



**UNIVERSIDAD DE BUENOS AIRES**

Facultad de Farmacia y Bioquímica

**ESTUDIO DE LA RESISTENCIA A *Sclerotinia sclerotiorum* EN LINEAS ENDOCRIADAS DE GIRASOL**

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área Biotecnología

**Lic. Jeremías Zubrzycki**

Directora de tesis: Norma Paniego

Director asistente de tesis: Gerardo Cervigni

Consejera de Estudios: Carina Rivolta

Instituto de Biotecnología, CICVyA, INTA Castelar

Buenos Aires, 2014

## **ESTUDIO DE LA RESISTENCIA A *Sclerotinia sclerotiorum* EN LINEAS ENDOCRIADAS DE GIRASOL**

### **RESUMEN**

La podredumbre húmeda del capítulo (PHC), causada por *Sclerotinia sclerotiorum*, es una de las limitante más importantes para el cultivo del girasol. Este trabajo tuvo como objetivo identificar las regiones del genoma de girasol asociadas a resistencia a PHC para sustentar el desarrollo de estrategias de mejoramiento asistido que permitan acelerar el proceso de creación de nuevos genotipos con resistencia a la enfermedad.

Para alcanzar este objetivo se trabajó con una población de 135 líneas recombinantes endocriadas (RILs) obtenidas del cruzamiento entre las líneas PAC2 (resistencia moderada a PHC) x RHA266 (susceptible a PHC), se estableció un protocolo para corroborar la pureza genética de las RILs entre ensayos, se saturó un mapa genético de girasol con marcadores funcionales de tipo mutaciones simples (SNP), se realizaron evaluaciones a campo y se llevaron adelante estudios para la identificación de las regiones responsables de la resistencia a PHC.

Para establecer el protocolo de pureza genética de genotipos de girasol se analizaron 40 marcadores de tipo SSR distribuidos en los 17 grupos de ligamiento. En una segunda etapa se seleccionaron diez marcadores capaces de discriminar de manera inequívoca a las RILs que componen la población, y que combinan tamaños de fragmentos y marca fluorescentes de tal manera que pueden resolverse claramente en una única corrida de electroforesis capilar. Este protocolo se implementó sobre todas las RILs después de cada proceso de multiplicación. Con el fin de automatizar y abaratar los costos de implementación del protocolo se desarrolló un ensayo multiplexado de PCR.

Para la saturación del mapa genético PAC2 x RHA266 con marcadores de tipo SNP se implementaron diferentes estrategias, inicialmente se utilizó la metodología de resolución de moléculas heteroduplex mediante cromatografía líquida de alta precisión en condiciones desnaturalizantes, mediante secuenciación de ADN y por último se diseñó un ensayo multiplexado de 384 marcadores SNPs usando la tecnología GoldenGate-Veracode (Illumina) basado en variantes alélicas presentes en genes candidatos asociados a respuestas a estreses bióticos y abióticos.

En total se sumaron sesenta y nueve SNPs a la matriz genotípica que contenía 327 AFLPs, 229 SSRs, 8 EST-SSRs y 2 InDels. Estos marcadores fueron ordenados en 17 grupos de ligamiento (GLs) con un valor mínimo de LOD = 4 y una frecuencia de recombinación

máxima de 0,35. El mapa resultante posee 2823,8 cM y una densidad media de 4,57 cM por *loci*.

La población de mapeo fue evaluada para resistencia a PHC en la EEA INTA Balcarce durante 4 campañas utilizando un diseño en bloques completos aleatorizados con tres repeticiones. Las ascosporas del patógeno fueron inoculadas sobre capítulos en estadio R 5.2 mediante aspersión mecánica con una concentración de 2500 esporas/mL. Se registraron los datos de incidencia (IEMAX), severidad (SEV21), intensidad (INT21) y período de incubación (PI).

El mapeo de QTLs de resistencia se llevó a cabo mediante el mapeo por intervalo compuesto. En todos los caracteres, la mayoría de los QTLs identificados mostraron coeficientes de determinación altos, en coincidencia con los moderados a altos valores de heredabilidades obtenidas. Estos parámetros indican que tanto la selección fenotípica como la indirecta serán eficientes.

Para el carácter IEMAX fueron identificados 9 QTLs. Cuatro de los mismos fueron consistentes sobre un segmento comprendido por genes candidatos SNPs ubicados contiguos sobre 10 cM en el GL 15. No sólo por su reproducibilidad, sino también por las magnitudes de sus  $R^2$  (hasta 24%), esta sección genómica es una de la más importante para la tolerancia a IEMAX.

Por otra parte, tanto para los 9 QTLs de SEV21 como para los 12 de INT21, su expresión está circunscripta a un ambiente determinado.

Para PI, fueron encontrados 6 QTLs, de los que una única región genómica sobre el GL 10 parece ser consistente en dos de las campañas consideradas.

A través del uso de marcadores compartidos, el análisis comparativo con otros mapas de girasol permitió corroborar la co-localización de QTLs para PHC en los grupos de ligamiento 1, 2, 8, 10, 11, 12, 15 y 16.

La dispersión en el genoma de los QTLs identificados para los caracteres, confirman la base genética amplia y compleja que hace a la resistencia al PHC del girasol. Este trabajo ha permitido profundizar el conocimiento de dichas regiones, resaltando la importancia de la contribución de las fuentes localizadas en los GL 10 y 15, que se presentan como candidatas a ser analizadas con más detalle para avanzar en la identificación de genes responsables que contribuyan a entender la complejidad de las bases genéticas que gobiernan la resistencia.

Asimismo, los marcadores asociados a los QTLs descritos en esta Tesis podrán ser usados por los programas de mejoramiento en el desarrollo de genotipos con mejor desempeño ante el PHC.

La integración de datos obtenidos por el grupo de trabajo en relación a la caracterización de nuevas fuentes de resistencia para PHC, que incluyen además el análisis de una población de mapeo por asociación constituida por 135 líneas de base genética amplia asociadas al programa de mejoramiento, permitió localizar otros factores asociados a la resistencia en el mapa genético de girasol. De esta manera, tres genes candidato asociados a niveles bajos de incidencia a PHC por mapeo de asociación se localizaron, en los grupos de ligamiento 10, 14, y 15 en las proximidades de los QTL definidos en esta Tesis.

A partir de este trabajo se logró acotar el estudio de la resistencia a PHC a un conjunto reducido de regiones del genoma de girasol sobre las cuales se puede avanzar a partir de distintas aproximaciones de la genómica y la genética para localizar los componentes funcionales responsables de la expresión del fenotipo de interés.

**Palabras claves:** *Sclerotinia sclerotiorum*, podredumbre húmeda del capítulo, girasol, mapeo de QTL, SNPs, genes candidatos.

## **GENETIC ANALYSIS OF PARTIAL RESISTANCE TO *Sclerotinia* HEAD ROT IN SUNFLOWER.**

### **ABSTRACT**

Sclerotinia head rot (SHR) caused by *Sclerotinia sclerotiorum* is one of the most important limiting factors in sunflower crop. This study aimed to identify regions of the sunflower genome associated with resistance to SHR, in order to support the development of assisted breeding strategies to accelerate the process of creating new genotypes with resistance to disease.

A population of 135 Recombinant Inbred Lines (RILs) from cross PAC2 (moderately resistant to SHR) with RHA266 (susceptible to SHR) was used in this study. First, we established a protocol to confirm the intertrial RIL's genetic purity and subsequently saturated a sunflower genetic map with SNP markers. We also conducted field evaluations. Finally, we performed studies to identify those regions responsible for SHR resistance.

For the protocol of genetic purity, 40 SSR markers, which were distributed in 17 linkage groups, were analyzed. At a second stage, 10 markers were selected. These selected markers were able to unequivocally discriminate RILs in this population and also combined fragment sizes and fluorescent marks in such a way that they can be clearly resolved in a single run of capillary electrophoresis. This protocol was implemented in every RIL after every multiplication process. In order to automate and reduce the costs of implementing the protocol, we developed a multiplex PCR assay.

Different strategies were implemented in order to saturate PAC2 x RHA266 genetic map with SNP markers. Initially, resolution of heteroduplex molecules was performed by high precision liquid chromatography under denaturing conditions, by DNA sequencing. Then, a 384-SNP-marker multiplex assay was designed, by using Golden Gate-Veracode (Illumina). This multiplex assay was based on allelic variants present in candidate genes associated with biotic and abiotic stress responses.

Sixty-nine SNPs were added to the genotypic array already containing 327 AFLPs, 229 SSRs, 8 EST-SSRs and 2 InDels. These markers were sorted into 17 linkage groups (LGs) with a minimum LOD = 4 and a maximum recombination frequency of 0.35. The resulting map has 2823.8 cM and has an average density of 4.57 cm loci.

The mapping population was evaluated for resistance to SHR in EEA INTA Balcarce over four seasons using a randomized complete-block design with three replications. Pathogen ascospores were inoculated over R 5.2 stage capitula by mechanical spraying with a 2500

spores/ml concentration. Incidence data (IEMAX), severity (SEV21), intensity (INT21) and incubation period (IP) were recorded.

Resistance QTL mapping was performed using the Composite Interval Mapping (CIM). In all characters, most of the identified QTLs showed high determination coefficients, in coincidence with moderate to high heritability obtained values. These parameters indicate that both phenotypic and indirect selection will be efficient.

Nine QTLs were identified for IEMAX. Four of them were consistent over a genomic region composed of adjacent SNP candidate genes located over 10 cM on LG 15. This is one of the most important genomic sections for IEMAX tolerance, not only for their reproducibility, but for their R<sup>2</sup> magnitudes (up to 24%).

Moreover, the expression of the 9 QTLs of SEV21 as well as of the 12 QTLs of INT21 was circumscribed to a given environment.

Six QTLs were found for IP, of which a single genomic region on LG 10 appears to be consistent in two of the considered seasons.

Through the use of sheared markers, comparative analyses with other sunflower maps allowed to corroborate QTL co-location for SHR in linkage groups 1, 2, 8, 10, 11, 12, 15 and 16.

The dispersion in the genome of the QTLs identified for the characters confirmed how vast and complex genetic base for sunflower resistance to SHR is. This thesis has deepened the knowledge of these regions and highlights the importance of the contribution of pointed sources in LG 10 and 15, which are presented as candidates to be analyzed in detail to advance in the identification of responsible genes that contribute to understand the complexity of resistance genetic basis.

Furthermore, the QTL's associated markers described in this Thesis could be used in improvement programs on genotypes development with better performance against SHR.

The integration of data obtained by this working group allowed the location of other resistance related factors in sunflower genetic map. Thus, 3 genes associated with low levels of incidence at SHR for association mapping candidate genes were located on linkage groups 10, 14, and 15 in the proximity of the QTL identified in this Thesis.

This work could help researchers to limit the study of SHR resistance to a reduced group of sunflower genome regions. From this reduced group, progress could be achieved by applying different approaches in genomics and genetics in order to locate functional components responsible for the expression of the Phenotype of interest.

Key words: *Sclerotinia sclerotiorum*, Sclerotinia Head Rot, Sunflower, QTL mapping, SNPs, candidate genes.

## **AGRADECIMIENTOS**

El presente trabajo ha sido realizado gracias al esfuerzo conjunto de muchas personas, que han colaborado tanto directa como indirectamente con la realización del mismo, es por eso que espero no olvidarme de nadie.

Un agradecimiento tan grande, que no me alcanzarían las hojas para plasmarlo, merece mi Directora de Tesis, Norma Paniego. Desde hace más de 5 años que depositó su confianza en mí, apoyando mis ideas, escuchándome o brindándome consejos oportunos y claros cuando más lo necesitaba. Aparte de su calidad humana se destaca por ser una excelente profesional a la que admiro. Por todo esto y mucho más, gracias!

Si hablar de calidad humana y laboral se trata, las siguientes personas son más ejemplos de ello. Ruth, Paula, Vero L., Caro, Luis y Dani T, que fueron y son compañía habitual de largas horas de charla en el laboratorio desde que llegué al INTA, donde uno se siente un “alumno”. Excelentes personas! Desde el primer momento brindaron su ayuda, tanto desde lo profesional como lo personal. Gracias totales.

A Corina, mi primer mentora, gracias!

A la gente de UGB, a esas manos mágicas que no se comprende cómo hacen para hacer todo lo que hacen. Sí, eso que parece un trabalenguas, es real. Y encima se hacen un tiempo para darte una mano. Apu, Vero Nishi, Pablo, Naty. Gracias!!

Después a ellos, los chicos del labo. Carlita y Fede que desde el principio formamos un grupo muy unido y capaz de entablar charlas de lo que sea. Gracias.

Acá faltó Seba, “el pibe”, pero para que no se enoje le hago un párrafo aparte, jaja. No, todo lo contrario, a pesar de que la gente piensa que peleamos mucho, tienen que saber que eso es “puro marketing”, el pibe es una excelente persona y amigo con el que conviví y compartí millones de cosas.

A los chicos (y no tan chicos) que se incorporaron recientemente, Moni, Juancho, “pulpi” y Mariana, se les agradece.

Me paso al labo de enfrente y tengo a muchos que agradecer. Al grupo “plantas II”, muchas gracias.

En el locutorio cuanta gente para agradecer! Esa gente que hace más amena la estadía frente a la compu. Resumo en un gran “gracias” por esas valiosas charlas y soporte moral.

A Gerardo, mi codirector, que me super ayudó en esta última etapa más que nada. Gracias por ser tan didáctico a la hora de expresar tu gran conocimiento.

A la gente de Balcarce. Carla, Facu, Silvio, Choco, Nereo. Gracias por hacer más llevadera mi estadía en esa ciudad y por la mano que siempre brindaron. Desde la primera vez que me tocó ir, siempre estuvieron presente para cuanto sea necesario.

A Julio Di Rienzo, por su ayuda en la parte estadística, por sus discusiones y sugerencias, y por estar dispuesto a explicarme en todas mis dudas.

A Maxi y a Sergio por ayudarme a resolver fácilmente todas mis dudas, por su ayuda y predisposición brindada siempre. Gracias.

A los chicos del invernáculo, Martin, Nacho, Matías y Agustín por estar siempre y sus colaboraciones en los ensayos.

A Mario (HP), un amigazo con una capacidad impresionante para hacerte sacar una sonrisa en los peores momentos, gracias, "¡¡evalo"!.

A todos los chicos becarios del INTA que forman un grupo excelente de personas con las que compartir, no solamente un almuerzo laboral, sino también eventos y reuniones. Desde los "plantólogos" hasta los "rojos".

A Carina Rivolta, mi consejera de estudios que siempre se mostró activa para resolverme cuanta duda tenga con la facultad. Gracias

Me vuelvo serio un rato, al INTA y sus directivos, por abrirme las puertas a esta prestigiosa institución donde he aprendido tanto y a la vez que he compartido lindos momentos con excelentes personas.

A CONICET por la beca que sin la cual no podría haber desarrollado esta tesis.

A Laura, que me bancó en la recta final y me ayudó a terminarla, gracias!

A mi familia porteña, mis tíos y primas. Gracias por darme una mano y hacer que extrañara menos a los que se quedaron en Misiones.

A mis padres, por su apoyo constante, por el cariño y la comprensión, por incentivar me a seguir mi camino, a mis hermanos, a toda mi familia.

A todos y muchos más, gracias, gracias y más gracias.

**A mis padres, Oscar y Loly, mis  
hermanos, Santi, Eli y Melanie, y  
mis abuelas, Julia y Elisa  
por acompañarme y alentarme durante estos años**

**Parte de los resultados presentados en este trabajo de Tesis fueron publicados:**

Carla V Filippi, Natalia Aguirre, Juan G Rivas, Jeremias Zubrzycki, Andrea Puebla, Diego Cordes, María V Moreno, Corina M Fusari, Daniel Alvarez, Ruth A Heinz, Horacio E Hopp, Norma B Paniego and Verónica V Lia. Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers Aceptado para su publicación en BMC Plant Biology. Fecha 27/11/2014

Filippi C\*, Zubrzycki\* J, Lia V, Heinz R, Paniego N, Hopp HE. Genetics and Genomics Applied to Sunflower Breeding. Sunflowers Growth Dev. Environ. Inlu. Pests/Diseases. Nova Science; 2014. p. 61-94. \*equal co-senior author.

Fusari CM, Lia V V., Nishinakamasu V, Zubrzycki JE, Puebla AF, Maligne AE, et al. Single nucleotide polymorphism genotyping by heteroduplex analysis in sunflower (*Helianthus annuus* L.). Mol. Breed. [Internet]. 2010 Jun 6 [cited 2012 Jun 9];28(1):73-89. Available from: <http://www.springerlink.com/index/10.1007/s11032-010-9462-9>.

## INDICE

<u>INTRODUCCIÓN GENERAL</u> .....	<u>1</u>
1. <u>EL GIRASOL</u> .....	<u>1</u>
2. <u>IMPORTANCIA ECONÓMICA</u> .....	<u>3</u>
3. <u>EL GIRASOL EN ARGENTINA</u> .....	<u>4</u>
4. <u>Sclerotinia Y OTRAS ENFERMEDADES QUE LIMITAN EL CULTIVO</u> .....	<u>6</u>
5. <u>CICLO DE LA ENFERMEDAD</u> .....	<u>8</u>
6. <u>SÍNTOMAS DE LA ENFERMEDAD</u> .....	<u>9</u>
7. <u>MECANISMOS DE ACCIÓN</u> .....	<u>10</u>
8. <u>CONTROL DE LA ENFERMEDAD</u> .....	<u>11</u>
9. <u>ESTUDIO DE LA RESISTENCIA</u> .....	<u>11</u>
10. <u>EL GENOMA DEL GIRASOL</u> .....	<u>12</u>
11. <u>APORTE DEL GRUPO INTA AL ESTUDIO GENÓMICO DEL GIRASOL</u> .....	<u>14</u>
<u>HIPÓTESIS DE TRABAJO</u> .....	<u>17</u>
<u>OBJETIVO GENERAL</u> .....	<u>17</u>
<u>OBJETIVOS ESPECÍFICOS</u> .....	<u>17</u>
<u>CAPÍTULO 1</u> .....	<u>18</u>
<u>INTRODUCCIÓN</u> .....	<u>19</u>
1. <u>LA POBLACIÓN DE MAPEO GENÉTICO</u> .....	<u>19</u>
2. <u>MARCADORES SNPs</u> .....	<u>19</u>
3. <u>MAPEO GENÉTICO</u> .....	<u>23</u>
<u>MATERIALES Y MÉTODOS</u> .....	<u>26</u>
1. <u>MATERIAL VEGETAL</u> .....	<u>26</u>
2. <u>CONTROL DE PUREZA GENÉTICA DE RILs DERIVADAS DE PAC2 x RHA266 y ESTRUCTURACIÓN DE LA POBLACIÓN</u> .....	<u>27</u>
3. <u>GENOTIPIFICACIÓN DE VARIANTES PRODUCIDAS POR MUTACIONES PUNTUALES. . .</u>	<u>32</u>
4. <u>MAPEO DE LIGAMIENTO</u> .....	<u>44</u>
<u>RESULTADOS</u> .....	<u>47</u>
1. <u>ESTRUCTURACIÓN DE POBLACIÓN DE RILs PARA MAPEO DE QTL PARA RESISTENCIA A PHC</u> .....	<u>47</u>
2. <u>CARACTERIZACIÓN GENOTÍPICA</u> .....	<u>50</u>
3. <u>CONSTRUCCIÓN DEL MAPA GENÉTICO</u> .....	<u>55</u>
<u>DISCUSIÓN</u> .....	<u>63</u>

1.	<a href="#"><u>LA POBLACIÓN DE MAPEO GENÉTICO</u></a>	<a href="#"><u>63</u></a>
2.	<a href="#"><u>CONTROL DE PUREZA</u></a>	<a href="#"><u>63</u></a>
3.	<a href="#"><u>SNPs</u></a>	<a href="#"><u>66</u></a>
4.	<a href="#"><u>MAPA DE LIGAMIENTO</u></a>	<a href="#"><u>70</u></a>
<a href="#"><u>CAPÍTULO 2</u></a>		<a href="#"><u>75</u></a>
<a href="#"><u>INTRODUCCIÓN</u></a>		<a href="#"><u>76</u></a>
1.	<a href="#"><u>CARACTERIZACIÓN GENÉTICA DE LA RESISTENCIA CUANTITATIVA A ENFERMEDADES</u></a>	<a href="#"><u>76</u></a>
2.	<a href="#"><u>FENOTIPIFICACIÓN</u></a>	<a href="#"><u>80</u></a>
3.	<a href="#"><u>MÉTODOS DE MAPEO DE QTLs</u></a>	<a href="#"><u>81</u></a>
4.	<a href="#"><u>GENES Y QTLs EN GIRASOL</u></a>	<a href="#"><u>82</u></a>
5.	<a href="#"><u>ALTERNATIVAS A LA MAPEO DE LIGAMIENTO TRADICIONAL</u></a>	<a href="#"><u>86</u></a>
<a href="#"><u>MATERIALES Y MÉTODOS</u></a>		<a href="#"><u>87</u></a>
1.	<a href="#"><u>MATERIAL VEGETAL</u></a>	<a href="#"><u>87</u></a>
2.	<a href="#"><u>CARACTERIZACIÓN FENOTÍPICA. ENSAYOS DE DESAFÍO DE LA POBLACIÓN CON <i>S. sclerotiorum</i> REALIZADOS EN EL CAMPO</u></a>	<a href="#"><u>87</u></a>
2.6	<a href="#"><u>ESTIMACIÓN DE LA HEREDABILIDAD</u></a>	<a href="#"><u>92</u></a>
<a href="#"><u>IDENTIFICACIÓN DE QTLs</u></a>		<a href="#"><u>92</u></a>
<a href="#"><u>RESULTADOS</u></a>		<a href="#"><u>96</u></a>
1.	<a href="#"><u>EVALUACIÓN FENOTÍPICA</u></a>	<a href="#"><u>96</u></a>
2.	<a href="#"><u>ESTIMACIÓN DE LA HEREDABILIDAD</u></a>	<a href="#"><u>100</u></a>
3.	<a href="#"><u>DETECCIÓN DE QTLs</u></a>	<a href="#"><u>101</u></a>
<a href="#"><u>DISCUSIÓN</u></a>		<a href="#"><u>118</u></a>
1.	<a href="#"><u>ANÁLISIS FENOTÍPICOS</u></a>	<a href="#"><u>118</u></a>
2.	<a href="#"><u>ANÁLISIS DE QTLs</u></a>	<a href="#"><u>126</u></a>
<a href="#"><u>CONCLUSIONES</u></a>		<a href="#"><u>137</u></a>
<a href="#"><u>PERSPECTIVAS</u></a>		<a href="#"><u>140</u></a>
<a href="#"><u>BIBLIOGRAFÍAS</u></a>		<a href="#"><u>142</u></a>
<a href="#"><u>APÉNDICE</u></a>		<a href="#"><u>171</u></a>

## ABREVIATURAS

°C	Grados Centígrados
ADN	Ácido desoxirribonucleico
ADT	<i>Assay Design Tool</i>
AFLP	Polimorfismos en la Longitud de Fragmentos Amplificados, del inglés: <i>Amplified Fragment Length Polymorphism</i>
ARN	Ácido ribonucleico
ARNt	ARN de transferencia
ASAGIR	Asociación Argentina de Girasol
ASO	Oligonucleótido Específico de Alelo, del inglés: <i>Allele-Specific Oligonucleotides</i>
BACs	Bacterial Artificial Chromosome
BAG-IM	Banco de Germoplasma Activo de la EEA Manfredi
BrEt	Bromuro de Etidio
CIM	Mapeo por Intervalos Compuestos, del inglés: <i>Composite Interval Mapping</i>
cM	centimorgan
CNV	Variante del número de copias, del inglés: <i>Copy Number Variation</i>
DH	Doble Haploides
dHPLC	Cromatografía Líquida de Alta Precisión, del inglés: <i>Denaturing High Performance Liquid Chromatography</i>
DPI	Días Post Inoculación
EBI	Instituto Europeo de Bioinformática, del inglés: <i>European Bioinformatics Institute</i>
EC	Electroforesis Capilar
ECM	Esterilidad Citoplasmática Masculina
eQTL	QTL de expresión
ESTs	Expressed Sequence Tags
ETI	Inmunidad activada por efectores, del inglés: <i>Effector-triggered Immunity</i>
FDR	Tasa de Falso Descubrimiento, del inglés: <i>False Discovery Rate</i>
GBS	Genotipificación por Secuenciación, del inglés: <i>Genotype by Sequencing</i>
GC	Genes Candidatos
GFPG, EEA	Grupo de Fitopatología de Girasol de la Estación Experimental
INTA Balcarce	Agropecuaria INTA Balcarce
GG-IB INTA Castelar	Genómica de Girasol del Instituto de Biotecnología de INTA Castelar
GL	Grupos de Ligamiento
GLPs	Proteínas del tipo Germinas, del inglés: <i>Germin Like Proteins</i>
GO	Ontología Génica, del inglés: <i>Gene Ontology</i>
Gpb	Giga Pares de Bases
GS	Selección Genómica, del inglés: <i>Genomic Selection</i>
GWAS	Estudio de Asociación del Genoma Completo, del inglés: <i>Genome-Wide Association Study</i>
Ha	Hectáreas
IE	Incidencia de la Enfermedad
IEMAX	Incidencia Máxima registrada en la Evaluación
IM	Mapeo por Intervalos, del inglés: <i>Interval Mapping</i>
InDel	Contracción de "inserción o delección"

INRA	Instituto Nacional de Investigación Agropecuaria de Francia, del francés: <i>l'Institut National de la Recherche Agronomique</i>
INT	Intensidad de la Enfermedad
INTA	Instituto Nacional de Tecnología Agropecuaria
LOD	<i>Logaritmo de Probabilidades, del inglés: Logarithm of Odds</i>
LSO	Oligonucleótido Locus Específico, del inglés: <i>Locus-Specific Oligonucleotides</i>
LTR	Long Terminal Repeats
MA	Mapeo por Asociación
MAS	Selección Asistida de Marcadores, del inglés: <i>Marker-Assisted Selection</i>
Mg	Miligramo
Min	Minutos
MI	Mililitro
Mpb	Mega Pares de Bases
MR	Medianamente Resistente
NCBI	Centro Nacional de Información Biotecnológica, del inglés: <i>National Center for Biotechnology Information</i>
ng	Nanogramos
NGS	Secuenciación de Nueva Generación, del inglés: <i>Next Generation Sequencing</i>
Nm	Nanómetros
PAMPs	Patrones Moleculares Asociados a Patógenos, del inglés: <i>Pathogen-associated Molecular Patterns</i>
Pb	Pares de Base
PCR	Reacción en cadena de la Polimerasa, del inglés: <i>Polymerase Chain Reaction</i>
Pg	Picogramos
PHC	Pudredumbre Húmeda del Capítulo
PHT	Pudredumbre Húmeda del Tallo
PI	Período de Incubación
PRRs	Receptores de Reconocimiento de Patrones, del inglés: <i>Pattern Recognition Receptors</i>
PTI	Inmunidad producida por PAMPs, del inglés: <i>PAMP-triggered Immunity</i>
QDR	Resistencia Cuantitativa a la Enfermedad, del inglés: <i>Quantitative disease Resistance</i>
QTL	Locus de Carácter Cuantitativo, del inglés: <i>Quantitative Trait Locus</i>
RAPD	Amplificación Aleatoria de ADN Polimórfico, del inglés: <i>Random Amplification of Polymorphic DNA</i>
RFLP	Polimorfismos de Longitud de Fragmentos de Restricción, del inglés: <i>Restriction Fragment Length Polymorphism</i>
RIL	Líneas Recombinantes Endocriadas, del inglés: <i>Recombinant Inbred Lines</i>
ROS	Especies Reactivas de Oxígeno, del inglés: <i>Reactive Oxygen Species</i>
Seg	Segundo
SEV	Severidad de la Enfermedad
SHR	Podredumbre húmeda del capítulo = Sclerotinia Head Rot
SIM	Mapeo por Intervalos Simples, del inglés: <i>Simple Interval Mapping</i>
SNPs	Polimorfismo de Nucleótido Simple, del inglés: <i>Single Nucleotide Polymorphism</i>
SOPA	<i>Stranded Conformational Polymorphism</i>

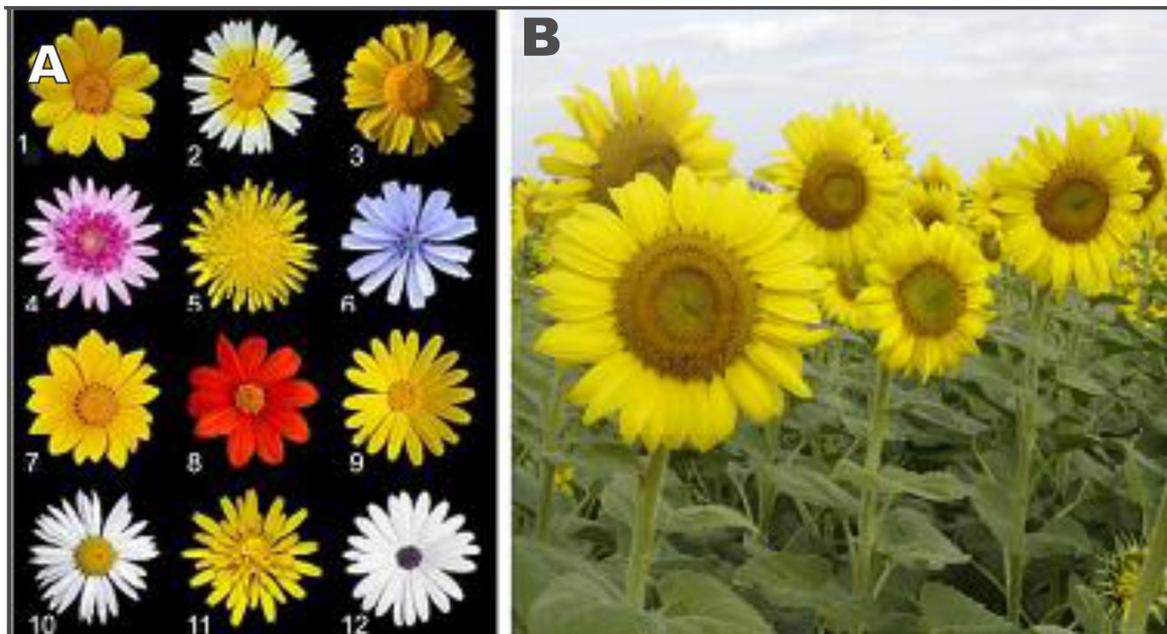
SSCP	Polimorfismo de Conformación de Cadena, del inglés: <i>Single Stranded Conformational Polymorphism</i>
SSRs	Simple Sequence Repeats
SUR	<i>Sunflower Unigene Repository</i>
USDA	Departamento de Agricultura de los Estados Unidos, del inglés: <i>United States Department of Agriculture</i>
UTR	Regiones no Traducidas, del inglés: <i>Untranslated Regions</i>
WGS	Whole Genome Sequencing
Mg	microgramo
MI	microlitro
$\mu$ M	microMolar

# INTRODUCCIÓN GENERAL

## 1. EL GIRASOL

El girasol cultivado (*Helianthus annuus* L.) es una especie perteneciente a la familia Asteraceae (Compositae), subfamilia Asteroideae, tribu *Helianthae*, subtribu *Helianthinae*, género *Helianthus* (Seiler. et al. 1997). La familia Compositae es una de las familias más grandes y diversas de todas las plantas angiospermas (Kallunki & Heywood 1994; Funk et al. 2005), conteniendo cerca de 2400 especies, un décimo de todas las especies con flor (Stevens 2010). Las plantas pertenecientes a esta familia se caracterizan por su estructura reproductiva denominada capítulo, donde las flores se disponen en forma sésil sobre un receptáculo ensanchado (Figura I.1) (Katinas et al. 2007).

La familia probablemente se originó en América del Norte hace 30 a 100 millones de años (Raven & Axelrod 1974) y sufrió una rápida y extensiva diversificación produciendo un gran espectro de taxones. Actualmente, más de 40 taxones económicamente importantes han sido domesticados a partir de especies silvestres pertenecientes a esta familia,



**FIGURA I.1. DIVERSIDAD DE FORMAS EN LOS CAPÍTULOOS DE LA FAMILIA ASTERACEAE.** a. Capítulos en las subfamilias Asteroideae y Cichoreideae: (1) camomila amarilla (*Anthemis tinctoria*, Asteroideae); (2) crisantemo silvestre (*Glebionis coronarium*, Asteroideae); (3) *Coleostephus myconis*(Asteroideae); (4) crisantemo (*Glebionis* sp., Asteroideae); (5) Cerraja (*Sonchus oleraceus*, Cichoreideae); (6) achicoria o radicheta (*Cichorium intybus*, Cichoreideae); (7) Gazania (*Gazania rigens*,Cichoreideae), (8) girasol mejicano (*Tithonia rotundifolia*, Asteroideae); (9) caléndula (*Calendula arvensis*, Asteroideae); (10) margarita (*Leucanthemum vulgare*, Asteroideae); (11) *Hieracium lachenalli* (Cichoreideae); (12) margarita africana (*Osteospermum ecklonis*, Asteroideae). b. Capítulos de girasol cultivado (*Helianthus annuus* spp.*macrocarpus*, Asteroideae).

incluyendo cultivos alimenticios (lechuga, achicoria, topinambur), aceiteros (girasol, cártamo), medicinales (manzanilla) y ornamentales (crisantemo, dalia, zinnia y copete) (Kesseli & Michelmore 1997). En (2010), Barreda et al. describieron un capítulo fósil y granos de polen del Eoceno que fueron hallados en la Patagonia, al sur de la Argentina. Este descubrimiento proveyó evidencia de que la familia evolucionó a partir de un ancestro probablemente originado en Sudamérica, cerca de 50 millones de años atrás (Barreda et al. 2010).

Actualmente, se encuentran representantes de esta familia en todos los ambientes y continentes, con excepción de la Antártida (Funk et al. 2005)

El género *Helianthus* comprende 36 especies perennes y 12 herbáceas anuales (Heiser et al. 1981), entre las cuales se incluyen plantas diploides, tetraploides, y hexaploides, todas con un número cromosómico básico  $x = 17$  (Heiser & Smith 1954). Su nombre deriva de las palabras griegas "*Helios*" que significa sol, y "*Anthus*" que significa flor; y le fue adjudicado por la característica que tiene la flor de mirar hacia el este hasta el período de antesis.

*H. annuus* está compuesto por tres subespecies: *H. annuus* ssp. *macrocarpus* (el girasol cultivado), *H. annuus* ssp. *lenticularis* y *H. annuus* ssp. *annuus*, ambos parientes silvestres del cultivado. Su dispersión geográfica es amplia, creciendo en climas templados y subtropicales.

El girasol cultivado, comenzó su domesticación aproximadamente en el centro-oeste de Norteamérica hace 3000-4000 años atrás gracias a las culturas pre-hispánicas (Crites 1993; Harter et al. 2004; Smith & Yarnell 2009; Bowers et al. 2012), las cuales lo usaron como alimento, consumiendo la semilla directamente o moliéndola para harina e incluso, como colorantes para propósitos ceremoniales (Mandel et al. 2013).

Desde entonces, el girasol ha sido usado también por sus propiedades medicinales, por sus características decorativas, para construcciones primitivas (Putt 1997) y principalmente para uso alimenticio (por consumo del aceite o directamente de las semillas).

Es una planta típicamente alógama y auto-incompatible, aunque su domesticación no sólo permitió la generación de autocompatibilidad, sino también el aumento del tamaño, vigor y número de semillas por capítulo. Esto convierte al girasol cultivado en una especie con características que facilitan la obtención de germoplasma autopolinizable, favoreciendo los programas de mejoramiento con la posibilidad del desarrollo de híbridos. Su tiempo

generacional es relativamente corto, de 70 a 100 días, y el rendimiento es de aproximadamente 1000 semillas por capítulo.

El mejoramiento convencional ha tenido éxito en aumentar el potencial de rendimiento de girasol y su estabilidad, así como en el control de la resistencia a algunas enfermedades fúngicas, plagas y malas hierbas parasitarias (Sala et al. 2012). La aparición y el desarrollo de marcadores moleculares y mapas genéticos han facilitado la comprensión de la base genética de diferentes características agronómicas (Kiani & Sarrafi 2010; Filippi et al. 2014).

Diversas especies silvestres del género *Helianthus* han sido importantes fuentes de genes a lo largo de la domesticación y mejoramiento del girasol cultivado, contribuyendo a la identificación de genes de resistencia a enfermedades, tolerancia a sequía, esterilidad citoplasmática masculina (ECM) y otras características de interés agronómico (Jan et al. 1998). Como ejemplo de ello, los cruzamientos interespecíficos *H. annuus* x *H. petiolaris* permitieron la combinación de la ECM de *H. petiolaris* y de los genes restauradores de la fertilidad de *H. annuus*, dando lugar al desarrollo de la industria de la semilla híbrida de girasol (Fick & Zimmer 1976).

## 2. IMPORTANCIA ECONÓMICA

El girasol es uno de los cultivos oleaginosos más importantes del mundo. El principal producto de la molienda de girasol y el que contribuye, casi excluyentemente, a la formación de su precio es el aceite (Dorrell & Vick 1997).

La semilla de girasol posee un 40% de aceite. A partir de la semilla, se obtiene aceite comestible de alta calidad, considerado además como Premium debido al bajo contenido de ácidos grasos saturados (linolénico) y alto contenido de ácidos grasos mono- y di-insaturados (oleico y linoleico).

En el mercado mundial de aceites, el de girasol constituye el cuarto aceite vegetal más importante a nivel mundial, después del aceite de soja, palma y canola, con una producción cercana al 13% del total (<http://www.fas.usda.gov/>). Hoy en día, se cultiva en más de 25 millones de hectáreas en todo el mundo, con una producción anual que oscila en torno a las 35/40 millones de toneladas anuales. Entre los principales países protagonistas se encuentran: la Federación Rusa, la Unión Europea y Ucrania, con aportes que oscilan entre 6,5 y 9,5 millones de toneladas cada uno, y el único país relevante en el hemisferio sur, la Argentina, con 3 a 3,5 millones de toneladas (<http://www.sunflowerusa.com/stats/world-supply/>).

Por otra parte, existe un interés creciente en el uso de proteínas de girasol en la nutrición humana debido a que las propiedades de las proteínas de girasol son comparables con los de soja y otras proteínas de leguminosas (González-Pérez et al. 2005; Zilic et al. 2010). En este contexto, el contenido de proteína remanente luego de la extracción de aceite de semillas, representa entre el 30 a 50% (Dorrell & Vick 1997).

El resto, principalmente harina de proteína vegetal, se vende como insumo forrajero para la producción de carnes y leche, principalmente. Otra característica de la semilla es su riqueza en fibras, con un contenido en lisina y un nivel calórico más bajo que la de soja, aunque con un contenido superior de metionina respecto a esta última (Paniego et al. 2007).

Además de su uso comestible, el aceite de girasol también es utilizado en la producción de agroquímicos, especialmente en pesticidas hidrofóbicos, surfactantes, adhesivos, plásticos, suavizantes, lubricantes y es, a su vez, un gran potencial para la producción de biodiesel (Paniego et al. 2007).

### **3. EL GIRASOL EN ARGENTINA**

El cultivo de girasol en Argentina se inició a mediados del siglo XIX en manos de inmigrantes rusos afincados en las provincias de Entre Ríos, Santa Fe y Buenos Aires, los cuales realizaron las primeras introducciones de semilla de poblaciones del sur de Rusia como "Mammoth ruso", "Gigante de la China" y "Gigante americano", de los que derivan las primeras variedades locales comerciales con el nombre de Klein, La Previsión 8 y Massaux. Posteriormente, con aportes de otras fuentes de germoplasma, incluidos materiales silvestres de origen ruso y local, se desarrollaron las variedades Manfredi INTA, Impira INTA, Córdoba INTA y Ñandubay INTA. El conjunto de estos materiales se denomina genéricamente "germoplasma argentino", es único en el mundo, posee características bien definidas y está adaptado a nuestras condiciones de cultivo.

En Argentina existen poblaciones naturalizadas de especies silvestres introducidas de *Helianthus annuus ssp. annuus*, *H. petiolaris* y *H. laetiflorus*. Existen además, en estado natural, formas originadas por cruzamientos espontáneos de estas especies con girasoles cultivados (Álvarez 2002).

Actualmente, en la República Argentina, el cultivo se extiende entre los 24 y 38 grados de longitud sur, abarcando una amplia gama de ambientes, desde el norte en la provincia del Chaco hasta el sur de la región Pampeana. En esta zona, el girasol constituye una alternativa importante para los productores (Figura I.2) (Díaz-Zorita et al. 2003). El

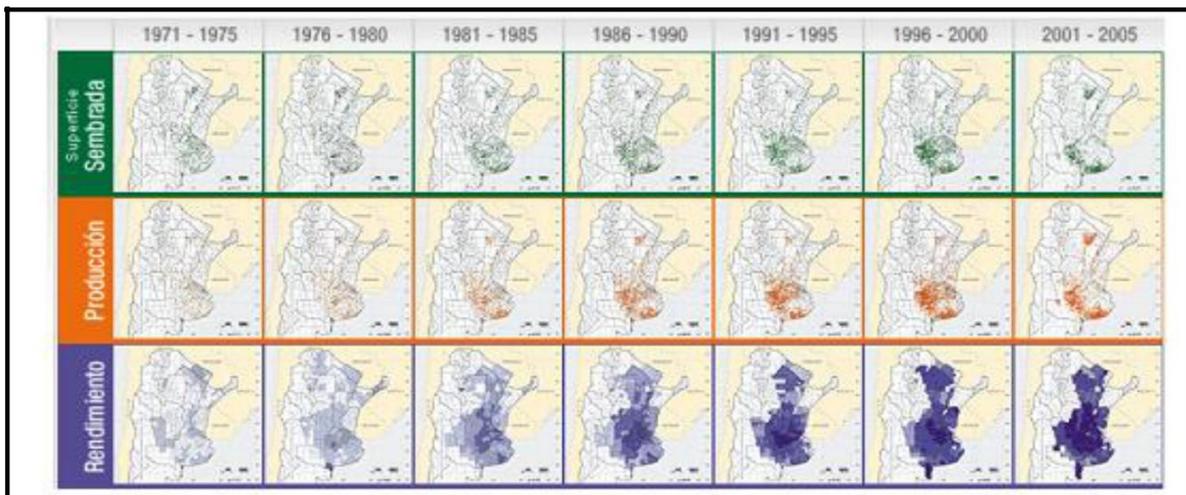
cultivo de girasol es el segundo cultivo oleaginoso de importancia dentro del mercado argentino, luego de la soja.

Actualmente, la producción de girasol en Argentina está prevista en 3 millones de toneladas para el año 2014 (<http://www.sunflowernsa.com/stats/world-supply/>). Hacia finales de los '90, Argentina logró una producción de 6,5 millones de toneladas, convirtiéndose así en el primer productor mundial del grano, con una exportación de 1,7 millones de toneladas de aceite sobre un total de 3 millones comercializadas mundialmente que la ubicó además como principal exportador a nivel mundial.

Luego de la caída de los precios a fines de los '90, y a raíz de la sobreoferta local, la situación del cultivo mejoró entre las campañas 2000-01 y 2005-06 (Figura I.2). La producción de girasol aumentó de 2,93 a 3,95 millones de toneladas, lo que implicó un 34,7% de crecimiento en seis años (Ingaramo 2005).

El área sembrada creció un 26,6%, aunque este aumento se produjo sobre zonas de mayor riesgo agroecológico o en tierras de menor calidad debido a un rápido cambio de las prácticas agrícolas de cultivos como la soja y el maíz, los cuales han incrementado ampliamente sus áreas cultivables como consecuencia de un mercado favorable y la mayor rentabilidad por parte de productos transgénicos con resistencia a herbicidas e insectos (Figura I.2). Por ello, el rendimiento sólo aumentó un 6,4%. Durante la campaña 2007 se produjeron 4,51 millones de toneladas, casi un millón de toneladas más que en la anterior, con una superficie de 2,64 millones de hectáreas (ha) cosechadas (Ingaramo 2005).

Este aumento del rendimiento, a pesar de que las zonas de siembra no hayan sido óptimas, se debió a las mejoras tecnológicas introducidas en la producción del cultivo. Sin embargo, en la campaña siguiente (2008), se produjo una fuerte baja en la superficie sembrada pasando de 2,6 a 1,82 millones de ha, disminuyendo drásticamente la producción. En la última campaña 2012/13 se sembraron 1,65 millones de ha, un 10,8% menos que en la campaña anterior, con una producción de 3,1 millones de toneladas (Asociación Argentina de Girasol). La participación de las distintas provincias en el cultivo de girasol ha ido variando desde 1986 hasta 2005 (superficie sembrada, Figura I.2). En la provincia de Buenos Aires, entre 2000/01 y 2006/07, el área ha aumentado un 23%. En Córdoba, en cambio, se redujo la superficie en un 68%, mientras que en La Pampa, se incrementó en un 30,5%. El Chaco y Santa Fe fueron las provincias que mayor aporte han hecho al crecimiento de la superficie de siembra. La primera ha pasado de 140.000 a 352.000 ha (152% de aumento) y la segunda, de 117.000 a 193.000 ha (un 64% más) (Asociación Argentina de Girasol).



**FIGURA I.2. MAPAS DE LA SUPERFICIE SEMBRADA, PRODUCCIÓN Y RENDIMIENTO DEL CULTIVO DE GIRASOL EN LA REPÚBLICA ARGENTINA DESDE EL AÑO 1971 HASTA 2005.** La superficie sembrada de girasol (verde), abarcaba entre 1971 y 1985 desde el norte del Chaco hasta el sur de la región Pampeana. Se observa un desplazamiento hacia el oeste de la región pampeana y zona chaqueña y una consolidación del sudeste de Buenos Aires como zona de producción del cultivo desde 1995 hasta 2005. A pesar de la reasignación de tierras, la producción (naranja, mayor intensidad = mayor producción) y el rendimiento (violeta, mayor intensidad = mayor rinde) se han ido incrementando. Tomado de "La Argentina en Mapas" (CONICET).

#### 4. Sclerotinia Y OTRAS ENFERMEDADES QUE LIMITAN EL CULTIVO.

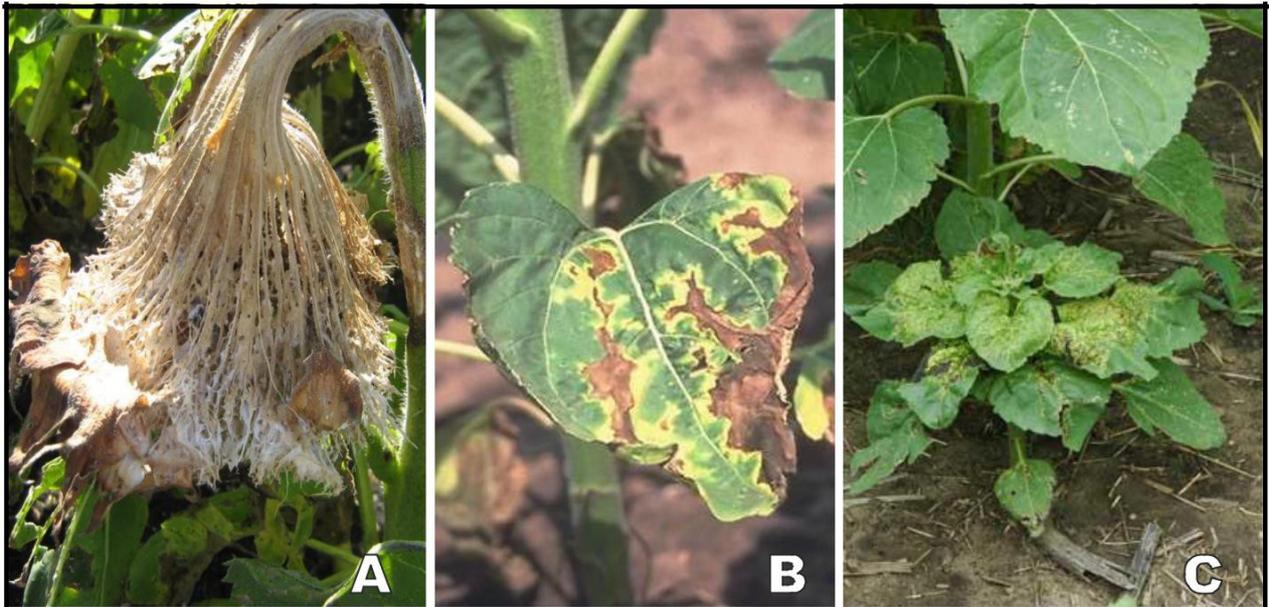
Entre los patógenos más importantes que afectan el girasol en Argentina se encuentran los de origen fúngico, siendo los más relevantes: *Sclerotinia sclerotiorum* (causando entre otras, la podredumbre húmeda del capítulo; PHC), *Verticillium dahliae* (secado anticipado y quebrado del tallo), *Plasmophara halstedii* (mildiu o enanismo) y *Puccinia helianthii* (roya negra) (Figura I.3).

*Sclerotinia sclerotiorum* (Lib.) de Bary pertenece a la clase Ascomycetes, orden Helotiales, familia Sclerotiniaceae, género *Sclerotinia*. Su importancia como patógeno se encuentra descrita en plantas data de hace muchos años (Libert 1837).

Este hongo necrotrófico es un fitopatógeno de difícil manejo (Pereyra & Bazzalo 1988; Troglia 2003) ya que es capaz de infectar más de 400 especies de plantas en el mundo incluyendo importantes cultivos (girasol, soja, colza, frijol comestible, garbanzo, maní, arveja seca, lentejas, lechuga, zanahoria, papa, tomate, y cebolla) (Boland & Hall 1994).

Además, la generación de estructuras de resistencias (llamadas esclerocios) a partir de la agregación y concentración de micelios permiten que el hongo perdure en condiciones que le son desfavorables (Huang & Dueck 1980).

Inicialmente identificado como patógeno del girasol por Fuckel en 1861 (Purdy 1979), *S. sclerotiorum* puede causar diversas enfermedades en el girasol:



**FIGURA 1.3. SÍNTOMAS DE LAS ENFERMEDADES CAUSADAS POR HONGOS EN GIRASOL.** A. Caída del capítulo, causada por *S. sclerotiorum*. B. Marchitamiento y secado prematuro, causado por *V. dahliae*. C. Enanismo causado por *P. halstedii*

Podredumbre húmeda de la base del tallo: la germinación miceliogénica de los esclerocios en el suelo puede infectar raíces y la base del tallo causando marchitez;

Podredumbre húmeda del Tallo: las ascosporas infectan hojas y la infección progresa hacia el pecíolo, produciendo degradación de la médula y quebrado del tallo;

Podredumbre húmeda del capítulo (PHC): las ascosporas infectan las flores tubulosas y la infección progresa produciendo la podredumbre del receptáculo.

*S. sclerotiorum*, se encuentra ampliamente distribuido y su aparición ha sido reportada en todas las regiones girasoleras del mundo (Bolton et al. 2006). En la Argentina, este patógeno afecta al cultivo en muchas regiones, pero la zona más expuesta es el sudeste y sur de la provincia de Buenos Aires (Pereyra & Bazzalo 1988). Aquí, la enfermedad que predomina es la PHC.

La PHC ocasiona pérdidas de producción por pudrición y destrucción de los capítulos y detrimento en la calidad del producto por cuerpos extraños y acidificación de la materia grasa. Esta enfermedad puede causar la pérdida total de la producción en lotes individuales (Pereyra, V.R. et al. 1994). El daño extensivo de los cultivos, junto con la falta de resistencia de los hospedantes y la dificultad de controlar las enfermedades causadas por este patógeno, han sido el impulso sostenido para la investigación en este tema en distintos países (Bolton et al. 2006).

## 5. CICLO DE LA ENFERMEDAD

El potencial máximo de reproducción y su capacidad de supervivencia hacen del esclerocio el componente central en la epidemiología de *S. sclerotiorum*. Los esclerocios en condiciones ambientales favorables (15-20°C y alta humedad) son capaces de germinar generando apotecios que subsecuentemente producen ascosporas, siendo estas últimas las que infectan los capítulos del girasol (Figura I.4). Los apotecios no germinan hasta que el canopeo de los cultivos se cierra favoreciendo el mantenimiento de la humedad en el suelo.

El apotecio consiste en un estípite que se origina a partir del esclerocio y un receptáculo con una capa himenial cóncava plana. Las capas del himenio tienen filas de ascos que son sacos cilíndricos que contienen 8 ascosporas binucleadas. Las ascosporas son descargadas por la fuerza de cada asco en condiciones ambientales favorables por más de 10 días en el campo a una tasa de 1600 ascosporas/hora (Clarkson et al. 2003). Las ascosporas están recubiertas por una capa mucilaginosa que les permite adherirse al sustrato sobre el cual

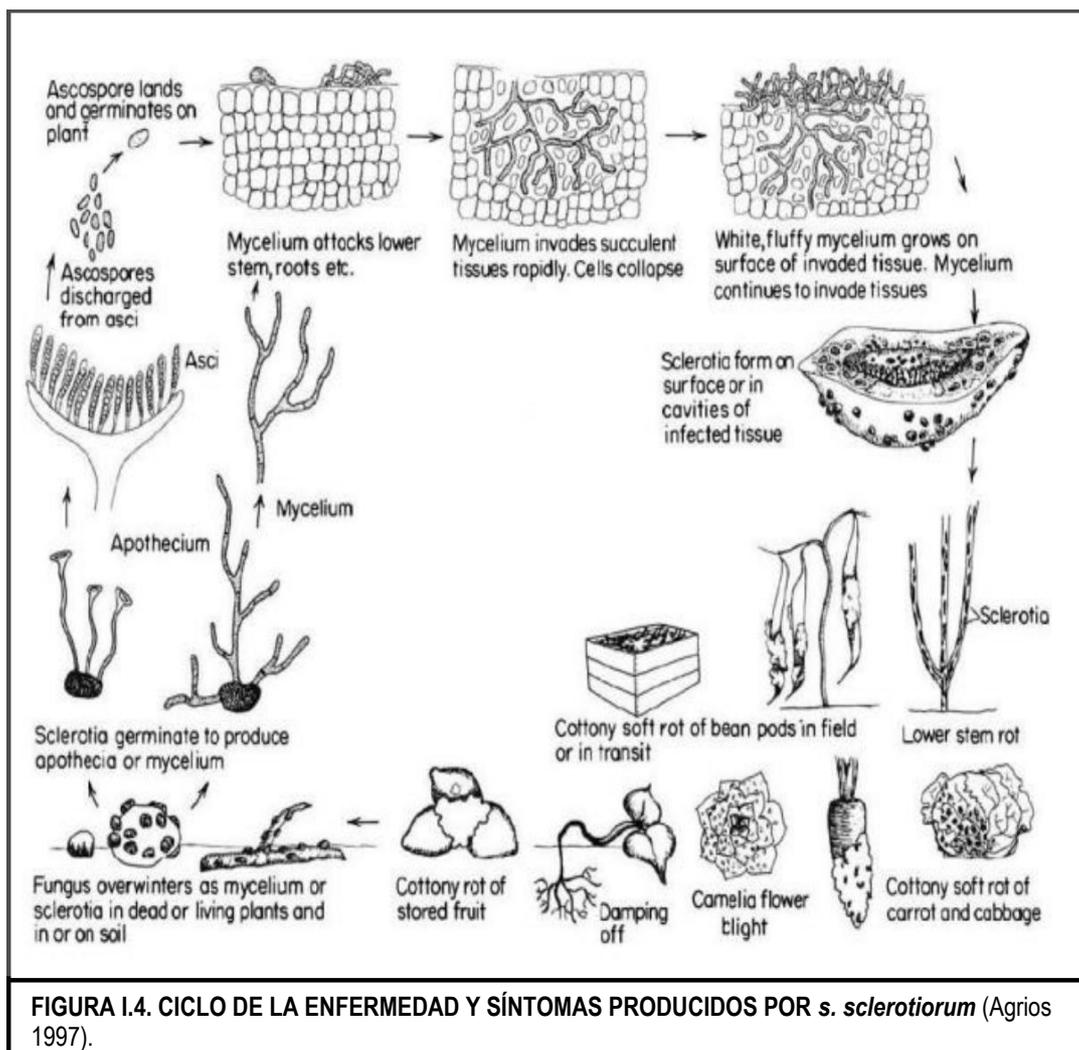


FIGURA I.4. CICLO DE LA ENFERMEDAD Y SÍNTOMAS PRODUCIDOS POR *s. sclerotiorum* (Agrios 1997).

se depositan. Pueden sobrevivir sobre los tejidos vegetales hasta dos semanas dependiendo de las condiciones ambientales, siendo la humedad relativa alta y la luz ultravioleta determinantes para su supervivencia (Clarkson et al. 2003). Las ascosporas pueden germinar en la superficie de tejidos sanos pero no pueden infectar a la planta sin una fuente exógena de nutrientes y una película de agua. Por lo tanto, los tejidos senescentes o necróticos generalmente sirven como fuente de nutrientes para iniciar la germinación, dando origen a una infección micelial de la planta hospedante (Abawi 1979; Lumsden 1969).

