L respectivamente. Su transpiración en sequía fue superior a los demás genotipos en la misma condición hídrica, y hasta superior a la alcanzada por el GT bien regado. Los genotipos GS, GT y Williams 82 alcanzaron valores similares de transpiración en situación de sequía que apenas superaron los 0,5 L.

Se graficó para cada genotipo el índice normalizado de transpiración (NTR) en función del contenido hídrico relativo (RSWC) (Figura 3.11). Los datos fueron ajustados a través de una regresión lineal segmentada, en donde los valores que superaron al parámetro X0 formaron un *plateau* en el cual las plantas en sequía transpiraron lo mismo que aquellas bien regadas. Por debajo de este parámetro, se ajustó una respuesta lineal entre el NTR y el contenido de agua relativo. El parámetro X0 sirve entonces como punto de inflexión, en donde comenzó a declinar la transpiración en respuesta a un contenido hídrico limitante. El valor estimado de este parámetro para cada genotipo se indica en el gráfico y corresponde al RSWC al cuál se cerraron los estomas, considerado como RSWC crítico (RSWC_C). Cuanto mayor sea el RSWC_C estimado, indica que ese genotipo responde más rápidamente ante la falta de agua en el suelo, cerrando los estomas minimizando las pérdidas de agua por transpiración.

El RSWCc estimado para los cuatro genotipos estuvo entre un 41 y 54%. El valor más bajo lo presentó el genotipo Williams, seguido por el PI 416937, el GT y luego el GS. Los R² ajustados para las distintas regresiones fueron relativamente altos para todos los genotipos con excepción del GS, que presentó un R² ajustado muy bajo (cercano a 35%). Esto se debió a la separación en dos grupos definidos que se mostró anteriormente. La prueba estadística realizada comprobó con un valor p <0,0001 que se debía rechazar la hipótesis nula correspondiente a ajustar un mismo X0 y pendiente para todos los genotipos. Es decir, la



(a) Transpiración acumulada promedio (L) para cada genotipo, el color indica el volumen transpirado.





Figura 3.11: Índice normalizado de transpiración (NTR) en función del contenido relativo de agua (RSWC) para cada genotipo. Los valores a la derecha de cada gráfico corresponden al RSWC_C (crítico) o X0. Además, se muestra el R² ajustado.

prueba estadística confirma que el modelo preferido es ajustar un parámetro X0 y pendiente inicial distinta para cada genotipo.

Dejando de lado el GS por el bajo ajuste que presentó, si se comparan los demás genotipos se puede observar que el GT y el PI 416937 fueron los que más lograron conservar el agua del suelo. Ambos genotipos cerraron sus estomas a un contenido hídrico cercano al 50 % de la CC limitando la transpiración, a diferencia del genotipo Williams 82 que continuó transpirando sin restricciones hasta alcanzar valores cercanos al 40 %. Esta disminución en la pérdida de agua favorece su conservación y podría demorar el tiempo en el que los genotipos tardan en alcanzar el punto de marchitez permanente.

Aunque resultaron muy similares, el RSWCc para el GT fue apenas superior al del genotipo PI 416937. Sin embargo, éste último resistió menos la sequía porque la transpiración en valores absolutos resultó superior a la del GT (Figura 3.10). Esto quiere decir que por más que ambos genotipos cierren los estomas para preservar el agua del suelo a un RSWC similar, ese RSWC se alcanzaría primero en el genotipo PI 416937 que en el GT.

3.3. Fotosíntesis y parámetros asociados

Cuando se midieron la tasa fotosintética y los parámetros asociados luego de tres días sin riego no se encontraron diferencias. Sin embargo, al cabo de 10 días sin riego las plantas destinadas a sequía habían perdido entre el 75 y 80 % de su contenido hídrico y las diferencias entre genotipos se hicieron evidentes (Figura 3.12). La WUE fue calculada como la relación entre la tasa fotosintética y la transpiración, y resultó significativamente mayor para el GT en condición de sequía con respecto a su control (Figura 3.13). Este resultado evidencia como el GT logra utilizar el agua de una manera más eficiente cuando esta es limitante, utilizando menos para fijar la misma cantidad de carbono. Los valores de WUE obtenidos por este genotipo en condiciones de sequía están en su mayoría por encima de 3 μ mol CO₂ mmol⁻²; llegando hasta 4,15 μ mol CO₂ mmol⁻² y se diferenciaron significativamente de las demás mediciones. Por el contrario, el GS en condiciones de sequía obtuvo WUE entre 2 y 2,7 μ mol CO₂ mmol⁻². En condiciones controladas de humedad, el GT presentó WUE en su mayoría cercanas a 2 μ mol CO₂ mmol⁻², mientras que el GS obtuvo valores entre 1,3 y 1,7 μ mol CO₂ mmol⁻².



Figura 3.12: Evolución del RSWC (contenido hídrico relativo) para los genotipos GS y GT (Susceptible y Tolerante).

Esta elevada eficiencia en el GT estuvo asociada con una menor transpiración producto de una menor conductancia estomática (g_s) (figuras 3.14a y 3.14b). Ambos genotipos disminuyeron su transpiración ante el estrés hídrico, sin embargo, la disminución fue significativamente mayor para el GT. Bajo condiciones de sequía este genotipo transpiró 3 mmol m⁻² s⁻¹, un 60 % menos que el GS, el cual alcanzó valores de hasta 7,2 mmol m⁻² s⁻¹. Esta misma tendencia se observó para la g_s , en donde ambos genotipos presentaron disminuciones frente al estrés, pero en el GT estas fueron mayores y significativas. Bajo condiciones control el GS alcanzó los valores más altos de g_s , que disminuyeron en respuesta a la sequía hasta niveles comparables con los obtenidos por el GT bajo condiciones irrigadas. Ambos resultados encontrados para la transpiración y la g_s explicarían la menor tasa fotosintética encontrada para el GT bajo condiciones de sequía (Figura 3.14c). Se ajustó para cada tratamiendo un modelo hiperbólico significativamente distinto (probado con una suma extra de cuadrados con un valor p < 0,05). Aunque los parámetros de los modelos



Figura 3.13: Eficiencia del uso del agua (WUE) en función de la irradiancia para los 4 tratamientos 10 días después del último riego. Los asteriscos indican diferencias significativas (p < 0, 05).

difirieron entre tratamientos, los valores obtenidos para la tasa fotosintética máxima variaron entre 21 y 26 μ mol CO₂ m⁻² s⁻¹ para todos los tratamientos, excepto para el GT bajo condiciones de sequía la cual alcanzó un máximo de 10,13 μ mol CO₂ m⁻² s⁻¹. Esta saturación a menores irradiancias es explicada por una menor difusión de CO₂ hacia la cámara subestomática, lo que también explica la menor concentración interna de CO₂ (C_i, Figura 3.14d).

3.4. Microscopía

Un detalle de las imágenes tomadas con el miscroscopio electrónico analizadas se puede ver en la figura 3.15. La densidad de células epidérmicas medida como número de células/mm² resultó distinta para cada genotipo y no varió entre las caras de la hoja. El GS presentó una menor cantidad de células epidérmicas, tanto en la cara abaxial (869 cel/mm²) como la adaxial (1006 cel/mm²), diferenciandose de los otros dos genotipos. El GT presentó una densidad de células epidérmicas mayor a 1700 cel/mm² en ambas caras de la hoja, diferenciándose del GS pero no del genotipo de referencia, Williams (Figura 3.16).

La densidad estomática varió significativamente entre genotipos y cara de la hoja (Figura 3.17a). La interacción entre ambos factores (cara y genotipo) resultó significativa, por lo que las diferencias entre genotipos no se mantuvieron a través de las distintas caras de la hoja. Para la cara abaxial el GT presentó una densidad estomática promedio de 385 estomas/mm² diferenciándose significativamente de los 139 y 218 estomas/mm² presentes en el GS y Williams, respectivamente. Tal como era de esperarse, la densidad estomática en la cara adaxial de la hoja fue menor; y esta no difirió significativamente entre genotipos, aunque el GT tendió a presentar una mayor densidad. La densidad estomática depende de la iniciación estomática y de la expansión celular, la cual es afectada por diversos factores ambientales, tal como la temperatura, el estrés lumínico o hidrico. Por el contrario, el índice estomático normaliza la variabilidad introducida por la expansión celular (Ecuación 3.1), ya que depende principalmente de la iniciación estomática. No se encontraron diferencias entre genotipos para este índice, alcanzando en promedio un 16 % y 6 % de la cara abaxial y adaxial, respectivamente, cubierta por estomas (Figura 3.17b).

$$SI(\%) = \frac{StomatalDensity}{StomatalDensity + EpidermalDensity} x100$$
(3.1)

Los estomas del GT, en comparación a los demás genotipos, resultaron un 16% más cortos en ambas caras de la hoja, alcanzando valores de 15 μ M mientras que el GS y Williams alzanzaron valores de 18 μ M (Figura 3.18b). El ancho de los estomas varió significativamente entre genotipos y caras de la hoja. Aquellos ubicados en la cara abaxial fueron un 11% más anchos que los de la cara adaxial, y los del GS más anchos que los de Williams. El GT presentó un ancho intermedio, lo cuál resultó en un índice de apertura estomático (SAI, del inglés *Stomatal Aperture Index*) mayor. El SAI se calculó como la relación entre el ancho y largo de los estomas (Legnaioli, Cuevas y Mas, 2009). Las diferencias encontradas entre genotipos para este índice se mantuvieron entre las caras de la hoja: el GT tuvo un mayor SAI (0,5 y 0,4 para la cara abaxial y adaxial, respectivamente), seguido por el GS y Williams.



Figura 3.14: Transpiración, conductancia estomática (g_s) , tasa fotosintética y concentración interna de CO_2 (Ci) en función de la irradiancia (PPFD). Las barras representan el error estándar de la media y los asteriscos diferencias significativas (p < 0.05). Para la tasa fotosintética se ajustó un modelo hiperbólico significativamente distinto para cada tratamiento.



Figura 3.15: Imágenes tomadas de la microscopía electrónica de la epidermis de los tres genotipos, en detalle se muestra un estoma.



Figura 3.16: Número de células epidérmicas para los genotipos contrastantes (GS y GT y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias significativas con un valor p< 0,05.

4. Discusión

El marchitamiento es el primer síntoma visible de la sequía ya que implica la pérdida de turgencia de los distintos tejidos. El rasgo de marchitamiento lento, (conocido en inglés como *slow wilting*), está asociado con una mayor tolerancia a la sequía ya que implica una menor pérdida, o una mayor capacidad de captación, de agua por parte de las plantas (Kunert y Vorster, 2020; King, Purcell y Brye, 2009; Hufstetler y col., 2007). Los resultados obtenidos permitieron concluir que el GT tolera la falta de agua por más tiempo, retrasando el marchitamiento. Esta tolerancia se debió principalmente a dos razones: el GT consumió menos agua durante el período evaluado, y además, toleró contenidos hídricos más restrictivos que Williams 82, GS y PI 416937 sin evidenciar síntomas de marchitamiento. Esta combinación de menor uso y mayor tolerancia determinó que recién pasados 18 días desde el último riego el GT comenzara a evidenciar síntomas de *wilting*, mientras que los demás genotipos ya tenían la mayoría de sus plantas muertas o en estados avanzados de desecación. Esto representa una gran ventaja a la hora de lidiar con sequías de larga duración o intensidad, ya que pareciera ser que el GT ante faltas hídricas severas es capaz de mantener el normal funcionamiento y homoeostasis celular por un período de tiempo superior a los demás genotipos.

Tal como se esperaba, el GT resultó ser el más tolerante frente a la sequía. Los demás genotipos respondieron de una manera bastante similar entre sí, empezando a evidenciar síntomas de marchitamiento a un RSWC $\sim 20\%$ y alcanzando un 100 en la escala de wilting alrededor de los 15 días. El genotipo Williams 82 respondió de la manera esperada encontrándose siempre dentro del grupo menos tolerante, sin embargo en algunas aspectos el genotipo PI 416937 respondió de manera contraria a lo esperado. Se esperaba que el genotipo PI 416937 tenga un marchitamiento lento, sin embargo, fue el primer genotipo en alcanzar un 100 en la escala de marchitamiento. Estos resultados encontrados coincidieron con aquellos obtenidos por King en el 2009 (King, Purcell y Brye, 2009). Cuando los autores utilizaron el PI 416937 en un ensayo conducidos en cámara de cultivo, este se marchitó más lentamente que aquellos genotipos considerados de rápido marchitamiento. Sin embargo, cuando se correlacionó este marchitamiento con el contenido hídrico no hubo diferencias entre genotipos. Parecería ser entonces, que el genotipo PI 416937 tiene algún mecanismo para disminuir su transpiración, pero que a un mismo contenido hídrico no presenta ninguna ventaja



Figura 3.17: Densidad estomática (a) e índice estomático (b) para los genotipos contrastantes (GS y GT) y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias con un valor p < 0.05 cuando el factor "Genotipo", o la interacción entre ambos factores resultó significativa. Asteriscos indican diferencias significativas entre efectos principales del factor "Cara de la hoja".



Figura 3.18: Ancho (a) y largo estomático (b) para los genotipos contrastantes (GS y GT) y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias con un valor p< 0,05 cuando el factor "Genotipo", o la interacción entre ambos factores resultó significativa. Asteriscos indican diferencias significativas entre efectos principales del factor "Cara de la hoja".



Figura 3.19: Índice de apertura estomático (SAI) para los genotipos contrastantes (GS y GT) y el de referencia (Williams) en la cara abaxial y adaxial de la hoja. Letras distintas indican diferencias con un valor p< 0,05 cuando el factor "Genotipo", o la interacción entre ambos factores resultó significativa. Asteriscos indican diferencias significativas entre efectos principales del factor Çara de la hoja".

por sobre los otros genotipos. Además, cabe destacar que algunos autores citan al PI 416937 como un genotipo de buen comportamiento ante la sequía pero no lo usan en sus experimentos (Pathan y col., 2014; Prince y col., 2015; Song y col., 2016; Zhao, Aleem y Sharmin, 2018). Otros autores si lo usan pero miden otro tipo de variables, como ser la conductancia epidérmica o la eficiencia en el uso del agua (Earl, 2002). Otros autores que lo usan someten las plantas a un tratmiento de deshidratación (descalzando plantas, o sacándolas de la solución hidropónica en la que crecían), no a una sequía en un sustrato (Shin y col., 2015; Sinclair, Zwieniecki y Holbrook, 2008).

También hay que tener en cuenta que el PI 416937 es una introducción que tiene más de 40 años. Los primeros trabajos que lo utilizaron y clasificaron como de buen comportamiento ante la sequía, los cuales muchos autores luego citan, tienen más de 25 años (Sloane, Patterson y Carter, 1990). En tiempos de mejoramiento vegetal, gracias a la ganancia genética, dos décadas son suficiente tiempo como para dejar obsoleto a un genotipo. Los avances que se han logrado en el transcurso de esos años a través de las distintas herramientas de biología molecular y *breeding* hacen que un genotipo comercial de la actualidad resulte ampliamente superior a uno que anteriormente se creía superior. Estos mismos resultados fueron encontrados en el 2020 por el grupo de trabajo de Stekeete (Steketee y col., 2020), donde evaluaron el marchitamiento de 162 genotipos. Los investigadores utilizaron al PI 416937 como control positivo, pero terminaron encontrando que 78 genotipos tenían un marchitamiento más lento que este.

Si consideramos estos resultados junto con los obtenidos en la medición de la transpiración vemos que el GT tardó más tiempo en percibir la falta de agua. Esta tolerancia a la falta de riego pudo haberse debido a dos razones distintas. O las plantas utilizaron menos agua, por lo que tardaron más tiempo en llegar a un contenido hídrico crítico en el cuál comienzan a evidenciar síntomas de respuesta a la sequía (menor transpiración, marchitamiento); o resistieron menores potenciales agua, pudiendo mantener la turgencia y el normal funcionamiento de la planta aún en condiciones hídricas limitantes (Zhao, Aleem y Sharmin, 2018). Pareciera ser que el GT combina ambas estrategias, dado que utilizó menos agua y soportó menores contenidos hídricos con un mejor estado. Utilizó menos agua que los demás genotipos, de hecho, en una situación control transpiró la misma cantidad que el genotipo PI en sequía.

Efectivamente, las plantas del GT utilizaron menos agua que las del GS dado que cerraron los estomas a RSWC mayores que el GS, esto se vio reflejado en un mayor RSWC_c y una menor tanspiración. Este cierre estomático pudo comprobarse al medirse directamente la g_s , la cuál resultó marcadamente menor para el GT en condiciones de sequía. Estos resultados se condicen con los encontrados en otros trabajos, en los cuales los genotipos tolerantes son los primeros en responder ante la falta de agua mientras los susceptibles tardan más en ajustar su respuesta (Hossain y col., 2014; Ray y Sinclair, 1997; Hufstetler y col., 2007). Además, este mecanismo de preservación del agua del suelo se relacionó con un menor wilting en el GT, al igual que lo descripto para maní por Sinclair (Sinclair y col., 2018). El GT además presentó una mayor eficiencia del uso del agua (WUE) bajo condiciones de sequía que los demás genotipos, un rasgo asociado a la tolerancia y deseado en programas de mejoramiento, especialmente si está asociado a buenos rendimientos (Fried, Narayanan y Fallen, 2019).

Los valores de densidad de células epidérmicas obtenidos para el GT (2000 células/mm²) en condiciones de buena disponibilidad hídrica son sorprendentemente altos en comparación con aquellos descriptos por Carrera en el 2021 (Carrera y col., 2021), más aún tratándose de plantas en situación control. En ese trabajo, los autores reportan densidades celulares epidérmicas de 1370 células/mm² bajo condiciones controladas de humedad y 1585 células/mm² en condicioens de sequía. Sin embargo, los resultados de densidad estomática encontrados en este trabajo concuerdan con este autor, en donde citan una mayor densidad estomática en la cara abaxial de la hoja. Tal como se encontró en el presente estudio para el GT, Xu y colaboradores (Xu y Zhou, 2008) hallaron una correlación positiva entre la densidad estomática y la WUE para el pasto Leymus chinensis, y una correlación negativa entre la densidad y el largo estomático en plantas de jojoba y Platanus acerifolia (Liu y col., 2006; Zhang, Xiangrong y Shoubing, 2004). También se ha asociado en papas transgénicas una mayor área asignada a estomas, medido como un mayor índice estomático y SAI, con una capacidad fotosintética aumentada y una mayor tolerancia a la sequía (Crocco y col., 2018). Sin embargo, tal como destaca Xu, en un principio varios autores reportaron un menor tamaño celular y una mayor densidad estomática como una adaptación a la sequía (características propias del GT); sin embargo, ellos reportan una alta plasticidad de estos caracteres frente al estrés, dependiendo de la intensidad del mismo. Describen una relación parabólica entre la densidad estomática y el potencial agua de las hojas, en donde sequías moderadas provocan un aumento en la densidad al disminuir el área foliar, pero sequías

severas disminuyen este parámetro al inhibir la iniciación estomática.

Los mecanismos presentes en el GT que le permitieron tolerar los distintos estreses impuestos en estos experimentos con una mejor *perfromance* que los demás genotipos pueden servir para distintos ambientes. Ante sequías de corta duración, pero recurrentes, este genotipo puede presentar una ventaja dado que conserva el agua del suelo. Es decir, cierra los estomas antes que otros genotipos, minimizando su transpiración y economizando el uso del agua. Este mecanismo tiene como desventaja una menor fotosíntesis y asimilación de carbono. Sin embargo, si las sequías son recurrentes esta capacidad le podría permitir contar con una mejor disponibilidad hídrica ante un nuevo evento de sequía. Asimismo, si consideramos una sequía severa de larga duración esta elevada sensibilidad estomática vuelve a ser una ventaja ya que tardaría más tiempo en llegar a contenidos hídricos nocivos. Además, el GT tolera menores contenidos hídricos sin pérdidas de turgencia. Se ha comprobado a través de modelos robustos de simulación que este *trait* permitiría incrementar los rendimientos en el 70 % de los años en Estados Unidos, acarreando mínimas pérdidas en años húmedos (Sinclair y col., 2010); mientras que otros autores afirman que esta característica aportaría tolerancia a la sequía sin comprometer el rendimiento (Kunert y Vorster, 2020).

5. Conclusiones

La sequía es uno de las principales adversidades abióticas que amenazan la productividad de los cultivos de grano. El mejoramiento para conseguir genotipos más eficientes en el uso del agua es complejo dado que las plantas pueden adoptar múltiples estrategias para hacerle frente al estrés y cuáles se seleccionen dependerá principalmente del ambiente objetivo. Es por esto que al estudiar la respuesta a la sequía es importante medir distintos parámetros, preferentemente en distintas fechas, ambientes climáticos y edáficos, y regiones de cultivo para estimar la potencialidad de un genotipo en función al ambiente ya que es sabido que las interacciones GxE (genotipo por ambiente) son importantes para la *performance* de los cultivos.

Partiendo de la base del germoplasma de la empresa GDM se pudieron identificar dos genotipos con respuestas contrastantes a la sequía (GT y GS). Estos genotipos fueron evaluados en distintos experimentos en los cuáles se incluyeron otros genotipos de respuesta conocida ampliamente citados en la bibliografía (PI 416937 y Williams 82). A través de esta evaluación se pudo comprobar la primer hipótesis, la cuál establece que el GT se diferencia del GS en diversos parámetros fisiológicos y morfológicos que le permiten utilizar el agua de una manera más eficiente minimizando el estrés provocado por la sequía. Se verificó además que esta respuesta contrastante era comparable a la de los genotipos de bibliografía. A diferencia del GS, el GT resistió significativamente más días en sequía sin presentar marchitamiento y pudo conservar el agua del suelo minimizando la transpiración al cerrar los estomas a contenidos hídricos más altos. Esta elevada sensibilidad estomática trajo aparejada una mayor eficiencia del uso del agua (WUE) a pesar de tener una menor tasa fotosintética. Estos resultados hacen del antagonismo entre genotipo tolerante y susceptible una interesante oportunidad para continuar estudiando las bases genéticas detrás de esta respuesta, con vistas al desarrollo de marcadores moleculares que sean posibles de incluirse en un plan de mejoramiento.

Capítulo 4

Estudio Transcriptómico

1. Introducción

La secuenciación del transcriptoma produce millones de lecturas a partir de muestras complejas de RNA y tiene diversas aplicaciones. Dependiendo de la profundidad de lectura alcanzada, es utilizada para medir la expresión génica, la anotación de nuevos transcriptos y la caracterización de variables de *splicing*, entre otros. La técnica se conoce generalmente como secuenciación del transcriptoma, o RNAseq en su versión abreviada, y hoy se utiliza de manera masiva gracias a la aparición de las plataformas de secuenciación de nueva generación (NGS, del inglés *Next Generation Sequencing*) que permitieron que sea rápido, accesible y confiable. Illumina es una de las plataformas más utilizadas y se basa en una secuenciación por síntesis. Para comenzar con este tipo de estudios, primero se debe realizar una extracción de RNA, luego se construye la *library*, se realiza la secuenciación propiamente dicha y por último, se analizan los datos obtenidos. Antes de avanzar al siguiente paso, cualquiera sea, se realizan controles de calidad (QC, Figura 4.1). El RNAseq permite conocer el nivel de expresión de todos los genes del tejido estudiado, dando una visión general de los procesos que estaban ocurriendo en el momento en el que se tomó la muestra. Al comparar distintos tratamientos se pueden identificar los genes específicos que respondieron ante dicho tratamiento.

La tolerancia a la sequía es un rasgo sumamente complejo, dado que esta gobernado por múltiples QTLs de baja heredabilidad y que, además, interaccionan con el ambiente (Tripathi y col., 2016; Alam Khan, 2018). Dada esta complejidad molecular y genética, el mejoramiento asistido por un gran panel de marcadores moleculares es una herramienta de gran utilidad. Estos marcadores pueden ser generados utilizando el polimorfismo presente en múltiples QTLs de interés en simultáneo, lo que hoy en día no representa un desafío tecnológico gracias las plataformas de high throughput como la SNPline, que permiten el análisis de cientos de miles de SNPs por día; o los chips prediseñados para el testeo de miles de SNPs específicos para una especie. Sin embargo, el desafío reside en el descubrimiento de aquellos genes asociados, o causantes, de la tolerancia. Estos genes pueden encontrarse de manera más inespecífica a través de estudios de asociación del genoma completo (GWAS del inglés, Genome Wide Association Studies), en donde se asocian rasgos fenotípicos de interés con polimorfismos presentes en el genoma. En este tipo de estudios lo importante es la asociación entre el polimorfismo y el fenotipo, sin importar el gen específico detrás del trait deseado. Por el contrario, el enfoque transcriptómico permite encontrar cuáles son los genes, y las rutas metabólicas, detrás de la tolerancia (Hua y col., 2018). La realización de estudios transcriptómicos con líneas contrastantes provee información detallada sobre genes diferencialmente expresados que podrían ayudar a la identificación de genes involucrados en la tolerancia a la sequía (Prince y col., 2015). A partir de estos estudios se pueden asociar distintos polimorfismos a la tolerancia ante un estrés en particular, para luego poder usarlos como marcadores moleculares en el mejoramiento.

Los genes que responden ante la sequía pueden dividirse en dos grandes categorías: los efectores y los reguladores (Buchanan, Gruissem y Jones, 2015; Zhao, Aleem y Sharmin, 2018; Thao y Tran, 2012). La primer categoría comprende genes que codifican para proteínas que actúan específicamente mitigando el estrés provocado por la sequía, como las proteínas LEA (late embryogenesis abundant proteins) que pro-



Figura 4.1: Esquema general de la secuenciación.

tegen a otras proteínas de la agregación provocada por la desecación o estreses osmóticos. Otro ejemplo son las proteínas involucradas en la biosíntesis de osmolitos, acuaporinas, chaperonas, antioxidantes y otras enzimas involucradas en distintas rutas metabólicas. Los genes reguladores son aquellos que codifican para proteínas que van a mitigar el estrés de manera indirecta, como los factores de transcripción, distintos receptores, quinasas, fosfatasas y proteínas que pliegan la calmodulina (proteína que regula la trasducción de la señal de calcio en todas las células eucariotas).

Existen diversos factores de transcripción que aumentan su expresión en respuesta a la sequía (Joshi y col., 2016). Según la base de datos de factores de transcripción de plantas (Plant TFDB), en soja las familias más grandes son la bHLH, MYB, bZIP (*basic leucine zipper domain*), ERF (*ethylene responsive factor*) y WRKY. Algunas están involucradas en la señalización de la vía del ABA en respuesta a la sequía, como los MYB, ERF , bZIP y NAC. Otras como la DREB (*dehydrationt responsive element binding protein*), actúan de manera independiente al ABA (Joshi y col., 2016).

En este capítulo se aborda el segundo objetivo específico planteado en el Capítulo 1, en donde se pretende dilucidar cuales son los genes y las vías metabólicas claves expresadas diferencialmente en respuesta a la sequía. Para cumplir con este objetivo se pondrá a prueba la segunda hipótesis, que establece que el GT presentará enriquecidas vías metabólicas propias de la tolerancia al estrés.

2. Materiales y Métodos

2.1. Diseño del experimento

El experimento se condujo según lo especificado en el Capítulo 2, manteniendo las plantas bien regadas hasta llegar al estado de V2. En ese momento se dejaron de regar las plantas destinadas a sequía, mientras que las destinadas a control se mantuvieron con una buena disponibilidad hídrica. Al cabo de 10 días, luego de realizar las mediciones de fotosíntesis, se tomaron las muestras seleccionando las mismas hojas que sobre las cuales se había trabajado (la segunda hoja trifoliada completamente expandida). Estas se guardaron en sobres individuales y se congelaron en N líquido inmediatamente. Luego, se las guardó a -80 °C hasta su posterior utilización.

2.2. Extracción del RNA y control de calidad

Se extrajo el RNA total de las muestras con el kit RNeasy Plant Mini Kit (QIAGEN) siguiendo las instrucciones del fabricante. Posteriormente, se verificó que las muestras cumplan con los parámetros de calidad requeridos para la secuenciación (masa, concentración, integridad y pureza). La concentración se cuantificó mediante el fluorómetro Qubit (Invitrogen Corporation, USA). Este instrumento mide la fluorescencia que emite un colorante específico, la cual se relaciona directamente con la concentración de ácidos nucleicos en solución. Esta señal de fluorescencia es convertida en una medición de concentración de RNA mediante la utilización de estándares de concentración conocida. Para la verificación de la integridad se realizó una electroforesis en gel de agarosa al 1,5 % bajo condiciones nativas. Para comprobar la pureza del RNA extraído se midió la absorbancia a 260, 280, 230 y 320 nm de 3500 ng disueltos en buffer TE. Los ácidos nucleicos tienen su pico de absorción a 260 nm, mientras que distintos contaminantes (proteínas, fenoles, guanidina, glicógeno, hidratos de carbono) tienen sus picos de absorción a 230 y 280 nm, y daños o polvo en las cubetas provocan un pico de absorción a 320 nm.

2.3. Secuenciación

Para la secuenciación del transcriptoma se utilizó la plataforma de NGS Illumina y se recurrió a la empresa Novogene, en California. Una vez que llegaron las muestras al laboratorio, se volvió a chequear su calidad (concentración, integridad y pureza) y aprobada esta instancia, se tomó 1 µg de RNA de cada muestra para la construcción de la librería (library). Para esto se utilizó el kit NEBNext®UltraTMRNA Library Prep Kit for Illumina® y se siguieron las indicaciones del fabricante. En primer lugar, para purificar el mRNA se utilizaron partículas magnéticas con oligo poly-T (una secuencia de timinas que se aparea con la cola poli-A del mRNA). La fragmentación se obtuvo gracias a la acción de los cationes divalentes en el NEBNext First Strand Synthesis Reaction Buffer (5X) y el sometimiento a altas temperaturas. La primera hebra del cDNA se sintetizó utilizando random primers y una retrotranscriptasa (M-MuLV Reverse Transcriptase); y la segunda hebra se sintetizó con una DNA polimerasa I. Se cortaron los extremos restantes, luego se adeniló el extremo 3' y, por último, se adicionaron los adaptadores (NEBNext Adaptor). Para seleccionar fragmentos de entre 150 y 200 pb se purificó la librería con el AMPure XP system (Beckman Coulter, Beverly, USA). Una vez que se contó con el cDNA con los adaptadores ligados y filtrado por longitud, se incubó con 3 µl de la enzima USER (NEB, USA) a 37 °C por 15 minutos, seguido por 5 minutos a 95 °C antes de la PCR. Luego, se corrió la PCR con una enzima de alta fidelidad (Pushion High Fidelity DNA polymerase), primers universales y primers Index (X). Finalmente, el producto de la PCR se filtró y se controló su calidad (Agilent Bioanalyzer 2100). El clustering (generación de grupos) de las muestras se realizó con un cBot Cluster Generation System siguiendo las instrucciones del fabricante en el Cluster Kit cBot-HS (Illumina). Luego se secuenciaron las muestras en una plataforma Illumina generando 40 millones de lecturas de entre 125 y 150 pb del tipo paired-end.

2.4. Procesamiento y análisis de los datos

Una vez recibidos los resultados de la secuenciación se procedió al análisis bioinformático utilizando distintos paquetes de R (R Core Team, 2020); los pasos básicos que se siguieron se detallan en la figura 4.2. El primer paso consistió en tomar las imágenes originales de la secuenciación de alto rendimiento y transformarlas en lecturas (*reads*) a través del software CASAVA. Estos archivos de datos crudos (*raw data*), se guardaron en formato FASTQ que además contienen información sobre la calidad de la lectura para cada base. El siguiente paso consistió en analizar la calidad de los datos obtenidos a través de tres variables distintas: la tasa de error, el contenido de GC y el filtrado de los datos. Se calculó para cada base el Q-score, que se basa en la probabilidad de que la lectura de esa base haya sido incorrecta (*error rate*). A medida que la probabilidad aumenta, el Q-score disminuye. Luego, se graficó el contenido relativo de cada base en función de la posición a lo largo de la lectura. El contenido de G/C y A/T deberían ser iguales y estables a lo largo de la secuenciación. En las primeras 6 bases se acepta una alta variación dado que se utilizan *random primers* para la construcción de las librerías de RNAseq. Para continuar con los análisis se eliminaron las lecturas que contenían fragmentos de los adaptadores, más del 10 % de bases indeterminadas o un Q-score bajo. Una vez que se contó con los datos de calidad, estos se mapearon al genoma de referencia de la soja.



Figura 4.2: Pasos seguidos en el análisis de datos y la metodología utilizada en cada uno.

Se alinearon las lecturas obtenidas al genoma de referencia disponible en Phytozome correspondiente a la versión 4 de la variedad Williams 82. Para realizar el alineamiento se utilizó el algoritmo HISAT2 que utiliza tres pasos básicos: primero alinea las lecturas que corresponden a un único exón en el genoma, luego las lecturas son segmentadas y mapeadas a exones adyacentes del genoma, y por último, se segmentan las lecturas y se mapean a 3 o más exones del genoma. El porcentaje de lecturas limpias (*clean reads*) mapeadas al genoma debe superar el 70%, mientas que aquellas lecturas que mapean a más de un sitio (*multiple mapped reads*) no deben superar el 10%. Otro aspecto a tener en cuenta es la distribución de las lecturas a lo largo de los cromosomas. La mayoría de las lecturas deberían mapear a regiones puramente exónicas. La aparición de lecturas que mapean a intrones puede ser producto de contaminación con pre-mRNA o por retención de intrones en eventos de splicing alternativo. También puede haber lecturas que mapean a sitios inter-génicos debido a una pobre anotación del genoma de referencia.

Para continuar con el análisis de los datos se calculó el nivel de expresión génica a través del cálculo del factor FPKM (*fragments per kilobase of transcript sequence per millions base pairs sequenced*). El factor FPKM es utilizado para determinar el nivel de expresión de un gen a partir de la cantidad de lecturas que mapean a éste, teniendo en cuenta su longitud y la profundidad de la secuenciación (Trapnell y col., 2010). Luego 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. Para asegurar la repetitividad del experimento y la veracidad de los resultados de los cálculos de expresión diferencial, la correlación entre réplicas biológicas debería superar el 0,92 y el R² superior a 0,8 (Hansen y col., 2011). Esto se verificó para las cuatro réplicas biológicas presentes por tratamiento. Otra manera de visualizar la coexpresión a través de los diagramas de Venn, que muestran el número de genes que se expresan únicamente para cada tratamiento y aquellos que se coexpresan entre muestras en las regiones superpuestas.

Para el cálculo de la expresión diferencial de genes se usó el paquete de R DESeq2, que tiene las herramientas estadísticas para calcular la expresión diferencial en cada muestra (Anders y Huber, 2010). Este algoritmo se basa en la distribución binomial negativa y los valores p obtenidos se corrigieron utilizando la técnica de Hochberg y Benjamini para disminuir la tasa de descubrimientos falsos (Benjamini y Hochberg, 1995). 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. Este análisis se realizó con el paquete de R cluster-Profiler (Yu y col., 2012) con dos bases de datos distintas, *Gene Ontology* (GO Ashburner y col., 2000; The Gene Ontology Consortium, 2021) y KEGG (Kyoto Encyclopedia of Genes and Genomes, Kanehisa y Goto, 2000). La primera es un sistema de clasificación de genes para todas las especies. Se agrupan genes de manera jerárquica en categorías correspondientes a su función. Las principales tres categorías que engloban a todas las demás son: componentes celulares, función molecular, procesos biológicos. Se consideraron categorías (*terms*) significativamente enriquecidos aquellos que obtuvieron un p-valor ajustado < 0, 05. Por otro lado, KEGG es una base de datos que permite identificar rutas metabólicas enriquecidas a partir de genes diferencialmente expresados. Los términos con un valor p ajustado < 0, 05 se consideraron significativamente enriquecido un análisis de enriquecimiento con la herramienta agriGO Du y col., 2010 y la posterior agrupación de las categorías con la herramienta REViGO (Supek y col., 2011), que permitió una mejor visualización de los datos.

3. Resultados

3.1. Controles de calidad

El RNA extraído cumplió con los requisitos de calidad necesarios para su secuenciación. Para comprobar su integridad se sembraron 500 ng del RNA total en un gel de agarosa (Figura 4.3). Luego, se midió su absorbancia y se observó que el único pico de absorción fue a 260 nm y el ratio 260/280 fue en todos los casos mayor a 2, lo que indicó una elevada pureza.



Figura 4.3: Gel de agarosa 1,5 % del RNA extraído. El nombre y número de cada muestra se encuentra en cada well.

Luego de la secuenciación se realizó el control de calidad de los datos obtenidos. Todos los datos cumplieron con los estándares de calidad mencionados anteriormente: una tasa de error menor al 1 % y un contenido de G/C similar a lo largo de las lecturas. Se filtraron las lecturas que contenían fragmentos de adaptadores y aquellas con más de un 10 % de indeterminaciones para realizar la alineación al genoma de referencia. Una vez realizado el mapeo se corroboró que aproximadamente el 95 % de las lecturas correspondieron a regiones exónicas y que estas se distribuyeron igual para todas las muestras a lo largo del genoma. Por último, se procedió al cálculo de la expresión de los genes a través del FPKM y se comprobó que la correlación entre réplicas biológicas fue mayor a 0,92.

Una vez asegurada la calidad de los datos se realizaron los cálculos de expresión diferencial de cada genotipo bajo estrés con respecto a su control, y con esos datos se realizó el análisis de enriquecimiento funcional.



Figura 4.4: Resultado del análisis de clusters de los FPKM para cada tratamiento, utilizando el log10(FPKM+1). El color rojo indica genes que disminuyeron su expresión y el verde genes que la aumentaron. El rango de colores del verde al rojo indica el nivel de expresión de mayor a menor.

3.2. Genes diferencialmente expresados

Al realizarse un análisis de clusters de los genes expresados para cada genotipo se pudieron encontrar grupos de genes con patrones de expresión similares a través de las distintas condiciones experimentales. En la figura 4.4 se puede observar como la expresión de algunos genes aumenta en respuesta a la sequía y como otros disminuyen en respuesta al estrés, sobretodo en el GT.

Para los análisis posteriores, en primer lugar se calcularon los genes diferencialmente expresados de cada genotipo con respecto a su control. Es decir, los resultados mostrados a continuación evidencian los genes que respondieron ante el estrés para cada genotipo en particular. Una vez realizado este análisis, se procedió a comparar la respuesta entre genotipos. Es por esto que se muestran primero los resultados para cada genotipo y luego una comparación entre ambos. El detalle de la totalidad de los genes diferencialmente expresados para cada genotipo puede verse en el Apéndice A.

3.2.1. Genotipo Tolerante

Para el genotipo tolerante la comparación entre condición sequía y control arrojó un total de 942 genes diferencialmente expresados, de los cuáles casi la mitad aumentó su expresión y la mitad restante la disminuyó (Figura 4.5). Como se verá más adelante, estos valores superaron ampliamente a los alcanzados por el GS, donde solo 421 genes cambiaron su expresión (Figura 4.8). Que la respuesta a la sequía en el GT involucre más genes nos permite hipotetizar que la respuesta inducida involucra una mayor cantidad de redes de señalización y complejidad. Pocos genes que disminuyeron su expresión pudieron ser mapeados a



Figura 4.5: Volcano plot. Expresión diferencial de genes para el genotipo tolerante. Nivel de significancia expresado como el -log10(padj) en función del nivel de expresión diferencial medido como el log2(Fold Change). Las etiquetas muestran el número de genes regulados hacia arriba (Up) y regulados hacia abajo (Down).

categorías de procesos biológicos (BP) de GO, la gran mayoría correspondió a categorías relacionadas con funciones moleculares (MF) y componentes celulares (CC). A continuación se describirán los genes que cambiaron su expresión en función de las anotaciones existentes para cada uno en las bases de datos con las que está enriquecido la anotación del genoma: PFAM (Mistry y col., 2021), PANTHER (Mi y col., 2021), UniPort (Bateman y col., 2021), KOG (Grigoriev y col., 2012) y GO (The Gene Ontology Consortium, 2021).

Genes que aumentaron su expresión

Los genes que aumentaron su expresión en el GT sometido a sequía no pudieron asociarse a ninguna categoría de la subontología procesos biológicos (BP) de GO, pero si a categorías de la subontología componentes celulares (CC) y funciones moleculares (MF) (Figura 4.6). Las categorías de componentes celulares guardaron relación con el DNA y su empaquetamiento: cromatina, nucleosoma, cromosoma, etc. Las categorías enriquecidas dentro de la subontología MF estuvieron relacionadas con la dimerización de proteínas, la actividad redox y la regulación de peptidasas. Particularmente, las categorías de dimerización de proteínas corresponden a la regulación de la transcripción mediada por el DNA, hubo distintos factores de transcripción, histonas y metiltransferasas que aumentaron su expresión. La categoría de dioxigenasas y actividad oxidorreductasa corresponde a lipooxigenasas, en su mayoría cloroplastídicas, algunas de las cuales están involucradas en oxidorreducción de coenzimas en respuesta a desecación, falta de agua, ácido abscício, entre otros. Además de las lipooxigenasas, hubo otros genes relacionados al mantenimiento del estado redox que aumentaron su expresión, como la tiorredoxina m y una peroxidasa. Las categorías que corresponden a la actividad de endopeptidasas son generalmente inhibidores de proteasas con dominos cystatin (de la familia cisteín proteasas) generalmente presentes en el apoplasto y la región extracelular de respuesta hiperosmótica, a estreses abióticos, nitrato y estrés del retículo endoplasmático; lo que tiene sentido ya que como se verá mas adelante disminuyen numerosos componentes del ribosoma. Dentro de la categoría de regulación de peptidasas, al igual que en la de endopeptidasas hay inhibidores de proteasas con dominios cystatin pero también hay inhibidores de tripsinas. Muchos de estos están situados en el citosol, específicamente en el retículo endoplasmático. Al igual que las anteriores responden a diversos tipos de estrés, entre ellos el oxidativo y la deprivación de agua. Por último, la categoría de actividad transferasa de grupos acilos y aminoacilos que resultó enriquecida contiene genes involucrados en la síntesis de glucosinolatos, ácidos grasos de cadena larga, cutícula, ceras, cisteínas, carotenoides, clorofilas, síntesis y acumulación de antocianinas, reorganización de la pared celular, respuesta a estímulos lumínicos (luz azul, rojo, rojo lejano,



Figura 4.6: Las 10 categorías funcionales más significativas enriquecidas para cada sub-ontología de GO (BP= procesos biológicos, CC= componente celular, MF= función molecular) para la comparación del GT en sequía con respecto a su control. En rojo (izquierda) aquellas que resultaron subrepresentadas y en verde (derecha) aquellas que resultaron sobrerepresentadas. El número dentro de la barra indica la cantidad de genes pertenecientes a esa categoría que resultaron enriquecidos. Los asteriscos indican enriquecimientos significativos (padj<0,05).

ultravioleta), alta intensidad lumínica y frío, entre otros.

Hubo otros genes que aumentaron su expresión pero no se vieron reflejados en las categorías enriquecidas de GO. Estos tuvieron que ver con procesos clave en la respuesta a la sequía tanto a nivel fisiológico y celular, como en la expresión de genes y la transducción de señales y se pueden ver resumidos en la figura 4.19, de la cual se hablará más adelante. En primer lugar, a nivel fisiológico y celular, se destacaron genes involucrados en la acumulación de solutos compatibles. Estos son compuestos orgánicos de bajo peso molecular que pueden acumularse para mantener el equilibrio osmótico sin interferir en el metabolismo celular. En el GT aumentó la expresión de enzimas involucradas en la síntesis de glicina beatina (fosfo-colina fosfatasa), mioinositol, prolina y rafinosa. Los carbohidratos como el mioinositol, la rafinosa y el galactinol (de los cuales el GT presentó enzimas claves de su síntesis sobreexpresadas) no solo sirven como reservas de energías, sino también regulan el potencial osmótico, estabilizan las macromoléculas y preservan la integridad de las membranas. Otro tipo de proteínas importantes en la respuesta a la sequía son las HSPs y las LEA, que actúan reteniendo el agua al ser altamente hidrofílicas o como chaperonas moleculares, y también resultaron sobreexpresadas en respuesta a la sequía en el GT. En segundo lugar, la falta de agua induce la expresión de numerosos genes a través de factores de transcripción y la activación de vías de señalización (véase apartado 3.3.2). Estas vías pueden ser: sistema regulador de dos componentes (phosphorelay system), las MAPK (mytogen activated protein kinases), y mensajeros secundarios como el Ca²⁺ y las ROS (reactive oxygen species). Estas vías de transducción de señales se vieron favorecidas en el GT bajo condiciones de sequía. En la figura 4.19 se puede ver como una de las categorías enriquecidas según REViGO fue el sistema de dos componentes. Este tipo de sistemas comprende un mecanismo de señalización constituido por una quinasa que ante un estímulo específico se autofosforila para luego transmitir esa señal al fosforilar una proteína reguladora de la respuesta. Además, aumentó la expresión de varios genes involucrados en la cascada de señalización por MAPK, como la AOS (allene oxidase synthase, parte del citocromo P450), proteínas de unión y transporte de calcio.

Genes que disminuyeron su expresión

En el genotipo tolerante sometido a sequía todos los genes que disminuyeron su expresión, según el enriquecimiento llevado a cabo con GO, correspondieron a categorías asociadas al ribosoma (Figura 4.6). Dentro de la subontología procesos biológicos se encontraron las categorías: biogénesis de los ribosomas y de los complejos ribonucleoproteicos; y el procesamiento y metabolismo del rRNA. En la subontología funciones moleculares se destacaron las categorías: constituyentes estructurales del ribosoma, actividad de moléculas estructurales y unión al RNA y al rRNA. Como era de esperarse, en la subontología componentes celulares también resultaron enriquecidas únicamente categorías ribosomales. El detalle de los genes estructurales que disminuyeron su expresión se puede ver en la figura 4.7. Dentro de las categorías mencionadas se encontraron genes relacionados con la transducción de señales, como el silenciamiento de la cromatina y la metilación del RNA en respuesta a estrés salino. No solo se vieron disminuidos genes que codifican para componentes estructurales del ribosoma, sino también aquellos que codifican para purinas, pirimidinas, ribosas y demás moléculas o enzimas involucradas en la transcripción y traducción del mensaje genético.

Aunque no se hayan visto como categorías enriquecidas en GO hubo muchos otros genes aislados que disminuyeron su expresión; sin embargo, son genes puntuales y no significa una disminución en toda la vía metabólica en la que participan. Dentro del metabolismo del carbono disminuyó la expresión del transportador de triosas fosato del cloroplasto que actúa como un regulador de la fotosíntesis y responde a sacarosa. También disminuyó la expresión de la hexoquinasa (una enzima que cataliza la fosforilación de la glucosa) y de la glucosa 6-fosfato deshidrogenasa (la primera enzima en la vía de las pentosas fosfato), entre otras. El gen que más disminuyó su expresión fue el que codifica para la enzima nitrato reductasa, que cataliza la primera reducción asimilatoria del nitrógeno. También disminuyó la expresión de la nitrito reductasa, sulfito reductasa y una nodulina involucrada en la síntesis de glucosinolatos. Asimismo, disminuyó la expresión de genes que podrían ayudar en la respuesa a la sequía: HSP (*heat shock proteins*), y cold shock proteins, calmodulina, proteínas relacionadas con la aclimatación al calor, una tiorredoxina, la ascorbato peroxidasa y otras peroxidasas. Otros genes que podrían estar asociados a una respuesta positiva frente al estrés hídrico, como la diferenciación entre las caras adaxiales y abaxiales de la hoja, modificación de la pared celular y plasmodesmos.La expresión de distintos canales se vio afectada, disminuyeron algunas acuaporinas y canales transportadores de iones, amonio, UDP-glucosa y aminoácidos.