La floración es considerada un factor crítico del hospedante asociado a esta enfermedad debido a que las partes de las flores senescentes sirven como fuente primaria de nutrientes (Turkington 1993; Inglis & Boland 1990). Como la floración de los cultivos ocurre en el momento que el canopeo ha cerrado ya el surco, las fuentes de nutrientes están disponibles en el momento en que las condiciones son más favorables para el crecimiento de este patógeno. En la podredumbre húmeda del capítulo, la infección se produce directamente sobre las flores senescentes (Trogia et al. 2005). Debido a que los esclerocios son el inóculo primario para el desarrollo de esta enfermedad en girasol, la densidad del inóculo presente en el suelo está directamente relacionada con la cantidad de enfermedad (Holley 1986). El micelio puede penetrar la cutícula de la planta hospedante mediante enzimas, por fuerza mecánica ejercida por los apresorios o por los estomas (Lumsden 1969; Lumsden 1973). Además de causar estrés hídrico, los estomas abiertos son utilizados por este patógeno para la emergencia de las hifas, para una colonización secundaria de la planta hospedante, así como también para la formación de esclerocios en la superficie de hospedante (Guimarães & Stotz 2004; Lumsden 1973).

## **6. SÍNTOMAS DE LA ENFERMEDAD**

Dado que *S. sclerotiorum* tiene un amplio rango de hospedantes los síntomas no son los mismos en todas las plantas infectadas por el patógeno. En girasol, los síntomas de la PHC se caracterizan por una necrosis floral seguida de manchas color marrón claro en el envés del capítulo y podredumbre húmeda del receptáculo; a medida que la infección progresa el patógeno va consumiendo los tejidos del capítulo hasta dejar sólo las fibras de los tejidos vasculares que dan al mismo el aspecto de escoba. A su vez, el micelio se va concentrando en diferentes puntos dando origen a los esclerocios en el receptáculo (Figura I.5).



**FIGURA I.5. SÍNTOMAS DE LA PODREDUMBRE HÚMEDA DEL CAPÍTULO EN GIRASOL.** A. Capítulo de girasol mostrando síntomas avanzados de la enfermedad. B. Esclerocios.

## 7. MECANISMOS DE ACCIÓN

En cuanto al mecanismo de acción de *S. sclerotiorum*, el ácido oxálico y numerosas enzimas poligalacturonasas se consideran los principales factores de patogenicidad del hongo. Se ha propuesto que el ácido oxálico llevaría a un debilitamiento de la planta debido a la disminución del pH del ambiente y la desregulación de la células oclusivas produciendo la apertura de los estomas y en consecuencia la penetración de las hifas de los tejidos vegetales Guimarães & Stotz (2004). Adicionalmente, se cree que estaría involucrado en el secuestro de  $\text{Ca}^{2+}$  de la pared celular, comprometiendo los mecanismos de defensa mediados por este catión y llevando a una necrosis extendida con pérdida de la integridad de la pared celular (Bateman & Beer 1965; Nordborg et al. 2002). Asimismo, el ácido oxálico también estaría involucrado en la supresión de la maquinaria oxidativa vegetal y en la activación de la muerte celular programada (ACP) (Hegedus & Rimmer 2005).

Para hacer frente a este patógeno, las respuestas de defensa de las plantas involucran diversos mecanismos, incluyendo defensas anatómicas, componentes antifúngicos preformados y compuestos post-infección (Rodríguez et al. 2004).

En genotipos resistentes de varias especies afectadas por *S. sclerotiorum* se ha detectado actividad oxalato oxidasa y/o superóxido dismutasa de proteínas germinas o tipos germinas (GLPs) (p. ej. (Carter & Thornburg 2000; Dumas et al. 1993; X. Hu et al. 2003).) Adicionalmente, la producción de fitoalexinas y compuestos fenólicos en los capítulos de girasol pareciera correlacionar con la resistencia a la infección de *S. sclerotiorum* (Prats et al. 2007).

En el caso de PHT en girasol, se ha descrito que la respuesta al ataque del hongo estaría relacionada con la inducción de la síntesis de proteínas de transferencia de lípidos (LTPs),

proteínas inhibidoras de poligalacturonasas (PGIPs) y defensinas (Hegedus et al. 2008; Regente et al. 1997; Regente & de la Canal 2000; Regente & de la Canal 2003; Urdangarín et al. 2000).

## **8. CONTROL DE LA ENFERMEDAD**

Se ha propuesto un programa de manejo integrado para combatir las enfermedades producidas por *S. sclerotiorum* en girasol que incluye métodos físicos, químicos y biológicos para el control.

Para este manejo se podría ajustar la densidad de plantas (Pereyra & Valetti 1993), rotar con especies no hospedantes (Garza et al. 2002; Troglia 2003); y utilizar funguicidas (Perez Fernandez & Corro Molas 2003) o especies de *Trichoderma* en momentos de mayor riesgo (Escande et al. 2002; Moschini et al. 2002). Sin embargo, el uso de cultivares con resistencia a PHC parece una de las estrategias más promisorias.

La obtención de líneas de girasol resistentes a PHC ha sido objetivo de la mayoría de los programas de mejoramiento. No obstante, el nivel de resistencia logrado no ha sido suficiente en ambientes muy favorables para la enfermedad. Varias especies silvestres de *Helianthus* han sido descritas como fuentes potenciales de genes de resistencia al hongo (Seiler et al. 1997), y se han utilizado para producir híbridos (Kräuter et al. 1991; Schnabl et al. 2002). Si bien algunas líneas derivadas de cruzamientos entre especies silvestres y cultivadas de girasol son prometedoras (Degener et al. 1999; Köhler & Friedt 1999; Ronicke et al. 2004), la incorporación de resistencia mediante técnicas moleculares permitiría reconocer, evaluar la eficiencia y combinar distintas fuentes de resistencia.

## **9. ESTUDIO DE LA RESISTENCIA**

Los factores determinantes de la resistencia a PHC se encuentran asociados a dos mecanismos principales: la entrada del patógeno a las células del hospedante y el desarrollo del micelio (Bioley et al. 1987). Esta resistencia ha sido caracterizada como multigénica, y el efecto individual de los genes involucra pequeños cambios a nivel fisiológico y anatómico (Castaño et al. 1993; Gulya et al. 1997). A su vez el desarrollo de la enfermedad es altamente dependiente de las condiciones ambientales por lo que la búsqueda de resistencia en ensayos a campo es casi siempre problemática especialmente en condiciones naturales de infección.

La resistencia a *S. sclerotiorum* en girasol ha sido descrita como cuantitativa para las tres formas de infección (raíz, tallo o capítulo), con diferentes genes de efecto menor que controlarían la resistencia en los diferentes órganos (Castaño et al. 1993; Bert et al. 2002).

Los resultados hallados en la literatura sugieren asociación significativa entre la infección en raíz, tallo y capítulo cuando se utiliza micelio como inóculo (Castaño et al. 1993). La resistencia no es específica de una raza (Thuault & Tourvieille De Labrouhe 1988) y la acción aditiva de los genes prevalece sobre la dominancia o epistasia (Vear & Tourvieille de Labrouhe 1988; Gentzbittel et al. 1998; Bert et al. 2002). Se han realizado continuos esfuerzos para desarrollar métodos capaces de una selección temprana y líneas con aptitud combinatoria para ayudar a obtener esto (Becelaere & Miller 2004), pero no existen datos que correlacionen la generación de una resistencia temprana o tardía para la resistencia (Micic, Hahn, Bauer, Schön, et al. 2005).

Debido a la complejidad de la base genética de la resistencia a este patógeno, que limita el desarrollo de germoplasma resistente por medio del mejoramiento clásico, los métodos actuales de mejoramiento contemplan la utilización de estrategias basadas en herramientas genómicas.

A pesar de que su genoma no está completamente secuenciado, en los últimos 10 años se ha producido un promisorio avance en los estudios genómicos y post-genómicos aplicados al girasol, lo cual ha incrementado significativamente la disponibilidad de herramientas moleculares, tales como marcadores moleculares microsatélites o SSRs (Simple Sequence Repeats) (Paniego et al. 2002; Tang et al. 2002; Tang et al. 2003), ESTs (Expressed Sequence Tags) (Fernandez et al. 2003; Gentzbittel et al. 1999), mapas genéticos de distinto nivel de saturación (Al-Chaarani et al. 2004; Z. Lai et al. 2005; S. Kiani et al. 2007; Paola Talia et al. 2010; Tang et al. 2002; Yu et al. 2003), mapas físicos (P Talia et al. 2010), colecciones genómicas de clones de BACs (Bacterial Artificial Chromosome) (Bouzidi et al. 2006; Feng et al. 2006; Al-Chaarani et al. 2002) (Bouzidi et al. 2006; Feng et al. 2006; Gentzbittel et al. 2002; Özdemir et al. 2004) y microarreglos de ADN (Alignan et al. 2006; Fernandez et al. 2008; Hewezi et al. 2006; Lai et al. 2006).

## 10. EL GENOMA DEL GIRASOL

La gran mayoría de las especies de la familia Compositae analizadas hasta la fecha poseen un contenido de ADN de aproximadamente 1,1 picogramos de ADN (1C) (Kesseli & Michelmore 1997) con excepción de algunas especies como *Leontodon taraxacoi* que contiene 0,3 pg (<http://compgenomics.ucdavis.edu/cwp/draft.php>).

El girasol cultivado es una especie diploide ( $2n = 2x = 34$ ) cuyo contenido de ADN promedio es de 3,7 pg (1C) ((Seiler. et al. 1997; Baack et al. 2005), estimado en 3,5 y 3,6 mil millones de bases (Baack et al. 2005) y con aproximadamente un 40% de contenido de citosina-guanina (Cavallini et al. 1986; Natali et al. 1993).

La duplicación del ADN, tanto a gran escala involucrando genomas y cromosomas enteros (poliploidización) como a pequeña escala comprendiendo a elementos transponibles (transposones), es responsable en última instancia del aumento en tamaño de los genomas vegetales (Morgante et al. 2007). Este último proceso ha dado lugar a la acumulación de muchas secuencias repetidas que son similares o idénticas a secuencias de otras partes del genoma, pero cuyo número de copias es mucho más alto que el alcanzado posiblemente a través de poliploidización. Se ha descrito en los últimos años el rol que los transposones en la regulación de genes (Xiao et al. 2008; Studer et al. 2011), y en la reestructuración física de los genomas, incluyendo reordenamientos cromosómicos de importante contribución a la especiación (Rieseberg 2001; Faria & Navarro 2010).

En girasol, diversos estudios de curvas de reasociación de ADN revelaron que, el genoma está organizado en fracciones altamente repetidas (representando 5% del genoma), medianamente repetidas (60%) y las llamadas secuencias únicas (comprendiendo el 35% restante) (Cavallini et al. 1986; Natali et al. 1993). Estos estudios no contemplan la composición de la secuencia de cada fracción por lo que no pudieron evaluar la aparición de formas raras de repeticiones dentro de las secuencias únicas, como los restos de retrotransposones. Lo último, llevó a Cavallini et al 2010 a caracterizar el componente repetitivo mediante el análisis de 1638 secuencias de ADN encontrando que la fracción de secuencias repetitivas ascienden al 62% de las secuencias, siendo el mayor componente de dicha fracción representado por retrotransposones LTR (Long Terminal Repeats), especialmente de la superfamilia *Gypsy*.

Las familias de transposones están escasamente representadas en girasol a diferencia de otras especies, en los que constituyen grandes fracciones del genoma (Meyers et al. 2001; Vicient et al. 2005; Vitte & Bennetzen 2006; Neumann et al. 2006). Una repetición (llamada "Contig 61"), cuya familia es de naturaleza desconocida ha demostrado ser la más repetida en el genoma de girasol (Cavallini et al. 2010)

Nuevos estudios sobre el componente repetitivo del genoma del girasol realizados empleando tecnologías de nueva secuenciación describieron que el genoma se encuentra compuesto por más del 81% de elementos transponibles (77% de las cuales eran retrotransposones LTR) Staton et al (2012). Por su parte, Natali et al (2013), usando diferentes enfoques computacionales para el ensamblado de lecturas provenientes de la secuenciación mediante la tecnología Illumina y 454, confirmaron que las secuencias pertenecientes a retrotransposones LTR eran la clase más abundante en el genoma de girasol, representando al menos 79,53% de las lecturas ensambladas (contigs).

La complejidad del genoma de girasol descrita en los párrafos anteriores, ha retrasado la disponibilidad de un genoma de referencia para esta especie que está en proceso desde 2010 por parte del Consorcio Compositae Genome Initiative.

Este consorcio abordó el estudio del genoma desde un enfoque híbrido que combina la secuenciación *de novo* del genoma completo (WGS, whole genome sequencing) empleando las plataformas de nueva generación de Illumina y Roche y la generación de mapas genéticos y físicos de alta densidad que puede servir como anclaje para el conjunto lineal de secuencias obtenidas de la secuenciación del genoma completo (Kane et al. 2011)

Un enfoque similar se ha utilizado con éxito para secuenciar el genoma del pavo (1,1 Gpb) (Dalloul et al. 2010), donde fueron capaces de asignar 917 Mpb (83% del genoma) a los diferentes cromosomas.

Se espera que la referencia del genoma completo de girasol genere un impacto positivo para los estudios de resecuenciación y aumentaría la utilidad de las bases de datos de ESTs (Zhao Lai et al. 2005; Church et al. 2007; Barker et al. 2008; Heesacker et al. 2008), revelando la ubicación de genes, el tamaño de la familia de genes, relaciones entre secuencias parálogas, secuencias promotoras e intrones, y sitios de empalme alternativo. La disponibilidad de un genoma de girasol también servirá como referencia útil para el estudio de la evolución del genoma en *Helianthus* (Rieseberg et al. 1993; Rieseberg et al. 1995; Rieseberg et al. 2003; Burke et al. 2004; Baack et al. 2005; Z. Lai et al. 2005; Ungerer et al. 2006; Heesacker et al. 2009), así como entre *Helianthus* y otras especies en la familia Compositae (Natali et al. 2006; Timms et al. 2006).

Como insumos parciales del estudio genómico completo de girasol se han desarrollado bases de datos del transcriptoma completa para dos genotipos de interés, el genotipo HA412 que es el elegido para la construcción de la referencia de genoma completo y el genotipo HaT13l de interés para los programas de mejoramiento de INRA. Dicha información está disponible para su uso por parte de la comunidad en el sitio <https://www.heliagene.org/HaT13l/> del INRA.

## **11. APORTE DEL GRUPO INTA AL ESTUDIO GENÓMICO DEL GIRASOL**

El grupo Genómica de Girasol del Instituto de Biotecnología de INTA Castelar (GG-IB INTA Castelar) conjuntamente con el Grupo de Fitopatología de Girasol de la Estación Experimental Agropecuaria INTA Balcarce (GFPG, EEA INTA Balcarce), el Banco de Germoplasma Activo de la EEA Manfredi (BAG-IM) y el Programa de Mejoramiento de Girasol, han abordado en los últimos años un análisis integral para identificar y

caracterizar nuevas fuentes de resistencia a estreses en germoplasma propio y germoplasma público internacional.

El grupo GG-IB se inició con el desarrollo de desarrolladas herramientas moleculares como una necesidad para apoyar las estrategias de mejoramiento asistido. Entre sus aportes se incluyen una colección de marcadores neutros SSRs altamente polimórficos (Paniego et al. 2002), un banco de ESTs órgano-específicos aislados de colecciones de ADN copia diferenciales (Fernandez et al. 2003), un mapa saturado de girasol cultivado (S. Kiani et al. 2007; Paola Talia et al. 2010), un microarreglo para estudios de expresión diferencial (Fernandez, Soria, et al. 2012), estrategias de genómica funcional para la validación de expresión diferencial asociados a genes relacionados con resistencia a estrés (Fernandez, Di Rienzo, et al. 2012).

El GFPG, EEA INTA Balcarce lleva adelante el diagnóstico y monitoreo de enfermedades, estudios epidemiológicos y de patogenicidad, identificación y caracterización de resistencia genética, evaluación de fungicidas y desarrollo de modelos predictivos. En particular, el GFPG realiza la evaluación del comportamiento sanitario de cultivares ante PHC en un trabajo coordinado con la Asociación Argentina de Girasol (ASAGIR) y con el BAG-IM (Dr. Alberto Escande, com. pers.).

El BAG-IM, que funciona desde hace más de 30 años alberga cerca de 1200 accesos que provienen de diversos orígenes geográficos y que han sido generados a partir de diferentes metodologías de cruzamiento el cual ha sido caracterizados recientemente desde el punto de vista molecular (Moreno et al. 2013). Uno de los objetivos principales del BAG-IM es, justamente, contribuir al sostenimiento e incremento de la producción nacional de girasol, disminuyendo la vulnerabilidad genética a las distintas enfermedades, a través de programas de mejoramiento. En el caso de PHC, estos programas centran su objetivo en la identificación de nuevas fuentes de resistencia para incrementar la variabilidad genética del germoplasma en cultivo (Ing. Agr. Daniel Álvarez, com. pers.).

Estos tres grupos han abordado en los últimos años un análisis integral para identificar y caracterizar nuevas fuentes de resistencia a la PHC a través de la exploración de germoplasma público y/o desarrollado por el INTA a partir de la utilización de herramientas de la genómica que incluyen tanto la estrategia de mapeo de QTLs conducida a partir de estudios de mapeo de ligamiento y mapeo por asociación (Fusari et al. 2012; Zubrzycki et al. 2012), como así también mediante estrategias basadas en la caracterización de perfiles transcripcionales y metabólicos para la identificación de genes candidatos asociados a la resistencia (Fernandez, Soria, et al. 2012; Peluffo et al. 2010).

Los trabajos de cooperación con el GFPG, EEA INTA Balcarce han permitido la identificación de QTL asociados a la tolerancia a PHC en poblaciones F<sub>2</sub>, obtenidas a partir del cruzamiento de líneas elite sensibles y tolerantes (Maringolo 2007).

Más recientemente, se han optimizado estrategias de análisis de SNPs en girasol de mediano a alto procesamiento (Fusari et al. 2010), así como estudios de desequilibrio de ligamiento para evaluar aplicaciones de estrategias de mapeo de asociación para caracteres complejos, como la resistencia a PHC (Fusari et al. 2012). Adicionalmente, se ha logrado establecer una plataforma bioinformática que integra herramientas dedicadas al procesamiento y análisis sistematizado de secuencias y aglutina los datos producidos por este proyecto y los datos disponibles internacionalmente para la especie ; (atgcsur.inta.gov.ar Fernandez et al. 2012b).

Por otro lado, se han llevado a cabo estudios básicos de transcriptómica y metabolómica orientados a la caracterización funcional de las respuestas de defensa de la planta ante el ataque de *S. sclerotiorum* (Peluffo 2010).

## HIPÓTESIS DE TRABAJO

Dado que la resistencia/tolerancia a *Sclerotinia sclerotiorum* es un carácter complejo y de baja heredabilidad, la base genética del mismo debería estar comprendida por genes/QTLs de efectos génicos menores, por lo que su detección sería facilitada por el uso de mapas genéticos saturados.

## OBJETIVO GENERAL

Saturar el mapa genético de girasol e identificar las regiones del genoma de girasol asociadas a resistencia a la podredumbre húmeda del capítulo (PHC) causada por el hongo *Sclerotinia sclerotiorum* en la etapa inicial de la infección.

## OBJETIVOS ESPECÍFICOS.

1. Desarrollar y mapear marcadores moleculares del tipo SNPs capaces de cubrir espacios existentes en el mapa genético de girasol.
2. Desarrollar e implementar estrategias eficientes de genotipificación automatizada de alto desempeño.
3. Evaluar la tolerancia al PHC de una población de RIL y mapear los QTLs de resistencia a dicha enfermedad.

# CAPÍTULO 1

## **Análisis Genómico de la Población de Mapeo PAC2 x RHA266**

# INTRODUCCIÓN

## 1. LA POBLACIÓN DE MAPEO GENÉTICO

Un elemento a considerar en estudios de mapeo genético y su posterior uso en la identificación de QTLs es el tipo de población a utilizar. Dependiendo del origen de las mismas, se determinan qué alelos evaluar, cuántos alelos por *locus* (mayormente 2) y el grado de resolución genético (Lauter et al. 2008). En especies agronómicas, la mayoría de los estudios publicados se basan en poblaciones estructuradas de mapeo de distinto tipo, incluyendo F<sub>2</sub> y poblaciones de líneas recombinantes endocriadas (RIL), poblaciones doble haploides (DH), entre otras. Sobre la base de estudios de simulación, una población de RIL bien desarrollada parece ser la estrategia más eficiente para el mapeo preciso de QTL (Ferreira et al. 2006).

Las poblaciones RILs son consideradas “inmortales” porque su genotipo ha sido fijado en cada marcador y cada individuo homocigota contiene una mezcla aleatoria de genotipos provenientes de las líneas parentales, que simplifican en gran medida los experimentos repetidos en el tiempo, y aseguran que la población pueda ser almacenada, difundida y utilizada para analizar diferentes caracteres fenotípicos en distintos ambientes.

Algunos estudios sostienen la ventaja de la utilización de RILs en comparación a usar poblaciones F<sub>2</sub> en el desarrollo de mapas genéticos saturados y la detección de QTLs (Austin & Lee 1996), se debe a que las RILs sufren meiosis múltiples, así, los mapas desarrollados presentan mayor poder de resolución y la arquitectura genética de los QTLs (la serie de parámetros que se estiman para un QTL) puede ser determinada con mayor exactitud (Burr & Burr 1991). Además, este tipo de población permiten obtener valores fenotipos más exactos de cada carácter debido a que el número semillas disponibles no es un factor limitante para evaluar numerosas repeticiones en diferentes ambientes o años. El grado de precisión en la medición de las distancias genéticas y la determinación del orden de los marcadores están directamente relacionados al número de individuos estudiados en la población de mapeo. Idealmente, una población de mapeo para la construcción de mapas de ligamiento debe contener como entre 50 individuos y 150 individuos (Collard et al. 2005).

## 2. MARCADORES SNPs

Los marcadores moleculares constituyen una herramienta útil para desarrollo de estudios genéticos y el mejoramiento vegetal. Dentro de ellos, los marcadores del tipo SNPs e InDels (tercera generación) se han ido popularizando a lo largo de la última década (Gupta et al. 2001).

Los SNPs representan diferencias en una única base nucleotídica entre dos secuencias de ADN, son estables de generación en generación y fueron identificados como el tipo más frecuente de polimorfismo en los genomas (Oraguzie et al. 2007). Son consideradas dentro de la categoría SNPs las pequeñas inserciones o deleciones de bases (Insertion/Deletions, InDels), dado que normalmente son identificadas conjuntamente con los SNPs.

En teoría, las variantes SNPs pueden involucrar las cuatro posibilidades determinadas por las bases nucleotídicas (A, T, G o C), pero en general, sólo dos se observan para un determinado locus. Por lo tanto, los SNPs son marcadores bialélicos, cuya información puede ser algo limitada comparada con el polialelismo de los marcadores microsatélites (SSRs). Esto, que en principio es una desventaja, se ve contrarrestado con la gran abundancia y estabilidad de los loci SNPs respecto de los loci SSRs (Oraguzie et al. 2007).

Generalmente, la frecuencia de SNPs en genomas vegetales oscila alrededor de 1 SNP cada 100-300 pares de bases (Ching et al. 2002; Van et al. 2005), mientras que para girasol, diferentes estudios evaluaron la frecuencia de SNPs determinando la ocurrencia de los mismos en 1 cada 39-143 pb (Liu & Burke 2006; Kolkman et al. 2007; Fusari et al. 2008; Pegadaraju et al. 2013). Su abundancia y ubicuidad hace de ellos marcadores moleculares ideales para construir mapas genéticos de alta densidad, realizar mapeo fino de QTL, asistir al mejoramiento, realizar estudios de asociación, y determinar relaciones genéticas entre individuos (Syvänen 2001).

El uso de los SNPs está cada vez más asociado a las aplicaciones que usualmente estaban circunscriptas a los microsatélites (Lijavetzky et al. 2007). Además, como los SNPs pueden encontrarse en regiones codificantes, constituyen una fuente de variación que puede estar relacionada directamente con la función de determinados genes, constituyendo así marcadores moleculares funcionales (Kwok 2001; Rafalski 2002).

## 2.1 *DESCUBRIMIENTO DE SNPs Y ESTRATEGIAS DE GENOTIPIFICACIÓN*

La detección y genotipificación de SNPs siguen siendo un desafío en especies para las que el genoma no se encuentra secuenciado y ensamblado (Seeb et al. 2011).

En general, la identificación de SNPs o InDels se ve favorecida en el caso de las plantas cultivadas, por la disponibilidad de líneas endocriadas que facilitan la tarea de identificación y asociación de los haplotipos de un gen candidato (definidos por un grupo de SNPs) con determinados fenotipos (Rafalski 2002).

En los últimos años, se han desarrollado numerosos proyectos de descubrimiento de SNPs a gran escala en especies cultivables, con el fin de realizar análisis de asociación que permitan el

mejoramiento genético de los cultivos (Feltus et al. 2006; Giancola et al. 2006; Monna et al. 2006; Rostoks, Mudie, et al. 2005; Song et al. 2009; Fusari et al. 2008; Fusari et al. 2010).

En especies no modelo, el enfoque clásico empleado para encontrar variantes nucleotídicas se centra en la búsqueda “*in vitro*” en bases de datos de colecciones de ESTs (Expressed Sequence Tags, secuencias que se expresan) y la secuenciación de las regiones candidatas, o búsquedas “*in vitro*” (Cho et al. 1999; Cordeiro et al. 2006; Kota et al. 2003; Z. Lai et al. 2005; Van et al. 2004; Van et al. 2005; Morales et al. 2004; Somers et al. 2003).

El gran esfuerzo realizado en la búsqueda y detección *in vitro* produjeron un número relativamente bajo de marcadores. De todas maneras, es el enfoque más apropiado cuando existe interés en detectar polimorfismos en genotipos específicos o genes candidatos, o en particular en SNPs sobre regiones no codificantes. En cambio, el descubrimiento *in silico* SNPs en colecciones de ESTs. Este método es el de elección para la identificación *de novo* de SNPs, ya que es posible identificar un gran número de ellos a bajo costo, aunque mayoritariamente en las regiones 3'UTR transcritas.

Hoy en día, existen numerosas alternativas metodológicas para la genotipificación de SNPs de mediano y alto rendimiento (Comai et al. 2004; Lin et al. 2009; Makridakis & Reichardt 2001; Ren et al. 2004; Tsuchihashi & Dracopoli 2002; J. Wang et al. 2005; Xiao & Oefner 2001). Estas permiten interrogar simultáneamente cientos a miles de SNPs en todo el genoma (Gupta et al. 2008). Cada uno de estos enfoques varía en términos de sensibilidad, reproducibilidad, exactitud, capacidad de multiplexado y rendimiento (Sylvänen 2001; Sobrino et al. 2005).

Entre las tecnologías actuales y más comúnmente empleadas para la genotipificación masiva de SNPs, se encuentra el genotipado de alta densidad mediante micromatrices de genes. Varios tipos de micromatrices como las provistas por Affymetrix o Illumina (tecnologías “GeneChip” e “Infinium BeadChip”, respectivamente), se han utilizado en cultivos para la interrogación de hasta 6,5 millones de SNPs por genotipo (Borevitz et al. 2003; Das et al. 2008; Kumar et al. 2007; Rostoks, Borevitz, et al. 2005; West et al. 2006; Song et al. 2013). La base de la identificación y genotipificación de polimorfismos es la hibridación diferencial del ARN o ADN en sondas de 25-50 pares de base de longitud debido a la presencia de SNPs o InDels. De esta manera, se han generado una gran cantidad de marcadores para su posterior uso en programas de mejoramiento.

En girasol, Bachlava et al 2012, reportaron el uso de una micromatriz Infinium de Illumina, conteniendo más de 10000 SNPs para evaluar un panel de diversas líneas de girasol, así como las especies salvajes relacionadas. Los resultados de este trabajo revelaron una alta frecuencia de

SNPs polimórficos (95%) y relativamente alto nivel de transferibilidad entre especies del género *Compositae*.

La tecnología de Illumina GoldenGate, por su parte, ofrece un ensayo de genotipado de SNP basado en la hibridación y la extensión alelo-específica. Permite el genotipado de grandes colecciones de muestras (hasta 480) con diferentes niveles de densidad de SNPs (48-3072). Esta tecnología ha mostrado un alto rendimiento con altas tasas de éxito y reproducibilidad (Ganal et al. 2009) en una variedad de estudios exitosos de diversidad y mapeo a través de una amplia gama de genomas vegetales con diferentes complejidades y conteniendo un gran número de genes parálogos (Hyten et al. 2008; Pavy et al. 2008; Close et al. 2009; Akhunov et al. 2009; Yang et al. 2010; Liu et al. 2012; Rostoks et al. 2006; Thomson et al. 2010; McMullen et al. 2009; Thomson et al. 2011).

En los últimos años, el uso de tecnologías de secuenciación de nueva generación para la genotipificación por secuenciación (GBS, genotype by sequencing) se está haciendo cada vez más importante, rentable, y una herramienta única para estudios en plantas y animales, incluyendo aquellos con genomas complejos que carecen de una secuencia de referencia. Esto se debe mayormente a la posibilidad de generar un gran conjunto de datos sin un esfuerzo a priori para identificar marcadores. Esta tecnología ha demostrado ser eficiente en estudios de mapeo genético y mejoramiento (Baird et al. 2008; Poland et al. 2012; Elshire et al. 2011).

Por otra parte, frente a las necesidades de reducir costos de genotipificación, se han desarrollado técnicas que involucran la detección de SNPs de manera eficiente y económica, a través de la formación de moléculas heterodúplex. Este principio de detección de SNPs ha sido utilizado en la práctica, a través de distintas aproximaciones experimentales, incluyendo hibridación de oligonucleótidos (Pease et al. 1994; Southern 1996), ensayos Taq-Man® (Higuchi et al. 1993; Livak et al. 1995), “molecular beacons (Barreiro et al. 2009; J. Wang et al. 2005)”, ensayos electrónicos de “dot blot” (Gilles et al. 1999), corte con enzimas endonucleasas específicas (Oleykowski et al. 1998) y cromatografía líquida de alto rendimiento en condiciones de desnaturalización parcial (dHPLC) (Xiao & Oefner 2001).

En particular, la técnica de dHPLC (denaturing High Performance Liquid Chromatography) resultó ser una alternativa útil para el análisis de heterodúplex y se ha utilizado en numerosas oportunidades mostrando alta sensibilidad y eficiencia (Giordano et al. 1999; Kota et al. 2001; Kota et al. 2008; Z. Lai et al. 2005; Randall et al. 2005; Wagner et al. 1999; Fusari et al. 2010). La principal ventaja de este método es que no necesita ninguna información a priori sobre el SNP, aunque es capaz de detectar su presencia o ausencia, no la naturaleza y localización de la mutación.

En este contexto actual de continua generación de altos niveles de información, resulta interesante determinar si los datos generados (en este caso SNPs) permitirán alcanzar los objetivos planteados. Por esta razón, es importante determinar el número óptimo SNPs, dependiendo de la aplicación ulterior. Para muchos estudios, especialmente aquellos que involucran a grandes poblaciones, resulta de importancia considerar el costo por muestra más que el costo por dato de genotipado, ya que un costo relativo elevado impide generalizar su uso en programas de mejoramiento aplicados (Thomson et al. 2011).

Mientras que las matrices de SNPs de alta densidad pueden ser útiles para GWAS y selección genómica (GS), un subconjunto de marcadores SNPs podría ser adecuada para mapeo de QTLs en poblaciones biparentales, mapeo por asociación de genes candidatos, Selección asistida por marcadores (MAS), y estudios de diversidad genética en poblacionales (Deulvot et al. 2010; McCouch et al. 2010; Boualaphanha et al. 2011; Chen et al. 2011; Kwon et al. 2011; Thomson et al. 2011). Esto fue demostrado por Moragues et al (2010), quienes encontraron que 384 SNPs proporcionan una buena combinación de la capacidad discriminante para la caracterización del germoplasma de cebada, después de comparar el costo y la efectividad del uso de diferentes subgrupos de 96 o 384 SNPs frente a un conjunto original de 1536 SNPs para estudios de diversidad.

### 3. MAPEO GENÉTICO

Un mapa genético es una representación del genoma de una especie, constituido por diversos *loci* (genes o marcadores), y distribuidos en grupos de ligamiento (GL) que se corresponden con los cromosomas de la especie. Este resulta del análisis de ligamiento de marcadores segregando en una población de mapeo acorde a las proporciones mendelianas.

Considerando la hipótesis nula de que un par de *loci* se hereda de forma independiente ( $r = 0,5$ ), la frecuencia de recombinación ( $r$ ) se estima inicialmente entre pares para luego ser ordenados en un GLs. Por esta razón, un buen mapa de ligamiento consiste de un conjunto de marcadores en segregación mendeliana que están eventualmente espaciados y abarcan el genoma completo.

En las últimas décadas, varios mapas de ligamiento genético difiriendo en longitud, densidad y tipo de marcadores moleculares, fueron desarrollados para el girasol cultivado. Gentzbittel et al (1994; 1995) han demostrado que la técnica de RFLP puede ser utilizada para identificar genotipos de girasol y que hay suficiente variabilidad para desarrollar un mapa marco de RFLP. Desde entonces, mapas más densos fueron desarrollados, basado en RFLPs (Berry et al. 1996; Berry et al. 1997; Jan et al. 1998; Gentzbittel et al. 1999), RAPDs (Rieseberg et al. 1993; Rieseberg 1998), AFLPs (Peerbolte et al. 1998; Berrios et al. 2000; Gedil et al. 2001; Al-Chaarani et al. 2002), DALPs (Langar et al. 2003), TRAPs, que utiliza la información de base de datos de

EST para generar marcadores polimórficos sobre secuencias de genes candidatos (J. Hu et al. 2003).

A partir de la década pasada, el desarrollo de microsatélites en girasol sumó aproximadamente 2040 marcadores (Dehmer & Friedt 1998; Paniego et al. 2002; Tang et al. 2002; Yu et al. 2002). Estos marcadores se incluyeron en los mapas genéticos existentes así como también en el desarrollo de nuevos mapas pudiéndose homologar la nomenclatura para denominar los GLs generados a partir del estudio de distintas poblaciones biparentales. Dentro de los mapas de ligamiento que permitieron referenciar los trabajos realizados en girasol se encuentran los originados a partir del estudio de las poblaciones RILs PAC2 x RHA266, RHA280 x RHA801 y PHA x PHAB (Berrios et al. 2000; Tang et al. 2002; Tang et al. 2003; Yu et al. 2003; Al-Chaarani et al. 2004; S. Kiani et al. 2007)

A partir de entonces, los mapas de referencia comenzaron a ser usados para incrementar la densidad de marcadores en ciertas regiones específicas de otros mapas. Todos los mapas anteriormente nombrados, además de mapas inéditos, pueden ser accedidos a través de la base de datos de girasol Cmap (<http://www.sunflower.uga.edu/cmap/>).

De esta manera, los SSRs se convirtieron rápidamente en los marcadores de elección para los análisis de ligamiento debido al hecho de que son altamente polimórficos, por lo general se heredan de manera codominante y, en la mayoría de los casos son cromosoma específico.

Actualmente, las tecnologías de secuenciación de nueva generación y plataformas de genotipado de alto rendimiento permiten interrogar simultáneamente cientos o miles de loci en los individuos que componen las poblaciones de mapeo (Gupta et al. 2008). Estas técnicas innovadoras y relativamente accesibles abrieron una perspectiva que hasta los últimos años no se había considerado en función de las limitaciones técnicas y presupuestarias. Por otro lado, la creciente disponibilidad de secuencias parciales o completas de genes y marcadores a partir de los proyectos genómicos, han permitido el desarrollo de miles de marcadores SNP y la utilización de éstos para el abordaje del mapeo genéticos cada vez más densos para la identificación de genes y polimorfismos responsables mediante el mapeo de QTLs.

Como se mencionó anteriormente Bachlava et al 2012 publicaron un ensayo de genotipificación de 10000 SNPs multiplexados usando la tecnología Infinium de Illumina. En este trabajo se utilizaron cuatro poblaciones de mapeo de girasol para producir un mapa consenso del genoma de girasol, que constituye el mapa más extenso reportado en la literatura hasta ahora (Bowers et al. 2012). Individualmente, estos mapas genéticos de alta densidad contienen entre 3500 y 5500 loci cada uno y a pesar de fueron altamente colineales, aún se observan regiones discontinuas de

hasta 26 cM de longitud carentes de marcadores. Esto fue subsanado a la hora de formar el mapa consenso de 10080 loci, lo que demuestra el valor de analizar simultáneamente distintas poblaciones de mapeo. Este mapa genético de alta densidad es una herramienta que a su vez, permitiría facilitar el ensamblado del genoma de girasol en proceso (Kane et al. 2011).

Por último, Bowers et al (Bowers et al. 2012) describieron la caracterización genotípica de un subconjunto de 69 líneas de la población de mapeo RHA280 x RHA801 utilizando una matriz de Affymetrix personalizada que contiene 2,4 millones de sondas, representando 86023 unigenes de girasol y especies relacionadas. El mapa generado previamente por Bachlava et al (2012), se usó para filtrar los datos de baja calidad y localizar 22481 unigenes, en el mapa genético de girasol. El mapa resultante contiene una fracción sustancial de todos los genes de girasol, y además de facilitar el ensamblado del genoma, permitiría la identificación de genes candidatos subyacentes a los QTLs.

# MATERIALES Y MÉTODOS

## 1. MATERIAL VEGETAL

En esta Tesis se trabajó con una población biparental de 135 líneas recombinantes endocriadas (RILs, de sus siglas en Inglés: *Recombinant Inbred Lines*), obtenidas a partir del cruzamiento de las líneas PAC2 x RHA266 de comportamiento medianamente resistente y susceptibles frente a la Podredumbre Húmeda del Capítulo (PHC) respectivamente, desarrolladas por la Dra. Veary y cedidas por el Dr. Gentzbittel, del Institut National de la Recherche Agronomique (INRA, Francia), en el ámbito de una cooperación académica entre este grupo y el Instituto de Biotecnología, INTA Castelar (Tabla 1.1).

TABLA 1.1. Origen de las líneas parentales de girasol utilizadas en el presente trabajo. USDA, United States Department of Agriculture (USA); INRA, Institut National de la Recherche Agronomique (France).		
LÍNEA	ORIGEN	INSTITUCIÓN
PAC2	<i>H. petiolaris</i> (restaurador de la fertilidad) x línea USDA (HA61)	INRA
RHA266	<i>H. annuus</i> salvaje x Pederovik	USDA

Normalmente el mantenimiento y multiplicación de las líneas o genotipos se realiza en la Estación Experimental Agropecuaria (EEA) Balcarce y contra-estación en la EEA Yuto (Jujuy). Para prevenir la fecundación cruzada y mantener la pureza de las líneas, los capítulos de cada línea fueron cubiertos con bolsas de autofecundación (Fitotech) en estadios R5. Las semillas cosechadas se guardaron secas a una temperatura de 4° C.

### 1.1 EXTRACCIÓN DE ADN

Hojas de plantas jóvenes (con 3 semanas de crecimiento y un estadio fenológico V2-3) fueron recolectadas en sobres acondicionados para tal fin y congeladas inmediatamente en nitrógeno líquido, finalmente el material vegetal se almacenó liofilizado a -20°C. La extracción de ADN se realizó a partir de 20 a 25 mg de material liofilizado al cual se le incorporaron 3 bolillas de tungsteno (QIAGEN, CAT 69989) para realizar la disrupción del tejido con el sistema TissueLyser (QIAGEN).

El ADN se extrajo mediante el sistema NUCLEOSPIN Plant II (Macherey-Nagel), de acuerdo a las especificaciones técnicas del proveedor.

### 1.2 EVALUACIÓN DEL ADN EXTRAÍDO

La integridad del ADN genómico total se evaluó mediante electroforesis en geles de agarosa al 1% (p/v) (TAE 1X: Tris-Acetato 20 mM, pH 8,5, EDTA 1mM) de acuerdo a lo descrito por Sambrook y col. (1989). La concentración de bromuro de etidio (BrEt) utilizada fue de 0,3 µg/ml y el volumen de muestra se mezcló con solución de siembra (1/6) bromophenolblue 0,25%. Para

la corrida electroforética se aplicó voltaje constante entre 1 y 5 V/cm. Para la visualización del ADN se empleó un sistema de adquisición de imágenes provisto con un transiluminador de luz UV (320 nm) con cámara incorporada (GEL DOC XR System, BioRad).

### 1.3 CUANTIFICACIÓN DE ADN

La cuantificación de ADN se realizó por medio de 2 metodologías:

- Espectrofluorimetría en placas de 96 pocillos (Fluorómetro Spectra MAX GEMINI EM, Molecular Devices, EEUU). Se utilizó el reactivo HOECHST 33258 (Intercalante de ADN, PolySciences, Cat #09460) de acuerdo al protocolo que se detalla a continuación. Se construyó una curva de calibración a partir de ADN genómico de timo de cabra en concentraciones conocidas: 22,5; 45; 90; 180; 360 y 720 ng/μl. Luego se mezclaron 2 μl de ADN, tanto del ADN incógnita como de la curva de calibración, con 200 μl de la solución de trabajo TNE-Hoechst (HOECHST 100 ng/ml y TNE 1X: Tris 10 mM, NaCl 200 mM, EDTA 1 mM, pH 7,4). El espectro se ajustó a una longitud de onda de excitación ( $\lambda$ ) de 365 nm y una longitud de onda de emisión ( $\lambda$ ) de 468 nm. La lectura obtenida para los datos de los estándares se ajustaron a una recta y las incógnitas se determinaron por interpolación.
- Espectrofotometría de luz UV (NanoDrop ND-100 spectrophotometer, NanoDrop Technologies). Este equipo estima la concentración a partir de la absorbancia (A) a 260 nm, asumiendo que  $A(260\text{ nm}) = 1$  equivale a 40 μg/ml de ADN; además, calcula la relación  $A_{260\text{ nm}}/A_{280\text{ nm}}$ , teniendo en cuenta que el pico máximo de absorción de las proteínas y otros contaminantes es 280 nm y la relación  $A_{260\text{ nm}}/A_{230\text{ nm}}$ , siendo la absorbancia a 230 nm, la correspondiente a distintos polifenoles; de este modo una relación cercana a 2 indica una buena calidad y pureza de la muestra de ADN. En base a la concentración calculada, el ADN extraído fue llevado a una concentración final de 30 μg/ml, a fin de establecer un estándar para realizar los pasos siguientes y evitar inconvenientes a causa del exceso o defecto en la concentración.

## 2. CONTROL DE PUREZA GENÉTICA DE RILs DERIVADAS DE PAC2 x RHA266 y ESTRUCTURACIÓN DE LA POBLACIÓN

Para analizar la pureza de las líneas endocriadas que componen las poblaciones de mapeo en estudio para resistencia a PHC se estableció el protocolo de calidad descrito a continuación.

Un mínimo de 5 semillas generadas a partir de la autofecundación de cada RIL, fueron sembradas en macetas y germinadas en invernáculo hasta obtener una plántula de aproximadamente 3 semanas. Una muestra de hojas verdes fue colectada, congelada, liofilizada y

almacenada a  $-20^{\circ}\text{C}$  para posteriormente realizar la extracción del ADN. De la misma manera se procedió luego de cada campaña de multiplicación.

### 2.1 SELECCIÓN DE SSRs Y DISEÑO DEL ENSAYO DE CONTROL DE PUREZA

Para el control de calidad genotípica, se seleccionaron en primera instancia 39 marcadores microsatélites localizados en el mapa descrito por Talia et al (2010) (descriptos en la Tabla I del apéndice) previamente mapeados en la población PAC2 x RHA266, representativos de los 17 GL del mapa de girasol y con bajo contenido de datos faltantes. La secuencia de los iniciadores así como la información inherente al fragmento amplificado, su ubicación y localización sobre mapas públicos, se encuentra disponible públicamente en la base de datos “Probes” del Centro Nacional de Información Biotecnológica (de sus siglas en inglés, NCBI, <http://www.ncbi.nlm.nih.gov/probe/>).

El ensayo de control de pureza se basó en la amplificación y la resolución de fragmentos mediante electroforesis capilar (EC) usando un analizador de ADN (Applied Biosystem, ABI 3130 XL).

Para explotar al máximo las posibilidades del sistema, economizar y ampliar la capacidad de procesamiento de muestras, se tuvieron en cuenta dos aspectos importantes: la marca de los oligonucleótidos y la preparación de las muestras. Se combinaron fragmentos amplificados de diferentes tamaños y conteniendo diferentes fluoróforos (Tabla 1.2), evitando así la superposición de fragmentos de la misma cantidad de pares de bases (pb) y del mismo color. Lo último permitió reunir en cada panel un grupo de hasta seis marcadores compatibles para aumentar la eficiencia de discriminación entre alelos en una sola corrida.

### 2.2 AMPLIFICACIÓN DE LOS SSRs SELECCIONADOS

Cada uno de los marcadores fue amplificado individualmente en las líneas objeto de multiplicación.

TABLA 1.2. Descripción de los fluoróforos empleados.		
FLUORÓFORO	COLOR EMITIDO	EMISIÓN DE FLUORESCENCIA (nm)
6-FAM	Azul	535
HEX	Verde	560
NED	amarillo	585
ROX	Rojo	604

La amplificación de los productos mediante reacción en cadena de la polimerasa (PCR) se realizó en el termociclador Eppendorf (Mastercycler epgradient S) en 12 µl volumen final y 30 ng de ADN genómico.

La mezcla de reacción se compuso de 2 mM MgCl<sub>2</sub>, 0,2 mM dNTPs, 1 U Polimerasa Taq Platinum (Invitrogen) y 0,25 mM de cada uno de los oligonucleótidos iniciadores específicos. Con el objeto de evitar amplificaciones espúreas, la reacción de amplificación fue llevada a cabo por enriquecimiento del templado mediante disminución progresiva de la temperatura de hibridación (“touch down”) (Paniego et al. 2002) de acuerdo al rango de temperatura óptima de unión de los iniciadores a probar a la secuencia específica del ADN (“annealing”).

Las condiciones generales de ciclado fueron:

- una desnaturalización inicial de 4 min a 94°C, 24 ciclos de (1) 30 seg a 94°C, (2) 30 seg a 64°C (-0,5°C/ciclo, touchdown 64-52) y (3) 30 seg a 72°C, seguido por 35 ciclos de (1) 30 seg a 94°C, (2) 30 seg a 52°C y (3) 30 seg a 72°C, con una extensión final 10 min a 72°C. La rampa de enfriamiento fue fijada a 50% (normalmente desciende 4,5 °C/seg) entre cada paso de los ciclos;

### 2.3 GENOTIPIFICACIÓN POR ELECTROFORESIS CAPILAR

En todos los casos, previo a la preparación de los paneles de co-electroforesis, los fragmentos se evaluaron en geles de agarosa para evitar que haya superposiciones de tamaños de fragmentos marcados con el mismo fluoróforo y se realizó la dilución correspondiente para que la intensidad de la señal de los distintos fluoróforos se encuentre dentro del rango de fluorescencia adecuado (entre 200 y 4.000 RFU). En este proceso se tuvieron en cuenta las distintas intensidades de emisión que presentan los fluoróforos utilizados. La señal emitida por 6-FAM es más alta que la emitida por HEX, y ésta que la emitida por NED.

Para el protocolo de preparación de las muestras para su resolución por EC, se realizó primeramente una optimización de la corrida electroforética con pocas muestras, y luego, en una segunda etapa, se realizó el ensayo con el total de las muestras a analizar.

Los fragmentos resueltos por EC fueron analizados de manera automática utilizando el programa GeneMapper® 4.0 (Applied Biosystem, EUA) y corregidos manualmente usando el mismo programa.

Los electroferogramas colectados fueron analizados automáticamente por comparación con una curva construida a partir del frente de corrida de un estándar interno de tamaño molecular marcado con el fluoróforo ROX, 6 carboxy-X-rhodamine (Applied Biosystems, N. Y. U.S.A.). Este estándar de tamaño molecular fue GeneScan ROX500 (-250). El mismo consta de 15 fragmentos

de: 35, 50, 75, 100, 139, 150, 160, 200, 300, 340, 350, 400, 450, 490 y 500 pb. El fragmento de 35 pb debe corregirse de manera manual en el análisis, ya que migra anormalmente y distorsiona los resultados de la curva estándar (Figura 1.1).

Para la genotipificación de marcadores y asignación de alelos se emplearon los criterios establecidos por Talia et al (2010).

#### 2.4 ANÁLISIS DE LA PUREZA GÉNICA

Para cada línea, el perfil multilocus de genotipos obtenidos fue procesado con la herramienta GenAlex (v.6.41) incorporada como extensión al programa Excel 2010 (Microsoft). Con el mismo se comparó el conjunto de genotipos provenientes de las líneas incógnitas (y que fueron objeto de multiplicación) con un set de 94 RILs genotipadas por Talia et al (2010) y que forman parte del set de datos originales. Esto permitió validar los materiales, mantener la homogeneidad de las líneas generación tras generación o resolver problemas de pérdida de etiquetas durante los procedimientos en el campo.

#### 2.5 ESTABLECIMIENTO DE UN PROTOCOLO DE CALIDAD

En una segunda instancia y a partir de los datos obtenidos, se seleccionaron 10 marcadores SSRs para desarrollar e implementar un protocolo de control de calidad automatizado y eficiente basado en una reacción de PCR multiplexada. La combinación de 10 marcadores aportando dos alelos informativos para la población, permitirá discriminar un individuo en una población de tamaño 1024 ( $2^{10}$ ).

Para desarrollar este protocolo los oligonucleótidos cebadores para cada marcador de los 40 evaluados originalmente, fueron analizados en conjunto con el programa Multiplex Manager ([www.multiplexmanager.com](http://www.multiplexmanager.com)). A partir de la información obtenida se optimizó la combinación de dichos cebadores marcados con los fluoróforos seleccionados para la técnica de

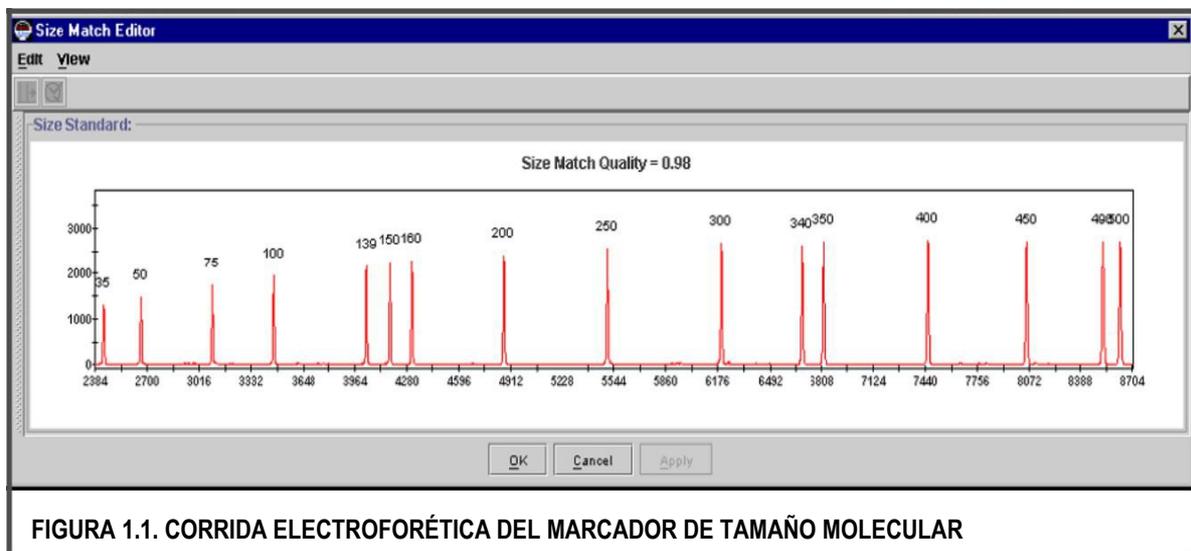


FIGURA 1.1. CORRIDA ELECTROFORÉTICA DEL MARCADOR DE TAMAÑO MOLECULAR

genotipificación por EC. El criterio de selección se basó en la longitud esperada de cada uno de los fragmentos amplificados, las temperaturas de hibridación de cada oligonucleótido con el templado y las posibles estructuras secundarias o complementariedades encontradas entre los iniciadores que conforman la mezcla. Brevemente, sobre este último punto el programa asigna un valor a cada alineamiento obtenido entre pares de iniciadores, asignando el valor +1 a las complementariedades entre bases y -1 a las bases no apareadas. De esta manera, determinando un valor umbral de corte de 5, es capaz de discernir entre pares de iniciadores que compondrán el multiplexado final. La Tabla 1.3 describe a los 10 SSRs seleccionados.

TABLA 1.3. Descripción de los oligonucleótidos seleccionados para el protocolo de control de pureza					
NOMBRE	OLIGONUCLEÓTIDOS INICIADORES			TAMAÑO ESPERADO	TAMAÑO ESPERADO
				PAC2 (pb)	RHA266 (pb)
HA1590	F: 5' 6-FAM R: 5'	GTGGGAGAGAAGATAGGAG GGGGATTATGCAACCTTGCC	3' 3'	66	68
HA3555	F: 5' 6-FAM R: 5'	GATATCTCTCATAAGTGCCG GGTCTTGTGATGACGAAGC	3' 3'	104	102
HA1752	F: 5' 6-FAM R: 5'	CCTTCTCCTCAAATCCCG CGTCATTGTCATTGAGAAGATCG	3' 3'	143	139
HA911	F: 5' 6-FAM R: 5'	CAAAGTTCACCTCGTTTTTC AAGTGGGAAGGTCTACGAGT	3' 3'	179	186
HA2190	F: 5' NED R: 5'	CAAACCCTAATCGCCCAATTGC GGAAGCTTGTGAATCTTGGAGGG	3' 3'	192	188
HA541	F: 5' HEX R: 5'	CAACACACGCTTTGTTGGGTG CGATAAGGAGGGAGGGGAAGATAG	3' 3'	229	213
ORS510	F: 5' 6-FAM R: 5'	CATCGCGTCCCTCTCTCTAA CCAACCATCACAGCAATCAG	3' 3'	249	258
ORS166	F: 5' 6-FAM R: 5'	CAGCCACATGCCCTCTGAC TGTTAAGAACCGCGACAACCTGC	3' 3'	327	339
ORS1146	F: 5' HEX R: 5'	GGCTCATCACTTGCATCTATTGT TGAAGACACCATCTCCAATGC	3' 3'	378	344
ORS767	F: 5' 6-FAM R: 5'	ACAGGAGCTTTCGGTCTTCA GGCACAGTCTCGTTCATATCC	3' 3'	371	369

Este sistema se evaluó mediante la co-amplificación de los fragmentos esperados utilizando el sistema PCR Type-it Microsatellite (Qiagen, Cat. N° 206243) modificado para las condiciones de trabajo establecidas en el laboratorio para amplificación de fragmentos de ADN de girasol. La reacción consistió en 0,2  $\mu$ M de cada iniciador disueltos en un volumen final de 10  $\mu$ l, conteniendo la "PCR Master Mix (Type-it Multiplex PCR Master Mix, que provee 3mM de  $MgCl_2$ )" y 50-100 ng de ADN.

Las condiciones consistieron en una desnaturalización inicial de 5 min a 95°C, 28 ciclos de (1) 30 seg a 95°C, (2) 90 seg a 60°C y (3) 30 seg a 72°C, con una extensión final de 30 min a 60°C.

De la misma manera que para los marcadores objeto de esta tesis, los productos obtenidos fueron teñidos con BrEt (0,3 µg/ml), evaluados en geles de agarosa al 3% (p/v), visualizados a través de un transiluminador UV (320 nm) y diluidos en 10 partes para ser analizados mediante EC.

Luego de diluir los productos obtenidos se procedió a analizar los electroferogramas resultados de la corrida electroforética capilar para PAC2 y RHA266 y corroborar los tamaños esperados para cada fragmento. Una vez determinado la eficacia del sistema de multiplexado, campaña tras campaña de multiplicación se procedió a realizar la verificación de identidad de las RILs mediante la co-amplificación de las líneas multiplicadas.

Los genotipos obtenidos luego del análisis de los electroferogramas con el programa GeneMapper, fueron acomodados en una tabla para realizar la comparación cruzada entre los individuos objetos de análisis y la base de genotipos original. Para ello se empleó la herramienta GenAlex v6.41) como fue descripto anteriormente.

En aquellos casos con perfiles heterocigotas o un patrón de SSRs diferentes al esperado, por comparación con la matriz de genotipos disponible, las semillas de una procedencia determinada fue descartada.

### **3. GENOTIPIFICACIÓN DE VARIANTES PRODUCIDAS POR MUTACIONES PUNTUALES.**

#### *3.1 GENOTIPIFICACIÓN DE SNPS MEDIANTE DHPLC*

##### **3.1.1 Selección de SNPs y diseño del ensayo de genotipificación**

Se seleccionaron 11 genes candidatos de un total de 31 involucrados en procesos de estrés biótico, los mismos fueron previamente caracterizadas en 19 RILs de girasol incluyendo los parentales de la población, PAC2 x RHA266, seleccionándose aquellas regiones candidatas que hayan resultado polimórficas para estas líneas en particular (Fusari et al. 2008) (Tabla 1.4).

Se rediseñaron oligonucleótidos para la amplificación de las 11 regiones con el programa Primer3 (<http://frodo.wi.mit.edu/>), de acuerdo a los requerimientos de la técnica de detección de heterodúplex (regiones entre 200 y 1000 pb).

La Tabla 1.5 describe los oligonucleótidos iniciadores utilizados para amplificar las regiones candidatas.

##### **3.1.2 Amplificación**

La amplificación de los productos mediante reacción en cadena de la polimerasa (PCR) se realizó en el termociclador Eppendorf (Mastercycler epgradient S) en 50 µl de volumen final usando entre 80-100 ng de ADN genómico.

TABLA 1.4. Descripción de los genes candidatos seleccionados				
GEN	SIMILITUD(BLASTx)	DESCRIPCIÓN	SNPs	INDELS (pb)
PGIP1	Proteína inhibidora de Poligalacturonasa ( <i>Vitis vinifera</i> ) Precursor de la proteína inhibidora de Poligalacturonasa ( <i>Actidina deliciosa</i> )	Defensa de la planta contra diversos patógenos que usan para romper la pared celular (Bishop 2005).	1	0
PGIP3			3	0 (0)
MADSB-TF3	Factor de transcripción MADS-BOX ( <i>Helianthus annuus</i> )	Factor de transcripción que actúa como regulador de varios aspectos del desarrollo de la planta (Dezar et al. 2003).	1	0 (0)
LIM	Dominio LIM de la proteína PLIM I b ( <i>Helianthus annuus</i> )	Factor de transcripción que juega un importante rol en la construcción del citoesqueleto y en la transducción de señales (Arnaud et al. 2007)	5	0 (0)
1ACCO1	1-aminociclopropano-1-ácido carboxílico oxidasa ( <i>Helianthus annuus</i> )	Cataliza la conversión de 1-aminociclopropano-1-ácido carboxílico a etileno, asociado al crecimiento de la planta, ya que produce la rápida división de las células (Liu et al. 1997)	1	0 (0)
ANT	Translocador de nucleótidos de adenina, precursor mitocondrial ( <i>Gossypium hirsutum</i> )	Transportador de membrana mitocondrial que participa en la integración del estrés celular y regulación la muerte celular programada (Lam et al. 2001).	9	0 (0)
GCvT	Proteína T del sistema de clivaje de glicina ( <i>Flaveria trinervia</i> )	El sistema de clivaje de glicina cataliza decarboxilación oxidativa de glicina en bacterias como en mitocondrias animales y vegetales (Lee et al. 2004).	2	0 (0)
CAM	Calmodulina ( <i>Morus nigra</i> )	Posee una función central en la señal mediada por calcio (Liu & Burke 2006).	17	1 (4)
GPX	Putativa peroxidasa glutatiónica ( <i>Medicago truncatula</i> )	Enzima antioxidante, factor importante en la protección contra daños oxidativos (Liu & Burke 2006).  Cataboliza la isomerización reversible de 6-fosfoglucosa y 6-fosfofructosa, una reacción esencial que precede a la biosíntesis de sacarosa (Liu & Burke 2006).	0	1 (6)
PGIC	Región acotada de Fosfoglucosa isomerasa citosólica ( <i>Stephanomeria tenuifolia</i> )		2	1 (1)
SCR1	Factor de transcripción tipo 1 ( <i>Castanea sativa</i> )		Genes reguladores tipo "Scarecrow", involucrado en la división asimétrica de la planta (Liu & Burke 2006).	2

En ambos casos, la mezcla de reacción se compuso de 2 mM MgCl<sub>2</sub>, 0,2 mM dNTPs, 1 U Polimerasa Taq Platinum (Invitrogen) y 0,25 mM de cada uno de los oligonucleótidos iniciadores específicos. La reacción de amplificación fue llevada a cabo mediante touch down de acuerdo al rango de temperatura óptima de apareamiento calculada para cada iniciador.

Las condiciones generales de ciclado fueron:

- un paso de 5 min a 94°C de desnaturalización inicial, 5 ciclos de (1) 60 seg a 94°C, (2) 60 seg a 60°C (-1°C/ciclo, touchdown 60-55) y (3) 90 seg a 72°C, seguido por 30 ciclos de

TABLA 1.5. Oligonucleótidos iniciadores para la evaluación de marcadores SNPs seleccionados				
NOMBRE	OLIGONUCLEÓTIDOS INICIADORES			
PGIP1	F:	5'	TTTGCCAGAGACGTTGACAG	3'
	R:	5'	CACCAAAGTTATATTCTTTTGTAAATG	3'
PGIP3	F:	5'	AAACTCACCGGCTCTATTCC	3'
	R:	5'	GACCAATCGTCAAACATAACT	3'
MADSB-TF3	F:	5'	TTGTTATCTATGGCTCGTCTGG	3'
	R:	5'	TGATACCGATCCAGCATGTC	3'
LIM	F:	5'	TCAAGCCAGAAAGAACATTT	3'
	R:	5'	AAGCTCTGTGGTATGCAGTT	3'
1ACCO-1	F:	5'	GCCGAAGTTACCGATATGGA	3'
	R:	5'	CCTTCATCAACTCCCTGCAT	3'
ANT	F:	5'	TTGCTGGTGCTTCTTCTCTAGTT	3'
	R:	5'	GGTACAAACCACGGTAGATGAAA	3'
GCvT	F:	5'	TCTTCTTCAAGCACGGACTAAAC	3'
	R:	5'	CAACACATAACCCCTGTGGAG	3'
CAM	F:	5'	AGAAACTCTGGGAAGTCGATTG	3'
	R:	5'	TCAGCGCAAATTACCCAAAT	3'
GPX	F:	5'	GACTCATTACCATCATCTTCTCTC	3'
	R:	5'	CTTCTGCACCAACATCACAATAGTA	3'
PGIC	F:	5'	GTCTTATCCGTTGACATTTTTCAGG	3'
	R:	5'	AGAGCTTGAGTGTAAAGGCAGGATAG	3'
SCR1	F:	5'	GGAATCCTGTCTGCTGATAAGT	3'
	R:	5'	TTCACTTGCGAAACAAGCTC	3'

- (1) 60 seg a 94°C, (2) 60 seg a 55°C y (3) 90 seg a 72°C, con una extensión final de 10 min a 72 °C. Los productos amplificados se visualizaron con luz UV (320 nm) luego
- de la electroforesis en geles de agarosa 1,5% (p/v) teñidos con BrEt (0,3 µg/ml) empleándose un marcador 1 Kb Ladder (Invitrogen) para estimar la longitud de los fragmentos según las recomendaciones del fabricante.

De la misma manera que para la cuantificación de ADN descrita anteriormente, se procedió a determinar la concentración de los fragmentos de ADN amplificados para detección de los SNPs.

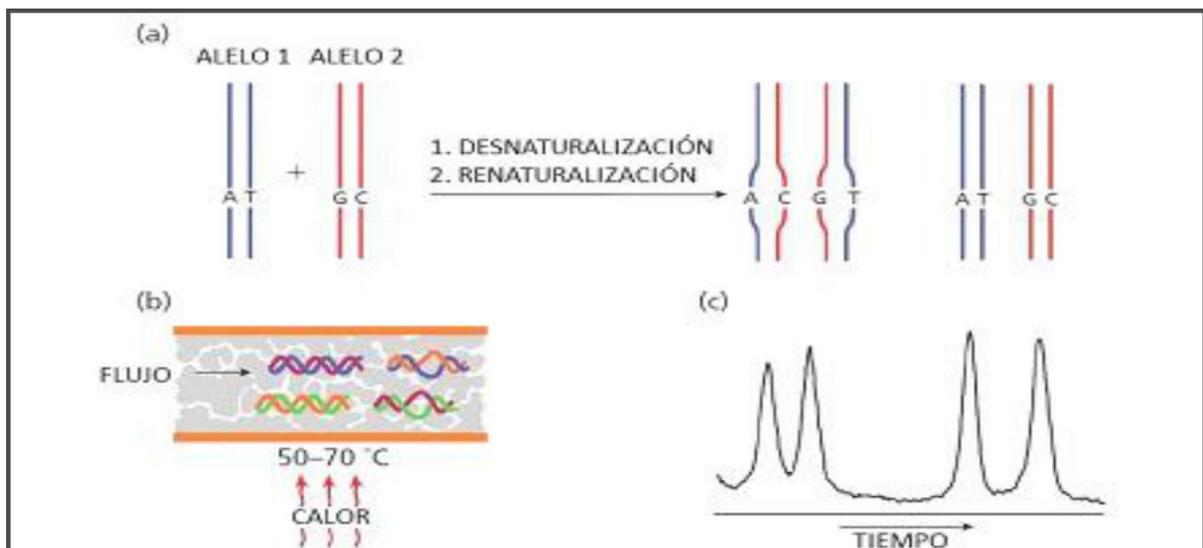
### 3.1.3 Genotipificación

Este método de genotipificación utilizado para la identificación de SNPs se basó en la resolución de moléculas heterodúplex mediante la técnica de HPLC en condiciones desnaturalizantes o dHPLC.

La detección de SNPs mediante esta estrategia implica la comparación de dos o más alelos de un locus dado (Figura 1.2), cuyos productos de amplificación específicos se mezclan y someten a un proceso de desnaturalización y lenta re-naturalización de las hebras de ADN. A partir de este proceso, por apareamiento aleatorio de las hebras complementarias se producen dos tipos de moléculas: las moléculas homodúplex, que son idénticas a las moléculas originales, y las moléculas heterodúplex, que resultan del apareamiento de una hebra proveniente del alelo A y la complementaria del alelo B, y viceversa. Estas últimas, poseen apareamiento imperfecto en los sitios variables (SNPs o indels).

El fundamento de este método es la retención diferencial de las moléculas homo y heterodúplex en su paso por una columna iónica de fase reversa en condiciones de desnaturalización parcial (Xiao & Oefner 2001). En estas condiciones las moléculas heterodúplex interaccionan con la columna más débilmente y presentan una elusión más rápida que las moléculas homodúplex.

Dado que tanto los parentales como la población analizadas en este trabajo son líneas altamente homocigotas, es necesario crear muestras heterocigotas para poder discriminar entre especies de homo y heterodúplex resultantes de la reasociación de los alelos en los procesos de desnaturalización y renaturalización. Por lo tanto, los productos de PCR de los parentales PAC2 y RHA266 se mezclaron entre sí (control positivo del HETERODÚPLEX) y cada una de las diferentes RILs con RHA266 para realizar el genotipado mediante la técnica de dHPLC. Además se realizaron los controles HOMODÚPLEX que contenían solamente cada una de las líneas por separado (tanto parentales como RILs).



**FIGURA 1.2. ESQUEMA REPRESENTATIVO DE LOS PASOS INVOLUCRADOS EN LA GENOTIPIFICACIÓN MEDIANTE LA TÉCNICA DHPLC.** Amplificación por PCR de los alelos seguida de la mezcla y desnaturalización-renaturalización para la formación de las especies Homodúplex y Heterodúplex(a), inyección de la mezcla en la columna a temperatura y gradiente específicos (b) y observación de la salida de los picos mediante un cromatograma (c). (Fuente: Modificado de Oefner (2000)).

Las muestras se sometieron a desnaturalización y renaturalización lenta mediante modificación gradual de la temperatura en un termociclador Eppendorf (Mastercycler epgradient S) corriendo el siguiente programa: 2 min a 95 °C, 95 °C descendiendo hasta 85 °C a razón de - 2°C/seg, 85 °C descendiendo hasta 25 °C a -0,1 °C/seg, 4 °C pausa. La rampa de enfriamiento se fijó en 45% en el primer paso descendiente y al 2% en el segundo.

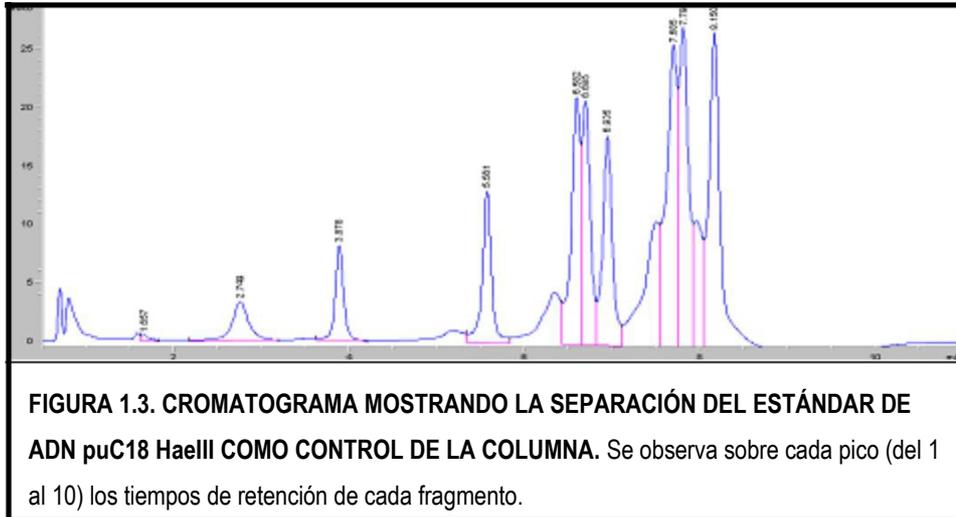
Para cada locus, las condiciones de temperatura y gradiente óptimas para la detección inequívoca de las especies homo y heterodúplex se encontraron evaluando las condiciones sugeridas *in silico* mediante el programa dHPLC Melt Program (<http://insertion.stanford.edu/melt.html>).

La temperatura y el gradiente óptimos para lograr una buena separación difirieron para cada locus ensayado de acuerdo a la composición nucleotídica de cada uno, pero todos fueron optimizados ajustando la temperatura entre los 54°C y 61°C.

Las corridas se realizaron en el sistema de HPLC 1100 Agilent Technologies, que comprende una unidad de desgasificación de solvente, una bomba binaria, un módulo de toma de muestra automático con control de temperatura (configurado a 4 °C durante la corrida de las muestras), un horno para la columna y un detector de longitud de onda variable. Las mezclas de los productos de PCR se inyectaron en la columna para ADN Varian Helix™ (3X50 mm; catálogo no. CP28353, Varian Inc.). La cromatografía se realizó a un flujo de 0,9 ml/min con un gradiente lineal de las soluciones Varian A y B (Varian Inc., nos. de catálogo 0393558101 y 0393558102, respectivamente) y un volumen de inyección entre 5 y 10 µl (aproximadamente, 100 ng de ADN total). Luego de 5 minutos de corrida, las muestras se detectaron a una longitud de onda de 260 nm. Antes de iniciar una serie de corridas y regularmente cada 50 muestras se incluyó una corrida de control de la técnica mediante la inyección del plásmido control pUC18 digerido con la enzima HaeIII (Varian Inc., catálogo no. CP28353), usando el método sugerido por el fabricante para asegurar la efectividad y reproducibilidad de la cromatografía. La Tabla 1.6 muestra las condiciones de corrida para la detección del plásmido control.

TABLA 1.6. Condiciones de corrida para la detección de fragmentos de plásmido puc18.			
TIEMPO (min)	% SOLUCIÓN B	FLUJO (ml/min)	PRESIÓN MÁXIMA (bar)
0	45	0,45	200
0,5	50	0,45	200
6	68	0,45	200
7	68	0,45	200
7,01	45	0,45	200
8,5	45	0,45	200
Temperatura: 50 °C			
Tiempo total de corrida: 11 minutos			

El cromatograma del control muestra el patrón correcto cuando se observa la separación de los picos 5 (257 pb) y 6 (267 pb), y de los picos 8 (434 pb) y 9 (458 pb), tal como se observa en la Figura 1.3.



Pudieron establecerse correctamente las condiciones de detección de heterodúplex y homodúplex para las siguientes regiones: CAM, SCR1, PGIC, LIM (Tabla 1.7).

Un conjunto de 93 RILs del total de 135 fueron seleccionadas y sus productos de PCR de para cada gen candidato fueron mezclados equimolecularmente con el parental RHA266, siguiendo luego la desnaturalización y renaturalización lenta. El análisis entonces consistió en analizar los cromatogramas de las mezclas RIL-RHA266. De este modo, cuando se observó un cromatograma similar al de la mezcla de los parentales, se asumió que dicha RIL posee el alelo correspondiente a PAC2. Por el contrario, la visualización de un pico similar al patrón de cualquiera de los homodúplex indicó que dicha línea posee el mismo alelo que RHA266 para la región correspondiente.

<b>TABLA 1.7. Condiciones de corrida para fragmentos puestos a punto. * a flujo y presión máxima constantes (0,9 ml/min y 200 bar)</b>				
SNP	TIEMPO (MIN)*			
	0	0,5	3	3,01
CAM (54 °C) % SOLUCIÓN	50	65,3	70,7	50
SCR1 (59 °C) % SOLUCIÓN	50	64,2	68,6	50
PGIC (55 °C) % SOLUCIÓN	50	54,9	60,3	50
LIM (57 °C) % SOLUCIÓN	50	64,2	68,6	50

Para descartar errores en los productos de amplificación de los genes candidatos para los que no se pudieron establecer las condiciones óptimas de detección de la especie heterodúplex y homodúplex con los parentales de la población (1ACCO-1, PGIP1, PGIP3, MADSB-TF3, ANT, GCvT y GPX), sus productos de amplificación fueron secuenciados en las líneas PAC2 y RHA266 y analizados mediante el programa Bioedit.

### 3.2 *DESARROLLO DE UN PANEL DE 384 SNPs PARA LA GENOTIPIFICACIÓN MASIVA UTILIZANDO LA PLATAFORMA VERACODE-BEADXPRESS DE ILLUMINA*

#### 3.2.1 Búsqueda y selección de SNPs para el ensayo de genotipificación “Sunflower Oligo Pool Assay (SOPA)”

La selección de SNPs para el diseño de un ensayo de genotipificación múltiple mediante la tecnología GoldenGate (Illumina) se basó en la selección de SNPs candidato a partir de dos fuentes de información.

En primer lugar, se realizó una búsqueda *in silico* de SNPs mediante el análisis de la base de datos “Sunflower Unigene Repository (SUR)” de 41013 unigenes desarrollados por Fernandez et al (2012) a partir del ensamblado de 133682 secuencias de EST provenientes de diferentes genotipos de girasol cultivado depositadas en GenBank (<http://www.ncbi.nlm.nih.gov/nucest>, revisión de diciembre de 2008).

Para llevar a cabo la identificación de SNPs sobre la base de datos de unigenes de girasol, se trabajó con el archivo de salida “.ace” del programa CAP3. Se analizaron los alineamientos de EST para cada uno de los contigs que conforman la base de datos. Para definir un SNP mediante esta metodología se establecieron como parámetros restrictivos que cada SNP este respaldado por:

- Mínimo de cobertura de cinco lecturas en la posición del SNP;
- Posiciones bialélicas con una representación de al menos 5%.

En segundo lugar, se seleccionaron un total de 119 genes candidatos (GCs) que fueron obtenidos a partir de bibliografía: un conjunto de 24 GCs relacionados a estrés biótico y abiótico fueron seleccionados a partir de un estudio de resecuenciación de más de 19 genotipos representativos del germoplasma cultivado girasol (Fusari et al. 2008); un segundo panel de 14 GCs relacionados con respuesta a estrés fueron seleccionados a partir de un estudio de mapeo de asociación (Fusari et al. 2012), mientras que los 81 GCs restantes fueron seleccionados a partir del estudio de diversidad nucleotídica en girasol publicado por Kolkman et al (Kolkman et al. 2007).

Se empleó la nomenclatura “SNPs *in silico*” y “SNPs *in vitro* acuñada por (Lepoittevin et al. 2010)” para aquellos SNPs descubiertos a partir de la búsqueda sobre ESTs o sobre genes candidatos resecuenciados, respectivamente.

Todas estas regiones genómicas seleccionadas fueron evaluadas mediante la aplicación secuencial de niveles estrictos de filtrado SNPs a fin de aumentar el éxito de genotipado considerando los consejos y requerimientos del proveedor Illumina para la síntesis de paneles de genotipificación adecuados para el sistema GoldenGate-Veracode (illumina). Estos son:

- la posición del SNPs debe estar al menos a 100 pb de distancia de los extremos de la secuencia;
- el contenido de GC de la secuencia consenso debe estar comprendida entre 40-60%, y carecer de elementos repetitivos en las secuencias flanqueantes;
- la separación mínima entre SNPs debe ser mayor a 60 nucleótidos

A partir de estos criterios se seleccionó un subconjunto de secuencias que contenían SNPs para ser evaluadas posteriormente utilizando el programa Assay Design Tool (ADT, [http://www.illumina.com/downloads/Illumina\\_Assay\\_Design\\_Tool.pdf](http://www.illumina.com/downloads/Illumina_Assay_Design_Tool.pdf)). Para el diseño del Oligo Pool Assay se consideraron aquellos SNPs cuyo valor de diseñabilidad fue  $\geq 0,6$ , valor de corte que establece Illumina para obtener un panel de marcadores con probabilidad alta de ser exitosos para la tecnología GoldenGate/Vera-Code. Finalmente, para la selección final de los 384 SNPs se consideró la anotación funcional establecida para los contigs, seleccionándose aquellos asociados a funciones de resistencia a estreses.

### 3.2.2 Cuantificación de ADN

Cada corrida del Sunflower Oligo Pool Assay (SOPA) fue realizada siguiendo el protocolo establecido por el fabricante. En primera instancia, la cuantificación del ADN fue hecha mediante espectrofluorometría con el equipo Nanodrop ND 3300 ([NanoDrop Technologies](#)). Para ello se utilizó el reactivo Picogreen (Intercalante de ADN, dsDNA reagent kit, Molecular Probes cat # P-11496) de acuerdo al siguiente protocolo. Se construyó una curva de calibración 2x empleando ADN genómico de timo de cabra en concentraciones de 0, 2, 10, 20, 100, 200, 400, 1000 y 2000 ng/ml de ADN. A cada una de las muestras incógnitas (y puntos de la curva) se le agregó un volumen de solución de trabajo Picogreen-TE (Picogreen 1:200 y TNE 1X: Tris 10 mM, EDTA 1 mM, pH 7,5). La lectura fue obtenida utilizando el método “Picogreen”, los datos de los estándares se ajustaron a una recta y las incógnitas se determinaron por interpolación.

### 3.2.3 Genotipificación

La corrida de evaluación de SNP se realizó utilizando el equipo lector de microarrays BeadXpress (Illumina). El SOPA desarrollado a partir de la tecnología Illumina, consta de un conjunto de oligonucleótidos que incluyen dos oligos alelo-específico (ASO, Allele-specific oligonucleotides) y un oligonucleótido locus específico (LSO, Locus-specific oligonucleotides, Figura 1.4). El proceso de trabajo incluye una etapa de hibridación de los oligos ASO y LSO con la secuencia blanco, seguido de un ciclo de extensión y ligación y posteriormente una amplificación con iniciadores marcados permite detectar la variante de cada uno de los locus analizados. Para ello, la tecnología Veracode usa perlas paramagnéticas de forma cilíndrica con una etiqueta holográfica que permite identificar el soporte que corresponde a cada SNP. Cada perla esta a su vez revestida de oligonucleótidos de secuencia única denominada “illumi code” cuya secuencia complementaria es parte del LSO y constituye la etiqueta que permite reconocer cada locus específicamente.

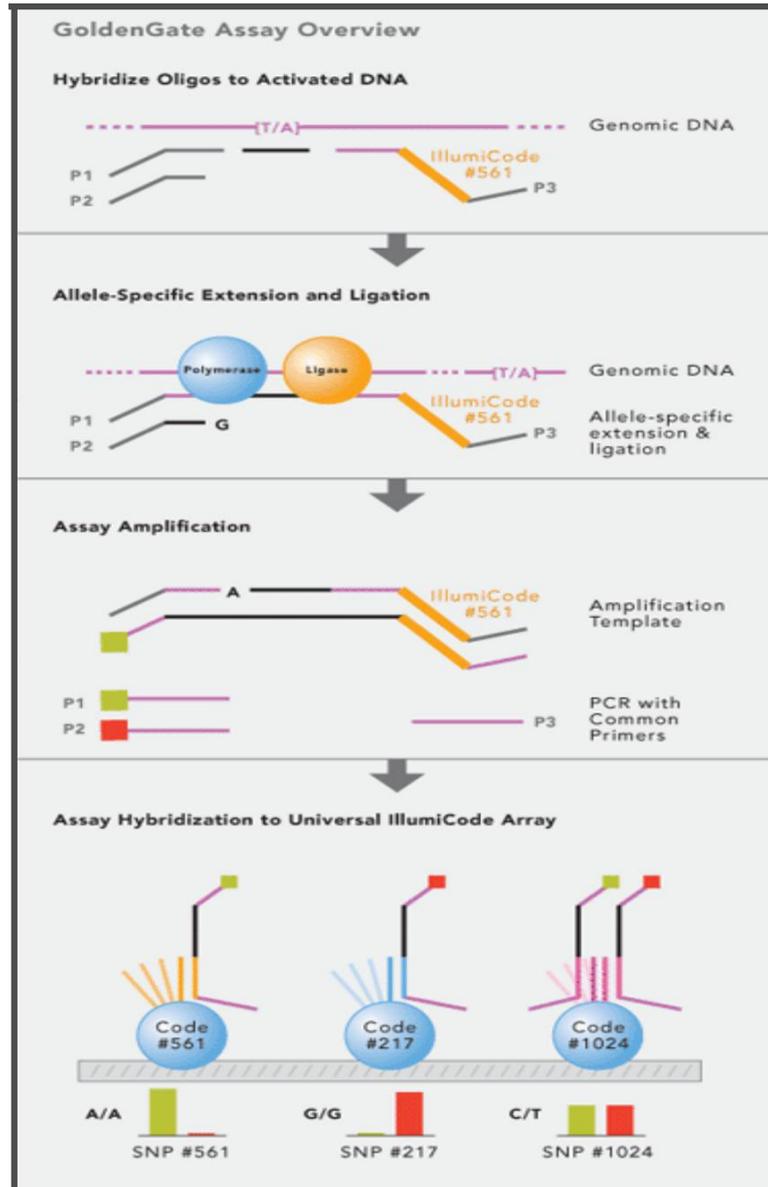
Durante el escaneo con el equipo BeadXPress se miden las intensidades de los fluoróforos para cada SNP, que a su vez está representado al menos treinta veces en el ensayo.

En la práctica, se utilizaron dos placas de 96 pocillos con 5 µl de templado cada uno, en una concentración de 50 ng/µL. Un total de 159 individuos comprendiendo las 135 RILs que componen la población de mapeo PAC2 x RHA266 y 24 muestras replicadas para evaluar la reproducibilidad del ensayo.

Se siguió el protocolo provisto por Illumina para llevar adelante la genotipificación. El mismo está dividido en 2 partes denominadas “pre-PCR” y “post-PCR”.

Durante la pre-PCR, el ADN genómico cuantificado y normalizado para cada muestra, es activado con biotina durante 30 minutos. El próximo paso involucra la hibridación alelo específica que combina en una placa el ADN genómico biotinilado con oligos interrogantes (query), reactivos de hibridación, y partículas o perlas paramagnéticas (beads). Esta placa se coloca en el bloque térmico durante 2 horas, donde los oligos se unen a las secuencias de interés de cada muestra y simultáneamente el ADN es capturado por las perlas.

La ligación y extensión enzimática de los oligos a 45°C se realiza mediante el agregado una “Master Mix” provista por Illumina e incubación por 15 minutos. Se agrega la ADN polimerasa recomendada (Taq Titanium, Clontech) a una master mix para PCR (provista por Illumina) y se las alícuota en cada placa de 96. Luego, se utiliza el producto de la extensión y ligación como templado para la PCR. Esta PCR utiliza 3 oligonucleótidos universales: 2 marcados con fluorocromos y 1 biotinilado. Este último permite capturar los productos de PCR y eluir las hebras que contienen las señales fluorescentes. Las muestras eluidas son transferidas a la placa



**FIGURA 1.4. ESQUEMA DEL ENSAYO GOLDENGATE/VERACODE (ILLUMINA).** De arriba hacia abajo, hibridación del oligonucleótido alelospecifico (ASO) y el oligonucleótido locus específico (LSO) con el ADN incógnita. Extensión y ligación del fragmento. Amplificación con tripletes de iniciadores universales marcados y reconocimiento de la señal mediante la captura de los fragmentos amplificados por parte de perlas con secuencias complementarias al LSO llamadas "illumi code". (Tomado de Illumina).

de PCR. En este proceso de ciclado de PCR se amplifican y marcan fluorescentemente los templados generados en el proceso de pre-PCR.

Las condiciones generales de ciclado fueron: una fase previa de 10 minutos a 37°C seguida por una desnaturalización inicial de 3 min a 95°C, 34 ciclos de (1) 35 seg a 95°C, (2) 35 seg a 56°C y (3) 2 minutos a 72°C, con una extensión final 10 min a 72°C y 5 minutos a 4 grados.

Durante la post-PCR, se agrega a la placa a de PCR el reactivo que contiene las perlas paramagnéticas y la mezcla se transfiere a una placa con filtro de 0,45  $\mu\text{m}$ . Esta última se incuba a temperatura ambiente durante 1 hora para que la hebra biotinilada se una a las perlas paramagnéticas, lo cual inmoviliza los productos de PCR de doble cadena.

Luego de filtrarse, las muestras son transferidas a una placa "VeraCode Bead (VBP)" que contienen las microbolillas de reconocimiento de SNPs y están listas para ser hibridadas a 45°C durante 45 minutos. Para ello se usa un incubador de placas con agitador (VeraCode Vortex Incubator).

Luego de realizar el lavado correspondiente, se procede a escanear la placa. El BeadXpress Reader utiliza láser para excitar los fluoróforos Cy3 y Cy5 de los productos de PCR de cadena simple unido a las perlas paramagnéticas. Luz las emisiones de estos compuestos fluorescentes son luego registrados en un archivo de datos. Los datos de fluorescencia se analizan para obtener resultados de genotipificación de Illumina utilizando el paquete software GenomeStudio (Illumina) Modulo "Genotyping".

Para el análisis de los datos de genotipificación obtenidos, es necesario crear primero tres archivos para el ingreso y reconocimiento de la muestra por parte del programa GenomeStudio. El primero, creado por el usuario, es conocido como "Sample sheet", describe la localización de cada una de las RILs en los pocillos de la placa. El segundo corresponde a cada una de las intensidades y colores de fluorescencia obtenidos para los SNPs del ensayo y el tercero es un archivo llamado "manifest", que se utiliza para vincular los dos anteriores. El "manifest" es propio de cada ensayo y contiene información de los 384 SNPs incluidos en el ensayo (como por ejemplo, las variantes alélicas esperadas para cada SNP).

Siguiendo las recomendaciones de Illumina, se consideran un valor de calidad denominado "call rate", definido como la tasa de individuos cuyos alelos en cuestión fueron correctamente clusterizados para cada SNP. Este valor fue determinado en  $>0,70$ .

El software agrupó las muestras en base a la relación de intensidad cy3/cy5 definiendo la pertenencia de cada SNP a una de tres categorías de genotipos. Por defecto, las muestras que caen dentro del área sombreada de rojo son nombradas "AA"; las que caen dentro del área violeta "AB" y las que caen dentro del área azul "BB". Al tratarse de RILs, muy poco o ninguna muestra debería ser heterocigota, lo cual se reduce a 2 áreas (roja y azul).

Cada RIL para el SNP seleccionado tiene un valor "GenCall" asociado a una medida de calidad que indica la confiabilidad de la asignación del alelo. Estos valores pueden variar entre 0 y 1. Los genotipos con valores bajos, son localizados lejos del centro del grupo.

Los valores de GenCall son calculados en base a la información obtenida de la agrupación de las muestras. Para obtener dicho valor, cada SNP es evaluado de acuerdo a las siguientes características de los clústeres:

- Ángulo;
- Dispersión;
- Superposición;
- Intensidad.

No existe una interpretación global del valor de GenCall, ya que éste depende de la agrupación de las muestras para cada SNP, que es afectada por muchas variables, incluyendo la calidad de las muestras.

El tamaño del área sombreada viene dado por un umbral de corte determinado manualmente en 0,25 según recomendaciones de estudios anteriores (Lepoittevin et al. 2010; Roorkiwal et al. 2013). Fuera de dicho área, no se asigna alelo a la muestra.

Cada área posee un ovalo, que representa a 2 desviaciones estándares. Por otro lado, se ordenaron los SNPs en base a un estadístico que evalúa la distancia entre clústeres como también su forma y calidad de fluorescencia (valor llamado "GenTrain Score"). Aquellos SNPs que no resolvieron correctamente la agrupación de clústeres y poseían un GenTrain Score <0,40 fueron descartados.

Los clústeres fueron inspeccionados visualmente para asegurar la alta calidad de los datos obtenidos, y cada SNP se volvió a comprobar manualmente. Los datos correspondientes a los genotipos obtenidos fueron exportados en forma de matriz y cada RIL recibió AA, AB o BB según si fue caracterizado homocigota para un alelo, heterocigota u homocigota para el otro alelo.

Además, el número de SNPs polimórficos y monomórficos, así como el número de SNP fallidos, se contabilizaron con el fin de predecir la tasa de éxito del ensayo de genotipificación (medida como la proporción de marcadores polimórficos sobre el total de marcadores ensayados). Una segunda medida calculada del éxito global del ensayo fue la tasa de conversión, definida por Fan et al (2003) como el número de SNPs polimórficos dividido por el número total de SNPs en el ensayo.

Mediante una prueba de  $\chi^2$  se evaluó el posible efecto del origen in silico o in vitro de los SPNPs sobre la tasa éxito y de conversión.

### 3.3 GENERACIÓN DE MATRICES

Luego de realizar la genotipificación de los marcadores SNPs, cada uno de los mismos se codificó en una matriz respetando el siguiente esquema:

- LOS MARCADORES QUE POSEÍAN EL MISMO PERFIL QUE PAC2, FUERON ANOTADOS COMO “1”;
- LOS MARCADORES QUE POSEÍAN EL MISMO PERFIL QUE RHA266, FUERON ANOTADOS COMO “2”;
- LOS MARCADORES QUE NO POSEÍAN PERFILES DEBIDO A FALLAS EN LA GENOTIPIFICACIÓN O ALELOS NULOS, FUERON ANOTADOS COMO “0”.

## 4. MAPEO DE LIGAMIENTO

LA MATRIZ GENERADA EN PASOS PREVIOS, CONSTÓ DE 135 INDIVIDUOS GENOTIPADOS Y UN TOTAL DE 735 MARCADORES. ADICIONALMENTE, SE INCORPORARON A ESTA MATRIZ, LOS GENOTIPOS OBTENIDOS PARA LAS RILs DE 4 genes candidatos asociados a resistencia a enfermedades producidas por *Sclerotinia* y que son objeto de estudios en trabajos complementarios a esta tesis (Flippi, Tesis en preparación, Fusari et al. 2008; Tabla 1.8).

TABLA 1.8. Descripción de los genes candidatos incorporados al mapa de ligamiento		
GEN	SIMILITUD(BLASTx)	DESCRIPCIÓN
RhoBP_B	Proteína con dominio p21-rho-binding	Función desconocida
GLP_3	Proteína tipo germina	Proteína apoplástica y glicosilada involucrada en defensa (Zimmermann et al. 2006)
GLP_4		
GLP_5		

### 4.1 ANÁLISIS DE LA SEGREGACIÓN Y CONSTRUCCIÓN DEL MAPA GENÉTICO

Los datos genotípicos provenientes de 135 RILs seleccionadas se utilizaron para producir un mapa genético formado por registros de alelos de SNPs más AFLPs y SSRs previamente mapeados por otros autores (Al-Chaarani et al. 2004; S. Kiani et al. 2007; Paola Talia et al. 2010; Zubrzycki 2009). Además, la inclusión de estos marcadores anclados en distintos mapas genéticos de la especie permite la correcta identificación de cada uno de los grupos de ligamiento (GLs) y determina su respectiva polaridad.

Los datos correspondientes a 739 marcadores polimórficos para los individuos que integran la población fueron manualmente analizados y aquellos marcadores y líneas que presentaron más de 50% de datos perdidos fueron descartados. De esta manera, la matriz final se utilizó para evaluar la segregación mendeliana mediante la técnica FDR (False Discovery Rate, Benjamín y

Hocheberg, (Benjamini & Hochberg 1995)) para  $P \leq 5\%$ , implementada en el programa GQmol (Schuster & Cruz 2004), disponible en <http://www.ufv.br/dbg/gqmol/gqmol.htm>). El método FDR propone controlar la razón de los falsos descubrimientos, siendo FDR la proporción de  $H_0$  verdaderas entre las  $H_0$  rechazadas; que es la proporción de falsos negativos (Q) (Benjamini & Yekutieli 2005). Solo aquellos marcadores que mostraron la segregación mendeliana esperada de 1:1 integraron los análisis ulteriores.

En este punto se tomó dos estrategias de mapeo con la finalidad realizar un contraste y validación de los mapas generados.

En primera instancia, se consideró la matriz completa de marcadores para llevar a cabo la construcción del mapa utilizando la versión de prueba del programa JoinMap v4.1 (Van Ooijen & Voorrips 2001).

En segunda instancia, se subdividió la matriz original en pequeñas matrices, cada una correspondiente a un GL. La misma contuvo a los marcadores previamente mapeados e informados en publicaciones anteriores (Al-Chaarani et al. 2004; Paola Talia et al. 2010). El orden de los mismos se utilizó como referencias inalterables para el mapeo de marcadores que no pudieron ser ubicados en trabajos previos de mapeo, y de todos los desarrollados en el presente trabajo de Tesis.

Este mapeo se llevó a cabo considerando un valor mínimo de  $LOD=4,0$ , una fracción máxima recombinación  $> 0,35$  (S. Kiani et al. 2007) estimada mediante máxima verosimilitud y transformada en distancia genética, expresada en cM, a través de la función de Kosambi (1943). La frecuencia de recombinación en la generación  $F_2$  se obtuvo corrigiendo la estimada en la población de RILs por el factor que considera múltiples generaciones de meiosis en autofecundación (Haldane & Waddington 1931)  $r = R/(2-2R)$ , donde  $R = nr/(nr + nn)$  es la frecuencia de recombinación no corregida y  $nr$  y  $nn$  es el número de genotipos recombinantes y no recombinantes, respectivamente (Tang et al. 2002).

Para el ordenamiento de los marcadores se utilizó la función de agrupamiento “Independencia de LOD”, que no se ve muy afectado por la segregación distorsionada de marcadores, lo que ocasiona menos ligamientos espurios.

Cada GL fue nombrado consecutivamente del 1 al 17 de manera correlacionada con los mapas públicos, correspondiéndose con cada uno de los 17 cromosomas del genoma haploide de girasol ( $x=17$ )

Las frecuencias de recombinaciones estimadas en las poblaciones de mapeo RILs fueron corregidas por el factor que considera múltiples generaciones de meiosis en autofecundación (Haldane & Waddington 1931)  $r = R/(2-2R)$ , donde  $R = nr/(nr + nn)$  es la frecuencia de

recombinación no corregida y nr y nn es el número de genotipos recombinantes y no recombinantes, respectivamente (Tang et al. 2002).

Para la representación gráfica de los diferentes GLs se utilizó el programa Mapchart 2,2 (Voorrips 2002).

# RESULTADOS

## 1. ESTRUCTURACIÓN DE POBLACIÓN DE RILs PARA MAPEO DE QTL PARA RESISTENCIA A PHC

Con el objeto de corroborar/mantener la identidad de las RILs durante las sucesivas campañas de multiplicación y fenotipificación, en cada línea se amplificaron 39 SSRs representativos de los 17 GL del mapa de girasol.

Del total de RIL, 14 fueron descartadas por poseer un perfil heterocigota para 3 y 7 marcadores, La Figura 1.5 muestra un esquema de los pasos llevados a cabo.

El análisis permitió ratificar que 68/80 individuos poseían el genotipo esperado al contrastarlos con la referencia de los genotipos de la matriz original. Las restantes 12 RILs no alcanzaron al 100% de identidad respecto de dicha referencia.

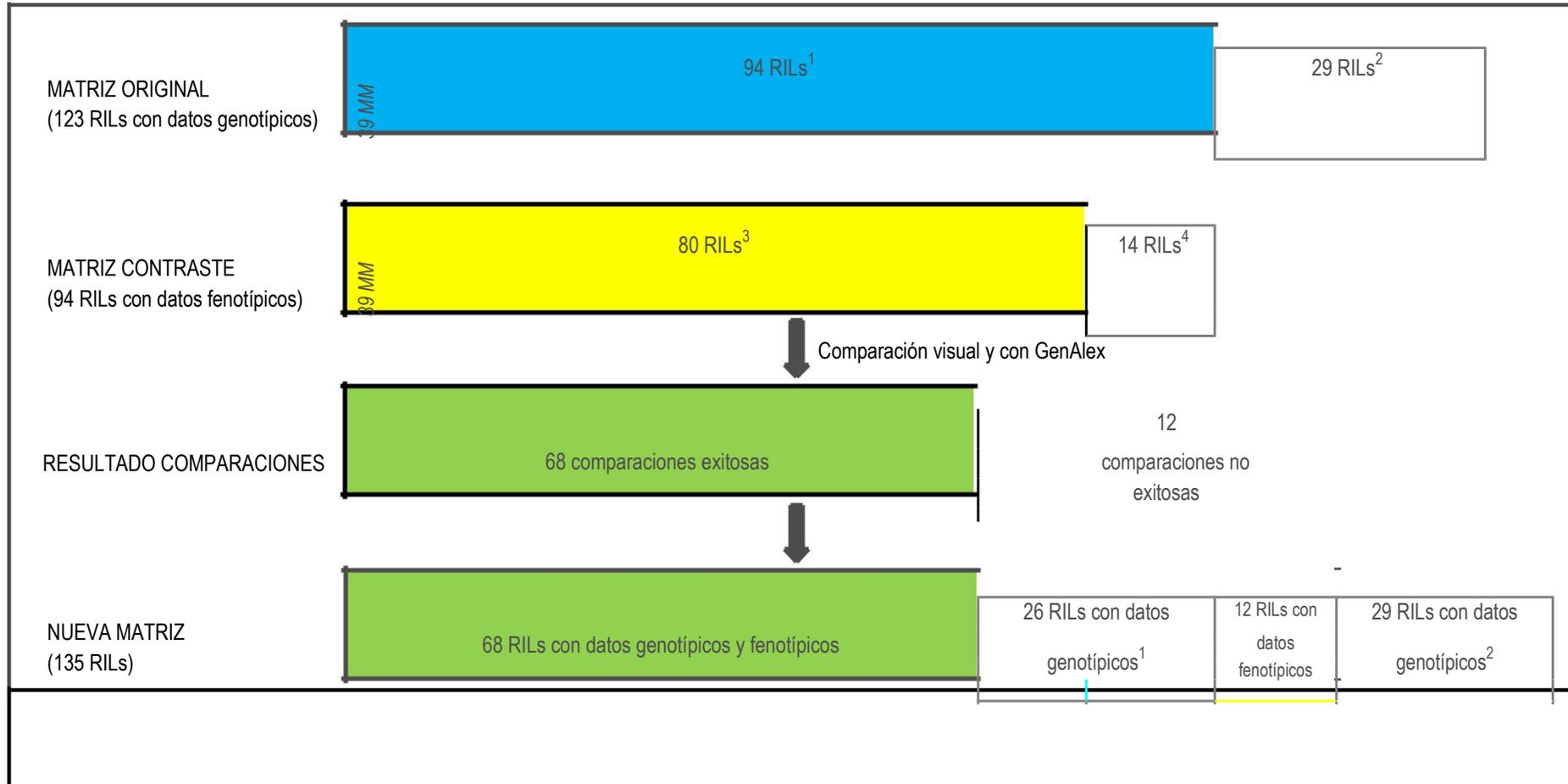
A partir de estos datos se reestructuró la población a una nueva versión conteniendo 135 individuos: 123 originales (68 con datos genotípicos y fenotípicos, 26 con datos genotípicos desarrollados por Talia et al (2010) y 29 con datos genotípicos originales (Al-Chaarani et al. 2004) + 12 con solamente datos fenotípicos.

Las 26 líneas que contenían solamente datos genotípicos, fueron genotipificadas en las próximas campañas de evaluación. Para el caso contrario, donde las 12 líneas solamente poseían datos fenotípicos, se procedió a genotipificarlas con los marcadores existentes.

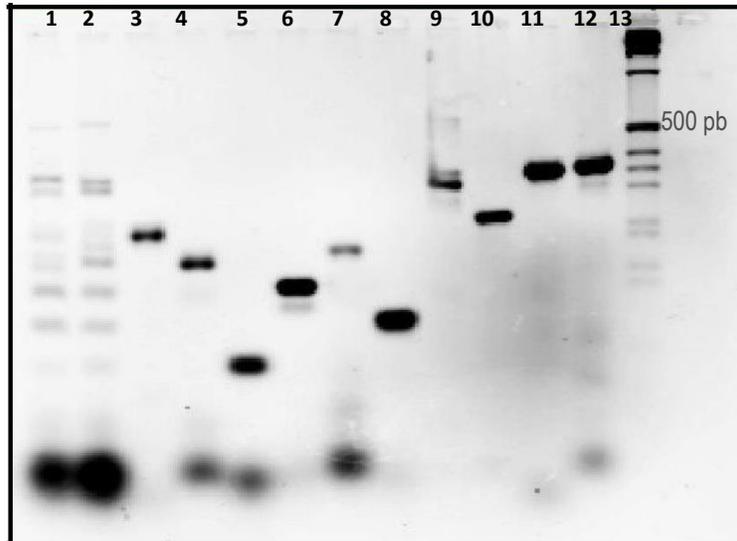
### 1.1 ESTABLECIMIENTO DE UN PROTOCOLO DE CALIDAD

Para sistematizar las futuras verificaciones de pureza en las RILs se implementó una técnica de co-amplificación de 10 marcadores multiplexados. La Figura 1.6 muestra el perfil de amplificación del ensayo de 10 SSR en las líneas parentales, La imagen corresponde a un gel de agarosa con concentración del 3% (p/v) donde puede observarse el alineamiento de cada uno de los fragmentos amplificados por separado con la co-amplificación de los mismos en las líneas parentales.

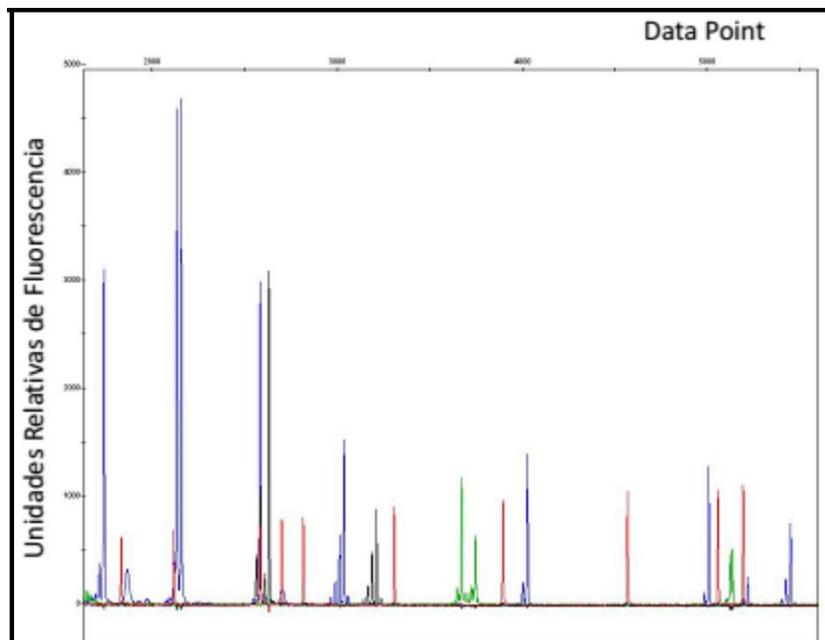
La Figura 1.7 muestra un electroferograma del producto obtenido de la corrida electroforética capilar para PAC2. La obtención de las longitudes de los fragmentos individuales mostrados en esta figura, corrobora los tamaños esperados para PAC2.



**FIGURA 1.5. ESQUEMA DE LOS PASOS REALIZADOS PARA LA REESTRUCTURACIÓN DE LA POBLACIÓN BIPARENTAL.** Referencias: 1) Ampliada por Talia et al (2010); 2) Datos genotípicos originales de Al-Chaarani et al (2004); 3) Perfil multilocus homocigota; 4) Perfil multilocus heterocigota (descartados).



**FIGURA 1.6. PERFILES DE CORRIDA DE LOS FRAGMENTOS EN GEL DE AGARSOSA CONCENTRACIÓN AL 3%.** Calles 1-2: producto de la amplificación multiplexada de las líneas parentales PAC2 y RHA266 respectivamente. Calles 3-12: fragmentos amplificados en PAC2 por separado, correspondientes a los marcadores: HA541, HA911, HA1590, HA1752, HA2190, HA3555, ORS166, ORS510, ORS1146 y ORS767. Calle 13: marcador de tamaño molecular.1Kb (Invitrogen)



**FIGURA 1.7. ELECTROFEROGRAMA DE UNA CORRIDA DE SSRs MULTIPLEXADA PARA PAC2.** Se observan los picos correspondientes a 10 microsatélites amplificados en una única reacción de PCR a partir de una muestra de girasol y el estándar interno de tamaño molecular (rojo)

## 2. CARACTERIZACIÓN GENOTÍPICA

### 2.1 GENOTIPIFICACIÓN DE SNPS MEDIANTE DHPLC

Para el desarrollo del protocolo de genotipificación de SNP vía dHPLC, en primera instancia y como medida de control de la integridad y capacidad de interacción de la columna Helix CP28353 (Varian Inc.), se obtuvo el patrón correcto de separación de picos para los fragmentos de ADN del plásmido pUC18 digerido con la enzima Hae III, estándar de corrida sugerido por el proveedor (Figura 1.3).

El desarrollo de esta técnica permitió definir 100 ng como la cantidad mínima de ADN, en un volumen de 5 a 10  $\mu$ l (alrededor de 20 ng/ $\mu$ l).

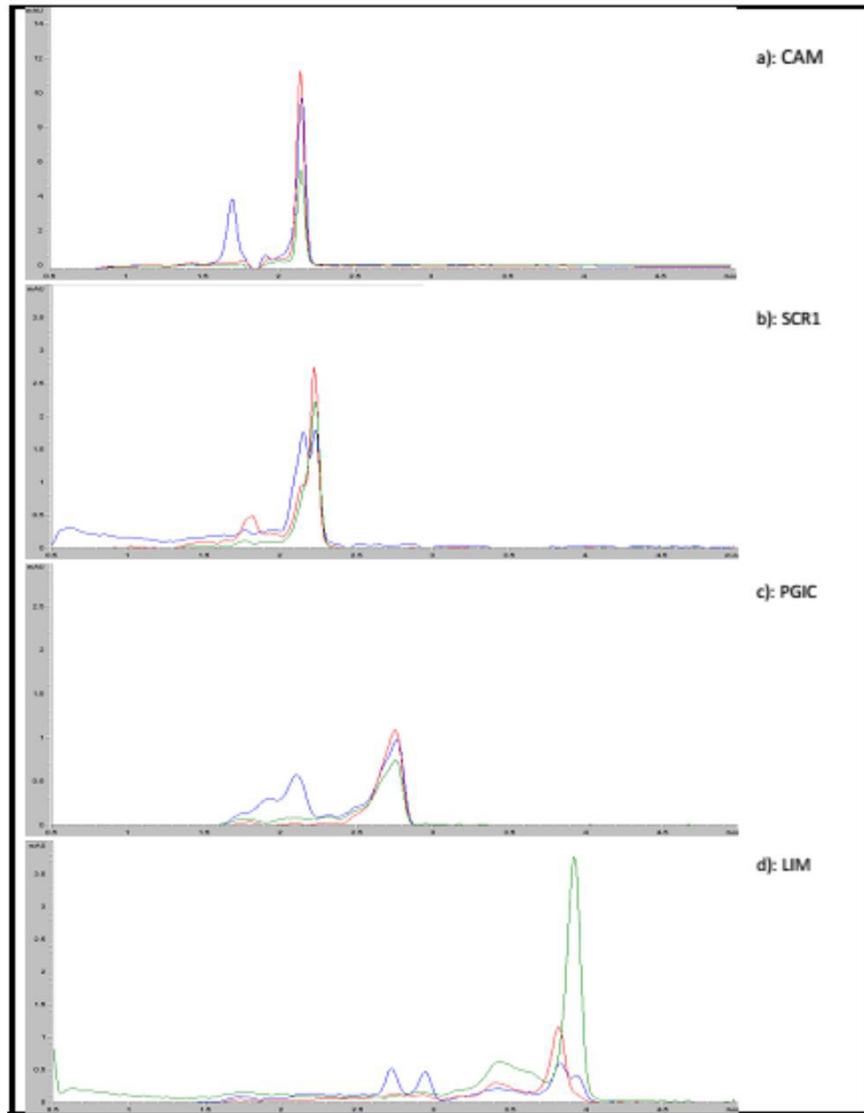
En este trabajo pudieron establecerse correctamente las condiciones de detección de heterodúplex y homodúplex para las siguientes regiones: CAM, SCR1, PGIC, LIM (Tabla 1.7).

La Figura 1.8 muestra los cromatogramas superpuestos de las corridas individuales para los GCs CAM, SCR1, PGIC, LIM amplificados en las líneas parentales y resueltos individualmente (controles homodúplex) o como una mezcla equimolar de los productos obtenidos para PAC2-RHA266 (control heterodúplex). En ellas puede observarse la presencia de solo un pico en las muestras homodúplex (RHA266 ó PAC2) o de dos a cuatro picos en los cromatogramas de las mezclas PAC2-RHA266 (heterodúplex). A cada uno de estos cromatogramas se les restó el cromatograma correspondiente al control negativo para ajustar a la línea de base

Un subconjunto de 93 RILs del total de la población fueron genotipadas, 9 de las cuales presentaron perfiles de cromatografía ambiguos y no pudieron ser asignados los alelos correspondientes (marcadas con \* en la Tabla II del apéndice). Para descartar la hipótesis de que pueda tratarse de un individuo heterocigota, se realizó un control extra, donde el producto de PCR de cada RIL se inyectó en la columna de dHPLC sin tratamiento previo. Este control no es sometido a ningún proceso de reordenamiento de hebras, por lo cual siempre debe dar un cromatograma de tipo “homodúplex”.