Figura 4.7: Diagrama de los componentes del ribosoma (KEGG gmx03010). El color corresponde al nivel de expresión para el genotipo tolerante bajo condiciones de sequía con respecto a su control.



Figura 4.8: Volcano plot. Expresión diferencial de genes para el GS. Se muestra el nivel de significancia como el -log10(padj) en función del nivel de expresión diferencial medido como el log2(Fold Change). Las etiquetas muestran el número de genes regulados hacia arriba (Up) y regulados hacia abajo (Down).

3.2.2. Genotipo Susceptible

La respuesta de un genotipo la estrés, cuando se compara con respecto a la situación control, puede explicarse a través de las rutas metabólicas que se activan como respuesta al tratamiento. En el caso del GS ante una condición de sequía aumentó significativamente la expresión de 182 genes y disminuyó la de 241, que según la base de datos Gene Ontology (GO) pudieron agruparse en 47 categorías distintas. En esta misma situación aumentó la expresión de 182 genes que se agruparon en 12 categorías (Figuras 4.8 y 4.9). Al igual que para el GT, a continuación se describirán los genes que cambiaron su expresión en función de las anotaciones existentes para cada uno en las bases de datos con las que está enriquecido la anotación del genoma: PFAM (Mistry y col., 2021), PANTHER (Mi y col., 2021), UniPort (Bateman y col., 2021), KOG (Grigoriev y col., 2012) y GO (The Gene Ontology Consortium, 2021).

Genes que aumentaron su expresión

Una de las primeras cosas que llama la atención sobre los genes diferencialmente expresados es que aquellos que aumentaron su expresión son un 35 % menos que aquellos que la disminuyeron (Figura 4.8). Dentro de estos, ninguno pudo mapearse a categorías funcionales de procesos biológicos (BP) de GO, solo algunos pudieron mapearse a compartimentos celulares (CC) y otros pocos a funciones moleculares (MF) específicas (Figura 4.9). De la subontología CC resultaron enriquecidas todas categorías correspondientes a la información genética y su procesamiento: nucleosoma, cromatina, cromosoma, empaquetamiento del DNA y complejo DNA-proteína. Todos los genes sobre-representados dentro de estas categorías correspondieron a histonas involucradas en el ensamblado de nucleosomas, y algunos, de respuesta a la falta de agua. También aumentó la expresión de una proteína HMG box (high mobility group), que está asociada con la regulación de procesos involucrados en la transferencia de la información genética como la replicación, transcripción y traducción. Esta proteína en particular responde a cambios osmóticos, temperatura y transporte de agua. De la subontología MF unas de las categorías más sobre-representadas estuvieron relacionadas con la actividad de dimerización de proteínas. Dentro de estas, los genes que aumentaron su expresión fueron, nuevamente, histonas y distintos factores de transcripción (de los cuales se hablará en el apartado 3.3.2). En las categorías de oxidorredución se encontraron mayoritariamente lipooxigenasas cloroplastídicas y una mioinositol oxigenasa (considerado un osmoprotector con actividad antioxidante). Estas lipooxigenasas regulan varios procesos, como el transporte de Ca^{2+} y otros iones, la síntesis de carotenoides, clorofilas, oxilipinas; y responden a distintos estímulos, como frío, falta de agua, daño mecánico, herbivoría, hongos, salinidad, etc. Otros genes aislados de interés redox que aumentaron su expresión, pero que no aparecieron anotados



Figura 4.9: Las 10 categorías funcionales más significativas enriquecidas para cada sub-ontología de GO (BP= procesos biológicos, CC= componente celular, MF= función molecular) para la comparación del GS en sequía con respecto a su control. En rojo (izquierda) aquellas que resultaron subrepresentadas y en verde (derecha) aquellas que resultaron sobrerepresentadas. El número dentro de la barra indica la cantidad de genes pertenecientes a esa categoría que resultaron enriquecidos. Los asteriscos indican enriquecimientos significativos (padj<0,05).

en GO, fueron: la tiorredoxina *m3*, una glutarredoxina, una glutatión transferasa de respuesta a proteínas mal plegadas, y una proteína resistente a la oxidación. Otra categoría que se vio enriquecida fue la actividad transferasa de grupos hexosas, dentro de la cuál se encontraron varios genes que codifican para la síntesis de enzimas con actividad UDP-glucosil o glucoronosil transferasas en el cloroplasto. Estas enzimas están involucradas en la glucosilación de zeatinas en el metabolismo de las citocininas.

Hubo otros genes que aumentaron su expresión pero no se incluyeron en categorías funcionales enriquecidas de GO dado que fueron genes aislados dentro de las vías metabólicas. Por ejemplo, genes del complejo de la ubiquitin ligasa involucrados en la proteólisis, proteínas LEA (late abundan embryogenesis) que evitan la agregación de proteínas en períodos de estrés y genes involucrados en la síntesis y transporte de distintos compuestos. De estos últimos, se destacan aquellos que codifican para la síntesis de triacilglicéridos en el cloroplasto y una enzima involucrada en la síntesis de ceras, pared celular, cutícula y ácidos grasos de cadena larga (Omega-hydroxypalmitate O-feruloyl transferase). También aumentó la expresión de enzimas involucradas en la síntesis de nucleótidos, de isoprenoides en la vía del ácido mevalónico y de compuestos fenólicos. Otros genes que se vieron sobre-expresados tuvieron que ver con la síntesis de transportadores y la señalización celular. El Ca²⁺ actúa como mensajero secundario en distintas vías de señalización y en la respuesta frente al estrés abiótico. En el GS aumentó la expresión de varios genes involucrados en su regulación frente a la sequía. Aumentó la expresión de sus transportadores, de la calmodulina y de proteínas de unión a la calmodulina (con motivo VQ y de respuesta a estrés). Otros transportadores que aumentaron su expresión fueron de iones (K⁺, Na⁺) y un transportador transmembrana de compuestos nitrogenados (amonio, aminoácidos y nucleótidos) de respuesta a hipoxia, ABA, SA y que está involucrado en la regulación del metabolismo del peróxido de hidrógeno. También aumentó la expresión de un transportador de respuesta a estrés de la familia ABC (ATP binding casette) que está involucrado en el movimiento estomático. Asimismo, aumentó la expresión de dos nodulinas, una de las cuáles transporta azúcares transmembrana. En lo que respecta a la señalización celular, aumentó la expresión de quinasas involucradas en la cascada de señalización mediada por MAPK, que responden a fitohormonas, proteínas mal plegadas, estrés oxidativo, falta de agua; una de las cuales está involucrada en el complejo de morfogénesis de estomas. Hubo otros genes que escaparon a la clasificación de GO pero resultaron de interés, como una proteína stay green que regula la degradación de la clorofila y una proteína condominio WD que está relacionada con la especialización de células de la epidermis, especialmente en la diferenciación de tricomas. Esta proteína responde a auxinas, etileno y regula positivamente la biosíntesis de flavonoides.

Genes que disminuyeron su expresión

Ante una situación de sequía el GS parecería tener procesos fundamentales del metabolismo primario disminuidos, como el metabolismo de nucleótidos y la asimilación del C, N y S. Además, se ven disminuidas vías que podrían mejorar la respuesta a la sequía, por ejemplo: la biosíntesis de diterpenos, fenilpropanoides y el procesamiento de proteínas en el retículo endoplasmático (Figura 4.9).

Aunque cuando se midió la fotosíntesis en el GS esta no resultó disminuida significativamente, hay varios genes pertenecientes a estas vías que si disminuyeron significativamente su expresión. Por lo que, aunque la fotosíntesis actual no resultó muy afectada, se podría asumir que sería una cuestión de tiempo hasta que esta comience a disminuir significativamente en estas condiciones. En lo que respecta a la etapa lumínica de la fotosíntesis, en el cloroplasto disminuyó la expresión de genes involucrados en las uniones Fe-S, que intervienen en el transporte de electrones entre fotosistemas (como la ferredoxina) y entre otras cosas, responden a la sacarosa (Figura 4.10). Los genes PsbP que disminuyeron su expresión son constituyentes de los fotosistemas y están presentes en más de una categoría funcional, se encuentran involucrados en procesos de oxidorreducción, la membrana fotosintética, el fotosistema II y complejo *oxygen evolving* del fotosistema II. También disminuyó la expresión de genes del complejo antena (Lhcb4, *light harvesting complex*). Continuando con la siguiente fase de la fotosíntesis, se encuentra que disminuyó la expresión de genes que codifican para enzimas involucradas en el ciclo de Calvin, especialmente una subunidad de la rubisco, la FBPasa (fructosa 1,6-bisfosfatasa), fosfoglicerato quinasa y la fosforibuloquinasa (Figura 4.11).

La FBPasa es una enzima altamente relacionada con el metabolismo del carbono, su isoforma cloroplastídica regula la salida de triosas fosfato desde el cloroplasto para la síntesis de sacarosa en el citosol. En situaciones de activa fotosíntesis, que determina un pH alcalino y un aumento de poder reductor en el cloroplasto, la FBPasa se torna muy activa favoreciendo la regeneración del aceptor de CO₂, o bien la síntesis de almidón transitorio en el cloroplasto si la salida de triosas fosfato al citosol está bloqueada por la sobre-



Figura 4.10: DAG (Directed acyclic graph) con los términos correspondientes a componentes celulares de GO que disminuyeron su expresión significativamente en el GS frente al estrés por sequía. El color sombreado de los términos corresponde a su nivel de expresión.



Figura 4.11: Principales reacciones involucradas en la fijación de C en organismos fotosintéticos (KEGG gmx 00710). El color sombreado de los genes corresponde a su nivel de expresión.

carga floemática de sacarosa. Además, la isoforma citosólica, por otra parte, participa en uno de los pasos claves de la guconeogénesis. En el GS frente a sequía disminuyó la expresión de la isoforma cloroplastídica. Por lo que podría intuirse que disminuyó la eficiencia de la fotosíntesis, principalmente por la falta de regeneración del aceptor. En las vías metabólicas del almidón y la sacarosa se puede ver como disminuyó notablemente la expresión de la β -amilasa, una de las primeras enzimas involucradas en la degradación del almidón Figura 4.12. Si disminuyó la degradación de almidón es lógico pensar que también disminuyeron los siguientes pasos en la obtención de energía a partir de la glucosa. Varias de las enzimas de la glucólisis resultaron afectadas. La fructosa bisfosfato aldolasa actúa rompiendo una hexosa en dos triosas fosfato y fue una de las enzimas que disminuyó su expresión. También resultó afectada la expresión de la gliceral-dehído 3-fosfato deshidrogenasa, enzima que cataliza la siguiente reacción de la glucólisis. Si se continúa con el proceso de obtención de energía a través de la degradación de glucosa, en el ciclo de Krebs disminuyó la expresión de las enzimas encargada de catalizar reacciones anapleróticas. Más adelante aún, encontramos que disminuyó la expresión del gen que codifica para la subunidad alfa de la ATP sintasa.

Siguiendo con el metabolismo de hidratos de carbono, es evidente que la tasa de crecimiento del GS estaba disminuyendo ante el estrés impuesto. Para que haya expansión celular es necesario que se relaje la pared celular y gane extensibilidad para poder separar las microfibrillas y permitir la inserción de polímeros recientemente sintetizados. En el GS disminuyó la expresión de varias familias enzimáticas involucradas en diferentes pasos de este proceso. Las dos principales categorías enriquecidas subrepresentadas correspondieron a enzimas con actividad hidrolasa que actúan en uniones glicosiladas (Figura 4.9). Los 30 genes que se encuentran en estas categorías actúan en la región extracelular, el apoplasto o la pared celular (Figura 4.10). La gran mayoría codifica para xiloglucano endo- β -transglicosilasas (XET) y expansinas, que conforman las dos principales familias enzimáticas encargadas de la regulación del entramado entre celulosa y xiloglucanos de la pared celular (Figura 4.12). También disminuyó la expresión de varias glicosilhidrolasas que modifican los polisacáridos que componen la pared. Además, disminuyó la expresión de una endoglucanasa involucrada en la degradación de la celulosa y un gen que codifica para su síntesis en



Figura 4.12: DAG (Directed acyclic graph) con los términos correspondientes a funciones moleculares de GO que disminuyeron su expresión significativamente en el GS frente al estrés por sequía. El color sombreado de los términos corresponde a su nivel de expresión.



Figura 4.13: Principales reacciones involucradas en el metabolismo del N (KEGG gmx 00910). El color del sombreado de los genes corresponde a su nivel de expresión.

tricomas y tallos. Entre los genes que más disminuyeron su expresión se encentraron genes involucrados en la morfogénesis celular y foliar; y genes de respuesta a auxinas. También disminuyeron glicoproteínas ricas en hidroxi-prolinas (componentes de la matriz extracelular), poligalacturonasas, y enzimas que modifican la pectina (pectin-esterasas y pectato-liasas). Cabe aclarar que en las paredes celulares de tipo I (presentes en las dicotiledóneas) el crecimiento celular está asociado a cambios en la matriz de pectinas. Todo esto crea un panorama desfavorable para la expansión celular y el crecimiento en el GS.

Otros procesos metabólicos fundamentales que se vieron afectados en el GS fueron la asimilación del N y del S, el transporte pasivo a través de membranas y la síntesis de nucleótidos. En la reducción asimilatoria del N, se encontró que disminuyó la expresión de la enzima nitrato reductasa, la isoforma cloroplastídica de la nitrito reductasa depndiente de ferredoxina, la asparagina sintetasa y la glutamato sintasa (Figura 4.13). Además, diminuyó la expresión de transportadores de aminoácidos transmembrana y de un receptor tipo quinasa de la nodulación. Por otro lado, las dos principales enzimas de la reducción asimilatoria del S disminuyeron su expresión. Una menor disponibilidad de N y S en la planta repercute en su capacidad fotosintética y su crecimiento. Las categorías de transporte transmembrana y actividad de canales corresponde en su mayoría a acuaporinas (Figura 4.12).

Así como disminuyó la actividad de procesos fundamentales, también disminuyo la expresión de genes de respuesta a estrés que podrían haber ayudado a combatir los efectos negativos de la sequía. Por ejemplo, disminuyó la expresión de genes de respuesta a estrés salino, frío, falta de agua, ácido abscísico y daños. Algunos de estos eran poligalacturonasas, genes de síntesis de quitina, inositol o cupinas. La categoría enriquecida actividad de canales correspondió enteramente a aquaporinas que disminuyeron su expresión

(Figura 4.9). El mantenimiento del estado redox es fundamental ante un estrés para evitar la producción de especies reactivas de oxígeno (ROS) que dañan a las células. En el GS, en lugar de aumentar, disminuyeron numerosos genes involucrados en este aspecto, por ejemplo: ferredoxinas, peroxidasas, tiorredoxinas (h2 y m3) y peroxirredoxinas (Prx Q). También disminuyó la expresión de una piridin-nucleótido disulfato oxidorreductasa que está involucrada en el mantenimiento redox de la célula y responde a distintos hidratos de carbono (como sacarosa o glucosa). Una de las consecuencias del estrés es el mal plegamiento de distintas proteínas, que deben ser eliminadas o plegadas correctamente; y al igual que para los genes involucrados en el estrés oxidativo, disminuyó la expresión de varios de estos en el GS. Específicamente, se vio reducida la expresión de distintas HSP (*heat shock proteins* 20, 70, 90) y CSP (*cold shock proteins*), chaperonas, calmodulina y receptores extracelulares que censan Ca²⁺, distintas ubiquitin ligasas, una proteína con actividad dehidrina (FK506 binding protein) que ayuda a tolerar el estrés hídrico.

Varias categorías enriquecidas guardaron relación con la síntesis de nucleótidos (y sus constituyentes) y en conjunto agruparon a 30 genes. Estos están involucrados en el mantenimiento de la homeostasis celular de cationes, metabolismo del H_2O_2 , reorganización de membranas tilacoidales, del cloroplasto y el aparato de Golgi. Además, intervienen en la síntesis de compuestos del metabolismo secundario (isoprenos, glucosinolatos, vía independiente del mevalónico), ácido saliscílico, y distintos hidratos de carbono (almidón, maltosa). Estos genes responden a distintos estímulos: hidratos de carbono (sacarosa, glucosa), frío.

3.3. Comparaciones entre genotipos contrastantes

3.3.1. Fitohormonas

Las células de las plantas detectan una gran variedad de señales que se originan en el ambiente que las rodean, o dentro de la misma planta, y usan esta información para regular su comportamiento. En todos estos casos, las respuestas de las células dependen de la transducción de señales, el proceso a partir del cuál se transfiere la señal desde el sitio en el que se percibió hasta el sitio donde se va a efectuar la respuesta. Las fitohormonas y los factores de transcripción juegan un rol fundamental en la respuesta ante distintos tipos de estrés ya que cada uno activa una vía de transducción de señales propia e induce la expresión de numerosos genes asociados.

Las auxinas son fitohormonas involucradas en el desarrollo de las plantas, regulan procesos como el tropismo, el desarrollo de los meristemas y la regeneración de tejidos luego de un daño. Sin embargo, los procesos mediados por auxinas mejor estudiados son la iniciación foliar y el entramado del tejido vascular. A diferencia del GS que reguló hacia abajo cuatro genes involucrados en la vía de transducción de señales mediados por auxinas, el GT solo presentó uno disminuido (Figura 4.14), que correspondió a la familia de proteínas SAUR (*small auxin up RNAs*) (Figura 4.15). Esta mayor depresión en la vía de las auxinas podría significar una menor elongación celular y crecimiento de las plantas pertenecientes al GS.

Las giberelinas regulan procesos como la germinación, la elongación del tallo y la iniciación floral, entre otros. Las proteínas GID2 (*gibberelin insensitive dwarf 2*) promueven la ubiquitinación y degradación de las proteínas DELLAs, las cuales reprimen la señalización mediada por giberelinas. En el genotipo tolerante aumenta la expresión de las GID2, por lo que debería haber menos represión del crecimiento mediado por las giberelinas (Figura 4.15). Por el contrario, el GS no presentó ningún gen de esta vía diferencialmente expresado (Figura 4.14).

Otro grupo de fitohormonas de interés son las citocininas. Estas son moléculas señalizadoras que actúan en diversos procesos como respuesta a estrés, división celular y la iniciación de brotes. La transducción de la señal se da a través de la fosforilación de un sistema de dos proteínas: una histidina quinasa y un regulador de la respuesta (RR). El receptor de citocininas ubicado en la membrana (CRE1, *cytokinin response*) es una histidina quinasa con un dominio extracelular al cual se unen las citocininas induciendo su autofosforilación. Luego, el grupo fosfato es transmitido hacia los AHPs a través del dominio receptor, quienes una vez fosforilados, se mueven hacia el núcleo donde transfieren el grupo fosfato a los ARRs (reguladores de respuesta). Estos factores de transcripción activan la expresión de genes de respuesta a citocininas amplificando la respuesta. Por el contrario, los ARRs de tipo A activados por los AHPs modulan negativamente la respuesta a citocininas. Esta vía de señalización, y los genes que responden a ella, se vio favorecida en



Figura 4.14: Transducción de señales mediadas por fitohormonas (KEGG gmx 04075).El color sombreado de los genes corresponde a su nivel de expresión para el GS frente a la sequía.



Figura 4.15: Transducción de señales mediadas por fitohormonas (KEGG gmx_04075). El color sombreado de los genes corresponde a su nivel de expresión para el GT frente a la sequía.

el GT ante una condición de sequía ya que aumentó la expresión de los receptores de membrana, de los reguladores de la respuesta, de los factores de transcripción del tipo B; y disminuyó la expresión de los factores de transcripción del tipo A que deprimen la respuesta a citocininas (Figura 4.15). En el GS solo se vio levemente aumentada la expresión de los ARRs de tipo B.

El ABA es conocido como la fitohormona del estrés dado que media la respuesta a distintos tipos de estreses abióticos como la sequía, salinidad, frío, calor, etc. Ante una sequía, esta fitohormona activa una cascada de señalización que termina por cerrar los estomas y activar la transcripción de numerosos genes. En primer lugar, el ABA se une a los receptores PYR/PYL formando un complejo que encaja en el sitio activo de las enzimas PP2C. Estas enzimas, a su vez, en ausencia de ABA inhiben a las SnRK2 al desfosforilarlas. Sin embargo, en presencia de ABA se liberan las SnRK2 y estas desencadenan dos procesos: regulan la actividad de canales iónicos en las células oclusivas y activan mediante fosforilaciones a los factores de transcripción ABF (ABA binding factor). La activación de los canales iónicos lleva a la despolarización de la membrana provocando una salida de iones de K⁺, una disminución en la turgencia y el consecuente cierre de los estomas. Los factores de transcripción ABF/AREB (ABA Responsive Element Binding protein) corresponden a la familia bZIP y se unen a los sitios de respuesta ABRE (ABA responsive element) de distintos genes de respuesta activando positivamente su expresión. Estos genes que contienen el elemento ABRE no solo son activados por los factores de transcripción del tipo AREB/ABF, también son activados por los AP2/ERF, MYB, NAC, NF-Y y WRKYs; todas familias que aumentaron su expresión significativamente en el GT (Figura 4.18). Para este genotipo, además, los receptores PYL/PYR fueron unos de los genes que más disminuyeron su expresión, alcanzando valores de hasta -2,8 Log2 (Fold Change). Al disminuir estos receptores, aumentó la expresión de las fosfatasas que inhiben (PP2C) y los factores de transcripción ABF; por lo que se encuentran numerosos genes de respuesta a ABA que aumentaron su expresión y cierre estomático. En lo que respecta a la síntesis del ABA, disminuyó la expresión de la NCED que es considerada una de las enzimas más importantes en su síntesis. El GS también varió la expresión de los receptores PYL/PYR y de las fosfatasas PP2C aunque en menor medida que el GT, y a diferencia de este no aumentó la expresión de los factores de transcripción ABF (Figura 4.14).

En el GS la vía de transducción de señales mediada por el etileno parecería estar favorecida por el aumento de la expresión del primer receptor (Figura 4.14). Sin embargo, en el GT el único gen de esta vía que aumentó su expresión fue el EBF1/2 (Figura 4.15). Estos genes reprimen la acción del etileno al generar un complejo de ubiquitin ligasas que degrada el factor de transcripción EIN3. Otro compuesto que compromete la estabilidad del EIN3 promoviendo su degradación es la glucosa, que guarda relación con lo anteriormente visto para el GT que parecería aumentar sus niveles al disminuir la expresión de la hexoquinasa. Por último, ambos genotipos respondieron igual en la vía del ácido jasmónico, aumentando la expresión del factor de transcripción MYC2. Esta familia de factores de transripción, junto con otras (MYB, NAC, HD ZIP, HD, HB, AP2), están involucrados en la respuesta del ABA bajo condiciones de estrés osmótico.

3.3.2. Factores de transcripción

Los factores de transcripción juegan un rol fundamental en la respuesta frente a distintos estímulos, ya que actúan activando o reprimiendo la expresión de distintos genes. Frente al estrés por sequía en los genotipos contrastantes 103 factores de transcripción cambiaron su expresión, de los cuales 53 fueron exclusivos para el GT, 31 para el GS y 19 fueron compartidos (Figura 4.16a). Ningún gen correspondiente a un factor de transcripción tuvo una expresión inversa en los genotipos contrastantes. Es decir, no hubo ningún gen que haya disminuido su expresión en un genotipo y la haya aumentado en el otro, y viceversa. Si se encontraron distintos niveles de expresión en los 19 genes coexpresados, pero siempre en el mismo sentido (Figura 4.16b).

MYB - Tanto los factores de transcripción MYB, como los relacionados con MYB (MYB *related*) y G2*like* tienen sitios de unión a MYB, una secuencia conservada de aminoácidos que puede repetirse de dos a cuatro veces en la misma proteína. Estas familias de factores de transcripción están involucradas en diversos procesos de las plantas: metabolismo secundario, transducción de señales mediadas por fitohormonas, forma celular, desarrollo de órganos y respuesta a estrés ambiental. La mayoría de ellos regula positivamente la transcripción. De aquellos que aumentaron su expresión en el GS frente a la sequía, uno está relacionado con la vía de señalización del estrés osmótico y es inducido por ABA. Por otro lado, el MYB que alcanzó la mayor expresión relativa (Glyma.19G219000.1) está directamente involucrado en la respuesta al estrés





(a) Diagrama de Venn para los factores de transcripción diferencialmente expresados en los genotipos contrastantes.

(b) Expresión diferencial de los distintos factores de transcripción para los genotipos contrastantes. Cada punto corresponde a un gen particular y las líneas unen el mismo gen a través de los genotipos.

Figura 4.16: Análisis preliminar de la expresión diferencial de factores de transcripción para los genotipos contrastantes.



Figura 4.17: Familias de factores de transcripción que cambiaron significativamente su expresión en el GS bajo condiciones de sequía.

abiótico, regula positivamente la expresión de genes que actúan en la síntesis y el transporte de la prolina. Además, activa la transcripción de genes involucrados en la detoxificación de ROS y genes asociados con el estrés (LEA3, DREB2A, RAB16A). De los G2-like que disminuyeron la expresión, uno esta asociado con la activación de la transcripción de genes de la absorción y la asimilación del N, la represión de la respuesta mediada por GA y la disminución de la disponibilidad de GA bioactivo. Esto guarda relación con la consecuente disminución de una de las enzimas que cataliza la última reacción en la bioactivación de las giberelinas y otras enzimas involucradas en su síntesis. De estas familias de factores de transcipción, en el GT también aumentó la expresión de varios genes de respuesta a distintas fitohormonas (ABA, GA, etileno, JA, auxinas) y a distintos tipos de estrés (osmótico, salino, hídrico, frío). Dentro de los más relacionados con la sequía, se destacó el Glyma.12G117700.1 de la familia G2-like que interviene en la cascada de señalización de MAPK, en la organización del cloroplasto, síntesis de clorofila y plástidos, y el metabolismo del H_2O_2 .

WRKY - La familia de factores de trascripción WRKY es una de las más numerosas que existe en las plantas, participan en diversos procesos fisiológicos y del desarrollo: desarrollo y dormición de semillas, germinación, senescencia, respuesta inmune, defensa contra patógenos e insectos. Recientemente se ha descubierto que además, participan en la transducción de señales hormonales y en la respuesta a estreses


Figura 4.18: Familias de factores de transcripción que cambiaron significativamente su expresión en el *GT* bajo condiciones de sequía.

abióticos. En el GS se sobreexpresaron tres factores de transcripción WRKY distintos (GmWRKY 28, 30 y 34), uno de los cuales presentó una variante de splicing distinta. En el GT aumentó la expresión de un solo factor de transcripción de esta familia y disminuyó la expresión de otros cuatro, dos de los cuales regulan negativamente la muerte celular programada.

NF-YB, HD-ZIP, Dof - La región conservada de los factores de transcripción NF-YB tiene homología estructural y de secuencia con el plegamiento H2B de las histonas. En el GS aumentó la expresión de uno de estos factores de transcripción en respuesta a la sequía, mientras que en el GT aumentó la expresión de dos. En ambos genotipos aumentó la expresión de factores de transcripción de la familia HD-ZIP, que suelen actuar como activadores de la transcripción promoviéndola. Estos están involucrados en el catabolismo de las clorofilas, la morfogenesis foliar y son de respuesta a citocininas. Otros factores de transcripción involucrados en la respuesta mediada por citocininas son los ARRR-B (Figura 4.14), que activan los reguladores de tipo A y terminan por promover la longevidad foliar. En el GS aumentó la expresión de uno de estos genes, mientras que en el GT ninguno. Asimismo, de la familia Dof, disminuyó la expresión de un gen asociado a la histogénesis del xilema y floema en el GS y ninguno en el GT.

NAC - Otra de las familias de factores de transcripción de respuesta a estrés abiótico más importante es la NAC que actúa modulando la expresión de genes ERD1 (*early response to dehydration 1*). De los cuatro que aumentaron su expresión en el GS, dos actúan río abajo de los factores MYC2 en la respuesta contra distintos patógenos mediada por JA y uno activa la transcripción de dos genes de respuesta a auxinas. Por el contrario, en el GT disminuyó la expresión de un NAC involucrado en la senescencia foliar y aumentó la expresión de tres NAC que regulan procesos que podrían ayudar en su tolerancia a la sequía. Estos son genes involucrados en la vía de señalización del ABA, de respuesta a falta de agua e involucrados en el desarrollo de tallos, xilema y la regulación de la muerte celular programada.

bHLH - De la familia bHLH la mayoría de los factores de transcripción que aumentaron su expresión fueron del tipo MYC, que promueven el crecimiento en respuesta a ABA, JA y activan las respuestas defensivas de las plantas en respuesta a patógenos, entre otros. Para esta familia no se encontraron grandes diferencias entre genotipos.

ERF - La familia de factores de transcripción ERF se divide en cuatro subfamilias, de las cuales ERF y DREB son las más estudiadas por su rol en la respuesta al estrés abiótico. En el GS aumentó la expresión de 8 de estos genes, dos de los cuales activan directamente la transcripción de genes relacionados con el estrés abiótico y regulan el cierre estomático en respuesta de la sequía; uno de estos lo hace al unirse a la secuencia DRE de los genes que activa. Los demás factores de transcripción que aumentaron su expresión están involucrados en las vías de transducción de señales activadas por estrés, algunas mediadas por etileno y otros se unen a regiones relacionados con la patogénesis. En el GT aumentó la expresión de cinco ERF,

tres de los cuales compartió con el GS. A difrencia de este, en donde solo disminuyó la expresión de uno de estos genes, en el GT disminuyó la expresión de cinco, la mayoría de los cuales actúan en la vía de señalización del ABA y etileno.

bZIP - La bZIP es una familia de factores de transcripción que además de regular el crecimiento y desarrollo de las plantas tiene un rol fundamental en la respuesta al estrés abiótico. En el GS disminuyeron dos genes bZIP que activan la transcripción de genes relacionados con el estrés hipoosmolar, como peroxidasas, síntesis de mioinositol, rafinosa y trehalosa. Además, median la acetilación de histonas para activar la transcripción inducida por auxinas. Los dos ejemplares de esta familia que aumentaron su expresión en respuesta a la sequía en el GS comparten esa relación con el estrés hipoosmolar, la inducción de peroxidasas, pero también inducen la síntesis de proteínas LEA. En el GT los factores de transcripción bZIP que aumentaron su expresión, en su mayoría, intervienen en la vía de señalización del ABA y del etileno, y responden a distintos estímulos (carbohidratos, superóxido, falta de agua, estrés salino, etc), de los cuáles ya se profundizó en la sección 3.3.1 por ser ABF/ AREB.

3.3.3. Agrupamiento de categorías: una visión generalizada

Se realizó un segundo enriquecimiento con la herramienta AgriGO para luego obtener una visión más generalizada de las categorías enriquecidas en ambos genotipos frente a la sequía con la herramienta REViGO, que permite agruparlas según similaridad semántica. De esta manera, se eliminan categorías redundantes y se logra que aquellas similares queden juntas. En resumen, para el GT se observó un enriquecimiento de categorías propias de la respuesta ante la sequía asociadas con la tolerancia (Figura 4.19). Por ejemplo, se destacó la síntesis de carbohidratos y alcoholes; muchos de los cuales pueden ser acumulados como osmolitos compatibles en el ajuste osmótico y tienen actividad antioxidante (alcoholes: mioinositol, manitol, pinitol; carbohidratos: rafinosa, glucosa, fructosa). También se destacó la síntesis de esteroides. Otra categoría útil en la respuesta a la sequía que se vio enriquecida en el GT fue la señalización y la transducción de señales que involucran la autofosforilación de histidinas quinasas con la posterior transferencia del grupo fosfato a un aspartato que luego fosforila a proteínas reguladoras. Estos tipos de sistemas de dos componentes son altamente estudiados por ser considerados promisorios para el desarrollo de plantas de soja tolerantes a la sequía. Los procesos de óxido reducción resultaron enriquecidos, lo que puede haber contribuido a el mantenimiento del homeostasis redox en las células. Así como se vieron enriquecidas categorías propias de una respuesta positiva hacia la sequía, también se vieron enriquecidas categorías que denotan un funcionamiento celular normal, como la fotosíntesis, el transporte de amonio y la síntesis de coenzimas. Por el contrario, para el GS aumentaron procesos de regulación de la calidad biológica, señalización en general y metabolismo del RNA (Figura 4.20). Estos resultados demuestran que, a diferencia del GT, el GS no parecería tener una respuesta estratégica frente a la sequía.

Las categorías enriquecidas encontradas con AgriGO que disminuyeron su expresión guardaron relación con aquellas encontradas con GO. Para el GT disminuyeron categorías relacionadas a la traducción, plegamiento de proteínas y biogénesis de ribosomas. Para el GS disminuyeron categorías asociadas a la síntesis de carbohidratos y nucleótidos, las cuales podrían ayudar en la tolerancia a la sequía mediante la síntesis de compuestos osmoprotectores. En ambos genotipos disminuyeron categorías asociadas al metabolismo primario, el cual está directamente asociado al rendimiento.

4. Discusión

Los mayores esfuerzos por comprender la respuesta a la sequía de las plantas se han hecho en *Arabidopsis*, mientras que los estudios en soja y otros cultivos de interés siguen siendo pocos en comparación. Además, muchos de los autores que han realizado estudios transcriptómicos en soja han impuesto tratamientos que podrían considerarse aproximaciones de laboratorio a la sequía observada en el campo, y tal como se ha visto en *Arabidopsis* los genes que responden a este tipo de estrés no son los mismos que responden ante una sequía proresiva (Huang y col., 2008). Por ejemplo, en el experimento llevado a cabo por Stolf-Moreira et al., (Stolf-Moreira y col., 2011) el tratamiento de sequía consistió en crecer las plantas en una solución hidropónica y retirarlas de la misma por un lapso de 100 minutos; esto más que una sequía podría consi-



Figura 4.19: Gráfico de dispersión de las categorías enriquecidas resultantes del análisis de AgriGO y REViGO para el GT en situación sequía con respecto a su control. Los ejes representan espacios semánticos que agrupan las categorías similares.

derarse una desecación. Otro caso similar es el de Hua (Hua y col., 2018) y el de Shin (Shin y col., 2015), en donde las plantas utilizadas para el estudio transcriptómico son retiradas del medio de crecimiento y se dejan desecar las raíces. La metodología utilizada en este trabajo es una aproximación que permite imponer un estrés más real sin descuidar las condiciones experimentales controladas necesarias para los enfoques moleculares. Además, a diferencia de muchos de los análisis transcriptómicos de la bibliografía que se han realizado a través de microarreglos (Hua y col., 2018) una técnica anterior, menos precisa y robusta que la secuenciación (Prince y col., 2015); en este trabajo se realizó un RNAseq.

Los genes promisorios en el mejoramiento de la soja frente al estrés por sequía pueden clasificarse en dos grandes grupos: los genes funcionales y los genes reguladores. Los primeros son aquellos que afectan directamente la tolerancia, como genes de síntesis de osmolitos, acuaporinas, HSP, LEAs, proteasas, transportadores iónicos, etc. El segundo grupo corresponde a genes involucrados en la regulación o la señalización de la respuesta frente al estrés, como factores de transcripción, sistemas de dos componentes, kinasas o moléculas de señalización (Thao y Tran, 2012; Fritsche-Neto y Borém, 2012). En este trabajo, el GT presentó tanto genes funcionales como reguladores aumentados, mientras que en el GS varios de estos disminuyeron su expresión. La mayor disponibilidad de genes específicos para mitigar el estrés impuesto por la sequía, en conjunto con el normal funcionamiento del metabolismo primario, explicarían su tolerancia.

Por ejemplo, dentro de los genes funcionales que aumentaron su expresión en el GT se encuentran aquellos involucrados en la acumulación de solutos compatibles. La acumulación de estos compuestos orgánicos de bajo peso molecular no solo tiene un efecto osmótico, sino también antioxidante. Por ejemplo, la glicina beatina previene la desestablización del fotosistema II y la inactivación de la rubisco. El mioinositol, la prolina y la rafinosa pueden detoxificar los radicales hidroxilos que provocan daños a macromoléculas, y en el GT se encontraron aumentadas enzimas involucradas en su síntesis en respuesta a la sequía (Buchanan, Gruissem y Jones, 2015; Hossain y col., 2014; Fuganti-Pagliarini y col., 2017). Además, en el GT se vio aumentada la expresión de tres genes de la familia HSP20 que responden a distintos estreses bióticos y abióticos. Estas chaperonas moleculares evitan la desnaturalización de otras proteínas y ayudan a mantener la conformación nativa de cadenas polipéptidas crecientes en una manera independiente de ATP (Lopes-



Figura 4.20: Gráfico de dispersión de las categorías enriquecidas resultantes del análisis de AgriGO y REViGO para el GS en situación sequía con respecto a su control. Los ejes representan espacios semánticos que agrupan las categorías similares.

Caitar y col., 2013). En el GS no se encontraron HSP, de esta ni otra familia, que hayan aumentado su expresión. Otro gen funcional que aumentó su expresión en el GT fue la dehidrina DHN-13 kDa (Glyma.19G114700). Esta proteína, y otras de la misma familia, han sido seleccionadas como indicadores para monitorear la respuesta temprana a la sequía en soja (Gallino y col., 2018; Fuganti-Pagliarini y col., 2017). Los autores también utilizan como indicadores de tolerancia una fototropina, una histona 2A y un factor de transcripción de la familia G2-like; los cuales también han resultado sobreexpresados en el GT pero no en el GS. Otra categoría de genes funcionales involucrados en la tolerancia a la sequía son las acuaporinas. Estos canales transportadores, por los cuáles fluye el agua en las células, ayudan en la tolerancia al estrés hídrico a través de dos estrategias distintas: minimizar la pérdida de agua frente una escasez hídrica en el suelo, o ante una excesiva demanda del ambiente (Shekoofa y Sinclair, 2018; Fuganti-Pagliarini y col., 2017). En el GS una de las categorías funcionales que más resultó disminuida fue la actividad de canales correspondiente a acuaporinas. Estos resultados evidencian como el GT activa diversos mecanismos para contrarrestar o minimizar los daños causados por la falta de agua, mientras que en el GS estos disminuyen o no cambian su expresión, aumentando su vulnerabilidad ante la sequía.

La actividad oxidorreductasa es fundamental en el control del estrés oxidativo provocado por la sequía. Ambos genotipos presentaron esta categoría sobre-representada, en donde se encontraron principalmente lipoxigenasas. Estas enzimas oxidan ácidos grasos insaturados en la vía de las oxilipinas, produciendo ácido jasmónico, volátiles de hoja verde y otros metabolitos involucrados en la respuesta a estreses bióticos y abióticos. El aumento de su expresión en respuesta a la sequía se ha obserbado en plantas de soja (Faustino y col., 2021), trigo (Xie y col., 2020), pimiento (Lim y col., 2015) y mijo (Zhang y col., 2021); confiriendo una mayor tolerancia al estrés. A diferencia del GS, el GT presentó una mayor cantidad de lipoxigenasas con expresión aumentada y, además, expresión de inhibidores de proteasas. Se ha encontrado que el aumento de estos inhibidores en respuesta a una sequía moderada afecta la supervivencia de insectos masticadores que se alimentan de estas plantas (Faustino y col., 2021). Este rasgo resulta de particular interés ya que podría indicar una mayor tolerancia del GT contra la herbivoría además de la sequía.

Dentro de los genes reguladores que aumentaron su expresión en el GT se pueden encontrar aquellos relacio-

nados con los sistemas de señalización de dos componentes, factores de transcripción, genes relacionados a fitohormonas y otros genes de interés. Se cree que el tipo de señalización de dos componentes en respuesta a la sequía podría utilizarse ventajosamente en el mejoramiento genético hacia la obtención de plantas tolerantes (Thao y Tran, 2012). En cuanto a las vías activadas por fitohoromonas que aumentaron su expresión en el GT en mayor medida que en el GS pueden citarse aquellas mediadas por ABA, auxinas y citocininas. En el GT se encontró aumentada la expresión de factores de transcripción, denominados factores de transcripción ABF, de la familia bZIP que se unen específicamente a los sitios ABRE (ABA Binding Responsive Element) de los genes RD29B (Response to Dessication 29B), del promotor del gen Dc3 (proteína LEA) y de la chalcona sintasa. Hay evidencia que sugiere que la proteína codificada por el gen RD29B protege a las plantas frente a distintos estreses abióticos, y que podría utilizarse para la producción de plantas tolerantes en regiones áridas y semiáridas (Msanne y col., 2011). También se lo ha utilizado como un gen marcador de respuesta a ABA inducido por sequía (Lim y col., 2015); y descripto como un gen con 'memoria' ya que ayuda a las plantas en el priming contra la sequía, aumentando más su expresión ante repeticiones del mismo estrés, o el contacto con bacterias de la rizosfera (Liu, Sikora y Park, 2020). La chalcona sintasa es una enzima clave en la síntesis de flavonoides, se ha encontrado que aumenta su expresión en respuesta a la sequía en plantas de trigo (Ma y col., 2014) y que su sobreexpresión confiere tolerancia a la sequía en tabaco (Hu y col., 2020). La sobreexpresión de estos factores de transcripción ABF confiere mayor tolerancia a la sequía y sensibilidad al ABA en plantas de Arabidopsis (Singh y Laxmi, 2015), al igual que lo observado en este trabajo para el GT. Por otro lado, en el GT aumentó la vía mediada por las citocinicas y se ha demostrado que un aumento en su síntesis resulta en una aumentada concentración de las enzimas involucradas en el ciclo ascorbato-glutatión, produciendo una marcada disminución en la senescencia provocada por la sequía y una mejor tolerancia ante este estrés (Aroca, 2012). Además, se ha encontrado que la biosíntesis de esteroides disminuye en respuesta a la sequía en Arabidopsis (Huang y col., 2008); sin embargo, en el GT fue uno de las vías que se vio enriquecida. Se ha demostrado que los esteroides y brasinesteroides juegan un rol importante en varios procesos de las plantas, entre ellos la tolerancia a la sequía, haciéndolos una estrategia promisoria para lograr genotipos con un mejor comportamiento ante distintos tipos de estrés (Cile Vriet, Russinova y Reuzeau, 2012; Tůmová y col., 2018). Una de las vías que diminuyó en el GS en mayor medida que en el GT fue la mediada por auxinas, y se ha demostrado en Arabidopsis esta fitohormona actúa positivamente en la respuesta a la seguía a través de la regulación de la arguitectura radical, la expresión de genes de respuesta a ABA, el metabolismo de ROS y la homeostasis celular (Shi y col., 2014). En conclusión, a diferencia del GS, el GT presentó enriquecidas distintas vías de señalización que le permiten activar mecanismos para tolerar el estrés de manera eficiente.

Al igual que lo obtenido para el GT en este trabajo, se ha visto en *Arabidopsis* que simulando un estrés por sequía severo o una sequía progresiva, algunos de los términos que disminuyeron su representación fueron aquellos relacionados con los ribosomas, sus componentes estructurales y su biosíntesis (Alqurashi y col., 2018; Huang y col., 2008). Esta disminución también se ha encontrado en soja sometida a inundación y sequía (Wang y col., 2017) y en plantas de arroz, implicando que hay una disminución general de la síntesis de proteínas frente al estrés (Hamzelou y col., 2020). Sin embargo, a diferencia de lo encontrado por Harb en donde la NCED es una de las enzimas que aumenta su expresión frente al estrés por sequía; en el GT esta fue una de las enzimas que más disminuyó su expresión (Harb y col., 2010). La sobreexpresión de esta enzima confiere una elevada resistencia a la sequía en plantas de tabaco transgénicas, debido a la acumulación de ABA y la posterior inducción de cierre estomático y expresión de genes de respuesta (Bao y col., 2016). Aún teniendo la expresión de esta enzima disminuida, el GT presentó sobre-representada la vía del ABA, cierre estomático e inducción de genes; por lo que podría pensarse en una mayor sensibilidad a la fitohormona.

La actividad hidrolasa, específicamente la actividad de enzimas β -amilasas, fue una de las categorías que más disminuyó su representación en el GS frente al estrés por sequía. Estas enzimas degradan el almidón produciendo maltosa, y se cree que su actividad es clave en la removilización de carbohidratos para proveer energía en hojas de plántulas de soja frente a estreses como sequía o inundación (Wang y col., 2017). Otra familia de genes que se vio deprimida en este genotipo fue la PbsP, que codifican para componentes estructurales de los fotosistemas. Se ha informado que estos genes PbsP están también involucrados en la señalización por MAPK (*mytogen activated protein kinases*), respuesta a diferentes tipos de estrés bióticos (hongos, bacterias) y abióticos (alta intensidad lumínica, frío), respuesta a sacarosa y ácido salicílico. Por otro lado, en la bibliografía se citan genes específicos que ayudan a tolerar el estrés que se encontraron deprimidos en el GS. Uno de estos fue una C3H4- *ring finger* con actividad E3 ligasa, la cual se asoció positivamente en *Arabidopsis* con el escape a la sequía mediado por ABA (Yang y col., 2016). Otro ejemplo

es una enzima (GDP-manosa 3,5-epimerasa) involucrada en la síntesis de ácido ascórbico (Ma y col., 2014), y una enzima involucrada en la síntesis de timina (vitamina B1) que ayuda a tolerar daños en el DNA. Las enzimas del tipo UDP-glucosil transferasas están involucradas en la síntesis de distintos compuestos, como flavonoides, fenilpropanoides, terpenos, esteroides y fitohormonas. Específicamente, estas enzimas transfieren grupos azúcar desde moléculas dadoras a aceptores específicos. Se ha visto que en *Arabidopsis* la UGT71b6 glicosila al ABA inactivándolo, por lo que los genotipos que la sobreexpresan presentan una disminuida tolerancia a la sequía (Liu y col., 2015). En el GS aumentaron varios genes pertenecientes a esta familia, lo que puede haber sido uno de los causantes de su baja tolerancia al estrés.

5. Conclusiones

La tolerancia a la sequía es un atributo gobernado por muchos genes, es por eso que la transcriptómica representa una herramienta valiosa para poder evaluar la respuesta de una manera integral. El RNAseq ayudó a comprender que la tolerancia a la sequía exhibida por el GT (Capítulo 3) podría estar ligada a la síntesis de compuestos osmoprotectores (carbohidratos, alcoholes, esteroides), a una mayor actividad antioxidante y una eficiente señalización; ya sea mediada por sistemas de dos componentes, factores de transcripción o fitohormonas. Por el contrario, el GS no exhibió una respuesta clara, no se destacaron procesos específicos de respuesta al estrés, lo que se condice con su comportamiento evaluado en el capítulo 3. Estos resultados podrían ayudar a entender los mecanismos detrás de la tolerancia exhibida, sin embargo, para poder utilizarlos en un programa de mejoramiento se deben seguir profundizando. Un siguiente paso en ese sentido sería el estudio de polimorfismos, presentes en estos genes de interés, ya sea en el genoma o en el transcriptoma. Este polimorfismo de secuencias podría explicar el diferente comportamiento exhibido por los genotipos estudiados.