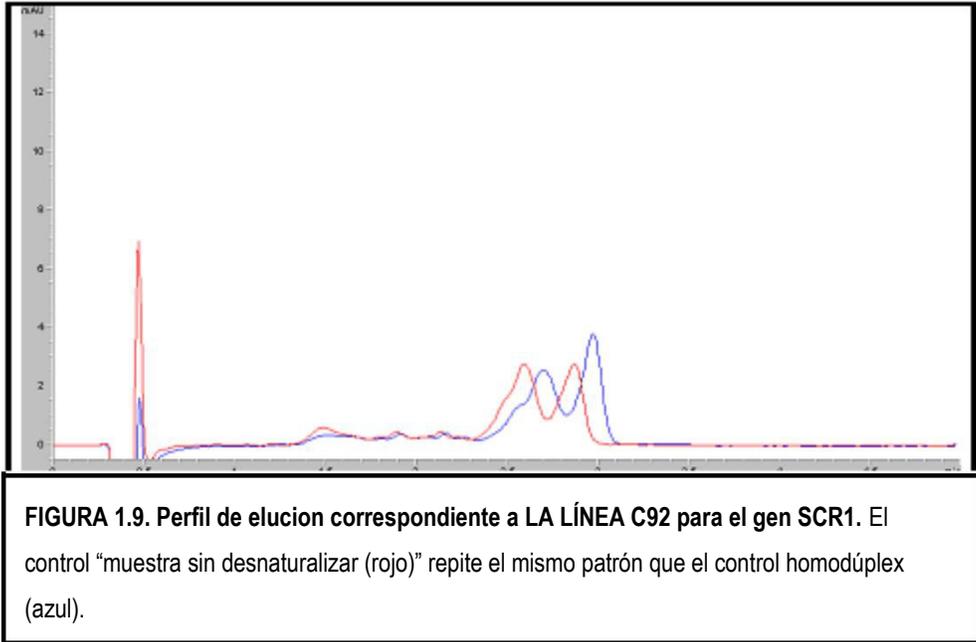
La Figura 1.9 muestra los resultados obtenidos con el individuo C92 para el locus SCR1 a manera de ejemplo. Aquí se observa que el control “muestra sin desnaturalizar” repite el mismo patrón que el control homodúplex. Por lo que se descarta la posibilidad de que estemos frente a un individuo heterocigota.

Para las 7 regiones restantes: 1ACCO-1, PGIP1, PGIP3, MADSB-TF3, ANT, GCvT y GPX, no se pudieron establecer las condiciones óptimas de detección de la especie heterodúplex y 50



**FIGURA 1.8. RESOLUCION DE SNPs DE DISTINTOS GENOTIPOS MEDIANTE dHPLC PARA LOS GENES S a) CAM; b) SCR1; c) PGIC Y d) LIM, MEDIANTE dHPLC. En cada cromatograma se observa: RHA266 homodúplex (rojo), PAC2 homodúplex (verde) y RHA266-PAC2 heterodúplex (azul).**

homodúplex con los parentales de la población. Para descartar errores en los productos de amplificación los siete fragmentos fueron secuenciados para las líneas PAC2 y RHA266. El análisis de las secuencias permitió concluir que en los loci secuenciados existían los distintos sitios polimórficos entre ambas líneas.

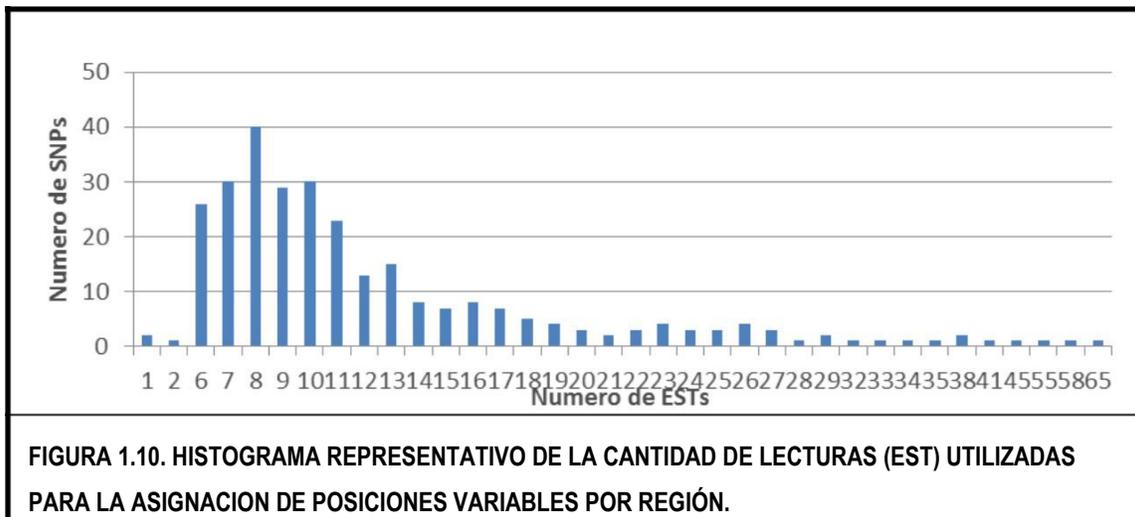


2.2 GENOTIPIFICACIÓN MEDIANTE ENSAYO 384 SNPs

2.2.1 Detección de SNPs y diseño del ensayo

A partir del análisis de los contigs de la base de datos SUR (Fernandez, Soria, et al. 2012), se identificaron 1200 SNPs luego de aplicar los filtros pertinentes descritos en la sección materiales y métodos. Las secuencias adyacentes a los SNP putativos identificados fueron analizadas con el programa ADT (Illumina) y 287 SNPs *in silico* con valores de diseñabilidad  $\leq 0,6$ , fueron seleccionados. La Figura 1.10 muestra el número total de ESTs analizados para la detección de 287 SNPs *in silico*.

De las 119 regiones genómicas seleccionadas a partir de un conjunto de fragmentos



resecuenciados, 22 de las mismas fueron descartadas ya que no cumplieron con los requisitos mínimos establecidos para su selección. De esta manera, 97 SNPs *in vitro* fueron incluidos en el ensayo SOPA, representando a 74 genes candidatos diferentes con valores de diseñabilidad de 1.

### 2.2.2 Genotipificación del ensayo

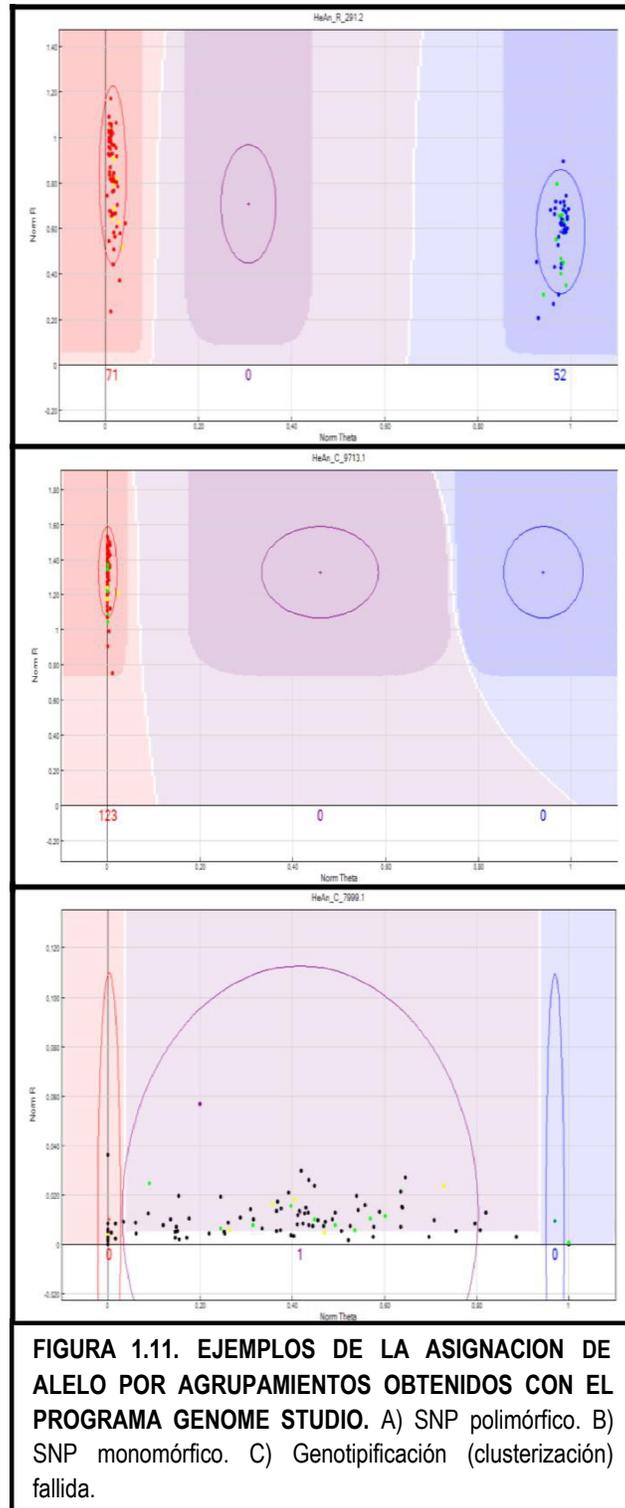
Para la genotipificación correspondiente, se ingresaron los 159 individuos al estudio. El análisis se realizó observando las agrupaciones en clústeres determinadas por el algoritmo propio del programa. En primer lugar, 36 individuos fueron eliminados de los análisis posteriores debido a un valor de “call rate” <0,70. De esta manera se prosiguió con 123 individuos en total (108 RILs + 8 PAC2 y 7 RHA266).

El ensayo de genotipificación en sí, consistió en evaluar los agrupamientos generados automáticamente por el software GenomeStudio. La Figura 1.11 permite observar las diferentes categorías obtenidas de asignación de alelos para SNPs. En ciertos casos, la separación en grupos aislados efectuada automáticamente no fue tan clara, presentándose solapamientos de clústeres, muestras sin asignar y genotipos aparentemente heterocigotas, entre otras. Esto tornó dificultosa la tarea de genotipificación, de tal manera que un total de 43 SNPs debieron ser ajustados manualmente, ampliando los óvalos que representan a 2 desviaciones estándares o desplazando el centro de los clústeres. Luego del reajuste, 10 SNPs pudieron ser claramente identificados como polimórficos y los restantes 33, no pudieron ser reasignados, por lo que se los categorizó como “falla en la genotipificación (clusterización)”.

Por último, la reproducibilidad fue corroborada en la totalidad de SNPs polimórficos evaluados, donde se obtuvo un 100% de identidad entre las 8 muestras parentales PAC2 y las 7 muestras RHA266, identificadas en la Figura 1.11 A (en este SNP a modo de ejemplo) como puntos amarillos y puntos verdes, respectivamente.

### 2.2.3 Caracterización del ensayo

La tasa de éxito en la genotipificación del ensayo SOPA fue determinado como el cociente del número de SNPs que fueron correctamente asignados a un genotipo (tanto monomórfico como polimórfico) y el número de SNPs que no pudieron ser clasificados, probablemente debido a una mala clusterización, valor de GenTrain <0.25 o por problemas de amplificación. En la Tabla 1.9 se presentan los números y porcentajes de SNPs monomórficos, polimórficos y fallidos, detallándose cada categoría (*in silico* e *in vitro*).



De los 384 SNPs analizados inicialmente, 282 fueron claramente asignados a un genotipo, obteniéndose una tasa de éxito del 73,43%, mientras que la tasa de conversión fue calculada en 16,67% (64/384).

TABLA 1.9. Tasa de éxito del ensayo SOPA.		
CATEGORÍA	# de SNPs ( <i>in vitro</i> / <i>in silico</i> )	% de SNPs ( <i>in vitro</i> / <i>in silico</i> )
Polimórfico	64 (26/38)	16,67 (26,80/13,24)
Monomórfico	218 (55/163)	56,77 (56,70/56,79)
Fallidos	102 (16/86)	26,56 (16,49 /29,97)
Total	384 (97/287)	100 (100/100)

Teniendo en cuenta las dos estrategias de búsqueda de SNPs, se observó una proporción levemente menor de datos perdidos para aquellos SNPs descubiertos *in vitro* (1,33%) respecto de los *in silico* (1,36%), totalizando cerca de 1,33% al considerarlos en conjunto.

Mediante un análisis de  $\chi^2$ , se encontraron diferencias significativas sobre la tasa de éxito ( $P [\chi^2 > 6,74] = 0,009, GL = 1$ ) y de conversión ( $P [\chi^2 > 9,60] = 0,009, GL = 1$ ) debidas al origen de los marcadores SNPs. En general los marcadores *in vitro* mostraron mayor proporción de marcadores claramente asignados a un genotipo respecto de SNPs *in silico*.

### 3. CONSTRUCCIÓN DEL MAPA GENÉTICO

La matriz de mapeo genético incluyó un total de 739 marcadores que resultaron polimórficos.

El análisis de segregación de los marcadores que componen la matriz de datos con el test estadístico FDR mostró que cinco marcadores presentaron segregación distorsionada ( $p < 0,05$ ), por lo que fueron descartados de futuros análisis.

El próximo paso consistió en descartar 21 individuos y 28 marcadores que poseían más del 50% de datos faltantes. Esto redujo sustancialmente el porcentaje de datos perdidos totales de un 27,7% al 17,2% quedando una matriz final compuesta por 114 RILs y 706 marcadores.

Como se describió en Materiales y Métodos, para la confección del mapa genético de ligamiento se utilizó la versión de prueba del programa JoinMap v4.1 (Van Ooijen & Voorrips 2001), estableciendo un porcentaje de recombinación máximo de 35% y un LOD mínimo de 4. Tanto el abordaje de mapeo considerando todos los GL al mismo tiempo, como el de GL de forma individual arrojaron resultados equiparables, corroborando el correcto agrupamiento de marcadores en GLs.

El ordenamiento de los marcadores mediante el establecimiento de un orden inicial fijo, generó mapas relativamente más cortos (y por ende más confiables) y sin reordenamientos

de bloques de marcadores, frecuentes en la estrategia que no implicaba información alguna sobre la disposición de marcadores. (Figura 1.12).

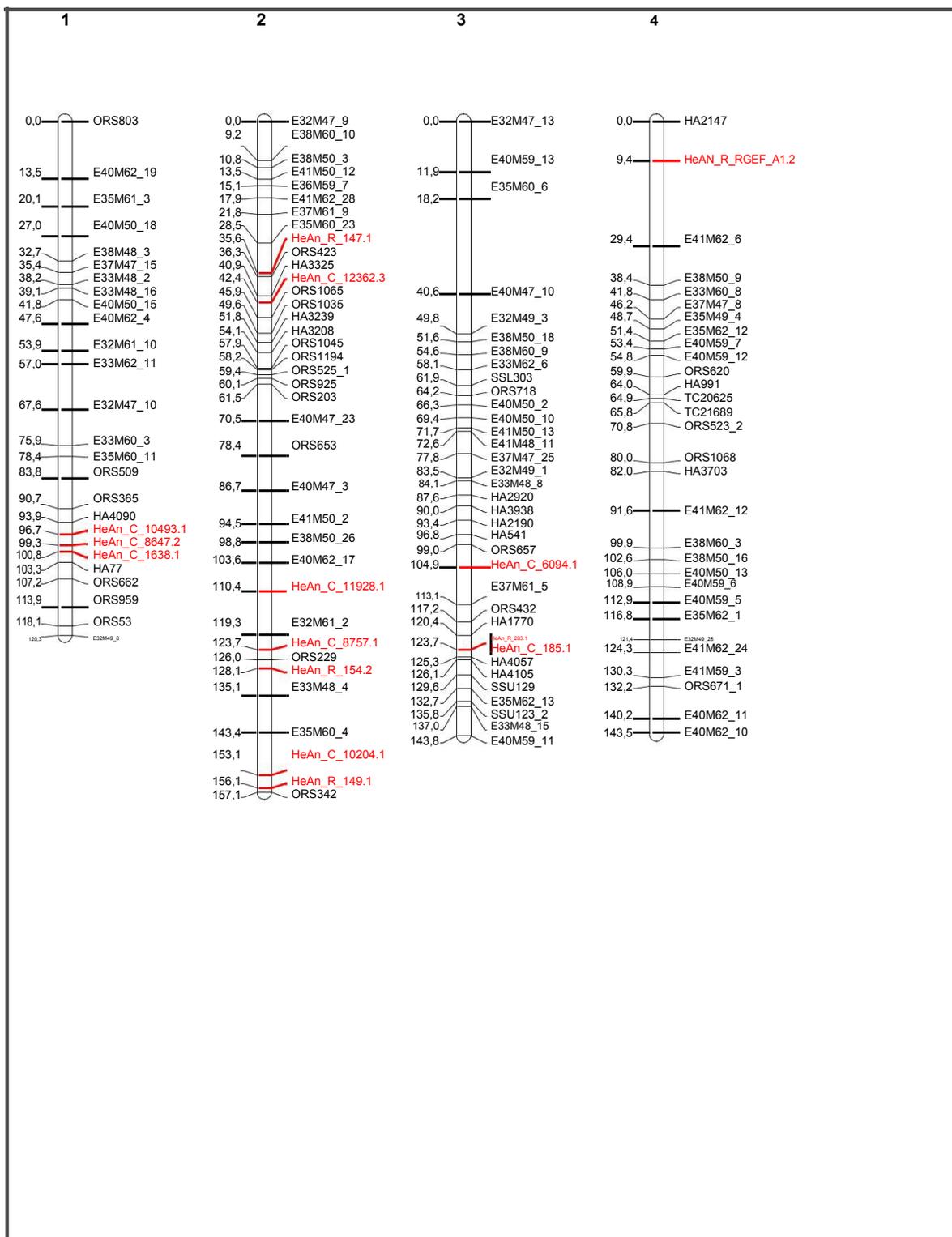
La Tabla 1.10 muestra la distribución y densidad de los marcadores en cada grupo de

<b>GRUPO DE LIGAMIENTO</b>	<b># AFLPs</b>	<b># SSRs</b>	<b># EST-SSRs</b>	<b># INDELS</b>	<b># SNPs</b>	<b># HeAns</b>	<b># locus</b>	<b>LONGITUD (cM)</b>	<b>Densidad (cM/locus)</b>
1	15	8	-	-	-	3	26	120,34	4,63
2	16	14	-	-	-	7	37	157,11	4,25
3	19	13	-	-	-	3	35	143,77	4,11
4	2-	7	2	-	-	1	30	143,5	4,78
5	16	12	-	-	-	4	32	176,26	5,51
6	15	7	-	-	2	2	26	125,55	4,83
7	16	8	-	-	-	1	25	103,15	4,13
8	19	2-	1	-	-	3	43	183,59	4,27
9	27	18	-	-	1	4	50	239,35	4,79
10	5-	2-	1	1	2	7	81	276,64	3,42
11	14	7	-	-	1	4	26	107,86	4,15
12	17	17	-	-	1	5	40	178	4,45
13	9	6	-	-	1	-	16	88,31	5,52
14	19	19	-	1	-	7	46	181,63	3,95
15	11	14	1	-	-	4	30	113,79	3,79
16	21	26	2	-	-	5	54	214,61	3,97
17	23	13	1	-	-	1	38	270,34	7,11
<b>TOTAL</b>	<b>327</b>	<b>229</b>	<b>8</b>	<b>2</b>	<b>8</b>	<b>61</b>	<b>635</b>	<b>2823,80</b>	<b>4,45</b>
<b>PROMEDIO</b>	<b>19,24</b>	<b>13,47</b>	<b>1,33</b>	<b>1,00</b>	<b>0,47</b>	<b>3,59</b>	<b>37,35</b>	<b>166,11</b>	<b>4,57</b>

ligamiento.

De esta manera, el mapa genético final se construyó con 635 marcadores incluyendo 327 AFLPs (ExMx<sub>x</sub>), 229 SSR (78 HAx, 119 ORSx y 32 cartisol), 8 EST-SSR y 2 InDels (TCx), y 69 SNPs (61 HeAn<sub>x</sub> y 8 genes candidatos localizados sobre los mapas descritos previamente. Un total de 71 marcadores (5 HeAn<sub>x</sub>) no pudieron ser agrupados bajo los parámetros de mapeo seleccionados en este trabajo.

Se obtuvo una cobertura homogénea de los grupos de ligamientos obteniendo un mapa de 2823,80 cM. El tamaño promedio para cada grupo de ligamiento fue 166 cM, con una densidad media aproximada de un *locus* cada 4,45 cM. Al considerar únicamente los SNPs, la distribución promedio fue de 1 SNP cada 41 cM en promedio.



**FIGURA 1.12. MAPA GENÉTICO DE LIGAMIENTO PARA LA POBLACIÓN PAC2 x RHA266.** Los grupos de ligamientos generados fueron nombrados acorde al trabajo de Talia et al (2010). Las distancias (cM) entre loci se detallan a la izquierda de cada grupo de ligamiento. Los nuevos marcadores incorporados a este mapa se destacan en rojo.

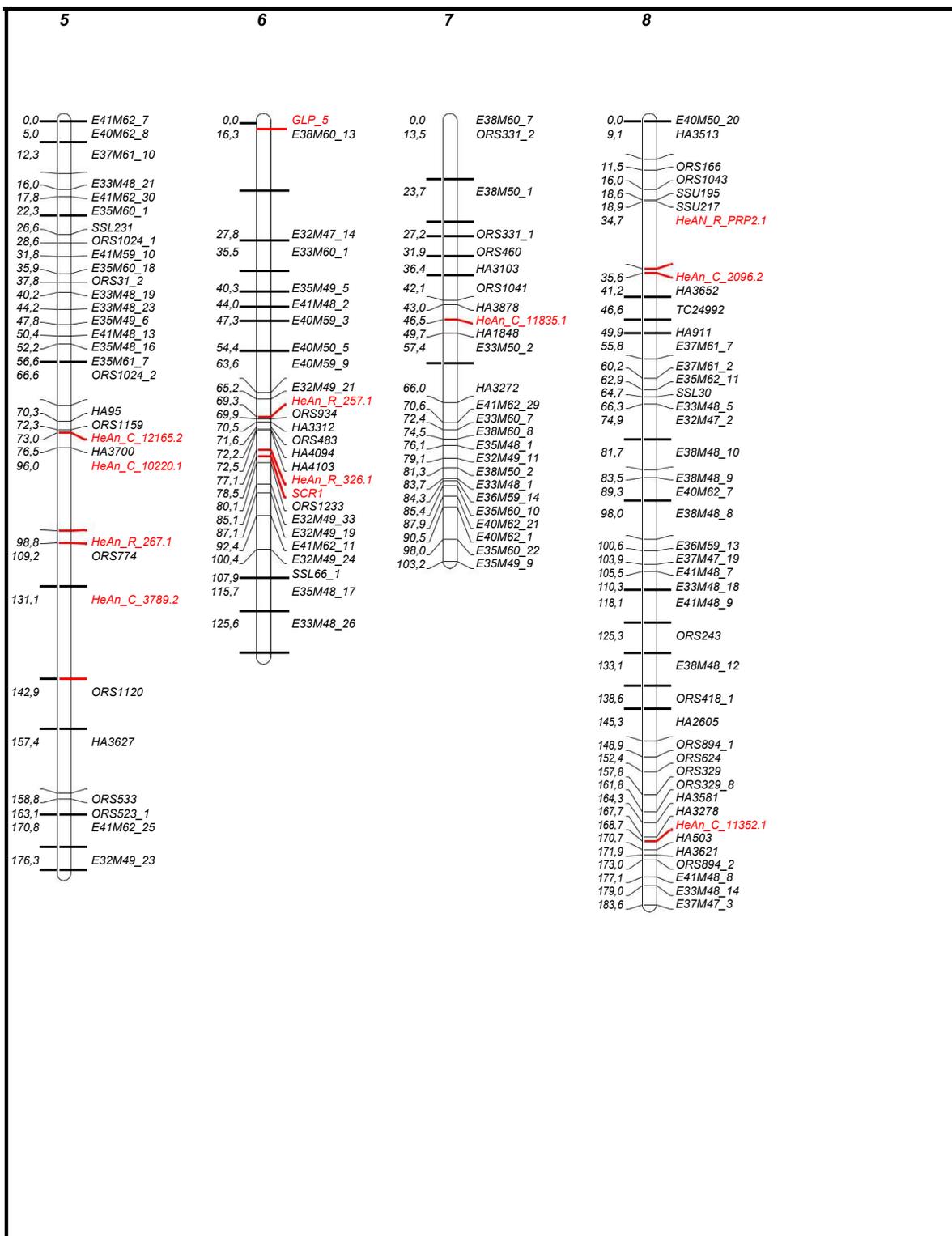


FIGURA 1.12 (cont.). MAPA GENÉTICO DE LIGAMIENTO PARA LA POBLACIÓN PAC2 x RHA266. Los grupos de ligamientos generados fueron nombrados acorde al trabajo de Talia et al (2010). Las distancias (cM) entre loci se detallan a la izquierda de cada grupo de ligamiento. Los nuevos marcadores incorporados a este mapa se destacan en rojo.

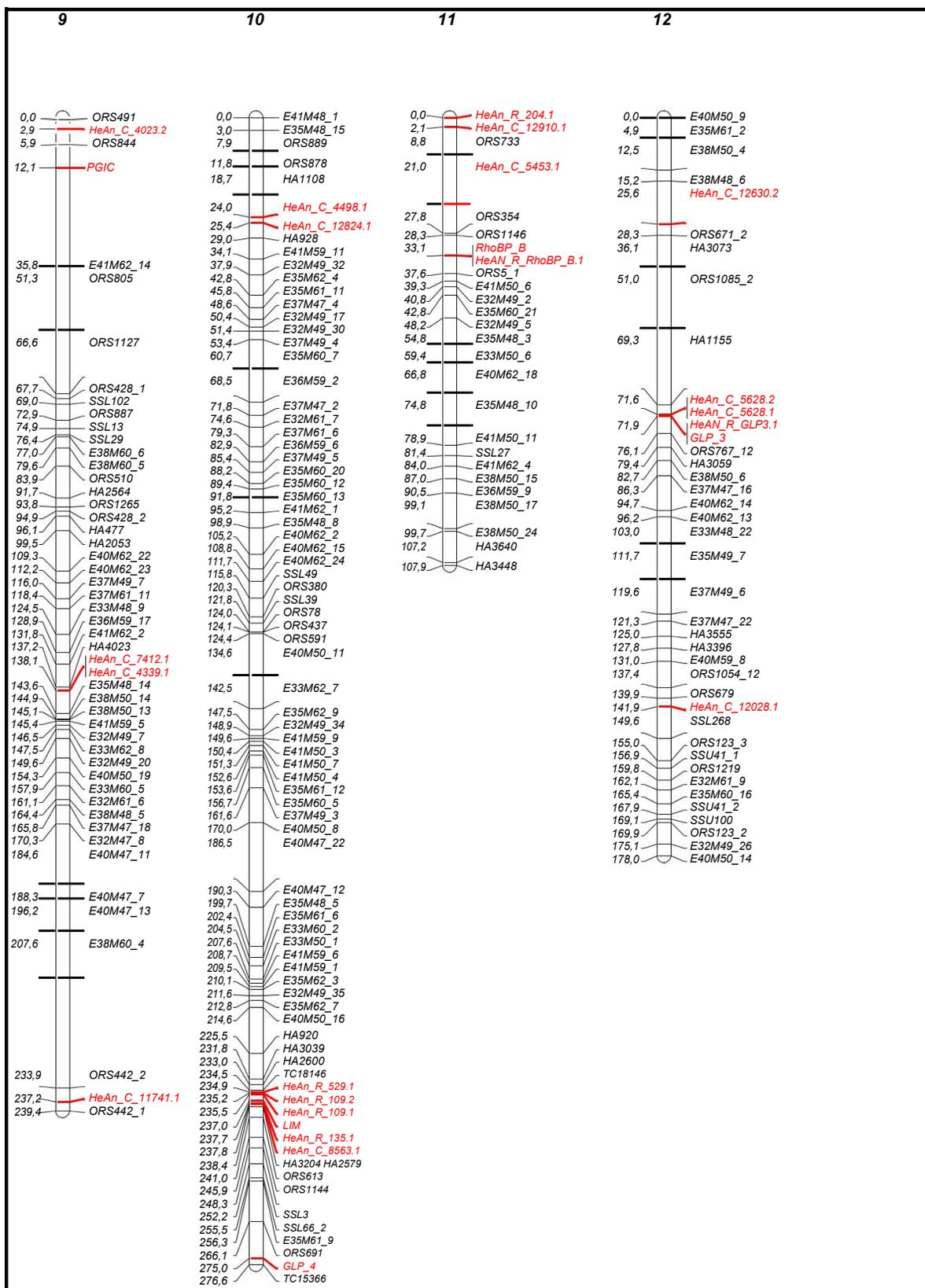


FIGURA 1.12 (cont.). MAPA GENÉTICO DE LIGAMIENTO PARA LA POBLACIÓN PAC2 x RHA266. Los grupos de ligamientos generados fueron nombrados acorde al trabajo de Talia et al (2010). Las distancias (cM) entre loci se detallan a la izquierda de cada grupo de ligamiento. Los nuevos marcadores incorporados a este mapa se destacan en rojo.

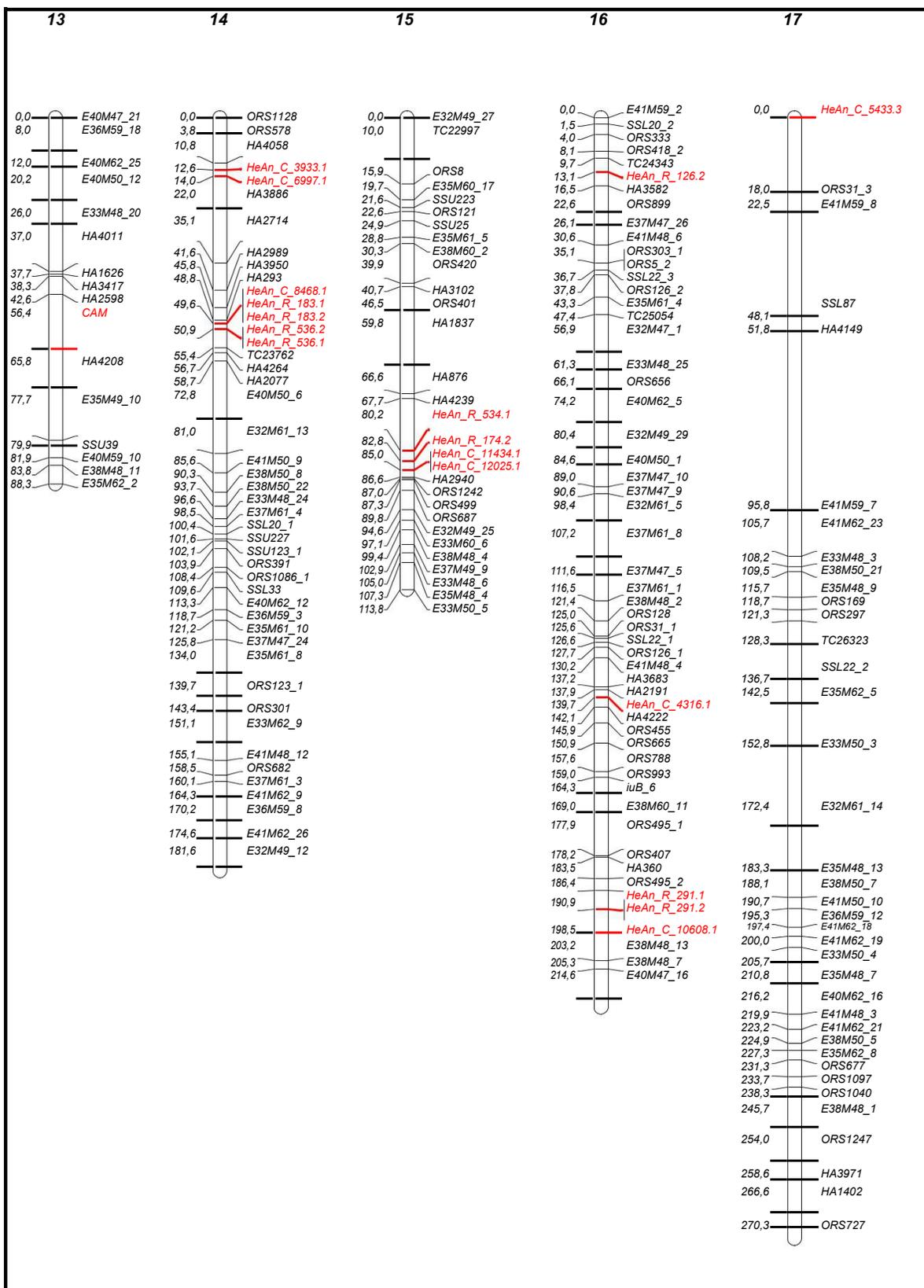


FIGURA 1.12 (cont.). MAPA GENÉTICO DE LIGAMIENTO PARA LA POBLACIÓN PAC2 x RHA266. Los grupos de ligamientos generados fueron nombrados acorde al trabajo de Talia et al (2010). Las distancias (cM) entre loci se detallan a la izquierda de cada grupo de ligamiento. Los nuevos marcadores incorporados a este mapa se destacan en rojo.

La longitud de los grupos fluctuó entre 88,31 (GL 13) a 276,64 cM (GL 10) conteniendo 16 y 81 marcadores, en los GL 13 y 10, respectivamente. La mayor distancia genética sin marcadores fue de 44 cM en el GL 17.

Ocho nuevos genes candidatos analizados mediante dHPLC (SCR1, LIM, PGIC, CAM, GLP\_3, GLP\_4, GLP\_5 y RhoBp\_B) se incorporaron como marcadores funcionales al mapa. Éstos se ubicaron en distintos GLs (Figura 12). El SCR1 y el GLP\_5 se mapearon en el GL 6. El locus PGIC se posicionó en la porción distal del GL 9. Por otro lado, los genes LIM y GLP\_4 se mapearon en el GL 10, al tiempo que RhoBP\_B, GLP\_3 y CAM se anclaron en los GL 11, 12 y 13 respectivamente.

Respecto de los SNPs genotipados a partir del ensayo multiplexado, se distribuyeron homogéneamente a lo largo de los GLs originados. Un 96,15% (25/26) de los SNPs *in vitro* polimórficos pudieron ser ubicados mientras que el porcentaje fue de 94,74% (36/38) para los SNPs *in silico*.

Por último, el proceso de construcción de un mapa genético con nuevos marcadores genotipados, permitió mapear 26 AFLPs y 21 SSRs (14 ORSx, 5 HAx, 1 SSUx y 1 TCx) que permanecían como no ligados en mapas anteriores.

### 3.1 CONFIRMACIÓN DEL REGISTRO DEL MARCADOR

Los GCs GLP\_3 y RhoBP\_B fueron genotipados de a pares en la población de RILs por las tecnologías dHPLC e Illumina GoldenGate (nombrados como “GLP\_3” y “RhoBP\_B” para la primera y “HeAn\_R\_GLP3.1” y “HeAn\_R\_RhoBP\_B.1”, para la segunda). Mediante el estudio de la localización de cada par sobre el mapa y la recombinación entre los GCs que lo integran, se planteó validar la obtención de genotipos por ambas técnicas y su utilización en el mapeo genético.

Tanto los GCs que integraban el par perteneciente a GLP\_3 como los mismos para RhoBP\_B mostraron una localización contigua sobre los GLs 12 y 11, respectivamente, producto de una recombinación menor al 3% entre ellos.

Del mismo modo, la búsqueda de SNPs para el desarrollo del ensayo SOPA permitió descubrir haplotipos con al menos 2 SNPs sobre su secuencia. Según el origen de las mismas, estas fueron nombradas como “HeAn\_C\_x” y “HeAn\_R\_x” (para SNPs identificados *in silico* e *in vitro*, respectivamente). Considerando una probabilidad nula o muy escasa de recombinación entre SNPs separados por unos pocos cientos de pares de bases, en casi la totalidad de los casos se

esperarían 2 haplotipos. Es así que los genotipos obtenidos para 5 pares de SNPs polimórficos (HeAn\_R\_109.1/.2, HeAn\_C\_5628.1/.2, HeAn\_R\_183.1/.2, HeAn\_R\_536.1/.2 y HeAn\_R\_291.1/.2), conformando un haplotipo cada uno, fueron evaluados.

En la totalidad de RILs evaluadas, por individuo, se obtuvo un 100% de identidad entre marcadores pertenecientes a un mismo amplicón o EST y la misma ubicación en el genoma.

Por último, al observar el mapa genético (Fig 1.12), puede notarse que un conjunto de SNPs identificados *in vitro* e *in silico*, co-localizaron sobre los GLs 3 y 14 (HeAn\_C\_185.1 - HeAn\_R\_283.1 y HeAn\_C\_8468.1 - HeAn\_R\_183.1/.2, respectivamente). Ambos conjuntos no presentaron recombinación entre los SNPs que lo integran. Dado que no se posee mayor información sobre SNP *in silico*, se llevó adelante un ensayo de alineamiento entre las secuencias de haplotipos que integran cada grupo con la finalidad de encontrar posibles correspondencias.

Únicamente el grupo localizado sobre el GL 14 (HeAn\_C\_8468.1 - HeAn\_R\_183.1/.2) mostró un alineamiento total en los extremos de sus secuencias de unas 30 pb de longitud. Aunque una visión más profunda de sus funciones putativas reveló que ambos haplotipos se han descrito como carboxiquinasa fosfoenolpiruvato. Debido a la falta información previa, la comparación hecha dentro de este grupo requiere de mayor análisis para validar totalmente si ambos SNPs pertenecen al mismo gen candidato. Esto último, no descarta la posibilidad de que exista ligamiento genético de marcadores distantes físicamente debido a la falta de recombinación (por ejemplo, regiones centroméricas).

# DISCUSIÓN

## 1. LA POBLACIÓN DE MAPEO GENÉTICO

En este trabajo se utilizó una población de 135 líneas recombinantes endocriadas (RILs) para la que se disponía de un mapa genético desarrollado en colaboración con el INRA (Instituto Nacional de Investigación Agronómica, Francia) (S. Kiani et al. 2007), con la finalidad de saturar los distintos grupos de ligamiento con marcadores de tipo SNPs e identificar las regiones del genoma de girasol asociadas a resistencia a *Sclerotinia sclerotiorum*. La población, que deriva del cruzamiento entre las líneas puras de girasol RHA266 (resistente a mildiu) y PAC2 (resistente a *S. sclerotiorum*) y expresa un comportamiento diferencial para la respuesta de resistencia al desarrollo de la PHC, ha sido utilizada en numerosos trabajos de mapeo genético y mapeo de QTLs para diferentes caracteres agronómicos (Al-Chaarani et al. 2004; Alfadil et al. 2007; Ebrahimi et al. 2008; Berrios et al. 2000; Hervé et al. 2001; S. Kiani et al. 2007; Kiani et al. 2008; Paola Talia et al. 2010) por lo que el desarrollo de la presente Tesis permitirá complementar la información generada por distintos grupos de trabajo.

## 2. CONTROL DE PUREZA

Para numerosas especies de fecundación cruzada, la contaminación genética es un problema establecido tanto en híbridos como en líneas elites y líneas recombinantes endocriadas. Esto se hace evidente cuando la pureza genética de la semilla híbrida influye directamente sobre la producción de un cultivo. Esta contaminación aumenta los niveles de endogamia, reduciendo la calidad genética y fisiológica de las semillas que en consecuencia disminuye la productividad de los cultivos. Por lo tanto, es de suma importancia evaluar la pureza genética en la producción de semillas para el comercio (Dongre & Parkhi 2012; Liwang et al. 2004; Garg et al. 2006). Por otra parte, definir la identidad, pureza y estabilidad de las líneas elites y RILs permite la protección de los derechos de aquellas instituciones que las desarrollaron. Desarrollar controles eficaces de calidad de la semilla permite garantizar la aplicabilidad de las líneas en cuestión en programas de mejoramiento o estudios de diversidad.

Tradicionalmente, la pureza genética de semillas híbridas en una variedad de cultivos ha sido evaluada basada en caracteres morfológicos en diferentes estadios del crecimiento o de acuerdo a sus niveles de resistencia a diferentes enfermedades (Moorthy et al. 2011; X. F. Wang et al. 2005). Sin embargo, las características morfológicas de los diferentes genotipos

suelen ser muy similares entre sí en algunas etapas, por lo que es difícil evaluar la pureza visualmente (Liu et al. 2007).

En los últimos años, los marcadores moleculares, se han convertido en una herramienta fundamental para la evaluación de la pureza y control de calidad de semillas en diversos cultivos. Estos han sustituido a técnicas alternativas que emplean isoenzimas o polimorfismos en proteínas (Tanksley & Jones 1981; X. F. Wang et al. 2005) demostrando ser más eficientes y menos demandantes de tiempos. De esta manera, la genotipificación del ADN (huella genética de ADN) ha abierto la posibilidad de identificar variación genómica a nivel de individuos, y de seguir el intercambio de ADN en la progenie, por lo que es una herramienta de gran alcance para la selección genotípica. Los diversos trabajos realizados incluyen arroz (Garg et al. 2006; Moorthy et al. 2011), algodón (Dongre & Parkhi 2012), maíz (Wu et al. 2006), cártamo (Naresh et al. 2009), y algunos cultivos hortícolas como tomate (Liu et al. 2007), pimientos (Mongkolporn et al. 2004) y calabaza (Ferriol et al. 2003).

En cuanto a girasol, Nandini y Chikkadevaiah (2005) realizaron una investigación utilizando híbridos, sus líneas parentales y variedades de polinización abierta para identificar y establecer las relaciones filogenéticas entre los genotipos y caracterizarlos basados caracteres morfométricos y marcadores RAPDs (*“Random Amplified Polymorphism”*). Este estudio permitió corroborar que las líneas morfológicamente muy similares no se encontraron relacionadas a nivel molecular y tales problemas asociados con la clasificación taxonómica enfatizaron la necesidad de herramientas complementarias para la identificación y caracterización de genotipos de girasol. De esta manera, establecieron un conjunto estándar de cebadores que se pueden utilizar para distinguir y caracterizar la mayoría de los genotipos comunes que sirven de ese modo como un suplemento útil a los datos morfológicos y agronómicos tradicionales para la protección de variedades vegetales.

Por otra parte, los marcadores SSRs han demostrado ser altamente eficiente para estudios del ADN y análisis de pedigrí (Weising et al. 1995), mapeo genético (McCouch et al. 1997; Röder et al. 1998), estudios poblacionales y evolutivos (Goldstein et al. 1999), dado al alto nivel de polimorfismo, reproducibilidad de los patrones generados, herencia codominante y fácil implementación. Estas características promovieron la elaboración de un amplio conjunto de SSRs para especies como arroz (McCouch 2002), cebada (Ramsay et al. 2000; Varshney et al. 2007) trigo (Hyten et al. 2008), maíz (Sharopova et al. n.d.), soja (Hwang et al. 2009), tomate

(He et al. 2003), papa (Ashkenazi et al. 2001), girasol (Paniego et al. 2002; Tang et al. 2002; Duca et al. 2014), entre otras.

Duca et al (2014) al analizaron el genotipo de 10 líneas de girasol cultivado en la República de Moldavia con 13 pares de oligonucleótidos correspondientes a SSRs públicos y concluyeron que siete marcadores fueron suficientes para discriminar inequívocamente entre las líneas parentales analizadas.

En el presente trabajo de Tesis, el control de calidad de pureza de las RILs utilizadas en sucesivos ensayo fue fundamental para asegurar la evaluación del carácter evitando mediciones espúreas derivadas de la pérdida de homogeneidad de los genotipos. En este sentido se dispuso el desarrollo e implementación de un protocolo de calidad para corroborar la identidad de las líneas que integran la población de mapeo y evaluar las sucesivas multiplicaciones de dicho material. Para ambos casos, se empleó la técnica de genotipificación a través de electroforesis capilar (EC) en secuenciador automático. Esto contrasta con la utilización de las técnicas de genotipificación manuales, mediante geles de poliacrilamida y la visualización con nitrato de plata, con SYBR® Gold o con nucleótidos marcados con radiactivo y detección por autorradiografía que son muy demandantes de tiempo.

La genotipificación de los 39 marcadores moleculares SSRs seleccionados para comprobar la identidad de las RILs alcanzó un alto nivel de procesividad utilizando la combinación en paneles de 2 a 6 marcadores compatibles en corridas de EC, previniendo la superposición de fragmentos del mismo tamaño y del mismo fluoróforo durante la co-electroforesis. El resultado de esta evaluación determinó la incorporación de 12 nuevos individuos con perfil homocigota, incrementándose la población 123 a 135 individuos.

Una desventaja importante de los SSRs como herramienta para genotipado es el bajo nivel de automatización comparado con marcadores como los SNP, lo que los hace demandantes de tiempo en aplicaciones de alto rendimiento. Parte integral de la aplicación de marcadores moleculares en fitomejoramiento e investigación son los métodos para mejorar la velocidad y eficiencia del genotipado de SSRs. La PCR multiplexada, donde varios marcadores se amplifican simultáneamente en la misma reacción, se presenta como una buena herramienta para aumentar la cantidad de información generada por ensayo, y al mismo tiempo reducir los costos de PCR y detección de los fragmentos (Henegariu et al. 1997).

El desarrollo de ensayos de PCR multiplexadas en plantas es difícil de abordar debido al gran tamaño del genoma y poliploidía de las mismas. Por ejemplo, la amplificación multiplexada de

SSRs en trigo, requirió una extensa optimización para superar la complejidad de un alohexaploide, compuesto por familias de genes parálogos y alrededor de un 75% de ADN repetitivo (Bennett 1995).

El desarrollo de ensayos de PCR multiplexadas en especies de cultivos con genomas diploides más pequeños también puede ser un reto (Mitchell et al. 1997; Masi et al. 2003) principalmente por las interacciones indeseables entre los oligonucleótido y la amplificación inespecífica (Elnifro et al. 2000). Sin embargo, hay antecedentes de PCR multiplexadas han sido desarrollada para colza (Mitchell et al. 1997), sorgo (Dean et al. 1999), soja (Narvel et al. 2000), maíz (Gethi et al. 2002), girasol (Tang et al. 2003), entre otros. La mayoría de estos paneles de SSR multiplexados incluyen un número limitado de marcadores y fueron diseñados principalmente para estudios de genotipificación, por ejemplo, para la identificación de líneas o variedades, la protección de la propiedad intelectual, y las pruebas de pureza de las semillas.

En el presente trabajo, se empleó el sistema PCR Type-it Microsatellite que se basa en la tecnología multiplex patentada por QIAGEN que utiliza una Taq Polimerasa “hot start” y un sistema de buffers optimizado que elimina eliminada la necesidad de realizar ajustes precisos sobre parámetros de la reacción de PCR tales como la concentración de  $Mg^{2+}$  y las condiciones de ciclado. Este procedimiento sencillo para la obtención de genotipos claros desarrollado a partir de una PCR multiplexada basada en 10 marcadores SSRs, mostró ser exitoso para el control de pureza de las semillas obtenidas cada campaña de multiplicación de las líneas. La misma, pudo discriminar entre individuos altamente emparentados como lo son las RILs y ha permitido reducir los costos y el tiempo necesario para llevar adelante esta tarea.

### **3. SNPs**

Para la genotipificación de genes candidato mediante la detección de SNPs se seleccionaron dos técnicas.

En primer lugar, se utilizó la técnica de resolución diferencial de heterodúplex mediante cromatografía líquida de alta precisión en condiciones de desnaturalización parcial (dHPLC). Con ella se pudieron genotipificar de manera inequívoca 4 genes candidatos de los 11 seleccionados originalmente, con un número mínimo de indeterminaciones. Para determinar si la causa de la indeterminación era la presencia de un locus heterocigota para estos individuos y en los loci analizados en particular, se realizaron corridas del fragmento de amplificación para cada uno de los individuos por separado en condiciones desnaturalizadas y parcialmente re-naturalizadas y no desnaturalizadas. Los resultados de la corrida de estos

fragmentos mostraron un patrón de resolución idéntico (Figura 1.9), con lo cual se descartó que se tratara de un locus heterocigota.

Para las siete regiones restantes para las que no pudieron establecerse las condiciones óptimas de detección de la especie heterodúplex y homodúplex con los parentales de la población se realizó la secuenciación de sus amplicones. El análisis de las secuencias permitió concluir que en los loci secuenciados se detectaban los distintos sitios polimórficos entre ambas líneas, con lo cual la falta de resolución de las mezclas mediante dHPLC puede deberse que no se hayan encontrados las condiciones óptimas de temperatura que permitan la identificación de las especies heterodúplex y homodúplex. No obstante, la técnica de dHPLC demostró ser un método útil y rápido para la identificación de SNPs e InDels a gran escala (Oefner 2000), una vez determinada las condiciones de corrida para cada uno de ellos. Entre las ventajas que ofrece el método, podemos mencionar la independencia del uso de iniciadores marcados necesarios para otra técnicas de genotipificación automática, permitiendo la adaptabilidad a sistemas semi-automatizados otorgando alta reproducibilidad y sensibilidad, y proporcionando un criterio objetivo de clasificación de las muestras (Abbas et al. 2004). Así mismo, el tiempo de procesamiento de las muestras es relativamente bajo dado que en general no requiere de pasos de purificación exhaustivas lo que favorece su uso para análisis de tipo poblacional teniendo en cuenta la reducción considerable del tiempo de procesamiento de las muestras (Xiao & Oefner 2001).

Numerosos estudios informaron acerca de su eficacia y especificidad en comparación con técnicas como PCR-RFLP (del Inglés: “Restriction Fragment Length Polymorphism”: polimorfismos de longitud de fragmentos de restricción) y SSCP (del Inglés: “Single Stranded Conformational Polymorphism”: polimorfismo de conformación de cadena) (Abbas et al. 2004; Jones et al. 1999; Yu et al. 2005). Costabile et al. (2006) evaluaron diferentes técnicas para el diagnóstico de enfermedades de inmunodeficiencia primaria, haciendo hincapié en la mayor exactitud del método dHPLC. Además, el método dHPLC es capaz de identificar los haplotipos en lugar de alelos individuales SNP, que pueden ser más informativos que los datos de un solo marcador (Rafalski 2002).

La temperatura óptima determinada para la detección inequívoca de las muestras heterodúplex en los fragmentos puestos a punto, oscilaron entre 54°C y 59°C. Esto sugiere que, a pesar de que la longitud del fragmento es una limitante para esta técnica, el conocimiento sobre la secuencia completa permite predecir las condiciones de

desnaturalización parcial. Como se describe en la Tabla 1.4, los genes candidatos sin InDels y con menor cantidad de SNPs sobre su haplotipo (LIM y SCR1), requirieron mayor temperatura para lograr una correcta identificación de los genotipos portadores de cada individuo de la población de RILs. Por el contrario, CAM y PGIC, cuyas secuencias presentan un mayor número de polimorfismos, requirieron temperaturas más bajas.

El método de dHPLC mostró ser más versátil que otras técnicas como la digestión de hetroduplex con la enzima Cel 1 (CEL1CH) para identificar múltiples SNPs por región, tal como lo demostró Fusari et al (2010), no obstante es complejo resolver mediante dHPLC regiones a las que se han asignado un numero alto de haplotipos por la gran cantidad de combinaciones de referencias que se deben hacer con las muestras incógnitas para poder asignar el alelo correspondiente .

Otra de las ventajas de la dHPLC es el muy bajo costo (0,8 U\$D por muestra genotificada) en comparación con cualquiera de los métodos de alto rendimiento disponibles para identificar SNPs hasta la fecha (Bagge & Lübberstedt 2008), que lo hace ideal para el mapeo de población biparental. De hecho, Kota et al. (2001) probaron el rendimiento del mapeo por dHPLC con cuatro ESTs previamente mapeados como marcadores RFLP obteniendo las mismas posiciones en el mapa de cebada. Luego, los mismos autores mapearon exitosamente 216 SNPs derivados de ESTs y cuatro InDels con la técnica dHPLC en tres poblaciones de mapeo (Kota et al. 2008). En girasol, dHPLC se ha utilizado en los trabajos de (Z. Lai et al. 2005; Lexer et al. 2005; Fusari et al. 2010) apoyando la viabilidad de la técnica.

La segunda técnica empleada para la genotipificación de SNPs fue la tecnología de Illumina GoldenGate. Esta ha mostrado un alto rendimiento con altas tasas de éxito y reproducibilidad (Ganal et al. 2009) en una variedad de estudios exitosos de diversidad y mapeo a través de una amplia gama de genomas vegetales con diferentes complejidades y conteniendo un gran número de genes parálogos (Hyten et al. 2008; Pavy et al. 2008; Close et al. 2009; Akhunov et al. 2009; Yang et al. 2010; Liu et al. 2012; Rostoks et al. 2006; Thomson et al. 2010; McMullen et al. 2009; Thomson et al. 2011)

En el presente trabajo de Tesis, se evaluó un panel de genotipificación masiva que contó con 384 SNPs para girasol seleccionados a partir de una base de datos local Fernandez et al (2012) (SNPs in silico) y 119 genes candidatos resecuenciados en distintos genotipos de girasol (SNPs in vitro, Kolkman et al. 2007; Fusari et al. 2008). Un total de 287 SNPs in silico y 97 SNPs in vitro fueron seleccionados en base a diferentes criterios de selección

recomendados para la selección de SNP *in silico* y al valor de diseñabilidad asignado por el programa ADT. (Illumina).

La tasa de éxito observada para este diseño fue de 73,43%, que en el orden de los valores descriptos en otras especies no modelo, tales como 66,1% en eucalipto (Grattapaglia et al. 2011), el 66,9% y el 68,75% en el pino marítimo (Lepoittevin et al. 2010; Chancerel et al. 2011) o el 70,3% para *Crassostrea gigas* (Lapègue et al. 2014). Estos valores, sin embargo, están muy lejos de los obtenidos en especies modelo, para lo cual se dispone de bases de datos genómicas más completas o genomas secuenciados por completo, y con frecuencia superan el 80% de tasa de éxito (Hyten et al. 2008; Rostoks, Mudie, et al. 2005; Deulvot et al. 2010; Yang et al. 2010).

En cuanto al nivel de polimorfismos alcanzado usando este ensayo SOPA en la población de mapeo en estudio fue de 16,67%, nuevamente este valor se ajustó a lo esperado para ensayos similares. Por ejemplo, valores de 19,5 y 12,5% fueron obtenidos en pino marítimo para la genotipificación de una población de polinización abierta y una población endocriada, respectivamente (Chancerel et al. 2011)). La tasa de conversión más alta descrita fue del 28,6% obtenida para la población de mapeo de QTLs en la conífera *Pinus lambertiana* (Jermstad et al. 2010)). Aunque este último valor fue superior, más del 35% de los SNPs polimórficos obtenidos presentaron una distorsión en la segregación significativa.

Al contrastar el éxito de la genotipificación en los SNPs clasificándolos por el método empleado para su identificación, se observó un menor porcentaje de SNPs *in vitro* que presentaron fallas en la genotipificación (16,5% frente a 30% de los SNPs *in silico*).

Las fallas en el genotipado de SNPs *in silico* son frecuentes, y podrían deberse a la baja calidad de las secuencias flanqueantes a los SNPs, a la presencia de SNPs no detectados en dichas secuencias, o a la presencia de una unión exón-intrón en las proximidades del SNP seleccionado (Wang et al. 2008; Lepoittevin et al. 2010).