Capítulo 5

Diseño de marcadores moleculares

1. Introducción

La selección asistida por marcadores moleculares representa una valiosa herramienta en los distintos programas de mejoramiento de cultivos. Esta permite la introgresión de rasgos de interés de manera rápida y eficiente ya que se minimizan los tiempos al poder evaluar los genotipos en estadíos tempranos y se utiliza una menor proporción de tierra destinada a ensayos a campo. Esta herramienta es ampliamente utilizada para caracteres gobernados por genes principales, como la resistencia a herbicidas o enfermedades. La tolerancia a sequía, sin embargo, es un rasgo complejo gobernado por muchos genes, por lo que resulta lógico pensar en una batería de marcadores moleculares a la hora de utilizarlos para la selección de genotipos en un programa de mejoramiento. Por ejemplo, para la generación de líneas de arroz tolerantes a sequía en el trabajo publicado por Dwivedi utilizaron 97 marcadores del tipo SSR (Dwivedi y col., 2021), mientras que en un trabajo publicado para mijo perlado utilizan 32 marcadores (Rani, Taunk y Yaday, 2021). Afortunadamente, la detección y evaluación de múltiples marcadores hoy en día es sencillamente abordada por las plataformas de secuenciación y genotipado de alto rendimiento, o high throughput, en conjunto con técnicas bioinformáticas, siendo parte de los programas de mejoramiento modernos (Abdelrahman y col., 2018). Por ejemplo, la plataforma SNP line (LGC Bioseaerch Technologies) permite analizar miles de SNPs en un día con posibilidad de escalarlo fácilmente. Las tecnologías de NGS (Next Generation Sequencing) han sido aplicadas en el mejoramiento de numerosos cultivos, para explicar el funcionamiento de rasgos genéticos complejos y para la detección de variaciones, ya sean SNPs, INDELs u otras (Polimorfismos de un solo nucleótido; einserciones y deleciones, respectivamente).

Los SNPs son los polimorfismos más abundantes en las células eucariotas, constituyendo una gran fuente de marcadores moleculares (Song y col., 2020). Además de ser abundantes, están bien distribuidos a lo largo del genoma por lo que pueden usarse para el genotipado de manera específica y con una alta reproducibilidad. Estos polimorfismos no solo pueden utilizarse en el mejoramiento como marcadores moleculares, sino también sirven para la construcción de mapas de ligamiento o el análisis de diversidad genética (Villordo-Pineda y col., 2015).

Hay diversas técnicas para la detección de SNPs: *Next Generation Sequencing* (NGS), *Expressed Sequence Tags* (EST), *Kompetitive Allele-Specific PCR* (KASP), etc. En leguminosas, independientemente de la manera en la que se generó la información, esta ha sido utilizada para el descubrimiento de QTLs y genes asociados en la tolerancia a distintos estreses abióticos. Por ejemplo, se han descripto SNPs en *Glycine soja* en el gen ERECTA, que codifica para una serin- treonina quinasa, asociados a la tolerancia a la sequía. Otro ejemplo son las variaciones encontradas para el gen *GmCHX1*, que codifica para un transportador de cationes, asociadas con la tolerancia a la sequía en soja (Abdelrahman y col., 2018).

La búsqueda de SNPs en el transcriptoma podría representar una alternativa ventajosa en comparación a la búsqueda en el genoma. Al analizarse únicamente las regiones exónicas se utilizarían los polimorfismos presentes en la región codificante de los genes, que tienen un mayor impacto que las regiones no codificiantes. Se podrían priorizar los polimorfismos encontrados en genes claves de la respuesta al estrés que resultaron expresados diferencialmente o son claves en vías metabólicas enriquecidas, aumentando el impacto aún más. Esta convergencia entre la búsqueda de QTLs (del inglés, *quantitative trait locus*) y expresión génica se denomina *expression QTL* (Bernardo y Woodbury, 2020). Por otro lado, la búsqueda de polimorfismos en la secuencia transcriptómica es una potencial herramienta en especies con genomas de gran tamaño, como la cebada o el trigo (Tanaka y col., 2019).

En este capítulo se abordará el último objetivo, buscando polimorfismos en la secuencia genómica y transcriptómica para el desarrollo de marcadores moleculares. Por un lado, se buscarán polimorfismos en la secuencia transcriptómica de los genes diferencialmente expresados en los genotipos contrastantes (GT y GS) frente al estrés por sequía. Por otro lado, se observará si los polimorfismos asociados al genotipo tolerante en la secuencia genómica pueden mapearse a genes que resultaron diferencialmente expresados. Una vez que se cuente con los SNPs candidatos se utilizarán para el desarrollo de marcadores moleculares con vistas a su utilización en programas de mejoramiento.

2. Materiales y métodos

2.1. Búsqueda de SNPs en la secuencia genómica

2.1.1. Preparación de las muestras

Para este análisis se utilizaron los genotipos cotrastantes, GT y GS, y el genotipo Williams; que se ha utilizado como ejemplo de genotipo de mal comportamiento en la bibliografía y la caracterización de los genotipos en el capítulo 3. Se tomaron 10 semillas pertenecientes a cada genotipo a las cuáles se les extrajo con un bisturí una fina lámina para formar un pool a partir del cuál se extrajo el ADN genómico. Se utilizó el kit de extracción Kleargene (LGC) que sigue los siguientes pasos: lisis celular por detergentes, pegado del ADN a una membrana de fibra de vidrio mediado por guanidina isotiocianato, lavado con buffers y elución. Posteriormente, se comprobó que las muestras cumplan con los parámetros de calidad requeridos para su análisis, masa superior a 1000ng y una elevada pureza (ratio ABS_{260/280}= 1,8-2).

El estudio genómico realizado fue un microarreglo que busca 6.000 SNPs en posiciones determinadas del genoma de soja comercializado por la empresa Illumina y diseñado por Song *et al.* (Song y col., 2020). Estos SNPs están cuidadosamente elegidos a partir de un estudio inicial que parte del estudio de 50.000 SNPs (Song y col., 2013).

2.1.2. Búsqueda de genes polimórficos diferencialmente expresados

Se buscaron tanto SNPs polimórficos entre las líneas GS y GT, como genes diferencialmente expresados en respuesta a la sequía (Capítulo 4), los cuales se agruparon de acuerdo a su ubicación en el genoma para la posterior construcción de mapas cromosómicos utilizando el paquete chromoMap de R. Dentro de los SNPs polimórficos, se buscaron aquellos que se correspondieran con genes diferencialmente expresados en respuesta a la sequía en cada una de las líneas. Se enriqueció la tabla de SNPs y el listado de transcriptos con expresión diferencial utilizando el listado de marcadores y anotaciones provisto por soybase.org.

2.2. Búsqueda de mutaciones en la secuencia transcriptómica

Las mutaciones producidas en la secuencia transcriptómica pueden generar cambios importantes ya que influyen directamente en la expresión y funcionalidad de las proteínas resultantes. Para la búsqueda de mutaciones en la secuencia transcriptómica, ya sean SNPs o INDELs se utilizó el software GATK (del inglés, *Genome Analysis ToolKit*, McKenna y col., 2010). Para el filtrado inicial se utilizaron los parámetros estándar del programa y se definieron otros (cluster:3, WindowSize:35; QD < 2.0; FS > 30.0; DP<10.)

Una vez identificadas las mutaciones se anotaron utilizando el software SnpEff (Cingolani y col., 2012), incluyendo la función e impacto del cambio aparecido. La función de la mutación puede ser con sentido (cambio de un aminoácido por otro), sin sentido (introducción de un codón stop) o silenciosa (sin cambio de fenotipo). El impacto puede clasificarse en alto (si SnpEff calcula que tendrá efectos deletéreos), moderado, bajo o modificador. Hay que tener en cuenta que el software SnpEff puede asignarle más de un efecto a una misma mutación. Por ejemplo, un SNP puede estar listado como que provoca efectos *downstream*, *upstream* y en una región intergénica. Esto se debe a que por default se toma una región de 5K pb de largo para buscar genes corriente arriba o abajo.

Para seleccionar las mutaciones de interés e impacto se aplicaron sucesivos filtros. En primer lugar, se filtraron aquellas mutaciones que e 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 GT, tomando aquellas variantes donde el GS (tanto en control como sequía) presentó la misma variante que la referencia pero el GT la alternativa, ya sea en ambas condiciones hídricas o solo bajo sequía, lo que podría demarcar eventos de splicing alternativo asociados a la respuesta al estrés. A continuación, se seleccionaron las mutaciones que el software SNPeff indicó como de impacto alto o moderado, y por último, aquellas que cayeran en genes diferencialmente expresados frente a la sequía en el GT. Solo se admitieron aquellos SNPs en donde las cuatro repeticiones biológicas de cada tratamiento presentaran la misma variante (no se aceptaron heterocigotas, ni diferencias entre repeticiones biológicas). Esto no quiere decir que los marcadores que se descartaron en estos pasos no sirvan, sino que de esta manera se puede realizar un filtro lo suficientemente restrictivo para llegar a un número de marcadores que sea utilizable en un plan de mejoramiento.

3. Resultados

3.1. SNPs en la secuencia genómica

De los 6000 SNPs evaluados en el SoySNP6k chip un 7 % no pudo determinarse, ya sea porque no se pudo leer la base correctamente o porque era una base no contemplada en el microarreglo. De los 5583 SNPs leídos, el 33,7 % fue común en los tres genotipos, un 17,7 % exclusivo al GS, un 5,2 % exclusivo al GT y un 10,7 % exclusivo a Williams Figura 5.1a. El GT compartió más variantes con Williams que el GS, teniendo un porcentaje de similaridad de 76,97 % vs. 58,10 %, por lo que al hacer un análisis de clusters el GT y Williams se agruparon por un lado y el GS por otro Figura 5.1b.



Figura 5.1: a: Diagrama de Venn mostrando los SNPs compartidos y exclusivos para cada genotipo. (GS= Genotipo Sensible, GT= Genotipo Tolerante). b: Dendrograma realizado a partir de la clusterización de los resultados del SoySNP6K chip. (GS= Genotipo Sensible, GT= Genotipo Tolerante, W=Williams)

Para seguir adelante en el diseño de los marcadores moleculares se tomaron los SNPs en donde el GT presentó una variante polimórfica con respecto al GS y Williams. Estos SNPs se cruzaron con los datos de expresión diferencial, seleccionando solo aquellos que cayeran dentro de la región de genes diferencialmente expresados para el GT, que no hayan cambiado su expresión en el GS. Hubo un único SNP que cumplió con todas las condiciones requeridas y específicamente, se ubicó en la región 5' UTR del gen Glyma.12G085900 (Cuadro 5.1). Este gen codifica para una fosfatasa que cataliza la degradación de 2-Carboxi-D-arabinitol 1-fosfato, una molécula que inhibe la rubisco, y aumentó casi un 30 % su expresión en respuesta a la sequía en el GT. Al caer el SNP en la región promotora del gen, se realizaron búsquedas de sitios de unión a factores de transcripción a través de la base de datos Jaspar (Castro-Mondragon y col., 2022), pero el cambio de base no resultó variaciones de los sitios de unión a factores de transcripción presentes en la región. Asimismo, se buscó en la base de datos Soybean Knowledge Base pero no se encontró que la secuencia alternativa derive en sitios de unión a algún miRNA que module la expresión del gen.

Cuadro 5.1: Información del SNP en la secuencia genómica que es polimórfico entre líneas y está en un gen diferencialmente expresado en el GT frente al estrés por sequía. (CHROMPOS= Número de cromosoma y posición donde cae el SNP, REF= alelo de referencia presente en GS y Williams, ALT= alelo alternativo presente en GT, log2FoldChange= nivel de expresión diferencial en el GT en respuesta a la sequía, padj= valor p ajustado de la exresión diferencial, locusName= identificación del locus polimórfico, chr= cromosoma donde se ubica, start= posición de inicio del gen, end= posición de finalización del gen, strand= hebra del ADN, descr= Descripción del gen).

CHROM_POS	REF	ALT	log2FoldChange	padj	locusName	chr	start	end	strand	descr
Gm12_6892713	Т	С	0.36463	0.020113	Glyma.12G085900	Gm12	6930928	6936598	+	Phosphoglycerate mutase family protein

3.2. Variantes en la secuencia transcriptómica

La búsqueda de mutaciones en la secuencia transcriptómica arrojó un total de 50.042 SNPs y 23.694 IN-DELs (inserciones o deleciones). A través del programa snpEff se calculó el impacto, función y región donde cayeron las mutaciones. Las mutaciones están buscadas tomando como referencia el genotipo Williams, y tal como se observó a través del 6K el GT presenta una mayor similaridad a este genotipo que el GS (Figura 5.1a). Por lo que, no debe llamar la atención que la cantidad de mutaciones encontradas para el primer genotipo sea menor que para el segundo.

Cuadro 5.2: Cantidad de SNPs e INDELs encontrados en el transcriptoma por cromosoma.

Cromosoma	SNPs	INDELs
Gm01	2108	951
Gm02	2228	1337
Gm03	2964	1039
Gm04	1525	1089
Gm05	2422	1080
Gm06	2759	1437
Gm07	2816	1248
Gm08	2803	1602
Gm09	2399	1135
Gm10	2301	1253
Gm11	1798	1197
Gm12	1696	1081
Gm 13	3180	1572
Gm 14	2045	1010
Gm 15	2751	1202
Gm 16	3170	1044
Gm 17	2056	1286
Gm 18	4869	1279
Gm 19	1980	1053
Gm 20	1665	1006

La mayoría de los SNPs e INDELs cayeron en la categoría "MODIFIER" que se refiere a cambios que probablemente no tengan impacto, o sea difícil de predecirlo (por ejemplo, un SNP que cae en una región intergénica). 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, o de impacto moderado ("MODERATE"), que indican cambios no disruptivos que podrían tener algún efecto en la efectividad de la proteína. Tan solo unos pocos SNPs (\sim 145 SNPs para el GS y 84 para el GT, \sim 1200 INDELs para ambos genotipos) fueron de alto impacto ("HIGH"), que son aquellos para los que se asume un impacto disruptivo en la proteína, causando el truncado, pérdida de función o degradación (Figura 5.2a, Figura 5.2b). Estos SNPs de alto impacto que podrían resultar de interés se encuentran detallados en el cuadro B.2 del Apéndice B. El impacto predicho para los SNPs guarda relación con la función y la región en donde cayeron. La gran mayoría provocó mutaciones silenciosas, mientras que muy pocos provocaron pérdida de sentido (Figura 5.2c).



Figura 5.2: Información del impacto y función predichos por SNPeff para las variantes encontradas en la secuencia transcriptómica.

Para seleccionar variaciones potenciales a utilizarse como marcadores moleculares se aplicaron una serie de filtros para lograr quedarse únicamente con aquellas polimórficas entre los genotipos de interés, asociadas al GT y relacionadas con la respuesta a la sequía de este genotipo (Figura 5.3). En primer lugar, se omitieron aquellas mutaciones que cayeran en contigs o scaffolds, quedando de esta manera únicamente las mutaciones que cayeron dentro de regiones conocidas de los distintos cromosomas (Cuadro 5.2). De esta manera, partiendo de 74.006 variantes entre los genotipos, se llegan a 73.436 que se distribuyeron de manera homogénea dentro de los 20 cromosomas, resultando en una cobertura adecuada e informativa. El siguiente paso fue seleccionar aquellas variantes asociadas al GT a través de dos búsquedas independientes. La primera búsqueda consistió en seleccionar las variantes en donde el GT presentó el alelo alternativo, ya sea bajo condiciones hídricas controladas o en sequía, y el GS el alelo de referencia. Esta búsqueda arrojó un total de 2.610 variantes. La segunda búsqueda, más restrictiva aún, consistió en seleccionar las variantes asociadas a la respuesta a la seguía en el GT. Es decir, aquellas donde sólo el GT en condiciones de seguía presentó el alelo alternativo, mientras que el GS (control o sequía) y el GT bien regado presentaron el alelo de referencia. Está búsqueda permitió encontrar 4 SNPs que podrían ser producto de splicing alternativo o algún otro mecanismo de introducción de variaciones en respuesta al estrés en el GT. De estas 2.614 variantes (2.610 de la primera búsqueda + 4 de la segunda búsqueda) se descartaron aquellas que no tuvieran un impacto 'ALTO' o 'MODERADO' adjudicado por SNPeff, lo que arrojó un total de 855 SNPs. Por último, se seleccionaron solo aquellos SNPs que cayeron en genes diferencialmente expresados en el GT frente al tratamiento de sequía (ver Capítulo 4), resultando en un panel de 25 SNPs de impacto moderado (Cuadro 5.3).



Figura 5.3: Pipeline utilizado para el filtrado de variantes. El número indica la cantidad de variantes resultantes luego de el filtro especificado en la derecha.

Los 25 SNPs encontrados generan variantes del tipo *missense* que se refiere a un cambio de base que provoca un aminoácido distinto. Estos se ubican en 16 genes distintos, presentando la mayoría de ellos 1 solo SNP, sin embargo algunos genes presentaron más de un SNP de interés. Estos genes fueron: Glyma.16G008900 (5 SNPs), Glyma.14G010500 (3 SNPs), Glyma.15G124700 (3 SNPs), Glyma.14G075000 (2 SNPs). Los 16 genes pueden clasificarse en tres grandes grupos: de respuesta al estrés, involucrados en la transferencia de la información genética, y una última categoría general (Cuadro 5.3).

En la primera categoría se agrupan 5 genes que contribuyen a la tolerancia a la sequía en el GT. El primero, corresponde al gen con más SNPs, el cual codifica para una serin-endopeptidasa anotada en la base PFAM y GO como un inhibidor de proteasas involucardo en la senescencia foliar (Glyma.16G008900, 5 SNPs). Este gen se expresó un 30 % más en el GT en condiciones de sequía con respecto a su control. El segundo gen dentro de esta categoría presenta más de un SNPs y corresponde a la enzima rafinosa sintasa involucrada en el metabolismo de la galactosa (Glyma.14G010500, 3 SNPs). Esta enzima cataliza el paso de galactinol a rafinosa y se expresó 2,17 veces más en el GT ante el estrés por sequía. El tercer gen dentro de esta categoría corresponde al mismo gen que resultó polimórfico en también en la secuencia genómica (Cuadro 5.1). Este codifica para la fosfatasa que participa en la activación de la rubisco, ya que cataliza la degradación de 2-Carboxi-D-arabinitol 1-fosfato, una molécula que la inhibe. Su expresión aumentó un $\sim 30\%$ frente a la sequía en el GT (Glyma.12G085900). Otro gen presente en esta categoría fue el Glyma.18G138300, el cual codifica para una enzima con actividadad oxidorreductasa, específicamente cataliza la producción de cetonas con consumo de poder reductor, actúa en el núcleo y cloroplastos en respuesta al frío y aumentó un 38 % su expresión en el GT en respuesta a la sequía. El último gen que podría incluirse en esta categoría corresponde a una fosfodiesterasa de nucleótidos cíclicos, la cuál aumentó un 60 % su expresión en el GT frente a sequía (Glyma.05G163100).

La segunda categoría esta compuesta por 5 genes involucrados en la transferencia de la información genética (transcripción y traducción del mensaje genético), los cuáles generalmente disminuyeron su expresión en respuesta a la sequía en el GT (ver Capítulo 4). El primer gen dentro de esta categoría corresponde al Glyma.14G075000 (2 SNPs) que codifica para la enzima que cataliza la activación de los tRNA, dejándolos listos para la adición de un grupo aminoacilo, pertenece a la familia de las flavodoxinas, análogas de las ferredoxinas. Este gen disminuyó un 30 % su expresión ante el estrés por sequía en el GT. También se encontraron un factor de iniciación de la transcripción de respuesta a hipoxia, el cuál disminuyó un 60 % su expresión; y un factor de elongación de la traducción, que disminuyó un 37 % su expresión frente al estrés por sequía (Glyma.18G096400 y Glyma.18G213900, respectivamente). Por úlitmo, se encontraron dos factores de transcripción con un sitio de unión MYB, pero a diferencia de los demás genes, estos aumentaron un 56 % y un 90 % su expresión frente a sequía en el GT (Glyma.09G167900 y Glyma.12G117700, respectivamente).

Se agruparon en la útlima categoría los 6 genes restantes. Tres de estos podrían estar involucrados en la respuesta favorable del GT ante la sequía, como el gen Glyma.15G124700 correspondiente a una proteína con actividad oxidorreductasa involucrada en el transporte electrónico de la etapa lumínica de la fotosíntesis con dominios de unión a NAD y FAD, el cuál aumentó un 30 % su expresión en el GT y acumuló 3 SNPs dentro de su secuencia codificante. Otro ejemplo es una aminotransferasa, específicamente una alanina-

glioxilato transaminasa, que actúa en mitocondrias y cloroplastos y está relacionada con la fotorrespiración y la respuesta a nitrógeno (Glyma.03G040600); la cuál se expresó \sim 2,5 veces más en respuesta a la sequía en el GT. Por último, el gen Glyma.17G194500 que codifica para un componente integral de la membrana, un transportador de solutos transmembrana. Actúa de manera simporte transportando péptidos y protones en contra de su gradiente; aumentó más de un 60 % su expresión en respuesta a la sequía. Luego, se encuentran los 4 genes restantes, que no tienen un rol claro en la respuesta a la sequía. El Glyma.10G088000, una enzima del tipo carboxipeptiadasa (hidroliza el residuo carboxilo de un aminoácido de una cadena polipeptídica) involucrada en la fotomorfogénesis, desarrollo de meristemas, raíces y la formación de tejido vascular en hojas. Se expresó 3,4 veces más en el GT frente a sequía. El Glyma.14G057400, una quinasa involucrada en la diferenciación de pelos radiculares, disminuyó su expresión en el GT (-30 %), presente en el cloroplasto y clasificado como un componente integral de la membrana. El Glyma.18G057900, que codifica para una epimerasa, una enzima oxidorreductasa que cataliza la formación de poder reductor en el citosol y duplicó su expresión en el GT en respuesta a la sequía.

Cuadro 5.3: Listado de los 25 SNPs seleccionados para ser utilizados como marcadores moleculares. CHROMPOS= Número 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 GS, ALT= alelo alternativo presente en el GT, EFFECT= efecto del SNP predicho por SNPeff, IMPACT= impacto del SNP predicho por SNPeff, log2FoldChange= Expresión diferencial en el GT en respuesta a la sequía, padj= valor p ajustado de la expresión diferencial, PFAM Names= Nombre del gen según la base Pfam, 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.

Nro	CHROM_POS	GeneName	REF	ALT	EFFECT	IMPACT	log2FoldChange	padj	PFAM Names	GO Biological Process Descriptions	GO Molecular Function Descriptions
1	Gm03_5211321	GLYMA.03G040600	А	С	missense_variant	MODERATE	1.39610	0.0002099	Aminotran_3;	photorespiration; cellular response to nitrogen levels;	alanine-glyoxylate transaminase activity; pyridoxal phosphate binding;
2	Gm05_35457627	GLYMA.05G163100	G	А	missense_variant	MODERATE	0.74366	0.0181680	CPDase;	tRNA splicing, via endonucleolytic cleavage and ligation; cyclic nucleotide metabolic process;	cyclic-nucleotide phosphodiesterase activity; 2',3'-cyclic-nucleotide 3'-phosphodiesterase activity;
3	Gm09_40118414	GLYMA.09G167900	А	Т	missense_variant	MODERATE	0.64129	0.0494590	Myb_DNA-binding;	regulation of transcription, DNA-templated; circadian rhythm;	DNA binding; DNA-binding transcription factor activity;
4	Gm10_11238659	GLYMA.10G088000	Т	С	missense_variant	MODERATE	1.75190	0.0000001	PA; Peptidase_M28; TFR_dimer;	photomorphogenesis; embryo development; flower development; negative regulation of flower development; leaf vascular tissue pattern formation; root development; meristem development;	carboxypeptidase activity; protein binding; metal ion binding;
5	Gm12_6914329	GLYMA.12G085900	G	С	missense_variant	MODERATE	0.36463	0.0201130	His_Phos_1;	NA	phosphatase activity; 2-carboxy-D-arabinitol-1-phosphatase activity;
6	Gm12_11931345	GLYMA.12G117700	С	G	missense_variant	MODERATE	0.93258	0.0161570	Myb_DNA-binding;	signal transduction; chloroplast organization; negative regulation of flower development; regulation of chlorophyll biosynthetic process; positive regulation of transcription, DNA-templated; negative regulation of leaf senescence;	DNA-binding transcription factor activity; protein binding; transcription regulatory region DNA binding;
7	Gm14_805931	GLYMA.14G010500	Т	G	missense_variant	MODERATE	1.12150	0.0001881	Raffinose_syn;	carbohydrate metabolic process;	galactinol-sucrose galactosyltransferase activity;
8	Gm14_808853	GLYMA.14G010500	С	Т	missense_variant	MODERATE	1.12150	0.0001881	Raffinose_syn;	carbohydrate metabolic process;	galactinol-sucrose galactosyltransferase activity;
9	Gm14_808929	GLYMA.14G010500	Α	Т	missense_variant	MODERATE	1.12150	0.0001881	Raffinose_syn;	carbohydrate metabolic process;	galactinol-sucrose galactosyltransferase activity;
10	Gm14_4505988	GLYMA.14G057400	G	С	missense_variant	MODERATE	-0.40956	0.0069068	LRR_8; LRRNT_2; Pkinase_Tyr;	protein phosphorylation; root hair cell differentiation;	protein serine/threonine kinase activity; protein binding; ATP binding;
11	Gm14_6168634	GLYMA.14G075000	Т	А	missense_variant	MODERATE	-0.43106	0.0176100	Flavodoxin_1; Radical_SAM; Wyosine_form;	tRNA processing; oxidation-reduction process;	FMN binding; metal ion binding; 4 iron, 4 sulfur cluster binding; tRNA-4-demethylwyosine synthase activity;
12	Gm14_6169775	GLYMA.14G075000	С	G	missense_variant	MODERATE	-0.43106	0.0176100	Flavodoxin_1; Radical_SAM; Wyosine_form;	tRNA processing; oxidation-reduction process;	FMN binding; metal ion binding; 4 iron, 4 sulfur cluster binding; tRNA-4-demethylwyosine synthase activity;
13	Gm15_9848408	GLYMA.15G124700	С	А	missense_variant	MODERATE	0.44425	0.0444920	Ferric_reduct; FAD_binding_8; NAD_binding_6;	ion transport; photosynthetic electron transport chain; oxidation-reduction process;	ferric-chelate reductase activity; metal ion binding;
14	Gm15_9852551	GLYMA.15G124700	С	G	missense_variant	MODERATE	0.44425	0.0444920	Ferric_reduct; FAD_binding_8; NAD_binding_6;	ion transport; photosynthetic electron transport chain; oxidation-reduction process;	ferric-chelate reductase activity; metal ion binding;
15	Gm15_9852719	GLYMA.15G124700	С	Т	missense_variant	MODERATE	0.44425	0.0444920	Ferric_reduct; FAD_binding_8; NAD_binding_6;	ion transport; photosynthetic electron transport chain; oxidation-reduction process;	ferric-chelate reductase activity; metal ion binding;
16	Gm16_772780	GLYMA.16G008900	G	С	missense_variant	MODERATE	0.38961	0.0315300	PA: Inhibitor_I9: Peptidase_S8:	proteolysis: leaf senescence: secondary shoot formation:	serine-type endopeptidase activity:
17	Gm16_773621	GLYMA.16G008900	С	Т	missense_variant	MODERATE	0.38961	0.0315300	PA; Inhibitor_I9; Peptidase_S8;	proteolysis; leaf senescence; secondary shoot formation;	serine-type endopeptidase activity;
18	Gm16_773822	GLYMA.16G008900	Т	С	missense_variant	MODERATE	0.38961	0.0315300	PA; Inhibitor_I9; Peptidase_S8;	proteolysis; leaf senescence; secondary shoot formation;	serine-type endopeptidase activity;
19	Gm16_774303	GLYMA.16G008900	Т	G	missense_variant	MODERATE	0.38961	0.0315300	PA; Inhibitor_I9; Peptidase_S8;	proteolysis; leaf senescence; secondary shoot formation;	serine-type endopeptidase activity;
20	Gm16_774990	GLYMA.16G008900	G	А	missense_variant	MODERATE	0.38961	0.0315300	PA; Inhibitor_I9; Peptidase_S8;	proteolysis; leaf senescence; secondary shoot formation;	serine-type endopeptidase activity;
21	Gm17_29090428	GLYMA.17G194500	А	G	missense_variant	MODERATE	0.70685	0.0123740	NA	NA	peptide:proton symporter activity; oligopeptide transmembrane transporter activity; low-affinity nitrate transmembrane transporter activity; peptide transmembrane transporter activity;
22	Gm18_5060538	GLYMA.18G057900	С	Т	missense_variant	MODERATE	0.91799	0.0065490	Epimerase;	NA	oxidoreductase activity; oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor; coenzyme binding;
23	Gm18_9803255	GLYMA.18G096400	Α	G	missense_variant	MODERATE	-0.75432	0.0388440	SUI1;	translational initiation; cellular response to hypoxia;	RNA binding; translation initiation factor activity;
24	Gm18_20438691	GLYMA.18G138300	A	Т	missense_variant	MODERATE	0.46589	0.0335350	ADH_N; ADH_zinc_N_2;	response to cold; oxidation-reduction process;	zinc ion binding; enone reductase activity; 2-alkenal reductase (NADP+) activity;
25	Gm18_50372038	GLYMA.18G213900	С	Α	missense_variant	MODERATE	-0.45694	0.0444920	EFP; Elong-fact-P_C; EFP_N;	translational elongation;	translation elongation factor activity;

4. Discusión

Las técnicas utilizadas para la búsqueda de variaciones generan gran cantidad de información y uno de los mayores desafíos es poder seleccionar un panel de SNPs lo suficientemente reducido como para que pueda utilizarse en planes de mejoramiento, pero que tengan impacto. Se encontraron en total más de 79.000 variaciones, 5.583 en la secuencia genómica a través de microarreglos y 73.436 en el total de la secuencia transcriptómica. A través de los sucesivos filtros impuestos se logró disminuir ese número a uno más adecuado e informativo. En primer lugar, el tamaño final de panel propuesto (26 SNPs, 25 provenientes de la secuencia transcriptómica + 1 proveniente de la secuencia genómica) es un número manejable y aplicable a un plan de mejoramiento. En segundo lugar, se trata de SNPs elegidos cuidadosamente y que guardan relación con la respuesta de tolerancia al estrés en el genotipo tolerante. Es importante aclarar que la gran mayoría de los SNPs que se filtraron en el camino no se descartan, estos pueden ser utilizados también si se cambian los criterios de búsqueda. Esto quiere decir, que las posibilidades de generación de marcadores moleculares a partir de este trabajo no se limita a los 26 SNPs aquí propuestos, sino que puede ampliarse en caso de ser necesario.

La búsqueda de SNPs a través de microarreglos es una técnica ampliamente utilizada en el cultivo de soja, de hecho en la Argentina se exige por parte del INASE (Instituto Nacional de Semillas) que cada nueva variedad registrada informe un panel de 4004 SNPs que surge a partir del microarreglo de 6K de Illumina. Sin embargo, los microarreglos tienen sus limitaciones. En primer lugar, sólo permiten búsquedas bialélicas, por lo que cuando hay varios loci indeterminados puede deberse a que la opción encontrada era una tercera no contemplada dentro del chip, o porque realmente hubo un error en la lectura de esa base. La otra gran limitante de los microarreglos predeterminados es que están diseñados tomando poblaciones específicas, que tal vez son genéticamente distantes de los materiales a evaluar, por lo que puede haber varios SNPs indeterminados, o una incapacidad de detectar la verdadera variabilidad dentro de las muestras. Esto ocurrió en cierta medida en este trabajo, donde se observó que el GT tenía un mayor porcentaje de similaridad con Williams que con el GS. La información del pedigree de estos materiales es clasificada y no se dispuso para este trabajo, pero seguramente estén más emparentadas entre sí que con Williams (genotipo de referencia, no comercial); y esta variabilidad no se ve reflejada a través del 6k.

La búsqueda de SNPs en la totalidad del transcriptoma a través de un RNAseq permite encontrar una gran cantidad de polimorfismos específicamente en regiones codificantes. Además, en este caso en particular se pudieron asociar estas variaciones a datos de expresión diferencial de genes, haciendo que los polimorfismos encontrados tengan también un impacto en la respuesta a la sequía. Esta técnica ya se ha utilizado en para la generación de marcadores de resistencia a chinches (Sabljic, 2020) y también sequía (Oliveira Vidal y col., 2012) en el cultivo de soja. Al buscar variantes en la totalidad del transcriptoma, esta técnica permitió encontrar 13 veces más SNPs que los encontrados con el microarreglo de Illumina. Como consecuencia, al aplicar los filtros elegidos para la selección de los SNPs para ser utilizados como marcadores moleculares, con esta técnica se llegó a 25 SNPs mientras que con la otra solo a uno.

Dos de los 16 genes con SNPs forman parte de las vías enriquecidas según la base de datos KEGG (Kyoto Encyclopedia of Genes and Genomes), y uno de las vías enriquecidas de GO (Gene Ontology) en el GT frente a sequía (ver Capítulo 4). Específicamente, el gen Glyma.03G040600 está presente dentro de 4 vías distintas: metabolismo de cisteínas y metioninas (gmx00270), metabolismo de glicina, serina y treonina (gmx00260), metabolismo de la alanina, aspartato y glutamato (gmx00250); y degradación de valina, isoleucina y leucina (gmx00280). Este gen está involucrado en el proceso de fotorrespiración, el cuál disminuye la eficiencia fotosintética al perderse parte del carbono fijado. Sin embargo, este proceso protege a las plantas del exceso de luz cuando hay una baja concentración de CO₂ en la hoja debido a un cierre estomático (Buchanan, Gruissem y Jones, 2015). El otro gen polimórfico presente en una vía enriquecida en el GT en respuesta a la seguía fue el Glyma.14G010500. Este gen acumuló 3 SNPs y está presente en la vía del metabolismo de la galactosa (gmx00052), su función es catalizar la reacción que produce rafinosa partiendo de galactinol, ambos compuestos son acumulados activamente por las plantas en respuesta a la sequía para mantener el potencial osmótico sin interferir con el metabolismo celular. La acumulación de rafinosa tiene un efecto osmótico y antioxidante, pudiendo detoxificar los radicales hidroxilos que provocan daños a macromoléculas (Buchanan, Gruissem y Jones, 2015; Hossain y col., 2014; Fuganti-Pagliarini y col., 2017). El gen Glyma.18G096400 se encontró dentro de la vía "rRNA binding" (GO:0003723), la cuál es parte de la vía RNA binding que resultó enriquecida en el GT frente a la sequía.

Los demás genes polimórficos, a pesar de haberse expresado diferencialmente, no estuvieron dentro de vías enriquecidas en el GT en respuesta a la seguía. Sin embargo, esto no quiere decir que no sean de interés en la respuesta de este genotipo al estrés hídrico. Por ejemplo el gen que presentó SNPs de interés tanto en la secuencia genómica como transcriptómica, Glyma.12G085900, codifica para una enzima que participa en la activación de la rubisco, ya que degrada un componente inhibitorio, el 2-carboxiarabinitol-1-fosfato. Este azúcar se une a la rubisco carbamilada inhibiendola durante períodos de oscuridad o baja intensidad lumínica (Milward, 2018; Servaites, 1990), por lo que el aumento de la proteína que lo degrada podría estar contribuyendo a disminuir su inhibición en los períodos de estrés donde su actividad se ve limitada no sólo por la baja concentración de CO₂ en la cámara subestomática. Otro gen que resultó polimórfico y está involucrado en la respuesta a la sequía es el Glyma.18G138300 que tiene actividad 2-alquenal reductasa. Se ha demostrado que plantas de tabaco que sobreexpresan esta enzima presentan una mayor tolerancia a la sequía, obteniendo mayor biomasa, capacidad fotosintética, contenido y fluorescencia de clorofila; y capacidad de detoxificar maldonaldehído y H_2O_2 que las plantas wild type (Wu y col., 2015). Por otro lado, el gen Glyma.05G163100 codifica para una fosfodiesterasa de nucleótidos cíclicos. Estas enzimas hidrolizan los enlaces fosfoester presente entre los grupos fosfato, dando como resultado nucleótidos 5' y 3' monofosfatos. La evidencia sugiere que el balance entre nucleótidos cíclicos y Ca2+ es crucial en la regulación de distintos efectores, como las fosfodiesterasas y los canales de nucleótidos cíclicos, de respuesta y señalización de procesos patogénicos y fisiológicos en plantas (Świeżawska-Boniecka y col., 2021). Los canales de nucleótidos cíclicos están implicados en la tolerancia a estrés biótico y abiótico en plantas, específicamente, se cree que estos podrían mediar la entrada de Ca^{2+} al citosol, el cuál previene la síntesis de oxidantes, protege las membranas, promueve la fotosíntesis y modula la expresión de genes de respuesta y fitohormonas (Jha, Sharma y Pandey, 2016). El gen que más SNPs acumuló (Glyma.16G008900) codifica para un inhibidor de proteasas involucrado en la senescencia foliar y aumentó un 30 % su expresión en el GT frente a la sequía. Esto concuerda con los resultados encontrados por en la tesis presentada por De Oliveira en donde este gen aumentó un 36 % su expresión en respuesta a la sequía en genotipos de soja transgénicos que sobreexpresan el factor de transcripción GmNAC081, un elemento de convergencia entre la senescencia y tolerancia a sequía (De Oliveira Ferreira, 2017). Este gen también ha sido identificado como candidato dentro de QTLs para el mejoramiento de la calidad, específicamente la cantidad de proteína y composición aminoácidica, de la soja (Gong y col., 2018). Los 4 genes restantes agrupados dentro de la categoría "transferencia de la información genética", no están específicamente dentro de las vías enriquecidas resultantes, pero guardan relación con estas. De los factores de transcripción con sitio de unión a MYB que aumentaron su expresión, el Glyma.12G117700 se ha destacado anteriormente (ver Subsubsección 3.3.2, Capítulo 4) porque interviene en la cascada de señalización de MAPK, en la organización del cloroplasto, síntesis de clorofila y plástidos, y el metabolismo del H_2O_2 .

Los genes polimórficos restantes tal vez no tienen una función tan evidente en la respuesta a la sequía como los mencionados anteriormente, pero aún así son relevantes. El proceso de selección de SNPs fue ideado para que todos los genes seleccionados estén estrictamente asociados al GT en condiciones de sequía. Por lo que, a pesar de no haber evidencia concreta sobre el accionar de estos genes en la tolerancia a la sequía en la bibliografía, podemos concluir que podrían ser importantes para este genotipo estudiado en particular. Por ejemplo, el Glyma.15G124700 está involucrado en la cadena de transporte electrónico de la etapa lumínica de la fotosíntesis generando poder reductor. En condiciones de estrés y cierre estomático, el normal funcionamiento de la cadena de transporte de electrones y la generación de poder reductor es un indicio de que no hay generación de estrés oxidativo que podría dañar a las células. Las serin-treonina quinasas como la proteína codificada por Glyma.14G057400, y los canales de transporte transmembrana, como la proteína codificada por Glyma.17G194500, están involucradas en procesos de señalización; por lo que podrían estar mediando la respuesta del GT a la sequía río arriba.

Un paso previo a la utilización de estos marcadores en planes de mejoramiento es su validación en poblaciones segregantes. Esto consiste en generar una población de mapeo mediante el cruzamiento de ambos genotipos, someterla a sequía, y luego evaluar si su comportamiento ante el estrés se condice con la aparición de estos SNPs. Se esperaría que aquellas plantas con mejor comportamiento tengan presentes en los SNPs los alelos asociados al GT. La generación de la población para realizar esta validación es un proceso largo que lleva varios años de cruzamientos y siembras a campo, por lo que quedó fuera del alcance de esta tesis. En caso de resultar validados los 26 marcadores, implementarlos dentro de un plan de mejoramiento podría ser económicamente viable. Correr esos 26 SNPs en una plataforma de detección masiva como la SNP Line (LGC) tendría un costo aproximado de 8,5 USD por muestra, y podría permitir seleccionar genotipos con mayor tolerancia al estrés sin necesidad de realizar ensayos específicos de sequía. Asimismo, no haría falta esperar a que la planta tenga un estado de desarrollo específico, pudiendo ahorrar tiempos y superficies cultivables. Las regiones flanqueantes de los SNPs a partir de las cuales se diseñan las sondas KASP se encuentra en el anexo (Apéndice B, Cuadro B.1).

En caso de querer agrandar el panel para lograr discriminar entre materiales específicos, hay una gran cantidad de SNPs que son interesantes también. Estos se han dejado de lado en el proceso de selección, no porque no funcionen, sino porque al partir de un número tan grande de polimorfismos se tomó un criterio estricto para alcanzar un número manejable y aplicable en planes de mejoramiento. Algunos ejemplos de variantes de interés que no fueron propuestas en el panel final fueron aquellas que tienen un impacto "HIGH" predicho por SNPeff. Estas suelen ser introducciones o pérdidas de codones stop que generan proteínas truncas y no funcionales. Un mRNA que genera una proteína truncada es degradado rápidamente por la célula (Lykke-Andersen y Jensen, 2015), por lo que es lógico que no haya resultado diferencialmente expresado en el GT. Otras variaciones que tienen un alto impacto son aquellas que generan pérdida o ganancia de un sitio de splicing. También se encontraron inserciones y deleciones que provocaron un cambio en el marco de lectura, ya sea por el corrimiento, pérdida o ganancia de este. Nuevamente, este tipo de variaciones provocan cambios que no se van a ver reflejados en la expresión diferencial, pero aún así son de sumo interés. Un resumen de posibles alternativas se detalla en el Cuadro B.2 presente en el Apéndice B.

Otra alternativa para obtener una validación de los marcadores de manera independiente sería evaluar la expresión de estos genes bajo distintas condiciones de sequía. Por ejemplo, se podrían diseñar nuevos experimentos en condiciones controladas (cámaras de cultivo, invernaderos) o a campo, en donde se desafíen los genotipos contrastantes al estrés por sequía y se evalúe la expresión de estos genes. Asimismo, otra posibilidad podría ser generar poblaciones mutantes de *Arabidopsis* para estos genes y evaluar su respuesta en comparación con los genotipos *wild type*. Estas evaluaciones permitirían acumular más evidencia independiente sobre el rol de estos genes en la tolerancia a la sequía.

5. Conclusiones

La selección asistida por marcadores moleculares es una herramienta poderosa que ayuda a ahorrar tiempo, superficie y dinero; ya que permite seleccionar genotipos con características deseadas sin necesidad de imponer el estrés y observar la respuesta para poder determinar su comportamiento. En este trabajo se propuso un panel de 26 marcadores de tolerancia a sequía seleccionados a partir de búsquedas en la secuencia genómica y transcriptómica de genotipos con comportamientos contrastantes a la sequía. Para llegar a marcadores de calidad se aplicaron filtros sucesivos que permitieron descartar marcadores heterocigotas o que no mapearon a lugares específicos dentro de los cromosomas. Luego, para seleccionar marcadores con impacto en la respuesta a la sequía, se seleccionaron aquellos asociados al GT que cayeran dentro de genes diferencialmente expresados. Se construyó un panel de 26 marcadores, de los cuales 11 caen en genes que guardan estrecha relación con la tolerancia a la sequía, y los restantes guardan relación con lo observado en el análisis transcriptómico (ver Capítulo 4).

Para poder utilizar estos marcadores en planes de mejoramiento se deben validar a través de pruebas en poblaciones de mapeo. En caso de que algunos marcadores no funcionen, se pueden cambiar (o flexibilizar) los criterios de selección y elegir marcadores distintos, o incluso ampliar el panel para que pueda discriminar entre materiales de interés.

Capítulo 6

Conclusiones Generales

1. Conclusiones finales

La soja es el cultivo más importante producido en la Argentina, y a su vez, el país es el tercer productor mundial de esta leguminosa por lo que las mejoras genéticas realizadas en nuestro país no son solo de interés nacional, sino mundial. La producción de este cultivo no se limita a la zona núcleo sino que se exploran diversos ambientes, siendo algunos más propensos a sufrir eventos de sequía. Además, debido al cambio climático, se estima que los eventos de sequía podrían exacerbarse y volverse más frecuentes incluso en ambientes donde antes no lo eran. Dadas estas condiciones, la generación de materiales de soja con tolerancia a la sequía es de suma importancia.

En este trabajo se partió de dos genotipos provenientes del germoplasma de una semillera nacional con comportamientos contrastantes a la sequía y se llegó a un panel de 26 marcadores moleculares de tolerancia a este estrés que, previo a una validación, podría utilizarse en programas de mejoramiento de manera sencilla y económicamente viable. El estudio previo del que surgieron estos genotipos fue realizado por el grupo de trabajo, en donde a través de un protocolo rápido de testeo se evaluó la respuesta a la sequía de 630 genotipos del programa de mejoramiento de la semillera GDM mediante la medición de características morfológicas (peso y longitud de parte aérea y longitud de raíz) y fisiológicas (días a VE y VC Fehr y Caviness, 1977). En base a los resultados obtenidos se seleccionaron 4 genotipos, dos de buen comportamiento y dos de mal comportamiento, a los que se sometió nuevamente a un estrés por sequía y se midió la expresión de tiorredoxinas (proteínas involucradas en la respuesta redox al estrés), seleccionándose aquellos con expresiones más contrastantes (Demicheli, 2017). Los genotipos seleccionados allí fueron denominados GS y GT, susceptible y tolerante respectivamente, y son los estudiados en este trabajo.

Como primer paso, este estudió comenzó con una profundización de la caracterización de la respuesta a la sequía de los dos genotipos seleccionados a modo de validación del protocolo rápido de testeo desarrollado en el trabajo anterior. Esta caracterización exhaustiva constó de la medición del marchitamiento, la fotosíntesis y parámetros asociados, la transpiración y construcción de un índice de transpiración normalizada; y por último, la caracterización epidérmica y estomática a través de microscopia electrónica. Todas estas mediciones se realizaron en los genotipos de interés (GS y GT) y genotipos de bibliografía (Williams 82 y PI416937). Los resultados de estos estudios permitieron validar el protocolo de testeo rápido desarrollado previamente, ya que permitió discriminar correctamente entre materiales según su comportamiento a la sequía, siendo la primera contribución de esta tesis. Los resultados se detallaron en el Capítulo 3 y se resumirán en la Sección 2 de este capítulo.

El entendimiento de la respuesta diferencial a la sequía de cada genotipo permitió diseñar un experimento que logre captar esta respuesta a nivel génico para estudiarla en profundidad. En el Capítulo 4 se detallan los procedimientos y resultados del estudio transcriptómico que develó los genes y vías metabólicas causantes de las respuestas fenotípicas observadas. En síntesis, se encontró que el GT presentó aumentada la expresión de genes clave en la respuesta al estrés como inhibidores de proteasas, enzimas involucradas en la acumulación de solutos compatibles que actúan como antioxidantes y genes involucrados en varios procesos de señalización (señalización de dos componentes, MAPK, Ca²⁺). Asimismo, presentó menos disminuida que el GS la vía de las auxinas y más estimulada la vía del ABA, de las citocininas y de las giberelinas. Por más que en el GS aumentó la expresión de algunos genes de respuesta a sequía interesantes, como por ejemplo HSP, LEAs o algunos involucrados en el estrés oxidativo, disminuyó la expresión de muchos genes claves. En primer lugar, vías del metabolismo primario que son esenciales para el normal funcionamiento de la célula se vieron comprometidas en el GS, como la asimilación de C, N o S; o el metabolismo de nucleótidos. Numerosos genes involucrados en la fotosíntesis disminuyeron su expresión. En la etapa liumínica disminuyó la expresión de componentes estructurales de los fotosistemas, de los complejos antenas y hasta de proteínas involucradas en el transporte electrónico. También hubo menor expresión de varias enzimas del ciclo de Calvin y de los demás procesos del metabolismo del C, como la degradación del almidón y la glucólisis. Además, disminuyó la expresión de varios genes involucrados en procesos de crecimiento y expansión celular. De esta manera, se pudo concluir que la tolerancia a la sequía exhibida por el GT estaba respaldada por varios procesos de evitación y mitigación del estrés, y la susceptibilidad del GS explicada por varios procesos fundamentales que se vieron afectados.

Por último, se buscaron polimorfismos (SNPs e INDELs) en la secuencia genómica y transcriptómica que estuvieran asociados a la respuesta a la sequía del GT Capítulo 5 llegándose así a un panel de 26 marcadores de tolerancia. Estos 26 marcadores corresponden a SNPs ubicados en 16 genes distintos, todos provocan variaciones del tipo *missense* que indican el cambio de un aminoácido por otro y tienen un impacto moderado predicho por SNPeff. Un solo marcador de los 26 propuestos corresponde a la secuencia genómica y cae en la región 5' UTR de uno de los genes que posee SNPs de interés en su secuencia codificante. La gran mayoría de los 25 marcadores derivados del SNP *calling* realizado del RNA-seq caen en genes de respuesta a la sequía, como por ejemplo inhibidores de proteasas, enzimas de síntesis de solutos compatibles, activadores de la rubisco o con actividad oxidorreductasa; que aumentaron su expresión en el GT frente a la sequía.

2. Cumplimiento de objetivos y verificación de hipótesis

Los objetivos planteados en este tesis se cumplieron en su totalidad, y los resultados asociados a cada uno de ellos se detallan a continuación:

Objetivo 1: Caracterizar la respuesta a la sequía de genotipos de soja con comportamientos contrastantes ante este estrés.

Para caracterizar la respuesta a la sequía de los genotipos contrastantes (GS y GT) se realizaron distintos experimentos detallados en el Capítulo 3 en donde se midieron parámetros de interés: marchitamiento, transpiración (acumulada y normalizada), fotosíntesis y demás parámetros asociados al intercambio gaseoso de las hojas; y por último, una caracterización estomática y epidérmica. A través de estos experimentos se identificó que, a diferencia del GS, el GT toleró mejor la sequía exhibiendo menor marchitamiento gracias a un mecanismo que le permitió conservar el agua del suelo. Este mecanismo se basó en una disminución de la transpiración relativa (NTR) a mayores contenidos hídricos relativos, causado por una mayor sensibilidad estomática (g_s), dando como resultado una mayor eficiencia de uso del agua (WUE) ante un evento de sequía. Estos resultados permitieron verificar la primera hipótesis: *"El genotipo tolerante se diferenciará del sensible en diversos parámetros fisiológicos y morfológicos que le permitirán utilizar el agua de una manera más eficiente minimizando el estrés provocado por la sequía"*.

Objetivo 2: Identificar los genes, y rutas metabólicas asociadas, con expresión diferencial ante el estrés por sequía en los genotipos contrastantes.

El segundo objetivo se logró cumplir a través de la realización de un RNAseq y posterior análisis bioinformático de los datos, detallado en el Capítulo 4. El cumplimiento del primer objetivo permitió diseñar un experimento para estudiar la respuesta contrastante a la sequía de los genotipos de interés para analizarla mediante el estudio de su transcriptoma. Este estudio permitió identificar 942 y 423 genes diferencialmente expresados en el GT y GS, respectivamente. Estos se agruparon en diferentes categorías enriquecidas, que guardaron relación con la respuesta de cada genotipo al estrés hídrico. El GT presentó varias vías clasificadas como de tolerancia aumentadas (señalización de dos componentes, síntesis de solutos compatibles y procesos de oxidorreducción), mientras que el GS presentó disminuidas vías de interés (transporte de cationes, metabolismo de carbohidratos). De esta manera quedó verificada la segunda hipótesis: *"El genotipo tolerante presentará enriquecidas rutas metabólicas propias de la tolerancia al estrés"*.

Objetivo 3: Analizar el polimorfismo de los principales genes de respuesta a la sequía y utilizarlo en el diseño de marcadores moleculares.

En el Capítulo 5 se buscaron polimorfismos tanto en la secuencia genómica como en la transcriptómica, a partir de las cuáles se aplicaron sucesivos filtros hasta llegar a un panel de SNPs que podrían ser utilizados como marcadores moleculares, cumpliéndose así el objetivo número 3. Los filtros impuestos permitieron identificar 26 variantes asociadas a la respuesta a la sequía del genotipo tolerante que cayeron en 16 genes distintos, de los cuales 5 tienen un rol fundamental en la tolerancia ante dicho estrés. Entre estos se destacaron genes de inhibición de la senescencia, síntesis de compuestos antioxidantes y actividad oxidorreductasa, degradación de compuestos inhibitorios de la rubisco, involucrados en procesos de señalización y factores de transcripción. De esta manera, se aceptó la tercera y última hipótesis: "Genes presentes en estas rutas metabólicas presentarán un polimorfismo estructural que permitirá su utilización como potenciales marcadores moleculares".

3. Perspectivas futuras

El trabajo aquí presentado aporta información valiosa para la evaluación y selección de materiales frente al estrés por sequía en soja. La evaluación fenotípica realizada en los genotipos contrastantes permitió validar un protocolo de testeo rápido y eficaz que podría ser aplicable en planes de mejoramiento, mientras que simultáneamente permitió una caracterización detallada de la respuesta a la sequía de los genotipos de interés. Por otro lado, el cruce entre la información transcriptómica y la búsqueda de variantes permitió construir un panel de marcadores de tolerancia a sequía, que podría escalarse con facilidad. Además, la información presentada aquí podría aplicarse en diversos enfoques, por ejemplo, si se quisiera contar con marcadores de susceptibilidad para lograr descartar materiales susceptibles.

El uso de marcadores moleculares en los programas de mejoramiento es una herramienta clave ya que permite aumentar la eficiencia de selección y la introgresión de rasgos de interés. La utilización de marcadores para rasgos gobernados por uno o pocos genes, como la resistencia a herbicidas y ciertas enfermedades o plagas, es más fácil y común. Sin embargo, para características poligénicas, como la tolerancia a la sequía, aún hay mucho camino por delante, ya que identificar los marcadores adecuados es complejo y requiere de un trabajo minucioso. A diferencia del enfoque propuesto por la selección genómica donde se toman marcadores sólo por que están asociados al fenotipo de interés, el trabajo aquí propuesto tiene un enfoque de genómica funcional. Esta perspectiva de trabajo tiene una fuerte base bioquímica, donde se busca que los marcadores propuestos estén involucrados en las vías metabólicas causantes de la respuesta tolerante.

Este trabajo sentó los precedentes para un trabajo futuro: la validación de los marcadores aquí propuestos. Para tal fin se podría desarrollar una población de mapeo del tipo RILs (líneas recombinantes endocríadas, del inglés Recombinant Inbreed Lines) a la cuál se debe someter a estrés por sequía y luego analizar con los marcadores aquí propuestos. Se esperaría que aquellos materiales tolerantes presenten los alelos asociados al GT, y aquellos susceptibles los alelos de referencia al igual que el GS. Sin embargo, la generación de poblaciones segregantes es un trabajo arduo y largo, por lo que otra alternativa podría ser validar los marcadores en una población comercial de genotipos con comportamientos conocidos a la sequía, tal como se hizo en Uruguay para validar marcadores de marchitamiento (Quero y col., 2021).

En conclusión, esta tesis aporta el punto de partida para trabajos futuros, ya sea validando los marcadores aquí propuestos o generando un nuevo panel de interés a partir de la información aquí provista.

Bibliografía

Abdelrahman, Mostafa y col. (2018). «Legume genetic resources and transcriptome dynamics under abiotic stress conditions». En: *Plant, Cell and Environment* 41, págs. 1972-1983.

Alam Khan, Mueen (2018). «Achievements and prospects of molecular breeding for drought tolerance in soybean [Glycine max (L.) MERR.]» En: *Genetika* 50.3, págs. 1095-1109.

Alqurashi, May y col. (2018). «Early responses to severe drought stress in the Arabidopsis thaliana cell suspension culture proteome». En: *Proteomes* 6.4.

- Anders, Simon y Wolfgang Huber (2010). «Differential expression analysis for sequence count data». En: *Genome Biology* 11.10, págs. 1-12.
- Aroca, Ricardo, ed. (2012). *Plant Responses to Drought Stress: From Morphological to Molecular Features*. Springer International Publishing.
- Ashburner, Michael y col. (2000). «Gene Ontology: tool for the unification of biology». En: *Nature genetics* 25.1, pág. 25.
- Bao, Gegen y col. (2016). «Co-expression of NCED and ALO improves vitamin C level and tolerance to drought and chilling in transgenic tobacco and stylo plants». En: *Plant Biotechnology Journal* 14.1, págs. 206-214.
- Bateman, Alex y col. (2021). «UniProt: the universal protein knowledgebase in 2021». En: *Nucleic Acids Research* 49.D1, págs. D480-D489.
- Benjamini, Yoav y Yosef Hochberg (1995). «Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing». En: *Journal of the Royal Statistical Society: Series B (Methodological)* 57.1, págs. 289-300.
- Bernardo, Rex y Stemma Press Woodbury (2020). *BREEDING for QUANTITATIVE TRAITS in PLANTS*. Ed. por Bernardo Rex. Third Edit. Minnesota: Stemma Press.
- Bhatnagar-Mathur, Pooja y col. (2009). «Differential antioxidative responses in transgenic peanut bear no relationship to their superior transpiration efficiency under drought stress». En: *Journal of Plant Physiology* 166, págs. 1207-1217.
- Bolsa de Comercio de Rosario (2022). Estimaciones Bolsa de Comercio de Rosario.
- Buchanan, Bob B., Wilhelm Gruissem y Russel L Jones, eds. (2015). Biochemistry and Molecular Biology of Plants. 2nd. John Wiley y Sons.
- Cardoso, Amanda A. y col. (2020). «Abscisic acid biosynthesis and signaling in plants: Key targets to improve water use efficiency and drought tolerance». En: *Applied Sciences (Switzerland)* 10.18.
- Carrera, Constanza S y col. (2021). «Leaf structure and ultrastructure changes induced by heat stress and drought during seed fi lling in fi eld-grown soybean and their relationship with grain yield». En: *An Acad Bras Cienc* 93.4, pág. 20191388.
- Casadebaig, Pierre, Philippe Debaeke y Jérémie Lecoeur (2008). «Thresholds for leaf expansion and transpiration response to soil water deficit in a range of sunflower genotypes». En: *European Journal of AgronomyJ. Agronomy* 28, págs. 646-654.
- Castro-Mondragon, Jaime A. y col. (2022). «JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles». En: *Nucleic Acids Research* 50.D1, págs. D165-D173.
- Chamarthi, Siva K. y col. (2021). «Identification and Confirmation of Loci Associated With Canopy Wilting in Soybean Using Genome-Wide Association Mapping». En: *Frontiers in Plant Science* 0, pág. 1371.
- Chan, R L, F Trucco y M E Otegui (2020). «Why are second-generation transgenic crops not yet available in the market?» En: *Journal of Experimental Botany* 71.22, págs. 6876-6880.
- Chaves, M M, J Flexas y C Pinheiro (2009). «Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell». En: Annals of Botany 103, págs. 551-560.

- Chen, Wei y col. (2016). «Identification and Comparative Analysis of Differential Gene Expression in Soybean Leaf Tissue under Drought and Flooding Stress Revealed by RNA-Seq». En: *Frontiers in Plant Science* 7.July, págs. 1-19.
- Cile Vriet, Cé, Eugenia Russinova y Christophe Reuzeau (2012). «Boosting Crop Yields with Plant Steroids». En: *The Plant Cell* 24, págs. 842-857.
- Cingolani, Pablo y col. (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». En: *Fly* 6.2, pág. 80.
- Crocco, Carlos D y col. (2018). «Heterologous Expression of AtBBX21 Enhances the Rate of Photosynthesis and Alleviates Photoinhibition in Solanum tuberosum 1[OPEN]». En: *Plant Physiology* 177, págs. 369-380.
- Cutler, Sean R. y col. (2010). «Abscisic Acid: Emergence of a Core Signaling Network». En: *Annual Review* of *Plant Biology* 61.1, págs. 651-679.
- De Oliveira Ferreira, Dalton (2017). «GmNAC081: Elemento de convergência para comunicação cruzada entre sensescência e tolerância à seca.» Tesis doct. Universidade Federal de Viçosa.
- Demicheli, Josefina (2017). «Papel de las tiorredoxinas en la resistencia a la sequía en soja». Tesis doct. Universidad de Buenos Aires, págs. 1-30.
- Du, Zhou y col. (2010). «agriGO: a GO analysis toolkit for the agricultural community». En: Nucleic acids research 38.suppl_2, págs. 64-70.
- Dwivedi, Priyanka y col. (2021). «Drought Tolerant near Isogenic Lines (NILs) of Pusa 44 Developed through Marker Assisted Introgression of qDTY2.1 and qDTY3.1 Enhances Yield under Reproductive Stage Drought Stress». En: *Agriculture* 11.64.
- Earl, Hugh J. (2002). «Stomatal and non-stomatal restrictions to carbon assimilation in soybean (Glycine max) lines differing in water use efficiency». En: *Environmental and Experimental Botany* 48.3, págs. 237-246.
- (2003). «A precise gravimetric method for simulating drought stress in pot experiments». En: Crop Science 43, págs. 1868-1873.
- Farooq, M y col. (2009). «Plant drought stress : effects, mechanisms and management». En: Agronomy for Sustainable Development 29, págs. 185-212.
- Faustino, Verônica Aparecida y col. (2021). «Soybean drought-stressed plants impair Anticarsia gemmatalis (Lepidoptera: Erebidae) midgut proteolytic activity and survival Interação bioquímica entre artrópodes herbívoros e o tomate sob estresse abiótico. View project Studies on the physiology of Braz». En: *Phytoparasitica* 49, págs. 491-500.
- Fehr, Walter R y Charles E Caviness (1977). «Stages of soybean development». En:
- Felipe de Mendiburu y Muhammad Yaseen (2020). agricolae: Statistical Procedures for Agricultural Research.
- Fried, Harrison Gregory, Sruthi Narayanan y Benjamin Fallen (2019). «Evaluation of soybean [Glycine max (L.) Merr.] genotypes for yield, water use efficiency, and root traits». En: *PLoS ONE* 14.2, págs. 1-18.
- Fritsche-Neto, Roberto y Aluízio Borém (2012). Plant Breeding for Abiotic Stress Tolerance. Ed. por Roberto Fritsche-Neto y Aluízio Borém. Springer Heidelberg New York Dordrecht London, págs. 1-52; 87-102.
- Fuganti-Pagliarini, Renata y col. (2017). «Characterization of Soybean Genetically Modified for Drought Tolerance in Field Conditions». En: *Frontiers in Plant Science* 8, pág. 448.
- Gabriela González, Fernanda y col. (2019). «Field-grown transgenic wheat expressing the sunflower gene HaHB4 significantly outyields the wild type». En: *Journal of Experimental Botany* 70.5, págs. 1669-1681.
- Gallino, Juan P. y col. (2018). «A dehydration-induced eukaryotic translation initiation factor iso4G identified in a slow wilting soybean cultivar enhances abiotic stress tolerance in Arabidopsis». En: *Frontiers in Plant Science* 9.March, págs. 1-22.
- Gong, Qian chun y col. (2018). «Meta-analysis of soybean amino acid QTLs and candidate gene mining». En: *Journal of Integrative Agriculture* 17.5, págs. 1074-1084.
- GraphPad. GraphPad Prism Software. San Diego, California USA.
- Grigoriev, Igor V. y col. (2012). «The Genome Portal of the Department of Energy Joint Genome Institute». En: *Nucleic Acids Research* 40.D1, págs. D26-D32.
- Ha, Chien Van y col. (2015). «Comparative analysis of root transcriptomes from two contrasting droughtresponsive Williams 82 and DT2008 soybean cultivars under normal and dehydration conditions.» En: *Frontiers in plant science* 6, pág. 551.
- Hamzelou, Sara y col. (2020). «Wild and cultivated species of rice have distinctive proteomic responses to drought». En: *International Journal of Molecular Sciences* 21.17, págs. 1-21.

- Hansen, Kasper D. y col. (2011). «Sequencing technology does not eliminate biological variability». En: Nature biotechnology 29.7, pág. 572.
- Harb, Amal y col. (2010). «Molecular and Physiological Analysis of Drought Stress in Arabidopsis Reveals Early Responses Leading to Acclimation in Plant Growth». En: *Plant physiology* 154, págs. 1254-1271.

Hernández, Xavier (2018). Santa Fe: por la sequía, habrá lotes de soja que no se cosecharán.

- Hossain, Md Mokter y col. (2014). «Differences between soybean genotypes in physiological response to sequential soil drying and rewetting». En: *The Crop Journal* 2.6, págs. 366-380.
- Hu, Ben y col. (2020). «Overexpression of chalcone synthase gene improves avonoid accumulation and drought tolerance in tobacco». En:
- Hua, Lei y col. (2018). «Genome-Wide Identification of Drought Response Genes in Soybean Seedlings and Development of Biomarkers for Early Diagnoses». En: *Plant Molecular Biology Reporter* 36.2, págs. 350-362.
- Huang, Daiqing y col. (2008). «The relationship of drought-related gene expression in Arabidopsis thaliana to hormonal and environmental factors». En: *Journal of Experimental Botany* 59.11, págs. 2991-3007.
- Hufstetler, E. Vicki y col. (2007). «Genotypic variation for three physiological traits affecting drought tolerance in soybean». En: *Crop Science* 47.1, págs. 25-35.
- Ibarguren, Milagros (2021). *Impacto de las precipitaciones en el rendimiento de la soja en Argentina*. Inf. téc. Bolsa de Cereales de Buenos Aires.
- Jenks, Matthew A, Paul M Hasegawa y S Mohan Jain, eds. (2007). Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops. Springer.
- Jha, Saroj K, Manisha Sharma y Girdhar K Pandey (2016). «Role of Cyclic Nucleotide Gated Channels in Stress Management in Plants». En: Current Genomics 17, págs. 315-329.
- Joshi, Rohit y col. (2016). «Transcription factors and plants response to drought stress: Current understanding and future directions». En: *Frontiers in Plant Science* 7.2016JULY.
- Jumrani, Kanchan y Virender Singh Bhatia (2019). «Identification of drought tolerant genotypes using physiological traits in soybean». En: *Physiology and Molecular Biology of Plants* 25.3, págs. 697-711.
- Kanehisa, Minoru y Susumu Goto (2000). «KEGG: kyoto encyclopedia of genes and genomes». En: Nucleic acids research 28.1, págs. 27-30.
- King, C. Andy y Larry C. Purcell (2017). «Evaluation of methods for estimating transpiration response to soil drying for container-grown plants». En: Crop Science 57.4, págs. 2143-2148.
- King, C. Andy, Larry C. Purcell y Kristofor R. Brye (2009). «Differential wilting among soybean genotypes in response to water deficit». En: Crop Science 49.1, págs. 290-298.
- Kunert, Karl y Barend J Vorster (2020). «In search for drought-tolerant soybean: is the slow-wilting phenotype more than just a curiosity?» En: *Journal of Experimental Botany* 71.2, págs. 457-460.
- La Nación (2018a). Por la sequía, en soja prevén pérdidas promedio de hasta el 46% en las tierras más ricas del país.
- (2018b). Sequía: en soja ya es la peor campaña del quinquenio.
- (2018c). Sequía: Etchevehere admite que las pérdidas son cuantiosas.
- Legnaioli, Tommaso, Juan Cuevas y Paloma Mas (2009). «TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought». En: *The EMBO Journal* 28.23, págs. 3745-3757.
- Lim, Chae Woo y col. (2015). «The pepper lipoxygenase CaLOX1 plays a role in osmotic, drought and high salinity stress response». En: *Plant and Cell Physiology* 56.5, págs. 930-942.
- Liu, Fulai y col. (2005). «Stomatal control and water use efficiency of soybean (Glycine max L. Merr.) during progressive soil drying». En: *Environmental and Experimental Botany* 54, págs. 33-40.
- Liu, S y col. (2006). «Stomatal Distribution and Character Analysis of Leaf Epidermis of Jujube under Drought Stress». En: *Journal of Anhui Agricultural Sciences* 34.7, pág. 1315.
- Liu, Wenshan, Edward Sikora y Sang-Wook Park (2020). «Plant growth-promoting rhizobacterium, Panebacillus polymyxa CR1, upregulates dehydration-responsive genes, RD29A and RD29B, during priming drought tolerance in Arabidopsis». En: *Plant Physiology and Biochemistry* 156, págs. 146-154.
- Liu, Zhen y col. (2015). «UDP-Glucosyltransferase71C5, a Major Glucosyltransferase, Mediates Abscisic Acid Homeostasis in Arabidopsis». En: *Plant physiology* 167, págs. 1659-1670.
- Lopes-Caitar, Valéria S y col. (2013). «Genome-wide analysis of the Hsp20 gene family in soybean: comprehensive sequence, genomic organization and expression profile analysis under abiotic and biotic stresses». En: *BMC Genomics* 14, págs. 1-17.
- Lykke-Andersen, Søren y Torben Heick Jensen (2015). «Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes». En: *Nature Reviews Molecular Cell Biology 2015 16:11* 16.11, págs. 665-677.

- Ma, Dongyun y col. (2014). «Expression of flavonoid biosynthesis genes and accumulation of flavonoid in wheat leaves in response to drought stress». En: *Plant Physiology and Biochemistry* 80, págs. 60-66.
- Manavalan, L. P. y col. (2009). «Physiological and Molecular Approaches to Improve Drought Resistance in Soybean». En: *Plant and Cell Physiology* 50.7, págs. 1260-1276.
- McKenna, Aaron y col. (2010). «The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data». En: *Genome research* 20.9, págs. 1297-1303.
- Mi, Huaiyu y col. (2021). «PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API». En: *Nucleic Acids Research* 49.D1, págs. D394-D403.
- Mian, M. a R y col. (1996). «Molecular markers associated with water use efficiency and leaf ash in soybean». En: *Crop Science* 36.5, págs. 1252-1257.
- Milward, Sara Eve (2018). «Interrogating plant Rubisco-Rubisco activase interactions». Tesis doct. Australian National University, pág. 290.
- Ministerio de Agroindustria (2018). Campaña 17/18 Informe Especial Sequía. Inf. téc.
- Mira, Cristian (2018). El efecto se sentirá más allá de las tranqueras.
- Miransari, Mohammad, ed. (2016). Abiotic and Biotic Stresses in soybean production. Academic Press.
- Mistry, Jaina y col. (2021). «Pfam: The protein families database in 2021». En: *Nucleic Acids Research* 49.D1, págs. D412-D419.
- Msanne, Joseph y col. (2011). «Characterization of abiotic stress-responsive Arabidopsis thaliana RD29A and RD29B genes and evaluation of trans... Cite this paper Related papers». En: *Planta* 234.1, págs. 97-107.
- Oliveira Vidal, Ramon y col. (2012). «Identification of SNPs in RNA-seq data of two cultivars of Glycine max (soybean) differing in drought resistance». En: *Genetics and Molecular Biology* 35, págs. 331-334.
- Passioura, John B. (2020). «Translational research in agriculture. Can we do it better?» En: Crop and Pasture Science 71.6, págs. 517-528.
- Pathan, S. M. y col. (2014). «Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought». En: *Journal of Agronomy and Crop Science* 200.3, págs. 231-236.
- Prince, Silvas J. y col. (2015). «Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting». En: *Plant Science* 240, págs. 65-78.
- Purcell, Larry y col. (2017). «Genomewide association mapping of canopy wilting in diverse soybean genotypes». En: *Theoretical and Applied Genetics* 130, págs. 2203-2217.
- Quero, Gastón y col. (2021). «An integrative analysis of yield stability for a GWAS in a small soybean breeding population». En: *Crop Science*, págs. 1-12.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Rani, Asha, Jyoti Taunk y Ram C Yadav (2021). «Development of advance pearl millet lines tolerant to terminal drought stress using marker-assisted selection». En: *Vegetos* 35.1, págs. 63-73.
- Ray, Jeffery D. y Thomas R. Sinclair (1997). «Stomatal Closure of Maize Hybrids in Response to Drying Soil». En: Crop Science 37.3, págs. 803-807.
- Ribichich, Karina F y col. (2020). «Successful field performance in warm and dry environments of soybean expressing the sunflower transcription factor HB4». En: *Journal of Experimental Botany* 71.10, págs. 3142-3156.
- Ries, Landon L. y col. (2012). «Physiological traits contributing to differential canopy wilting in soybean under drought». En: *Crop Science* 52.1, págs. 272-281.
- Sabljic, Ivana (2020). «Bases bioquímicas y moleculares de la resistencia al ataque de las chinches en plantas de soja (Glycine max L.)» Tesis doct. Universidad de Buenos Aires.
- Sanchez, Diego H y col. (2012). «Comparative metabolomics of drought acclimation in model and forage legumes». En: *Plant, Cell and Environment* 35, págs. 136-149.
- Sanità de Toppi, Luigi y Barbara Pawlik-skowrońska (2003). Abiotic Stresses in Plants. Ed. por Luigi Sanita di Toppi y Barbara Pawlik-skowrońska. January 2003. Kluwer Academic Publishers, pág. 244.
- Satorre, Emilio H. y col. (2003). *Producción de granos: bases funcionales para su manejo*. Buenos Aires: Editorial Facultad de Agronomía.
- Schneider, Caroline A, Wayne S Rasband y Kevin W Eliceiri (2012). «NIH Image to ImageJ: 25 years of image analysis». En: *Nature Methods 2012 9:7* 9.7, págs. 671-675.
- Servaites, Jerome C (1990). Inhibition of Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase by 2-Carboxyarabinitol-1-Phosphate1. Inf. téc., págs. 867-870.
- Shekoofa, Avat y Thomas Sinclair (2018). «Aquaporin Activity to Improve Crop Drought Tolerance». En: *Cells* 7.9, pág. 123.
- Shi, Haitao y col. (2014). «Modulation of auxin content in Arabidopsis confers improved drought stress resistance». En: *Plant Physiology and Biochemistry* 82, págs. 209-217.

- Shin, Jin Hee y col. (2015). «Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance». En: *BMC Plant Biology* 15.1.
- Sinclair, Thomas R (2017). «Soybean». En: Water-Conservation Traits to Increase Crop Yields in Water- deficit Environments: Case Studies. Ed. por Thomas R. Sinclair. Springer International Publishing. Cap. Soybean, págs. 17-26.
- Sinclair, Thomas R., Maciej A. Zwieniecki y Noel Michele Holbrook (2008). «Low leaf hydraulic conductance associated with drought tolerance in soybean». En: *Physiologia Plantarum* 132.4, págs. 446-451.
- Sinclair, Thomas R. y col. (2010). «Assessment across the united states of the benefits of altered soybean drought traits». En: Agronomy Journal 102.2, págs. 475-482.
- Sinclair, Thomas R. y col. (2018). «Identification of virginia-type peanut genotypes for water-deficit conditions based on early decrease in transpiration rate with soil drying». En: *Crop Science* 58.6, págs. 2607-2612.
- Singh, Dhriti y Ashverya Laxmi (2015). «Transcriptional regulation of drought response: A tortuous network of transcriptional factors». En: *Frontiers in Plant Science* 6.0CTOBER, págs. 1-11.
- Sloane, Richard J., Robert P. Patterson y Thomas E. Carter (1990). «Field Drought Tolerance of a Soybean Plant Introduction». En: *Crop Science* 30, págs. 118-123.
- Song, Li y col. (2016). «Genome-wide transcriptome analysis of soybean primary root under varying waterdeficit conditions». En: BMC Genomics 17, págs. 1-17.
- Song, Q y col. (2013). «Development and Evaluation of SoySNP50K, a High-Density Genotyping Array for Soybean». En: *PLoS ONE* 8.1, pág. 54985.
- Song, Qijian y col. (2020). «Soybean BARCSoySNP6K: An assay for soybean genetics and breeding research». En: *The Plant Journal* 104, págs. 800-811.
- Steketee, Clinton J y col. (2020). «Genome-Wide Association Analyses Reveal Genomic Regions Controlling Canopy Wilting in Soybean». En: G3: Genes, Genomes, Genetics 10, págs. 1413-1425.
- Stolf-Moreira, Renata y col. (2011). «Transcriptional Profiles of Roots of Different Soybean Genotypes Subjected to Drought Stress». En: *Plant Molecular Biology Reporter* 29.1, págs. 19-34.
- Supek, Fran y col. (2011). «REVIGO summarizes and visualizes long lists of gene ontology terms». En: *PloS one* 6.7, e21800.
- Świeżawska-Boniecka, Brygida y col. (2021). «Cross Talk Between Cyclic Nucleotides and Calcium Signaling Pathways in Plants–Achievements and Prospects». En: Frontiers in Plant Science 12.
- Tanaka, Tsuyoshi y col. (2019). «Development of genome-wide SNP markers for barley via reference-Based RNA-seq analysis». En: *Frontiers in Plant Science* 10.May, págs. 1-9.
- Thao, Nguyen Phuong y Lam Son Phan Tran (2012). «Potentials toward genetic engineering of droughttolerant soybean». En: *Critical Reviews in Biotechnology* 32.4, págs. 349-362.
- The Gene Ontology Consortium (2021). «The Gene Ontology resource: enriching a GOld mine». En: *Nucleic acids research* 49.D1, págs. D325-D334.
- Trapnell, Cole y col. (2010). «Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation». En: *Nature biotechnology* 28.5, págs. 511-515.
- Tripathi, Prateek y col. (2016). «A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes». En: *BMC Genomics 2016 17:1* 17.1, págs. 1-22.
- Tůmová, Lenka y col. (2018). «Drought-tolerant and drought-sensitive genotypes of maize (Zea mays L.) differ in contents of endogenous brassinosteroids and their drought-induced changes». En: *PLoS ONE* 13.5, págs. 1-22.
- Villordo-Pineda, Emiliano y col. (2015). «Identification of novel drought-tolerant-associated SNPs in common bean (Phaseolus vulgaris)». En: Frontiers in Plant Science 6.JULY, págs. 1-9.
- Wang, Xin y col. (2017). «Organ-specific proteomics of soybean seedlings under flooding and drought stresses». En: *Journal of Proteomics* 162, págs. 62-72.
- Wu, Xi y col. (2015). «Influence of overexpression of 2-alkenal reductase gene on drought resistance in tobacco.» En: Acta Botanica Boreali-Occidentalia Sinica 35.6, págs. 1166-1172.
- Xie, Haicui y col. (2020). «Aphid fecundity and defenses in wheat exposed to a combination of heat and drought stress». En: *Journal of Experimental Botany* 71.9, pág. 2713.
- Xu, Zhenzhu y Guangsheng Zhou (2008). «Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass». En: *Journal of Experimental Botany* 59.12, págs. 3317-3325.

Yang, Liang y col. (2016). «Arabidopsis C3HC4-RING finger E3 ubiquitin ligase AtAIRP4 positively regulates stress-responsive abscisic acid signaling». En: *Journal of Integrative Plant Biology* 58.1, págs. 67-80.

- Ybran, Romina G. y Gabriel A. Lacelli (2016). Informe estadístico mercado de la soja. Inf. téc. INTA.
- Yu, Guangchuang y col. (2012). «clusterProfiler: an R package for comparing biological themes among gene clusters». En: *Omics : a journal of integrative biology* 16.5, págs. 284-287.

- Zhang, Hao, Wang Xiangrong y Wang Shoubing (2004). «A study on stomatal traits of Platanus acerifolia under urban stress». En: *Fu dan xue bao. Zi ran ke xue ban = Journal of Fudan University. Natural science* 43.4, págs. 651-656.
- Zhang, Qianxiang y col. (2021). «The Responses of the Lipoxygenase Gene Family to Salt and Drought Stress in Foxtail Millet (Setaria italica)». En: *life* 11.1169, págs. 1-14.
- Zhao, Tuanjie, Muqadas Aleem y Ripa Akter Sharmin (2018). «Adaptation to Water Stress in Soybean: Morphology to Genetics». En: *Plant, Abiotic Stress and Responses to Climate Change*. Ed. por Violeta Andjelkovic. IntechOpen. Cap. 3, págs. 33-68.

Anexos

Anexo A

Genes diferencialmente expresados

Cuadro A.1: Genes con expresión diferencial significativa para el GS (genotipo susceptible) en respuesta a la sequía. GeneID= ID del gen en la notación Wm82.a4.v1, padj= valor p ajustado de la expresión diferencial, log2FoldChange= expresión diferencial en logaritmo, PFAM Description= Descripción del gen según la base de datos PFAM.

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.04G222100	0.0000000	-3.36370	Pollen allergen; Rare lipoprotein A (RlpA)-like double-psi beta-barrel
Glyma.17G008500	0.0000089	-3.06830	NA
Glyma.06G131000	0.0000141	-3.02490	NA
Glyma.11G044800	0.0000000	-2.88970	Xyloglucan endo-transglycosylase (XET) C-terminus; Glycosyl hydrolases family 16
Glyma.02G109100	0.0000102	-2.85460	Rare lipoprotein A (RlpA)-like double-psi beta-barrel; Pollen allergen
Glvma.15G057600	0.0000002	-2.85020	NA
Glyma.06G143300	0.0000135	-2.82070	Pollen allergen; Rare lipoprotein A (RlpA)-like double-psi beta-barrel
Glyma.12G035700	0.0001186	-2.70090	Auxin responsive protein
Glyma.12G195500	0.0002778	-2.69990	Peroxidase
Glyma.10G251400	0.0000029	-2.61920	P21-Rho-binding domain
Glyma.12G233400	0.0000018	-2.61340	PF03005
Glyma.03G233800	0.0001828	-2.53940	F-box domain
Glyma.18G269800	0.0010757	-2.53430	NA
Glyma.13G084000	0.0023511	-2.36660	Oxidoreductase NAD-binding domain ; Mo-co oxidoreductase dimerisation domain; Oxidoreductase molybdopterin binding domain; Cytochrome b5-like Heme/Steroid binding domain; Oxidoreductase FAD-binding domain
Glyma.07G032000	0.0000001	-2.34690	Kelch motif; F-box domain
Glyma.10G196500 Glyma.16G007700 Glyma.18G289100 Glyma.19G215500 Glyma.02G049200	$\begin{array}{c} 0.0014386\\ 0.0008845\\ 0.0000001\\ 0.0022513\\ 0.0044643 \end{array}$	-2.29980 -2.29710 -2.25270 -2.25100 -2.22710	Protein of unknown function (DUF1635) Cupin Hsp70 protein NA Auxin responsive protein
Glyma 15G079300	0.0021115	-2 22700	GDSL-like Lipase/Acylhydrolase
Glyma.07G132600	0.0004313	-2.21780	Rare lipoprotein A (RIpA)-like double-psi
Glyma 08G171000	0.0103720	-2 15680	NA
Glyma 01G113100	0.0106530	-2.15000	NA
Glyma.13G046900	0.0104030	-2.12990	NA
Glyma 16G173400	0.0000092	-2 12270	NA
Glyma.04G036800	0.0015336	-2.12100	Protein of unknown function, DUF617
Glyma.02G132100	0.0007544	-2.11690	Nitrite and sulphite reductase 4Fe-4S domain;
-			Nitrite/Sulfite reductase ferredoxin-like half
			domain

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G007100 Glyma.11G105800	$0.0130670 \\ 0.0027132$	-2.11560 -2.11170	Auxin responsive protein DnaJ domain
Glyma.06G008200 Glyma.20G058000 Glyma.14G195000 Glyma.06C026300	$\begin{array}{c} 0.0131120\\ 0.0154840\\ 0.0154840\\ 0.0120810 \end{array}$	-2.08770 -2.08190 -2.08140 2.07800	Glycosyl hydrolase family 9 Hydroxymethylglutaryl-coenzyme A reductase Asparagine synthase Clycosyl ovidese N terminus: Domein of
Glyma.04G167800	0.0120810	-2.07890	unknown function (DUF1929) Pollen allergen; Rare lipoprotein A (RlpA)-like double-psi beta-barrel
Glyma.19G002100 Glyma.19G190200	0.0187920 0.0129390	-2.02150 -2.00940	Myb-like DNA-binding domain Protein tyrosine kinase; Leucine Rich Repeat;
Glyma.20G141800	0.0196730	-2.00840	Leucine rich repeat N-terminal domain Glycosyl hydrolases family 16; Xyloglucan
Glyma.04G169600	0.0218160	-1.99220	endo-transglycosylase (XET) C-terminus Gibberellin regulated protein Duridina puelostida disulphida avideraduatesa
Oryma.000127400	0.0010557	-1.99030	Glutamate synthase central domain; Pyridine nucleotide-disulphide oxidoreductase; GXGXG motif; Glutamine amidotransferases class-II; Conserved region in glutamate synthase
Glyma.06G089300	0.0015033	-1.99000	NA Cysteine rich secretory protein family
Glyma.08G203000 Glyma.15G271000	0.0210720 0.0125250 0.0094374	-1.98980 -1.98840 -1.98810	Major intrinsic protein Glycosyl hydrolase family 1
Glyma.16G138500	0.0223020	-1.98510	NA
Glyma.12G225700 Glyma.17G100900 Glyma.11G225200	$\begin{array}{c} 0.0233890 \\ 0.0236020 \\ 0.0007187 \end{array}$	-1.97850 -1.97570 -1.97310	NA NA PRONE (Plant-specific Rop nucleotide
Glyma.06G044400 Glyma.13G075700	0.0002188 0.0023511	-1.97280 -1.97010	exchanger) Gibberellin regulated protein Protease inhibitor/seed storage/LTP family
Glyma.06G283800 Glyma 07G132700	0.0288350 0.0198870	-1.93790 -1 92830	PF03005 NA
Glyma.13G348000 Glyma.10G107100	0.0227500 0.0218160	-1.92390 -1.92250	Domain of unknown function (DUF3511) NAD-dependent glycerol-3-phosphate dehydrogenase C-terminus; NAD-dependent glycerol 3 phosphate dehydrogenase N terminus
Glyma.10G029600	0.0184040	-1.91420	NA
Glyma.14G050400	0.0323100	-1.90810	Partial alpha/beta-hydrolase lipase region; alpha/beta hydrolase fold
Glyma.17G164800 Glyma.08G269600	0.0012271 0.0093468	-1.88190 -1.87940	NA Strictosidine synthase
Glyma.10G013900	0.0360300	-1.86350	Glycosyl hydrolase family 3 N terminal domain; Glycosyl hydrolase family 3 C-terminal domain Myb like DNA binding domain
Glyma.14G126200	0.0405190	-1.85580	Protein of unknown function (DUF642)
Glyma.13G200300 Glyma.03G063900	0.0434440	-1.85000	NA Transmembrane amino acid transporter protein
Glyma.08G273600 Glyma.11G225700	0.020000000000000000000000000000000000	-1.83530	NA Protein of unknown function (DUF620)
Glyma.02G112900	0.0009665	-1.82170	Protein of unknown function, DUF538
Glyma.17G103300 Glyma.18G116400 Glyma.08G033000 Glyma.08G094300	$\begin{array}{c} 0.0442310\\ 0.0080006\\ 0.0434980\\ 0.0302050\end{array}$	-1.81580 -1.81570 -1.81420 -1.81310	Weak chloroplast movement under blue light Glycosyl hydrolases family 28 Pectinesterase NA
Glyma.18G148200 Glyma.18G291700 Glyma.12G228100	0.0183050 0.0061374	-1.81130 -1.80680	UDP-glucoronosyl and UDP-glucosyl transferase Heavy-metal-associated domain
Glyma.12G228100 Glyma.11G160000	0.0260180	-1.80000 -1.79960	Rare lipoprotein A (RlpA)-like double-psi
Glyma.15G018100	0.0011775	-1.79100	beta-barrel; Pollen allergen Major intrinsic protein
Glyma.17G113000	0.0329820	-1.78730	Plastocyanin-like domain
			(Conntinua en la pág. siguiente)

Cuadro A.1: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.19G050800	$\begin{array}{c} 0.0191790\\ 0.0361750\\ 0.0280660\end{array}$	-1.78660	GDSL-like Lipase/Acylhydrolase
Glyma.09G221000		-1.78640	Auxin responsive protein
Glyma.18G116800		-1.78470	Leucine Rich Repeat; Leucine rich repeat
Glyma.13G293500	0.0156990	-1.78080	Fasciclin domain
Glyma.16G138400 Glyma.01G196800	0.0256370 0.0427870	-1.77650 -1.77360	NA Xyloglucan endo-transglycosylase (XET) C-terminus; Glycosyl hydrolases family 16
Glyma.11G034000	$\begin{array}{c} 0.0183050\\ 0.0398290\\ 0.0156990 \end{array}$	-1.76420	Major intrinsic protein
Glyma.05G191000		-1.75660	NA
Glyma.08G345100		-1.75520	NA
Glyma.19G258400	0.0017382	-1.73670	NA
Glyma.15G012100	0.0000145	-1.72880	2OG-Fe(II) oxygenase superfamily
Glyma.13G249100	0.0150260	-1.72770	NA
Glyma.06G039100	0.0108650	-1.71210	NA
Glyma.18G272300	0.0010182	-1.70760	Cytochrome P450
Glyma.15G267200	$\begin{array}{c} 0.0465620\\ 0.0004927\\ 0.0010182\\ 0.0428440\\ 0.0428440 \end{array}$	-1.70130	Plant protein of unknown function (DUF946)
Glyma.15G026000		-1.69960	Domain of unknown function (DUF3511)
Glyma.08G162100		-1.69810	Eukaryotic aspartyl protease
Glyma.02G189000		-1.69360	Ycf1
Glyma.19G037900		-1.68530	bZIP transcription factor
Glyma.06G298600	0.0210720	-1.68010	Wound-induced protein
Glyma.14G073601 Glyma.11G146500 Glyma.16G056300 Glyma.10G279200	$\begin{array}{c} 0.0398210\\ 0.0001010\\ 0.0068459\\ 0.0428440 \end{array}$	-1.67700 -1.67620 -1.67270 -1.66610	Major intrinsic protein Ribonucleotide reductase, small chain Family of unknown function (DUF716)
Glyma.02G242700	$\begin{array}{c} 0.0210720\\ 0.0414900\\ 0.0335730\\ 0.0092795\\ 0.0401760 \end{array}$	-1.66100	NA
Glyma.02G275900		-1.65120	NA
Glyma.19G047000		-1.62910	Ribulose bisphosphate carboxylase, small chain
Glyma.01G081300		-1.59150	Glycosyl hydrolases family 31
Glyma.01G032000		-1.59090	NHL repeat
Glyma.04G183700	$\begin{array}{c} 0.0265070\\ 0.0104420\\ 0.0427870\\ 0.0189220\\ 0.002200\end{array}$	-1.58950	Dof domain, zinc finger
Glyma.10G094600		-1.58640	NA
Glyma.16G066700		-1.58010	NA
Glyma.04G010300		-1.55150	bZIP transcription factor
Glyma.09G126400	$\begin{array}{c} 0.0002208\\ 0.0323100\\ 0.0081724\\ 0.0134700\\ 0.0187920\\ 0.0007039 \end{array}$	-1.54900	NA
Glyma.08G271100		-1.54770	Zinc finger, C3HC4 type (RING finger)
Glyma.17G108200		-1.53570	Glycosyl hydrolase family 79, N-terminal domain
Glyma.08G150800		-1.52520	Tetrapyrrole (Corrin/Porphyrin) Methylases
Glyma.18G251800		-1.52000	NA
Glyma.08G139800		-1.50700	NA
Glyma.20G115600	$\begin{array}{c} 0.0000322\\ 0.0101350\\ 0.0001186\\ 0.0439330\\ 0.0330440 \end{array}$	-1.50650	B-box zinc finger; CCT motif
Glyma.18G284900		-1.50400	BAG domain; IQ calmodulin-binding motif
Glyma.10G274300		-1.49810	CCT motif; B-box zinc finger
Glyma.12G207600		-1.49600	Fasciclin domain
Glyma.15G067200		-1.48740	V-ATPase subunit H
Glyma.18G127400 Glyma.18G153500	0.0015033 0.0428440	-1.48050 -1.46110	Thioredoxin Pollen allergen; Rare lipoprotein A (RlpA)-like double-psi beta-barrel
Glyma.08G060300 Glyma.19G079600 Glyma.08G270300	$\begin{array}{c} 0.0000129\\ 0.0301190\\ 0.0401760\end{array}$	-1.45530 -1.45160 -1.44680	PB1 domain NA ABC transporter transmembrane region; ABC transporter
Glyma.03G014300	$\begin{array}{c} 0.0330440\\ 0.0265070\\ 0.0478890\\ 0.0221130\\ 0.0448240 \end{array}$	-1.44570	NA
Glyma.02G059600		-1.44400	EF hand
Glyma.03G001000		-1.44020	Protein of unknown function (DUF1191)
Glyma.13G045300		-1.43610	NA
Glyma.10G050600		-1.42100	NA
Glyma.14G212100	0.0102010	-1.41770	Protein kinase domain; Leucine rich repeat
Glyma.18G276800 Glyma.06G110400 Glyma.11G039400 Glyma.19G112700	$\begin{array}{c} 0.0024191\\ 0.0000662\\ 0.0402130\\ 0.0378370 \end{array}$	-1.41000 -1.40590 -1.39030 -1.38750	N-terminal domain Transmembrane amino acid transporter protein 'Cold-shock' DNA-binding domain; Zinc knuckle Glycosyl hydrolase family 14 NA