La presencia de complejidad en el genoma de girasol, representada principalmente por secuencias parálogas y genes multicopia, puede ser un factor limitante en el éxito del ensayo (Akhunov et al. 2009; Sandve et al. 2010; Shen et al. 2005; Fan et al. 2006). Choi et al. (2007), analizando 6 genotipos en soja, describieron fallas en la genotipificación de SNPs cuando los iniciadores se diseñaron sobre los extremos 3' de las secuencias ESTs o cuando el diseño provino de secuencias de baja calidad como STS (del Inglés: "*Sequence-tagged sites*"). Los mismos concluyeron que la amplificación de múltiples productos de PCR puede deberse

probablemente a la detección de secuencias parálogas. En cambio, Shirasawa et al (2010) sostienen que la falla en genotipificación de los SNPs definidos *in silico*, también puede deberse a la existencia de genes de múltiples copia, ocasionando distorsiones de la señal fluorescente e incorrecta clusterización a la hora de asignar los alelos.

Por otra parte, al considerar la tasa de conversión, un mayor porcentaje de SNPs *in vitro* resultaron polimórficos cuando se comparan ambos grupos (26,8% *in vitro* frente a 13,24% obtenido para SNPs *in silico*). Estos porcentajes son similares a los obtenidos por Chancerel et al (2011) para las dos poblaciones de pino marítimo ya descritas.

Una consideración que debe ser tomada en cuenta en la valoración de los parámetros analizados para definir la calidad del ensayo SOPA es que a diferencia de la tasa de éxito, que es una medida inherente del ensayo, la tasa de conversión puede variar considerablemente dependiendo de la población utilizada. Como consecuencia, el número de SNPs polimórficos obtenidos utilizando poblaciones biparentales puede ser relativamente más bajo para los ensayos cuyos SNPs no fueron seleccionados *a priori* por su polimorfismo en las líneas parentales.

En cuanto a la proporción de datos faltantes, nuevamente los SNPs seleccionados *in vitro* mostraron una menor proporción cuando se comparan con SNPs obtenidos *in silico*. Sin embargo, considerando la totalidad de SNPs polimórficos, el porcentaje total de los datos faltantes es muy bajo (menos de 1,3%). Este hecho, sumado a la muy buena reproducibilidad de las medidas obtenida mediante la corroboración del 100% de identidad entre los genotipos replicados de PAC2 y RHA266 que se usaron como control en el ensayo, muestran la solidez del ensayo de genotipificación Veracode (Illumina) y de los datos generados.

#### 4. MAPA DE LIGAMIENTO

En este trabajo se incorporaron marcadores SNP asociados a genes candidato para resistencia a estreses en girasol a la matriz de genotipos de una población pública (Al-Chaarani et al. 2004; S. Kiani et al. 2007; Paola Talia et al. 2010; Zubrzycki 2009) que es tomada como referencia tanto para localización de marcadores como para estudios de QTL por INRA (Francia) e INTA (Argentina). La matriz completa sumó un total de 739 marcadores. A partir del análisis de la segregación de los alelos con la prueba estadística FDR, se descartaron cinco marcadores que presentaban distorsión de la segregación dado que este tipo de datos influye negativamente en la construcción del mapa genético (Zhang et al. 2002) y en la detección de QTLs ocasionando asociaciones espurias

Distorsiones de este tipo han sido reportadas previamente en poblaciones de girasol y en otros cultivos (Leon et al. 2000; Vogl & Xu 2000; Winter et al. 2000) y pueden deberse a errores en la genotipificación de los marcadores, a artefactos de la PCR o a algún tipo de selección que haya ocurrido durante el desarrollo de las poblaciones (Leon et al. 2000).

En una segunda etapa se eliminaron 21 individuos y 28 marcadores que poseían más del 50% de datos faltantes. Esta decisión redujo sustancialmente el porcentaje de datos perdidos totales a un valor final de 17,2%. El hecho de no obtener amplificación para estas regiones puede asociarse a problemas en la reacción de PCR debido a la calidad de ADN, sin embargo también hay que considerar la existencia de alelos nulos que en girasol ha sido reportada con una frecuencia de aproximadamente 9% (Yu et al. 2002; Tang et al. 2002). Se presume que la presencia de alelos nulos se debe a polimorfismos en las zonas de hibridación de los oligonucleótidos (Mogg et al. 2002).

El presente trabajo describe la construcción de un mapa de ligamiento genético optimizado para la población PAC2 x RHA266 que incluyó 566 marcadores previamente mapeados (327 AFLPs, 229 SSRs, 8 EST-SSRs, 2 InDels) y 69 genes candidatos (61 HeAn\_x genotipados mediante el ensayo SOPA y 8 SNPs genotipados con dHPLC). La inclusión de SNP permitió que 26 marcadores AFLPs y 21 SSRs incluidos en la matriz original que no habían podido ser incorporados al mapa, se localizaran con valores estadísticos significativos sobre todos los GLs, excepto sobre los GLs 11 y 13.

Por otro lado, 71 marcadores, de los cuales cinco son SNP pertenecientes al ensayo SOPA, no pudieron ser agrupados bajo los parámetros de mapeo implementados en este trabajo.

El tamaño del mapa obtenido fue de 2823,80 cM, extendiendo la versión del mapa en 900 cM respecto del publicado por Talia et al (2010) utilizando un grupo de 94 RILs de la misma población. Comparado con otros mapas de ligamiento desarrollados para girasol, el presente mapa genético es más grande respecto de los construidos a partir de 94 RILs derivadas de la cruce entre RHA280 x RHA801 basado en SSR/InDels (Tang et al. 2002) o del elaborado usando SSRs y SNPs, descrito por Lai 2005, cuyas longitudes fueron 1368,3 y 1349,3 cM respectivamente.

Cabe destacar que la comparación entre mapas ha de considerarse con la debida cautela, ya que el uso distintos programas informáticos empleados para su construcción (JoinMap, Carthagene, Mapmaker, R/qtl) como el ajuste de los diferentes parámetros (LOD, frecuencia de recombinación, etc) pueden influir determinadamente sobre el mapa resultado. Además,

esta variación puede ser resultado de la combinación de diferentes tipos de marcadores que muestran patrones de distribución diferentes (Sebastian et al. 2000; Mei et al. 2004; Syed et al. 2006; Zhang et al. 2006).

Otros efectos a considerar como el tamaño de la población, la agrupación de marcadores, los efectos distorsionadores que aporta la segregación de marcadores o procesos involucrados en la generación de un mapa robustos con pocos marcadores que sirvan de anclaje, puede contribuir en dichas diferencias (Cervera et al. 2001; Zhang et al. 2006).

La incorporación de los 69 nuevos marcadores se dio a lo largo de todo el genoma. La localización sobre el GL 13 del gen candidato CAM, una calmodulina (Liu & Burke 2006), permitió llenar un intervalo de 23 cM. Además, el gen candidato asociado a la biosíntesis de sacarosa, PGIC (Liu & Burke 2006), se ancló a un pequeño grupo de marcadores situados en la región distal del LG9 en el mapa obtenido por Tang et al. (2003), con un subgrupo de marcadores en GL 9 obtenidos por Kiani et al. (2007) y HeAn\_C\_4023.2, resultado de la genotipificación masiva del panel de 384 SNPs (Figura 1.12). SCR1 ha sido descrito como un factor de transcripción involucrado en la división asimétrica de la planta (Liu & Burke 2006). Este se localizó sobre el GL 6, muy próximo a la ubicación del HeAn\_R\_326.1, indicando un posible clúster de genes. Por otra parte, el gen candidato LIM pudo ser ubicado en GL 10, donde se encontraron seis QTLs que explican el 13% de la resistencia de girasol a *S. sclerotiorum* hongo patógeno (Maringolo 2007). Este gen es un factor de transcripción putativo que podría estar implicado en el mecanismo de defensa, aunque también se describe como involucrado en la formación del citoesqueleto (Arnaud et al. 2007).

Por su parte, los 11 de los GCs (HeAn\_R\_149.1, HeAn\_C\_10204.1, HeAN\_R\_RGEF\_A1.2, GLP5, HeAn\_C\_4023.2, PGIC, HeAn\_C\_11741.1, HeAn\_R\_204.1, HeAn\_C\_12910.1, HeAn\_R\_126.2 y HeAn\_C\_5433.3) se ubicaron en las regiones distales de los GL 2, 4, 6, 9, 10, 11 y 17, contribuyendo de esta manera al aumento en el tamaño del mapa obtenido.

El valor de cobertura de mapa fue de 1 marcador en promedio cada 4,45 cM, que es equiparable al de los mapas publicados por Tang et al (2003) (3,1), Talia et al (2010) (3,6), Alchaarani et al (2004) (7,9). Por otra parte, el número de intervalos entre marcadores superiores a 25 cM observados fue de 3, indicando buen nivel de saturación para los objetivos propuestos.

El mayor intervalo sin marcadores fue detectado en el GL 17 (44 cM) y es factible que se haya originado por la localización del gen candidato HeAn\_C\_5433.3.

En cuanto al orden en el que se posicionaron los cromosomas, para el presente trabajo se tomó como referencias inalterables para el mapeo de marcadores la posición de los marcadores en trabajo desarrollado por Talia et al (2010). De esta manera se generó un mapa similar en cuanto a la disposición de los marcadores ya determinados en los mapas públicos y se evitó las diferencias en el orden de los marcadores causadas por variaciones locales en la frecuencia de recombinación, por la segregación de marcadores específicos en un solo parental y por rupturas en la sintenia de los parentales.

En este trabajo se observa una amplia distribución de los marcadores en los GLs. Si bien los marcadores mapeados se observan homogéneamente distribuidos en algunas regiones del genoma y agrupados en otras, el fenómeno de formación de clusters fue mínimo de tal modo que no obstaculizó el desarrollo del mapa genético. Esta observación coincide con lo reportado por Flores Berrios et al. (Berrios et al. 2000) utilizando la misma población que en este trabajo y marcadores AFLPs, y con lo descrito en el mapa genético de girasol desarrollado por Tang et al. (Tang et al. 2002). Estos autores reportaron que los marcadores microsatélites se encontraron distribuidos formando un denso cluster en los GL 4, 10 y 17 y un moderado agrupamiento en los GL 2, 3, 5, 13 y 15.

El fenómeno de formación de agrupamientos o clusters ha sido atribuido a la frecuente asociación de los SSR con retrotransposones y con otros elementos de ADN repetitivo, a la alta densidad de dichos elementos en regiones centroméricas y a la reducida recombinación en las regiones centroméricas (Areshchenkova & Ganal 1999; Cuadrado & Schwarzacher 1998; Künzel et al. 2000; Ramsay et al. 1999; Ramsay et al. 2000). Este fenómeno se ha observado en varias especies, tales como algunas coníferas y cebada (Paglia et al. 1998; Ramsay et al. 2000), pero no en otras como soja y arroz (Cregan et al. 1999; Temnykh et al. 2000).

En girasol, estudios de hibridación in situ, usando como sondas secuencias repetidas derivadas de retroelementos de *H. annuus*, permitieron la detección de secuencias de las familias de retroelementos Ty1/copia-like y Ty3/gypsy-like dispersas a lo largo de todos los cromosomas en todas las especies estudiadas. Sin embargo, las secuencias del retrotrasposón Ty3/gypsy-like fueron localizadas preferencialmente en las regiones centroméricas en la mayoría de la especies estudiadas, mientras que las secuencias Ty1/copia-like fueron menos representadas o ausentes alrededor del centrómero y más abundantes en los extremos de los cromosomas en *H. annuus* (Natali et al. 2006; Santini et al. 2002; Natali et al. 2013; Staton et al. 2012).

El desarrollo de mapas genéticos es el primer paso hacia la detección de factores que controlan la expresión de características de importancia agronómica. La identificación de marcadores o genes ligados a QTLs que controlan características de importancia agronómica, permite su utilización en programas de mejoramiento asistido por marcadores, facilitando la selección de genotipos superiores en forma eficiente, reduciendo el tiempo y el espacio necesarios. En este contexto la información generada por el mapa de ligamiento presentado en este trabajo, es útil para la identificación y validación de diversos QTLs. De esta manera este mapa sirve como punto de partida a futuros trabajos del grupo de investigación, tendientes a la identificación de regiones que controlan características de importancia agronómica y el análisis comparativo de genomas de especies relacionadas.

# CAPÍTULO 2

**Estudio de la resistencia parcial del girasol  
a la Podredumbre Húmeda de Capítulo  
causada por *Sclerotinia sclerotiorum***

# INTRODUCCIÓN

## 1. CARACTERIZACIÓN GENÉTICA DE LA RESISTENCIA CUANTITATIVA A ENFERMEDADES

En los últimos años los principales objetivos del mejoramiento de girasol en el mundo han sido la resistencia a estrés, fundamentalmente a los producidos por las enfermedades más comunes y de más impacto sobre el cultivo, y a factores abióticos como la sequía, considerando la importancia de disponer de genotipos adaptados a los desafíos que enfrenta el cultivo relativos al desplazamiento hacia regiones marginales y los efectos del cambio climático (Paniego et al. 2012).

En el caso de la resistencia a estrés biótico, uno de los objetivos más extendidos es la de identificar resistencias a enfermedades fúngicas, siendo marchitez por *Verticillium* (*Verticillium dahliae*), mildiú (*Plasmopara halstedii*), podredumbre húmeda del capítulo y tallo (*Sclerotinia sclerotiorum*), roya negra (*Puccinia helianthi*), mancha de la hoja (*Alternaria helianthi*), entre otras las que más impacto tienen sobre el cultivo en distintas regiones de Argentina y en el mundo ((Virányi 2008; Paniego et al. 2002).

Existen diversas fuentes de resistencia en girasol presentes en especies silvestres, variedades tradicionales y poblaciones de polinización abierta, que han sido introgresadas en el germoplasma cultivado (Seiler 2010; Gulya et al. 2010), no obstante existen aún fuentes inéditas que deben ser exploradas y caracterizadas genéticamente para ser incorporadas al germoplasma cultivado aumentando la diversidad de la respuesta de resistencia a patógenos y las posibilidades de superar las limitantes producidas por este factor en el cultivo.

La resistencia a enfermedades en plantas ha sido categorizada históricamente en dos grupos, la resistencia incompleta asociada a muchos genes de efecto parcial, y la resistencia completa asociadas a un gen simple, ambos tipos de resistencias son distinguibles inequívocamente en sus formas extremas (Vanderplank 1969; Poland et al. 2009). En los últimos años, esta categorización tan contundente se encuentra siendo revisada a partir de la avance de los conocimientos sobre mecanismos de resistencia en plantas (Poland et al. 2009). La primera respuesta de resistencia está mediada por receptores de superficie celular y receptores intracelulares de reconocimiento de patrones (PRRs: *pattern recognition receptors*), que perciben elicitores (o inductores de respuesta) microbianos conservados, denominados

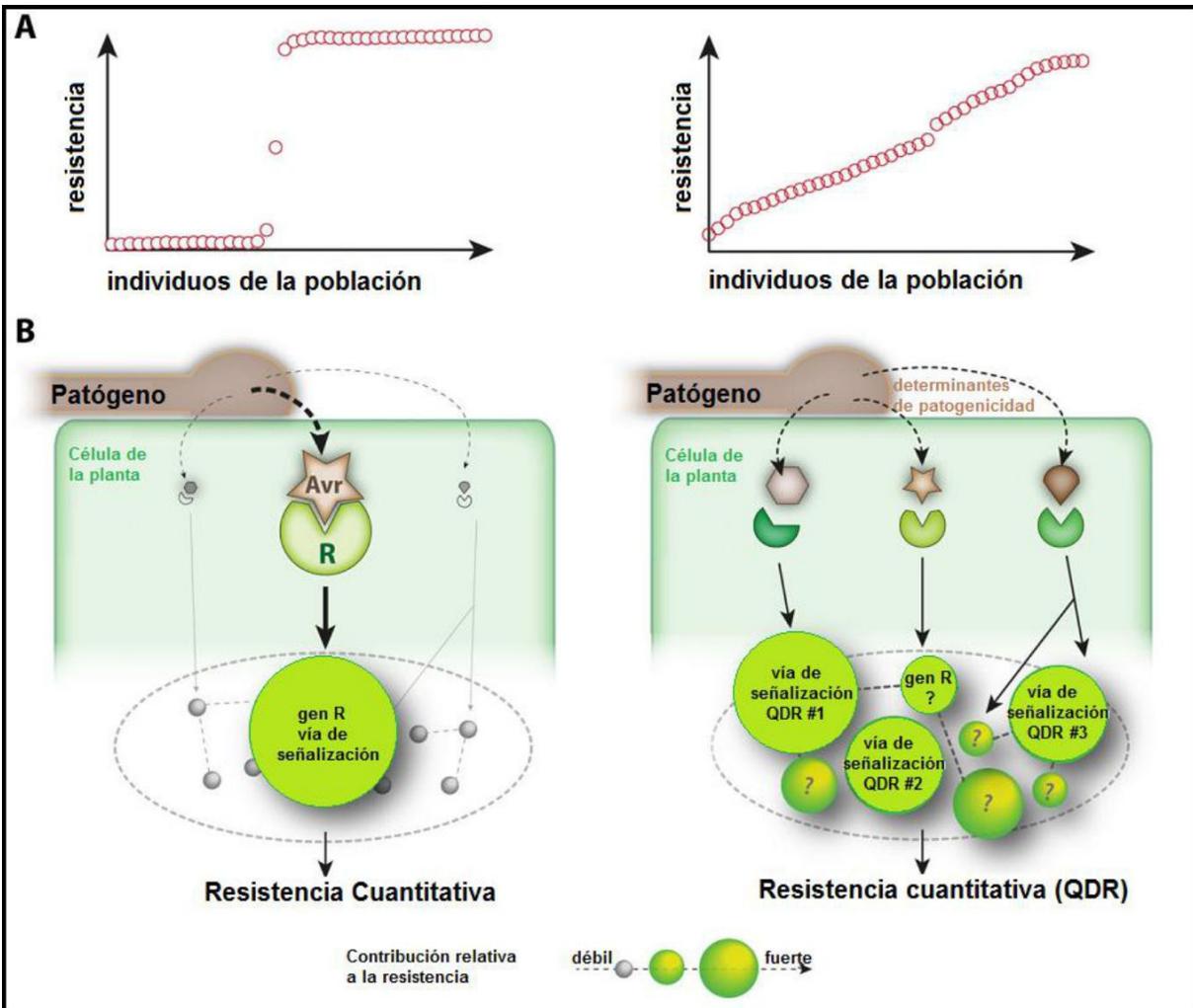
patrones moleculares asociados a patógenos (PAMPs: *pathogen-associated molecular patterns*) (Poland et al. 2009; Roux et al. 2014; Nimchuk et al. 2003).

La percepción de estos elicitores inicia cascadas de eventos de señalización y de transcripción, conocido como inmunidad producida por PAMPs (PTI: PAMP-triggered immunity). Ciertos patógenos adaptados secretan moléculas efectoras capaces de suprimir la PTI, pero que también pueden ser reconocidas por proteínas resistentes (R). Esto da inicio a la inmunidad activada por efectores (ETI: effector-triggered immunity) o defensa cualitativa de la planta, y típicamente produce fenotipos de resistencia completa contra la enfermedad. Por el contrario, el reconocimiento de un solo PAMP tiene típicamente una contribución más débil a la resistencia global de la planta (Nimchuk et al. 2003). En general, en ausencia de resistencia cualitativa, el fenómeno resistencia incompleta se observa más a menudo, lo que lleva a una reducción de la enfermedad en lugar de ausencia de la misma. Es por esta razón que se suele hacer referencia a la resistencia a enfermedades como cuantitativa (QDR: Quantitative disease resistance)(Poland et al. 2009) (Figura 2.1).

La definición de QDR está basada en la ocurrencia de una distribución continua de sus caracteres de resistencia, por lo que, dentro de una población determinada, la identificación de clases fenotípicas es extremadamente difícil. La base genética de estos caracteres está conformada por muchos genes de efectos pequeños o poligenes. Geldermann (1975) propuso denominar a los *loci* controladores de estos caracteres cuantitativos como QTLs (QTL – *Quantitative Trait Loci*). La arquitectura molecular de una QDR entonces, puede deberse a una red intrincada, integración de múltiples vías de respuesta a varios determinantes moleculares patogénicos y señales ambientales (Mukhtar et al. 2011).

Si bien, con base en los resultados fenotípicos aparentemente distintos se tiende a establecer a la QDR y a la resistencia cualitativa como mutuamente excluyentes, diversos factores pueden introducir algún grado de variación cuantitativa en PTI y ETI (Vetter et al. 2012). Entre ellos, el ambiente, el tamaño de la población, la precisión y la resolución de la metodología de fenotipificación, los genotipos de patógenos, mutaciones naturales en los genes R y PRRs, pueden alterar la fuerza de resistencia conferida (Poland et al. 2009; Lorang et al. 2007; McDonald & Linde 2002; Mukhtar et al. 2011).

Por lo tanto, la resistencia de las plantas a enfermedades es intrínsecamente dual, con mecanismos y componentes moleculares QDR específicos, y sus conexiones con otras vías de señalización, aún no descubiertas (Poland et al. 2009).



**FIGURA 2.1. MECANISMOS DE RESISTENCIA A ENFERMEDADES EN LAS PLANTAS.** La figura ilustra la naturaleza cualitativa (izquierda) y cuantitativa (derecha). A) Los fenotipos de resistencia a la enfermedad sigue una distribución binaria 'susceptible o resistencia' para la resistencia cualitativa, y una distribución para la resistencia cuantitativa. B) La resistencia cualitativa depende de la percepción de un solo efector (Avr) por parte de genes R de la planta, mientras que la resistencia cuantitativa depende de la integración de múltiples vías de percepción activas simultáneamente, cada uno con una contribución relativamente menor al fenotipo de resistencia. Adaptado de Roux et al (2014).

Los receptores conocidos que intervienen en la resistencia normalmente pertenecen a dos clases: Los PRRs que frecuentemente contienen, o están asociados con quinasas no-RD (sin arginina ni aspartato)(Huard-Chauveau et al. 2013), y genes R con regiones de unión a nucleótidos con repeticiones ricas en leucina (NBS-LRR: nucleotide binding site-leucine-rich repeat) y carentes de dominios quinasa (Nimchuk et al. 2003).

Aunque numerosos QTLs de resistencia a enfermedades han sido identificados en plantas, los genes y mecanismos subyacentes a QDRs siguen siendo en gran medida desconocidos. Como para la mayoría de los rasgos cuantitativos, la identificación de genes QDR se ha centrado en

QTLs de gran efecto (explicando más del 20% de la varianza fenotípica), queda por 78

determinar si los genes QDR de efecto menor ampliarán la gama de funciones moleculares asociados con la resistencia a enfermedades. Además, aunque se ha encontrado interacción entre vías PTI y ETI, las conexiones de estas con las QDRs siguen siendo escasas y la ubicación de los componentes identificados en las vías de resistencia sigue siendo enigmática.

ETI se basa en el reconocimiento específico de efectores de patógenos altamente variables por el gen R correspondiente, por lo general funciona solamente contra los aislamientos específicos de una especie de patógenos. En algunos casos raros, son eficaces contra múltiples cepas y especies, tal como el gen *RPW8*, que confiere resistencia contra múltiples patógenos en *Arabidopsis* (Xiao et al. 2001). Por el contrario, los genes QDR suelen proporcionar resistencia de amplio espectro. El transportador ABC *Lr34* es un ejemplo de ello, que confiere resistencia en el trigo tanto contra patógenos fúngicos ascomicetes como basidiomicetes (Krattinger et al. 2009). Además, QDR ha sido observado frente a patógenos virales, bacterianos y eucariotas (en particular, hongos, nematodos y patógenos oomycete) (Lacombe et al. 2010). Una vez más, la resistencia de amplio espectro que ofrecen los genes QDR puede estar relacionada con las vías de percepción y respuesta a patrones microbianos bien conservados (Poland et al. 2009).

El reconocimiento de los efectores de patógenos por proteínas vegetales NB-LRR se asocia mayormente con la resistencia a patógenos biotróficos, aunque puede desencadenar específicamente la susceptibilidad a algunos necrótrofos (Lorang et al. 2007).

Por otra parte, entre los numerosos QTLs de resistencia contra diversos patógenos necrotrofico y hemibiotróficos que han sido reportados, algunos son también eficaces contra patógenos biotróficos obligados. Es el caso, por ejemplo, del gen *Lr34* del trigo, que confiere resistencia a patógenos hemibiotróficos, así como patógenos biotróficos obligados como *Blumeria graminis* (Krattinger et al. 2009). La mayoría de los agentes patógenos contra los cuales la QDR es la forma predominante de resistencia de la planta, son capaces de infectar una amplia gama de especies hospedadoras. Son ejemplo de ello *Sclerotinia sclerotiorum* (Perchepped et al. 2010) y *Botrytis cinerea* (Denby et al. 2004), cada uno infectando varios cientos de especies de plantas.

De esta forma, la eficiencia de la resistencia está basada en genes R y PRRs con múltiples alelos funcionales, donde las diferencias por SNPs e InDels son los principales patrones de variación. SNPs sobre genes R y loci PRRs pueden introducir diferencias cuantitativas en el fenotipo de resistencia. La naturaleza y la posición de SNPs en un locus dado pueden conducir,

por lo tanto una pérdida completa de la resistencia (resistencia cualitativa) o variaciones cuantitativas en el nivel de resistencia (QDR). Por lo tanto, SNPs e InDels son polimorfismos que pueden regir la resistencia cualitativa o cuantitativa (Mukhtar et al. 2011; Roux et al. 2014).

Otros patrones de variación genética que parecen estar asociado con genes QDR (Cook et al. 2012) son las variantes en el número de copias del gen (CNV: copy number variation).

Manosalva et al 2009, asociaron el silenciamiento de un número de copias de un gen dentro de un clúster ubicado en un QTL que codifica para una proteína del tipo germina con QDR a *Rhizoctonia solani* en arroz.

Genes duplicados y mecanismos de regulación en *cis* y *trans* pueden introducir variaciones cuantitativas en las vías de señalización de resistencia. La relación entre la abundancia de transcritos y el fenotipo QDR se ha establecido recientemente a través de estudios de la abundancia de transcritos en el genoma completo y mapeo de QTL de expresión (eQTL) (Chen et al. 2010).

## 2. FENOTIPIFICACIÓN

El mapeo de *loci* de resistencia cuantitativa típicamente produce múltiples QTLs de pequeño a moderado efecto, explicando entre 5% y 30% de la varianza fenotípica (Collard et al. 2005). Esta complejidad plantea retos importantes para la caracterización funcional de QTLs de resistencia.

En primer lugar, es necesario medir de manera precisa la contribución fenotípica de los genes QDRs individuales que subyacen a un QTL y luego mapearlo con precisión en el genoma. El progreso tecnológico reciente sobre las plataformas de genotipificación basados en genómica de nueva generación y las plataformas automatizadas de fenotipificación (McCouch et al. 2010), ofrecen perspectivas prometedoras en estos dos campos.

Son múltiples los caracteres cuantitativos que pueden medirse para caracterizar una QDR en una interacción huésped-patógeno, como el índice de enfermedad, tamaño de la lesión, período de incubación (generalmente en patógenos filamentosos), el número de unidades formadoras de colonia (patógenos bacterianos), ensayos inmunoenzimáticos (patógenos virales), etc. La elección de una característica verdaderamente cuantitativa, que se pueda medir con precisión y fiabilidad, puede ser esencial tanto para la precisión del mapeo de QTL como para la validación de la función del gen QDR (Roux et al. 2014).

Cuanto más se disecciona el rasgo QDR, podemos esperar un acceso más directo a sus bases genéticas. Definiendo la ruta más corta entre las bases genéticas de la QDR y la característica medida, el uso de marcadores moleculares, debería ampliar la varianza fenotípica explicada por genes QDR. La evaluación simultánea de múltiples caracteres QDR debería aumentar aún más nuestra capacidad de cuantificar con precisión la contribución fenotípica de genes QDR. Por tanto, existe necesidad de un conocimiento detallado de las interacciones planta-microorganismo a nivel subcelular y molecular. La comprensión de cómo los patógenos manipulan la expresión génica del huésped y la arquitectura de la célula huésped puede, por ejemplo, proporcionar lecturas precisas para la evaluación de caracteres (Bergelson & Roux 2010).

### 3. MÉTODOS DE MAPEO DE QTLS

El primer paso para el mapeo genético de un locus que controla un carácter cuantitativo o QTL, es la existencia de una asociación estadística significativa entre la segregación del carácter y la segregación de un marcador genético en una población determinada (Jensen 1989). La identificación de QTLS basada únicamente en la evaluación convencional de los caracteres no es posible. Actualmente, la utilización de mapas de ligamiento genético de marcadores moleculares, obtenido de poblaciones segregantes biparentales, permite identificar el genotipo de los QTLS, su posición genómica, sus efectos y contribución específica en la determinación del carácter. El proceso de construcción de mapas de ligamiento y la realización de análisis de QTL para identificar regiones genómicas asociadas con rasgos que se conoce como mapeo de QTLS (McCouch & Doerge 1995; Mohan et al. 1997). El procedimiento básico y simplificado de análisis de QTLS para el carácter en cuestión es similar al seguido para la construcción del mapa del mapa de ligamiento. Entre los métodos usados en el mapeo de QTLS, el denominado mapeo por intervalos simples (SIM: simple interval mapping, (Lander & Botstein 1989)) significó un gran avance metodológico debido a su mayor poder de detección comparada con la metodología de asociación por marcas simples. Sin embargo, este método mapea QTLS de a uno por vez sin considerar otros QTLS (mapeados o no), por lo que el efecto de estos últimos es equivalente al efecto ambiental, obteniéndose una reducción en el valor de LOD. Como consecuencia, el poder de detección puede estar comprometido, y las estimaciones de la localización y efectos de los QTLS pueden estar sesgadas (Jansen, 1993). Por lo tanto, puesto que un carácter cuantitativo por definición está controlado por varios genes, este método tiene escaso valor. SIM, es el principio en que se basan todos los demás métodos desarrollados posteriormente. Entre ellos, el mapeo por

intervalos compuestos (CIM: composite interval mapping) desarrollado independientemente por diferentes autores (Zeng 1994; Zeng 1993; Jansen & Stam 1994) que consiste en fijar modelos de un único QTL (uno por cada intervalo) y el uso de marcadores seleccionados como cofactores para la regresión del fenotipo sobre el genotipo. De ese modo se consigue eliminar los efectos de posibles QTLs en otras partes del genoma. Este último es uno de los más utilizados y permite no sólo el mapeo de la región genómica involucrada en el control del carácter, sino también los genes efectivamente involucrados en dicho control, si el mapa genético posee buen grado de saturación y, si además, marcadores génicos o funcionales fueron utilizados en su construcción. La identificación de un gen puede permitir la identificación de vías regulatorias o bioquímicas responsables de influenciar las variaciones en ese carácter.

De esta manera, en los intervalos entre pares de marcadores adyacentes, se calcula el valor de LOD para presencia de un locus influyendo en el carácter en estudio (probabilidad de la presencia de un QTL frente a la probabilidad de la ausencia del mismo). Si el valor de LOD obtenido supera un valor umbral crítico, un QTL puede declararse como “real”. El resultado es una valoración de las regiones cromosómicas afectando este carácter (De Vienne 2003).

La significación estadística de los QTLs se obtiene calculando un valor umbral de LOD mediante diferentes métodos (Doerge & Churchill 1996). La mayoría de los programas informáticos utilizan el método empírico desarrollado por Churchill y Doerge (1994) basado en la teoría de la permutación. Con el fin de destruir posibles relaciones entre los valores fenotípicos y genotípicos se permutan aleatoriamente los datos cuantitativos entre los individuos de la población de mapeo. Este método es válido para cualquier diseño experimental, sin exigencias en cuanto al tamaño poblacional, y con cualquier densidad de mapa, pudiendo ser aplicado a datos que no siguen una distribución normal (Lander & Kruglyak 1995). La precisión con la que se calculan los valores umbrales estará determinado por el número de permutaciones. Un número de 1.000 permutaciones se considera adecuado para un nivel de significación  $\alpha = 0,05$  (Doerge & Churchill 1996).

#### **4. GENES Y QTLs EN GIRASOL**

El estrés biótico es considerado uno de los factores más importantes que afectan girasol en todo el mundo. El girasol cultivado es susceptible de varias enfermedades fúngicas y bacterianas, siendo la gravedad de la infección dependiente en gran medida de las condiciones ambientales. Entre las enfermedades que afectan a la mayor parte del girasol zonas de cultivo

se incluyen la marchitez, podredumbre del tallo y podredumbre del capítulo (causada principalmente por *Sclerotinia sclerotiorum*), mildiu (*Plasmopara halstedii*), el cancro del tallo (*Phomopsis helianthi*), roya (*Puccinia helianthi*) y marchitez por *Verticillium* (*Verticillium dahliae*).

Una contribución relevante para el estudio de los factores de resistencia a enfermedades de girasol es el trabajo exhaustivo realizado por Radwan et al., (2008) de mapeo de 167 loci vinculados con genes R NBS-LRR a lo largo de los 17 GLs. Dichos loci NBS-LRR se distribuyeron principalmente en clúster o singletons en el extremo superior e inferior de la mayoría de los grupos de ligamiento. Radwan et al. (2008) describieron dos grandes conjuntos en los GLs 8 y 13, mientras que conjuntos más pequeños fueron mapeados en los GLs 1, 4, 9 y 15. Curiosamente, la mayoría de los genes R mapeados en girasol hasta ahora, se encuentran en los mismos GLs (Tabla 2.1).

**TABLA 2.1. Genes de resistencia a diferentes enfermedades en girasol.**

CARÁCTER	GEN	UBICACIÓN	FUENTE BIBLIOGRÁFICA
<b>Resistencia a mildiu (<i>Plasmopara halstedii</i>)</b>	<i>Pl1</i>	GL 8	Zimmer & Fick, 1974; Gedil et al., 2001; Bertero de Romano et al., 2010
	<i>Pl2</i>	GL 8	Zimmer & Fick, 1974; Bertero de Romano et al., 2010
	<i>Pl4</i>	?	Sackston, 1990
	<i>Pl5</i>	GL 13	Bert et al., 2001; Radwan et al., 2004
	<i>Pl6</i>	GL 8	Miller & Gulya, 1991; Bouzidi et al., 2002; Vear et al., 2003; Panković et al., 2007
	<i>Pl7</i>	GL 8	Miller & Gulya, 1991; Bert et al., 2001;
	<i>Pl8</i>	GL 13	Miller & Gulya, 1991; Radwan et al., 2004; Bachlava et al., 2011
	<i>Pl9</i>	?	Sackston, 1990; Molinero-Ruiz et al., 2003
	<i>Pl13</i>	GL 1	Mulpuri et al., 2009; Bertero de Romano et al., 2010
	<i>Pl14</i>	GL 1	Bachlava et al., 2011
	<i>Pl15</i>	GL 8	Bertero de Romano et al., 2010
	<i>PlArg</i>	GL 1	Dussle et al., 2004; Wieckhorst et al., 2010; Imerovski et al., 2011
<b>Resistencia a roya negra (<i>Puccinia helianthii</i>)</b>	<i>R1</i>	GL 8	Lawson et al., 1996
	<i>R<sub>adv</sub></i>	GL 13	Lawson et al., 1998; Bachlava et al., 2009
	<i>R4</i>	GL 13	Qi et al., 2011; Miller et al., 1988
	<i>R2</i>	GL 9	Lawson et al., 2010
	<i>R3</i>	?	Goulter 1990
	<i>R5</i>	GL 2	Qi et al., 2011
	<i>Pu6</i>	?	Yang 1989
	<i>R11</i>	GL 13	Qi et al., 2012
<i>R12</i>	?	Gong et al., 2013	
<b>Marchitez por <i>Verticillium dahliae</i> Podredumbre Húmeda del Capítulo (<i>S. sclerotiorum</i>)</b>	<i>V1</i>	?	Gulya, 2007
	<i>HaRIC_B</i>	GL 11	Fusari et al., 2012

En relación con la resistencia del cultivo a QDRs como las ocasionada por *Alternaria helianthi* y *Sclerotinia sclerotiorum*, se han realizado varios estudios basados en la búsqueda de QTLs. Para la primera, tres QTLs de efecto mayor fueron localizados sobre los GLs 4, 10 y 17 que muestra efectos aditivos significativos (Galella et al. 2010). En cuanto a *S. sclerotiorum*, objeto de estudio en el presente trabajo de Tesis, numerosos trabajos han sido publicados. Mestries et al, (1998) identificaron 2 QTLs para la resistencia a la PHC basados en un mapa genético con marcadores RFLPs e isoenzimas. Cada QTL individual explicó entre 9% y 48% de la variabilidad fenotípica, confirmando la naturaleza poligénica de este carácter. Simultáneamente, Gentzbittel et al, (1998) informaron de varios QTLs para resistencia a la penetración o la extensión del micelio de *S. sclerotiorum* en diferentes tejidos. Estos últimos autores también encontraron una proteína serin-treonina kinase-like (PK) que cosegrega con un QTL en el GL 1, aparentemente específico de la línea endocriada PAC1. Este QTL explicó el 50% de la variación en una población derivada de la cruce de PAC1 con una línea muy susceptible y el 15% en una cruce con una línea altamente resistente. Sin embargo, este QTL no se ha encontrado en otras poblaciones, incluso cuando la fuente de resistencia provenía de la selección recurrente de una población originada por PAC1 y cuando los padres mostraron polimorfismo para el gen PK.

Luego, Bert et al, (2002) detectaron 15 QTLs para la reacción a *Sclerotinia* a lo largo de varios grupos de ligamiento, mientras que Bert et al, (2004) describieron siete QTLs para la resistencia al ataque de *S. sclerotiorum* a la yema terminal. Cada uno de ellos explicó menos de 10% de la variación fenotípica. En cuanto a resistencia a PHC, detectaron 4 QTLs, cada uno explicando hasta el 20% de la variación. Todo los QTLs encontrados, fueron producto del análisis de seis cruzamientos diferentes entre líneas puras cultivadas y se encontraron distribuidos en por lo menos 14 grupos de ligamiento.

Micic et al, (2004) identificaron nueve QTLs para la lesión de la hoja, ocho para la lesión del tallo, y seis para la velocidad de crecimiento del hongo. Para la lesión del tallo dos regiones genómicas explicaron aproximadamente el 26,5% de la varianza genotípica (Micic, Hahn, Bauer, Melchinger, et al. 2005). De tres a cuatro QTLs putativos se detectaron para la lesión del tallo, lesión de la hoja y la velocidad de crecimiento de los hongos que explican entre 40,8% y 72,7% de la varianza genotípica. Dos QTLs para la lesión del tallo mostraron grandes efectos genéticos y corroboraron los hallazgos anteriores obtenidos del cruzamiento de NDBLOSsel (resistente) x CM625 (susceptible) (Micic, Hahn, Bauer, Schön, et al. 2005).

Rönicke et al, (Rönicke et al. 2005) por su parte, identificaron cinco QTLs para la longitud de la lesión y PHC explicando 10,6 a 17,1% de la varianza fenotípica total. Los QTLs más frecuentes fueron asociados al gen recesivo de ramificación *b1* en el grupo de ligamiento 10. Este hecho se ratifica con lo observado por Vear et al (Vear et al. 2008; Vear & Grezes-Besset 2010) quienes notaron que las líneas ramificadas poseen un mejor comportamiento frente a este patógeno que las líneas no ramificadas concluyendo que hay una asociación genética entre un QTL para la resistencia y el gen de ramificación en lugar de un efecto directo del fenotipo ramificado.

En Argentina se han identificado QTLs para resistencia a la PHC en líneas de buen comportamiento local, los cuales han sido localizados fundamentalmente en el grupo de ligamiento 10 del mapa genético de girasol (Maringolo 2007).

Yue et al, (2008) identificaron nueve y siete QTLs para la incidencia y la severidad de la enfermedad, respectivamente. Cada QTL explicó entre 8,4 y 34,5% de la varianza fenotípica, lo que sugiere la base poligénica de la resistencia a la pudrición del capítulo. Cinco de estos QTLs se identificaron en más de un experimento, y cada uno de ellos explicó más de 12,9% de la varianza fenotípica. Aunque existía una correlación positiva entre los dos índices de la enfermedad, la mayoría de los QTLs se encontraban en diferentes regiones cromosómicas, lo que nuevamente sugiere una base genética diferente para ambos fenómenos.

Davar et al, (2010) detectaron 7 QTLs para el porcentaje de área necrótica, localizada en 7 grupos de ligamiento, cuyos efectos fueron pequeños a moderados.

Los estudios de QTL para la resistencia a *S. sclerotiorum* han demostrado que los QTL detectados a partir de diferentes cruzamientos parecen ser diferentes, lo que sugiere que la piramidización de QTLs para la resistencia es posible. Sin embargo, como cualquier otro carácter cuantitativo, es necesario la confirmación de la posición del QTL mediante el mapeo fino o mapeo por asociación antes que su utilización en la selección asistida por marcadores (MAS).

La posibilidad de aumentar el nivel de resistencia de las líneas de élite por introgresión de varios QTL con la ayuda de MAS y herramientas de cultivo de embrión, sin afectar a los componentes del rendimiento, constituye una manera simple con respecto a reducir el impacto de esta compleja enfermedad.

## 5. ALTERNATIVAS A LA MAPEO DE LIGAMIENTO TRADICIONAL

Alternativas al mapeo de QTL a partir del análisis de poblaciones biparentales han surgido recientemente, incluyendo el mapeo por asociación (MA) en sus estrategias de mapeo de genes candidatos y genoma completo (genome-wide association, GWA), que utiliza la diversidad genética natural y los numerosos eventos de recombinación ocurridos en las poblaciones a lo largo de su historia para lograr mapeo a nivel de genes individuales (Zhu et al. 2008; Bergelson & Roux 2010; Rafalski 2010). Actualmente, tanto mapeo de ligamiento como mapeo por asociación, o la combinación de ambos son las dos estrategias más utilizadas a la hora de disecar caracteres de base genética compleja.

En la última década, las tecnologías de secuenciación de nueva generación (next generation sequencing, NGS) ha facilitado el acceso a numerosos datos genómicos, permitiendo tanto mapeo de ligamiento y GWA en diversas especies vegetales.

Un ejemplo de la potencia combinada de ambas técnicas ha permitido mapear finamente genes QDRs en *Arabidopsis*. Marcadores genéticos asociados con la variación natural de QDR a *Xanthomonas campestris* fueron localizados en la región o dentro del gen *RKS1*, gen clonado previamente a través de un enfoque de mapeo de QTL tradicional (Huard-Chauveau et al. 2013).

En el caso de girasol, el primer trabajo de mapeo por asociación descrito en la literatura se centró en el estudio de la resistencia a la PHC utilizando un enfoque basado en genes candidatos sobre una población de 94 líneas endogámicas que pertenecen al "Programa de Mejoramiento de Girasol" de INTA (Fusari et al. 2012). Dado que no existen mecanismos biológicos o vías bioquímicas identificadas para la resistencia a PHC, la selección de genes candidatos se basó en estudios de transcripción de perfiles anteriores en girasol (Giacomelli et al. 2010) y *Brassica napus* (Zhao et al. 2007) infectados con *S. sclerotiorum* y derivaron en el estudio de 16 genes candidatos que fueron genotipados en la población de mapeo por asociación. Las asociaciones entre la incidencia a PHC y los haplotipos de esos 16 genes candidatos permitieron identificar una asociación significativa entre el gen candidato *HaRIC\_B* y la incidencia a PHC, que representa una reducción de la incidencia de SHR de aproximadamente 20% (Fusari et al. 2012).

# MATERIALES Y MÉTODOS

## 1. MATERIAL VEGETAL

El material vegetal de las 135 RILs derivadas del cruzamiento entre PAC2 x RHA266 empleados en esta etapa fue descrito en el Capítulo I, en la sección correspondiente.

## 2. CARACTERIZACIÓN FENOTÍPICA. ENSAYOS DE DESAFÍO DE LA POBLACIÓN CON *S. sclerotiorum* REALIZADOS EN EL CAMPO.

A fin de caracterizar fenotípicamente la resistencia a PHC y dado lo complejo del carácter, durante el período comprendido entre diciembre de 2009 y marzo de 2014 se llevaron a cabo 5 ensayos a campo en condiciones de riego en la EEA INTA Balcarce situada a 37°45'43,9" Sur, 58°18'27,4" Oeste a 125 m.s.n.m. El sur de Buenos Aires tiene la característica de ser una zona de alto impacto para la enfermedad PHC.

Exceptuando la campaña 2010-2011, se sembraron 135 líneas y los parentales.

Por último, con el objeto de evaluar la variabilidad inherente al lote y a modo de control de infección, en cada campaña se sembraron materiales testigos en un lote contiguo. Estos estaban compuestos por híbridos comerciales (PARAISO 22, SOLFLOR, ACA 885, ACA 861 y DEKALB 3820) y líneas endocriadas elites (HA89, RHA266 y PAC2) de comportamiento caracterizado frente al patógeno *S. sclerotiorum* (Al-Chaarani et al. 2002; Al-Chaarani et al. 2004; Gentzbittel et al. 1995; S. Kiani et al. 2007; S. P. Kiani et al. 2007; Kiani et al. 2008; Alfadil et al. 2007)

### 2.1 PRE-SIEMBRA

Previo a la siembra del material vegetal, cada semilla fue tratada con una solución de ácido giberélico en una concentración de 100 partes por millón por un período de 30 minutos según lo establecido por Pereyra (1993). Lo último permite garantizar la ruptura de la dormición y asegurar un alto poder germinativo.

Además, las semillas se trataron con una dilución de Metalaxil M (METALAXIL-M APRON GOLD, SYNGENTA AGRO S.A., Balcarce, Argentina) en agua corriente (3:1, v:v) como tratamiento preventivo frente a mildiu. Siguiendo la dosis recomendada del principio activo que indica 600 cm<sup>3</sup> de formulado cada 100 kg de semillas, aproximadamente 0,006 ml por gramo de semilla fue rociado hasta lograr una cobertura homogénea.

## 2.2 SIEMBRA

En todas las campañas las líneas se dispusieron en el campo utilizando un diseño en bloques completos aleatorizados con tres repeticiones. La Tabla 2.2 muestra las fechas de siembra correspondientes a cada campaña.

CAMPAÑA	FECHA DE SIEMBRA
09-10	29/10/2009
11-12	05/12/2011
12-13	05/12/2012
13-14	09/12/2013

Los materiales testigos fueron sembrados durante dos semanas previas y 2 dos semanas posteriores a la fecha de siembra propia del ensayo para que su floración coincidiera con la de las RILs.

Debido a que la siembra se realizó mecánicamente, a razón de al menos 5 semillas por golpe, fue preciso realizar el raleo post emergencia de las plántulas. Entonces, la unidad experimental en estos ensayos consistió en un surco de 10 m. con una distancia entre plantas de 0,25 m., totalizando aproximadamente 30 plantas por surco y separado de la parcela contigua por 0.75 m.

El control de malezas se realizó con escardillo, cuando las plantas estaban en el estadio R1.

## 2.3 INOCULACIÓN

Las ascosporas necesarias para producir el inóculo fueron cedidas por el grupo de fitopatología de girasol de la Unidad Integrada INTA Balcarce. El protocolo de inoculación empleado es el descrito por Vear & Tourvieille (1984). Se realizó inoculación utilizando aspersión mecánica (Escande et al. 2002) con una solución de 2500 ascosporas/mL medidas previamente con hemocitómetro (cámara de Neubauer). La misma se efectuó cuando el 50% de los capítulos de la parcela alcanzo el estadio R5.2 (Schneiter & Miller 1981), período durante el cual el girasol presenta mayor susceptibilidad frente al patógeno (Pedraza & Pereyra 2004). Inmediatamente después de la inoculación, se procedió al tapado de capítulos con bolsas de papel tipo Kraft (EMBAIRES S.A., Buenos Aires, Argentina) y se mantuvo el tapado durante los siguientes 10 días.

Para favorecer el desarrollo y progreso de la enfermedad y evitar la desecación del inóculo, luego a cada inoculación, el ensayo fue regado por aspersión y mantenido diariamente con períodos de riego de 15 minutos en las horas de mayor temperatura.

Respecto de los controles, en paralelo al ensayo, cada fecha de inoculación se procedió a realizar la infección de los mismos siguiendo el protocolo ya descrito. A su vez, para las campañas 2012-13 y 2013-14, en una parcela contigua a los controles desafiados, plantas sin inóculo se cubrieron con bolsas de papel tipo Kraft durante 10 días para emular las mismas condiciones que los controles desafiados. Es así como cada control contó con parcelas “inoculadas” y “no inoculadas”.

#### 2.4 *EVALUACIÓN FENOTÍPICA*

En los ensayos de evaluación fenotípica, se midieron dos variables a distintos tiempos post-inoculación sobre los 135 individuos de la población de RILs. Las variables evaluadas fueron:

- 1) Incidencia de la enfermedad (IE): definida como el cociente entre el número de plantas enfermas y el número total de plantas evaluadas. Bajos valores de Incidencia indican un alto nivel de resistencia (Castaño et al. 1993).
- 2) Severidad de la enfermedad (SEV): definida como la proporción del área del capítulo con síntomas externos de PHC respecto de la superficie total de la inflorescencia promediada entre las plantas enfermas. Bajos valores de Severidad indican un alto nivel de resistencia del material en cuestión para dicha variable.

Estas dos variables, representarían etapas diferentes del mecanismo de resistencia. Incidencia se vincula con las primeras barreras del hospedante, mientras que Severidad está involucrada con la resistencia a la dispersión del hongo provista por la planta.

A partir de estos registros se obtuvieron otras variables que se desprenden de las anteriores, entre ellas:

- 3) Intensidad de la enfermedad (INT): definida como el promedio del porcentaje afectado del capítulo por línea. De la misma manera que para SEV, valores bajos de esta variable indican niveles de resistencia superiores
- 4) Período de incubación (PI): definida como el tiempo promedio medido en días post inoculación en el cual aparecen los primeros síntomas. Altos valores de esta variable se condicen con altos niveles de resistencia.

La evaluación de las variables citadas se registró visualmente para cada parcela a 10, 14, 17, 19, 21, 24 y 28 días post inoculación (DPI), exceptuándose la primera campaña, donde no se siguió este esquema sino que se evaluó todo el ensayo a días fijos

En ocasiones, la determinación de las variables a los puntos establecidos se vio dificultada o no fue posible realizarla, debido a características propias de la línea (como presencia de manchas, capítulos duros o haber alcanzado la madurez fisiológica -estadio fenológico R9-) que impidieron visualizar síntomas donde en un punto anterior fueron detectados. De esta manera, se consideró la variable “IEMAX” como la incidencia máxima registrada en las evaluaciones realizadas. Para descartar evaluaciones erróneas, todas las observaciones correspondientes a IEMAX iguales o inferiores a 21 DPI fueron eliminadas.

En cuanto a las variables SEV e INT, se determinó empíricamente que para evitar cometer errores ligados a la subjetividad, la observación realizada al DPI 21, es la óptima para determinar diferencias significativas entre líneas, ya que las evaluaciones posteriores a este punto denotaban un retroceso en el porcentaje del capítulo infectado en la mayoría de las líneas por las causas anteriormente detalladas. Las variables entonces fueron nombradas como “SEV21” e “INT21”.

Como SEV resulta del cociente entre la sumatoria de las severidades registradas y el total de plantas enfermas, aquellas parcelas que al DPI 21 no presentaban síntomas (y por ende no habían plantas enfermas) fueron asignadas con un valor de “0,01” para evitar el error matemático de dividir por 0.

Cabe aclarar que para los controles, únicamente se contempló la IE máxima (IEMAX).

## 2.5 ANÁLISIS DE LOS DATOS

El análisis de los datos fenotípicos para determinar el efecto de las RILs (genotipo) se llevó a cabo empleando el software INFOSTAT (Di Rienzo et al. 2014). Las parcelas que contenían menos de 5 plantas inoculadas, fueron descartadas.

En primer lugar, y a fin de describir los aspectos fundamentales del conjunto de datos obtenidos a partir de las variables que caracterizan la resistencia a la PHC, se procedió a calcular algunas de las medidas de posición y dispersión de los mismos. Las medidas estimadas fueron la media aritmética, los valores máximos y mínimos, el rango y el coeficiente de variación.

Uno de los pasos previos a la comparación simultánea de más de un conjunto de datos fue corroborar que SEV21 e INT21 cumplieren con la condición de normalidad y homogeneidad de variancias (homocedasticidad). La normalidad de los residuos fue probada utilizando la prueba de Shapiro-Wilk, mientras que se realizó la prueba de Levene de homogeneidad de

variancias. Esta prueba determina que dos o más variancias son homogéneas cuando la probabilidad asociada a  $F_0$  (estadístico de Levene) es mayor a un nivel de significancia establecido que, en nuestro caso fue  $\alpha=0,05$ .

Si bien la transformación recomendada para ajustar datos correspondientes a proporciones

(como lo son estas variables) es  $(\sqrt{p(1-p)})$  -donde “p” es la proporción-, los valores de los

datos fenotípicos no se transformaron, debido a que estudios previos (Mutschler et al. 1996) han sugerido que los datos no transformados son mejores estimadores en el análisis de QTL. Además, la transformación de los datos puede disminuir diferencias entre individuos debido al acortamiento de las colas de los datos hacia el centro de la distribución (Mutschler et al. 1996; Byrne et al. 1998)

Verificada la homocedasticidad, se procedió realizar el análisis conjunto mediante modelo lineales mixtos. Para ello se elaboró, primero, el modelo teórico estadístico-matemático que facilita el estudio y comprensión de cada una de las variables “y” (dependiente) consideradas, como sigue:

$$Y_{ijkl} = \mu + RIL + Campaña_i + Bloque_j + fecha\ de\ inoculación_k + Genotipo_l + Genotipo \times fecha\ de\ inoculación_{(kl)} + Error_{ijkl}$$

Donde  $Y_{ijk}$  es el promedio de la variable considerada en la campaña “i”, el bloque “j”, fecha de inoculación (k), el genotipo “l” y  $\mu$  es la media general de dicha variable.

La variable IEMAX, se trata de una distribución de probabilidad discreta que cuenta el número de plantas enfermas en una secuencia de n ensayos (plantas totales) por lo que responde a una distribución binomial. En cambio la variable PI responde a una distribución de probabilidad gamma.

Para evaluar estas dos variables se utilizaron modelos lineales generalizados mixtos que utilizan una función de enlace determinada (“logit” y “gamma inversa” para IEMAX y PI respectivamente) para vincular la media de la distribución de las  $Y_{ijkl}$  con el predictor lineal descrito anteriormente

Para este trabajo de Tesis, todos los análisis llevados a cabo, se tuvo en cuenta a los términos del modelo “Campaña”, “Repeticiones (bloques)”, “Fecha de inoculación” e interacción de “Fecha de inoculación x RIL”, como efectos aleatorios, dado que dichas fuentes de variación pueden considerarse como muestras tomadas en un conjunto de ambientes, repeticiones y fecha de inoculaciones con posibilidades de ser utilizados para efectuar las evaluaciones. Al

término “RIL” del modelo, en cambio, se los consideró como efectos fijos; esto último significa que las inferencias realizadas a partir de sus comportamientos serán sólo aplicables a los materiales genéticos evaluados.

El hecho de no contar con la totalidad de los genotipos en los ensayos evaluados determinó que los datos estuvieran desbalanceados.

Se obtuvieron las medias ajustadas por líneas para la totalidad de campañas consideradas y de manera individual (quitando al término “campaña” del modelo).

Con el método DGC (Di Rienzo, Guzman y Casanoves) incorporado en el programa InfoStat, se realizaron comparaciones múltiples para identificar diferencias significativas ( $\alpha= 0,05$ ) entre las medias de cada genotipo. DGC utiliza la técnica multivariada del análisis de conglomerados sobre una matriz de distancia obtenida a partir de las medias muestrales para realizar las comparaciones múltiples.

Luego, se calcularon los coeficientes de correlación de Pearson ( $r$ ) para evaluar las asociaciones de las variables que califican a la resistencia a la PHC.

## 2.6 ESTIMACIÓN DE LA HEREDABILIDAD

Dado que estos datos se encuentran desbalanceados, para estimar la heredabilidad de los caracteres en sentido amplio ( $H^2$ ), se emplearon modelos lineales mixtos y modelos lineales mixtos generalizados, donde se introdujo el término “RIL” como efecto aleatorio en cada uno de los modelos anteriormente evaluados. El porcentaje del total de la varianza genotípica permitió estimar la heredabilidad correspondiente.