Cuadro A.1: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.09G149200	$\begin{array}{c} 0.0409720\\ 0.0184290\\ 0.0030223\\ 0.0478420\\ 0.0004741 \end{array}$	-1.38070	2OG-Fe(II) oxygenase superfamily
Glyma.05G182100		-1.37810	NA
Glyma.08G178200		-1.36990	NA
Glyma.11G133700		-1.36770	Myb-like DNA-binding domain
Glyma.16G040400		-1.35980	NA
Glyma.11G049300 Glyma.10G180100 Glyma.06G169100 Glyma.12G210600 Glyma.06G213300	$\begin{array}{c} 0.0164600\\ 0.0198870\\ 0.0134700\\ 0.0061374\\ 0.0427870 \end{array}$	-1.34300 -1.33730 -1.33650 -1.33130 -1.32540	Protein of unknown function DUF260 AUX/IAA family Phosphoesterase family Phosphoenolpyruvate carboxylase PAZ domain; Domain of unknown function (DUF1785); Piwi domain
Glyma.16G154200 Glyma.08G332900	$0.0377050 \\ 0.0183050$	-1.31970 -1.31900	Syntaxin; SNARE domain Hsp90 protein; Histidine kinase-, DNA gyrase B-, and HSP90_like ATPase
Glyma.11G133800	$\begin{array}{c} 0.0023511\\ 0.0486860\\ 0.0062350\end{array}$	-1.31420	Sodium:sulfate symporter transmembrane region
Glyma.16G068900		-1.30770	GDSL-like Lipase/Acylhydrolase
Glyma.06G185300		-1.30090	NA
Glyma.08G210700 Glyma.06G195000	0.0023511 0.0081949	-1.29320 -1.29170	Kelch motif; F-box domain Rare lipoprotein A (RlpA)-like double-psi beta-barrel: Pollen allergen
Glyma.18G238000	$\begin{array}{c} 0.0108030\\ 0.0007039\\ 0.0455150\end{array}$	-1.28960	Heavy-metal-associated domain
Glyma.09G207100		-1.28900	NA
Glyma.20G102200		-1.27920	GRAM domain
Glyma.13G242800	$\begin{array}{c} 0.0118830 \\ 0.0333760 \end{array}$	-1.26820	NA
Glyma.06G077600		-1.25700	Ion transport protein; Cyclic nucleotide-binding
Glyma.11G221000	$\begin{array}{c} 0.0022449\\ 0.0475970\\ 0.0130670 \end{array}$	-1.23930	NA
Glyma.08G303800		-1.23790	NA
Glyma.09G242900		-1.23490	Protein kinase domain
Glyma.11G176000	$\begin{array}{c} 0.0015933\\ 0.0044125\\ 0.0023788\\ 0.0323100\\ 0.0424090\end{array}$	-1.20630	Thylakoid soluble phosphoprotein TSP9
Glyma.11G014800		-1.19900	AP2 domain
Glyma.17G228400		-1.19660	NA
Glyma.14G218700		-1.19220	NA
Glyma.13G063000 Glyma.06G051800 Glyma.18G040900 Glyma.12G232000	0.0434980 0.0366750 0.0493900 0.0134700	-1.17930 -1.14430 -1.14210 -1.14010	Polyketide cyclase / dehydrase and lipid transport NUDIX domain ATP synthase alpha/beta family, nucleotide-binding domain; ATP synthase alpha/beta chain, C terminal domain; ATP
Glyma.05G219900	0.0475970	-1.13800	NA
Glyma.12G056000	0.0095180	-1.13680	Protein kinase domain
Glyma.07G263700 Glyma.20G112600 Glyma.13G046200	$\begin{array}{c} 0.0056128\\ 0.0260180\\ 0.0057071 \end{array}$	-1.11280 -1.10350 -1.08410	NA Pectate lyase Ribulose-1,5-bisphosphate carboxylase small subunit; Ribulose bisphosphate carboxylase,
Glyma.18G297500 Glyma.15G051500	$0.0335730 \\ 0.0146510$	-1.07660 -1.06840	Regulator of Vps4 activity in the MVB pathway Glycosyl hydrolase family 3 N terminal domain; Glycosyl hydrolase family 3 C-terminal domain
Glyma.18G028400	$\begin{array}{c} 0.0123930 \\ 0.0042926 \end{array}$	-1.06730	NA
Glyma.13G093700		-1.03870	Protein kinase domain; Universal stress protein
Glyma.18G206400	0.0493770	-1.03610	Protein kinase domain; Leucine Rich Repeat; Leucine rich repeat N-terminal domain
Glyma.17G096600	0.0335730 0.0268710	-1.03340	HPP family
Glyma.08G177400		-1.02920	Sodium:sulfate symporter transmembrane region
Glyma.12G058000 Glyma.08G014200	0.0129390 0.0251800	-1.02110 -1.02020	Sodium:sulfate symporter transmembrane region Tubulin/FtsZ family, GTPase domain; Tubulin C-terminal domain
Glyma.20G002000	$\begin{array}{c} 0.0224340\\ 0.0100380\\ 0.0486890 \end{array}$	-1.01780	Major Facilitator Superfamily
Glyma.19G244600		-1.00210	NAD dependent epimerase/dehydratase family
Glyma.10G260900		-1.00010	NA

Cuadro A.1: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.12G169600	$\begin{array}{c} 0.0146510\\ 0.0225070\\ 0.0184040 \end{array}$	-0.99133	2Fe-2S iron-sulfur cluster binding domain
Glyma.05G160900		-0.98658	NA
Glyma.03G086200		-0.97792	Tetratricopeptide repeat; FKBP-type
Glyma.19G190800	0.0017382	-0.97782	Redoxin
Glyma.01G203400	0.0087495	-0.97638	Glycosyl hydrolase family 14
Glyma.06G024100	0.0017382	-0.96888	Tubulin/FtsZ family, GTPase domain; Tubulin
Glyma.10G166600 Glyma.19G041375 Glyma.19G047100 Glyma.19G046800	$\begin{array}{c} 0.0244880\\ 0.0018262\\ 0.0010046\\ 0.0053217 \end{array}$	-0.95963 -0.95670 -0.95105 -0.93067	C-terminal domain Protein kinase domain NA NA Ribulose bisphosphate carboxylase, small chain; Ribulose-1,5-bisphosphate carboxylase small subunit
Glyma.11G039700	$\begin{array}{c} 0.0100380\\ 0.0339930\\ 0.0242440\\ 0.0183050\\ 0.0330060 \end{array}$	-0.92980	NA
Glyma.15G261900		-0.92312	Phosphoglycerate kinase
Glyma.08G126800		-0.89501	2Fe-2S iron-sulfur cluster binding domain
Glyma.08G304200		-0.89356	PsbP
Glyma.17G084300		-0.89311	Rhodanese-like domain
Glyma.19G147900 Glyma.06G174200 Glyma.14G011600	$\begin{array}{c} 0.0413490 \\ 0.0109590 \\ 0.0306850 \end{array}$	-0.88246 -0.88100 -0.87502	NA haloacid dehalogenase-like hydrolase Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase; Hsp90 protein
Glyma.03G078700	$0.0127280 \\ 0.0179340$	-0.87320	Major intrinsic protein
Glyma.20G013600		-0.87065	DnaJ domain
Glyma.19G046600	0.0108650	-0.86127	Ribulose bisphosphate carboxylase, small chain; Ribulose-1,5-bisphosphate carboxylase small
Glyma.17G046500	0.0290240	-0.86010	AhpC/TSA family
Glyma.07G208750	0.0286920	-0.85180	NA
Glyma.12G089100	0.0079717	-0.84168	Myb-like DNA-binding domain
Glyma.12G089100	0.0079717	-0.84168	Myb-like DNA-binding domain
Glyma.03G194800 Glyma.17G037900	0.0302050 0.0428440	-0.84159 -0.84062	alpha/beta hydrolase fold Pyridine nucleotide-disulphide oxidoreductase, dimerisation domain; Pyridine nucleotide-disulphide oxidoreductase; Pyridine pugleotide disulphide oxidoreductase
Glyma.14G058200 Glyma.16G217900 Glyma.10G255500	$\begin{array}{c} 0.0231260\\ 0.0153520\\ 0.0168970 \end{array}$	-0.83149 -0.82432 -0.81822	Helix-loop-helix DNA-binding domain Glycosyl hydrolase family 14 Tubulin C-terminal domain; Tubulin/FtsZ family, GTPase domain
Glyma.10G040600	$\begin{array}{c} 0.0339930\\ 0.0158560\\ 0.0288350\\ 0.0403050\\ 0.0398840 \end{array}$	-0.81609	Psb28 protein
Glyma.11G111400		-0.79276	Fructose-bisphosphate aldolase class-I
Glyma.13G051000		-0.78128	ATP-sulfurylase
Glyma.13G113200		-0.77843	AhpC/TSA family
Glyma.01G010200		-0.77803	Phosphoribulokinase / Uridine kinase family
Glyma.13G164000	$\begin{array}{c} 0.0062350\\ 0.0401760\\ 0.0080006\\ 0.0130670\\ 0.0196730\\ \end{array}$	-0.77308	NA
Glyma.17G256102		-0.76755	NA
Glyma.02G165000		-0.76186	NA
Glyma.13G127200		-0.75688	NA
Glyma.11G246600		-0.75115	Hsp20/alpha crystallin family
Glyma.20G065300 Glyma.08G046400	$0.0486860 \\ 0.0480660$	-0.74045 -0.73872	Exostosin family Dynamin central region; Dynamin family; Dynamin GTPace affector domain
Glyma.01G229800	$\begin{array}{c} 0.0428440 \\ 0.0144550 \\ 0.0129390 \end{array}$	-0.73097	NA
Glyma.01G181200		-0.71095	SHNi-TPR
Glyma.10G251500		-0.70033	Thi4 family
Glyma.05G168400	$\begin{array}{c} 0.0191790 \\ 0.0103720 \\ 0.0194380 \end{array}$	-0.69831	2Fe-2S iron-sulfur cluster binding domain
Glyma.01G233600		-0.69822	NA
Glyma.11G110000		-0.68942	NAD dependent epimerase/dehydratase family;
Glyma.07G182300	$0.0156990 \\ 0.0227500$	-0.68257	Fructose-1-6-bisphosphatase
Glyma.03G073740		-0.67822	NA

Cuadro A.1: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G311600	$0.0225070 \\ 0.0367510$	-0.66611	PPR repeat
Glyma.02G038800		-0.65697	ABC transporter transmembrane region; ABC
Glyma.14G029000 Glyma.17G133600 Glyma.16G220200	0.0403050 0.0299930 0.0333760	-0.64568 -0.63587 -0.58476	transporter Protein kinase domain Tetratricopeptide repeat Protein kinase domain
Glyma.10G166400	$\begin{array}{c} 0.0493900\\ 0.0475970\\ 0.0416080\\ 0.0416080\\ 0.0227500 \end{array}$	-0.57516	PPR repeat
Glyma.18G030200		-0.54523	NA
Glyma.03G219300		0.47700	bZIP transcription factor
Glyma.03G219300		0.47700	bZIP transcription factor
Glyma.13G245900		0.62511	B3 DNA binding domain; CW-type Zinc Finger
Glyma.12G132400	$\begin{array}{c} 0.0331580\\ 0.0296850\\ 0.0409720\\ 0.0259730\\ 0.0063915 \end{array}$	0.67733	TIR domain; NB-ARC domain
Glyma.09G198300		0.68969	HMG (high mobility group) box
Glyma.16G140800		0.73159	Leucine Rich Repeat
Glyma.11G155677		0.73459	NA
Glyma.17G255602		0.75070	NA
Glyma.18G017300	$\begin{array}{c} 0.0431550\\ 0.0290240\\ 0.0493900\\ 0.0439790\\ 0.0366750 \end{array}$	0.76404	TLD
Glyma.17G068300		0.78362	Zinc finger, C3HC4 type (RING finger)
Glyma.14G217100		0.80812	Core histone H2A/H2B/H3/H4
Glyma.02G280000		0.81972	Domain of unknown function (DUF702)
Glyma.13G208900		0.83431	NA
Glyma.08G126000 Glyma.09G195400 Glyma.17G030600	0.0134050 0.0331580 0.0242440	0.85171 0.85707 0.87955	Ankyrin repeat; Protein tyrosine kinase Diacylglycerol acyltransferase Response regulator receiver domain; Myb-like DNA-binding domain
Glyma.05G109500	0.0323100	0.88258	domain
Glyma.13G246800	0.0475970	0.88380	NA
Glyma.13G147000	0.0281690	0.89253	Core histone H2A/H2B/H3/H4
Glyma.03G070300	0.0109590	0.90688	Serine carboxypeptidase
Glyma.14G078300	0.0365000	0.90691	NA
Glyma.18G107000	0.0236020	0.91583	Phosphoribosyl transferase domain
Glyma.15G186400	0.0409280	0.93402	NA
Glyma.08G271900	$\begin{array}{c} 0.0281690\\ 0.0139090\\ 0.0399120\\ 0.0397440\\ 0.0158990 \end{array}$	0.94725	Helix-loop-helix DNA-binding domain
Glyma.01G096600		0.95956	Helix-loop-helix DNA-binding domain
Glyma.09G250200		0.96134	NUDIX domain
Glyma.17G090500		0.97944	Helix-loop-helix DNA-binding domain
Glyma.04G184400		0.99303	UvrB/uvrC motif; Protein of unknown function
Glyma.13G043900 Glyma.09G044500	0.0361750 0.0139090	$0.99405 \\ 1.01170$	(DUF525) Phosphate-induced protein 1 conserved region ABC transporter transmembrane region; ABC
Glyma.12G181500 Glyma.17G109300 Glyma.13G244600	0.0017001 0.0156990 0.0475970	$1.01590 \\ 1.02290 \\ 1.03960$	transporter NA NA NA
Glyma.04G228000	$\begin{array}{c} 0.0105840\\ 0.0044643\\ 0.0428440\\ 0.0106530\\ 0.0052725 \end{array}$	1.04950	WD domain, G-beta repeat
Glyma.19G187100		1.05650	UDP-glucoronosyl and UDP-glucosyl transferase
Glyma.05G055100		1.05670	NA
Glyma.01G230200		1.08320	NA
Glyma.18G014200		1.09130	ANTH domain
Glyma.07G027400	$\begin{array}{c} 0.0216240\\ 0.0433540\\ 0.0183050\\ 0.0019622\\ 0.0192860 \end{array}$	1.10000	NA
Glyma.10G062451		1.10780	NA
Glyma.20G184800		1.11280	NA
Glyma.03G187000		1.11580	UDP-glucoronosyl and UDP-glucosyl transferase
Glyma.07G213700		1.12060	Thioredoxin
Glyma.01G018400 Glyma.14G205600 Glyma.18G003100 Glyma.03G157432 Glyma.08G091900	$\begin{array}{c} 0.0002092\\ 0.0064595\\ 0.0083120\\ 0.0180960\\ 0.0007532 \end{array}$	$\begin{array}{c} 1.15340 \\ 1.16040 \\ 1.16050 \\ 1.16650 \\ 1.16720 \end{array}$	Helix-loop-helix DNA-binding domain AP2 domain UDP-glucoronosyl and UDP-glucosyl transferase NA K+ potassium transporter
Glyma.06G130200	$\begin{array}{c} 0.0374840\\ 0.0263640\\ 0.0192860\\ 0.0242440 \end{array}$	1.17570	NA
Glyma.16G020900		1.18680	NA
Glyma.07G023700		1.19030	NAD dependent epimerase/dehydratase family
Glyma.10G112600		1.19040	NA

Cuadro A.1: Continuación
Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G287000	0.0100010	1.20530	Glutaredoxin
Glyma.11G190500 Glyma.03G187500 Glyma.06G312300	$\begin{array}{c} 0.0129670\\ 0.0232240\\ 0.0009665\end{array}$	1.21630 1.22690 1.22690	NA UDP-glucoronosyl and UDP-glucosyl transferase ABC transporter transmembrane region; ABC
Glyma.13G319400 Glyma.09G204500	0.0009665 0.0448380	1.24030 1.24680	transporter NA Helix-loop-helix DNA-binding domain
Glyma.15G254000 Glyma.08G022900 Glyma.18G127333 Glyma.11G242600 Glyma.03G258100	$\begin{array}{c} 0.0414900\\ 0.0054223\\ 0.0163220\\ 0.0244880\\ 0.0150260\end{array}$	1.25990 1.26630 1.26820 1.28570 1.28580	No apical meristem (NAM) protein Zinc finger, C3HC4 type (RING finger) NA NA Helix-loop-helix DNA-binding domain
Glyma.11G027400 Glyma.18G069802 Glyma.13G243400 Glyma.04G170100 Glyma.06G099000	$\begin{array}{c} 0.0139550\\ 0.0220300\\ 0.0005571\\ 0.0163220\\ 0.0026873 \end{array}$	$\begin{array}{c} 1.29510 \\ 1.30140 \\ 1.32500 \\ 1.32810 \\ 1.32890 \end{array}$	NA NA Enoyl-CoA hydratase/isomerase family Myb-like DNA-binding domain Glutathione S-transferase, N-terminal domain
Glyma.03G120700 Glyma.09G082900 Glyma.14G089063 Glyma.12G203100 Glyma.11G229100	$\begin{array}{c} 0.0493900\\ 0.0331580\\ 0.0251800\\ 0.0134700\\ 0.0439330\end{array}$	1.35480 1.35520 1.36480 1.36650 1.36780	VQ motif UDP-glucoronosyl and UDP-glucosyl transferase NA AP2 domain NA
Glyma.19G198100 Glyma.19G117200 Glyma.03G174300 Glyma.03G185100	0.0292260 0.0194380 0.0016789 0.0121810	1.38080 1.38210 1.40040 1.40080	NA Sodium/calcium exchanger protein Exo70 exocyst complex subunit Nodulin-like
Glyma.17G063200 Glyma.12G051100 Glyma.04G255400 Glyma.03G053500 Glyma.13G279900 Glyma.05G128400	$\begin{array}{c} 0.0003313\\ 0.0367510\\ 0.0193980\\ 0.0493770\\ 0.0191790\\ 0.0268710\end{array}$	1.40130 1.44650 1.45490 1.47780 1.48530 1.48530	NA Cellulose synthase NA No apical meristem (NAM) protein NA
Glyma.08G082400 Glyma.08G082400 Glyma.06G190200	$\begin{array}{c} 0.0205820\\ 0.0205820\\ 0.0191790 \end{array}$	1.48800 1.48800 1.50810	WRKY DNA -binding domain WRKY DNA -binding domain Homeobox associated leucine zipper; Homeobox
Glyma.11G075500 Glyma.04G061500	0.0040462 0.0046722	1.51990 1.53070	domain linker histone H1 and H5 family NAF domain; Protein kinase domain
Glyma.19G174600	0.0399120	1.53140	Plant protein 1589 of unknown function (A thal
Glyma.11G053400 Glyma.15G116800 Glyma.13G347800	$\begin{array}{c} 0.0049275\\ 0.0059006\\ 0.0455150\end{array}$	1.53290 1.53390 1.53520	UDP-glucoronosyl and UDP-glucosyl transferase Domain of unknown function (DUF588)
Glyma 13G32300	0.0253900	1.53920	Ammonium Transporter Family
Glyma.20G094500 Glyma.06G290000 Glyma.08G033800 Glyma.06G114000	$\begin{array}{c} 0.0001974\\ 0.0022712\\ 0.0302050\\ 0.0314390\\ 0.0134700 \end{array}$	1.56170 1.57830 1.58090 1.58390	Glycosyl transferase family 8 AP2 domain Protein phosphatase 2C No apical meristem (NAM) protein
Glyma.16G159300 Glyma.13G030300 Glyma.10G198200 Glyma.08G057900 Glyma.13G283100	$\begin{array}{c} 0.0428440\\ 0.0156990\\ 0.0280660\\ 0.0239400\\ 0.0047187\end{array}$	$\begin{array}{c} 1.58860\\ 1.59570\\ 1.59630\\ 1.59870\\ 1.60200\end{array}$	Leucine Rich Repeat Lipoxygenase; PLAT/LH2 domain NA NA S-locus glycoprotein family; Protein tyrosine kinase: D-mannose binding lectin: PAN-like
			domain
Glyma.08G071300 Glyma.14G171500 Glyma.16G202000 Glyma.04G040000 Glyma.03G210500	$\begin{array}{c} 0.0151540\\ 0.0194040\\ 0.0022294\\ 0.0268740\\ 0.0398520\end{array}$	$\begin{array}{c} 1.61370 \\ 1.62140 \\ 1.62310 \\ 1.62740 \\ 1.62800 \end{array}$	Phosphate-induced protein 1 conserved region AP2 domain Embryo-specific protein 3, (ATS3) NA Glutaredoxin
Glyma.03G044500 Glyma.05G182300 Glyma.02G118500	$\begin{array}{c} 0.0338320\\ 0.0248160\\ 0.0224340\\ 0.0103720\end{array}$	1.62800 1.62810 1.64230 1.64970	Dirigent-like protein NA Protein of unknown function (DUF1675)

Cuadro A.1: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.02G104600 Glyma.13G247200	$\begin{array}{c} 0.0063312 \\ 0.0312600 \end{array}$	1.66710 1.67750	UDP-glucoronosyl and UDP-glucosyl transferase Myb-like DNA-binding domain
Glyma.07G027200	0.0011630	1.68520	Helix-loop-helix DNA-binding domain
Glyma.08G141000	0.0102010	1.69540	Histone-like transcription factor (CBF/NF-Y) and
Glyma.08G141000	0.0104190	1.69540	Histone-like transcription factor (CBF/NF-Y) and
Glyma.06G043000	0.0096787	1.70060	Asparaginase
Glyma.09G225400	0.0281690	1.70820	FAD binding domain ; Cytokinin dehydrogenase
Glyma.12G175800	0.0494840	1.71300	NA Hydroxymethylglutaryl coopgyme A synthese N
Glyma.010215500	0.0010400	1./1/50	terminal; Hydroxymethylglutaryl-coenzyme A
Glyma.06G097300	0.0161500	1.72700	Transferase family
Glyma.12G226600	0.0374840	1.73770	AP2 domain
Glyma.08G340300 Glyma.11G000500 Glyma.07G262700	$\begin{array}{c} 0.0162510\\ 0.0042926\\ 0.0434980\end{array}$	1.74020 1.74880 1.75870	Protein kinase domain; Protein tyrosine kinase UDP-glucoronosyl and UDP-glucosyl transferase WRKY DNA -binding domain
Glyma.03G196300 Glyma.18G246100	$0.0415220 \\ 0.0283540$	1.76020 1.76380	NA NA
Glyma.09G145600	0.0074409	1.76650	Leucine Rich Repeat; Leucine rich repeat
Glvma.13G275800	0.0000646	1.77060	N-terminal domain Plastocyanin-like domain
Glyma.15G008300	0.0019814	1.77180	NA AP2 domain
Glyma.13G298000	0.0100330	1.78760	Protein of unknown function (DUF688)
Glyma.15G099600	0.0238040	1.79890	NA NAE domain: Protain kinasa domain
Glyma.020234800	0.0390940	1.80510	Family of unknown function (DUF706)
Glyma.18G035000 Glyma.10G047100	0.0015933 0.0360770	1.81680 1.82080	Protein phosphatase 2C Uncharacterised protein family (UPF0014)
Glyma.06G183800	0.0139550	1.82850	U-box domain
Glyma.17G257200	0.0016407	1.84050	NA Possible lysine decarboxylase
Glyma.14G212700 Glyma.19G009900	0.0027283	1.85040	Eukaryotic aspartyl protease Sugar efflux transporter for intercellular exchange
Glyma.05G204100	0.0374640	1.87600	NA
Glyma.11G222600 Glyma.03G181000	0.0024674	1.91060	Protein phosphatase 2C
Glyma.15G161500	0.0150260	1.92580	Protein tyrosine kinase
Glyma.04G1/4533	0.0130120	1.95040	NA Destain choordataan 20
Glyma.19G069200 Glyma.18G216800	0.0097743	1.96690	NA
Glyma.06G062100 Glyma.10G010100	0.0000521 0.0027132	1.97630 1.98170	Protein kinase domain; NAF domain NA
Glyma.15G252000	0.0053217	1.99250	Glutathione S-transferase, N-terminal domain;
C1 04C050051	0.0150260	1 00 400	Glutathione S-transferase, C-terminal domain
Glyma.04G050851 Glyma.17G222300	0.0150260	1.99400	NA WRKY DNA -binding domain
Glyma.18G020600	0.0096787	2.01030	NA U
Glyma.12G225500	0.0013309	2.01380	Plastocyanin-like domain
Glyma.12G205700 Glyma.06G299900	0.0036907 0.0014386	2.03270	NA Myb-like DNA-binding domain
Glyma.19G160100	0.0104030	2.06330	EF hand
Glyma.01G131500 Glyma.05G211700	0.0000000	2.06350 2.15170	Aminotransferase class I and II
Glyma.18G013800	0.0000521	2.15210	Zinc finger, C3HC4 type (RING finger)
Glyma.10G282200 Glyma.17G090900	0.0018262	2.16430 2.19120	Armadillo/beta-catenin-like repeat; U-box
Glyma.12G149100	0.0078912	2.19290	domain No apical meristem (NAM) protein
			(Conntinua en la pág. siguiente)

Cuadro A.1: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.19G181600	0.0092795	2.19340	NA
Glyma.02G213700	0.0003669	2.21610	Carbonic anhydrase
Glyma.07G044900	0.0001186	2.21610	Protein of unknown function (DUF581)
Glyma.05G212300	0.0012035	2.29720	Tubby C 2
Glyma.08G176300	0.0000553	2.29990	NA
Glyma.17G127900	0.0004741	2.31030	NA
Glyma.15G250050	0.0001186	2.36330	NA
Glyma.19G219000	0.0000027	2.47110	Myb-like DNA-binding domain
Glyma.08G293366	0.0012035	2.50130	NA
Glyma.03G213300	0.0010182	2.50370	TspO/MBR family
Glyma.13G234500	0.0002275	2.51450	Glutaredoxin
Glyma.13G236500	$\begin{array}{c} 0.0002092\\ 0.0002984\\ 0.0000060\\ 0.0002984\\ 0.0000029\end{array}$	2.51810	AP2 domain
Glyma.12G117000		2.56070	AP2 domain
Glyma.15G250100		2.57790	NA
Glyma.19G181700		2.67200	NA
Glyma.07G090400		2.67410	NA
Glyma.11G243400	$\begin{array}{c} 0.0000553\\ 0.0000755\end{array}$	2.68560	Zinc finger, C3HC4 type (RING finger)
Glyma.01G060300		2.77400	Protein of unknown function (DUF1675)

Cuadro A.1: Continuación

Cuadro A.2: Genes con expresión diferencial significativa para el GT (genotipo tolerante) en respuesta a la sequía. GeneID= ID del gen en la notación Wm82.a4.v1, padj= valor p ajustado de la expresión diferencial, log2FoldChange= expresión diferencial en logaritmo, PFAM Description= Descripción del gen según la base de datos PFAM.

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.13G084000	0.0000000	-2.79960	Oxidoreductase NAD-binding domain ; Mo-co
-			oxidoreductase dimerisation domain;
			Oxidoreductase molybdopterin binding domain;
			Cytochrome b5-like Heme/Steroid binding
Glyma.07G155500 Glyma.09G216600	$0.0000000 \\ 0.0000001$	-2.66460 -2.47040	domain; Oxidoreductase FAD-binding domain Polyketide cyclase / dehydrase and lipid transport NA
Glyma.16G138400	0.0000000	-2.43790	NA
Glyma.18G289100	0.0000000	-2.37190	Hsp70 protein
Glyma.07G043100 Glyma.16G138500	$0.0000002 \\ 0.0000034$	-2.35310 -2.25010	NA NA
Glyma.05G103300	0.0000011	-1.99630	Protein of unknown function DUF260
Glyma.10G251400	0.0000931	-1.96890	P21-Rho-binding domain
Glyma.18G206000	0.0000290	-1.94440	Polykeude cyclase / denydrase and lipid transport
Glyma.02G275900	0.0000003	-1.82880	NA
Glyma.06G044400	0.0000341	-1.81540	Gibberellin regulated protein
Giyina.150000100	0.0000038	-1.80300	Al Pase family associated with various central
Glyma 07G043100	0.0000222	-1 74790	$\Delta $ $\Delta $ $\Delta $
Glyma.01G162500	0.00000001	-1.74390	NA
Glyma 11G095600	0.0017597	-1 68740	NA
Glyma.09G232700	0.0000562	-1.67270	Phosphate-induced protein 1 conserved region
Glyma.20G186500	0.0007508	-1.65520	Myb-like DNA-binding domain
Glyma.02G132100	0.0000284	-1.63680	Nitrite and sulphite reductase 4Fe-4S domain;
			Nitrite/Sulfite reductase ferredoxin-like half
			domain
Glyma.13G250700	0.0000000	-1.61900	NA
Glyma.01G124700	0.0004495	-1.58010	Polyketide cyclase / dehydrase and lipid transport
Glyma.05G184500	0.0009655	-1.57630	WRKY DNA -binding domain
Glyma.06G30/500	0.0042129	-1.5/480	Voltage-dependent anion channel
Glyma.05G034500	0.0046350	-1.56//0	Gibberellin regulated protein
Giyina.00G051800	0.0047033	-1.55950	Myd-like DNA-dinding domain
Glyma.12G228100	0.0048074	-1.55900	NA
Glyma.01G208200	0.0049096	-1.552/0	Najor intrinsic protein
Glyma 05G206200	0.0052862	-1.54810	Tubby C 2
Glyma.16G214500	0.0036532	-1.50420	TIR domain: EF hand

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G137000	$\begin{array}{c} 0.0038219\\ 0.0012421\\ 0.0094811\\ 0.0039914\\ 0.000002\end{array}$	-1.48310	2OG-Fe(II) oxygenase superfamily
Glyma.04G010300		-1.47400	bZIP transcription factor
Glyma.10G161000		-1.46960	Ndr family
Glyma.18G062300		-1.46040	AAA domain (Cdc48 subfamily)
Glyma.12G031300		-1.45900	Noduin-like
Glyma.14G056300	0.0112630	-1.44920	Peroxidase; Polyketide cyclase / dehydrase and
Glyma.08G096200 Glyma.04G052000 Glyma.08G332900	0.0010548 0.0013205 0.0006597	-1.43820 -1.43640 -1.42470	lipid transport Retinal pigment epithelial membrane protein NA Hsp90 protein; Histidine kinase-, DNA gyrase B-,
Glyma.08G261500	0.0000036	-1.42000	and HSP90-like ATPase NA
Glyma.02G028600 Glyma.15G052600 Glyma.06G158700 Glyma.04G042400	0.0121450 0.0160480 0.0013187 0.0097129	-1.41530 -1.39330 -1.36850 -1.36490	Putative peptidoglycan binding domain; Matrixin Peroxidase Auxin responsive protein D-mannose binding lectin; S-locus glycoprotein family: Protein kinase domain
Glyma.14G176400	0.0202070	-1.36340	Phosphate-induced protein 1 conserved region
Glyma.11G132802	$\begin{array}{c} 0.0208660\\ 0.0003820\\ 0.0014186\\ 0.0214680\\ 0.0202070\\ \end{array}$	-1.36080	NA
Glyma.11G097000		-1.35580	Transmembrane amino acid transporter protein
Glyma.07G165900		-1.35540	NA
Glyma.11G133300		-1.35060	Aldehyde dehydrogenase family
Glyma.18G065700		-1.34540	NA
Glyma.13G206700	0.0091542	-1.34020	Triose-phosphate Transporter family; EamA-like
Glyma.17G258200	$\begin{array}{c} 0.0089037\\ 0.0098208\\ 0.0008987\\ 0.0003021 \end{array}$	-1.33460	Gibberellin regulated protein
Glyma.06G102100		-1.33420	Phosphate-induced protein 1 conserved region
Glyma.11G140500		-1.33040	Hsp70 protein
Glyma.20G195900		-1.32830	AP2 domain
Glyma.09G272101 Glyma.05G042200 Glyma.20G137700	0.0086252 0.0039289 0.0237870	-1.32520 -1.31810 -1.31390	NA NA Protein tyrosine kinase; Salt stress response/antifungal
Glyma.06G264900	0.0006668 0.0000003	-1.30990	NA
Glyma.13G138600		-1.29350	Ubiquitin family
Glyma.17G153300 Glyma.11G253900	0.0023049 0.0261740	-1.29280 -1.29030	POT family Glycosyl hydrolases family 16; Xyloglucan endo-transglycosylase (XET) C-terminus
Glyma.18G289600	$\begin{array}{c} 0.0334280\\ 0.0207560\\ 0.0347380\end{array}$	-1.28630	Hsp70 protein
Glyma.18G284900		-1.27590	BAG domain; IQ calmodulin-binding motif
Glyma.06G050500		-1.27420	NA
Glyma.12G216300	0.0356220	-1.27340	NA
Glyma.02G002100	0.0351290	-1.27200	EF hand
Glyma.16G142200	0.0007262	-1.26830	GNS1/SUR4 family
Glyma.08G142400	0.0000001	-1.26750	WRKY DNA -binding domain
Glyma.18G276800	0.0000003	-1.26500	Transmembrane amino acid transporter protein
Glyma.12G035700	$\begin{array}{c} 0.0358520\\ 0.0047717\\ 0.0091054\\ 0.0100200\\ 0.0378930 \end{array}$	-1.26420	Auxin responsive protein
Glyma.01G206600		-1.26330	AP2 domain
Glyma.13G051800		-1.26220	UDP-glucoronosyl and UDP-glucosyl transferase
Glyma.01G203400		-1.25830	Glycosyl hydrolase family 14
Glyma.01G181102		-1.25020	NA
Glyma.18G287900	$\begin{array}{c} 0.0000003\\ 0.0351290\\ 0.0395640\\ 0.0090455\\ 0.0009282 \end{array}$	-1.24680	Hsp70 protein
Glyma.U035800		-1.24490	Helix-loop-helix DNA-binding domain
Glyma.06G031400		-1.24330	Myb-like DNA-binding domain
Glyma.06G264700		-1.23870	NA
Glyma.01G230900		-1.22920	Cytochrome c
Glyma.02G204700	$\begin{array}{c} 0.0309650\\ 0.0028759\\ 0.0005667\\ 0.0463330\\ 0.0462240 \end{array}$	-1.22880	Serine carboxypeptidase
Glyma.08G345000		-1.22660	MazG nucleotide pyrophosphohydrolase domain
Glyma.06G264800		-1.22640	NA
Glyma.09G277800		-1.22550	Peroxidase
Glyma.18G030300		-1.22540	NA
Glyma.04G173200	$\begin{array}{c} 0.0462240\\ 0.0326240\\ 0.0414920\end{array}$	-1.21650	NA
Glyma.02G225600		-1.21500	NA
Glyma.03G091100		-1.21240	NA

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.07G032000 Glyma.02G028700	$\begin{array}{c} 0.0250330\\ 0.0179190 \end{array}$	-1.20810 -1.20210	Kelch motif; F-box domain Putative peptidoglycan binding domain; Matrixin
Glyma.01G217400 Glyma.01G217400 Glyma.01G217400 Glyma.01G217400 Glyma.07G167000	$\begin{array}{c} 0.0175950\\ 0.0175950\\ 0.0175950\\ 0.0175950\\ 0.0175950\\ 0.0377300 \end{array}$	-1.20130 -1.20130 -1.20130 -1.20130 -1.20130 -1.19440	HSF-type DNA-binding HSF-type DNA-binding HSF-type DNA-binding HSF-type DNA-binding NA
Glyma.08G150800 Glyma.14G168800 Glyma.17G011600 Glyma.01G211200 Glyma.11G143100	$\begin{array}{c} 0.0000001\\ 0.0116620\\ 0.0132500\\ 0.0000026\\ 0.0478990 \end{array}$	-1.19310 -1.18660 -1.17970 -1.17920 -1.17920	Tetrapyrrole (Corrin/Porphyrin) Methylases NA Voltage-dependent anion channel NA Major intrinsic protein
Glyma.06G063800 Glyma.12G021700 Glyma.12G184700 Glyma.01G215600 Glyma.07G269600	$\begin{array}{c} 0.0236580\\ 0.0055661\\ 0.0058731\\ 0.0469410\\ 0.0091568\end{array}$	-1.16720 -1.16460 -1.14750 -1.14540 -1.14130	PPPDE putative peptidase domain Hexokinase; Hexokinase Myb-like DNA-binding domain NA Ribosomal protein S19
Glyma.13G133900 Glyma.15G062700 Glyma.03G120900 Glyma.13G093700	$\begin{array}{c} 0.0002927\\ 0.0338540\\ 0.0084021\\ 0.0000485 \end{array}$	-1.14020 -1.12730 -1.12370 -1.11390	Ribosomal L29 protein Cysteine-rich secretory protein family NAC domain; UBA/TS-N domain Protein kinase domain; Universal stress protein family
Glyma.07G138000	0.0225600	-1.10960	NA
Glyma.18G297500 Glyma.11G178300 Glyma.13G229700 Glyma.17G154100 Glyma.13G242800	$\begin{array}{c} 0.0176660\\ 0.0141220\\ 0.0374860\\ 0.0220110\\ 0.0462240 \end{array}$	-1.10530 -1.10060 -1.09990 -1.09260 -1.08940	Regulator of Vps4 activity in the MVB pathway Protein tyrosine kinase Protein of unknown function (DUF640) No apical meristem (NAM) protein NA
Glyma.14G011600	0.0004393	-1.08660	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase: Hsp90 protein
Glyma.18G251900 Glyma.17G145400 Glyma.14G160100 Glyma.03G162900	$\begin{array}{c} 0.0084516\\ 0.0494280\\ 0.0334200\\ 0.0005921 \end{array}$	-1.08640 -1.08420 -1.08300 -1.08190	NA AP2 domain Sugar efflux transporter for intercellular exchange Cofilin/tropomyosin-type actin-binding protein
Glyma.19G147800 Glyma.10G274300 Glyma.09G274000 Glyma.04G227900 Glyma.05G063600	$\begin{array}{c} 0.0304790\\ 0.0250330\\ 0.0121820\\ 0.0399850\\ 0.0025139 \end{array}$	-1.07430 -1.06750 -1.06600 -1.05790 -1.05530	Protein phosphatase 2C CCT motif; B-box zinc finger WRKY DNA -binding domain NA AP2 domain
Glyma.12G074400 Glyma.20G013600 Glyma.18G272300 Glyma.17G227800	$\begin{array}{c} 0.0048511\\ 0.0038620\\ 0.0469410\\ 0.0357040 \end{array}$	-1.05470 -1.04860 -1.04580 -1.04440	CCT motif DnaJ domain Cytochrome P450 Glycosyl hydrolases family 32 C terminal:
Gryma.17G227600	0.0337040	-1.0+++0	Glycosyl hydrolases family 32 C terminal domain
Glyma.08G210700	0.0032030	-1.03530	Kelch motif; F-box domain
Glyma.06G319460 Glyma.01G026200 Glyma.15G051500	$\begin{array}{c} 0.0070814\\ 0.0472260\\ 0.0261630 \end{array}$	-1.03390 -1.03000 -1.02940	NA MatE Glycosyl hydrolase family 3 N terminal domain; Glycosyl hydrolase family 3 C-terminal domain
Glyma.08G339500 Glyma.15G072800	$0.0227550 \\ 0.0342900$	-1.02320 -1.01670	NA Fatty acid desaturase; Sphingolipid Delta4-desaturase (DES)
Glyma.08G020600	0.0032582	-1.01600	ATP synthase (E/31 kDa) subunit
Glyma.15G030800	0.0421490	-1.01400	NA
Glyma.18G189600 Glyma.06G124800	$0.0397390 \\ 0.0301430$	-1.00610 -1.00480	POT family NA
Glyma.16G017200	0.0424300	-1.00220	NA
Glyma.14G158500 Glyma.14G103100	0.0214810 0.0176780	-1.00060 -0.99960	NA WRKY DNA -binding domain
Glyma.06G139300	0.0187880	-0.97620	NA Cur/Mat matchalism DLD damas durt anomalism
Gryma.15G001200	0.0430990	-0.97310	Cys/wei metadonsm PLP-dependent enzyme

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G019200	0.0047055	-0.97297	Protein of unknown function (DUF632); Protein of unknown function (DUF630)
Glyma.18G011000 Glyma.18G209600	$\begin{array}{c} 0.0202270\\ 0.0021762\end{array}$	-0.97246 -0.96993	Protein tyrosine kinase D-ala D-ala ligase C-terminus; D-ala D-ala ligase
Glyma.17G138900 Glyma.20G248100	0.0388620 0.0187880	-0.96707 -0.96565	N-terminus NA Helix-loop-helix DNA-binding domain
Glyma 08G339600	0.0058731	-0.96333	NA
Glyma.08G139800	0.0225600	-0.95752	NA NA
Glyma.03G146100 Glyma.02G091500	0.0025354	-0.95344	Activator of Hsp90 ATPase, N-terminal; Activator of Hsp90 ATPase homolog 1-like protein Protein of unknown function (DUF506)
Glyma.02G200300	0.0023494	-0.94847	Ribosomal S17
Glyma.19G164400	0.0017099	-0.94746	Cofilin/tropomyosin-type actin-binding protein
Glyma.18G272200 Glyma.17G154800	0.0325090	-0.94655 -0.94117	NA
Glyma.07G215500 Glyma.04G225500	0.0472260 0.0061206	-0.93853 -0.93637	UDP-glucoronosyl and UDP-glucosyl transferase NA
Glyma.08G136200	0.0495570	-0.93485	Annexin
Glyma.04G055600	0.0058731	-0.92972	NA NA
Glyma.08G261450	0.0351290	-0.92784	NA Ea S matchaliam appaainted domains Opinalinate
Giyilia.03G042400	0.0000050	-0.92017	synthetase A protein
Glyma 20G169600	0.0195290	-0.90556	NA
Glyma.16G042000	0.0193290	-0.90063	short chain dehydrogenase
Glyma.16G214800	0.0041737	-0.90036	TIR domain; EF hand PB1 domain
Glyma.13G361700	0.0342900	-0.89368	20G-Fe(II) oxygenase superfamily
Glyma.17G096600	0.0015326	-0.89185	HPP family
Glyma.19G035800	0.0026515	-0.89036	UDP-glucoronosyl and UDP-glucosyl transferase
Glyma.12G192200	0.0208070	-0.89033	Protein kinase domain
Glyma.07G080400	0.0007481	-0.88210	C2 domain; Phospholipase D Active site motif; Phospholipase D C terminal
Glyma.18G011900	0.0010995	-0.88185	60Kd inner membrane protein
Glyma.08G077300 Glyma.12G210600	0.0000581	-0.87751 -0.87389	Utp11 protein Phosphoenolpyruvate carboxylase
Glyma.17G162000	0.0331750	-0.87388	Protein of unknown function DUF260
Glyma.07G248000	0.0342900	-0.87255	NA
Glyma.03G086200	0.0027382	-0.86915	Tetratricopeptide repeat; FKBP-type
Glvma.04G199900	0.0022324	-0.85804	Helix-loop-helix DNA-binding domain
Glyma.01G240000	0.0020783	-0.85467	Ribosomal protein S8
Glyma.07G043800 Glyma.11G225400	0.0046162	-0.85271 -0.84609	NA
Glyma.17G018600	0.0072100	-0.84577	Ribosomal protein S8
Glyma.03G070700	0.0064395	-0.84558	Mitochondrial carrier protein
Glyma.03G128300 Glyma.11G049300	0.0392860	-0.84272 -0.84173	NA Protein of unknown function DUF260
Glyma.16G016600	0.0003514	-0.84089	No apical meristem (NAM) protein
Glyma.13G126400	0.0049358	-0.83330	Ribosomal protein L14p/L23e
Glyma.13G165200 Glyma.02G096800	0.0001138	-0.82692 -0.82642	Ribosomal protein S17 Glucose-6-phosphate dehydrogenase. C-terminal
Gryma.020090000	0.0002271	0.02012	domain: Glucose-6-phosphate dehydrogenase.
			NAD binding domain
Glyma.13G051000 Glyma.02G125900	0.0364710 0.0384340	-0.82102	ATP-sulfurylase NA
Glyma 06G101500	0.0045818	_0.81577	NA
Glyma.12G063000	0.0275800	-0.81493	Tim10/DDP family zinc finger
Glyma.18G067700 Glyma.14G017400	0.0320120	-0.80723	Mitochondrial carrier protein Calponin homology (CH) domain
Glyma.11G002400	0.0099660	-0.80290	Ribosomal protein S10p/S20e
			(Conntinua en la pág. siguiente)

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
$C1_{vm_0} 06C164900$	0 0008834	0.00100	Fibrillorin
Glyma 15C050000	0.0098834	-0.80180	FIDIIIIafiii Usen70 meetain
Glyma 12C072100	0.0027834	-0.79915	Derovidose
Glyma.12G0/3100	0.0000080	-0./9890	Peroxidase
Glyma.15G009900	0.0022552	-0.79795	INA NA
Glyma.20G200800	0.0199510	-0./9/64	NA
Glyma.10G151200	0.0019829	-0.79282	Ribosomal L22e protein family
Glyma.01G081100	0.0351290	-0.78657	AP2 domain
Glyma.01G081100	0.0351290	-0.78657	AP2 domain
Glyma.18G284600	0.0000193	-0.78248	Glucose-6-phosphate dehydrogenase, NAD
2			binding domain: Glucose-6-phosphate
			bilding domain, Ordeose-o-phosphate
G1 15G000000	0.04702(0	0.701(0	dehydrogenase, C-terminal domain
Glyma.15G098800	0.04/0260	-0./8162	NA
Glyma.19G134800	0.0485510	-0.78152	NA
Glyma.10G076400	0.0055602	-0.78079	Ribosomal S17
Glvma.05G051600	0.0027168	-0.77800	Tetratricopeptide repeat
Glvma.10G193200	0.0022324	-0.77763	TCP-1/cpn60 chaperonin family
Glvma.15G010900	0.0017643	-0.77587	Cytochrome c oxidase subunit Vb
C1 20C125200	0.00000(1	0.77500	NIA
Glyma.20G135200	0.0000061	-0.77520	
Glyma.08G100000	0.0139850	-0.77345	Ribosomal L38e protein family
Glyma.14G111400	0.0388620	-0.77330	Plant protein of unknown function (DUF868)
Glyma.20G237000	0.0006050	-0.77328	Ribosomal L22e protein family
Glyma.08G095900	0.0035124	-0.77240	NA
Glyma.11G030500	0.0339820	-0.76244	Tetratricopeptide repeat
Glyma 18G096400	0.0388440	-0.75432	NA
Glyma 12G230300	0.0485830	-075244	NA
Glyma 01G165900	0.0421490	-0 74903	Armadillo/beta-catenin-like repeat
Glyma.08G360200	0.0252450	-0.74660	No apical meristem (NAM) protein
	0.0050450	0.74660	
Glyma.08G360200	0.0252450	-0.74660	No apical meristem (NAM) protein
Glyma.03G142200	0.0023135	-0.74253	Ribosomal protein S10p/S20e
Glyma.06G092600	0.0137290	-0.74219	PPR repeat
Glyma. 10G224200	0.002/312	-0./41/2	
Glyma.19G030900	0.0254170	-0./3465	ICP family transcription factor
Glyma.03G172300	0.0041737	-0.73142	RNA recognition motif. (a.k.a. RRM, RBD, or
2			RNP domain)
Glyma 01G199600	0.0301420	-0 72829	Protein phosphatase 2C
Glyma 11G003600	0.0125450	-072479	Ribosomal protein S8
Glyma.03G084000	0.0160480	-0.72465	Cleavage site for pathogenic type III effector
0171111102 000 1000	0.0100100	0.72105	avimulance fector Avim
Glyma 16G074600	0.0260420	0 72442	Avirulence factor Avi Dibosomal protein L 13a
Gryma.100074000	0.0200420	-0.72442	Ribbsoniai protein E15e
Glyma.15G177100	0.0091542	-0.72315	GDP-fucose protein O-fucosyltransferase
Glyma.15G010300	0.0001610	-0.72158	Biotin-requiring enzyme
Glyma.19G238800	0.0202070	-0.71842	Ribosomal protein S11
Glyma.03G216600	0.0007508	-0.71760	Ribosomal protein L19e
Glyma.U031221	0.0177090	-0.71564	NA
Glyma 11G249700	0.0043827	-0.71225	Ribosomal protein L 22p/L 17e
Glyma 07G136500	0.0424300	-0.71207	Regulator of chromosome condensation (RCC1)
Grynna.or Greecee	0.0121500	0.71207	regulator of enformesonic condensation (recei)
C_{1}	0.0106170	0 70170	Pibesemel protein I 7A o/I 20o/S12o/Codd45
Glyma.18G281200	0.0106170	-0./01/0	Ribosomai protein L/Ae/L30e/S12e/Gadd45
			family
Glyma.04G004500	0.0424300	-0.70028	Nodulin-like
Glyma.17G063100	0.0245470	-0.70014	NA
Glyma 11G237200	0.0113530	-0 69529	Fibrillarin
Glyma 15G106200	0.0107400	-0 69440	NA
Glyma 02G263200	0.0019435	-0.69379	KOW motif: Ribosomal family S4e: RS4NT
020200200	5.0017100	0.07017	(NUC022) domain: \$4 domain
Glyma 02G272700	0.0006113	0.60047	(NUC025) dollalli, 54 dollalli Probable molybdopterin binding domain: MoeA
Oryma.02O272700	0.0000113	-0.09047	riobable morybuopterin binding domain; MoeA
			C-terminal region (domain IV); MoeA
			N-terminal region (domain I and II)
Glyma.09G052100	0.0139240	-0.68884	Ribosomal LŽ7e protein family; KÓW motif
Glyma 10C0/5000	0.0268670	0.68550	Ribosomal I 29 protein
Glyma 08C165600	0.0208070	-0.06339	Helicase conserved C terminal domain:
Gryma.000105000	0.0100000	-0.00392	
			DEAD/DEAH box helicase