### *IDENTIFICACIÓN DE QTLS*

La detección y mapeo de QTL asociados a la resistencia a PHC se efectuó utilizando la información generada en las dos etapas anteriores: las medias ajustadas por líneas obtenidas en la caracterización fenotípica y los datos del mapa generados en la caracterización genotípica.

Se empleó la metodología de mapeo por intervalos compuestos (CIM, Composite Interval Mapping) (Zeng 1994) combina el mapeo por intervalos (IM, Interval Mapping) (Lander & Botstein 1989) y la regresión múltiple. El CIM permite incrementar el poder de detección de QTLs controlando la probable variabilidad genética aportada por regiones genómicas fuera del intervalo genético en el cual se realizan los test de hipótesis. Esto se consigue usando

como cofactores aquellos marcadores asociados al carácter e identificados por el proceso de regresión múltiple. La presencia de un QTL se infiere a partir del valor del estadístico LR (Likelihood Ratio) o su equivalente valor de LOD. El nivel de significancia o de corte a partir del cual un QTL se determinó como válido se obtuvo mediante la realización de 1000 permutaciones para  $\alpha = 5\%$  (Doerge & Churchill 1996). En este proceso los datos fenotípicos son permutados con respecto a los datos genotípicos 1000 veces, y para cada permuta se lleva cabo el proceso de mapeo y el mayor valor de LOD es retenido. Finalmente, los 1000 valores de LOD son ordenados de mayor a menor y, si el valor de LOD obtenido en la estrategia de mapeo está comprendido entre los mayores 50 valores la presencia del QTL puede declararse real con un 95% de confianza. En el presente trabajo el mapeo de QTL se realizó a través del método CIM, seleccionando en el programa QTL Cartographer 2.5 (Wang et al. 2011) el modelo “6” de CIM en el que los cofactores son seleccionados automáticamente, a cada 10 cM y en número máximo de 5. Como método de regresión múltiple se seleccionó la de avance y retroceso con 5% de probabilidad para entrada y salida de los marcadores o variables independiente, la velocidad de avance fue de 0.5 cM (es decir, cada 0.5 cM el software realizó las pruebas de hipótesis:  $H1: a \neq 0$  /  $H0: a = 0$ ). El máximo valor de LOD y superior al del corte, indicó la posición del QTL (cromosoma e intervalo genético) y sus parámetros, efecto aditivo (a) y coeficiente de determinación múltiple ( $R^2$ ) que indica el porcentaje de la variabilidad fenotípica explicada por cada QTL.

### 2.7 ANÁLISIS IN SILICO DE LOS MARCADORES QUE SUBYACEN A LOS QTL DETECTADOS

La falta de un genoma de referencia para girasol dificulta la posibilidad de analizar las regiones genómicas asociadas a los QTL con mayor detalle. No obstante, y a los fines de aproximarnos al estudio de la región, emprendimos un análisis observacional de los marcadores más cercanos a los QTL mediante la exploración de la anotación funcional de los mismos relevada en el repositorio ATGC-SUR y confirmada a partir de nuevas búsquedas o bien asignando funciones a las secuencias anónimas correspondientes a SSRs mediante inferencias soportadas por búsquedas en repositorios públicos como NCBI.

Las secuencias de los SNPs (HeAn\_x) fueron obtenidas del modo descrito en la sección correspondiente al diseño del ensayo.

A su vez, las secuencias de los marcadores SSRs (HAX y ORSx) situados en regiones de interés, fueron obtenidas de la base de datos dbSTS (Sequence Tag Site de NCBI

<http://www.ncbi.nlm.nih.gov/dbSTS/>. Dos marcadores SSRs (ORS121 y SSU25) fueron descartados del análisis por no poseer la secuencia del fragmento amplificado.

Los marcadores del tipo AFLPs fueron descartados para este análisis ya que al ser anónimas no representan secuencias locus específicas.

Por último, aunque el GC GLP4 ha sido descrito y resecuenciado en estudios previos (Fusari 2010), se optó por incluir la secuencia obtenida de dicho trabajo en el análisis con la finalidad de actualizar (si así lo fuera) la caracterización del mismo.

De este modo, se generó un archivo “multifasta” conteniendo 39 secuencias correspondientes a 15 HeAn\_x, GLP4 y 23 SSRs (12 HAx y 11 ORSx). Este archivo fue insumo (“query”) para utilizar la herramienta “blastx” del programa Blast2Go (en su versión básica 3.0) y así comparar la traducción de cada secuencia en 6 proteínas hipotéticas (1 por cada marco de lectura) contra una base de datos no redundantes de las mismas (E-valor de  $1e^{-10}$ ). Una segunda evaluación fue llevada a cabo sobre las mismas secuencias con otra herramienta del programa Blast2Go, llamada “InterProScan”. Ésta misma también puede encontrarse libremente en el enlace: <http://www.ebi.ac.uk/interpro/>. Brevemente, InterProScan es una herramienta del Instituto Europeo de Bioinformática (EBI, de sus siglas en inglés) que permite realizar escaneos de secuencias nucleotídicas o aminoacídicas contra modelos predictivos provistos por diferentes bases de datos que integran el consorcio.

Los valores de búsqueda para realizar la evaluación, fueron configurados por defecto.

Los datos obtenidos de ambas herramientas, fueron fusionados y anotados con el mismo programa cuando superaron el umbral inferior de valor E igual o mayor a  $1e^{-10}$ . De esta manera se determinó la anotación génica y posible función de los marcadores asociados.

En este punto, cabe tener en cuenta dos consideraciones:

Si bien el desarrollo del panel de genotipificación de 384 SNPs con la tecnología VeraCode BeadXpress de Illumina implicó la selección de unigenes con ontología génica relacionada a estreses bióticos y abióticos, se optó por incluirlos nuevamente en este análisis para corroborarla mediante la comparación con bases de datos más actualizadas. La anotación para los HeAn\_x se encuentran depositadas en la base de datos SUR: <http://atgc-sur.inta.gob.ar/>.

Los marcadores SSRs en su mayoría se presentarían como secuencias genómicas (y por ende, sin función asociada), se evaluaron sus secuencias de igual manera.



Se realizó otra comparación con la herramienta “blast” incorporada en la base de datos “heliagene (<https://www.heliagene.org>)” correspondiente a la secuenciación del transcriptoma de la línea Ha412, seleccionada para su secuenciación completa por parte del Consorcio de Secuenciación de Compuestas. Los parámetros para llevar adelante este análisis fueron configurados del modo descrito anteriormente. En esta comparación fueron consideradas como válidas las alineaciones que mostraron una identidad superior al 90% y cubriendo parte de la secuencia query.

Por último y con la intención de encontrar características asociadas, con la herramienta Cmap se procedió a inspeccionar visualmente la zona flanqueante de los marcadores SSRs aquí seleccionados sobre los mapas públicos depositados en el siguiente enlace: [www.sunflower.uga.edu/cgi-bin/cmap/viewer](http://www.sunflower.uga.edu/cgi-bin/cmap/viewer).

# RESULTADOS

## 1. EVALUACIÓN FENOTÍPICA

Se evaluaron 4 variables que caracterizan el nivel de resistencia a PHC en diferentes estadios de las líneas que integran la población PAC2 x RHA266

Como fuera mencionado, la campaña 2009-2010 no fue evaluada en los días post inoculación 10, 14, 17, 19, 21, 24 y 28, sino que todo el ensayo fue evaluado a días fijos, promoviendo la falta de datos en algunas de las líneas para cada variable evaluada.

Otra característica de esta campaña es que la siembra se realizó con anticipación a la fecha normal del año, originando un ciclo más largo y por ende fechas de evaluaciones que superaron al DPI 28.

La Tabla 2.3 presenta los valores promedio, desvíos y errores estándares, los coeficientes de variación, los valores mínimos y máximos de la población y medidas descriptivas de la curva de distribución, considerando cada campaña por separado.

La media de IEMAX para la campaña 2011-12 difiere de las restantes y esto puede deberse a las condiciones climáticas durante el período de inoculación. La Tabla III del apéndice muestra un

<b>CAMPAÑA</b>	<b>VARIABLE</b>	<b>MEDIA</b>	<b>D.E.</b>	<b>E.E.</b>	<b>C.V.</b>	<b>MIN</b>	<b>MAX</b>	<b>ASIMETRÍA</b>	<b>CURTOSIS</b>
09-10	IEMAX	0,51	0,27	0,02	52,45	0,00	1,00	-0,17	-0,92
	SEV21	0,18	0,14	0,01	80,21	0,05	1,00	3,20	13,72
	INT21	0,03	0,05	0,00	133,68	0,00	0,36	3,93	20,52
	PI	28,03	4,48	0,31	16,00	16,00	42,00	0,02	0,05
11-12	IEMAX	0,11	0,18	0,01	161,56	0,00	0,75	1,78	2,26
	SEV21	0,30	0,27	0,03	92,76	0,01	1,00	0,90	-0,10
	INT21	0,03	0,06	0,00	220,20	0,00	0,36	3,01	9,93
	PI	20,26	3,56	0,41	17,57	14,00	28,00	0,39	-0,51
12-13	IEMAX	0,65	0,24	0,02	37,35	0,00	1,00	-0,55	-0,17
	SEV21	0,51	0,21	0,01	40,99	0,01	1,00	0,06	-0,20
	INT21	0,28	0,19	0,01	67,25	0,00	0,88	0,60	-0,27
	PI	19,50	2,70	0,18	13,87	14,00	28,00	0,31	-0,19
13-14	IEMAX	0,48	0,28	0,02	59,57	0,00	1,00	0,09	-0,97
	SEV21	0,29	0,17	0,02	61,00	0,01	0,80	0,26	-0,50
	INT21	0,12	0,13	0,01	111,12	0,00	0,70	1,41	1,82
	PI	20,92	2,82	0,24	13,47	14,00	28,00	-0,13	-0,41

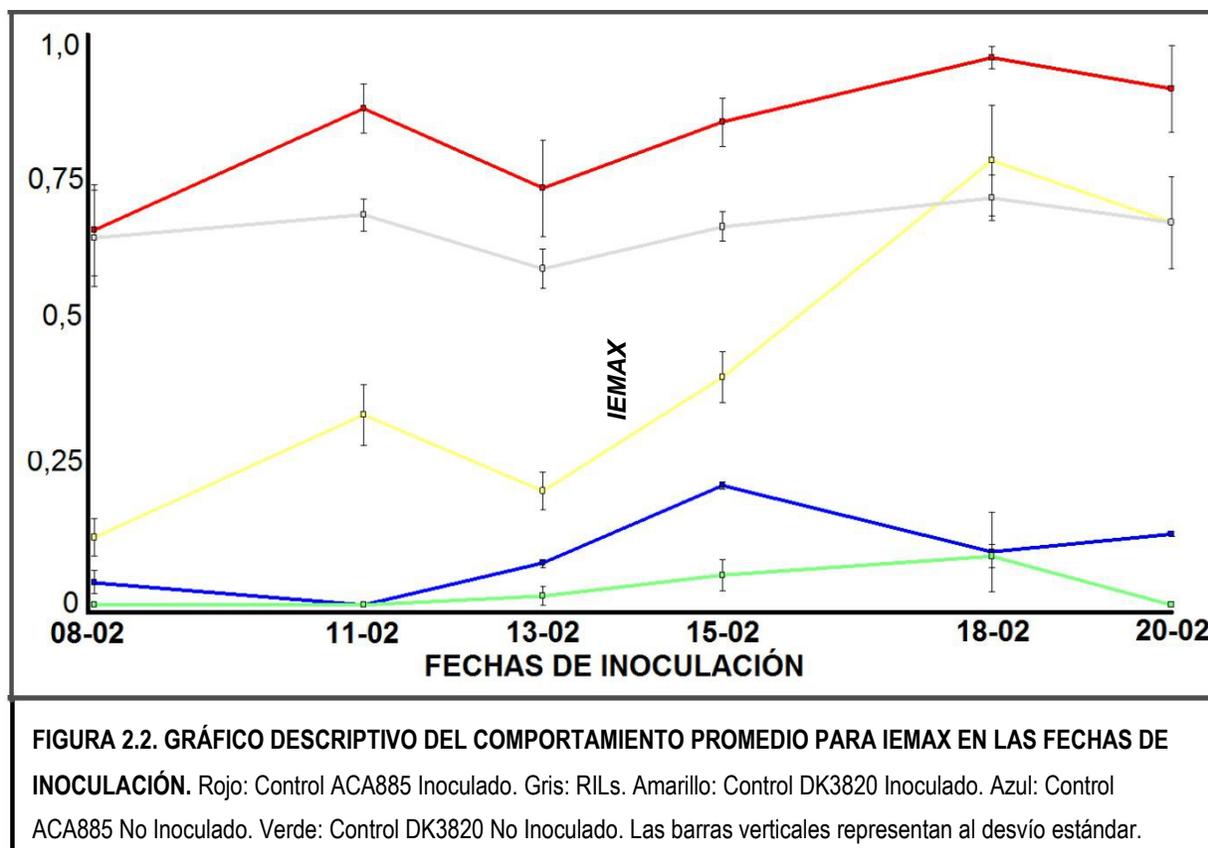
extracto de las condiciones ambientales para dicho período en las diferentes campañas. Como puede observarse, las altas temperaturas máximas que se presentaron en la campaña 2011-12, sumadas a la escases de precipitaciones (sobre todo en los primeros 10 días donde se realizaron las inoculaciones) afectaron al grado de ataque y patogenicidad del hongo, hecho que se refleja en las medidas obtenidas (asimetría y curtosis elevadas y positivas).

Asimismo, considerando a INT DPI-21 como un buen descriptor de las condiciones ambientales reinantes a ese día post inoculación, se observa una clara diferencia (excepto para la campaña 2009-10) que avala lo anteriormente expuesto. Por otra parte, el hecho de que la campaña 2009-10 fuera sembrada con antelación ocasionó que al DPI 21 muchas líneas no expresaran síntomas y mostraran períodos de incubación más altos.

Cada ensayo contó con híbridos o líneas para llevar adelante el control de la inoculación. Únicamente el ensayo 2012-13 contó con suficientes parcelas inoculadas cada fecha. Esto se debió principalmente a que para los demás ensayos no pudieron coordinarse los estadios fenológicos R5.2 de los controles con las fechas de inoculación.

Debido a que la variable IEMAX fue evaluada tanto en las RILs como en los controles sembrados en cada campaña, fue posible realizar un gráfico descriptivo del comportamiento medio y errores estándares para las RILs y controles inoculados el mismo día. La Figura 2.2 muestra un gráfico de puntos perteneciente al ensayo llevado a cabo en la campaña 2012-13, siendo los híbridos comerciales ACA885 y DK3820, de comportamiento susceptible y medianamente resistente a PHC, respectivamente. En ella puede observarse que si bien el comportamiento es dentro de lo esperado, se percibe una interacción ocasionada por las condiciones climáticas de los días de inoculación (evidente en el control "DK 3820 INOC"). Esto permitió corroborar la correcta incorporación del término de interacción "Genotipo x fecha de Inoculación" al predictor lineal del modelo para así evaluar la varianza aportada por este término.

Como fue mencionado en la sección "materiales y métodos", la variable IEMAX posee una distribución binomial, mientras que PI responde a una distribución gama inversa. Ambas fueron evaluadas empleando modelos lineales generalizados mixtos. Los datos pertenecientes a SEV21 e INT21 fueron analizados con modelos lineales mixtos sin transformar. Los test de Shapiro y Levene permitieron comprobar la normalidad y la homogeneidad de varianza (datos no mostrados).



Del total de RIL sembradas en todas las campañas, en 36 de ellas no fue posible obtener datos por problemas de germinación, escasos de plantas en la parcela, sequedad o la imposibilidad de medir las variables por características propias de la línea como manchas, capítulos deformes, entre otras.

Los resultados obtenidos de las medias ajustadas de cada variable para las 99 líneas RILs y ambos parentales se presentan en la Tabla IV del apéndice. Puede observarse que los análisis realizados mediante el empleo de modelos lineales mixtos (para SEV21 e INT21) y modelos lineales generalizados mixtos (para IEMAX y PI) detectaron diferencias significativas ( $p < 0,001$ ) entre líneas, lo cual destaca la variabilidad presente en la población y su utilidad para selección y mapeo de QTLs de resistencia. Las diferencias observadas se reflejan en los histogramas de las medias ajustadas por campaña y de manera conjunto para todas las variables analizadas (Figura 2.3).

De acuerdo a estos resultados, exceptuando la campaña 2011-12, los valores de incidencia máxima fueron variables. Del mismo modo que en el análisis exploratorio inicial, se hace evidente las condiciones que la campaña 2011-12 presentó condiciones menos favorables para el desarrollo de la enfermedad.

En cuanto, SEV21 y PI, dado a que estas variables dependen de las plantas enfermas únicamente, se observó una amplia variabilidad de valores. Para INT21, las dos últimas campañas muestran distribuciones más amplias que las correspondientes a las dos campañas iniciales. Nuevamente, esto puede deberse a que, como fue detallado anteriormente, la campaña 2009-10 fue sembrada con antelación en la fecha respectiva del año y para el momento de evaluación al DPI 21, pocas líneas presentaban síntomas. Para el caso de la campaña 2011-12, este acotamiento en la variabilidad de la respuesta, puede estar explicada por las condiciones ambientales que retrasaron o impidieron la aparición de síntomas.

El paso siguiente consistió en un análisis de correlación. Para ello, se tomaron los valores ajustados en el conjunto de campañas para revisar asociación entre las características evaluadas. La Tabla 2.4 muestra los coeficientes de correlación de Pearson ( $r$ ) por debajo de la diagonal principal y los valores  $p$  de significación mientras que la Figura 2.4 muestra una matriz de diagramas de dispersión entre las variables.

	IEMAX	SEV21	INT21	PI
IEMAX	1	<0,00001	<0,00001	<0,00001
SEV21	0,41	1	<0,00001	<0,00001
INT21	0,64	0,8	1	<0,00001
PI	-0,45	-0,67	-0,64	1

Todos los valores de los coeficientes de Pearson obtenidos resultaron ser significativos probablemente a que el análisis contempló una muestra grande. Por esta razón, se procedió clasificar a los coeficientes de correlación en 3 categorías: baja-moderada-alta para su correcta interpretación. Se encontró una moderada correlación negativa entre PI y las variables SEV21 e INT21 ( $r = -0,67$  y  $-0,64$ , respectivamente). Mientras que INT21 mostró una correlación positiva con IEMAX y SEV21 ( $r = 0,64$  y  $0,8$ , respectivamente). Esto último es esperable, ya que el cálculo de la intensidad (% del capítulo afectado/plantas inoculadas) resulta de multiplicar incidencia (plantas enfermas/plantas inoculadas) y severidad (% del capítulo afectado/plantas enfermas).

Por último, los valores más bajos de correlación fueron obtenidos para dos características asociadas con IEMAX, ellas fueron SEV21 y PI, con  $r = 0,41$  y  $r = -0,45$ , respectivamente.

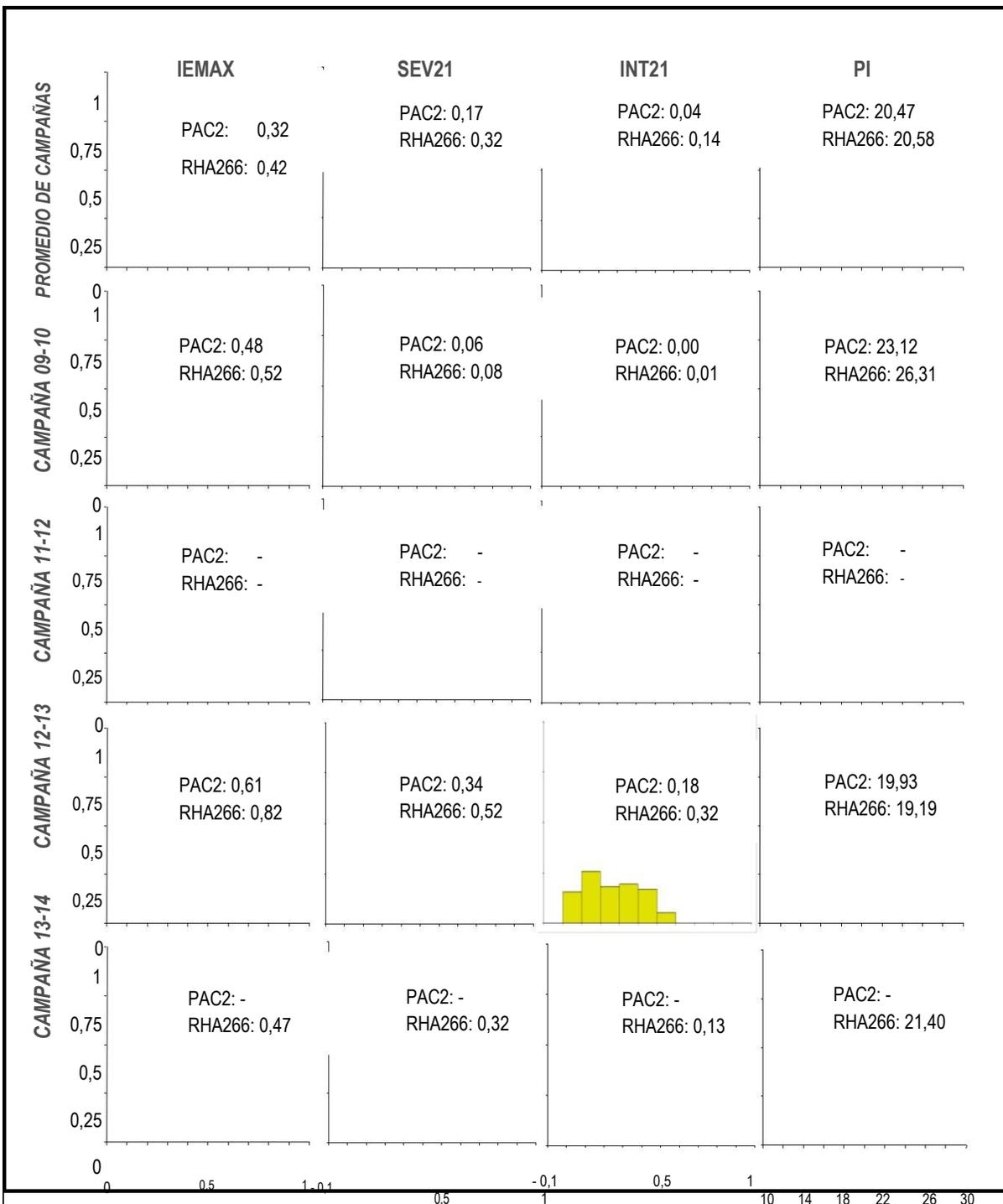


FIGURA 2.3. HISTOGRAMA DE FRECUENCIA DE DISTRIBUCIÓN MEDIAS AJUSTADAS. Se observan las mismas para cada variable evaluada en cada campaña y el promedio de los ensayos.

## 2. ESTIMACIÓN DE LA HEREDABILIDAD

La Tabla 2.5 representa los valores obtenidos de heredabilidad y sus respectivos desvíos estándares para las variables IEMAX, SEV21, INT21 y PI. Aunque el genotipo de las RILs es

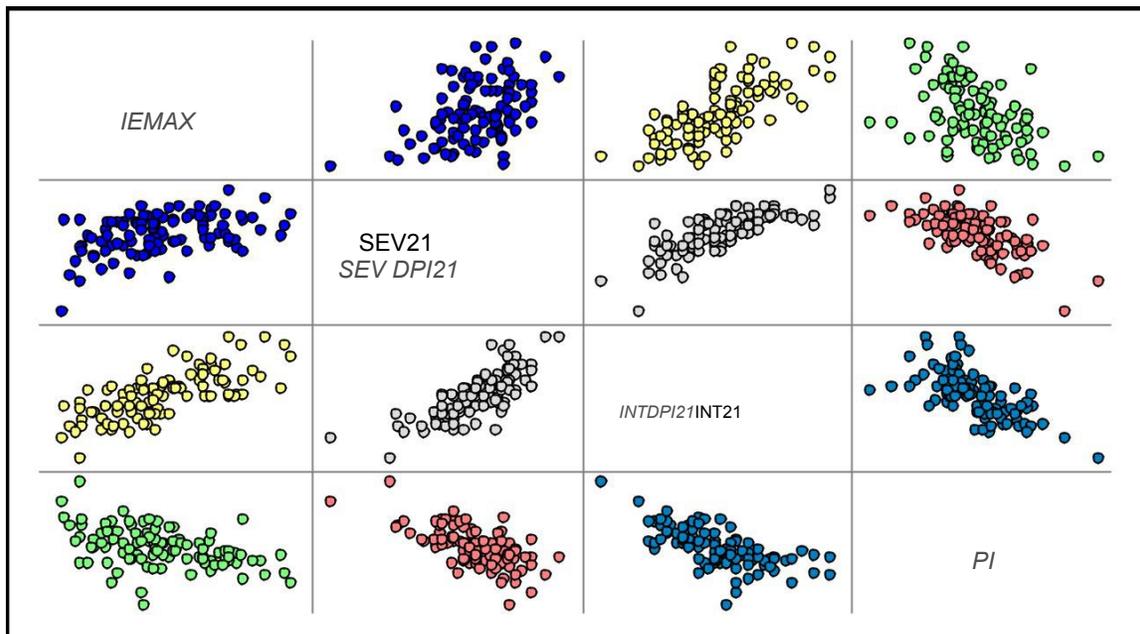
esencialmente homocigoto, los valores de heredabilidad se expresan en sentido amplio debido a la presencia de efectos epistáticos cuya variancia no es posible estimar en este delineamiento experimental. Como puede observarse, las heredabilidades oscilaron entre 1,86 (para INT21) de la primera campaña hasta 59,9% (para la misma variables pero última campaña). Tal discrepancia en valores indica el fuerte efecto ambiental que influye en los valores fenotípicos, y por tanto, en la eficiencia de la selección clásica como en la detección de QTLs.

**TABLA 2.5. Estimación de la heredabilidad ( $H^2$ ) por carácter y desvío estándar (DE)**

CAMPAÑA	IEMAX		SEV21		INT21		PI	
	$H^2$ (%)	DE						
09-10	45,61	2,91	20,4	0,05	1,86	0,01	13,74	1,87
11-12	7,34	0,51	43,36	0,19	2,87	0,01	45,48	2,45
12-13	34,98	3,12	20,49	0,1	26,49	0,1	37,94	1,73
13-14	31,49	2,29	50,52	0,13	59,9	0,1	52,45	2,05

### 3. DETECCIÓN DE QTLs

El mapeo de QTLs fue realizado usando, como vectores fenotípicos las medias ajustadas por líneas obtenidas, considerando cada por campaña de forma individual o promediando el promedio de las campañas, y el mapa genético obtenido y presentado en el Capítulo 1 de esta tesis. La identificación de QTLs se realizó a través del método mapeo por intervalos compuestos, implementado en el programa QTLCartographer (Wang et al. 2011). En este



**FIGURA 2.4. MATRIZ DE DISPERSIÓN DE DATOS**

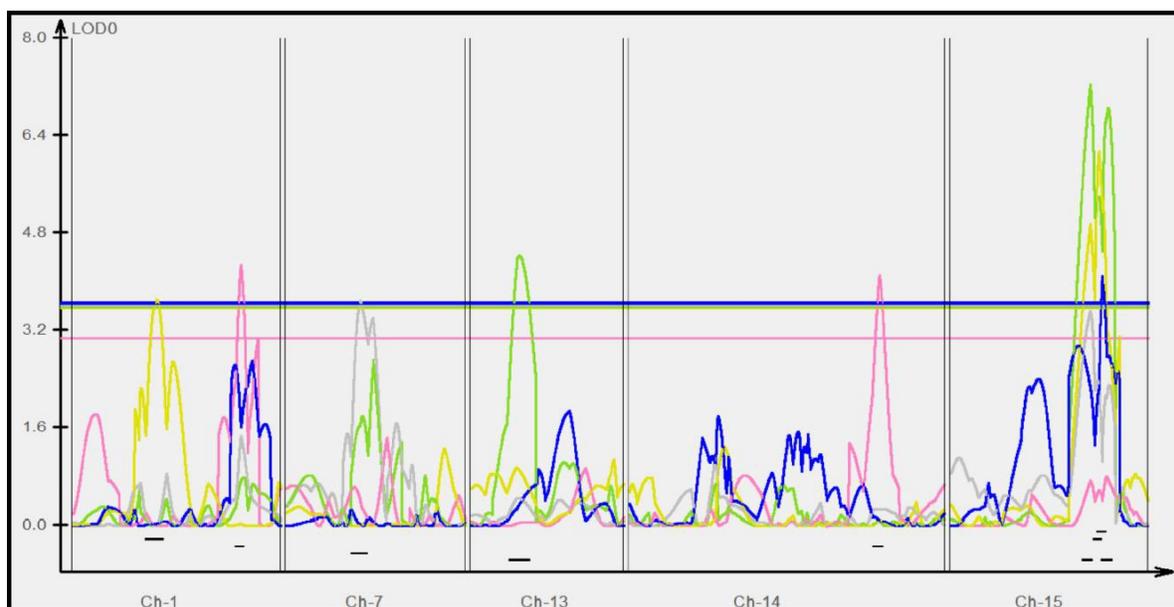
trabajo, la búsqueda de QTLs se realizó a nivel genómico.

En la Tabla 2.6 se muestran los valores de LOD críticos, por encima de los cuales un QTL puede declararse como real con un 95% de confianza ( $\alpha = 5\%$ ). Los valores obtenidos para todos los caracteres, excepto para INT21 de la CAMPAÑA 2009-10, fueron próximo o superiores a 3. Esto significa que la hipótesis alternativa ( $H_1: a \neq 0$ ), existe un QTL en la posición indicada, es 1000 veces más probable que la ausencia del mismo ( $H_0: a = 0$ ).

TABLA 2.6. Valores de LOD críticos obtenidos de 1000 permutaciones				
	IEMAX	SEV21	INT21	PI
EN PROMEDIO	3.61	3.33	3.26	3.26
CAMPAÑA 09-10	3.63	3.30	2.48	3.54
CAMPAÑA 11-12	3.08	3.69	3.05	3.69
CAMPAÑA 12-13	3.60	3.37	3.68	3.25
CAMPAÑA 13-14	3.66	3.87	3.38	3.57

### 3.1 QTLs ASOCIADOS CON LA CARACTERÍSTICA IEMAX.

Para el caso de la característica IEMAX, se detectaron valores de LOD superiores al crítico en cinco GLs (Figura 2.5 y Tabla 2.7) denotando la presencia de nueve QTLs vinculados con las variables consideradas. El valor de LOD informado para ellos, osciló entre 3,70 y 7,25, mientras que los intervalos de confianza del 95% se extendieron entre 3,3 y 12,2 cM.



**FIGURA 2.5. GRÁFICO DESCRIPTIVO DE POSICIONAMIENTO Y MAGNITUD DE LOS QTLs DETECTADOS POR CIM PARA IEMAX.** Sobre el eje y se disponen los valores de LOD mientras que en el eje x se describe la posición en los grupos de ligamientos ("Ch-" en la Figura). Las líneas horizontales representan los umbrales de LOD por sobre los cuales un QTL es considerado real. Referencia de colores: Gris: campaña 2009-10; Rosado: campaña 2011-12; Amarillo: campaña 2012-13; Azul: campaña 2013-14; Verde: Promedio de las campañas.

TABLA 2.7. Descripción de los QTLs detectados para IEMAX. *Distancia a punto de mayor LOD (cM)									
VARIABLE	NOMBRE	GL	cM	LOD SCORE	EFFECTO ADITIVO	R <sup>2</sup>	TAMAÑO INTERVALO (cM)	MARCADOR CERCANO	cM a >LOD*
IEMAX	qIEMAX/13	13	27,5	4,434	0,0754	0,1594	9,7	E33M48_20	1,51
	qIEMAX/15a	15	80,2	7,2523	-0,0921	0,2263	4,5	HeAn_R_534.1	0,03
	qIEMAX/15b	15	90,8	6,8677	-0,0933	0,2387	4,5	ORS687	0,97
IEMAX(09/10)	qIEMAX.09-10/7	7	42,1	3,697	0,0934	0,1245	12,2	ORS1041	0,03
IEMAX(11/12)	qIEMAX.11-12/1	1	96,7	4,2933	-0,0402	0,1515	3,3	HeAn_C_10493.1	0
	qIEMAX.11-12/14	14	143,4	4,1085	-0,0392	0,1606	4,2	ORS301	0,03
IEMAX(12/13)	qIEMAX.12-13/1	1	48,6	3,7105	0,0795	0,1361	8,3	E40M62_4	0,97
	qIEMAX.12-13/15	15	85	6,1546	-0,0953	0,1977	4	HeAn_C_12025.1/ HeAn_C_11434.1	0,05
IEMAX(13/14)	qIEMAX.13-14/15	15	87	4,0964	-0,0973	0,1471	3,3	ORS1242	0

Sobre el GL 1 en las campañas 2011/12 y 2012/13 se identificaron dos QTLs, (qIEMAX.11-12/1 y qIEMAX.12-13/1), siendo los marcadores más próximos HeAn\_C\_10493.1 y E40M62\_4, respectivamente. Sus efectos aditivos se estimaron en -0,0402 para el primero y + 0,0795 para el segundo, indicando que HeAn\_C\_10493.1 es un locus con alelos capaces de disminuir un en -0,0402 y E40M62\_4 elevar un + 0,0795 unidades los valores de IEMAX. Cada QTL fue capaz de explicar el 15,15% y 13,61% de la variación fenotípica observada (R<sup>2</sup>).

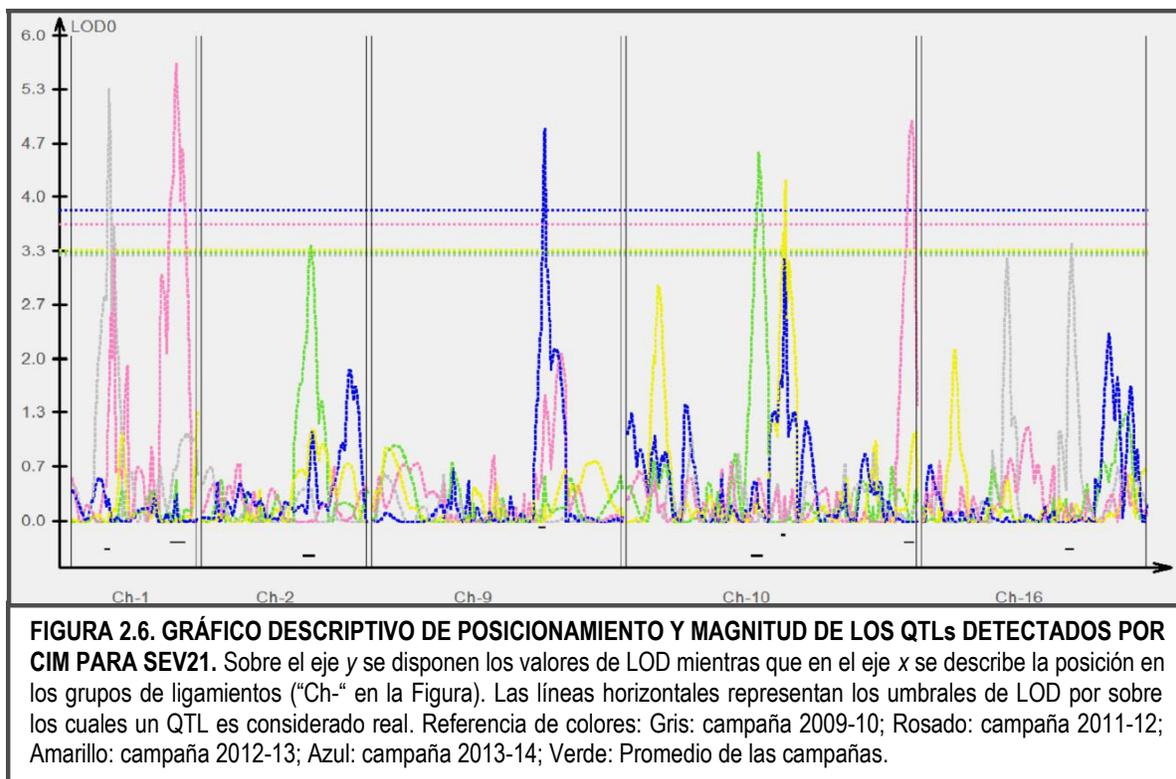
Sobre el GL 7, se detectó un QTL asociado con el SSR ORS1041 (qIEMAX.09-10/7), cuyo efecto aditivo se estimó en +0,0934 y su coeficiente de determinación en 12,45%. En el GL 14, el QTL qIEMAX.11-12/14, próximo del SSR ORS301, posee un efecto aditivo de -0,04 y explica el 16,06% de la varianza fenotípica. Por otra parte, sobre el GL 15 se identificaron dos QTLs, uno en la campaña 2012/13 (qIEMAX.12-13/15) en la posición 85 cM y el segundo en la campaña 2013/14 (qIEMAX.13-14/15) a 87 cM. Ambos con efectos aditivos y de R<sup>2</sup> similares en magnitud.

Cuando se consideró el promedio de las campañas, estos dos QTLs se detectaron nuevamente, qIEMAX/15a y qIEMAX/15b, y en posiciones similares a las identificadas en las campañas individuales mencionadas sugiriendo que esta sección genómica es una de las más importante para la tolerancia a IEMAX. No sólo por su reproducibilidad sino también por la magnitud de su R<sup>2</sup>, que en ambos casos supera el 20% siendo por tanto QTLs de efecto mayor. Además, dichos QTLs se localizaron en un segmento genómico que abarca aproximadamente 10 cM (80,2 cM - 90,8 cM) y comprende una zona poblada de marcadores SNPs funcionales (HeAn\_R\_534.1, HeAn\_R\_174.2, HeAn\_C\_11434.1 y HeAn\_C\_12025.1).

En el GL 13 se halló un nuevo QTL para el promedio de las campañas (qIEMAX/13) de efecto moderado (R<sup>2</sup>=16%) y un efecto aditivo de +0,0754.

### 3.2 QTLs ASOCIADOS CON LA CARACTERÍSTICA SEVERIDAD

Para esta característica, fueron detectadas 9 regiones genómicas, involucrada en la mayoría de los análisis realizados. Estos QTLs fueron identificados en los GLs 1, 2, 9, 10 y 16 (LODs= 3,41 – 5,67 e IC95%= 0,7-11 cM, Figura 2.6 y Tabla 2.8).



Sobre el GL 1, se ubicaron tres QTLs provenientes de diferentes campañas, nombrados como qSEV21.09-10/1, qSEV21.11-12/1a y qSEV21.11-12/1b y cuyos marcadores más próximos fueron E37M47\_15, HeAn\_C\_8647.2 y HA77, respectivamente. Debido a la magnitud del coeficiente de determinación de cada uno de ellos, estas regiones pueden considerarse de efecto mediano a grande. Los efectos aditivos estimados fueron de -0,04 para el primero, y -0,19 para los dos últimos. Ambos se encuentran a una distancia de 5,5 cM y pertenecen al mismo intervalo de confianza del 95%.

Sobre el GL 2 y GL9, se identificaron los QTLs qSEV21/2, con un efecto aditivo de -0,05 y un  $R^2$  de 12,05% y qSEV21.13-14/9, con -0,07 y 20,84% para los mismos parámetros.

Sobre el GL10, tres QTLs (qSEV21/10, qSEV21.11-12/10 y qSEV21.12-13/10 cercanos a ORS591, GLP\_4 y E41M50\_3, respectivamente) explican entre el 16,2% y el 30% cada uno. El valor de aditividad para qSEV21.11-12/10 fue estimado en +0,16, mientras que para qSEV21/10 y SEV21(12/13)-10 fue de -0,05 y -0,07, respectivamente.

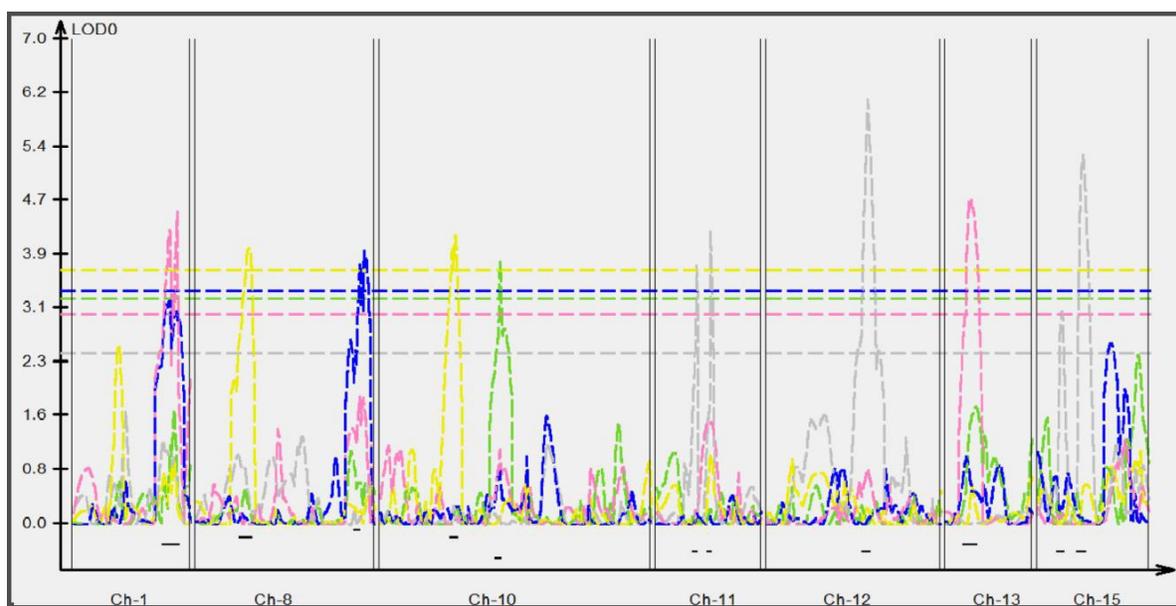
VARIABLE	NOMBRE	GL	cM	LOD SCORE	EFEECTO	R <sup>2</sup>	TAMAÑO INTERVALO (cM)	MARCADOR CERCANO	cM a >LOD*
					ADITIVO				
SEV21	qSEV21/2 qSEV21/10	2	102,8	3,4146	-0,0478	0,1205	7,7	E40M62_17 ORS591	0,78
		10	124,4	4,5697	-0,0564	0,1697	6,6		0,03
SEV21(09/10)	qSEV21.09-10/1 qSEV21.09-10/16	1	35,4	5,3559	-0,0445	0,2492	2,2	E37M47_15 HA4222	0,01
		16	142,1	3,4409	0,0366	0,139	5,4		0,03
SEV21(11/12)	qSEV21.11-12/1a qSEV21.11-12/1b	1	99,3	5,672	-0,1882	0,2912	11	HeAn_C_8647.2 HA77	0,01
		1	104,8	4,6323	-0,1872	0,281	11		1,52
	qSEV21.11-12/10	10	271,1	4,954	0,1644	0,2965	6	GLP_4	3,88
SEV21(12/13)	qSEV21.12-13/10	10	150,4	4,2265	-0,079	0,162	0,7	E41M50_3	0,05
SEV21(13/14)	qSEV21.13-14/9	9	164,4	4,8614	-0,0735	0,2084	2,3	E38M48_5	0,02

Por último, sobre el GL 16 para el QTL qSEV21.09-10/16 muy próximo al marcador HA4222, se estimó una variación fenotípica del 13,9% y un efecto aditivo de +0,04.

### 3.3 QTLs ASOCIADOS CON LA CARACTERÍSTICA INTENSIDAD

Un total de 12 QTLs se identificaron a lo largo de los GLs 1, 8, 10, 11, 12, 13 y 15 (LODs= 3,08 – 6,13 e IC95%= 2-13,2 cM, Figura 2.7 y Tabla 2.9).

Sobre el GL 1, dos QTLs (qINT21.11-12/1a, próximo al SNP HeAn\_C\_8647.2 y qINT21.11-



**FIGURA 2.7. GRÁFICO DESCRIPTIVO DE POSICIONAMIENTO Y MAGNITUD DE LOS QTLs DETECTADOS POR CIM PARA INT21.** Sobre el eje y se disponen los valores de LOD mientras que en el eje x se describe la posición en los grupos de ligamientos ("Ch-" en la Figura). Las líneas horizontales representan los umbrales de LOD por sobre los cuales un QTL es considerado real. Referencia de colores: Gris: campaña 2009-10; Rosado: campaña 2011-12; Amarillo: campaña 2012-13; Azul: campaña 2013-14; Verde: Promedio de las campañas.

12/1b, cercano al SSR ORS662) que pertenecen al mismo intervalo de confianza del 95%. Estos explican el 15,65% y 18,66% de la varianza, siendo sus efectos aditivos de -0,02 cada uno.

Sobre el GL8, se sitúan los QTLs qINT21.12-13/8 y qINT21.13-14/8. El primero cuyo efecto aditivo se calculó en -0,06 explicaría 16,17% de la variación fenotípica y el segundo, cuyo

TABLA 2.9. Descripción de los QTLs detectados para INT21. *Distancia a punto de mayor LOD (cM)									
VARIABLE	NOMBRE	GL	cM	LOD SCORE	EFEECTO ADITIVO	R <sup>2</sup>	TAMAÑO INTERVALO (cM)	MARCADOR CERCANO	cM a >LOD*
INT21	qINT21/10	10	121,8	3,8106	-0,0348	0,1487	3	SSL39	0,04
INT21(09/10)	qINT21.09-10/11a	11	41,3	3,7364	0,0203	0,1523	2,5	E32M49_2	0,55
	qINT21.09-10/11b	11	54,8	4,2459	-0,021	0,1738	2	E35M48_3	0,05
	qINT21.09-10/12	12	103	6,1303	0,017	0,2345	5,2	E33M48_22	0,05
	qINT21.09-10/15a	15	24,6	3,077	-0,0142	0,1205	4,8	SSU25	0,26
	qINT21.09-10/15b	15	46,5	5,3331	0,02	0,2305	6,8	ORS401	0,02
INT21(11/12)	qINT21.11-12/1a	1	98,7	4,252	-0,0164	0,1565	11,9	HeAn_C_8647.2	0,61
	qINT21.11-12/1b	1	106,8	4,5172	-0,0187	0,1866	11,9	ORS662	0,43
	qINT21.11-12/13	13	24,3	4,7124	0,0185	0,205	10,7	E33M48_20	1,69
INT21(12/13)	qINT21.12-13/8	8	54,5	3,9945	-0,0617	0,1617	10,7	E37M61_7	1,27
	qINT21.12-13/10	10	76,1	4,1708	-0,0702	0,2014	10,8	E32M61_7	1,47
INT21(13/14)	qINT21.13-14/8	8	167,7	3,7828	-0,0498	0,1476	13,2	HA3278	0,01

valor de aditividad se estimó en -0,05, el 14,76%.

Sobre el GL10, se hallaron los QTLs qINT21/10 y qINT21.12-13/10. El primero asociado con un efecto aditivo de -0,03 y explicando 14,87% de la variación fenotípica mientras que el segundo, cuyo efecto aditivo fue estimado en -0,07, explica 20,14% de la variación.

Sobre el GL 11, se mapearon el QTL qINT21.09-10/11a y el qINT21.09-10/11b. El primero con un efecto aditivo de +0,02 explica el 15,23% de la variación fenotípica mientras que el efecto aditivo del segundo fue de -0,02 y explica 17,38% de la variación.

Sobre los GLs 12 y 13, se identificaron los QTLs, qINT21.09-10/12 y qINT21.11-12/13. Para ambos el efecto aditivo fue calculado en +0,02 y un coeficiente de determinación de 23,45% y 20,5% respectivamente. Por último, en el GL 15 se ubicaron dos QTLs, el qINT21.09-10/15a (cercano al SSR SSU25) y el qINT21.09-10/15b (cercano al SSR ORS401). El primero con un efecto aditivo de -0,01 y explica el 12,05% de la variación obtenida, mientras que para el segundo el valor del efecto aditivo se estimó en +0,02 y explicó un 23,05% de la varianza fenotípica.

### 3.4 QTLs ASOCIADOS CON LA CARACTERÍSTICA PERÍODO DE INCUBACIÓN

Para esta característica, 6 QTLs fueron descriptos sobre 4 GLs (GLs 3, 8, 10 y 15). Los valores de LOD obtenidos oscilaron entre 3,48 y 4,82 y los intervalos de confianza del 95% entre 2,7 y 12,7 cM (Figura 2.8 y Tabla 2.10).

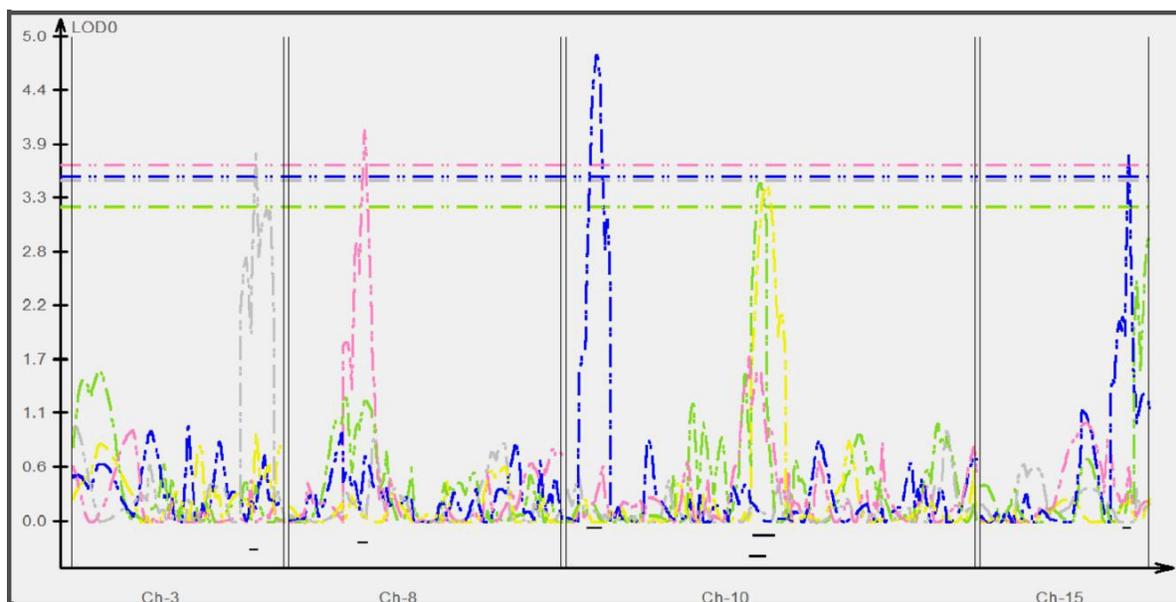
Sobre el GL 3, se situó el QTL qPI(09-10)-3, próximo al SNP HeAn\_C\_185.1, su efecto aditivo fue calculado en +0,84 unidades de la variable y capta el 19,53% de la variación total.

Sobre el GL 8 se identificó el QTL qPI.11-12/8 cercano a HA911, su efecto aditivo es +1,77 DPIs explicando el 21,26% de la variación fenotípica.

Ubicados en el GL 10, los QTLs qPI.12-13/10 y qPI.13-14/10 (próximo a E40M50\_11, E40M50\_11 y HA1108, respectivamente) cuyos efectos aditivos oscilaron entre 0,85 y 1,18 DPIs explican entre el 16,61% y 19,53% de la variación.

Por último, sobre el GL 15, se identificó el QTL qPI.13-14/15, su efecto aditivo fue de +0,98 DPIs a la característica y explica un 19,09% de la variación. La región genómica qPI.12-13/10 se identificó nuevamente usando el promedio de las campañas, en este caso denominada qPI/10.

La Figura 2.9 resume los diversos QTLs encontrados sobre los GLs que integran el mapa,



**FIGURA 2.8. GRÁFICO DESCRIPTIVO DE POSICIONAMIENTO Y MAGNITUD DE LOS QTLs DETECTADOS POR CIM PARA PI.** Sobre el eje y se disponen los valores de LOD mientras que en el eje x se describe la posición en los grupos de ligamentos ("Ch-" en la Figura). Las líneas horizontales representan los umbrales de LOD por sobre los cuales un QTL es considerado real. Referencia de colores: Gris: campaña 2009-10; Rosado: campaña 2011-12; Amarillo: campaña 2012-13; Azul: campaña 2013-14; Verde: Promedio de las campañas.

exceptuando los GLs 4, 5 y 6 donde no ha sido posible comprobar la presencia de regiones genómicas asociadas a alguna característica evaluada.

**TABLA 2.10. Descripción de los QTLs detectados para INT21. \*Distancia a punto de mayor LOD (cM)**

VARIABLE	NOMBRE	GL	cM	LOD SCORE	EFEECTO ADITIVO	R <sup>2</sup>	TAMAÑO INTERVALO (cM)	MARCADOR CERCANO	cM a >LOD*
PI	qPI/10	10	130,4	3,5063	0,8502	0,1953	9	E40M50_11	4,23
PI(09/10)	qPI.09-10/3	3	123,7	3,8047	0,8426	0,1509	3,2	HeAn_C_185.1	0,04
PI(11/12)	qPI.11-12/8	8	50	4,0454	1,7703	0,2126	4,9	HA911	0,06
PI(12/13)	qPI.12-13/10	10	134,6	3,4846	1,0761	0,1661	12,7	E40M50_11	0,03
PI(13/14)	qPI.13-14/10	10	19,7	4,8218	1,1768	0,1909	7,4	HA1108	0,99
	qPI.13-14/15	15	99,4	3,7788	0,9758	0,1344	2,7	E38M48_4	0,04

### 3.5 EFICIENCIA DE LA SELECCIÓN INDIRECTA

La Tabla 2.11 muestra los p-valores obtenidos para la prueba de medias realizadas. Puede observarse que para 2/21 marcadores asociados evaluados, las medias no fueron significativamente diferentes. Para los 19 restantes, PAC2 (de comportamiento medianamente resistente) aportó los alelos que determinaron una mayor tolerancia en 16 casos (84,21% considerando al marcador HeAn\_C\_8647.2 asociado a dos QTLs). El otro 15,79% fue aportado por los alelos pertenecientes al parental susceptible (RHA266). Estos últimos marcadores asociados, son los que aportaron un efecto aditivo significativo incrementó la media de la incidencia, severidad e intensidad en las campañas señaladas.