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.13G128000	0.0273360	-0.68238	Beta-ketoacyl synthase, N-terminal domain;
Glyma.10G018900 Glyma.20G171600	$0.0224760 \\ 0.0121070$	-0.68126 -0.68004	Beta-ketoacyl synthase, C-terminal domain Ribosome biogenesis protein Nop16 Ras family
Glyma.06G138700	$\begin{array}{c} 0.0012396\\ 0.0375650\\ 0.0469830\\ 0.0275000\\ 0.0250330\end{array}$	-0.67980	Ribosomal protein L6
Glyma.02G047500		-0.67971	'Cold-shock' DNA-binding domain; Zinc knuckle
Glyma.04G007500		-0.67950	ATP synthase subunit C
Glyma.02G095000		-0.67921	NA
Glyma.10G039900		-0.67858	Ribosomal protein L14p/L23e
Glyma.20G160500	0.0113940	-0.67684	NA
Glyma.05G183300	0.0198450	-0.67621	NA
Glyma.10G049400	0.0181370	-0.67589	Ribosomal protein S12/S23
Glyma.12G181400	0.0042166	-0.67549	NA
Glyma.14G101400	0.0088183	-0.67343	Protein phosphatase 2C
Glyma.03G176800	$\begin{array}{c} 0.0339290\\ 0.0027834\\ 0.0199510\\ 0.0019829\\ 0.0010273 \end{array}$	-0.66983	Adenylate kinase
Glyma.15G015500		-0.66964	Ribosomal protein L37e
Glyma.12G201900		-0.66950	Ribosomal protein S7p/S5e
Glyma.07G214900		-0.66864	Ribosomal protein S9/S16
Glyma.13G357900		-0.66759	Ribosomal protein L37e
Glyma.20G132651	$0.0178600 \\ 0.0079395$	-0.66533	NA
Glyma.14G061100		-0.65801	Ribosomal protein L10; 60s Acidic ribosomal
Glyma.18G020000 Glyma.12G182800 Glyma.02G255500	$\begin{array}{c} 0.0345250\\ 0.0202520\\ 0.0052260\end{array}$	-0.65763 -0.65462 -0.65361	Fibrillarin 60s Acidic ribosomal protein Ribosomal protein L10; 60s Acidic ribosomal protein
Glyma.01G221200	$\begin{array}{c} 0.0356630\\ 0.0078805\\ 0.0268160\\ 0.0055220\\ 0.0006597 \end{array}$	-0.65209	PLAC8 family
Glyma.17G105900		-0.65177	Ribosomal protein S17
Glyma.06G077800		-0.65125	UAA transporter family
Glyma.19G132800		-0.65000	EF hand
Glyma.16G065900		-0.64960	PQ loop repeat
Glyma.01G222100	$\begin{array}{c} 0.0047055\\ 0.0361670\\ 0.0249620\\ 0.0032660\\ 0.0047055\end{array}$	-0.64658	Plectin/S10 domain
Glyma.09G269000		-0.64639	Cornichon protein
Glyma.15G267100		-0.64637	Ribosomal protein S2
Glyma.02G236900		-0.64362	NA
Glyma.20G207500		-0.64237	Ribosomal protein L14
Glyma.07G012800	0.0124660	-0.64217	Ribosomal L28e protein family
Glyma.03G111300	0.0085714	-0.64117	Protein of unknown function (DUF632); Protein
Glyma.01G242902 Glyma.01G026100 Glyma.07G198600	$\begin{array}{c} 0.0094639\\ 0.0029638\\ 0.0251250\end{array}$	-0.64110 -0.64016 -0.64012	of unknown function (DUF630) NA Ribosomal protein L18e/L15 Ribosomal protein S6e
Glyma.15G075600	0.0063148	-0.63784	Ribosomal protein S6e
Glyma.13G352100	0.0214810	-0.63763	Ribosomal L28e protein family
Glyma.19G028700	0.0009492	-0.63761	Ribosomal protein S19e
Glyma.03G125800	0.0413650	-0.63713	Chaperonin 10 Kd subunit
Glyma.08G131500	0.0211870	-0.63695	60s Acidic ribosomal protein
Glyma.15G152900	$\begin{array}{c} 0.0049096\\ 0.0110730\\ 0.0207560\\ 0.0067680\\ 0.0378380 \end{array}$	-0.63295	C2 domain
Glyma.08G181300		-0.63236	Ribosomal protein L44
Glyma.15G053700		-0.63187	Protein phosphatase 2C
Glyma.10G004300		-0.62761	Ribosomal protein L14
Glyma.15G258800		-0.62729	Auxin responsive protein
Glyma.01G007200 Glyma.05G022200	0.0339290 0.0250330	-0.62644 -0.62506	NA Clp amino terminal domain; AAA domain (Cdc48 subfamily); ATPase family associated with various cellular activities (AAA); C-terminal, D2-small domain, of ClpB protein
Glyma.09G234300	$\begin{array}{c} 0.0007508 \\ 0.0479210 \\ 0.0350500 \end{array}$	-0.62386	Adenylate kinase; Adenylate kinase, active site lid
Glyma.19G203300		-0.62325	Ribosomal protein S28e
Glyma.20G197100		-0.62325	TCP-1/cpn60 chaperonin family
Glyma.05G061400 Glyma.02G107100 Glyma.10G002400	0.0023257 0.0239870 0.0046608	-0.62261 -0.62247 -0.62203	Ribosomal L15 Ribosomal protein L34e Ribosomal protein L7Ae/L30e/S12e/Gadd45 family

Cuadro A.2: Continuación

Glyma.12G131700 0.0199510 -0.62131 NA Glyma.05G132500 0.0009873 -0.62028 Ribosomal protein 1.35Ae Glyma.04G091700 0.001755 -0.62017 MYND finger; Ubiquitin carboxyl-terminal hydrolase Glyma.16G05000 0.009655 -0.61839 Glyma.16G05000 0.0099055 Glyma.16G21040 0.0170130 -0.61839 NA Glyma.16G21040 0.00089055 -0.61839 NA Glyma.16G21040 0.010705660 -0.61183 NA Glyma.16G210400 0.0026228 -0.61092 Ubiquint family: Ribosomal protein S2ra Glyma.030086400 0.002628 -0.61092 Ribosomal protein S2c6 Glyma.030086400 0.0025310 -0.60328 Ribosomal protein S2c6 Glyma.030086400 0.0025310 -0.60328 Ribosomal L29 protein Glyma.07G149500 0.011380 -0.60338 Mitochondrial carrier protein Glyma.07G149500 0.011380 -0.60338 Mitochondrial carrier protein Glyma.07G149500 0.011380 -0.60338 Mitochondrial carrier protein	Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Gym.abs/G02120 0.0001755 0.02017 Glyma.04G091700 0.0001755 0.0217 MYND finger: Ubiquitin carboxyl-terminal hydrolase Glyma.17G098600 0.0137290 -0.61934 Cytochrome c Glyma.16G05000 0.0009655 -0.61934 Cytochrome c Glyma.16G05000 0.0009655 -0.6183 NA Glyma.16G20400 0.0007655 -0.6183 NA Glyma.16G20400 0.0017660 -0.61188 NA Glyma.16G20400 0.0026828 -0.61092 Ubiquitin family; Ribosomal protein S26 Glyma.03G066400 0.0027312 -0.60852 Ribosomal protein S26 Glyma.03G086400 0.0027312 -0.60852 Ribosomal protein S26 Glyma.03G086400 0.0027312 -0.60852 Ribosomal protein S26 Glyma.03G086400 0.002732 -0.60852 Ribosomal protein S26 Glyma.03G086400 0.0038935 -0.60133 Ribosomal protein S26 Glyma.03G207500 0.0038935 -0.60133 Ribosomal protein S27 Glyma.03G207500 0.0047055 -0.60133	Glyma.12G131700 Glyma.05G132500	0.0199510	-0.62131	NA Ribosomal protein I 35Ae
bydrolase bydrolase bydrolase Glyma. J7G098600 0.0137290 -0.61934 SY domain; Ribosomal protein S4/S9 N-terminal Glyma. J6G050000 0.0009655 -0.61934 SY domain; Ribosomal protein Glyma. J6G050000 0.00176600 -0.61189 NA Glyma. J30369700 0.0176600 -0.61189 NA Glyma. J6G050000 0.0176600 -0.61189 NA Glyma. J30369700 0.0176602 -0.61923 Ubiquitin family; Ribosomal protein S27a Glyma.J30369700 0.0212320 -0.61922 Ubiquitin family; Ribosomal protein S26 Glyma.J30368400 0.0027312 -0.60920 Ribosomal protein S26 Glyma.J30264700 0.0028514 -0.60831 Ribosomal protein S26 Glyma.J30264700 0.0028542 -0.60831 Ribosomal protein S26 Glyma.J3026000 0.011380 -0.60183 Mitchonfrai carrier protein Glyma.J60045300 0.0103803 -0.60183 Mibosomal protein S27 Glyma.J60045300 0.0195150 -0.59348 Ribosomal protein S27 Glyma.J60045300	Glyma.04G091700	0.0001755	-0.62017	MYND finger; Ubiquitin carboxyl-terminal
Glyma, 170,0800 0.01470520 -0.61849 34 domain; Ribosomal protein S4/S9 N-terminal domain Glyma, 13630800 0.0099655 -0.61849 34 domain; Ribosomal protein S4/S9 N-terminal domain Glyma, 136318200 0.0099005 -0.61839 NA Glyma, 136318200 0.0098005 -0.61839 NA Glyma, 136359700 0.0176660 -0.61183 NA Glyma, 136359700 0.0176660 -0.61183 NA Glyma, 026024400 0.012320 -0.61050 NA Glyma, 026034400 0.042650 -0.60922 Picfoldin subunit Glyma, 026044700 0.0198574 -0.60822 Ribosomal protein S26 Glyma, 02604700 0.0098574 -0.60820 NA Glyma, 02604700 0.0098574 -0.60832 Ribosomal protein S26 Glyma, 02604700 0.0098574 -0.60832 Na Glyma, 02604700 0.0098574 -0.60338 Mitochondrial carrier protein Glyma, 0270500 0.0110380 -0.60338 Ribosomal protein L21 Giyma, 027077500 0.0342410 -0.59441 </td <td></td> <td>0.0476600</td> <td>0 (1024</td> <td>hydrolase</td>		0.0476600	0 (1024	hydrolase
Glyma.16G050000 0.0009655 -0.61839 NA Glyma.13G318200 0.0089005 -0.61288 Koticic ribosomal protein 1.7Ae/I.30e/S12e/Gadd45 Glyma.13G369700 0.0176660 -0.61188 NA Glyma.13G369700 0.0026828 -0.61092 Ubiquitin family: Ribosomal protein S27a Glyma.03G086400 0.0027312 -0.60050 NA Glyma.03G086400 0.00027312 -0.60050 NA Glyma.03G086400 0.00027312 -0.60050 NA Glyma.03G086400 0.0002324 -0.60829 NA Glyma.04G084700 0.00423810 -0.60409 NA Glyma.07G149500 0.010389 -0.60138 Mitochondrial carrier protein Glyma.07G640600 0.0151550 -0.59848 Ribosomal protein L32 Glyma.07G060000 0.0151550 -0.59848 Ribosomal protein L32 Glyma.07G060000 0.0151550 -0.59848 Ribosomal protein L22 Glyma.07G060000 0.0151550 -0.59747 Methy transferase domain Glyma.07G0600000 0.0151550 -0	Glyma.17G098600	0.0476620 0.0137290	-0.61934 -0.61849	S4 domain; Ribosomal protein S4/S9 N-terminal
Glyma.13G318200 0.0089005 -0.61288 60s Acidic ribosomal protein Glyma.13G318200 0.00170130 -0.61188 NA Glyma.13G3260900 0.012320 -0.61188 Ribosomal protein L7Ac/L30c/S12c/Gadd45 Glyma.03G260900 0.0212320 -0.61083 Ribosomal protein SZee Glyma.03G08400 0.0027312 -0.60929 Prefoldin submit Glyma.03G08400 0.002321 -0.60851 Ribosomal protein SZee Glyma.03G08400 0.002321 -0.60851 Ribosomal protein SZee Glyma.03G08400 0.002323 -0.60851 Ribosomal protein SZe Glyma.07G149500 0.011380 -0.60833 Mitochondrial carrier protein Glyma.17G149500 Glyma.16030500 0.0047055 -0.61138 Ribosomal protein L21e Glyma.17G080000 0.015150 -0.59941 Ribosomal protein S27 Glyma.16030500 0.0444920 -0.59774 Methyltransferase domain Glyma.1604633900 0.0444920 -0.59749 Ribosomal protein L22p/L17e Glyma.04G133900 Glyma.04G13900 0.0444921 <td>Glyma 16G050000</td> <td>0.0009655</td> <td>-0.61839</td> <td>domain NA</td>	Glyma 16G050000	0.0009655	-0.61839	domain NA
Glyma.13G369700 0.0170130 -0.61189 NA Glyma.13G369700 0.0170660 -0.61183 Ribosomal protein L7Ae/L30e/S12e/Gadd45 family Glyma.03G260900 0.022832 -0.61050 NA Glyma.03G3086400 0.0027312 -0.60922 Ribosomal protein S8e Glyma.03G204700 0.0025814 -0.60920 Ribosomal protein S2e Glyma.03G241100 0.0005814 -0.60823 NA Glyma.03G241100 0.0005824 -0.60829 NA Glyma.14G208600 0.00138935 -0.60829 NA Glyma.07G149500 0.01038935 -0.60133 Mitochondrial carrier protein Glyma.07G149500 0.014755 -0.60183 Ribosomal protein L22 Glyma.04G038900 0.0444920 -0.5974 Glyma.07G060000 0.015150 -0.59888 Ribosomal protein L22 Glyma.132025800 0.00489140 -0.59494 Glyma.07500 0.0342410 -0.59734 Methyltransfrase domain Glyma.132025300 0.0444920 -0.59734 Glyma.07517400 0.04489140 -0.597349 <	Glyma.13G318200	0.0089005	-0.61268	60s Acidic ribosomal protein
Glyma.15G129800 0.0026828 -0.61030 The protein Driver Drive	Glyma.16G210400 Glyma.13G369700	0.0170130	-0.61189 -0.61183	NA Ribosomal protein L7Ae/L30e/S12e/Gadd45
			0101100	family
	Glyma.15G129800 Glyma.03G260900	0.0026828 0.0212320	-0.61092	Ubiquitin family; Ribosomal protein S27a
Glyma.03C086400 0.0027312 -0.60920 Ribosomal protein S8e Glyma.20C240700 0.0098374 -0.60851 Ribosomal protein S2ce Glyma.20C240700 0.009252 -0.60829 NA Glyma.20C240700 0.0003232 -0.60409 NA Glyma.12C26800 0.0038935 -0.60138 Ribosomal protein L32 Glyma.02C26800 0.0365670 -0.60169 Thioredoxin Glyma.04C038900 0.0355670 -0.60169 Thioredoxin Glyma.04G038000 0.0151550 -0.59941 Ribosomal protein L21e Glyma.04G038000 0.01382410 -0.59774 Methyltransferase domain Glyma.13C005000 0.01382410 -0.59774 Methyltransferase domain Glyma.08G151900 0.002485 -0.59349 Ribosomal protein L27/L17e Glyma.08G151900 0.002485 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, Glyma.02G174300 0.002645 -0.59349 Ribosomal protein L44 Glyma.02G17500 0.0047921 -0.59320 Ribosomal protein L45 Glyma.02G19500 0	Glyma.09G034400	0.0426650	-0.60992	Prefoldin subunit
Glyma.140208000 0.0035310 -0.00822 Kibosomal L29 protein Glyma.036241100 0.00053874 -0.60829 NA Glyma.102640700 0.00053874 -0.60839 NA Glyma.102640700 0.00053874 -0.60839 NA Glyma.07G149500 0.0110380 -0.60133 Mitochondrial carrier protein Glyma.04G038900 0.0365670 -0.60169 Thioredoxin Glyma.04G04500 0.0199510 -0.59888 Ribosomal protein L21e Glyma.04G04500 0.0151550 -0.59888 Ribosomal protein L22p/L17e Glyma.04G04500 0.0135970 -0.59610 Hsp20/alpha crystallin family Glyma.17G08300 0.0449210 -0.59747 Methyltransferase domain Glyma.05G174300 0.0047921 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, Glyma.05G174300 0.0047921 -0.59232 Ribosomal protein L44 Gima.036215300 Glyma.0238700 0.001476240 -0.59232 Ribosomal protein L44 Gima.036215300 0.0026515 Glyma.026215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Gifma.02320000 0.0026434 -0.	Glyma.03G086400	0.0027312	-0.60920	Ribosomal protein S8e
Giyma.03C241100 0.0005232 -0.60829 NA Giyma.12G084700 0.0423810 -0.60409 NA Giyma.12G084700 0.0110380 -0.60338 Mitochondrial carrier protein Giyma.14G038900 0.00356570 -0.60153 Ribosomal protein L32 Giyma.04G038900 0.0047055 -0.60153 Ribosomal protein L21e Giyma.04G133900 0.0444920 -0.59784 Ribosomal protein L27 Giyma.05025800 0.0135970 -0.59849 Ribosomal protein L22p/L17e Giyma.08G0151900 0.002485 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Giyma.05G174300 0.0047921 -0.59322 Ribosomal protein L44 Cigma.05G17500 Giyma.03G17500 0.0047921 -0.59323 Ribosomal protein L44 Cigma.03G17500 0.00476240 -0.59323 Giyma.03G174300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Cigma.03G17300 0.002651 -0.59165 Giyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Cigma.03G17300 0.002651 -0.59874 Giyma.04G17300 0.0024563 -0.58678 Ribosomal	Glyma.20G240700	0.0095874	-0.60852	Ribosomal L29 protein
Giyma.126084700 0.0423810 -0.60409 NA Giyma.116226800 0.0110380 -0.60338 Mitochondrial carrier protein Giyma.046038900 0.0365670 -0.60163 Ribosomal protein L32 Giyma.042038900 0.0365670 -0.60163 Ribosomal protein L21e Giyma.046038900 0.0151550 -0.59941 Ribosomal protein S27 Giyma.046038900 0.0151550 -0.59888 Ribosomal L18p/L5e family Giyma.04603800 0.0151550 -0.59874 Methyltransferase domain Giyma.136005500 0.01382410 -0.59577 Ribosomal protein L22p/L17e Giyma.04603800 0.0382410 -0.59578 Ribosomal protein L22p/L17e Giyma.046013900 0.044920 -0.59474 Methyltransferase domain Giyma.136007500 0.0382410 -0.59575 Ribosomal protein L22p/L17e Giyma.07603500 0.044921 -0.59329 Ribosomal protein L22p/L17e Giyma.076174300 0.0047921 -0.59322 Ribosomal protein L22p/L17e Giyma.076219500 0.0093614 -0.59232 Ribosomal protein L44 Giyma.026038700 0.001664 -0.59230 Ribosomal protein L18e/L15 Giyma.026038700 0.001664 -0.59165 Ribosomal protein L18e/L15 Giyma.026038700 0.0026415 -0.59058 Ribosomal protein L18e/L15 Giyma.026038700 0.002643 -0.58676 Ribosomal protein L18e/L15 Giyma.046211800 0.002243 -0.58676 Ribosomal protein L18e/L15 Giyma.046229000 0.002643 -0.58676 Ribosomal protein L18e/L15 Giyma.046229000 0.002643 -0.58676 Ribosomal protein L18e/L15 Giyma.046229000 0.002643 -0.58874 NA Giyma.046229000 0.002643 -0.58874 RA Giyma.046121800 0.002643 -0.58875 Ribosomal protein L18e/L15 Giyma.046121800 0.002643 -0.58876 Ribosomal protein L18e/L15 Giyma.046029000 0.002643 -0.58874 RA Giyma.046029000 0.002643 -0.58878 RA Giyma.046042000 0.002643 -0.58878 RA Giyma.046042000 0.002643 -0.58878 RA Giyma.046042000 0.00275670 -0.58437 Ribosomal protein S27 Giyma.046042000 0.00276870 -0.58421 Ribosomal protein S24 Giyma.046042000 0.00276870 -0.58421 Ribosomal protein S24 Giyma.046043000 0.0332470 -0.56849 NA Giyma.0460420400 0.0032470 -0.56849 NA Giyma.046043000 0.0332470 -0.56342 NA Giyma.046043000 0.0332470 -0.56342 NA Giyma.046043000 0.0332470 -0.56342 NA Giyma.046043000 0.0332470 -0.56342 NA Giyma.046	Glyma.03G241100	0.0005232	-0.60829	NA
Glyma.102(5149500 0.0013803 -0.60338 Mitochondral carrier protein Glyma.12226800 0.0038935 -0.60169 Thioredoxin Glyma.020207500 0.0047055 -0.60169 Ribosomal protein L21e Glyma.032027500 0.0047055 -0.60169 Ribosomal protein S27 Glyma.046113900 0.0151550 -0.59888 Ribosomal protein S27 Glyma.136205800 0.0138970 -0.59610 Hsp20/alpha crystallin family Glyma.136007500 0.0382410 -0.59578 Ribosomal protein L22p/L17e Glyma.08G151900 0.00489140 -0.59349 Ribosomal protein L22p/L17e Glyma.08G151900 0.0047921 -0.59322 Ribosomal protein L22p/L17e Glyma.03G215300 0.0047921 -0.59232 Ribosomal protein L44 Glyma.03G20500 0.001664 -0.59238 Ribosomal protein L44 Glyma.03G215300 0.0026435 -0.59058 Ribosomal protein L18e/L15 Glyma.04221800 0.0026436 -0.58874 NA Glyma.156250500 0.0072003 -0.58874 NA Glyma.16624600 0.0026438 -0.58686 Sacosmal protein L18e/L15	Glyma.12G084700	0.0423810	-0.60409	NA
Giyma.04G038900 0.0365670 -0.60153 Ribosomal protein L21e Giyma.07G060000 0.0199510 -0.59941 Ribosomal protein S27 Giyma.16G046500 0.0199510 -0.59848 Ribosomal protein S27 Giyma.16G046500 0.0135570 -0.59848 Ribosomal protein S27 Giyma.16G07500 0.0382410 -0.59977 Methyltransferase domain Giyma.170083500 0.0489140 -0.59949 Ribosomal protein L7AcL30e/S12e/Gadd45 Giyma.05G174300 0.0020485 -0.59349 Ribosomal protein L6e ; Ribosomal protein Giyma.05G174300 0.0047921 -0.593222 Go Acidic ribosomal protein CA Giyma.05G174300 0.0047921 -0.593222 Ribosomal protein L44 Ciyma.05G174300 0.0047921 Giyma.05G174300 0.0047921 -0.59232 Ribosomal protein L18c/L15 Giyma.05G174300 0.0020464 -0.59058 Ribosomal protein L18c/L15 Giyma.05G174300 0.0026014 -0.59058 Ribosomal protein L18c/L15 Giyma.05G174300 0.002000 -0.58874 NA Giyma.17021500 0.007003 -0.58874 NA Giyma.16G1700 0.0376120 -0.58876 Ribosomal pro	Glyma.0/G149500 Glyma.11G226800	0.0110380	-0.60338 -0.60183	Ribosomal protein L32
Glyma.16020/500 0.004/055 -0.60135 Ribosomal protein S27 Glyma.06046500 0.0199510 -0.59941 Ribosomal protein S27 Glyma.04G133900 0.0444920 -0.59774 Methyltransferase domain Glyma.162025800 0.0135970 -0.59848 Ribosomal protein L22p/L17e Glyma.16205500 0.0382410 -0.59477 Ribosomal protein L7Ae/L30e/S12e/Gadd45 family Glyma.05G174300 0.0047921 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.05G174300 0.0047921 -0.59232 Ribosomal protein L18e/L15 Glyma.02G038700 0.00476240 -0.59232 Ribosomal protein L18e/L15 Glyma.02G038700 0.0010664 -0.59168 Ribosomal protein L18e/L15 Glyma.02G23500 0.002200515 -0.59058 Ribosomal protein L18e/L15 Glyma.19G117300 0.0230060 -0.58874 NA Glyma.19G117300 0.022433 -0.58676 Ribosomal protein L18e/L15 Glyma.19G105700 0.0376120 -0.58676 Ribosomal protein L6e; Ribosomal protein L6, N-terminal domain Glyma.106228000 0.0275670 -0.58421 Ribosomal protein L6; Ribosoma	Glyma.04G038900	0.0365670	-0.60169	Thioredoxin
Glyma.07G060000 0.0151550 -0.59888 Ribosomal L18p/L5e family Glyma.04G133900 0.0444920 -0.59774 Methyltransferase domain Glyma.13G205800 0.0135970 -0.59610 Hsp20/alpha crystallin family Glyma.17G083500 0.0382410 -0.59575 Ribosomal protein L22p/L17e Glyma.07G06000 0.0382410 -0.59479 Ribosomal protein L22p/L17e Glyma.07G07500 0.0047921 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.07G19300 0.0047921 -0.59322 Go S Acidic ribosomal protein L44 Glyma.07G219300 0.0093614 -0.59232 Ribosomal protein L18e/L15 Glyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.046229000 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.1062219000 0.0026438 -0.58874 NA Glyma.106022000 0.0037612 -0.58606 Go Acidic ribosomal protein Glyma.10605000 0.0376120 -0.58676 Ribosomal protein L37e Glyma.10615700 0.0376120 -0.58676 <td>Glyma.03G207500 Glyma.16G046500</td> <td>0.0047055</td> <td>-0.60153 -0.59941</td> <td>Ribosomal protein L21e Ribosomal protein S27</td>	Glyma.03G207500 Glyma.16G046500	0.0047055	-0.60153 -0.59941	Ribosomal protein L21e Ribosomal protein S27
GIýma.04G133900 0.0444920 -0.59774 Methyltransferåse domain Glýma.13G205800 0.0135970 -0.59610 Hsp20/alpha crystallin family Glýma.13G205800 0.0382410 -0.59575 Ribosomal protein L22p/L17e Glýma.08G151900 0.0020485 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.05G174300 0.0047921 -0.59322 Glo Acidic ribosomal protein L4e Glyma.07G219500 0.0093614 -0.59232 Ribosomal protein L44 Glyma.07G219500 0.0093614 -0.59238 Glyma.03G215300 0.0002615 -0.59058 Ribosomal protein L18e/L15 Glyma.03G215300 0.0072003 -0.58997 Glyma.13G21800 0.026615 -0.59058 Ribosomal protein L18e/L15 Glyma.19G117300 0.026438 -0.58867 Glyma.19G117300 0.026438 -0.58868 60s Acidic ribosomal protein Glyma.146062900 0.021633 -0.58847 Glyma.06G128000 0.0376120 -0.58876 Ribosomal protein S27 Glyma.04G229000 0.0216370 -0.58847 Glyma.06G128000 0.0326240 -0.58234 KObsomal protein L37e Glyma.04G04900 0.025670 <td>Glvma.07G060000</td> <td>0.0151550</td> <td>-0.59888</td> <td>Ribosomal L18p/L5e family</td>	Glvma.07G060000	0.0151550	-0.59888	Ribosomal L18p/L5e family
Glyma.13G203800 0.0135970 -0.59610 Hsp20/appa crystalin Tamity Glyma.13G203800 0.0382410 -0.59575 Ribosomal protein L22pL/L7e Glyma.08G151900 0.0020485 -0.59349 Ribosomal protein L22pL/L7e Glyma.08G151900 0.0047921 -0.59328 Ribosomal protein L6e ; Ribosomal protein Glyma.07G219500 0.0476240 -0.59232 Ribosomal protein L44 Glyma.07G219500 0.0093614 -0.59232 Ribosomal protein L8e/L15 Glyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.03G215300 0.0026515 -0.58057 Ribosomal protein L18e/L15 Glyma.19G211800 0.0316120 -0.58628 Eukaryotic protein Glyma.04G229000 0.0026438 -0.58638 NA Glyma.04G229000 0.0276670 -0.58437 Ribosomal protein S27 Glyma.04G128000 0.0326240 -0.58437 Ribosomal protein S27 Glyma.04G128000 0.0326240 -0.58437 Ribosomal protein L6e, r.Ribosomal protein L6, R149040 Glyma.04G128000 0.0326240 <td>Glyma.04G133900</td> <td>0.0444920</td> <td>-0.59774</td> <td>Methyltransferase domain</td>	Glyma.04G133900	0.0444920	-0.59774	Methyltransferase domain
Glyma.17G083500 0.0489140 -0.59449 Ribosomal protein L7Ae/L30e/S12e/Gadd45 filyma.08G151900 0.0020485 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.05G174300 0.0047921 -0.59322 Ribosomal protein L4e ; Ribosomal protein L44 Glyma.07G219500 0.0076240 -0.59232 Ribosomal protein L18e/L15 Glyma.07G219500 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.19G117300 0.0230060 -0.58877 NP Glyma.19G117300 0.0220060 -0.58876 Ribosomal protein L18e/L15 Glyma.07G229500 0.007203 -0.58876 Ribosomal protein L18e/L15 Glyma.19G117300 0.0230060 -0.58876 Ribosomal protein S27 Glyma.04G229000 0.002243 -0.58676 Ribosomal protein L37e Glyma.07G04900 0.0275670 -0.58875 Ribosomal protein L4e ; Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.07G04900 0.0275670 -0.58437 Ribosomal protein L37e Glyma.07G04900 0.026240 -0.58284 CS domain Glyma.12G024500 0.0376120 -0.58284 CS domain <t< td=""><td>Glyma.13G205800 Glyma.18G007500</td><td>0.0135970</td><td>-0.59610 -0.59575</td><td>Ribosomal protein L22p/L17e</td></t<>	Glyma.13G205800 Glyma.18G007500	0.0135970	-0.59610 -0.59575	Ribosomal protein L22p/L17e
family Glyma.08G151900 0.0020485 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.05G174300 0.0047921 -0.59322 60s Acidic ribosomal protein 144 Glyma.07G219500 0.0076240 -0.59232 Ribosomal protein L44 144 Glyma.07G219500 0.0026515 -0.59058 Ribosomal protein L18e/L15 15 Glyma.103C250500 0.0072003 -0.58897 TCP-1/cpn60 chaperonin family 164 Glyma.19G117300 0.026515 -0.58876 Ribosomal protein L18e/L15 15 Glyma.04G229000 0.0026438 -0.58876 Ribosomal protein L18e/L15 16 Glyma.09G105700 0.0376120 -0.58876 Ribosomal protein L18e/L15 17 Glyma.09G240600 0.0002243 -0.58623 Eukaryotic porin 18/ Glyma.09G12000 0.216330 -0.58847 NA 18/ Glyma.07G004900 0.0216570 -0.58847 Ribosomal protein L37e 19/ Glyma.06G128000 0.0326240 -0.58235 KOW motif 19/	Glyma.17G083500	0.0489140	-0.59449	Ribosomal protein L7Ae/L30e/S12e/Gadd45
Glyma.08G151900 0.0020485 -0.59349 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.05G174300 0.00476240 -0.59232 Gos Acidic ribosomal protein L44 Glyma.02G038700 0.0093614 -0.59232 Ribosomal protein L184/L15 Glyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.13G25500 0.0072003 -0.588997 TCP-1/cpn60 chaperonin family Glyma.19G211800 0.026604 -0.58765 Ribosomal protein L18e/L15 Glyma.19G211800 0.0216438 -0.58676 Ribosomal protein S27 Glyma.19G240600 0.0026438 -0.58868 60s Acidic ribosomal protein L37e Glyma.09G240600 0.002703 -0.58847 Ribosomal protein L37e Glyma.09G240600 0.0026438 -0.58868 NA Glyma.09G240600 0.0216330 -0.58847 Ribosomal protein L37e Glyma.09G240600 0.00275670 -0.58847 Ribosomal protein L46e ; Ribosomal protein L6, N-terminal domain Glyma.05G141700 0.0326240 -0.58284 CS domain Glyma.106G128000 0.0326240 -0.58284 CS domain Glyma				family
Glyma.05G174300 0.0047921 -0.59322 60s Acidic ribosomal protein Glyma.08G075500 0.0476240 -0.59232 Ribosomal protein L44 Glyma.02G038700 0.0010664 -0.59216 Ribosomal protein L18e/L15 Glyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.03G215300 0.0026515 -0.59058 Ribosomal protein L18e/L15 Glyma.19G211800 0.0220060 -0.58897 TCP-1/cpn60 chaperonin family Glyma.19G211800 0.024438 -0.58686 60s Acidic ribosomal protein L18e/L15 Glyma.19G105700 0.0376120 -0.58676 Ribosomal protein S27 Glyma.19G0240600 0.002243 -0.58823 Eukaryotic porin Glyma.09G240600 0.0275670 -0.58427 Ribosomal protein L37e Glyma.08G141700 0.0156990 -0.58235 KOW motif Glyma.12G024500 0.0478990 -0.58235 KOW motif Klosomal protein Staperotein Staperotein L4, Noterminal domain Glyma.12G0617700 0.0042169 -0.58774	Glyma.08G151900	0.0020485	-0.59349	Ribosomal protein L6e ; Ribosomal protein L6,
	Glyma.05G174300	0.0047921	-0.59322	60s Acidic ribosomal protein
Glyma.070219300 0.0093014 -0.39230 Ribosomal protein L18e/L15 Glyma.03G215300 0.0010664 -0.59165 Ribosomal protein L18e/L15 Glyma.19G215300 0.0072003 -0.58765 Ribosomal protein L18e/L15 Glyma.19G211800 0.0418840 -0.58765 Ribosomal protein L18e/L15 Glyma.19G211800 0.0026438 -0.58765 Ribosomal protein S27 Glyma.04G229000 0.002243 -0.58868 60s Acidic ribosomal protein Glyma.04G2000 0.002243 -0.58868 NA Glyma.04G624000 0.002243 -0.58886 NA Glyma.04G024000 0.0275670 -0.588437 Ribosomal protein L37e Glyma.06G128000 0.0326240 -0.58284 CS domain Glyma.15G068500 0.0478990 -0.58235 KOW motif Glyma.12G024500 0.0162680 -0.58182 Ribosomal S17 Glyma.12G024500 0.0022648 -0.58182 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.15G068500 0.0478990 -0.58235 KOW motif Ribosomal S17 Glyma.12G024500 0.0162680 -0.57704 Ribosomal protein	Glyma.08G075500	0.0476240	-0.59232	Ribosomal protein L44
	Glyma.02G038700	0.0093014	-0.59165	Ribosomal protein L18e/L15
	Glyma.03G215300	0.0026515	-0.59058	Ribosomal protein L18e/L15
Glyma.19G117300 0.0250000 -0.58765 Ribosomal protein L18e/L15 Glyma.19G211800 0.0418840 -0.58765 Ribosomal protein S27 Glyma.19G105700 0.0376120 -0.58676 Ribosomal protein S27 Glyma.09G240600 0.0002243 -0.58623 Eukaryotic porin Glyma.14G062900 0.0275670 -0.58437 Ribosomal protein L37e Glyma.07G004900 0.0275670 -0.58421 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.06G128000 0.0326240 -0.58284 CS domain Glyma.12G017700 0.0002168 -0.58182 Ribosomal S17 Glyma.12G024500 0.0162680 -0.58182 Ribosomal protein S12/S23 Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.19G178800 0.0042129 -0.57029 Ribosomal protein S8 Glyma.11G234300 0.0239870 -0.56491 NA Glyma.10G034000 0.0322470 -0.56191 NA Glyma.10G034000 0.0322470 -0.56342 NA Glyma.10G034000 0.0322470 -0.56342 NA Glyma.10G160300	Glyma.15G250500	0.0072003	-0.58997	TCP-1/cpn60 chaperonin family
Glyma.04G229000 0.0026438 -0.58680 60s Acidic ribosomal protein Glyma.19G105700 0.0376120 -0.58676 Ribosomal protein S27 Glyma.09G240600 0.0002243 -0.58623 Eukaryotic porin Glyma.14G062900 0.0216330 -0.58863 NA Glyma.07G004900 0.0275670 -0.58437 Ribosomal protein L37e Glyma.08G141700 0.0156990 -0.58284 CS domain Glyma.15G068500 0.0478990 -0.58284 CS domain Glyma.12G024500 0.002168 -0.58182 Ribosomal S17 Glyma.12G024500 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.04G063000 0.0239870 -0.57029 Ribosomal protein S8 Glyma.17G143200 0.0252270 -0.56919 NA Glyma.10G034000 0.0324470 -0.56382 NA Glyma.10G03600 0.0322470 -0.56307 Ribosomal L39 protein Glyma.10G03600 0.0322470 -0.56347 ptKB family carbohydrate kinase Glyma.10G160300 0.0351030	Glyma.19G211800	0.0230000	-0.58765	Ribosomal protein L18e/L15
	Glyma.04G229000	0.0026438	-0.58680	60s Acidic ribosomal protein
Glyma.040602900 0.00216330 -0.58286 NA Glyma.07G004900 0.0275670 -0.58437 Ribosomal protein L37e Glyma.08G141700 0.0156990 -0.58284 CS domain Glyma.06G128000 0.0326240 -0.58284 CS domain Glyma.15G068500 0.0478990 -0.58235 KOW motif Glyma.12G024500 0.0002168 -0.58182 Ribosomal S17 Glyma.12G024500 0.0162680 -0.57704 Ribosomal protein S12/S23 Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.17G143200 0.0252270 -0.56919 NA Glyma.11G234300 0.0449990 -0.56849 NA Glyma.10G034000 0.0332470 -0.56382 NA Glyma.10G160300 0.0351030 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.03G245300 0.0252270 -0.56307 Ribosomal protein S24e Glyma.03G245300 0.0252270 -0.56307 Ribosomal protein S24e<	Glyma.19G105700 Glyma.09G240600	0.0376120	-0.58676	Ribosomal protein S27 Fukarvotic porin
Glyma.07G004900 0.0275670 -0.58437 Ribosomal protein L37e Glyma.08G141700 0.0156990 -0.58421 Ribosomal protein L6e ; Ribosomal protein L6, N-terminal domain Glyma.06G128000 0.0326240 -0.58284 CS domain Glyma.15G068500 0.0478990 -0.58235 KOW motif Glyma.20G017700 0.0002168 -0.58182 Ribosomal S17 Glyma.12G024500 0.0162680 -0.58125 Translocon-associated protein (TRAP), alpha subunit Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.17G143200 0.0259270 -0.56919 NA Glyma.10G034000 0.032470 -0.56849 NA Glyma.10G034000 0.0322470 -0.56842 NA Glyma.10G160300 0.0322470 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.03G245300 0.0252270 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.56757 Ribosomal protein S24e <tr< td=""><td>Glyma.14G062900</td><td>0.0216330</td><td>-0.58586</td><td>NA</td></tr<>	Glyma.14G062900	0.0216330	-0.58586	NA
Glyma.06G128000 0.0326240 -0.58284 CS domain Glyma.15G068500 0.0478990 -0.58235 KOW motif Glyma.20G017700 0.0002168 -0.58182 Ribosomal S17 Glyma.12G024500 0.0162680 -0.58125 Translocon-associated protein (TRAP), alpha subunit Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.17G143200 0.02239870 -0.56849 NA Glyma.10G034000 0.0322470 -0.56849 NA Glyma.10G034000 0.0322470 -0.56382 NA Glyma.10G160300 0.0322470 -0.56382 NA Glyma.10G16300 0.0351030 -0.56382 NA Glyma.10G16300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56041 50S ribosomal protein S24e Glyma.03G245300 0.0252270 -0.56757 Ribosomal protein S24e Glyma.03G245300 0.0252270 -0.56382 NA Glyma.03G245300 0.0252270 -0.56041 50	Glyma.07G004900 Glyma.08G141700	0.0275670	-0.58437 -0.58421	Ribosomal protein L3/e Ribosomal protein L6e : Ribosomal protein L6.
		010120770	0.00121	N-terminal domain
Glyma.15G068500 0.0478990 -0.58235 KOW mothf Glyma.20G017700 0.0002168 -0.58182 Ribosomal S17 Glyma.12G024500 0.0162680 -0.58125 Translocon-associated protein (TRAP), alpha subunit Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.17G143200 0.0252270 -0.56919 NA Glyma.10G034000 0.0332470 -0.56649 NA Glyma.10G034000 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56641 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e	Glyma.06G128000	0.0326240	-0.58284	CS domain
Glyma.12G024500 0.0162680 -0.58125 Translocon-associated protein (TRAP), alpha subunit Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.17G143200 0.0252270 -0.56849 NA Glyma.10G034000 0.0322470 -0.56649 NA Glyma.10G034000 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G663000 0.032270 -0.56641 50S ribosomal protein S24e Glyma.10G160300 0.0322470 -0.56377 Ribosomal protein S24e Glyma.04G663000 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G6160300 0.0252270 -0.56757 Ribosomal protein S24e Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55767 Ribosomal protein S24e	Glyma.15G068500 Glyma 20G017700	0.0478990	-0.58235 -0.58182	KOW motif Ribosomal S17
subunit subunit Glyma.19G178800 0.0042129 -0.57704 Ribosomal protein S12/S23 Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.04G063000 0.0239870 -0.57029 Ribosomal L39 protein Glyma.17G143200 0.0252270 -0.56919 NA Glyma.10G034000 0.0332470 -0.56849 NA Glyma.10G034000 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55764 Ribosomal protein S24e	Glyma.12G024500	0.0162680	-0.58125	Translocon-associated protein (TRAP), alpha
Glyma.07G255500 0.0039289 -0.57113 Ribosomal protein S8 Glyma.04G063000 0.0239870 -0.57029 Ribosomal L39 protein Glyma.17G143200 0.0252270 -0.56919 NA Glyma.11G234300 0.0449990 -0.56849 NA Glyma.10G034000 0.0332470 -0.56547 pfkB family carbohydrate kinase Glyma.10G160300 0.0351030 -0.56382 NA Glyma.04G219100 0.0456050 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55647 Ribosomal protein S24e	Glyma.19G178800	0.0042129	-0.57704	subunit Ribosomal protein S12/S23
Glyma.04G063000 0.0239870 -0.57029 Ribosomal L39 protein Glyma.17G143200 0.0252270 -0.56919 NA Glyma.11G234300 0.0449990 -0.56849 NA Glyma.10G034000 0.0332470 -0.56547 pfkB family carbohydrate kinase Glyma.10G034000 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56037 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56047 S0S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55647 Ribosomal protein S24e	Glyma.07G255500	0.0039289	-0.57113	Ribosomal protein S8
Glyma.11G234300 0.0439200 -0.56849 NA Glyma.10G034000 0.0332470 -0.56849 NA Glyma.19G047300 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.5757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55647 Ribosomal protein L14p/L23e	Glyma.04G063000 Glyma 17G143200	0.0239870	-0.57029	Ribosomal L39 protein
Glyma.10G034000 0.0332470 -0.56547 pfkB family carbohydrate kinase Glyma.19G047300 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55664 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55664 Ribosomal protein L14p/L23e	Glyma.11G234300	0.0449990	-0.56849	NA
Glyma.19G047300 0.0382410 -0.56382 NA Glyma.10G160300 0.0351030 -0.56307 Ribosomal protein S24e Glyma.04G219100 0.0456050 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55664 Ribosomal protein L14p/L23e	Glyma.10G034000	0.0332470	-0.56547	ptkB family carbohydrate kinase
Glyma.04G219100 0.0456050 -0.56041 50S ribosome-binding GTPase; TGS domain Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55664 Ribosomal protein L14p/L23e	Glyma.19G047300 Glyma.10G160300	0.0382410 0.0351030	-0.56382	NA Ribosomal protein \$24e
Glyma.03G245300 0.0252270 -0.55757 Ribosomal protein S24e Glyma.03G166700 0.0164990 -0.55664 Ribosomal protein L14p/L23e	Glyma.04G219100	0.0456050	-0.56041	50S ribosome-binding GTPase; TGS domain
	Glyma.03G245300 Glyma.03G166700	0.0252270	-0.55757 -0.55664	Ribosomal protein S24e Ribosomal protein L14p/L23e

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.15G051000	0.0107310	-0.55664	Ribosomal protein L44
Glyma.04G052800	0.0055661	-0.55614	Domain of unknown function (DUF3511)
Glyma.13G140100	0.0359220	-0.55573	Ribosomal protein L11, N-terminal domain;
Glyma.14G056900	0.0065616	-0.55422	RS4NT (NUC023) domain; S4 domain; KOW
- ,			motif; Ribosomal family S4e
Glyma.09G024000	0.0336040	-0.55395	Ribosomal protein L13
Glyma.05G186900	0.0234220	-0.55341	Ribonucleotide reductase, all-alpha domain;
			Ribonucleotide reductase, barrel domain; ATP
Glyma 02G253500	0.0452200	0 55126	cone domain
Glyma.12G041300	0.0034074	-0.55106	S25 ribosomal protein
Glyma.10G221600	0.0259480	-0.54810	Ribosomal protein \$13/\$18
Glyma.08G346500	0.0127720	-0.54440	Ribosomal protein S8e
Glyma.17G026800	0.0407030	-0.54237	NA
Glyma.05G238000 Glyma.02G049300	0.042/690	-0.53972	NA Ribosomal Proteins I 2 RNA binding domain:
Olyma.020047500	0.0017400	-0.55771	Ribosomal Proteins L2, River binding domain,
Glyma.05G234800	0.0103160	-0.53765	Ribosomal protein S19
Glyma.13G051200	0.0202070	-0.53699	Ribosomal protein L32
Glyma.17G100600	0.0106510	-0.53538	lactate/malate dehydrogenase, alpha/beta
			C-terminal domain; lactate/malate
			dehydrogenase, NAD binding domain
Glyma.04G226900	0.0495570	-0.53409	Ribosomal protein L6
Glyma.13G217200 Glyma.04G020800	0.0422000	-0.53065	20G-Fe(II) oxygenase superfamily
Glyma.15G003700	0.0162460	-0.52839	NA
Glyma.15G271300	0.0357040	-0.52779	Ribosomal protein L6e ; Ribosomal protein L6,
2			N-terminal domain
Glyma.02G002300	0.0452200	-0.52773	Ribosomal protein L7Ae/L30e/S12e/Gadd45
Clama 16C120700	0.0107920	0.52670	family Bibasenal Destains I 2, DNA binding demains
Glyma.16G129700	0.0107850	-0.32079	Ribosomal Proteins L2, RNA binding domain;
Glvma.03G250800	0.0387510	-0.52404	Beta-eliminating lyase
Glyma.08G196500	0.0330070	-0.52283	Ribosomal L28e protein family
Glyma.08G222100	0.0018141	-0.52248	Ribosomal L30 N-terminal domain; Ribosomal
			protein L30p/L7e
Glyma.07G068750	0.0374620	-0.52197	NA United a second C terminal demains
Glyma.04G145400	0.0116820	-0.52125	Helicase conserved C-terminal domain;
Glvma.13G151600	0.0251250	-0.51928	NA
Glyma.02G017800	0.0360940	-0.51795	Ribosomal L22e protein family
Glyma.20G018500	0.0213650	-0.51603	SGS domain ; Siah interacting protein, N terminal
•			; CS domain
Glyma.09G156600	0.0221590	-0.51566	Ribosomal protein S3, C-terminal domain; KH
Gluma 15G120100	0.0257040	0 51262	domain Brafaldin subunit
Glyma.19G154900	0.0337040	-0.51203	NA
Glyma.08G087000	0.0212320	-0.50680	NA
Glyma.17G004300	0.0185310	-0.50602	Ribosomal protein S19
Glyma.01G173050	0.0419560	-0.50554	NA Dibasemal matrix 174 a /1 20a /212a /Cadd45
Glyma.08G257100	0.0127390	-0.50408	family
Glyma.02G036000	0.0134720	-0.50390	Ubiquitin family: Ribosomal protein S27a
Glyma.20G168100	0.0377800	-0.50310	NA
Glyma.09G094200	0.0476240	-0.50088	Ribosomal protein S17
Glyma.04G221400	0.0077631	-0.50007	Ribosomal protein S26e
Glyma 06G090100	0.0062500	-0.49698 -0.49683	NA
Glyma.02G167900	0.0133730	-0.49489	Ribosomal protein S7p/S5e
Glyma.06G138400	0.0419800	-0.49438	Ribosomal protein L6
Glyma.09G031800	0.0142720	-0.49358	S1 RNA binding domain
			(Conntinua en la pág siguiente)