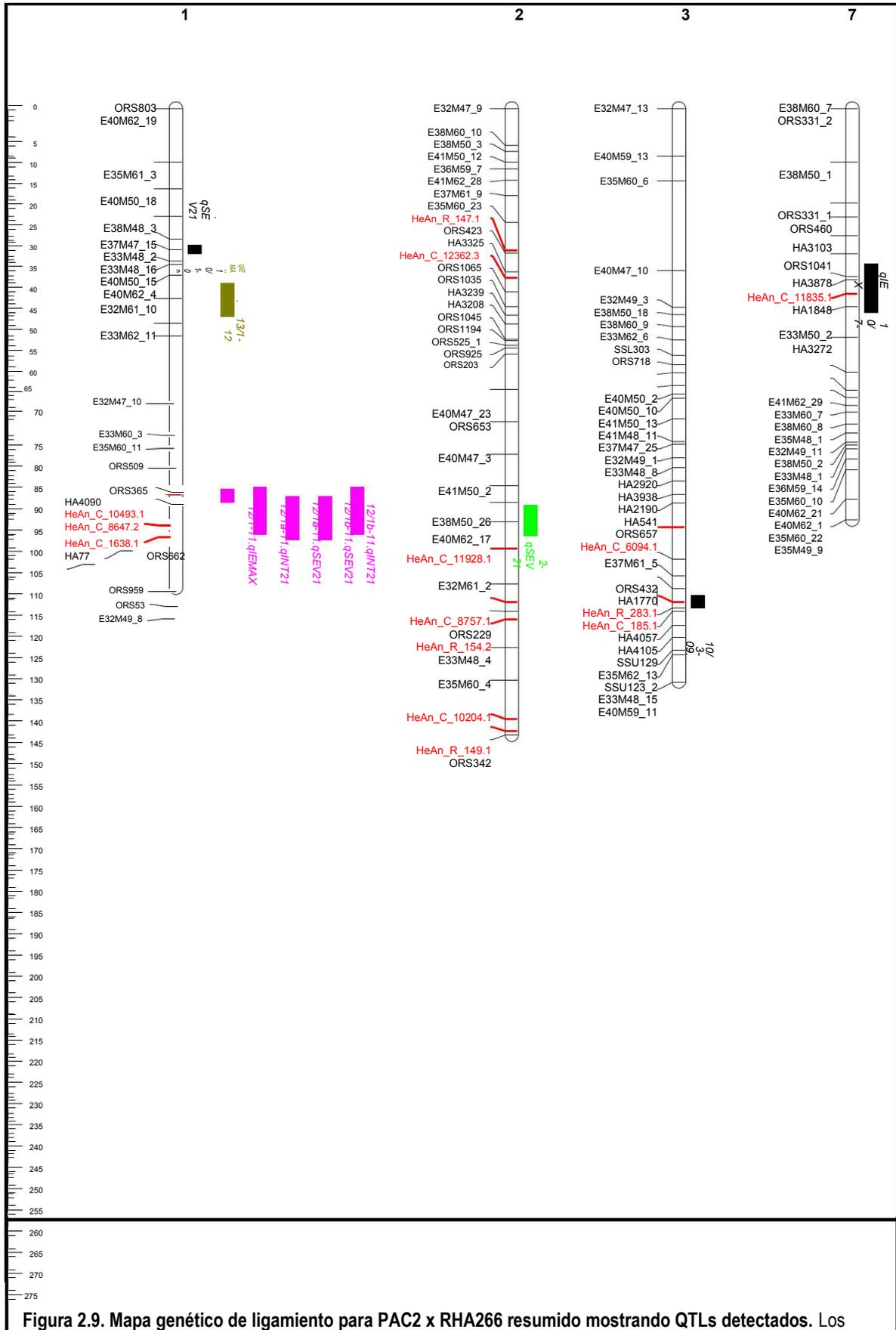
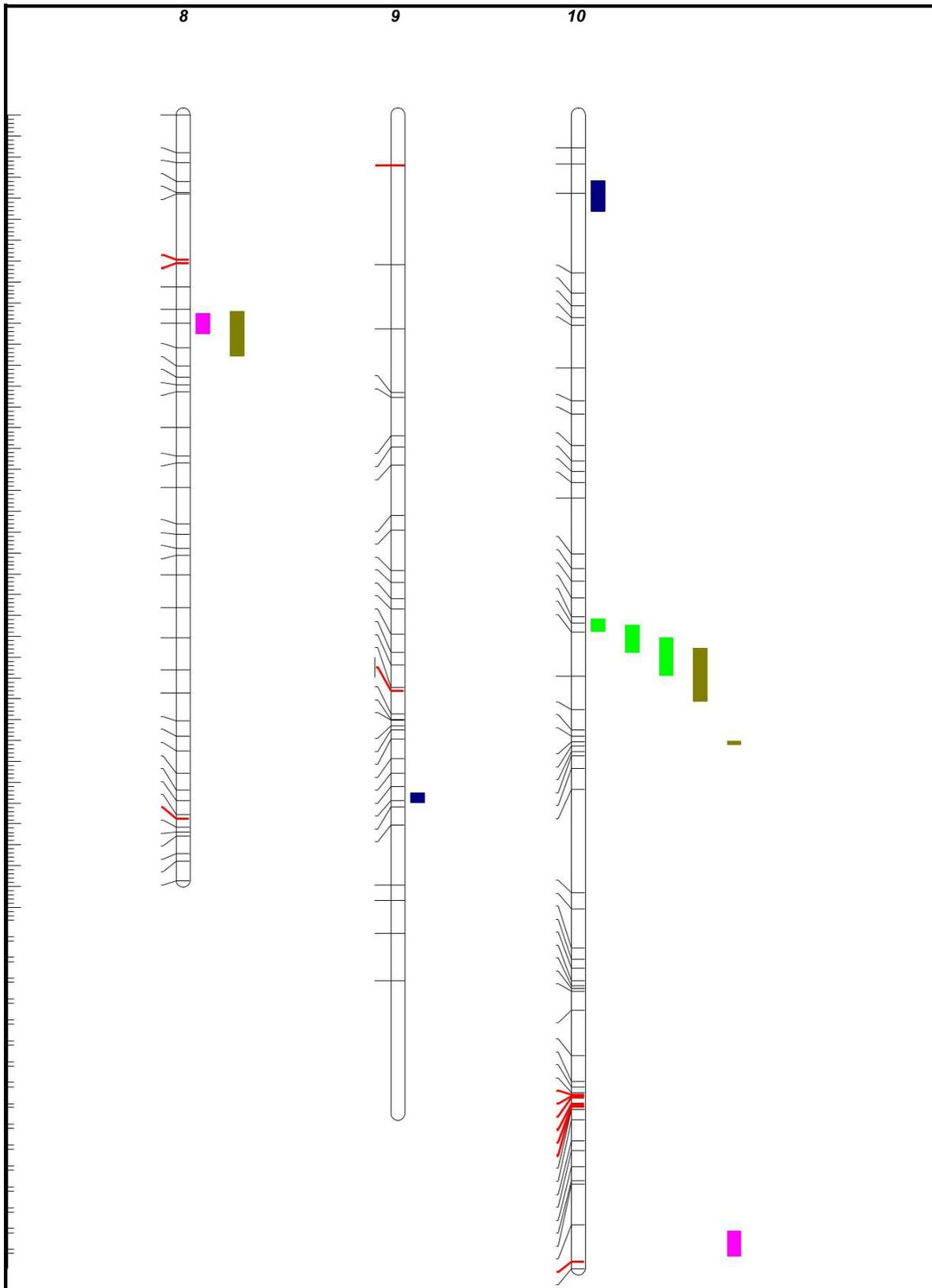


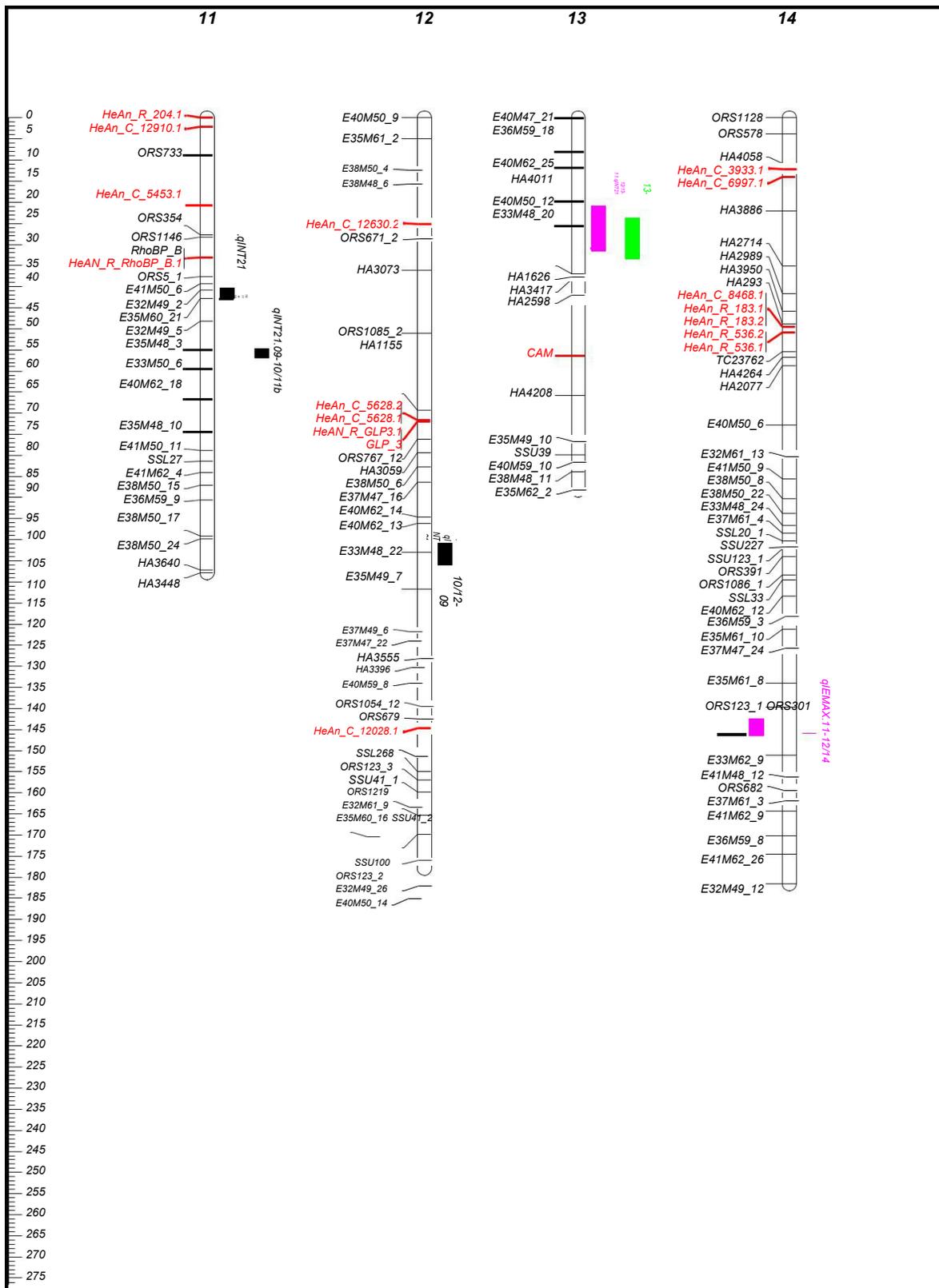
Figura 2.9. Mapa genético de ligamiento para PAC2 x RHA266 resumido mostrando QTLs detectados. Los

marcadores molecular mapeados se listan a la izquierda. En rojo se presentan los 69 SNPs mapeados en este trabajo. Los intervalos determinados por los QTLs detectados son señalados por las barras verticales. La leyenda de cada barra indica la característica para la que fue detectada el QTL.



**Figura 2.9 (cont.). Mapa genético de ligamiento para PAC2 x RHA266 resumido mostrando QTLs detectados.**

Los marcadores moleculares mapeados se listan a la izquierda. En rojo se presentan los 69 SNPs mapeados en este trabajo. Los intervalos determinados por los QTLs detectados son señalados por las barras verticales. La leyenda de cada barra indica la característica para la que fue detectada el QTL.



**Figura 2.9 (cont.). Mapa genético de ligamiento para PAC2 x RHA266 resumido mostrando QTLs detectados.** Los marcadores molecular mapeados se listan a la izquierda. En rojo se presentan los 69 SNPs mapeados en este trabajo. Los intervalos determinados por los QTLs detectados son señalados por las barras verticales. La leyenda de cada barra indica la característica para la que fue detectada el QTL.

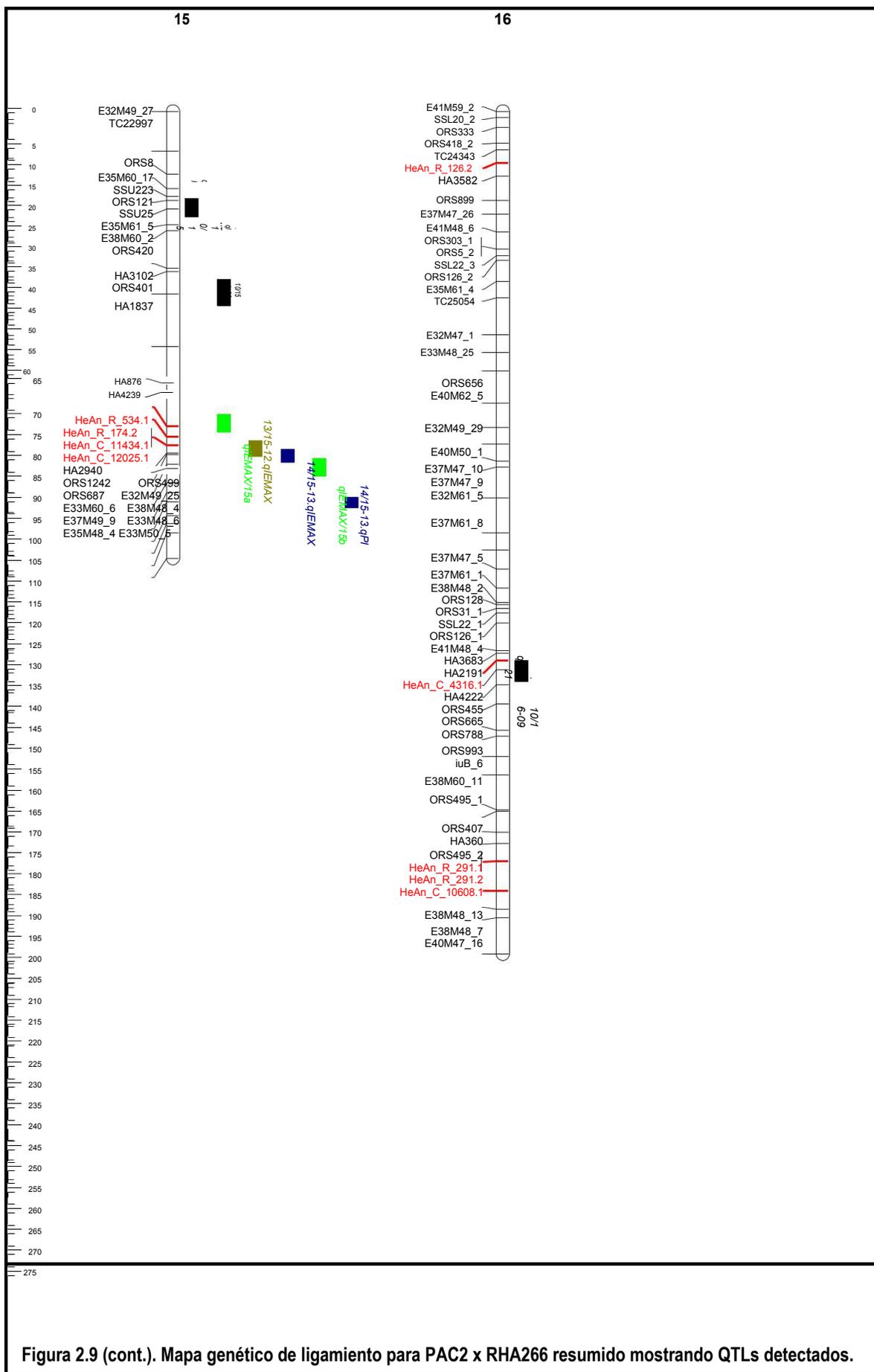


Figura 2.9 (cont.). Mapa genético de ligamiento para PAC2 x RHA266 resumido mostrando QTLs detectados.

Los marcadores molecular mapeados se listan a la izquierda 1.12. En rojo se presentan los 69 SNPs mapeados en este trabajo. Los intervalos determinados por los QTLs detectados son señalados por las barras verticales. La leyenda de cada barra indica la característica para la que fue detectada el QTL.

TABLA 2.11. Prueba de medias entre alelos aportados por los parentales. ns = no significativo.						
QTL	MARCADOR CERCANO	LG	MEDIA ALELO PAC2	MEDIA ALELO RHA266	p-valor	Aporte
qIEMAX/15a	HeAn_R_534.1	15	0,31	0,46	0,0001	PAC2
qIEMAX/15b	ORS687	15	0,31	0,46	0,0001	PAC2
qIEMAX.09-10/7	ORS1041	7	0,6	0,45	0,0097	RHA266
qIEMAX.11-12/1	HeAn_C_10493.1	1	0,03	0,12	0,0008	PAC2
qIEMAX.11-12/14	ORS301	14	0,04	0,12	0,0088	PAC2
qIEMAX.12-13/15	HeAn_C_12025.1/ HeAn_C_11434.1	15	0,55	0,75	<0,0001	PAC2
qIEMAX.13-14/15	ORS1242	15	0,35	0,57	0,0003	PAC2
qSEV21/10	ORS591	10	0,27	0,37	0,0021	PAC2
qSEV21.09-10/16	HA4222	16	0,21	0,15	0,044	RHA266
qSEV21.11-12/10	GLP_4	10	0,4	0,23	0,083 ns	RHA266
qSEV21.11-12/1b	HA77	1	0,19	0,42	0,0171	PAC2
qINT21.11-12/1a	HeAn_C_8647.2	1	0,15	0,44	0,0002	PAC2
qINT21/10	SSL39	10	0,08	0,15	0,0038	PAC2
qINT21.09-10/15b	ORS401	15	0,04	0,02	0,0294	RHA266
qINT21.09-10/15a	SSU25	15	0,03	0,04	0,379 ns	PAC2
qINT21.11-12/1a	HeAn_C_8647.2	1	0,02	0,05	0,0013	PAC2
qINT21.11-12/1b	ORS662	1	0,02	0,05	0,0286	PAC2
qINT21.13-14/8	HA3278	8	0,09	0,23	0,005	PAC2
qPI.09-10/3	HeAn_C_185.1	3	23,54	22,41	0,0213	PAC2
qPI.11-12/8	HA911	8	20,76	17,82	0,012	PAC2
qPI.13-14/10	HA1108	10	21,91	19,62	0,0037	PAC2

### 3.6 ANÁLISIS *IN SILICO* DE LAS REGIONES GENÓMICAS COMPRENDIDAS POR LOS QTLs DETECTADOS

Un total de 39 marcadores (16 SNPs y 23 SSRs) fueron seleccionados dada su ubicación subyacente a los intervalos de confianza determinados por los QTLs (Tabla 2.12). El análisis *in silico* de sus secuencias (query) permitió caracterizar a cada uno de estos marcadores (siempre y cuando haya sido posible) describiendo su ontología génica (GO) y predecir su función putativa.

TABLA 2.12.Descripción de marcadores seleccionados para el análisis <i>in silico</i>			
MARCADOR EN INTERVALO	LG	cM	QTL ASOCIADO
HeAn_C_10493.1	1	96,7	qIEMAX.11-12/1 qINT21.11-12/1a qINT21.11-12/1b
HeAn_C_8647.21		99,31	qIEMAX.11-12/1 qINT21.11-12/1a qSEV21.11-12/1a qSEV21.11-12/1b
HeAn_C_1638.1	1	100,8	qINT21.11-12/1a qSEV21.11-12/1a qSEV21.11-12/1b qINT21.11-12/1b
HA77	1	103,28	
ORS662	1	107,23	qSEV21.11-12/1a qSEV21.11-12/1b
HeAn_C_11928.1	2	110,37	qSEV21/2
HeAn_R_283.1	3	123,65	qPI.09-10/3
HeAn_C_185.1	3	123,66	
HA4057	3	125,26	
HA4105	3	126,09	
ORS1041	7	42,07	qIEMAX.09-10/7
HA3878	7	43,02	
HeAn_C_11835.1	7	46,47	
HA1848	7	49,7	
HA911	8	49,94	qPI.11-12/8 qINT21.12-13/8
HeAn_C_11352.1	8	168,74	qINT21.13-14/8
HA503	8	170,73	
HA3621	8	171,89	
ORS894_2	8	173	
HA1108	10	18,71	qPI.13-14/10
HeAn_C_4498.1	10	24	
HeAn_C_12824.1	10	25,39	
ORS437	10	124,08	qINT21/10 qSEV21/10
ORS591	10	124,43	qINT21/10 qSEV21/10 qPI/10 qPI.12-13/10
ORS691	10	266,1	qSEV21.11-12/10
GLP_4	10	274,98	
ORS301	14	143,43	qIEMAX.11-12/14
HA3102	15	40,68	qINT21.09-10/15b
ORS401	15	46,52	
HeAn_R_534	115	80,23	qIEMAX/15 <sup>a</sup>
HeAn_R_174.2	15	82,78	
HeAn_C_11434.1	15	84,95	qIEMAX/15a qIEMAX.12-13/15
HeAn_C_12025.1	15	84,95	
HA2940	15	86,55	qIEMAX.12-13/15 qIEMAX.13-14/15
ORS1242	15	87	qIEMAX.12-13/15 qIEMAX.13-14/15 qIEMAX/15b
ORS499	15	87,32	qIEMAX.13-14/15 qIEMAX/15b
ORS687	15	89,83	qIEMAX/15b
HA2191	16	137,87	qSEV21.09-10/16
HeAn_C_4316.1	16	139,72	

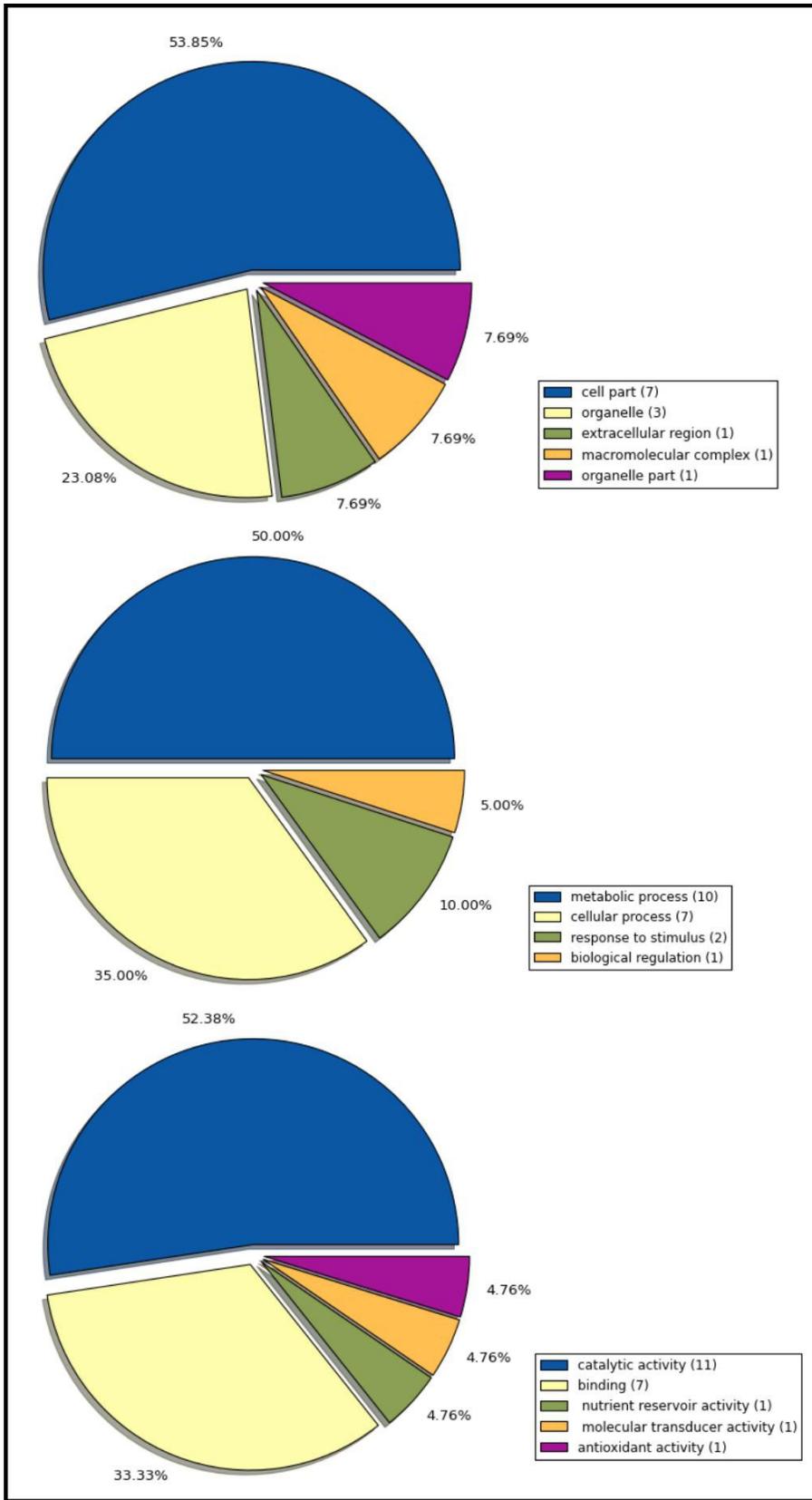
En primer lugar, el programa Blast2Go (y sus herramientas Blast e InterProScan) empleado comparó de manera automática e individualmente las seis secuencias de aminoácidos (obtenidas de la traducción de la secuencia query en sus 6 marcos de lectura hipotéticos) contra una base de datos de proteínas redundante y obtuvo una lista resultados positivos de dicha comparación (hits), todos ellos superando los umbrales descriptos en la sección de materiales y métodos. Únicamente se encontraron hits para 16 secuencias query, todos ellos

provenientes de marcadores SNPs y el paso siguiente de asignación de términos GO fue exitoso para 15 de estos (HeAn\_C\_11352.1 no presentó términos GO asociados).

En segundo lugar, las 39 secuencias query fueron evaluadas con la herramienta blast incorporada en la base de datos de “Heliagene” y los parámetros establecidos en la sección materiales y métodos. Los resultados obtenidos permitieron corroborar las anotaciones GO asignadas en el análisis previo para 15 de los 16 GCs conteniendo SNPs seleccionados.

De esta manera, ambas aproximaciones permitieron la asignación de términos de ontología génica representando un concepto clave en biología molecular: la función molecular de los productos génicos; su rol en los procesos biológicos; y su localización en componentes celulares. La Tabla V del apéndice resume los resultados obtenidos mientras que la Figura 2.10 muestra la distribución de los marcadores SNPs evaluados en las 3 categorías mencionadas y secciones pertenecientes a cada una de ellas (solo el primer nivel de cada categoría es mostrado). Dado que este agrupamiento se realiza teniendo en cuenta los términos GOs obtenidos para cada secuencia, es posible que la secuencia un mismo gen candidato query, se vea contenida en diferentes subcategorías (dado a que estas no son mutuamente excluyentes).

Los análisis efectuados demostraron que de los 16 GCs query, 7 de los mismos se vieron involucrados en procesos de oxido-reducción en vías de respuesta a estrés oxidativo, con una función descrita de oxidoreductasa o deshidrogenasa (HeAn\_C\_10493.1, HeAn\_C\_12025.1, HeAn\_R\_534.1, HeAn\_C\_1638.1, HeAn\_C\_4316.1, HeAn\_C\_11835.1 y HeAn\_R\_283.1). Los genes candidatos HeAn\_C\_12025.1, HeAn\_C\_11434.1 y HeAn\_C\_185.1 fueron caracterizados con funciones putativas quinasa, mientras que GLP\_4, identificada como una proteína del tipo Germina, se asoció con términos GOs referidos a la unión de iones metálicos y a estructuras externas de encapsulamiento. Por su parte HeAn\_C\_4498.1, fue descrito como un factor de transcripción con dominio BHLH y a los genes HeAn\_R\_174.2 y HeAn\_C\_8647.2, se los asoció con funciones de unión a adenilribonucleótidos. HeAn\_C\_11928.1 y HeAn\_C\_12824.1 se describieron como una proteína vinculada al fotosistema II y una enzima con actividad metiltransferasa. HeAn\_C\_11352.1, no obtuvo asignación de términos GO, pero su hit en BLAST correspondió a un transportador de prolina en *A. thaliana*.



**FIGURA 2.10. DISTRIBUCIÓN DE ANOTACIONES DE ONTOLOGÍA GÉNICA.** Porcentajes de términos GO asignados por categoría: Componentes Celulares (arriba), Procesos Biológicos (medio) y Función Molecular

Por otra parte, la inclusión de los marcadores SSRs en el análisis de comparación efectuado contra la base de datos de Heliagene, permitió la obtención de hits para 11 de ellos (Tabla VI del apéndice). En este punto, se debe tener en cuenta que para los hits encontrados el porcentaje promedio de cobertura de la secuencia query fue de 38%, aunque mostraron una alta identidad y valores E inferiores a  $7,09e^{-14}$ . Esto puede deberse a que la comparación involucra regiones genómicas contra regiones transcritas (provenientes del ensayo de genotipificación del transcriptoma de la línea Ha412) y las diferencias entre secuencias puede estar dada por la presencia de intrones en la secuencia del SSR.

A cinco de estas secuencias query se les pudo asignar términos de ontología génica tanto GO como InterProScan, excepto una (ORS437) que no contó con anotación GO. De ellas se destacan las asignaciones de los marcadores HA1108 y HA3102 con actividad peroxidasa (involucrada en respuesta a estreses oxidativos) y factor de transcripción putativo, respectivamente. Por su parte, HA4057 predijo una actividad ligasa arginina-ARNt, HA911 se vinculó con un motivo de unión a zinc y ORS437 con una transferasa de fosfatidilinositol.

Por último, para los 23 SSRs objeto de este estudio se llevó un análisis exhaustivo de la zona aledaña a la localización de dicho marcador en los mapas previamente publicados. La Tabla VII del apéndice, resume las características encontradas.

Aproximadamente la mitad (11/23) de los SSRs analizados, no se encontraron localizados sobre los mapas públicos o no se describieron previamente QTLs o genes cercanos al marcador. Para los restantes, la mayoría de las características encontradas se relacionan con caracteres morfológicos. Tal es el caso de los marcadores ORS662, ORS1041, HA1848, ORS894\_2, ORS437, ORS691 y ORS499 que se vieron situados cercanos a regiones involucradas en características como tamaño, peso, largo del grano, días a floración, número de lígulas, número de flores radiales, entre otras (Burke et al. 2005; Wills & Burke 2007; Ebrahimi et al. 2008; S. Kiani et al. 2007; Lexer et al. 2005; Micic, Hahn, Bauer, Melchinger, et al. 2005).

En otros casos, el estudio de las zonas flanqueantes a los SSRs permitió descubrir genes o QTLs implicados en la resistencia a hongos necrotróficos como *Phoma macdonaldii* (ORS894\_2, HA1108, HA3102, ORS401, ORS687), en la tasa de crecimiento de la lesión ocasionada por un patógeno fúngico (ORS662) o en la concentración de calcio (ORS591), implicado en el ataque del hongo (Micic et al. 2004; Alfadil et al. 2007).

# DISCUSIÓN

## 1. ANALISIS FENOTÍPICOS

Los ensayos de caracterización fenotípica se realizaron en el campo experimental de la EEA Balcarce y como primer paso de este proceso se tomaron las consideraciones y medidas necesarias para garantizar la disponibilidad de un número mínimo de plantas por parcela que asegurará el éxito de los ensayos. Con este propósito, las semillas almacenadas y preservadas a 8°C fueron tratadas 5 días antes de la siembra con una solución de ácido giberélico con el fin de romper la dormición. Distintos autores han descripto que el tratamiento con bajas concentraciones de ácido giberélico aumenta significativamente la energía germinativa (Debeaujon & Koornneef 2000), resultando particularmente efectivo en la promoción de la germinación, tanto en los casos en los que el origen de la dormición es la cubierta seminal o es el embrión (Roberts 1988).

Además, se realizó un tratamiento preventivo para enfermedades fúngicas que son de alta ocurrencia en la zona del ensayo como mildiu, con una dilución de Metalaxil M. Si bien este tratamiento fue efectivo para controlar la enfermedad, ocasionalmente se detectaron plantas enanas (signo evidente de la infección por mildiu) debido probablemente a la aparición de ciertas razas de *Plasmophara haldsteii* con resistencia al Metalaxil M (Molinero-Ruiz & Dominguez 2005).

El protocolo de inoculación asistida consistió en asperjar ascosporas del hongo en el momento de mayor susceptibilidad de la planta que es cuando las inflorescencias están en estadio R5.2. Llegado este momento es necesario que la parcela tuviera un número considerable de plantas en este estadio, no obstante debido a la variabilidad propia de las RILs, cada parcela alcanzó el estadio fenológico R5.2 en diferentes momentos. Esto ocasionó un desfase en las fechas de inoculación que abarcó aproximadamente de 2 a 3 semanas en cada campaña.

Diversos estudios han demostrado que la inoculación asistida es el método más adecuado para evaluar la resistencia a *S. sclerotiorum* en líneas e híbridos comerciales frente a alternativas como el test de micelios, infección seminatural o la inyección de ácido oxálico en la cara dorsal del capítulo (Van & Miller 2004; Castaño et al. 2001; Vear et al. 2007). Además, Vear & Tourvieille (1988; 1984) y han demostrado que los resultados obtenidos mediante inoculación asistida reproducen fielmente la infección natural en el patosistema girasol-*Sclerotinia*.

Debe considerarse, sin embargo, que la inoculación de cerca de 2.500 ascosporas fue 5 veces menor a la empleada en los estudios desarrollados por Castaño et al (2001) y Vear et al (2007). El uso de una mayor concentración de ascosporas en el inóculo, sumado al cubrimiento de los capítulos inoculados con bolsas de papel, generaría un ambiente más severo de infección y el consecuente incremento en la incidencia. Esta condición asegura contar con un número elevado de plantas enfermas necesario para estimar precisión el período de incubación de la enfermedad (Castaño et al. 1993), sin embargo dichos valores se traducirían en una reducción de variabilidad para la incidencia de la enfermedad.

Se sabe que los mecanismos de resistencia a PHC pueden ser separados en dos fases. La primera se inicia con la infección y es cuando actúan las primeras barreras del hospedante, y la segunda, donde la planta se enfrenta a la infección sistémica del patógeno (Castaño et al. 2001). Considerando estas etapas, se seleccionaron para su medición las variables incidencia (IE), severidad (SEV), intensidad (INT) y período de inoculación (PI) en función de su capacidad para describir los dos mecanismos planteados.

A partir de las limitantes observadas durante las mediciones efectuadas en los ensayos en el campo derivadas de la presencia de manchas sobre el capítulo y la imposibilidad de obtener un registro dado que la línea llegó a madurez fisiológica, entre otras que impidieron o dificultaron el registro visual de las variables a los días 10, 14, 17, 19, 21, 24 y 28 post inoculación (DPI), se decidió usar únicamente en los estudios de comportamiento fenotípico y el análisis de QTL la variable incidencia máxima ("IEMAX").

En cuanto a SEV e INT, fueron consideradas las nuevas variables "SEV21" e "INT21" como la severidad e intensidad de la enfermedad PHC al día 21 post-inoculación (DPI 21). Evaluaciones posteriores a este punto mostraron un retroceso en el porcentaje del capítulo infectado en la mayoría de las líneas por las causas ya detalladas. Es así que se determinó empíricamente al DPI 21 como óptimo para evaluar estas variables y evitar errores propios de la subjetividad.

Dos consideraciones debieron hacerse sobre la campaña 2009-2010. La primera, fue que las líneas en esta campaña no fueron evaluadas siguiendo el esquema planteado de 10, 14, 17, 19, 21, 24 y 28 DPIs. Por el contrario, todo el ensayo fue evaluado a días fijos del calendario. Entonces, debido al desfase en la fecha de inoculación ya mencionado, el día de evaluación se correspondió a diferentes DPIs para cada una de las parcelas (líneas) evaluadas y produjo menor cantidad de datos por DPI.

El segundo punto destacado fue que esta campaña se sembró con antelación (en la fecha relativa del año) respecto de las demás campañas. Esto ocasionó un ciclo más largo y por ende fechas de evaluaciones que superaron al DPI 28.

Por otra parte, los ensayos a campo realizados durante el transcurso de este trabajo de Tesis se previeron de manera tal que las RILs que integran cada uno de ellos, llegaran a estar presentes (y por lo tanto ser evaluadas) en todas las campañas. Estos ensayos “equilibrados” tenían por objeto simplificar los análisis estadísticos y permitir la comparación de todos los materiales en la totalidad de campañas (ambientes), generando así la mayor información posible.

Esta premisa no pudo ser cumplida en su totalidad debido a varios factores que impidieron la obtención de un registro completo para toda la población a lo largo de todas las campañas. Entre las limitantes se pueden mencionar la ausencia de una cantidad mínima de semillas necesaria para completar al menos una de las repeticiones, el bajo poder germinativo de las semillas de las RILs (aun habiendo sido tratadas con ácido giberélico), el proceso de reestructuración de la población por pérdida de homogeneidad de las líneas ocurridas durante las multiplicaciones que se llevó en paralelo a las campañas de fenotipificación, la disminución en el número total de plantas por parcela debido a la presencia de predadores naturales y enfermedades que afectan a la integridad de la planta. En efecto, si bien las variables fueron analizadas en numerosos genotipos que estuvieron presentes a lo largo de todas las campañas, para otros solamente pudieron ser registradas en sólo en un ensayo y/o una sola repetición. Este hecho provocó que los análisis conjuntos de la variancia a través de ambientes fueran desbalanceados, incrementando la variancia correspondiente al error experimental.

El análisis exploratorio y preliminar de los datos (Tabla 2.3) permitió corroborar la distribución continua que presentan los caracteres SEV21 e INT21 en la población segregante utilizada en este trabajo y establecer relaciones entre las características estudiadas.

El análisis de correlaciones mostró resultados significativos para todos los valores de los coeficientes de Pearson obtenidos (Tabla 2.4). Dado a que este hecho es probable que se deba al gran tamaño de la muestra analizada, se procedió clasificar a los coeficientes de correlación en 3 categorías (baja-moderada-alta) para su correcta interpretación. Es así que se encontró altas asociaciones entre INT21 e IEMAX ( $r = +0,64$ ) e INT21 y SEV21 ( $+0,8$ ). Lo último es esperable, ya que ambas variables consideran para su cálculo el porcentaje infectado del

capítulo. Una correlación moderada fue observada entre la variable PI y las variables que dependen del porcentaje del capítulo infectado ( $r = -0,67$  para SEV21 y  $-0,64$  para INT21). Por último, las características asociadas con IEMAX fueron las más débilmente relacionadas ( $r = 0,41$  para SEV21 y  $r = -0,45$  para PI).

El análisis descriptivo, es un estudio que complementó a la información descriptiva de las variables y proveyó una fuente adicional de información de las asociaciones estadísticas entre estas variables consideradas.

El estudio exploratorio de los datos permitió también, observar valores promedios más uniformes en las variables registradas para las campañas 2012-13 y 2013-14. En cambio, los valores de IEMAX e INT21 registrados en la campaña 2011-12 fueron claramente inferiores al resto de las campañas lo que puede relacionarse con las condiciones imperantes durante el experimento. Se ha descrito que el desarrollo de PHC, se ve favorecido a un nivel determinado de humedad atmosférica, precipitaciones y temperatura media en el momento de inoculación (VEAR & TOURVIEILLE DE LABROUHE 1984). Es por eso que el protocolo de inoculación, indica la necesidad de cubrir el capítulo y suplementar con riego para asegurar el ambiente de humedad necesaria para el desarrollo de la enfermedad. La campaña en cuestión muestra valores promedio superiores a las demás campañas en el lapso en que se realizaron las inoculaciones (Tabla III del Apéndice). El mayor valor promedio de temperatura diaria media ( $23,7^{\circ}\text{C}$ ) y menores valores de precipitaciones ( $0,4-0,8$  mm) durante gran parte del período de inoculación de la campaña 2011-12 habrían tenido gran influencia sobre el escaso desarrollo de la enfermedad. Esto se ve avalado por un estudio realizado por Abawi et al (1979) quienes determinaron que las temperaturas medias deberían oscilar entre  $16$  y  $22^{\circ}\text{C}$ , para un crecimiento adecuado del micelio de *Sclerotinia*.

En cuanto a la humedad relativa promedio, se percibe que fue similar para todos los ensayos, por lo que se asume que esta no ha sido un factor determinante en las variables comentadas.

Del estudio anterior se concluye que el ambiente modifica la magnitud de la resistencia de los materiales respecto de la PHC, por lo que se justifica la realización de experimentos en distintos ambientes para compensar sus efectos (Castaño et al 2001).

En cuanto a la campaña 2009-10, los valores promedios obtenidos para las variables INT21 y PI reflejaron el hecho de que la misma tuvo un ciclo más largo y por lo tanto se retrasó la aparición de los síntomas.

Estas situaciones descritas, tanto para las campañas 2009-10 y 2011-12, afectaron el grado de ataque del hongo de modo tal que se vieron reflejadas en la asimetría y curtosis (elevadas y positivas) que describen la distribución.

Las mismas consideraciones para cada campaña (ambiente) debieron ser tenidas en cuenta para los microambientes definidos por las fechas de inoculación. Para su estudio se utilizaron los valores de IEMAX obtenidos para los híbridos sembrados en la campaña 2012-13.

En primer lugar, cabe aclarar que únicamente se contó con registros para los híbridos comerciales de la campaña 2012-13 ya que no pudieron coordinarse los estadios fenológicos R5.2 de los materiales controles con las fecha de inoculación del ensayo de líneas de la población biparental en los otros tres ensayos.

La Figura 2.2 resume gráficamente el análisis realizado. En ella se observa el comportamiento medio por fecha de inoculación de los materiales clasificados según su origen en control susceptible ACA885 inoculado y no inoculado, control tolerante DK3820 inoculado y no inoculado y RILs. El comportamiento promedio de las RILs por fecha de inoculación fue el esperado, con valores de IEMAX delimitados por el control susceptible inoculado y los demás controles excepto para las dos últimas fechas (18 y 20 de febrero del 2012), donde los valores medios del control resistente inoculado, superaron a la media de las RILs. Este cambio en las magnitudes de respuesta entre las RILs y los controles (interacción) pudo ser detectarse gracias a la siembra y evaluación de los controles pertinentes. Nuevamente, esta interacción detectada entre fechas de inoculación es atribuible mayormente a la temperatura media y precipitaciones de dichas fechas.

En los estudios llevados a cabo con híbridos comerciales la homogeneidad genética de los materiales permite llegar con el número necesario de plantas al estadio R5.2 haciendo factible la inoculación en una única fecha (Castaño et al 2001 y Vear et al 2007), Por el contrario, en los ensayos llevados a cabo en la presente Tesis, la variabilidad de las RILs impidió aplicar la misma metodología por lo que se optó por la incorporación de los términos “fecha de inoculación” y “Genotipo x fecha de inoculación (interacción)” en los modelos empleados para calcular las medias ajustadas en los pasos siguientes.

De esta forma, el nuevo modelo permitió:

- Subsanan la desventaja de no poseer registros de los materiales controles para todas las campañas, lo que hubiera posibilitado la utilización como covariable los valores de IEMAX.
- Evitar que los valores ajustados para cada RIL obtenidos estuvieran confundidos con los efectos propios del microambiente generado por las fechas de inoculación y de la interacción del mismo con los genotipos.

Un hecho que avala la inclusión de los nuevos términos es que independientemente de la variable analizada, un gran porcentaje de la varianza fenotípica fue explicada por ellos (Tabla VIII del apéndice). En todos los casos, la campaña aportó el porcentaje más alto (entre 36,5 y 51,3%), seguido por la interacción y la fecha de inoculación con porcentajes similares (entre 2,6 y 14,4%).

Por otra parte, la proporción de la varianza fenotípica atribuida al genotipo (es decir su heredabilidad en sentido amplio ( $H^2$ ) fue calculada por ensayo. Si bien el genotipo de las RIL es homocigoto, la heredabilidad se expresa en sentido amplio ya que la varianza de los efectos epistáticos fue desconsiderada, (Tabla 2.5) En general la  $H^2$  calculada para la mayoría de los caracteres evaluados osciló entre 13,8 y 60%. Excepto para IEMAX de la campaña 2011-12 e INT21 de las campañas 2009-10 y 2011-12, lo cual es altamente esperable ya que las heredabilidades son validadas para la población en estudio, el ambiente en donde se desarrolló el experimento y del método estadístico para estimarla (Falconer & Mackay 1996).

La estimación de la heredabilidad está influenciada por la diversidad genética de los padres usados para obtener la población. Una población derivada de cruces con padres muy divergentes incrementaría la variación genética. Otro factor que afecta a la heredabilidad es el número de individuos en una población ya que si todos los individuos son evaluados, la verdadera varianza genética de una población puede ser determinada (Fehr 1987).

Para el caso del presente trabajo, la evaluación de los parentales para los caracteres de resistencia a PHC demostró que, si bien los mismos se comportan según lo esperado por estudios previos de caracterización, los valores mostrados no resultaron ser extremos. Esto es debido a la influencia ambiental que enmascara los verdaderos efectos genéticos, de baja magnitud y varianza como lo reflejan algunas de la heredabilidades estimadas (Tabla 2.5). Los valores de  $H^2$  para las variables de resistencia a *S. sclerotiorum* aquí presentados son comparables con los obtenidos por lo Mestries et al (1998), Castaño et al (2001), Bert et al

(2002) y levemente inferiores a los obtenidos por Bert et al (2004) y Micic et al (2005), estos últimos para podredumbre húmeda del tallo (PHT).

El empleo de modelos lineales mixtos (para SEV21 e INT21) y modelos lineales generalizados mixtos (para IEMAX y PI) resultó ser eficiente para el cálculo de las medias ajustadas de 99 líneas RILs.

Las diferencias observadas se presentaron en histogramas por campaña y de manera conjunta para todas las variables analizadas (Figura 2.3).

De acuerdo a estos resultados, se hace evidente que la enfermedad resulta dependiente del ambiente. Al observar la IEMAX de las líneas año a año, todas excepto la campaña 2011-12, mostraron valores variables. Del mismo modo que en el análisis exploratorio inicial, se hace evidente las condiciones que la campaña 2011-12 presentó condiciones menos favorables para el desarrollo de la enfermedad.

Los histogramas de SEV21 e INT21 por año muestran una distribución más normal, respecto a los observados para el carácter anterior. Esto se corresponde con lo esperado, dado que este carácter sigue una distribución normal. Un caso particular se da con la SEV21 de la campaña 2011-12, cuya distribución parece ser soportada por una bi-modal.

Para INT21, las dos últimas campañas muestran distribuciones más amplias que las correspondientes a las dos campañas iniciales. Para las campañas 2009-10 y 2011-12 el desplazamiento de la curva hacia menores valores de intensidad pudo deberse a los factores que se vieron involucrados en estos ensayos y que fueron ampliamente discutidos a lo largo de esta Tesis.

Puede observarse que tanto SEV21 y PI, que dependen de las plantas enfermas únicamente para su registro, se observó una amplia variabilidad de valores.

La segregación transgresiva de los datos fenotípicos respecto del rango establecido por los parentales en cada una de las variables, ofrece información sobre la naturaleza de los genes que controlan la expresión de dicho carácter (Eibach, 1990). Asimismo, los histogramas que presenten desviaciones de igual magnitud hacia ambos lados del rango parental, indican que hay un porcentaje relativamente alto de efectos genéticos aditivos.

Por otra parte, todos los análisis modelando la variable respuesta detectaron diferencias significativas ( $p < 0,001$ ) entre líneas (Tabla IV del apéndice).

En este punto cabe aclarar que las comparaciones múltiples para identificar diferencias significativas entre las medias de cada genotipo fueron realizadas con el método DGC (Di Rienzo, Guzman y Casanoves) que utiliza la técnica multivariada del análisis de conglomerados sobre una matriz de distancia obtenida a partir de las medias muestrales (Di Rienzo et al. 2014). A diferencia de métodos de comparación de medias como Tukey, este método agrupa las medias que no presentan diferencias significativas en conjuntos mutuamente excluyentes, evitando solapamientos. Mayormente, se identificaron dos grupos de líneas significativamente diferentes en sus medias fueron obtenidos (Tabla IV del apéndice). Esto, que resulta en mayor practicidad en la lectura a la hora de encontrar líneas de comportamiento contrastante, pero impidió que las líneas parentales PAC2 y RHA266 cuyas medias ajustadas no resultaron extremas, fueran incluidas en grupos diferentes en todos los análisis.

El análisis con los modelos descriptos y la descripción de los resultados obtenidos, reflejó la alta variabilidad genética presente en las RILs para las variables IEMAX, SEV21, INT21 y PI, por lo que pueden considerarse buenas descriptoras del comportamiento de la población frente a PHC, destacando su utilidad para selección y mapeo de QTLs de resistencia.

La variable incidencia determina la proporción de plantas infectas por parcela. Esta variable, también conocida como “porcentaje de ataque” ha demostrado ser eficiente en los trabajos realizados por Castaño et al 2001, Gentzbittel 1998, Bert 2002, 2004, Yue 2008.

Por su parte la severidad registra el porcentaje promedio de infección por parte del hongo sobre el capítulo. Ella ha sido estudiada exitosamente en un trabajo previo realizado por Yue 2008. Una alternativa fue empleada en los estudios llevados a cabo por Van Beacelare (2004) y Ronicke 2005, donde se obtuvo un registro a partir de una escala discreta (del 1 al 5 y del 1 al 8, respectivamente) indicando desde la ausencia de síntomas hasta capítulos completamente infectados.

En cuanto a la intensidad, que se desprende de la anterior al considerar el promedio de infección del capítulo en todas las plantas de la parcela, esta fue estudiada por Hampel et al (1981), considerando una escala de valores entre 1 y 4 para su evaluación.

Por último, período de incubación o su equivalente “período de latencia”, mide el tiempo promedio transcurrido entre la inoculación y la aparición de síntomas. Una variante de la misma ha sido utilizada por Gentzbittel et al (1998), Castaño et al (2001), Bert et al (2002), Bert et al (2004). Los mismos mostraron el uso efectivo de considerar el “índice de latencia”

(la relación entre el período de incubación de la línea respecto del mismo para los controles susceptible) para el cálculo de medias ajustadas.

## 2. ANÁLISIS DE QTLs

Los caracteres de resistencia a PHC han sido descritos como rasgos cuantitativos, determinados por acción combinada de numerosos genes y efectos ambientales significativos sobre su expresión (Castaño et al. 1993; Gulya et al. 1997).

La posibilidad de manipular los niveles de resistencia a PHC del cultivo depende de la comprensión de las complejas bases genéticas que gobiernan estos caracteres.

Una estrategia que permite detectar variantes génicas responsables de una parte de la variabilidad fenotípica observada, se basa en el análisis del genoma en busca de QTLs. Este tipo de análisis provee información del número de loci y magnitud de sus efectos, la ocurrencia de interacción entre ellos (epistasia) y la presencia de interacciones genotipo x ambiente, cuando se evalúa el carácter en más de un ambiente.

En este trabajo de Tesis se llevó a cabo un análisis de QTLs para caracteres implicados en la resistencia a la PHC a partir de la evaluación de características fenotípicas como la IEMAX, SEV21, INT21 y PI. Las mismas fueron evaluados en una población de RILs y en cinco ambientes (cuatro campañas y un ambiente que considera el promedio de las RILs en todas las campañas) con la metodología de mapeo por intervalos compuestos (CIM).

A diferencia de metodologías como el análisis por marcadores simples o mapeo por intervalos, CIM permite controlar la variabilidad genética en regiones exteriores al intervalo considerado mediante la incorporación de cofactores (marcadores asociados significativamente al carácter evaluado) y de esta forma incrementar el poder de detección de los QTL (Zeng, et al 1994). De esta manera, el programa informático empleado (QTLCartographer) combinó la segregación de las medias ajustadas para cada variable por ambiente y la segregación de los marcadores en un intervalo considerado en búsqueda de asociación estadística. Las asociaciones significativas, se vieron reflejadas en el mapa genético como una región de máxima probabilidad, QTL (Jensen, 1989).

Por otro lado, la imposibilidad de detectar otras posibles regiones del genoma vinculadas a la expresión de la resistencia a PHC, pudieron deberse a la existencia varios loci de efectos menores cuya detección requiera de metodologías de mapeo especiales, un aumento del

tamaño poblacional evaluado o un incremento en el número de marcadores moleculares utilizados (Beavis et al 1998)

### 2.1 IDENTIFICACIÓN DE QTLs PARA EL CARÁCTER IEMAX

Los valores de incidencia fueron registrados como el último valor registrado para esta característica o el valor más alto registrado para ella. Es así que el análisis de IEMAX por ambiente, resultó en la identificación de 9 QTLs con impacto en la característica (Figura 2.5, Tabla 2.7). Cuatro de los 9 QTLs fueron localizados con altos valores de LOD de manera consistente sobre un segmento comprendido por 10 cM en el GL 15 para las campañas 2012-13 y 2013-14 y cuando se consideró el promedio de las campañas (qIEMAX.12-13/15, qIEMAX.13-14/15, qIEMAX/15a y qIEMAX/15b). Las magnitudes de sus  $R^2$  (14,7-23,9%) determinaron que esta sección genómica sea una de la más importante para la tolerancia a IEMAX, no sólo por su reproducibilidad.

Los 5 QTLs restantes pudieron explicar valores explicar porcentajes moderados de la variación total ( $R^2 = 12,4-16,1\%$ ) pero no presentaron consistencia en distintas. La presencia de cada uno de ellos en una única campaña puede indicar que se evidencia en ambientes particulares, indicando una posible manifestación significativa de la interacción Genotipo x Ambiente.

Estudios previos sobre el ataque de *S. sclerotiorum* a capítulos realizados por Bert 2002, 2004, Yue 2008 donde identificaron QTLs para la incidencia de la PHC, informaron valores de  $R^2$  para esos QTL similares a los estimados en esta tesis.

En este contexto, todos los QTLs detectados, excepto tres (localizados en el GL 1, 3 y 7), presentaron efectos aditivos significativos de signo negativo, indicando que disminuyen la incidencia máxima, en magnitudes que van desde 0,04 hasta 0,10 (qIEMAX.11-12/14 y qIEMAX.12-13/15, respectivamente).

En todos los casos estudiados, los alelos de resistencia provinieron de la línea parental PAC2, caracterizada como medianamente resistente (MR).

En este punto cabe una consideración. Para el estudio del aporte de alelos de cada parental a los marcadores asociados al QTL, únicamente fueron considerados los marcadores locus específicos, por lo que los AFLPs (anónimos) no fueron tenidos en cuenta para este análisis.

## 2.2 IDENTIFICACIÓN DE QTLs PARA EL CARÁCTER SEV21

La severidad fue caracterizada en el día 21 post inoculación (DPI 21). Nueve regiones genómicas fueron asociadas con el carácter SEV21 (Figura 2.6, Tabla 2.8). Si bien los QTLs no se caracterizaron por ser consistentes a lo largo de los diferentes ambientes considerados, puede observarse que los valores que determinan el porcentaje de la varianza explicada fueron considerablemente superiores en 3 QTLs ubicados en el GL 1 ((qSEV21.09-10/1, qSEV21.11-12/1a, qSEV21.11-12/1b) y un QTL en el GL 10 (qSEV21.11-12/10). En todos los casos los valores de  $R^2$  superaron el 25%.

Esto podría indicar que, aunque su expresión está circunscripta a un ambiente determinado, la magnitud de sus efectos génicos lo tornan QTLs importantes respecto de la tolerancia.

Lo último también se demuestra por la dimensión de los valores de sus efectos aditivos que disminuirían significativamente (en unidades de la variable) entre 0,16 y 0,19 el grado de SEV21 para 3 de ellos (qSEV21.11-12/1a, qSEV21.11-12/1b) y qSEV21.11-12/10).

Los QTLs restantes mostraron valores de  $R^2$  y aditividad menores (incluso positivos para el efecto aditivo).

Estudios previos realizados Yue 2008 donde identificaron QTLs para la severidad de la PHC, informaron valores altos de  $R^2$ , coincidentes a los estimados en esta tesis e inferiores a los detectados por Ronicke 2005.

En los casos estudiados, los valores significativos de resistencia provinieron del parental MR (PAC2), excepto para el caso de qSEV21.09-10/16.

## 2.3 IDENTIFICACIÓN DE QTLs PARA EL CARÁCTER INT21

La intensidad, al igual que la característica anterior, fue registrada al DPI 21. Un total de 12 QTLs fueron detectados para la variable INT21 de manera no consistente (Figura 2.7, Tabla 2.9) entre campañas. Si bien sobre el GL 1 puede observarse una región genómica donde aparentemente se solapan cuatro QTLs (obtenidos de la campaña 2011-12 y 2013-14), solamente dos de ellos fueron considerados como verdaderos (qINT21.11-12/1a y qINT21.11-12/1b) y se encontraron a una distancia menor de 10 cM.

Esto cercanía relativa entre QTLs fue vista nuevamente para la campaña 2009-10 en dos ocasiones, donde los QTLs se encontraban cercanos a 13 y 22 cM aproximadamente.