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.07G023900	0.0467590	-0.49342	NA
Glyma.09G017300	0.0305480	-0.49240	Myb-like DNA-binding domain
Giyina.09G017500	0.0303480	-0.49240	Myd-like DNA-dinding domain
Glyma.01G004300	0.0309530	-0.49231	NA lastate/malate.dahudroganase.almha/hate
Giyina.10G000500	0.0559820	-0.48955	C taminal languing hastata (an lata
			C-terminal domain; lactate/malate
Glyma 06G161200	0.0476240	0 48874	dehydrogenase, NAD binding domain
Glyma.16G126200	0.0076546	-0.48708	Proteasome subunit A N-terminal signature:
5			Proteasome subunit
Glyma.11G241000	0.0156990	-0.48577	DnaJ domain; DnaJ C terminal domain; DnaJ
			central domain
Glvma.07G020000	0.0099660	-0.48438	Ribosomal protein L30p/L7e: Ribosomal L30
01/11/07/0020000	0.0077000	0110100	N-terminal domain
Glyma.19G020400	0.0419560	-0.47880	WD domain, G-beta repeat
Glyma.13G099200	0.0449390	-0.47813	Methyltransferase domain; Methyltransferase
~			involved in Williams-Beuren syndrome
Glyma.03G104300	0.0290760	-0.47314	20G-Fe(II) oxygenase superfamily
Giyina.100148500	0.0208100	-0.47140	Ribosoinai protein S27
Glyma.03G189300	0.0424300	-0.47050	Ribosomal protein L24e
Glyma 16G087700	0.0339820	-0.47045	NA Ribosomal protein S8e
Glyma.01G029200	0.0311470	-0.46179	Ubiquitin family; Ribosomal protein S27a
Glyma.17G032200	0.0342900	-0.45974	Ribosomal L29e protein family
Glyma.11G014600	0.0331470	-0.45886	De-etiolated protein 1 Det1
Glyma.20G206300	0.0335760	-0.45700	POT family
Glyma.18G213900	0.0444920	-0.45694	Elongation factor P, C-terminal; Elongation factor
			P (EF-P) OB domain; Elongation factor P (EF-P)
GI 10 GI 15000	0.00000	0.45500	KOW-like domain
Glyma 11G00000	0.0208670	-0.45580	Ribosomal protein S10p/S20e
Olyma.110099000	0.0330040	-0.43391	
Glyma.02G103600 Glyma.11G231800	0.0436660	-0.45322	Ribosomal protein S9/S16
Glyma.10G133800	0.0156990	-0.44927	Ribosomal protein S3. C-terminal domain: KH
-)			domain
Glyma.09G114000	0.0382620	-0.44810	NA
Glyma.08G067000	0.0410640	-0.44732	NA
Glyma.04G011100	0.0235150	-0.44637	S25 ribosomal protein
Glyma.13G319300	0.0207910	-0.44592	MORN repeat
Glyma 19G210600	0.0131330	-0.43882	CS domain
Glyma.14G001000	0.0268800	-0.43830	GHMP kinases C terminal ; GHMP kinases N
2			terminal domain
Glyma 03G213802	0.0145090	-0.43617	NA
Glyma.18G255700	0.0062564	-0.43548	Eukaryotic porin
Glyma.03G190500	0.0089005	-0.43537	Enolase, C-terminal TIM barrel domain; Enolase,
G1 00 G000500	0.0000050	0.40465	N-terminal domain
Glyma.09G023700 Glyma.10G048600	0.0292050	-0.43467	Diquitin family; Ribosomal protein S2/a
Glyma.100048000	0.0140730	-0.43431	Ribosomai protein L4/L1 family
Glyma.13G300300	0.0493640	-0.43381	Ribosomal protein S/p/S5e S4 domain: Pibosomal protein S4/S0 N terminal
Grynna.050028000	0.0093014	-0.43347	domain, Kibosoinai protein 54/59 N-terminai
Glyma.04G056000	0.0378930	-0.43156	YGGT family
Glyma.14G075000	0.0176100	-0.43106	Radical SAM superfamily; Flavodoxin; Wyosine
			base formation
Glyma.19G232500	0.0214240	-0.42355	BRCA1 C Terminus (BRCT) domain; Pescadillo
			N-terminus
Glyma.06G109300	0.0330350	-0.42328	Ribosomal protein S13/S18
Glyma.02G220000	0.0318680	-0.41590	Ribosomal protein L16p/L10e
Glyma 1/G099200	0.0466500	-0.41103	Nucleosome assembly protein (NAP) Protein tyroging kingse: Lauging Righ Parast
Oryma.140037400	0.0009008	-0.40930	Lausing righ repeat N terring 1 downing
Glyma 12G030200	0.0189440	-0 40933	Plant specific enkaryotic initiation factor 4R
	5.5157110	0.10/00	

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.02G273100 Glyma.18G043600	$0.0320120 \\ 0.0178600$	-0.40630 -0.39642	Hydroxymethylglutaryl-coenzyme A reductase Pyridoxal-dependent decarboxylase conserved
Glyma.19G189600 Glyma.14G025900 Glyma.13G217100	$\begin{array}{c} 0.0491900 \\ 0.0290760 \\ 0.0378380 \end{array}$	-0.39537 -0.38527 -0.38199	Ribosomal protein L24e LSM domain Ribosomal L18ae/LX protein domain
Glyma.12G039900 Glyma.08G209100 Glyma.12G085900 Glyma.05G013500 Glyma.16G008900	$\begin{array}{c} 0.0325090\\ 0.0234010\\ 0.0201130\\ 0.0443580\\ 0.0315300 \end{array}$	-0.33509 -0.30195 0.36463 0.38760 0.38961	NAC domain TLC ATP/ADP transporter Histidine phosphatase superfamily (branch 1) TFIIE alpha subunit NA
Glyma.20G144400 Glyma.10G298000	0.0495570 0.0156990	0.39992 0.42132	NA Starch binding domain; Pyruvate phosphate
Glyma.19G131400 Glyma.15G200800 Glyma.17G255602	$\begin{array}{c} 0.0439220\\ 0.0248980\\ 0.0359220 \end{array}$	$\begin{array}{c} 0.42280 \\ 0.42818 \\ 0.43643 \end{array}$	Lycopene cyclase protein Methyltransferase domain NA
Glyma.04G028400 Glyma.15G124700	0.0239870 0.0444920	0.44424 0.44425	ATP-dependent protease La (LON) domain Ferric reductase like transmembrane component; FAD-binding domain; Ferric reductase NAD
Glyma.18G286100 Glyma.12G002500 Glyma.02G019200	$0.0341890 \\ 0.0329140 \\ 0.0145090$	0.44467 0.44472 0.45231	binding domain Cation efflux family NA NA
Glyma.10G215200 Glyma.14G129700 Glyma.05G205100 Glyma.07G177100 Glyma.18G138300	$\begin{array}{c} 0.0234600\\ 0.0261740\\ 0.0418840\\ 0.0093614\\ 0.0335350\end{array}$	$\begin{array}{c} 0.45337 \\ 0.45642 \\ 0.45807 \\ 0.46583 \\ 0.46589 \end{array}$	Zinc finger, C2H2 type Enhancer of polycomb-like Methyl-CpG binding domain PPR repeat Zinc-binding dehydrogenase; Alcohol
Glyma.13G192200 Glyma.07G012500	$0.0393350 \\ 0.0232040$	0.47045 0.47130	dehydrogenase GroES-like domain NA Glu/Leu/Phe/Val dehydrogenase, dimerisation domain; Glutamate/Leucine/Phenylalanine/Valine
Glyma.14G190300 Glyma.12G111000 Glyma.19G173300	0.0449400 0.0310120 0.0343570	0.47338 0.47541 0.47618	dehydrogenase NA Lanthionine synthetase C-like protein NA
Glyma.20G127800	0.0125280	0.48194	Partial alpha/beta-hydrolase lipase region;
Glyma.06G307800 Glyma.05G139800 Glyma.12G215500 Glyma.19G228100	0.0374620 0.0319020 0.0162460 0.0047118	$\begin{array}{c} 0.48435 \\ 0.48535 \\ 0.48880 \\ 0.49208 \end{array}$	alpha/beta hydrolase fold Dienelactone hydrolase family NADH dehydrogenase transmembrane subunit Core histone H2A/H2B/H3/H4 Redoxin; NHL repeat; haloacid dehalogenase-like hydrolase
Glyma.13G365500 Glyma.15G230500 Glyma.01G018400 Glyma.05G147400 Glyma.16G140800	$\begin{array}{c} 0.0352440\\ 0.0438750\\ 0.0112630\\ 0.0424340\\ 0.0399110 \end{array}$	0.49211 0.49315 0.49397 0.49401 0.49716	Methyl-CpG binding domain Exonuclease Helix-loop-helix DNA-binding domain Sel1 repeat; MYND finger Leucine Rich Repeat
Glyma.04G029700 Glyma.13G030500 Glyma.11G178100 Glyma.06G185200 Glyma.17G204700	$\begin{array}{c} 0.0407440\\ 0.0055325\\ 0.0468360\\ 0.0452300\\ 0.0234340 \end{array}$	0.49778 0.49981 0.50097 0.50111 0.50697	Protein kinase domain NA Core histone H2A/H2B/H3/H4 Protein of unknown function, DUF393 Enhancer of polycomb-like
Glyma.09G147500 Glyma.U031117 Glyma.01G081600 Glyma.20G141100 Glyma.02G238700	$\begin{array}{c} 0.0494590\\ 0.0485830\\ 0.0050490\\ 0.0235900\\ 0.0066636\end{array}$	$\begin{array}{c} 0.50945\\ 0.51505\\ 0.51920\\ 0.52190\\ 0.52370 \end{array}$	NA NA POT family Kinase/pyrophosphorylase NA
Glyma.05G150600 Glyma.12G043500	$\begin{array}{c} 0.0341890 \\ 0.0085864 \end{array}$	0.53216 0.54290	UDP-glucoronosyl and UDP-glucosyl transferase

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.17G137600	0.0113720	0.54876	Elongation factor Tu domain 2; Elongation factor G C-terminus; Elongation factor G, domain IV; Elongation factor Tu GTP binding domain
Glyma.05G051300	$\begin{array}{c} 0.0212320 \\ 0.0374860 \end{array}$	0.55129	NA
Glyma.08G084600		0.55458	NAD dependent epimerase/dehydratase family
Glyma.14G198700	$\begin{array}{c} 0.0290760\\ 0.0437930\\ 0.0339820\\ 0.0146120\\ 0.0081836 \end{array}$	0.55478	Calcineurin-like phosphoesterase
Glyma.16G199400		0.55845	Protein kinase domain
Glyma.06G121400		0.56188	Thioredoxin
Glyma.07G022700		0.56673	NA
Glyma.14G070400		0.57444	ATP-dependent protease La (LON) domain
Glyma.10G185400	$\begin{array}{c} 0.0043027\\ 0.0148980\\ 0.0195230\\ 0.0489120\\ 0.0024015 \end{array}$	0.57696	NA
Glyma.15G197200		0.57742	mTERF
Glyma.11G141600		0.58082	Core histone H2A/H2B/H3/H4
Glyma.03G211067		0.58127	NA
Glyma.18G037000		0.58279	Protein of unknown function (DUF1117)
Glyma.11G152124	$\begin{array}{c} 0.0042129\\ 0.0275320\\ 0.0024085\\ 0.0286750\\ 0.0378930 \end{array}$	0.58447	NA
Glyma.13G002200		0.58910	Domain of unknown function DUF221
Glyma.02G223600		0.59299	Zinc finger, C3HC4 type (RING finger)
Glyma.14G031800		0.59469	PsbP
Glyma.08G126000		0.59615	Ankyrin repeat; Protein tyrosine kinase
Glyma.13G315900 Glyma.07G097700 Glyma.04G191500 Glyma.12G090100 Glyma.12G001000	$\begin{array}{c} 0.0102870\\ 0.0023882\\ 0.0276920\\ 0.0423810\\ 0.0001855 \end{array}$	$\begin{array}{c} 0.60144\\ 0.60219\\ 0.60366\\ 0.60774\\ 0.61659\end{array}$	Protein of unknown function (DUF3049) Zinc finger, C3HC4 type (RING finger) NA Protein of unknown function (DUF3049) DnaJ domain
Glyma.08G361200 Glyma.10G047200 Glyma.08G063700	$\begin{array}{c} 0.0252270\\ 0.0378380\\ 0.0085687\end{array}$	$0.61705 \\ 0.61754 \\ 0.61895$	Sugar (and other) transporter NA FAD dependent oxidoreductase; Squalene
Glyma.09G217900	0.0469830	$0.62401 \\ 0.62494$	Uncharacterized protein family, UPF0114
Glyma.20G043800	0.0032715		MoaE protein
Glyma.13G235400	$\begin{array}{c} 0.0339820\\ 0.0120380\\ 0.0360360\\ 0.0229050\\ 0.0045209 \end{array}$	0.62599	DnaJ domain; DnaJ C terminal domain
Glyma.13G163700		0.62672	Uncharacterized protein family (UPF0051)
Glyma.10G043200		0.63009	NA
Glyma.01G196300		0.63041	Aminotransferase class IV
Glyma.15G253700		0.63208	Photosystem II 10 kDa polypeptide PsbR
Glyma.17G133200 Glyma.09G167900 Glyma.09G167900 Glyma.06G291300 Glyma.17G061600	$\begin{array}{c} 0.0047118\\ 0.0494590\\ 0.0494590\\ 0.0492670\\ 0.0329140 \end{array}$	$\begin{array}{c} 0.63925\\ 0.64129\\ 0.64129\\ 0.64129\\ 0.64142\\ 0.64705\end{array}$	NA Myb-like DNA-binding domain Myb-like DNA-binding domain NA Protein of unknown function (DUF1517)
Glyma.09G235600	$\begin{array}{c} 0.0047055\\ 0.0047055\\ 0.0422600\\ 0.0368710\\ 0.0418840 \end{array}$	0.64751	Myb-like DNA-binding domain
Glyma.09G235600		0.64751	Myb-like DNA-binding domain
Glyma.12G085500		0.65070	Prephenate dehydratase
Glyma.19G007800		0.65070	Trypsin
Glyma.13G253800		0.65114	SelR domain
Glyma.02G145100 Glyma.09G018600	0.0000127 0.0245470	0.65242 0.65799	STAS domain; Sulfate transporter family Ferric reductase NAD binding domain; Ferric reductase like transmembrane component; FAD-binding domain
Glyma.19G194992	$\begin{array}{c} 0.0279750\\ 0.0441960\\ 0.0133990 \end{array}$	0.65861	NA
Glyma.02G131700		0.66156	bZIP transcription factor
Glyma.08G101800		0.66335	Ankyrin repeat
Glyma.01G084700	$\begin{array}{c} 0.0160480\\ 0.0291760\\ 0.0000284\\ 0.0374620\\ 0.0268670 \end{array}$	0.66462	ATP-dependent Clp protease adaptor protein ClpS
Glyma.19G204100		0.66798	TRAF-type zinc finger
Glyma.06G043400		0.66981	F-box domain
Glyma.03G204800		0.67123	ABC-2 type transporter; ABC transporter
Glyma.15G245600		0.67182	C2 domain
Glyma.09G161900	$\begin{array}{c} 0.0112170\\ 0.0328960\\ 0.0487210\\ 0.0160910\\ 0.0125240 \end{array}$	0.67219	BRO1-like domain
Glyma.10G007900		0.67236	Enoyl-CoA hydratase/isomerase family
Glyma.13G324800		0.68117	Protein of unknown function (DUF707)
Glyma.09G282900		0.68145	Cytochrome P450
Glyma.01G161200		0.68194	NA
Glyma.01G213700	0.0047921	0.68248	Aminotransferase class I and II

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.11G090000	0.0071994	0.68347	EamA-like transporter family
Glyma.01G133501 Glyma.06G146900	0.0414630	0.68880	NA ABC transporter
Glyma.05G207700	0.0252270	0.69161	haloacid dehalogenase-like hydrolase
Glyma.06G004800	0.0009655	0.69624	Phosphatidylinositol-specific phospholipase C, X domain
Glyma.08G271900	0.0476240	0.69680	Helix-loop-helix DNA-binding domain
Glyma.10G255200 Glyma.17G194500	0.0110620	0.69809	NA
Glyma.06G300900	0.0339820	0.70886	B-box zinc finger
Glyma.05G007200	0.0006113	0.70984	Trypsin
Glyma.01G115100 Glyma.15G084600	0.0441960	0.71052	NA
Glyma.01G006200	0.0064395	0.71723	NA
Glyma.01G152700	0.0041737	0.71737	Uncharacterized protein family, UPF0114
Glyma.16G156400 Glyma.13G243400	0.0000047	0.71800	Zinc finger, C2H2 type Enovl-CoA hydratase/isomerase family
Glyma.07G226567	0.0027490	0.72070	NA
Glyma.01G186300	0.0084709	0.72102	Remorin, C-terminal region
Grynna.17G090700	0.0122800	0.72514	domain
Glyma.14G215100	0.0281740	0.72592	linker histone H1 and H5 family
Glyma.04G144900	0.0254050	0.72780	E1-E2 ATPase; haloacid dehalogenase-like
Glvma.17G029000	0.0374620	0.73592	NA
Glyma.17G152600	0.0036984	0.73858	Transferase family
Glyma.14G006900	0.0013205	0.73968	GRAS domain family
Glyma.12G187400 Glyma.09G091900	0.0489140	0.74011	CAF1 family ribonuclease
Glyma.13G301700	0.0072039	0.74055	Wound-induced protein
Glyma.13G304500	0.0361670	0.74279	Terpene synthase, N-terminal domain; Terpene
Glyma.05G163100	0.0181680	0.74366	cyclic phosphodiesterase-like protein
Glyma.02G131500	0.0468530	0.74506	Carboxyl transferase domain
Glyma.08G109700 Glyma.07G019500	0.0341820	0.74651	DnaJ domain START domain: Homeobox, domain
Glyma.07G019500	0.0275800	0.75541	START domain; Homeobox domain
Glyma.09G019602	0.0309650	0.76180	NA
Glyma.04G228400	0.0020842	0.76690	Sigma 54 modulation protein / S30EA ribosomal protein
Glyma.08G055000	0.0060572	0.76783	ABC transporter
Glyma.19G025900	0.0102350	0.76857	NA
Glyma.20G037400	0.0000931	0.77441	Zinc finger, C3HC4 type (RING finger)
Glyma.09G190600	0.0320200	0.77956	HSF-type DNA-binding
Glyma.13G043900 Glyma.18G036200	0.0332470	0.78309	Domain of unknown function (DUF296)
Glyma.02G063300	0.0495650	0.78660	alpha/beta hydrolase fold
Glyma.11G118100	0.010///0	0.78/9/	NA
Glyma.10G168200 Glyma 20G179800	0.0058731	0.78822	Ammonium Transporter Family
Glyma.13G130700	0.0399310	0.79223	NA
Glyma.17G094400 Glyma.13G149100	0.0000002	0.79389	Myb-like DNA-binding domain
Glyma 04G166000	0.02+3070	0.79010	Myh like DNA hinding domain
Glyma.08G220000	0.0328000	0.79727	NA
Glyma.09G168300	0.0082864	0.79732	Glycosyl hydrolase family 14
Glyma.17G109300	0.0441960	0.80239	NA
Glyma.14G053000	0.0476240	0.80266	NA
Glyma.08G213000	0.0436660	0.80439	NA No opical manistam (NAM) sustain
Glyma.10G107500	0.0160910	0.80694	Protein kinase domain
Glyma.04G089700	0.0160480	0.81542	Phosphoenolpyruvate carboxykinase

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.07G086400 Glyma.07G039900 Glyma.17G036500 Glyma.13G261400 Glyma.15G241300	$\begin{array}{c} 0.0000200\\ 0.0277290\\ 0.0001185\\ 0.0145720\\ 0.0219480 \end{array}$	0.81783 0.81877 0.82768 0.83092 0.83173	NA Lipoxygenase; PLAT/LH2 domain NA Metallo-beta-lactamase superfamily Transferase family
Glyma.02G202100 Glyma.02G285100 Glyma.10G168700 Glyma.20G136200	$\begin{array}{c} 0.0037615\\ 0.0275490\\ 0.0489140\\ 0.0257750 \end{array}$	0.83357 0.83656 0.83814 0.83898	NA NA NA Pyruvate kinase, barrel domain; Pyruvate kinase, alpha/beta domain
Glyma.09G198300	0.0000000	0.84145	HMG (high mobility group) box
Glyma.07G198200 Glyma.13G202300	0.0478990 0.0428140	0.84259 0.84667	Telomerase activating protein Est1 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B
Glyma.13G202300	0.0428140	0.84667	CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B
Glyma.13G227500 Glyma.08G308200	0.0043827 0.0030293	0.84837 0.85180	NA NA
Glyma.06G116200	0.0082118	0.85474	WD40-like Beta Propeller Repeat
Glyma.15G038500 Glyma.07G237700	0.0023184 0.0058955	0.85580 0.85795	ENTH domain GHMP kinases C terminal ; Galactokinase galactose-binding signature; GHMP kinases N
Glyma.10G142700	0.0331470	0.86129	Proteolipid membrane potential modulator
Glyma.10G112700	0.0427960	0.86320	Formate-tetranydrofolate ligase
Gryma.05G050000	0.0005170	0.00402	domain
Glyma.17G138500	0.0202070	0.86598	Domain of unknown function (DUF3357); Glycosyl hydrolases family 32 C terminal; Glycosyl hydrolases family 32 N-terminal domain
Glyma.13G166200	0.0199370	0.86679	NA
Glyma.15G176000	0.0080839	0.87011	NA
Glyma.10G001700 Glyma.11G141800 Glyma.10G062451 Glyma.13G142000 Glyma.18G039100	$\begin{array}{c} 0.0131950\\ 0.0000271\\ 0.0085186\\ 0.0086252\\ 0.0003783 \end{array}$	0.87087 0.87160 0.87339 0.87789 0.87789 0.87809	NAF domain; Protein kinase domain Core histone H2A/H2B/H3/H4 NA NA Armadillo/beta-catenin-like repeat
Glyma.01G022600 Glyma.13G326500 Glyma.13G330400 Glyma.08G335500	$\begin{array}{c} 0.0010377\\ 0.0414360\\ 0.0301640\\ 0.000026\end{array}$	$0.87854 \\ 0.88368 \\ 0.88580 \\ 0.89037$	U-box domain; Protein tyrosine kinase Zinc finger, C3HC4 type (RING finger) PAS fold; Protein kinase domain RbcX protein
Glyma.16G089000	0.0084709	0.89171	1-deoxy-D-xylulose 5-phosphate reductoisomerase C-terminal; 1-deoxy-D-xylulose 5-phosphate reductoisomerase
Glyma.20G239700	0.0003514	0.89278	PsbP
Glyma.06G177500 Glyma.08G056400 Glyma.04G184400	$\begin{array}{c} 0.0237860\\ 0.0343570\\ 0.0210740 \end{array}$	0.89684 0.90204 0.90315	NA Calcineurin-like phosphoesterase UvrB/uvrC motif; Protein of unknown function
Glyma.10G274400	0.0212320	0.90538	(DUF525) FAE1/Type III polyketide synthase-like protein; 3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal
Glyma.11G035966	0.0249620	0.90597	NA
Glyma.20G007900	0.0142550	0.90793	AMP-binding enzyme
Glyma.04G057500 Glyma.08G091900	$0.0332530 \\ 0.0467590$	0.91373 0.91378	BRCA1 C Terminus (BRCT) domain K+ potassium transporter
Glyma.09G254302 Glyma.11G239400	$0.0000109 \\ 0.0009240$	0.91419 0.91466	NA PLAT/LH2 domain

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.18G057900 Glyma.14G029700 Glyma.02G042400	$\begin{array}{c} 0.0065490 \\ 0.0268800 \\ 0.0222340 \end{array}$	0.91799 0.91985 0.92303	NAD dependent epimerase/dehydratase family NA Strictosidine synthase
Glyma.20G130300 Glyma.10G167800 Glyma.11G238700 Glyma.07G241000 Glyma.05G097800	$\begin{array}{c} 0.0007508\\ 0.0000862\\ 0.0420180\\ 0.0000369\\ 0.0008506\end{array}$	$\begin{array}{c} 0.92451\\ 0.92458\\ 0.92474\\ 0.92792\\ 0.92945\end{array}$	NA Ammonium Transporter Family Protein of unknown function (DUF581) NA FAD dependent oxidoreductase
Glyma.12G117700 Glyma.15G151300 Glyma.05G156300	0.0161570 0.0403450 0.0022736	0.93258 0.93272 0.93981	Myb-like DNA-binding domain NA Pheophorbide a oxygenase; Rieske [2Fe-2S]
Glyma.02G229900 Glyma.07G044900	$0.0237420 \\ 0.0378380$	0.94078 0.94216	domain Protein tyrosine kinase Protein of unknown function (DUF581)
Glyma.13G236500 Glyma.18G208300 Glyma.05G218400 Glyma.10G246000 Glyma.12G200500	$\begin{array}{c} 0.0341890\\ 0.0439190\\ 0.0006226\\ 0.0142750\\ 0.0111200\\ \end{array}$	0.94848 0.95140 0.95318 0.95362 0.95380	AP2 domain UDP-glucoronosyl and UDP-glucosyl transferase Glutaredoxin Endonuclease/Exonuclease/phosphatase family Wound-induced protein
Glyma.08G320400 Glyma.06G179800 Glyma.13G237500 Glyma.12G200200 Glyma.19G210402	$\begin{array}{c} 0.0028583\\ 0.0109340\\ 0.0357040\\ 0.0000057\\ 0.0310120\\ \end{array}$	0.95454 0.95693 0.96267 0.96379 0.96592	NA short chain dehydrogenase NA Thylakoid formation protein NA
Glyma.01G057700 Glyma.10G061800 Glyma.20G180800 Glyma.05G072100 Clyma.105001000	$\begin{array}{c} 0.0206200\\ 0.0327580\\ 0.0418840\\ 0.0476240\\ 0.0176660\end{array}$	0.97240 0.97581 0.97761 0.98106	Replication factor-A C terminal domain Thaumatin family Aromatic amino acid lyase NA
Glyma.10G094000 Glyma.11G036032 Glyma.09G073050 Glyma.06G136500	0.0176660 0.0202030 0.0427860 0.0000075	0.98111 0.98169 0.98232 0.98318	NA NA Sigma 54 modulation protein / S30EA ribosomal
Glyma.08G351932 Glyma.14G085600	0.0006099 0.0301170	0.98843 0.99251	NA
Glyma.19G144100	0.0000066	0.99263	Leucine rich repeat N-terminal domain; Protein kinase domain; Leucine Rich Repeat
Glyma.11G053100	0.0223100	0.99348	WRKY DNA -binding domain; Plant zinc cluster domain
Glyma.11G053100	0.0223100	0.99348	WRKY DNA -binding domain; Plant zinc cluster domain
Glyma.07G027400	0.0004973	0.99880	NA
Glyma.09G256100 Glyma.17G063200 Glyma.07G060400	$\begin{array}{c} 0.0307390 \\ 0.0006113 \\ 0.0000694 \end{array}$	0.99998 1.00440 1.00610	Glycosyl hydrolases family 28 Core histone H2A/H2B/H3/H4 bZIP transcription factor; G-box binding protein
Glyma.07G060400	0.0000694	1.00610	MFMR bZIP transcription factor; G-box binding protein
Glyma.07G060400	0.0000694	1.00610	MFMR bZIP transcription factor; G-box binding protein MFMR
Glyma.14G077100 Glyma.19G212000 Glyma.13G275500 Glyma.11G164700 Glyma.17G105600	$\begin{array}{c} 0.0489140\\ 0.0076839\\ 0.0254170\\ 0.0022736\\ 0.0137070\end{array}$	1.00740 1.00900 1.00910 1.01000 1.01420	Protein of unknown function, DUF607 Zinc finger, C3HC4 type (RING finger) NA NAD dependent epimerase/dehydratase family
Glyma.17G105000 Glyma.19G117600 Glyma.08G062300 Glyma.18G104000 Glyma.10G206600 Glyma.01G098700	0.0137970 0.0043827 0.0113940 0.0497260 0.0423300 0.0357040	1.01420 1.02050 1.03160 1.03350 1.03580 1.04190	Core-2/I-Branching enzyme (2R)-phospho-3-sulfolactate synthase (ComA) Transferase family NA NA
Glyma.13G321600 Glyma.02G128400 Glyma.18G015500	0.0211270 0.0355850 0.0003514	1.04570 1.04910 1.04950	Core histone H2A/H2B/H3/H4 NA Lanthionine synthetase C-like protein

Cuadro A.2: Continuación

Glyma.15(22) 300 0.0098212 1.05040 UDP-glucoronosyl and UDP-glucosyl transferase Glyma.14(G118100 0.000283 1.05100 NA Glyma.14(G118100 0.000283 1.05100 NA Glyma.14(G118100 0.002283 1.05120 NA Glyma.09(203650 0.0223100 1.05300 NA Glyma.09(213000 0.0226850 1.05720 NA Glyma.09(203650 0.0206860 NA NA Glyma.07(235300 0.0206860 NA NA Glyma.13(2254300 0.0399850 1.06550 NA Glyma.07(2000 0.0444200 1.06660 NA Glyma.13(2254300 0.0000004 1.07520 NA Glyma.15(2019500 0.0000000 1.08400 ABC transporter transmembrane region; ABC transporter disma.08(213000 0.0104760 1.08400 Glyma.15(21000 0.0104760 1.08400 NA Giyma.08(217600 0.0206230 Glyma.05(211100 0.0106724 1.09590 NA Giyma.07(254000 0.0206230 Info30 Giyma.07(Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.120096500 0.0153780 1.05100 NA Glyma.146118100 0.0009282 1.05120 NA Glyma.116122700 0.0000003 1.05300 Cytochrome P450 Glyma.116122700 0.00202825 1.05204 NA Glyma.090(140400 0.0223100 1.05420 NA Glyma.090(140400 0.0223100 1.05420 NA Glyma.016203560 0.023836 1.05720 NA Glyma.016213200 0.0251250 1.05940 NA Glyma.016213200 0.0251250 1.05940 NA Glyma.016213700 0.025420 1.06560 NA Glyma.016210700 0.0344200 1.06680 NA Glyma.016210700 0.0344200 1.06680 NA Glyma.016210700 0.0344200 1.06780 NA Glyma.016210700 0.000000 1.07530 NA Glyma.01621500 0.0000000 1.07530 NA Glyma.01621500 0.0000000 1.07530 NA Glyma.01520500 0.0000000 1.07500 NA Glyma.01520500 0.0000000 1.07500 NA Glyma.015021500 0.0000000 1.078400 ABC transporter transmembrane region; ABC transporter transporter transmembrane region; ABC transporter transmembrane region; ABC Glyma.01617960 0.026230 1.10520 Pathogenesis-related protein BC v 1 family Glyma.016167800 0.000020 1.11750 Cold acclimation protein WCOR413 Glyma.016167800 0.000002 1.12540 ACT domain Glyma.016167800 0.000002 1.12540 ACT domain Glyma.01604200 0.0001730 1.12500 Frontoscel strase fam	Glyma.15G221300	0.0098212	1.05040	UDP-glucoronosyl and UDP-glucosyl transferase
Glyma. 14(1)8100 0.0009282 1.05120 NA Glyma. 14(1)8100 0.0009282 1.05120 NA Glyma.15(083000 0.0424300 1.05300 NA Glyma.09(1)8400 0.022100 1.05420 NA Glyma.09(1)8400 0.022100 1.05420 NA Glyma.09(1)8400 0.022150 1.05420 NA Glyma.09(1)8000 0.027950 1.05420 NA Glyma.07(2)55300 0.0206860 1.06650 NA Glyma.13(2)54300 0.039850 1.06550 NA Glyma.13(2)54300 0.039850 1.06550 NA Glyma.13(2)54300 0.00004 1.06660 NA Glyma.07(2)5500 0.0000667 1.07620 NA Glyma.03(2)500 0.0000667 1.07620 NA Glyma.03(2)500 0.0000667 1.07620 NA Glyma.03(2)500 0.000006 1.08910 Hpt domain Glyma.05(2)1100 0.0544200 1.06660 NA Glyma.05(2)1100 0.05420 1.06660 NA Glyma.05(2)100 0.034200 1.06890 NA Glyma.05(2)1100 0.000000 1.08900 NA Glyma.05(2)1100 0.000000 1.08900 NA Glyma.05(2)1100 0.000000 1.08900 NA Glyma.05(2)1100 0.000000 1.0920 NA Glyma.05(2)1100 0.000000 1.0920 NA Glyma.05(2)1100 0.000000 1.0722 Cold acclimation protein WCOR413 Glyma.16(2)2600 0.0225270 1.10910 Transferase family 17; X8 domain Glyma.18(2)58000 0.000051 1.1720 Cold acclimation protein WCOR413 Glyma.18(2)58000 0.000051 1.1720 Cold acclimation protein WCOR413 Glyma.18(2)58000 0.000051 1.1720 Rdf1 carsoporter Glyma.18(2)58000 0.000051 1.1720 Rdf1 carsoporter Glyma.18(2)58000 0.000051 1.12500 Rdf1 carsoporter Glyma.18(2)58000 0.00252270 1.10910 Transferase family 17; X8 domain Glyma.18(2)58000 0.00252270 1.1910 Transferase family 17; X8 domain Glyma.18(2)58000 0.00252270 1.1910 Transferase family 16 Glyma.16(1)68100 0.000002 1.13200 ACT domain Glyma.16(1)68100 0.000002 1.13200 FACT domain Glyma.03(0)65500 0.11730 1.12400 NA Glyma.03(0)65500 0.01730 1.12400 NA Glyma.03(0)65500 0.01730 1.12400 NA Glyma.03(0)65500 0.01731 1.12400 NA Glyma.03(0)65500 0.017361 1.12400 NA Glyma.03(0)65500 0.017384 1.12160 SCC ACT domain Glyma.03(0)65500 0.0070814 1.14160 20G CCAT binding transcription factor (CBF-B/NF-YA) subunit B Glyma.03(0)65500 0.0070814 1.14160 ACT domain Glyma.03(0)60400 0.000002 1.13450 (CCAT binding transcription factor (CBF-B/NF-YA) subunit B Glyma.03(0)6000 0.000002 1.13450 (CCAT	Glyma.12G096500	0.0153780	1.05100	NA
Cityma.09(203660 0.0023300 1.03300 NA Cityma.09(203660 0.0023836 1.05720 NA Cityma.09(203660 0.0023836 1.05720 NA Cityma.07(225500 0.020660 NA NA Cityma.07(225500 0.020660 NA NA Cityma.07(225500 0.020660 NA NA Cityma.07(20500 0.0279950 1.06570 NA Cityma.07(20500 0.0279950 1.06570 NA Cityma.07(20500 0.0000004 1.07530 NA Cityma.07(20500 0.0000000 1.07520 NA Cityma.07(20500 0.0000000 1.08400 NA Cityma.07(20500 0.0000000 1.08400 NA Cityma.07(20500 0.0000000 1.09500 NA Cityma.15(019500 0.026230 1.0520 Peroxidase Cityma.15(20100 0.0026230 1.0520 Peroxidase Cityma.105(19400 0.022700 1.0720 Cold ascimation protein WCOR413 Cityma.107(256400 0.0225270 1.0750 Gityoosyl hydrolases family 17; X8 domain	Glyma.14G118100	0.0009282	1.05120	NA Cutochrome P450
Glýma.09CJ.000560 0.0223100 1.05420 NA Glyma.04CJ.031200 0.0251250 1.05700 NA Glyma.04CJ.031200 0.0251250 1.05700 NA Glyma.04CJ.031200 0.0251250 1.05600 NA Glyma.10CJ.03400 0.00279590 1.06507 NA Glyma.10CJ.04700 0.000004 1.06560 NA Glyma.07G08600 0.0001704 1.06880 NA Glyma.07G08600 0.000006 1.07510 NA Glyma.07G086100 0.000000 1.06800 NA Glyma.05C21100 0.000000 1.06900 Pathogenesis-related protein Bet v I family Glyma.05C21100 0.000000 1.0720 Cold acclimation protein WCOR413 Glyma.05C21100 0.000000 1.10720 Cold acclimation protein WCOR413 Glyma.01G167800 0.0206230 1.10720 Cold acclimation protein WCOR413 Glyma.11G122600 0.0000000 1.10720 Cold acclimation protein WCOR413 Glyma.11G122600 0.0001540 1.13800 Mamonium Transporter Family	Glyma.15G083000	0.0000003	1.05360	NA
Glyma.096203650 0.0023836 1.05720 NA Glyma.076255300 0.0206860 1.06050 NA Glyma.136254300 0.0399850 1.06570 NA Glyma.1362254300 0.0399850 1.065670 NA Glyma.1362254300 0.0001704 1.06880 NA Glyma.076086300 0.0001704 1.06880 NA Glyma.076086300 0.0001704 1.06880 NA Glyma.086245300 0.000000 1.07200 NA Glyma.086245300 0.0005667 1.07620 NA Glyma.086221100 0.0104760 1.08800 NA Glyma.086221100 0.0104760 1.08800 NA Glyma.086221100 0.0104760 1.08800 NA Glyma.086221100 0.0104760 1.08900 NA Glyma.056211100 0.0105320 1.09360 NA Glyma.056221100 0.0104760 1.08900 NA Glyma.056221100 0.0125320 1.09360 NA Glyma.0562200 0.000492 1.09630 NA Glyma.056200 0.000492 1.09630 NA Glyma.056200 0.000492 1.09630 NA Glyma.056194400 0.02026130 1.09500 NA Glyma.056194400 0.020263 1.10550 Peroxidase Glyma.056194400 0.02025270 1.10750 Glycosyl hydrolases family 17; X8 domain Glyma.076269400 0.0229220 1.10750 Glycosyl hydrolases family 17; X8 domain Glyma.106209200 0.0005420 1.11380 ABC transporter Glyma.076269400 0.0229220 1.10750 Glycosyl hydrolases family 17; X8 domain Glyma.106167800 0.0001881 1.12150 Raftinose synthase or seed imbibition protein Sip1 Glyma.146010500 0.0001841 1.12150 Raftinose synthase or seed imbibition protein Sip1 Glyma.146010500 0.000055 1.12460 NA Glyma.146010500 0.000055 1.12460 NA Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor CCBH-B/N-YA ysubunt B Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor CCBH-B/N-YA ysubunt B Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor CCBH-B/N-YA ysubunt B Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor CCBH-B/N-YA ysubunt B Glyma.036080600 0.017541 1.14990 Cytochase family 16 Glyma.036006500 0.0257410 1.15300 NA Glyma.036018000 0.007814 1.14160 200-FCHJ oxygenase superfamily Glyma.04501800 0.007855 1.15140 Matt Glyma.04501800 0.007874 1.15190 CAAT-binding transcription factor Glyma.04501800 0.007874 1.14190 Cytochase family 1 Glyma.04501800 0.007855 1.15140 Matt G	Glyma.09G140400	0.0223100	1.05420	NA
Glyma.01G131200 0.0251250 1.05940 NA Glyma.01G25300 0.020680 1.06550 NA Glyma.180005900 0.0279590 1.06570 NA Glyma.18005900 0.0279590 1.06580 NA Glyma.010210700 0.0344200 1.06680 NA Glyma.1960114700 0.000004 1.07530 NA Glyma.196021500 0.000000 1.08400 NA Glyma.156019500 0.0104760 1.08910 Hpt domain Glyma.05G211100 0.015320 1.09601 NA Glyma.05G21100 0.0102420 1.09601 NA Glyma.05G21100 0.020233 1.09601 NA Glyma.05G21900 0.000000 1.10720 Cold acclimation protein WCOR413 Glyma.01G167800 0.0202270 1.10910 Transferase family Glyma.14G010500 0.0001731 1.12300 ACT domain Glyma.14G010500 0.0001732 1.1240 Anmonium Transporter Glyma.14G010000 0.00000551 1.12100 Act doma	Glyma.09G203650	0.0023836	1.05720	NA
Glyma, 180023900 0.0230000 1.00030 NA Glyma, 130254300 0.039880 1.00530 PB1 domain Glyma, 130254300 0.0394820 1.00530 PB1 domain Glyma, 130210700 0.0394820 1.00530 NA Glyma, 1302110700 0.0394820 1.00530 NA Glyma, 13021500 0.000000 1.07530 NA Glyma, 15G019500 0.0104760 1.08400 ABC transporter transmembrane region; ABC transporter Glyma, 05G211100 0.0155320 1.09600 NA Regulator of chromosome condensation (RCC1) repeat Glyma, 05G194400 0.0206230 1.10520 Peroxidase Glyma, 17G36500 0.0229220 1.10750 Glyma, 07G369400 0.02292920 1.10750 Glyma, 07G369400 0.02292920 1.10750 Glyma, 07G369400 0.0005521 1.2460 NA Glyma, 01G167800 0.005540 1.11380 ABC transporter Glyma, 01G167800 0.0001730 1.12160 NA Glyma, 01G167800 0.002551 1.2460 AC domain Glyma, 01G167800 0.0001730 1.12160	Glyma.01G131200	0.0251250	1.05940	NA
Giyma.136254300 0.0399850 1.06590 PB1 domain Giyma.017G086300 0.0001704 1.06680 NA Giyma.1362110700 0.0344200 1.06680 NA Giyma.136211000 0.000064 1.07530 NA Giyma.13621100 0.0003667 1.07620 NA Giyma.1362211100 0.0104760 1.08910 Hpt domain Giyma.053211100 0.0155320 1.09360 NA Giyma.0532600 0.0009492 1.09630 NA Giyma.056194400 0.026230 1.10520 Peroxidase Giyma.13622600 0.000900 1.0720 Cold acclimation protein WCOR413 Giyma.116122600 0.000000 1.0720 Gid acclimation protein WCOR413 Giyma.126258000 0.0252270 1.10910 Transferase family 17: X8 domain Giyma.116122600 0.0001730 1.12150 Raffnose synthase or seed imbibition protein Giyma.116167800 0.0091542 1.12160 NA Giyma.016167800 0.0091542 1.1280 Rechomain Giyma.186	Glyma.18G005900	0.0200800	1.06570	NA
	Glyma.13G254300	0.0399850	1.06590	PB1 domain
Glyma.07G086300 0.0001704 1.06880 NA Glyma.19G021500 0.000000 1.07520 NA Glyma.19G021500 0.000000 1.08400 ABC transporter transmembrane region; ABC transporter Glyma.15G019500 0.0104760 1.08910 Hpt domain Glyma.05G211100 0.0155320 1.09620 Pathogenesis-related protein Bet v I family Glyma.05G194400 0.026230 1.10520 Peroxidase Glyma.05G194400 0.022630 1.10520 Peroxidase Glyma.11G122600 0.000000 1.10720 Cold acclimation protein WCOR413 Glyma.08G179600 0.0229220 1.10720 Glyda cold acclimation protein WCOR413 Glyma.11G122600 0.000000 1.10720 Glyda cold acclimation protein WCOR413 Glyma.14G010500 0.000181 1.12150 Rafinose synthase or seed imbibition protein Sip1 Glyma.14G016500 0.0001730 1.12300 ACT domain Glyma.01G167800 0.091542 1.1240 NA Glyma.14G010000 0.000002 1.12450 ACT domain Glyma.14G010000<	Glyma.01G210700	0.0344200	1.06660	NA
Glyma. 1951 14301 0.0000004 1.0730 NA Glyma. 196021500 0.0000000 1.08400 NA Transporter transmembrane region; ABC Glyma. 156019500 0.0104760 1.09620 NA Transporter Glyma. 056121100 0.0155320 1.09620 Pathogenesis-related protein Bet v I family Glyma.0561294400 0.0261150 1.09620 Pathogenesis-related protein Bet v I family Glyma.056194400 0.020230 1.10520 Peroxidase Glyma.11G122600 0.0000000 1.10720 Cold acclimation protein WCOR413 Glyma.146010500 0.00055400 1.10720 Cold acclimation protein WCOR413 Glyma.146010500 0.0001881 1.12150 Raffinose synthase or seed imbibition protein Sip1 Glyma.146010500 0.0001730 1.12300 ACT domain Glyma.146010000 0.000002 1.13820 Protease inhibitor/seed storage/LTP family Glyma.146010000 0.000002 1.13830 Glyma.147 Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor CCBF_B/RY-FA) subuni B	Glyma.07G086300	0.0001704	1.06880	NA
Glyma.19G021500 0.0000000 1.08400 ABC transporter transmembrane region; ABC transporter Glyma.15G019500 0.0104760 1.08910 Hpt domain Glyma.05G211100 0.0155320 1.09360 NA Glyma.05G194400 0.0026150 1.09690 NA Glyma.08G179600 0.0206230 1.09690 Regulator of chromosome condensation (RCC1) repeat Glyma.01G12500 0.0000000 1.10720 Cold acclimation protein WCOR413 Glyma.01G12500 0.00025270 1.10910 Transforter Glyma.10G126900 0.0206230 1.10750 Glycosyl hydrolases family 17; X8 domain Glyma.10G129200 0.005490 1.11380 ABC transporter Glyma.10G167800 0.0091542 1.12150 Rafinose synthase or seed imbibition protein Sip1 Glyma.1664200 0.0227410 1.12320 ACT domain Glyma.146010000 0.000002 1.1380 CCAAT-binding transcription factor Glyma.146010000 0.000002 1.1380 CCAAT-binding transcription factor Glyma.146010000 0.0000002 1.13450 CCAAT-binding tran	Glyma 08G245300	0.0000004	1.07530	NA
Glyma.15G019500 0.0104760 1.08910 Hpt domain Glyma.05G211100 0.0155320 1.09360 NA Glyma.05G230400 0.0372840 1.09620 Pathogenesis-related protein Bet v I family Glyma.05G194400 0.0261150 1.09690 Regulator of chromosome condensation (RCC1) repeat repeat repeat Glyma.05G19400 0.022920 Glyma.16(2269400 0.022920 1.0750 Glycosy hydrolases family 17; X8 domain Glyma.14G209200 0.0005490 1.11800 ABC transporter Glyma.14G010500 0.0001881 1.12150 Raffinose synthase or seed imbibition protein Sip1 Glyma.01G167800 0.0091542 1.1260 AA Glyma.16168100 0.0000055 1.12406 Ammonium Transporter Family Glyma.16168100 0.0000002 1.12500 ACT domain Glyma.146010000 0.000002 1.12800 Protease inhibitor/seed storage/LTP family Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor Glyma.146010000 0.000002 1.13450 <t< td=""><td>Glyma.19G021500</td><td>0.0000000</td><td>1.08400</td><td>ABC transporter transmembrane region; ABC</td></t<>	Glyma.19G021500	0.0000000	1.08400	ABC transporter transmembrane region; ABC
				transporter
	Glyma.15G019500	0.0104760	1.08910	Hpt domain
	Glyma.05G211100	0.0155320	1.09360	NA Dethe server is related anothin Detail I family
Glyma.05G194400 0.0261150 1.09690 Regulator of chromosome condensation (RCC1) repeat Glyma.08G179600 0.0206230 1.10520 Peroxidase Glyma.011G122600 0.000000 1.10720 Cold acclimation protein WCOR413 Glyma.011G122600 0.00229920 1.10750 Glycosyl hydrolases family 17; X8 domain Glyma.101G209200 0.0055490 1.11380 ABC transporter Glyma.101G167800 0.0091542 1.12150 Rafinose synthase or seed imbibition protein Sip1 Sip1 NA Glyma.01G167800 0.001730 1.12300 Glyma.10164200 0.00187910 1.12780 NA Glyma.0164200 0.0187910 1.12850 Glyma.146010000 0.0000002 1.13400 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B Glyma.146010000 0.000002 1.13450 CCAAT-binding transcription factor Glyma.146010000 0.000002 1.13430 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B Glyma.146010000 0.000002 1.13430 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B <	Glyma.08G230400 Glyma.17G032600	0.0372840	1.09620	NA
Glyma.08G179600 0.0206230 1.10520 Peroxidase Glyma.11G122600 0.0000000 1.10720 Cold acclimation protein WCOR413 Glyma.107G269400 0.0252270 1.10750 Glycosyl hydrolases family 17; X8 domain Glyma.102209200 0.005430 1.11380 ABC transporter Glyma.102209200 0.0005490 1.11380 ABC transporter Glyma.10200200 0.0001881 1.12150 Raffinose synthase or seed imbibition protein Sip1 Glyma.10G46200 0.0011730 1.12300 ACT domain Glyma.01G167800 0.0201422 1.12800 RAMmonium Transporter Family Glyma.01G168100 0.0020420 1.12800 Brotease inhibitor/seed storage/LTP family Glyma.14G010000 0.020002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.0125240	Glyma.05G194400	0.0261150	1.09690	Regulator of chromosome condensation (RCC1)
	-			repeat
	Glyma.08G179600	0.0206230	1.10520	Peroxidase
Glyma.16269400 0.0252270 1.10910 Glycosyl nydrolases family Glyma.162258000 0.0252270 1.10910 Transferase family Glyma.14G010500 0.0001881 1.12150 Raffinose synthase or seed imbibition protein Sip1 Glyma.11G046200 0.0011730 1.12300 ACT domain Glyma.01G167800 0.0091542 1.12400 Ammonium Transporter Family Glyma.01G0168100 0.0000055 1.12450 Ammonium Transporter Family Glyma.01G065500 0.0220420 1.12850 Glycosyl hydrolases family 16 Glyma.14G010000 0.000002 1.13280 Protease inhibitor/seed storage/LTP family Glyma.14G010000 0.000002 1.13450 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B Glyma.14G010000 0.000002 1.13450 Glyma.14G010000 0.000002 1.13830 Myo-inositol-1-phosphate synthase Glyma.03G009600 0.0070814 1.4160 20G-Fe(II) oxygenase superfamily Glyma.03G005400 0.004807 1.4590 Calcineurin-like phosphoesterase Glyma.03G005400 0.001781 1.14900 Glycosyl hydrolase family 1 Gl	Glyma.11G122600	0.0000000	1.10720	Cold acclimation protein WCOR413
Glyma.10G209200 0.0065490 1.11380 ABC transporter Glyma.14G010500 0.000181 1.12150 Raffinose synthase or seed imbibition protein Sip1 Sip1 Sip1 Sip1 Glyma.11G046200 0.0011730 1.12300 ACT domain Glyma.01G167800 0.0091542 1.12400 Ammonium Transporter Family Glyma.01G168100 0.0000055 1.12400 Ammonium Transporter Family Glyma.01G161000 0.00257410 1.1280 Rotease inhibitor/seed storage/LTP family Glyma.14G010000 0.000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor Glyma.18G018600 0.0125240 1.1380 Myo-inositol-1-phosphate synthase Glyma.03G009600 0.0070814 1.4160 20G-Fe(II) oxygenase superfamily Glyma.03G005400 0.00489120 1.14790 Calcineurin-like phosphoesterase Glyma.03G25500 0.0217841	Glyma.0/G269400 Glyma 18G258000	0.0229920	1.10/50	Glycosyl hydrolases family 1/; X8 domain Transferase family
	Glyma.10G209200	0.0065490	1.11380	ABC transporter
	Glyma.14G010500	0.0001881	1.12150	Raffinose synthase or seed imbibition protein
				Sip1
	Glyma.01G167800	0.0091542	1.12160	NA
	Glyma 10G168100	0.0011730	1.12300	ACT domain Ammonium Transporter Family
	Glyma.07G005300	0.0187910	1.12780	NA
	Glyma.03G065500	0.0220420	1.12850	Glycosyl hydrolases family 16
Glyma.14G010000 0.0000002 1.13450 CCCAAT-binding transcription factor (CBF-B/NF-YA) subunit BGlyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit BGlyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit BGlyma.18G018600 0.0125240 1.13830 Myo-inositol-1-phosphate synthase; Myo-inositol-1-phosphate synthaseGlyma.03G009600 0.0070814 1.14160 $2OG-Fe(II)$ oxygenase superfamily Glyma.07G047500Glyma.03G009600 0.0070814 1.14470 Transferase family Calcineurin-like phosphoesteraseGlyma.05G147100 0.0048972 1.14470 Transferase familyGlyma.05G147100 0.0007781 1.14990 Cytochrome P450Glyma.03G005400 0.0009655 1.15140 MatEGlyma.03G225300 0.0219480 1.15530 NAGlyma.05G180500 0.0176660 1.16280 NAGlyma.06G295800 0.0341890 1.6500 NAGlyma.16G145300 0.010730 1.16780 Serine carboxypeptidaseGlyma.13G319400 0.0037874 1.16780 Serine carboxypeptidaseGlyma.13G319400 0.002351 1.17220 NAGlyma.13G19400 0.002325320 1.17280 NAGlyma.04G046600 0.0423820 1.17280 NAGlyma.04G046600 0.002322 1.17280 NAGlyma.04G046600 0.000797 1.17400 NAGlyma.02G22000 0.000002 1.1	Glyma.18G294200	0.0257410	1.13280	Protease inhibitor/seed storage/LTP family
Glyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor Glyma.14G010000 0.0000002 1.13450 CCAAT-binding transcription factor Glyma.18G018600 0.0125240 1.13830 Myo-inositol-1-phosphate synthase; Myo-inositol-1-phosphate synthase Myo-inositol-1-phosphate synthase Glyma.03G009600 0.0070814 1.14160 2OG-Fe(II) oxygenase superfamily Glyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.05G147100 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.03G005400 0.0009655 1.15140 MatE Glyma.03G005400 0.009655 1.15140 MatE Glyma.03G005400 0.0017781 1.14590 Cytochrome P450 Glyma.03G005400 0.0037874 1.15310 Glycosyl hydrolase family 1 Glyma.03G180500 0.0176660 1.16280 NA Glyma.03G180500 0.0176660 1.16790 Raffinose synthase or seed imbibition protein Glyma.11G016600 0.0254580 1.17220 NA Serine carboxypeptidase Glyma.11G016600 0.0254580 1.17220 NA	Glyma.14G010000	0.0000002	1.13450	CCAAI-binding transcription factor
Glyma.14G010000 0.0000002 1.13450 CCBF-B/NF-YA) subunit B Glyma.18G018600 0.0125240 1.13830 Myo-inositol-1-phosphate synthase; Myo-inositol-1-phosphate synthase; Myo-inositol-1-phosphate synthase Glyma.03G009600 0.0070814 1.14160 2OG-Fe(II) oxygenase superfamily Glyma.03G009600 0.000489120 1.14170 Transferase family Glyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.03G11700 0.0004807 1.14590 Cytochrome P450 Glyma.03G005400 0.0009655 1.15140 MatE Glyma.03G005400 0.0037874 1.15310 Glycosyl hydrolase family 1 Glyma.03G0258700 0.037874 1.15300 NA Glyma.03G180500 0.0176660 1.16280 NA Glyma.02G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.11G016600 0.0254580 1.17220 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.04G046600 0.000797 1.17400 NA Glyma.02G250200<	Glyma.14G010000	0.0000002	1,13450	(CBF-B/NF-YA) subunit B CCAAT-binding transcription factor
				(CBF-B/NF-YA) subunit B
	Glyma.14G010000	0.0000002	1.13450	CCAAT-binding transcription factor
Glyma.18G018600 0.0125240 1.13830 Myo-inositol-1-phosphate synthase; Myo-inositol-1-phosphate synthase Glyma.03G009600 0.0070814 1.14160 2OG-Fe(II) oxygenase superfamily Glyma.08G311900 0.0489120 1.14170 Transferase family Glyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.07G047500 0.0061442 1.1460 AP2 domain Glyma.05G147100 0.0017781 1.14900 Cytochrome P450 Glyma.07G258700 0.0037874 1.15310 Glycosyl hydrolase family 1 Glyma.03G225300 0.0219480 1.15530 NA Glyma.03G03000 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domain Glyma.04G295800 0.0341890 1.16500 NA Glyma.02G303300 0.0037874 1.16790 Raffnose synthase or seed imbibition protein Sip1 Glyma.116016600 0.0254580 1.17220 NA Glyma.116016600 0.0423820 1.17280 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.04G046600 0.0423820 1.17260 NA	Clama 19C019600	0.0125240	1 12020	(CBF-B/NF-YA) subunit B
Glyma.03G009600 0.0070814 1.14160 2OG-Fe(II) oxygenase superfamily Glyma.08G311900 0.0489120 1.14170 Transferase family Glyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.05G147100 0.0061442 1.14900 AP2 domain Glyma.03G005400 0.0009655 1.15140 MatE Glyma.03G005400 0.0009655 1.15140 MatE Glyma.03G225300 0.0219480 1.15310 Glycosyl hydrolase family 1 Glyma.03G0300 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domain Glyma.05G180500 0.0176660 1.16280 NA Glyma.06G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.11G016600 0.0254580 1.17220 NA Glyma.11G016600 0.0254580 1.17230 NA Glyma.13G319400 0.0013351 1.17280 NA Glyma.15G190500 0.0000797 1.17400 NA Glyma.15G190500 0.0000797 1.17260 Protein phosphatase 2C	Glyma.18G018600	0.0125240	1.15850	Myo-mositol-1-phosphate synthase;
Glyma.03G009600 0.0070814 1.14160 $2OG-Fe(II)$ oxygenase superfamilyGlyma.08G311900 0.0489120 1.14170 Transferase familyGlyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesteraseGlyma.12G117000 0.0061442 1.14960 AP2 domainGlyma.03G005400 0.0009655 1.15140 MatEGlyma.03G0258700 0.0037874 1.15310 Glycosyl hydrolase family 1Glyma.03G225300 0.0219480 1.15530 NAGlyma.03G225300 0.0219480 1.15530 NAGlyma.03G30300 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domainGlyma.05G180500 0.0176660 1.16280 NAGlyma.06G295800 0.0341890 1.16500 NAGlyma.16G145300 0.0110730 1.16780 Serine carboxypeptidaseGlyma.11G016600 0.0254580 1.17220 NAGlyma.13G319400 0.0013351 1.17230 NAGlyma.16G046600 0.0423820 1.17280 NAGlyma.04G046600 0.000797 1.17400 NAGlyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C		0.0050014	1 1 1 1 6	Myo-mositor-i-phosphate synthase
Glyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.07G047500 0.0004807 1.14590 Calcineurin-like phosphoesterase Glyma.05G147100 0.0017781 1.14900 AP2 domain Glyma.03G005400 0.0009655 1.15140 MatE Glyma.03G0258700 0.0037874 1.15310 Glycosyl hydrolase family 1 Glyma.03G225300 0.0219480 1.15530 NA Glyma.03G0300 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domain Glyma.06G295800 0.0341890 1.16500 NA Glyma.06G295800 0.0341890 1.16780 Serine carboxypeptidase Glyma.06G295800 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Sip1 Sip1 Sip1 Sip1 Glyma.11G016600 0.0254580 1.17220 NA Glyma.13G319400 0.0013351 1.17280 NA Glyma.15G190500. 0.0000797 1.17400 NA Glyma.02G250200 0.000002 1.17560 Protein phosphatase 2C	Glyma.03G009600 Glyma.08G311900	0.0070814	1.14160	20G-Fe(II) oxygenase superfamily
Glýma.12G117000 0.0061442 1.14960 AP2 domainGlyma.05G147100 0.0017781 1.14990 Cytochrome P450Glyma.03G005400 0.0009655 1.15140 MatEGlyma.07G258700 0.0037874 1.15310 Glycosyl hydrolase family 1Glyma.03G225300 0.0219480 1.15530 NAGlyma.13G030300 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domainGlyma.05G180500 0.0176660 1.16280 NAGlyma.06G295800 0.0341890 1.16500 NAGlyma.16G145300 0.0110730 1.16780 Serine carboxypeptidaseGlyma.16G145300 0.0037874 1.16790 Raffinose synthase or seed imbibition proteinSip1 $Sip1$ $Sip1$ Serine carboxypeptidaseGlyma.11G016600 0.0254580 1.17220 NAGlyma.04G046600 0.0423820 1.17280 NAGlyma.04G046600 0.0000797 1.17400 NAGlyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.07G047500	0.0004807	1.14590	Calcineurin-like phosphoesterase
Glyma.05G14/100 0.0017/81 1.14990 Cytochrome P450 Glyma.03G005400 0.0009655 1.15140 MatE Glyma.07G258700 0.0037874 1.15310 Glycosyl hydrolase family 1 Glyma.03G225300 0.0219480 1.15530 NA Glyma.03G025400 0.0219480 1.15530 NA Glyma.03G225300 0.0219480 1.15530 NA Glyma.03G25500 0.0176660 1.16280 NA Glyma.05G180500 0.0176660 1.16280 NA Glyma.06G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.16G145300 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Sip1 Serine carboxypeptidase Sip1 Glyma.11G016600 0.0254580 1.17220 NA Glyma.13G319400 0.0013351 1.17230 NA Glyma.15G190500 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.12G117000	0.0061442	1.14960	AP2 domain
Glyma.03G005400 0.0009655 1.15140 MatE Glyma.07G258700 0.0037874 1.15310 Glycosyl hydrolase family 1 Glyma.03G225300 0.0219480 1.15530 NA Glyma.13G030300 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domain Glyma.05G180500 0.0176660 1.16280 NA Glyma.06G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.02G303300 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Sip1 Sip1 Serine carboxypeptidase Sip1 Glyma.13G319400 0.0013351 1.17220 NA Sip1 Glyma.04G046600 0.0423820 1.17280 NA Serine carboxypeptidase Glyma.15G190500 0.0000797 1.17400 NA Serine carboxypeptidase Serine carboxypeptidase Glyma.04G046600 0.0423820 1.17280 NA Serine carboxypeptidase Serine carboxypeptidase Glyma.02G250200 0.000002 1.1780 NA Serine carboxypeptidase Serine carboxypept	Glyma.05G14/100	0.0017781	1.14990	Cytochrome P450
Glyma.03G225300 0.0037874 1.15510 Glycosyn hydrolase rainity i Glyma.03G225300 0.0219480 1.15530 NA Glyma.05G180500 0.0176660 1.16280 NA Glyma.06G295800 0.0341890 1.16500 NA Glyma.06G295800 0.0341890 1.16700 Raffinose synthase or seed imbibition protein sip1 Glyma.16G145300 0.0110730 1.16790 Raffinose synthase or seed imbibition protein sip1 Glyma.11G016600 0.0254580 1.17220 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.15G190500. 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.03G005400	0.0009655	1.15140	MatE
Glyma.13G030300 0.0456050 1.15940 Lipoxygenase; PLAT/LH2 domain Glyma.05G180500 0.0176660 1.16280 NA Glyma.06G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.02G303300 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Serine carboxypeptidase Sip1 Glyma.13G319400 0.0013351 1.17220 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.15G190500 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.03G225300	0.0037874	1.15530	NA
Glyma.05G180500 0.0176660 1.16280 NA Glyma.06G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.02G303300 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Sip1 Serine carboxypeptidase Sip1 Glyma.11G016600 0.0254580 1.17220 NA Glyma.13G319400 0.0013351 1.17230 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.15G190500 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.13G030300	0.0456050	1.15940	Lipoxygenase; PLAT/LH2 domain
Glyma.06G295800 0.0341890 1.16500 NA Glyma.16G145300 0.0110730 1.16780 Serine carboxypeptidase Glyma.02G303300 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Sip1 Glyma.13G319400 0.0013351 1.17220 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.15G190500. 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.05G180500	0.0176660	1.16280	NA
Glyma.10G145300 0.0110730 1.16780 Serme carboxypeptidase Glyma.02G303300 0.0037874 1.16790 Raffinose synthase or seed imbibition protein Sip1 Sip1 Sip1 Sip1 Glyma.13G319400 0.0013351 1.17220 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.15G190500. 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.06G295800	0.0341890	1.16500	NA Social control to the second
Glyma.11G016600 0.0254580 1.17220 NA Glyma.13G319400 0.0013351 1.17230 NA Glyma.04G046600 0.0423820 1.17280 NA Glyma.15G190500. 0.0000797 1.17400 NA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.10G145300 Glyma.02G303300	0.00110730	1.16780	Raffinose synthase or seed imbibition protein
Glyma.11G0166000.02545801.17220NÅGlyma.13G3194000.00133511.17230NAGlyma.04G0466000.04238201.17280NAGlyma.15G1905000.00007971.17400NAGlyma.02G2502000.00000021.17560Protein phosphatase 2C			1.10,90	Sip1
Glyma.13G3194000.00133511.17230NAGlyma.04G0466000.04238201.17280NAGlyma.15G190500.0.00007971.17400NAGlyma.02G2502000.00000021.17560Protein phosphatase 2C	Glyma.11G016600	0.0254580	1.17220	NÀ
Glyma.04G0466000.04238201.17280NAGlyma.15G190500.0.00007971.17400NAGlyma.02G2502000.00000021.17560Protein phosphatase 2C	Glyma.13G319400	0.0013351	1.17230	NA
Glyma.02G250200 0.0000002 1.17400 INA Glyma.02G250200 0.0000002 1.17560 Protein phosphatase 2C	Glyma.04G046600	0.0423820	1.17280	NA
	Glyma.02G250200	0.0000002	1.17560	Protein phosphatase 2C

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G190200	0.0000261	1.17680	Homeobox associated leucine zipper; Homeobox
Glyma.06G312300	0.0018056	1.18070	domain ABC transporter transmembrane region; ABC transporter
Glyma.09G144800	0.0339820	1.18160	2OG-Fe(II) oxygenase superfamily
Glyma.06G050300	0.0033001	1.18000	Zinc finger C-x8-C-x5-C-x3-H type (and similar)
Glyma.10G151100 Glyma.07G240900	0.0045774 0.0030363	1.18830	NA NA
Glyma.18G005800	0.0000612	1.19830	Protease inhibitor/seed storage/LTP family
Glyma.12G041200	0.0199510	1.19940	NA Muh lika DNA hinding domain
Glyma.10G261000	0.0000204	1.20470	NA
Glyma.18G184551	0.0280950	1.20610	NA
Glyma.07G213700 Glyma.05G152500	0.0022736	1.21560	Thioredoxin Transferase family
Glyma.03G026000	0.0325090	1.21930	Protein of unknown function (DUF1442)
Glyma.11G222600 Glyma.17G076100	0.0023993	1.22370	Glycosyl hydrolases family 18
Glyma.18G055900	0.0388240	1.23660	Nodulin-like
Glyma.03G044500	0.0311470	1.24000	Dirigent-like protein Helix loop helix DNA hinding domain
Glyma.12G217400	0.0201130	1.24040	BURP domain
Glyma.08G293366	0.0420770	1.24410	NA
Glyma.13G340000 Glyma.06G295900	0.0000835	1.24480	Aldehyde dehydrogenase family NA
Glyma.18G184600	0.0176660	1.24850	NA
Glyma.08G177000	0.0003021	1.24890	Zinc finger, C3HC4 type (RING finger)
Glyma.01G129400	0.0008180	1.25340	Aminotransferase class-III
Glyma.04G039300 Glyma.04G039300	0.0131400	1.25740	bZIP transcription factor bZIP transcription factor
Glyma.04G039300	0.0131400	1.25740	bZIP transcription factor
Glyma 08G141000	0.0131400	1.23740	Histone-like transcription factor (CBE/NE-Y) and
Glyma.000141000	0.0017701	1.23750	archaeal histone
Glyma.08G141000	0.0017781	1.25750	Histone-like transcription factor (CBF/NF-Y) and
Glyma.08G083700	0.0212320	1.26380	Protein of unknown function (DUF561)
Glyma.05G215700 Glyma 10G224300	0.0000777 0.0335380	1.26590	NA ABA/WDS induced protein
Glyma.03G181000	0.0361670	1.26820	NA
Glyma.12G135300	0.0113530	1.27150	ABC transporter; ABC transporter
Glvma.01G241500	0.0055220	1.27230	transmembrane region NA
Glyma.07G067500	0.0000588	1.28040	Protein of unknown function, DUF604
Glyma 15G243400	0.00000090	1.28270	Ar 2 domain Cytochrome P450
Glyma.11G123000	0.0000998	1.30200	WD domain, G-beta repeat
Glyma.08G215500 Glyma.02G156800	0.0131400 0.0176740	1.30290	Helix-loop-helix DNA-binding domain
Glyma.06G084200	0.0110650	1.31140	Phosphorylase superfamily
Glyma.01G142400	0.0095651	1.31590	Glycosyl hydrolases family 18
Glyma.11G082300	0.0000001	1.32390	NA
Glyma.12G234700 Glyma.11G075500	0.0058245 0.0000000	1.33300 1.33310	Trypsin and protease inhibitor linker histone H1 and H5 family
Glyma.08G071300	0.0110620	1.33610	Phosphate-induced protein 1 conserved region
Glyma.08G195100 Glyma 04G054400	0.0068273 0.0225600	1.33770	Putative Phosphatase Hsp20/alpha crystallin family
Glyma.02G234800	0.0122860	1.34150	NAF domain; Protein kinase domain
Glyma.06G216600	0.0234600	1.34210	Zn-inger in Kan binding protein and others
Glyma.13G275600	0.0012368	1.35170	NA

Cuadro A.2: Continuación

-

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.05G157400 Glyma.12G109800 Glyma.19G185600	$\begin{array}{c} 0.0202270\\ 0.0007314\\ 0.0141220 \end{array}$	$ \begin{array}{r} 1.35260 \\ 1.35320 \\ 1.35440 \end{array} $	AP2 domain O-methyltransferase; Dimerisation domain Major Facilitator Superfamily; Nodulin-like
Glyma.08G277200 Glyma.03G003400 Glyma.03G003400 Glyma.08G083800 Glyma.14G212700	$\begin{array}{c} 0.0084709\\ 0.0002605\\ 0.0002605\\ 0.0000621\\ 0.0132500 \end{array}$	1.35510 1.35650 1.35650 1.35650 1.35650 1.35980	NA bZIP transcription factor bZIP transcription factor NA Eukaryotic aspartyl protease
Glyma.19G219100 Glyma.13G247200 Glyma.07G066900 Glyma.05G180600	$\begin{array}{c} 0.0004973\\ 0.0000001\\ 0.0208660\\ 0.0094811 \end{array}$	$\begin{array}{c} 1.36040 \\ 1.36060 \\ 1.36280 \\ 1.36350 \end{array}$	Glycosyl transferase family 8 Myb-like DNA-binding domain Glycosyl hydrolases family 28 Myo-inositol-1-phosphate synthase;
Glyma.01G060300	0.0202070	1.36620	Protein of unknown function (DUF1675)
Glyma.12G092900 Glyma.16G094100 Glyma.18G027100	0.0100570 0.0072100 0.0000000	1.37370 1.37880 1.38240	short chain dehydrogenase Glycosyltransferase family 92 Phosphatidylinositol-specific phospholipase C, Y domain; C2 domain; Phosphatidylinositol-specific phospholipase C, X domain; Phosphoinositide-specific phospholipase
Glyma.12G099700 Glyma.02G179100	0.0010995 0.0079395	$1.39050 \\ 1.39070$	Miro-like protein
Glyma.08G246400 Glyma.02G180200 Glyma.03G040600 Glyma.17G252900 Glyma.11G145100	0.0003514 0.0160910 0.0002099 0.0074797 0.0000305	$\begin{array}{c} 1.39130 \\ 1.39280 \\ 1.39610 \\ 1.40490 \\ 1.40540 \end{array}$	Protein of unknown function (DUF1675) NA NA NA Possible lysine decarboxylase
Glyma.03G126600 Glyma.10G208400 Glyma.03G221700 Glyma.04G021000 Glyma.04G056500	$\begin{array}{c} 0.0000020\\ 0.0002584\\ 0.0138280\\ 0.0010633\\ 0.0031829 \end{array}$	$\begin{array}{c} 1.40850 \\ 1.41300 \\ 1.42240 \\ 1.42570 \\ 1.42660 \end{array}$	F-box domain DVL family Myb-like DNA-binding domain NA NA
Glyma.09G155500 Glyma.06G062100 Glyma.02G243600 Glyma.20G057500 Glyma.02G104600	$\begin{array}{c} 0.0072039\\ 0.0039289\\ 0.0000000\\ 0.0037874\\ 0.0001722 \end{array}$	1.43180 1.43230 1.43250 1.45130 1.45870	Trypsin and protease inhibitor Protein kinase domain; NAF domain NA UDP-glucoronosyl and UDP-glucosyl transferase UDP-glucoronosyl and UDP-glucosyl transferase
Glyma.06G043000 Glyma.08G033800 Glyma.01G021000	0.0011372 0.0000284 0.0041737	1.46230 1.46370 1.46870	Asparaginase Protein phosphatase 2C Alcohol dehydrogenase GroES-like domain; Zinc-binding dehydrogenase
Glyma.08G195000 Glyma.01G152900	0.0093614 0.0088234	1.47340 1.48010	NA
Glyma.16G037900 Glyma.01G068200 Glyma.01G132500 Glyma.12G149100 Glyma.13G283100	$\begin{array}{c} 0.0000000\\ 0.0060915\\ 0.0001433\\ 0.0023918\\ 0.0008404 \end{array}$	$\begin{array}{c} 1.48900 \\ 1.49360 \\ 1.49510 \\ 1.50610 \\ 1.50830 \end{array}$	NA 2OG-Fe(II) oxygenase superfamily Protease inhibitor/seed storage/LTP family No apical meristem (NAM) protein S-locus glycoprotein family; Protein tyrosine kinase; D-mannose binding lectin; PAN-like domain
Glyma.06G183800 Glyma.06G304900 Glyma.15G116800 Glyma.08G176300 Glyma.15G055400	$\begin{array}{c} 0.0062500\\ 0.0000077\\ 0.0004415\\ 0.0002718\\ 0.0023799 \end{array}$	$\begin{array}{c} 1.52040 \\ 1.52670 \\ 1.53160 \\ 1.53180 \\ 1.53920 \end{array}$	U-box domain Miro-like protein Domain of unknown function (DUF588) NA Cleavage site for pathogenic type III effector avirulence factor Avr
Glyma.18G106300 Glyma.12G205700 Glyma.13G175700 Glyma.05G183200	$\begin{array}{c} 0.0017630\\ 0.0049769\\ 0.0027125\\ 0.0017154 \end{array}$	$\begin{array}{c} 1.53990 \\ 1.54100 \\ 1.54760 \\ 1.55250 \end{array}$	NA NA Hsp20/alpha crystallin family Histone-like transcription factor (CBF/NF-Y) and archaeal histone
Glyma.18G071700	0.0001301	1.57290	Transmembrane amino acid transporter protein (Conntinua en la pág. siguiente)

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions
Glyma.06G295700	0.0000526	1.57960	O-methyltransferase; Dimerisation domain
Glyma.06G088100	0.0000228	1.59080	Hpt domain Zing finger C3HC4 type (PING finger)
Glyma.01G204800	0.0003831	1.59540	Pollen allergen; Rare lipoprotein A (RlpA)-like
			double-psi beta-barrel
Glyma.17G116500	0.0000000	1.59560	Asp/Glu/Hydantoin racemase
Glyma.08G104100 Glyma.14G038300	0.0000271	1.59790	Cytochrome P450
Glyma.13G090700	0.00000777	1.60720	NA
Glyma.16G202500	0.0014897	1.60920	Embryo-specific protein 3, (ATS3)
Glyma 02G101200	0.0003482	1.61860	WAY2 C terminal domain: Eatty agid
Giyilia.05G101200	0.0050505	1.01800	wAA2 C-terminal domain, Fatty actu hydroxylase superfamily
Glyma.07G050300	0.0000060	1.62140	S1/P1 Nuclease
Glyma.02G291700	0.0029853	1.62320	Senescence regulator Protein of unknown function, DUE538
Glyma.03G185100	0.0000222	1.64230	Nodulin-like
Glyma.02G088900	0.0000612	1.64330	Protease inhibitor/seed storage/LTP family
Glyma.02G118500	0.0000057	1.64650	Protein of unknown function (DUF1675)
Glyma.04G255400 Glyma.19G181700	0.0000017	1.66100	NA
Glyma.05G112900	0.0000034	1.66520	NA
Glyma.03G213300	0.0018922	1.67090	TspO/MBR family
Glyma.04G061500 Glyma.06G154400	0.0000034	1.68300	NAF domain; Protein kinase domain
Glyma.06G154400	0.0000369	1.69220	No apical meristem (NAM) protein
Glyma.02G279700	0.0001133	1.71250	NA
Glyma.17G090900	0.0000017	1.72240	Armadillo/beta-catenin-like repeat; U-box
Glyma 17G222000	0.0005813	1 72910	domain FamA-like transporter family
Glyma.11G236500	0.0008569	1.73530	NA
Glyma.18G035000	0.0000000	1.74040	Protein phosphatase 2C
Giyilia.100088000	0.0000001	1.75190	Pentidase family M28
Glyma 13G347800	0.0002244	1 76880	
Glyma.20G167500	0.00002244	1.77770	ABA/WDS induced protein
Glyma.15G250100	0.0000009	1.79820	NA Unabaracterized protein family (UPE0014)
Glyma.07G139400	0.0000230	1.81460	Protease inhibitor/seed storage/LTP family
Glyma.13G088700	0.0000341	1.84610	Annexin
Glyma.06G170300	0.0000257	1.87590	Calcineurin-like phosphoesterase
Glyma.08G103500 Glyma.01G131500	0.0000000	1.88080	NA Protein kinase domain: NAF domain
Glyma.05G045800	0.0001420	1.91660	NA
Glyma.02G213700	0.0000001	1.91960	Carbonic anhydrase
Glyma. 19G0/6500 Glyma 08G009900	0.0000025	1.92960	NA Sugar efflux transporter for intercellular exchange
Glyma.18G129800	0.0000701	1.96680	NA
Glyma.16G043400	0.0000098	1.97180	GDA1/CD39 (nucleoside phosphatase) family
Glyma.03G220751 Glyma.19G069200	0.0000619	1.98730	NA Protain phosphatase 2C
Glyma.04G050300	0.0000000	2.00160	Zinc finger C-x8-C-x5-C-x3-H type (and similar)
Glyma.05G074200	0.0000000	2.04060	NA Blacteouanin like domain
Glyma 11C025100	0.0000000	2.00440	AD2 domain
Glyma.14G162100	0.0000024	2.11500 2.19930	Protein phosphatase 2C
Glyma.06G097300	0.0000000	2.21260	Transferase family
Glyma.0/G090400 Glyma.17G127900	0.0000000	2.26160 2.33340	NA NA
Glyma 08G068800	0.0000010	2 34580	Hsp20/alpha crystallin family
Glyma.07G015600	0.0000000	2.37460	Hpt domain
Glyma.10G047100 Glyma 20G191751	0.0000001	2.39740	Uncharacterised protein family (UPF0014)
Glyma.08G199300	0.0000001	2.45560	Family of unknown function (DUF706)
			(Conntinua en la não aiquiente)

Cuadro A.2: Continuación

Gene_ID	padj	log2FoldChange	PFAM_Descriptions	
Glyma.08G120200	0.0000001	2.47700	Major intrinsic protein	
Glyma.10G198200 Glyma.10G157800	0.0000000	2.57170 2.60620	NA Hpt domain	

Cuadro A.2: Continuación

Anexo B

Información complementaria de SNPs

Cuadro B.1: Tabla con información de secuencias para el pedido de sondas KASP específicas de cada SNP. La secuencia comprende las regiones flanqueantes del SNP (50 pares de bases upstream y downstream), con el alelo de referencia y el alternativo entre corchetes, tal como debe proveerse a la empresa LGC para el diseño de las sondas KASP.

CHROM_POS	GeneName	KASP
Gm12_6914329 Gm14_805931 Gm14_808853 Gm14_808929 Gm16_772780	GLYMA.12G085900 GLYMA.14G010500 GLYMA.14G010500 GLYMA.14G010500 GLYMA.16G008900	CAGGTCTGTTCTTGTGGTTGCTCACAATGCTGTTAATCAGGCCCTTGTTG[G/C]CACAGCAATTGGTACCTGTTTCTCCCTCTAGGAAAATTTTTATGATGAAAAA ATTCGACACCATTGCTCAATTCCTTCACGGGAAACTATTGTGAAAACTTCA[7/G]ATTCTCTGGTTTTCAAGGTAACTGGGATGCATCCTTTGGAAAGGTAA TGCTCTGAAATCTCCTTCGAGAAGGGGTAAGAAAACTGCATAGGTAGCAG[C/T]ACCTTTGGTCTCCTCCTCCAATATCAGAGCCATTGTGTGTCTTCAACGA TCAGAGCCATTGTGTGGCTTCAACGAGCAAGAACTGCGTCTCAATGGGGAT[A/T]TCTTGTCCCCACAAGTGCCCATTCTCTCGCGCCATTCCACCACACTCTGAACGG AGAAAGACGTGTCTCCACAGCAATTCATTCAAATGGAATGATGAATATGCA[G/C]CACTCAATGTAACAGCAATTGGACTCCTCCACAACATGAACTGCATCCATC
Gm16_773621 Gm16_773822 Gm16_774303 Gm16_774990 Gm05_35457627	GLYMA.16G008900 GLYMA.16G008900 GLYMA.16G008900 GLYMA.16G008900 GLYMA.05G163100	CAGAATCTGCGACGTAAGATGGTTCGGTCCACGGCTAGAAAAGGAAGCAA[C/T]TTGAGGAGCAGAAGGAGAACCACCAATCACAGTTCCTCTAAACTCAATCG CTCTTCACCGTTGGCTTCAGTGTTGGCCATAATCATTCCCAACCCACGA[T/C]AAGTTTTACTGCACTACCTTTTTCAACTCTGGCATTCCCTCTATCAC GAAGGACCAGGTCCAGAGTTTCCGGCGGGGCGCAGAGAGAG
Gm18_20438691 Gm09_40118414 Gm12_11931345 Gm18_9803255 Gm14_6168634	GLYMA.18G138300 GLYMA.09G167900 GLYMA.12G117700 GLYMA.18G096400 GLYMA.14G075000	TCTCCTCCACAACTTCACAACtcacttcccttccttccttccttccttcCTCC[A/T]CCTCTCCACTTTCTCGCAAATTTCCCCCTCACTTTCCGGGAAAAAAAA
Gm14_6169775 Gm18_50372038 Gm03_5211321 Gm10_11238659 Gm14_4505988	GLYMA.14G075000 GLYMA.18G213900 GLYMA.03G040600 GLYMA.10G088000 GLYMA.14G057400	AATGTGTTTTCTGCTGGAGGCATCACACAAATCCAGTAGGGAAAAGTTGG[C/G]AGTGGCAGATGGATGATCCAATAGAGATTGTAAATTCTGCAATAGACCTG ACTCACTGAGAAATCGAGGTTTAGAAGAAGAAGAAGGGGTTCGCATTGGGGAGA[C/A]GAGAGGGGAATGAAGAGGAAGCGCTGAAATTTGACATTGGAGAAGAAAAA TGGTTACTATGGAAATGTCTTCAGAATTACACCTCCCTTGTGTTCACCA[A/C]AGAAGATGCAGGTTAGACTCGGTCTCAATGTCTATTTGTTTAACACTA CCGCCTCGTTAACCGCACGCTATGTGGTGAGAACCATTTCACCACCCCCGGG[T/C]ITCAAACAAAAAACAGTACAGTACAACATCTGGGCGCTACCTGCTTCCTACCCGGC AATAATTTTCTCTTTTGAAAGGCATATTGGTCAAGGTAAAGGTACTCGT[G/C]GCTCACAACATTCTAATCTCTCCTGCTGCTGACAATTACCAGGTTAGACCATTAAC
Gm15_9848408 Gm15_9852551 Gm15_9852719 Gm17_29090428 Gm18_5060538	GLYMA.15G124700 GLYMA.15G124700 GLYMA.15G124700 GLYMA.17G194500 GLYMA.18G057900	TGCAACCACTGAACCACCAAAGATAACAACACTTGCAACCATGCAGATTA[C/A]ATATAGTAGTACCCATGACCACCATGTTTCTATGTGAAATGGGGCTATGT TAAATTATCTTTAAGAATAGAAATCCATTAAAGAGAACTCACTTCTTGTA[C/G]AAGGTGTCCTTCCATTGCCCATGCAACAACATAGAGGAGTCCCTGGACAG TATATAGCGGAAAGAACAGAAC

Cuadro B.2: SNPs e INDELs alternativos a utilizarse como marcadores moleculares. CHROMPOS= Cromosoma y posición donde se encuentra, REF= alelo o variante de referencia presente en el GS, ALT= alelo o variante alternativa presente en el GT, EFFECT= efecto predicho por SNPeff, IMPACT= impacto predicho por SNPeff, GO Molecular Function Description= Descripción de la función moleular cumplida por el gen anotada en Gene Ontology.

CHROM_POS	REF	ALT	GeneName	EFFECT	IMPACT	GO Molecular Function Descriptions
Gm01_33703355 Gm02_3475812	C	T	GLYMA.01G099600 GLYMA.02G037500	splice_acceptor_variant&intron_variant	HIGH	NA hydrolase activity
Gm06_46302500	G	Ť	GLYMA.06G276100	stop_gained	HIGH	molecular function;
Gm08_3179000	T	C	GLYMA.08G040400	splice_acceptor_variant&intron_variant	HIGH	aminoacyl-tRNA hydrolase activity;
Giii13_40377432	ſ	C	GLI MA.15G514800	spice_acceptor_variant&intron_variant	HIGH	NA ADD his Jin at
Gm16_33551090	T	A	GLYMA.16G033900 GLYMA.16G173200	stop_gained stop_lost&splice_region_variant	HIGH	ADP binding; NA
Gm19_49139307 Gm19_51070407	A C	G T	GLYMA.19G237000 GLYMA.19G262700	stop_lost&splice_region_variant stop_gained	HIGH HIGH	branched-chain-amino-acid transaminase activity; DNA binding; DNA-binding transcription factor activity; protein binding;
Gm10.2310187	TGC	т	GLYMA 10G026400	frameshift variant	HIGH	transcription regulatory region DNA binding; GTP binding; hydrolase activity;
Gm10 2344378	C	CA	GLYMA 10G026900	frameshift variant	HIGH	copper ion hinding: ribosomal large subunit hinding:
Gm10_46297992	ĞAC	Ğ	GLYMA.10G232800	frameshift_variant	HIGH	protein serine/threonine kinase activity; MAP kinase kinase kinase
Gm11_1998212	GT	G	GLYMA.11G027800	frameshift_variant	HIGH	RNA binding;
Gm11_37120234 Gm13_37373560	CG	C	GLYMA.11G227500 GLYMA.13G278600	frameshift_variant	HIGH	molecular function; beta-glucosidase activity: coniferin beta-glucosidase activity: scopolin
01115_57575500	0	00	GET MA.15G278000	nancsint_valan	mon	beta-glucosidase activity; connerni beta-glucosidase activity; scoponii
Gm14 6524835	TCTCTA	т	GLYMA.14G078300	frameshift variant	HIGH	NA
Gm17_29106816	TG	Т	GLYMA.17G194600	frameshift_variant	HIGH	peptide:proton symporter activity; oligopeptide transmembrane transporter
						activity; low-affinity nitrate transmembrane transporter activity; peptide
Gm18_8067526	А	AAC	GLYMA.18G082800	frameshift_variant	HIGH	nucleotide binding; protein binding; ATP binding; ADP binding;
Gm20.2106419 Gm02.15774574	GC	Ģ	GLYMA.20G020400	frameshift_variant	HIGH MODER ATE	NA melocular function:
GIII02_13/74374	т	TTCC	GL1 MA.02G149200	disruptive_inframe_deletion	MODERATE	DNA kinding
Gm03_45118984	TCATTGC	T	GLYMA.03G241000	disruptive_inframe_deletion	MODERATE	carbohydrate:proton symporter activity; carbohydrate transmembrane
Gm03_45422118	CCAG	С	GLYMA.03G245200	disruptive_inframe_deletion	MODERATE	transporter activity; protein binding; sequence-specific DNA binding; transcription regulatory
Gm04 9262715	TGGA	т	GLYMA 04G101000	conservative inframe deletion	MODERATE	region DNA binding; protein kinase binding;
Gm05_36083432	TCCAAGC	Ť	GLYMA.05G170200	conservative_inframe_deletion	MODERATE	RNA binding; protein binding;
Gm06_46414222	TCGA	T	GLYMA.06G277000	disruptive_inframe_deletion	MODERATE	ubiquitin-protein transferase activity; protein binding;
Gm06_4/4//94/	GACA	G	GLY MA.06G290000	conservative_inframe_deletion	MODERATE	DNA binding; DNA-binding transcription factor activity; protein binding; sequence-specific DNA binding; transcription regulatory region DNA binding;
Gm07_35382451	А	AATGAAG	GLYMA.07G182200	disruptive_inframe_insertion	MODERATE	NA
Gm07_35382507 Gm09_39765183	C	CGAAGAT	GLYMA.07G182200 GLYMA.09G164600	disruptive_inframe_insertion	MODERATE	NA DNA hinding: protein hinding:
Gm09_37703183	GGCATCT	G	GLYMA 09G247400	disruptive inframe deletion	MODERATE	mPNA binding
Gm10_1066624	A	ATCT	GLYMA.10G011300	disruptive_inframe_insertion	MODERATE	DNA-binding transcription factor activity; protein binding; sequence-specific DNA binding; transcription regulatory region DNA
Gm10 11777895	AGCG	А	GLYMA 10G089300	conservative inframe deletion	MODERATE	binding; molecular function;
Gm11_9986609	CĂĞĂ	Ĉ	GLYMA.11G131200	conservative_inframe_deletion	MODERATE	DNA-binding transcription activator activity, RNA polymerase II-specific;
						DNA-binding transcription factor activity; protein binding; transcription
Gm11 39141955	А	ATGGGCC	GLYMA 11G251100	disruptive inframe insertion	MODERATE	regulatory region DNA binding; protein dimerization activity; lipid binding;
Gm13 25965080	Δ	AACAAAG	GLVMA 13G152576	disruptive inframe insertion	MODERATE	molecular function:
Gm13_26106528	ĉ	CTAGGGT	GLYMA.13G152882	conservative_inframe_insertion	MODERATE	DNA binding; TBP-class protein binding; protein heterodimerization
Gm14_1168726	AAGG	А	GLYMA.14G016300	disruptive_inframe_deletion	MODERATE	DNA binding; DNA-binding transcription factor activity; protein binding;
Gm14_1347647	TGAAGAA	Т	GLYMA.14G018700	disruptive_inframe_deletion	MODERATE	peptidyl-prolyl cis-trans isomerase activity; ribosome binding; protein
Gm15_5170764	т	TCCA	GLVMA 15G067900	disruptive inframe insertion	MODER ATE	folding chaperone; DNA-binding transcription factor activity: protein binding: zinc ion
0111020110101		reen	0211111100007700	distiprive annune ansertion	MODERATE	binding; sequence-specific DNA binding; transcription regulatory region DNA binding:
Gm16_1206917	Т	TGGTGAC	GLYMA.16G013900	conservative_inframe_insertion	MODERATE	L-alanine:2-oxoglutarate aminotransferase activity; copper ion binding;
Gm16_32883119	CACA	C	GLYMA 16G167500	conservative_inframe_deletion	MODERATE	ATP binding; pyridoxal phosphate binding; DNA binding: DNA-binding transcription factor activity:
Gm17_31780176	GGAT	Ğ	GLYMA.17G200600	splice_region_variant&disruptive_inframe_deletion	MODERATE	molecular function;
Gm17_31/801/6 Gm17_38919065	ATTC	G	GLYMA.17G200600 GLYMA.17G232700	disruptive_inframe_deletion	MODERATE	protein binding: ATP binding: sequence-specific DNA binding:
						transcription regulatory region DNA binding;
Gm18_5047885	Т	TGAA	GLYMA.18G057500	disruptive_inframe_insertion	MODERATE	phosphoprotein phosphatase activity; metal ion binding;
Gm18_6790256	TGCTGGC	Т	GLYMA.18G072500	conservative_inframe_deletion	MODERATE	aspartic-type endopeptidase activity; polyubiquitin modification-dependent
Gm18_6984009	AGCTGAT	А	GLYMA.18G074100	conservative_inframe_deletion	MODERATE	ATP binding; unfolded protein binding;
Gm18_23333984	CTGT	С	GLYMA.18G144100	conservative_inframe_deletion	MODERATE	inorganic phosphate transmembrane transporter activity; low-affinity
Gm18_49878246	CCAACCT	с	GLYMA.18G210100	disruptive_inframe_deletion	MODERATE	phosphate transmembrane transporter activity; symporter activity; ubiquitin binding;
Gm19 49042932	Т	TGAA	GLYMA 19G235800	disruptive inframe insertion	MODERATE	molecular function:
Gm19_50869063	Ť	TGCA	GLYMA.19G259800	disruptive_inframe_insertion	MODERATE	structural constituent of ribosome; protein kinase activator activity;
Gm19_51059227	GAACCTC	G	GLYMA 19G262500	disruptive_inframe_deletion	MODERATE	ribonucleoprotein complex binding; translation initiation factor activity; transferase activity;
Gm19_51085335	G	GCGACCA	GLYMA.19G263000	disruptive_inframe_insertion	MODERATE	dioxygenase activity;