En cuanto a los valores de  $R^2$ , los mismos oscilaron entre 15 y 20% para todos los QTLs encontrados, excepto para dos (qINT21.09-10/12 y qINT21.09-10/15b), que explicaron 128

varianzas superiores al 23% y sus efectos fueron aditivos fueron positivos. En este caso, el aporte de alelos significativos de resistencia a INT21 (en la campaña 2009-10 solamente), proviene del parental susceptible RHA266 (S), como puede verse para el marcador ORS401 (Tabla 2.11).

Para todos los demás QTL, los efectos aditivos fueron relativamente bajos y mayormente negativos (disminuyen los valores de INT21 en las unidades expresadas). En cuanto al aporte de alelos significativos de resistencia al marcador asociado, estos fueron realizados por PAC2.

#### 2.4 IDENTIFICACIÓN DE QTLS PARA EL CARÁCTER PI

El período de Incidencia fue registrado para el lapso de tiempo en detectar síntomas visibles de la PHC, desde el momento de inoculación. Para esta característica se detectaron 6 QTLs (Figura 2.8 y Tabla 2.10). Una única región genómica parece ser consistente al detectarse en la campaña 2012-13 y cuando se consideró el promedio de todas las campañas (qPI.12-13/10 y qPI/10, respectivamente). Esta se encuentra en el LG 10, entre los 130 y 134 cM aproximadamente. Sus valores de  $R^2$  son moderados (16-20%) y sus efectos aditivos aportarían un aumento de entre 0,85 y 1,07 unidades de la variable, repercutiendo favorablemente en retrasar el tiempo de aparición de síntomas.

Los QTLs qPI.11-12/8 y qPI.13-14/10, son regiones genómica cuya varianza explicada y efecto aditivo se destacan por sobre el resto (21,2%+1,77 para el primero y 19% y 1,17 para el segundo, respectivamente). Esto indicaría la presencia de un QTL que únicamente se expresó en las condiciones ambientales de la campaña 2011-12 o 2013-14.

Para los demás QTLs, todos se vieron con valores relativos más bajos tanto de  $R^2$  como aditividad, sin exceder a los valores 16,6% y 1,07 para el primero y el segundo, respectivamente.

Valores similares de  $R^2$  a los presentados para PI fueron hallados en estudios realizados por Bert 2002 y 2004.

Por último, todos los alelos de resistencia fueron significativamente aportados por PAC2 en los estudios realizados.

#### 2.5 CO-LOCALIZACIÓN DE QTLS PARA LAS VARIABLES ESTUDIADAS

Un estudio de comparación fue llevado a cabo para evaluar regiones genómicas implicadas en todas las variables analizadas, donde co-localizan QTLs encontrados para las diferentes características. De esta forma, al superponer las Figuras 2.5, 2.6, 2.7 y 2.8 y Tablas 2.7, 2.8, 2.9

y 2.10 (no mostrado), se obtienen las regiones candidatas implicadas en la asociación con más de un carácter.

Esta comparación permitió encontrar dos regiones del genoma que podrían verse involucradas en la determinación de resistencia para diferentes variables dentro de un mismo ensayo.

Debe ser tomada en cuenta una consideración. Las longitudes de las zonas halladas no superan 10 cM. La misma puede contener cientos de miles de pares de base de longitud, pero esta aproximación es válida para realizar un estudio de los marcadores subyacentes y determinar posibles genes candidatos cuya expresión determine la resistencia en más de una característica evaluada por ambiente.

La primera región localizada, comprende a las características IEMAX, SEV21 e INT21 de la campaña 2011-12. Esta se ubicó en el GL 1, entre los 96,7 cM y 106,7 cM.

Por otra parte, sobre el GL 10, se encontró una región comprendida entre los 121,8 cM y 124,4 cM que es consistente para el promedio de las campañas en las características SEV21 e INT21.

#### 2.6 *CO-LOCALIZACIÓN DE QTLs POR COMPARACIÓN CON TRABAJOS PREVIOS EN GIRASOL*

La presencia de QTLs estables y no consistente se encuentra con frecuencia en muchos trabajos sobre resistencia a PHC, donde se informa sobre algunos QTLs de efecto mayor que aparecen en más de un ambiente y otros QTL que se encuentran solo en ambientes específicos o bajo determinadas condiciones de cultivo (Rönicke et al. 2005; Bert et al. 2004; Bert et al. 2002; Yu et al. 2008; Mestries et al. 1998; Gentzbittel et al. 1998; Micic et al. 2004; Micic, Hahn, Bauer, Melchinger, et al. 2005; Micic, Hahn, Bauer, Schön, et al. 2005).

En esta Tesis, los QTLs detectados para todas las variables en diferentes ambientes, fueron localizados en 36 regiones genómicas ampliamente distribuidas a lo largo del genoma, excepto en los GL 4, 5, 6 y 17. Esto restringe la comparación a los 13 GLs restantes (Figura 2.9).

La comparación entre los marcadores asociados a estos QTLs con los resultados previos publicados resulta difícil. Esto se debe a que el mapa de comparación puede contener diferentes conjuntos de marcadores polimórficos, que en la población biparental aquí analizada no se encuentran. Sin embargo, a través de marcadores moleculares comunes a ambos estudios, los QTLs detectados pueden ser asociados de manera indirecta haciendo uso de la herramienta Cmap.

De esta manera, el presente trabajo de Tesis se comparó con estudios previos de QTLs realizados en girasol para la detección de QTLs de resistencia a diferentes enfermedades, en búsqueda de *c* entre las regiones genómicas identificadas en distintos cruzamientos.

Este análisis identifica a la región inferior del GL 10 como portadora de resistencia a PHC en tres de cuatro poblaciones estudiadas de manera independiente. Ronicke et al (2005) y Yu et al (2008) identificaron QTLs de resistencia en el GL 10 cerca del marcador ORS613 (ubicado a 245 cM en el mapa de PAC2 x RHA266). Maringolo (2007) describe un QTL asociado a los marcadores HA3204 y HA479 ubicados en la región distal y en esta Tesis, donde se definió un QTL de resistencia para SEV21 de la campaña 2011-12 en la misma región cromosómica localizada a 270 cM del mapa PAC2 x RHA266 .

Por otra parte, cercano al locus que alberga el gen de ramificación (*b1*) en la región central del GL 10, Bert et al (2002), Mestries et al. (1998) y Yue et al (2008) identificaron QTLs para las características incidencia, índice de latencia (similar a PI) e índice de lesión en el capítulo (similar a SEV21). En esta Tesis, gracias al posicionamiento que brindó el marcador ORS437 (125 cM, PAC2 x RHA266), se detectaron cinco QTLs de resistencia en la zona aledaña a la putativa ubicación del gen *b1* en el mapa generado. Estos QTLs fueron detectados tanto para el ambiente que considera el promedio de las campañas como para la campaña 2012-13 y se vincularon con todas las características evaluadas, excepto IEMAX.

Finalmente, sobre el GL 10 también, fueron encontrados numerosos QTLs vinculados con la resistencia a *Phoma helianthi*. El marcador HA1108 (19 cM, PAC2 x RHA266) permitió vincular un QTL detectado (qPI.13-14/10) con un QTL referido a resistencia a necrosis del tallo producida por este patógeno (Bert et al 2002).

En el GL 15, en la región comprendida por los QTLs más importantes para IEMAX identificados en este trabajo en cuanto a las magnitudes de sus efectos se refiere (qIEMAX.12-13/15, qIEMAX.13-14/15, qIEMAX/15a y qIEMAX/15b en la región que va de 77 a 92 cM), Micic 2004 encontraron QTLs importantes formando un clúster que determinan la resistencia para la tasa de crecimiento del patógeno y el tamaño de la lesión tanto en hoja como en tallo. El hecho de estos QTLs se encuentren asociados a marcadores SNPs posicionados en esta Tesis, tornó más interesante el estudio de las bases genéticas implicadas en la determinación de la resistencia en esta porción del genoma.

Más alejado del clúster mencionado, se encontró consistencia entre el QTL qINT21(11-12)-15b y el encontrado por Bert 2004 para resistencia a la mancha negra del tallo, ocasionada por *Phoma macdonaldii*.

Sobre el GL 1 en cambio, en torno al marcador ORS662 (107 cM, PAC2 x RHA266, donde se posicionaron numerosos QTLs en las campaña 2011-12 para IEMAX, SEV21 e INT21) se encontró una región genómica que afecta la tasa del crecimiento del hongo *Sclerotinia* en PHT (Micic et al. 2004) y la severidad (León, com. pers.).

Por su parte, en el GL 2, en medio de una amplia región delimitada por los marcadores ORS925 (60 cM) y ORS229 (125 cM), se posicionó un QTL referido a severidad, detectada cuando fue considerado el promedio de las campañas. Aquí también fueron encontrados dos QTLs referidos a incidencia y severidad por Yue 2008.

Distante de estos, pero también sobre el GL 2, se encontraron regiones genómicas vinculadas con la resistencia a la expansión de la lesión en el tallo (Micic et al 2004) producida por *Sclerotinia*.

Sobre el GL11, consistentemente con lo descubierto por Micic et 2004, se asoció un QTL referido a la tasa de crecimiento de *Sclerotinia* con el propio detectado INT21 en la campaña 2009-10.

Un único QTL (qSEV21.09-10/16) detectado en el GL 16 fue co-localizado mediante referencia cruzada con un QTL descrito por (Micic et al. 2004) que determina la resistencia para la lesión provocada por *Sclerotinia* al tallo de la planta.

Para el GL8, un QTL detectado para la intensidad en la campaña 2013-14 co-localizó con regiones genómicas detectadas para la resistencia a *Phoma helianthi* (Bert 2002).

Otros QTLs fueron descritos previamente sobre los GL 1 (León, com. pers, Mestries et al 1998, Ronicke et al 2005 y Micic et al 2004), GL 3 (Bert et al 2004, yue et al 2008), GL 4 (Mestries et al 1998) GL 7 (Yue et al. 2008), GL8 (Mestries et al. 1998 Micic et al. 2005) GL 9 (Mestries et al. 1998, Yue et al. 2008, Micic et al. 2004), GL 10 (Micic et al. 2004), GL 12 (Yue et al 2008) y GL 13 (Micic et al 2004), pero su posición no pudo ser referida y corroborada en el presente mapa dada la falta de marcadores comunes.

Los resultados de esta Tesis junto con los antecedentes disponibles, reafirman la complejidad que existe en expresión de la resistencia y la alta sensibilidad al ambiente que tiene este carácter.

El soporte que brinda la comparación cruzada con trabajos previos, permite focalizar los esfuerzos para dirigir la disección genética de una región en particular del genoma del girasol que haya sido detectada consistentemente. Esto implicaría por ejemplo, acotar estas regiones a través de la incorporación de más marcadores moleculares, y así identificar el marcador que maximice la predicción de la variación fenotípica.

Otra alternativa dada la falta de un genoma de referencia, implicaría realizar un estudio de las funciones putativas de los marcadores asociados, labor que se acelera si estos marcadores se desarrollaron a partir de secuencias funcionales.

## 2.7 ANÁLISIS IN SILICO

En relación al análisis funcional de secuencias que contienen marcadores asociados a QTLs de resistencia, y de co-localización de estas secuencias en mapas públicos se detectaron distintas clases de productos de genes con funciones asignadas que están directamente relacionadas con respuestas a defensa a patógenos como son las quinzasas, enzimas asociadas a procesos de óxido-reducción (peroxidazas, oxidoreductasas, deshidrogenasas), acumulación de prolina, proteínas de tipo germinas y otras menos relacionados con factores de transcripción de tipo BHLH y homocysteine s-methyltransferasas (Tabla V del apéndice).

Entre las clases de productos directamente relacionados con respuesta a patógenos encontramos tres genes candidatos que fueron categorizados con funciones putativas de quinasa. Los GCs HeAn\_C\_12025.1, HeAn\_C\_11434.1 están asociados a la característica IEMAX en múltiples campañas y se localizaron en el GL 15, y el HeAn\_C\_185.1 asociado con PI en la campaña 2009-10) sobre el GL 3.

Diversos estudios han demostrado que genes que codifican para receptores que intervienen en la resistencia mediada por el reconocimiento de patrones propios de los patógenos (PRRs), frecuentemente contienen o se asocian con quinzasas no-RD (sin arginina ni aspartato) (Huard-Chauveau et al. 2013). Estos genes a los que llamamos QDR (o implicados en la resistencia a caracteres cuantitativos), están comúnmente involucrados en el reconocimiento de un rango amplio de patógenos necrotróficos. Por ejemplo, Zipfel, et al. (2004) trabajó con el gen de *Arabidopsis FLS2*, un receptor quinasa-LRR (RLK) que participan en la percepción de la flagelina patogénica. Estos autores han establecido una base mecánica de reconocimiento de flagelina y la resistencia basal en las plantas. Las mutaciones en *FLS2* parecen producir un fenotipo que en condiciones naturales modifica levemente (cuantitativamente) la severidad

de la enfermedad y la colonización bacteriana. Diferentes alelos FLS2 se han encontrado para genes ortólogos en varias especies de Brassicáceas (Dunning et al. 2007).

Otros siete GCs localizados en los GLs 1 (HeAn\_C\_10493.1, HeAn\_C\_1638.), 3 (HeAn\_R\_283.1), 7 (HeAn\_C\_11835.1), 15 (HeAn\_C\_12025.1, HeAn\_R\_534.1) y 16 (HeAn\_C\_4316.1) se asociaron a diversos QTLs expresados a lo largo de todas las campañas, excepto la campaña 2013-14. Estos GC se vieron involucrados en vías de respuesta a estrés oxidativo, con una función descrita de oxidoreductasa o deshidrogenasa. Esto último indicaría, tanto la activación de la vía de señalización del estrés en respuesta a *S. sclerotiorum*, como la activación de las enzimas involucradas en la degradación de H<sub>2</sub>O<sub>2</sub> para la detoxificación de las células. Esta hipótesis es sustentada por los experimentos de perfiles metabólicos y por la mayor actividad catalasa observada en las flores de líneas moderadamente resistente sugiriendo que estas serían capaz de contrarrestar el efecto inhibitorio producido por el ácido oxálico sobre el estrés oxidativo (Cessna et al. 2000; Peluffo 2007).

HeAn\_C\_11352.1, ubicado sobre el GL 8, asociado con qINT21.13-14/8, no obtuvo asignación de términos GO, pero la secuencia contra la que obtuvo alineamiento significativo en BLAST, correspondió a un transportador de prolina en *A. thaliana*. Se ha descrito que la acumulación de prolina desempeña una función adaptativa en la tolerancia a diferentes estreses por parte la planta. Estreses bióticos y abióticos son capaces de inducir estrés oxidativo, estudios in vitro demostraron que la prolina podría estar envuelta en la eliminación de especies reactivas de oxígeno (ROS) (Smirnoff & Cumbes 1989). Una hipótesis propone que la prolina funciona como chaperona estabilizando la estructura de las proteínas, y que la acumulación de la misma puede proporcionar una manera estabilizar el pH citosólico y equilibrar el estado redox celular (Smirnoff & Cumbes 1989), mientras que Maggio et al (2002) sugiere que la acumulación de prolina puede ser parte de la señal de estrés involucrada en respuestas adaptativas.

El gen candidato GLP\_4, ha sido identificado como una proteína del tipo Germina y fue localizado en el GL10 próximo a un QTL referido a severidad en la campaña 2011-12. Este GC fue caracterizado en colecciones de ADNc diferenciales para respuesta al patógeno en líneas Resistentes y Susceptibles para el día 4 post inoculación (Peluffo 2007). Las proteínas de tipo germinas constituyen una familia multigénica con actividades de oxalato oxidasa, proteínas transportadoras de lípidos, inhibidores de poligalacturonasas, proteínas PR, proteasas, proteínas de unión a ADN, etc. En el caso de la GLP\_4, su función corresponde a términos GOs

referidos a la unión de iones metálicos y a estructuras externas de encapsulamiento. En girasol se aislaron y caracterizado filogenéticamente las secuencias de nueve miembros putativos de esta familia y caracterizado los niveles de expresión en distintos órganos y respuesta al patógeno, en el caso de la GLP\_4 su expresión fue ubicua (Ehrenbolger et al, 2012).

En genotipos resistentes de varias especies afectadas por *S. sclerotiorum* se ha detectado actividad oxalato oxidasa y/o superóxido dismutasa de proteínas germinas o tipos germinas (GLPs) (Carter & Thornburg 2000; Dumas et al. 1993; X. Hu et al. 2003).

Para los genes relacionados indirectamente con la resistencia a patógeno, se encontraron HeAn\_C\_4498.1 y HeAn\_C\_12824.1, ambos ubicados en el GL 10 asociados a PI de la campaña 2013-14. El primero fue descrito como un factor de transcripción con dominio BHLH y el segundo como una enzima con actividad homocisteína s-metiltransferasa (HMT). Ambos GCs han sido recientemente asociados a respuestas de defensa en el caso de BHLH ha sido descrito como un regulador de la respuesta inmune de las plantas (Fang et al. 2014). En el caso de HMT, enzima asociada al metabolismo de la homocisteína, fue inducida en gran medida en presencia del efector *SnTox3* del patógeno *Stagonospora nodorum*, al mismo tiempo que otros componentes de la respuesta de defensa como las proteínas PR y algunos metabolitos secundarios también aumentaron su nivel de expresión e inhibiendo la patogenicidad de este hongo necrotrófico, sugiriendo que la homocisteína tiene un papel importante en la en la defensa de la planta.

Por otra parte, el análisis de las secuencias correspondientes a marcadores SSRs contra la base de datos Heliagene obtuvo 11 hits (Tabla VI del apéndice). Cabe aclarar que el bajo porcentaje de cobertura de la secuencia incógnita y la alta identidad obtenida en las regiones solapantes pueden deberse a que la comparación involucra regiones genómicas y regiones transcriptas, por lo que este hecho puede estar dado por la presencia de intrones en la secuencia del SSR.

Entre las asignaciones de términos GO a las secuencias genómicas asociadas a repeticiones de tipo SSRs, se destacan las asignaciones de los marcadores HA1108 y HA3102 con actividad peroxidasa (involucrada en respuesta a estreses oxidativos ya comentada) y factor de transcripción putativo, respectivamente. Por su parte, HA4057 predijo una actividad ligasa arginina-ARNt, HA911 se vinculó con un motivo de unión a zinc y ORS437 con una transferasa de fosfatidilinositol.

Por último, para los 23 SSRs objeto de este estudio se llevó un análisis exhaustivo mediante la exploración visual de la zona aledaña a la localización de dicho marcador en los mapas previamente publicados. La Tabla VII del apéndice, resume las características encontradas. Este análisis fue eficiente para complementar la tarea de comparación entre QTLs identificados sobre la misma porción del genoma, en esta Tesis y trabajos publicados.

En términos generales, debido al tamaño de los QTLs identificados, y al posible número de genes que se encuentran incluidos en ellos (muchos de ellos con funciones desconocidas), hacer alguna sugerencia en cuanto a cuáles genes (marcadores) podrían ser los responsables de esos QTL se torna dificultoso. Por lo que el análisis *in silico* de las secuencias aquí presentado, es una mera aproximación a la determinación de sus funciones putativas o procesos en los que se pudieran ver envueltos.

# CONCLUSIONES

1. En estudios de mejoramiento molecular es importante mantener la homogeneidad de los materiales para asegurar resultados consistentes y reproducibles. En este trabajo se estableció un protocolo de validación de la homogeneidad de las líneas a partir de un ensayo multiplexado de 10 SSR que se resuelven de manera inequívoca en corridas automatizadas de fragmentos fluorescentes mediante electroforesis capilar.
2. La inclusión de marcadores SNP identificados sobre genes candidato relacionados a procesos de respuesta a estrés al mapa en estudio, reforzó la identificación de QTLs asociados a resistencia a PHC en los grupos de ligamiento 1, 3, 7, 8, 15 y 16 aportando datos semilla para profundizar el estudio del rol de esas regiones en la resistencia a PHC.
3. Los parámetros fenotípicos evaluados han mostrado segregación transgresiva y variación continua en la progenie, lo que sugiere una herencia poligénica y alta influencia de los factores ambientales sobre estos caracteres, y pone de manifiesto la necesidad de la inclusión de nuevos términos referidos en la modelización empleando modelos lineales generalizados mixtos y modelos lineales mixtos para la obtención de medias ajustadas confiables. Dada esta complejidad en el estudio, las variables IEMAX, SEV21, INT21 y PI seleccionadas, son considerarse buenas descriptoras del comportamiento de la población frente a PHC, destacando su utilidad para selección y mapeo de QTLs de resistencia. Estas variables han demostrado una moderada asociación entre sí, siendo las correlaciones que involucran a la incidencia máxima, las más bajas. La mayor correlación se da entre severidad e intensidad, vinculadas por el registro del porcentaje infectado del capítulo.
4. La heredabilidad osciló en un rango de moderado a alto, excepto para IEMAX de la campaña 2011-12 e INT21 de las campañas 2009-10 y 2011-12. Dada la diversidad genética de los padres usados para obtener la población y la influencia ambiental que enmascara los verdaderos efectos genéticos, de baja magnitud y varianza, estos valores fueron esperables.
5. Se identificaron nueve QTLs a lo largo de 5 GLs para la incidencia máxima en los ambientes evaluados. Sobre el GL15 se detectaron consistentemente 4 QTLs de efecto moderado-alto, que explicaron hasta el 24% de la varianza total. Nueve QTLs fueron descritos sobre 5 GLs para la severidad al 21° día post inoculación en todas las campañas. En los grupos 1 y 10 se detectaron 4 QTLs que tuvieron valores de  $R^2$  altos circunscriptos

únicamente a un ambiente, no obstante estos QTL no mostraron estabilidad a lo largo de los ensayos, Para la intensidad al 21° día post inoculación en todas las campañas, se encontraron un total de diez QTLs en 7 GLs con valores de  $R^2$  de entre 15 y 20%, mientras que 2 QTLs explicaron varianzas superiores al 23% cada uno y sus efectos fueron aditivos fueron positivos. Para período de incubación a lo largo de los diferentes ensayos se detectaron 6 QTLs. Una única región genómica con valores moderados de  $R^2$  resultó ser estable al detectarse en la campaña 2012-13 y se mantuvo cuando se consideró el promedio de todas. Dos QTLs se destacaron por poseer valores de  $R^2$  y aditividad superiores al resto. Dos regiones genómicas fueron identificadas por verse involucradas en la determinación de resistencia para diferentes variables dentro de un mismo ensayo sobre los GLs 1 y 10.

6. Por medio del uso de marcadores comunes entre mapas de estudios previos y de esta Tesis, se detectaron QTLs de resistencia a Sclerotinia que co-localizan sobre los GLs 1, 2, 8, 10, 11, 12, 15 y 16. Este hecho refuerza la fiabilidad de los QTLs encontrado en esta Tesis. Asimismo, otros QTLs fueron encontrados en trabajos previos de resistencia a patógenos sobre casi todos los GLs, pero su posición no pudo ser referida ni corroborada en el presente mapa dada la falta de marcadores comunes.

7. Los resultados de esta Tesis reafirman una base genética compleja subyacente a la resistencia y la alta sensibilidad al ambiente que tienen estos caracteres. En este entorno, el estudio de detección de QTLs realizado en este trabajo complementado por la información disponible de trabajos similares referidos a patógenos que afectan a girasol, permitió destacar dos regiones sobre los GLs 10 y 15 candidatas a ser analizadas en profundidad con la finalidad de direccionar investigaciones futuras hacia la identificación de genes responsables que permitan desenmarañar la complejidad de las base genéticas que gobiernan la resistencia.

# PERSPECTIVAS

Los resultados presentados en este trabajo de Tesis profundizaron en el conocimiento referido a las bases moleculares que determinan la expresión de resistencia a PHC en girasol. Por un lado, la tesis aportó regiones inéditas del genoma vinculadas a la expresión de resistencia. Si bien es necesario realizar una validación de los QTLs encontrados y lograr acotar su tamaño, los resultados obtenidos permiten anticipar la selección de nuevos genes candidatos asociados a la resistencia que pueden ser transferidos a los programas de mejoramiento para acelerar el proceso de obtención de líneas experimentales con mejor comportamiento a PHC.

Se considera que los QTL puedan ser utilizados en planes de mejoramiento a través de selección asistida por marcadores (MAS) o alternativamente, la selección recurrente asistida por marcadores (MARS).

En paralelo, el estudio combinado de la abundancia de transcritos de genes candidato asociados a la resistencia cuantitativa a la PHC, con los niveles de resistencia evaluados en la población, permitirá aportar nuevas evidencias al estudio de las regiones genómicas y los procesos biológicos asociados a la respuesta de defensa de la planta a partir de la localización genética de QTL y de QTL de expresión. Evidencia de lo mencionado es un trabajo realizado por el grupo en paralelo al desarrollo de esta Tesis, que analiza la expresión diferencial de los genes representados en el repositorio ATGC-SUR, en el que se identifican siete genes candidato localizados en la mapa PAC2 x RHA266 que presentan señales expresión diferencial estadísticamente significativa.

**Dra Norma Paniago**  
Directora

**Dr. Gerardo Cervigni**  
Director Asistente

**Dra. Carina Rivolta**  
Consejera de Estudios

**Lic. Jeremías Zubrzycki**  
Doctorando

# BIBLIOGRAFÍAS

- Abawi, G.S., 1979. Epidemiology of Diseases Caused by Sclerotinia Species. *Phytopathology*, 69(8), p.899. Available at: [http://www.apsnet.org/publications/phytopathology/backissues/Documents/1979Abstracts/Phyto69\\_899.htm](http://www.apsnet.org/publications/phytopathology/backissues/Documents/1979Abstracts/Phyto69_899.htm) [Accessed July 4, 2014].
- Abbas, A. et al., 2004. Assessment of DHPLC usefulness in the genotyping of GSTP1 exon 5 SNP: comparison to the PCR-RFLP method. *Journal of biochemical and biophysical methods*, 59(2), pp.121–126. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15163523> [Accessed June 5, 2011].
- Agrios, G.N., 1997. *Plant pathology*, San Diego: Academic Press.
- Akhunov, E., Nicolet, C. & Dvorak, J., 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 119(3), pp.507–517. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2715469&tool=pmcentrez&rendertype=abstract>.
- Al-Chaarani, G.R. et al., 2002. A QTL analysis of sunflower partial resistance to downy mildew ( *Plasmopara halstedii*) and black stem ( *Phoma macdonaldii*) by the use of recombinant inbred lines (RILs). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 104(2-3), pp.490–496. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12582723> [Accessed May 29, 2013].
- Al-Chaarani, G.R.R. et al., 2004. Genotypic variation and identification of QTLs for agronomic traits, using AFLP and SSR markers in RILs of sunflower (*Helianthus annuus* L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 109(7), pp.1353–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15365625> [Accessed May 22, 2013].
- Alfadil, T.A. et al., 2007. QTL mapping of partial resistance to *Phoma* basal stem and root necrosis in sunflower (*Helianthus annuus* L.). *Plant Science*, 172(4), pp.815–823. Available at: <http://www.sciencedirect.com/science/article/pii/S0168945206003645> [Accessed July 5, 2014].
- Alignan, M. et al., 2006. A cDNA microarray approach to decipher sunflower (*Helianthus annuus*) responses to the necrotrophic fungus *Phoma macdonaldii*. *The New phytologist*, 170(3), pp.523–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16626474> [Accessed July 4, 2014].
- Álvarez, D., 2002. Germoplasma de Girasol: Un Recurso Irreemplazable. *Girasol*, pp.148–151.
- Areshchenkova, T. & Ganai, M.W., 1999. Long tomato microsatellites are predominantly associated with centromeric regions. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 42(3), pp.536–44. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10382301> [Accessed July 7, 2014].
- Arnaud, D. et al., 2007. Genome-wide analysis of LIM gene family in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*. *DNA research*: an international journal for rapid publication of reports on genes and genomes, 14(3), pp.103–16. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2779900&tool=pmcentrez&rendertype=abstract> [Accessed June 17, 2014].
- Ashkenazi, V. et al., 2001. Development of microsatellite markers in potato and their use in phylogenetic and fingerprinting analyses. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 44(1), pp.50–62. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11269356> [Accessed November 30, 2014].
- Austin, D.F. & Lee, M., 1996. Comparative mapping in F2:3 and F 6:7 generations of quantitative trait loci for grain yield and yield components in maize. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 92(7), pp.817–26. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24166546> [Accessed July 5, 2014].

- Baack, E.J., Whitney, K.D. & Rieseberg, L.H., 2005. Hybridization and genome size evolution: timing and magnitude of nuclear DNA content increases in *Helianthus* homoploid hybrid species. *The New Phytologist*, 167(2), pp.623–630. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2442926&tool=pmcentrez&rendertype=abstract>.
- Bachlava, E. et al., 2012. SNP discovery and development of a high-density genotyping array for sunflower. *PloS one*, 7(1), p.e29814. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3251610&tool=pmcentrez&rendertype=abstract> [Accessed April 23, 2013].
- Bagge, M. & Lübberstedt, T., 2008. Functional markers in wheat: technical and economic aspects. *Molecular Breeding*, 22(3), pp.319–328. Available at: <http://link.springer.com/10.1007/s11032-008-9190-6> [Accessed May 6, 2011].
- Baird, N.A. et al., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. J. C. Fay, ed. *PloS one*, 3(10), p.e3376. Available at: <http://dx.plos.org/10.1371/journal.pone.0003376> [Accessed July 9, 2014].
- Barker, M.S. et al., 2008. Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Molecular biology and evolution*, 25(11), pp.2445–55. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2727391&tool=pmcentrez&rendertype=abstract> [Accessed November 13, 2014].
- Barreda, V.D. et al., 2010. Eocene Patagonia fossils of the daisy family. *Science (New York, N.Y.)*, 329(5999), p.1621.
- Barreiro, L.B., Henriques, R. & Mhlanga, M.M., 2009. High-throughput SNP genotyping: combining tag SNPs and molecular beacons. *Methods in molecular biology (Clifton, N.J.)*, 578, pp.255–76. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19768600> [Accessed July 5, 2014].
- Bateman, D.F. & Beer, S. V., 1965. Simultaneous production and synergistic action of oxalic acid and polygalacturonase during pathogenesis by *Sclerotinia rolfsii*. *Phytopathology*, 55, pp.204–211.
- Becelaere, G. Van & Miller, J.F., 2004. Combining Ability for Resistance to *Sclerotinia* Head Rot in Sunflower. *Agronomie*, 89, pp.1542–1545.
- Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B Methodological*, 57(1), pp.289–300. Available at: <http://www.jstor.org/stable/2346101>.
- Benjamini, Y. & Yekutieli, D., 2005. Quantitative trait Loci analysis using the false discovery rate. *Genetics*, 171(2), pp.783–90. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1456787&tool=pmcentrez&rendertype=abstract> [Accessed May 27, 2014].
- Bennett, M.D., 1995. The development and use of genomic in situ hybridization (GISH) as a new tool in plant biosystematics. In *Bransham PE, Bennett MD (eds) Kew Chromosome Conference IV, Royal Botanical Gardens, Kew, UK*. pp. 167–183.
- Bergelson, J. & Roux, F., 2010. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature reviews. Genetics*, 11(12), pp.867–79. Available at: <http://dx.doi.org/10.1038/nrg2896> [Accessed September 24, 2014].
- Berrios, F. et al., 2000. AFLP mapping of QTLs for in vitro organogenesis traits using recombinant inbred lines in sunflower (*Helianthus annuus* L.). *TAG Theoretical and Applied Genetics*, 33(8), pp.1299–1306. Available at: <http://link.springer.com/10.1007/s001220051610>.
- Berry et al., 1997. Presentation of the Advanta sunflower RFLP linkage map for public research. *Proc. 19th Sunflower Res. Workshop, Fargo, ND. 9–10 Jan. 1997*, pp.113–118.
- Berry, S.T. et al., 1996. Construction of a high density, composite RFLP linkage map for cultivated sunflower *Helianthus annuus*. *Proceedings of the 14th International Sunflower Conference - Beijing, China. 12–20 June 1996*, 2, pp.1150–1160.

- Bert, P.F. et al., 2002. Comparative genetic analysis of quantitative traits in sunflower ( *Helianthus annuus* L.) 1. QTL involved in resistance to *Sclerotinia sclerotiorum* and *Diaporthe helianthi*. *Tag Theoretical And Applied Genetics Theoretische Und Angewandte Genetik*, 105(6-7), pp.985–993. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12582925>.
- Bert, P.F. et al., 2004. Comparative genetic analysis of quantitative traits in sunflower (*Helianthus annuus* L.) 3. Characterisation of QTL involved in resistance to *Sclerotinia sclerotiorum* and *Phoma macdonaldi*. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 109(4), pp.865–74. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15141292> [Accessed May 21, 2013].
- Bioley, J.P., Jay, M. & Fiasson, J.L., 1987. Biochemical mechanisms of the secondary mechanism in the reaction of cultivated plants to fungi pathogens. In *Sclerotinia sclerotiorum in sunflower - Study of a biochemical way for looking for resistance*. *Inf. Tech. CETIOM*, 101, str., pp. 7–19.
- Bishop, J.G., 2005. Directed mutagenesis confirms the functional importance of positively selected sites in polygalacturonase inhibitor protein. *Molecular biology and evolution*, 22(7), pp.1531–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15829620> [Accessed June 1, 2014].
- Boland, G.J. & Hall, R., 1994. Index of plant hosts of *Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 16(2), pp.93–108. Available at: <http://www.tandfonline.com/doi/abs/10.1080/07060669409500766> [Accessed May 27, 2014].
- Bolton, M.D., Thomma, B.P.H.J. & Nelson, B.D., 2006. *Sclerotinia sclerotiorum* (Lib.) de Bary: biology and molecular traits of a cosmopolitan pathogen. *Molecular plant pathology*, 7(1), pp.1–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20507424> [Accessed June 18, 2014].
- Borevitz, J.O. et al., 2003. Large-scale identification of single-feature polymorphisms in complex genomes. *Genome research*, 13(3), pp.513–23. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=430246&tool=pmcentrez&rendertype=abstract> [Accessed November 13, 2014].
- Boualaphanha, C. et al., 2011. Use of New Generation Single Nucleotide Polymorphism Genotyping for Rapid Development of Near-Isogenic Lines in Rice. *Crop Science*, 51(5), p.2067. Available at: <https://www.crops.org/publications/cs/abstracts/51/5/2067> [Accessed May 28, 2014].
- Bouzidi, M.F. et al., 2006. A sunflower BAC library suitable for PCR screening and physical mapping of targeted genomic regions. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 113(1), pp.81–89. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16783592>.
- Bowers, J.E. et al., 2012. Development of an Ultra-Dense Genetic Map of the Sunflower Genome Based on Single-Feature Polymorphisms S.-B. Wu, ed. *PLoS ONE*, 7(12), p.e51360. Available at: <http://dx.plos.org/10.1371/journal.pone.0051360> [Accessed April 22, 2013].
- Burke, J.J.M. et al., 2004. Comparative mapping and rapid karyotypic evolution in the genus *helianthus*. *Genetics*, 167(1), pp.449–457. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1470840&tool=pmcentrez&rendertype=abstract> [Accessed May 13, 2014].
- Burke, J.M., Knapp, S.J. & Rieseberg, L.H., 2005. Genetic consequences of selection during the evolution of cultivated sunflower. *Genetics*, 171(4), pp.1933–1940. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1456116&tool=pmcentrez&rendertype=abstract>.
- Burr, B. & Burr, F.A., 1991. Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. *Trends in genetics*: *TIG*, 7(2), pp.55–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2035192> [Accessed July 5, 2014].
- Byrne, P.F. et al., 1998. Maize Silk Maysin Concentration and Corn Earworm Antibiosis: QTLs and Genetic Mechanisms. *Crop Science*, 38(2), p.461. Available at:

- [http://www.researchgate.net/publication/250117881\\_Maize\\_Silk\\_Maysin\\_Concentration\\_and\\_Corn\\_Earworm\\_Antibiosis\\_QTLs\\_and\\_Genetic\\_Mechanisms](http://www.researchgate.net/publication/250117881_Maize_Silk_Maysin_Concentration_and_Corn_Earworm_Antibiosis_QTLs_and_Genetic_Mechanisms) [Accessed December 1, 2014].
- Carter, C. & Thornburg, R.W., 2000. Tobacco nectarin I. Purification and characterization as a germin-like, manganese superoxide dismutase implicated in the defense of floral reproductive tissues. *The Journal of biological chemistry*, 275(47), pp.36726–33. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10952990> [Accessed July 4, 2014].
- Castañó, F. et al., 1993. Resistance of sunflower inbred lines to various forms of attack by *Sclerotinia sclerotiorum* and relations with some morphological characters. *Euphytica*, 68(1-2), pp.85–98. Available at: <http://link.springer.com/10.1007/BF00024158> [Accessed May 23, 2013].
- Castañó, F., Vear, F. & Labrouhe, D.T. De, 2001. The genetics of resistance in sunflower capitula to *Sclerotinia sclerotiorum* measured by mycelium infections combined with ascospore tests. , pp.373–380.
- Cavallini, A. et al., 2010. Analysis of transposons and repeat composition of the sunflower (*Helianthus annuus* L.) genome. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 120(3), pp.491–508. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19826774> [Accessed November 26, 2014].
- Cavallini, A. et al., 1986. Nuclear DNA changes within *Helianthus annuus* L.: cytophotometric, karyological and biochemical analyses. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 73(1), pp.20–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24240742> [Accessed November 27, 2014].
- Cervera, M.T. et al., 2001. Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics*, 158(2), pp.787–809. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1461694&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- Cessna, S.G. et al., 2000. Oxalic acid, a pathogenicity factor for *Sclerotinia sclerotiorum*, suppresses the oxidative burst of the host plant. *The Plant cell*, 12(11), pp.2191–200. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=150167&tool=pmcentrez&rendertype=abstract> [Accessed December 1, 2014].
- Chancerel, E. et al., 2011. Development and implementation of a highly-multiplexed SNP array for genetic mapping in maritime pine and comparative mapping with loblolly pine. *BMC genomics*, 12(1), p.368. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3146957&tool=pmcentrez&rendertype=abstract> [Accessed March 21, 2014].
- Chen, H. et al., 2011. Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 123(6), pp.869–79. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21681488> [Accessed June 12, 2014].
- Chen, X. et al., 2010. An eQTL analysis of partial resistance to *Puccinia hordei* in barley. *PloS one*, 5(1), p.e8598. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2798965&tool=pmcentrez&rendertype=abstract> [Accessed July 26, 2010].
- Ching, A. et al., 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC genetics*, 3, p.19. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=130040&tool=pmcentrez&rendertype=abstract>.
- Cho, R.J. et al., 1999. Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature genetics*, 23(2), pp.203–7. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2323813&tool=pmcentrez&rendertype=abstract>.

- Choi, I.-Y. et al., 2007. A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics*, 176(1), pp.685–96. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1893076&tool=pmcentrez&rendertype=abstract> [Accessed June 23, 2014].
- Church, S.A. et al., 2007. Using variable rate models to identify genes under selection in sequence pairs: their validity and limitations for EST sequences. *Journal of molecular evolution*, 64(2), pp.171–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17200807> [Accessed November 30, 2014].
- Churchill, G.A. & Doerge, R.W., 1994. Empirical threshold values for quantitative trait mapping. *Genetics*, 138(3), pp.963–971. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1206241&tool=pmcentrez&rendertype=abstract>.
- Clarkson, J.P. et al., 2003. Ascospore release and survival in *Sclerotinia sclerotiorum*. *Mycological Research*, 107(2), pp.213–222. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0953756208611673> [Accessed July 3, 2014].
- Close, T.J. et al., 2009. Development and implementation of high-throughput SNP genotyping in barley. *BMC genomics*, 10, p.582. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2797026&tool=pmcentrez&rendertype=abstract> [Accessed March 27, 2014].
- Collard, B.C.Y. et al., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142(1-2), pp.169–196. Available at: <http://link.springer.com/10.1007/s10681-005-1681-5>.
- Comai, L. et al., 2004. Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *The Plant journal: for cell and molecular biology*, 37(5), pp.778–86. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14871304> [Accessed July 5, 2014].
- Cook, D.E. et al., 2012. Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science (New York, N.Y.)*, 338(6111), pp.1206–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23065905> [Accessed September 10, 2014].
- Cordeiro, G.M. et al., 2006. Characterisation of single nucleotide polymorphisms in sugarcane ESTs. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 113(2), pp.331–343. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16791699> [Accessed March 5, 2012].
- Costabile, M., Quach, A. & Ferrante, A., 2006. Molecular approaches in the diagnosis of primary immunodeficiency diseases. *Human mutation*, 27(12), pp.1163–73. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16960849> [Accessed November 30, 2014].
- Cregan, P.B. et al., 1999. *An Integrated Genetic Linkage Map of the Soybean Genome*, Available at: <http://digitalcommons.unl.edu/agronomyfacpub/20> [Accessed July 8, 2014].
- Crites, 1993. Domesticated Sunflower in Fifth Millennium B.P. Temporal Context: New Evidence from Middle Tennessee. *American Antiquity*, 58, pp.146–148.
- Cuadrado, A. & Schwarzacher, T., 1998. The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma*, 107(8), pp.587–94. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9933412> [Accessed July 7, 2014].
- Dalloul, R.A. et al., 2010. Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. *PLoS biology*, 8(9). Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2935454&tool=pmcentrez&rendertype=abstract>.
- Das, S. et al., 2008. Detection and validation of single feature polymorphisms in cowpea (*Vigna unguiculata* L. Walp) using a soybean genome array. *BMC genomics*, 9, p.107. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2270837&tool=pmcentrez&rendertype=abstract> [Accessed November 25, 2014].

- DAVAR, R. & Darvishzadeh, R., 2010. QTL mapping of partial resistance to basal stem rot in sunflower using recombinant inbred lines. *Phytopathologia ...*, pp.330–341. Available at: <http://www.amati.fupress.net/index.php/pm/article/view/8374> [Accessed May 21, 2013].
- Dean, R.E. et al., 1999. Genetic Redundancy and Diversity among “Orange” Accessions in the U.S. National Sorghum Collection as Assessed with Simple Sequence Repeat (SSR) Markers. *Crop Science*, 39(4), p.1215. Available at: <https://www.crops.org/publications/cs/abstracts/39/4/CS0390041215?access=0&view=pdf> [Accessed November 30, 2014].
- Debeaujon, I. & Koornneef, M., 2000. Gibberellin requirement for Arabidopsis seed germination is determined both by testa characteristics and embryonic abscisic acid. *Plant physiology*, 122(2), pp.415–24. Available at: [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58878&tool=pmcentrez&render\\_type=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58878&tool=pmcentrez&render_type=abstract) [Accessed November 27, 2014].
- Degener, J., Melchinger, A.E. & Hahn, V., 1999. Optimal allocation of resources in evaluating current sunflower inbred lines for resistance to Sclerotinia. *Plant Breeding*, 118(2), pp.157–160. Available at: <http://doi.wiley.com/10.1046/j.1439-0523.1999.118002157.x> [Accessed July 4, 2014].
- Dehmer, K.J. & Friedt, W., 1998. Evaluation of different microsatellite motifs for analysing genetic relationships in cultivated sunflower (*Helianthus annuus* L.). *Plant Breeding*, 117(1), pp.45–48. Available at: <http://doi.wiley.com/10.1111/j.1439-0523.1998.tb01446.x>.
- Denby, K.J., Kumar, P. & Kliebenstein, D.J., 2004. Identification of Botrytis cinerea susceptibility loci in Arabidopsis thaliana. *The Plant journal: for cell and molecular biology*, 38(3), pp.473–86. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15086796> [Accessed November 24, 2014].
- Deulvot, C. et al., 2010. Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. *BMC genomics*, 11, p.468. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3091664&tool=pmcentrez&rendertype=abstract>.
- Dezar, C. a et al., 2003. Identification of three MADS-box genes expressed in sunflower capitulum. *Journal of experimental botany*, 54(387), pp.1637–1639. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12730268> [Accessed June 5, 2011].
- Díaz-Zorita, M., Duarte, G. & Plante, E., 2003. El Cultivo de Girasol. In *ASAGIR*.
- Doerge, R.W. & Churchill, G.A., 1996. Permutation Tests for Multiple Loci Affecting a Quantitative Character. *Genetics*, 142(1), pp.285–294. Available at: <http://www.genetics.org/content/142/1/285.abstract> [Accessed December 1, 2014].
- Dongre, A. & Parkhi, V., 2012. Identification of Cotton Hybrid through the Combination of PCR Based RAPD, ISSR and Microsatellite Markers. *Journal of Plant Biochemistry and Biotechnology*, 14(1), pp.53–55. Available at: <http://www.springerlink.com/index/10.1007/BF03263226> [Accessed November 30, 2014].
- Dorrell, D.G. & Vick, B.A., 1997. Properties and processing of oilseed sunflower. In *Schneiter, A.A. (Eds), Sunflower technology and production American Society of Agronomy, Madison, Wisconsin*. pp. 709–744.
- Duca, M. et al., 2014. Microsatellite Marker Application in Sunflower ( *Helianthus Annuus* L.) Fingerprinting. *Biotechnology & Biotechnological Equipment*, 27(3), pp.3772–3775. Available at: <http://www.tandfonline.com/doi/abs/10.5504/BBEQ.2013.0021> [Accessed November 25, 2014].
- Dumas, B. et al., 1993. Identification of barley oxalate oxidase as a germin-like protein. *Comptes rendus de l'Académie des sciences. Série III, Sciences de la vie*, 316(8), pp.793–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8044704> [Accessed July 4, 2014].

- Dunning, F.M. et al., 2007. Identification and mutational analysis of Arabidopsis FLS2 leucine-rich repeat domain residues that contribute to flagellin perception. *The Plant cell*, 19(10), pp.3297–313. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2174712&tool=pmcentrez&rendertype=abstract> [Accessed November 21, 2014].
- Ebrahimi, A. et al., 2008. QTL mapping of seed-quality traits in sunflower recombinant inbred lines under different water regimes. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 51(8), pp.599–615. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18650950> [Accessed May 29, 2013].
- Elnifro, E.M. et al., 2000. Multiplex PCR: optimization and application in diagnostic virology. *Clinical microbiology reviews*, 13(4), pp.559–70. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=88949&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- Elshire, R.J. et al., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. L. Orban, ed. *PloS one*, 6(5), p.e19379. Available at: <http://dx.plos.org/10.1371/journal.pone.0019379> [Accessed July 9, 2014].
- Escande, A.R., Laich, F.S. & Pedraza, M. V., 2002. Field testing of honeybee-dispersed *Trichoderma* spp. to manage sunflower head rot (*Sclerotinia sclerotiorum*). *Plant Pathology*, 51(3), pp.346–351.
- Falconer & Mackay, 1996. *Quantitative genetics*. Longman.,
- Fan et al., 2006. Highly parallel genomic assays. *Nature Reviews Genetics*, 7(8), pp.632–644. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16847463> [Accessed June 5, 2014].
- FAN, J.-B. et al., 2003. Highly Parallel SNP Genotyping. *Cold Spring Harbor Symposia on Quantitative Biology*, 68(0), pp.69–78. Available at: <http://symposium.cshlp.org/content/68/69.short> [Accessed July 5, 2014].
- Faria, R. & Navarro, A., 2010. Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends in ecology & evolution*, 25(11), pp.660–9. Available at: <http://www.sciencedirect.com/science/article/pii/S0169534710001795> [Accessed August 20, 2014].
- Fehr, W.R., 1987. Principles of cultivar development. Volume 2. Crop species. Available at: <http://www.cabdirect.org/abstracts/19871663199.html> [Accessed December 1, 2014].
- Feltus, F.A. et al., 2006. A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. *Plant physiology*, 140(4), pp.1183–91. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1435799&tool=pmcentrez&rendertype=abstract> [Accessed July 4, 2014].
- Feng, J. et al., 2006. Construction of BAC and BIBAC libraries from sunflower and identification of linkage group-specific clones by overgo hybridization. *Theoretical and applied ...*, 113(1), pp.23–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16612648> [Accessed July 4, 2014].
- Fernandez, P., Di Rienzo, J., et al., 2012. Development and validation of a high density sunflower microarray for functional studies on biotic and abiotic stresses. In ISA, ed. *18th International Sunflower Conference*. Mar del Plata, Argentina. Available at: [http://www.asagir.org.ar/asagir2008/buscar\\_congreso.asp?e=1](http://www.asagir.org.ar/asagir2008/buscar_congreso.asp?e=1).
- Fernandez, P., Soria, M., et al., 2012. Development, characterization and experimental validation of a cultivated sunflower (*Helianthus annuus* L.) gene expression oligonucleotide microarray. *PloS one*, 7(10), p.e45899. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3482228&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].

- Fernandez, P. et al., 2003. Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project. *BMC Genomics*, 4(1), p.40. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=270089&tool=pmcentrez&rendertype=abstract>.
- Fernandez, P. et al., 2008. Transcriptomic identification of candidate genes involved in sunflower responses to chilling and salt stresses based on cDNA microarray analysis. *BMC Plant Biology*, 8(1), p.11. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18221554>.
- Ferreira, A. et al., 2006. Estimating the effects of population size and type on the accuracy of genetic maps. *Genetics and Molecular Biology*, 29(1), pp.187–192. Available at: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1415-47572006000100033&lng=en&nrm=iso&tlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572006000100033&lng=en&nrm=iso&tlng=en) [Accessed July 7, 2014].
- Ferriol, M., Picó, B. & Nuez, F., 2003. Genetic diversity of a germplasm collection of Cucurbita pepo using SRAP and AFLP markers. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 107(2), pp.271–82. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12845442> [Accessed November 30, 2014].
- Fick, G.N. & Zimmer, D.E., 1976. No Title. In I. S. Assoc., ed. *7th Int.Sunflower Conf. Krasnadar*, Paris, France, pp. 253–258.
- Filippi, C. et al., 2014. Genetics and Genomics Applied to Sunflower Breeding. In *Sunflowers: Growth and Development, Environmental Influences and Pests/Diseases*. Nova Science, pp. 61–94.
- Funk, V.A. et al., 2005. Everywhere but Antarctica: using a supertree to understand the diversity and distribution of the Compositae. In *Biol Skr.* pp. 343–374.
- Fusari, C.M. et al., 2012. Association mapping in sunflower for sclerotinia head rot resistance. *BMC plant biology*, 12, p.93. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22708963> [Accessed July 5, 2014].
- Fusari, C.M. et al., 2008. Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. *BMC plant biology*, 8, p.7. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2266750&tool=pmcentrez&rendertype=abstract> [Accessed May 24, 2013].
- Fusari, C.M. et al., 2010. Single nucleotide polymorphism genotyping by heteroduplex analysis in sunflower (*Helianthus annuus* L.). *Molecular Breeding*, 28(1), pp.73–89. Available at: <http://www.springerlink.com/index/10.1007/s11032-010-9462-9> [Accessed June 9, 2012].
- Galella, M.T., León, A. & Bazzalo, M.E., 2010. Taller Asagir 2010. Mar del Plata, Argentina. In *Taller Asagir*. Mar del Plata, Argentina.
- Ganal, M.W., Altmann, T. & Röder, M.S., 2009. SNP identification in crop plants. *Current opinion in plant biology*, 12(2), pp.211–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19186095>.
- Garg, A. et al., 2006. Utility of a fertility restorer gene linked marker for testing genetic purity of hybrid seeds in rice (*Oryza sativa* L.). *Seed Science and Technology*, 34(1), pp.9–18. Available at: [http://www.ingentaconnect.com/content/ista/sst/2006/00000034/00000001/art00002?cr\\_awler=true](http://www.ingentaconnect.com/content/ista/sst/2006/00000034/00000001/art00002?cr_awler=true) [Accessed November 30, 2014].
- Garza, J.A.G. et al., 2002. Influence of crop rotation and tillage on production of apothecia by *Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 24(2), pp.137–143. Available at: <http://www.tandfonline.com/doi/abs/10.1080/07060660309506988> [Accessed July 4, 2014].
- Gedil, M.A. et al., 2001. Candidate disease resistance genes in sunflower cloned using conserved nucleotide-binding site motifs: genetic mapping and linkage to the downy mildew resistance gene Pl1. *Genome / National Research Council Canada = Génome / Conseil national de*

- recherches Canada*, 44(2), pp.205–12. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11341730> [Accessed June 7, 2013].
- Geldermann, H., 1975. Investigations on inheritance of quantitative characters in animals by gene markers I. Methods. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 46(7), pp.319–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24420173> [Accessed December 1, 2014].
- Gentzbittel, L. et al., 2002. A bacterial artificial chromosome (BAC) library for sunflower, and identification of clones containing genes for putative transmembrane receptors. *Molecular genetics and genomics*: *MGG*, 266(6), pp.979–87. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11862492> [Accessed July 4, 2014].
- Gentzbittel, L. et al., 1999. A composite map of expressed sequences and phenotypic traits of the sunflower (*Helianthus annuus* L.) genome. *TAG Theoretical and Applied Genetics*, 99(1-2), pp.218–234. Available at: <http://link.springer.com/10.1007/s001220051228> [Accessed November 4, 2014].
- Gentzbittel, L. et al., 1998. Cloning of molecular markers for disease resistance in sunflower, *Helianthus annuus* L. *TAG Theoretical and Applied Genetics*, 96(3-4), pp.519–525. Available at: <http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s001220050769> [Accessed May 21, 2013].
- Gentzbittel, L. et al., 1995. Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annuus* L.). *Theoretical and Applied Genetics*, 90(7-8), pp.1079–1086. Available at: <http://www.springerlink.com/index/10.1007/BF00222925>.
- Gentzbittel, L. et al., 1994. RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Helianthus annuus* L.: evidence for distinct restorer and maintainer germplasm. *Theoretical and Applied ...*, pp.419–425. Available at: <http://link.springer.com/article/10.1007/BF00225376>.
- Gethi, J.G. et al., 2002. SSR Variation in Important U.S. Maize Inbred Lines. *Crop Science*, 42(3), p.951. Available at: <https://www.crops.org/publications/cs/abstracts/42/3/951> [Accessed November 30, 2014].
- Giacomelli, J.I. et al., 2010. Expression analyses indicate the involvement of sunflower WRKY transcription factors in stress responses, and phylogenetic reconstructions reveal the existence of a novel clade in the Asteraceae. *Plant Sci*, 178, pp.398–410.
- Giancola, S. et al., 2006. Utilization of the three high-throughput SNP genotyping methods, the GOOD assay, Amplifluor and TaqMan, in diploid and polyploid plants. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 112(6), pp.1115–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16453133> [Accessed June 3, 2014].
- Gilles, P.N. et al., 1999. Single nucleotide polymorphic discrimination by an electronic dot blot assay on semiconductor microchips. *Nature biotechnology*, 17(4), pp.365–70. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10207885> [Accessed July 5, 2014].
- Giordano, M. et al., 1999. Identification by denaturing high-performance liquid chromatography of numerous polymorphisms in a candidate region for multiple sclerosis susceptibility. *Genomics*, 56(3), pp.247–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10087191> [Accessed July 4, 2014].
- Goldstein, D.B. et al., 1999. The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics*, 151(2), pp.797–801. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1460481&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- González-Pérez, S. et al., 2005. Emulsion properties of sunflower (*Helianthus annuus*) proteins. *Journal of agricultural and food chemistry*, 53(6), pp.2261–2267. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15769166>.

- Grattapaglia, D. et al., 2011. High- throughput SNP genotyping in the highly heterozygous genome of Eucalyptus: assay success, polymorphism and transferability across species. *BMC plant biology*, 11(1), p.65. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3090336&tool=pmcentrez&rendertype=abstract> [Accessed April 1, 2014].
- Guimarães, R.L. & Stotz, H.U., 2004. Oxalate production by *Sclerotinia sclerotiorum* deregulates guard cells during infection. *Plant physiology*, 136(3), pp.3703–11. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=527168&tool=pmcentrez&rendertype=abstract> [Accessed July 4, 2014].
- Gulya, T., Rashid, K.Y. & Marisevic, S.M., 1997. Sunflower technology and Production. In D. R. Berglund, ed. *Sunflower diseases*. American Society of Agronomy Monograph, pp. 263–379.
- Gulya, T.J., Marek, L.F. & GavriloVA, V., 2010. Disease resistance in cultivated sunflower derived from public germplasm collections. In *International Symposium "Sunflower Breeding on Resistance to Diseases"*. Krasnodart, Russia, p. 7.
- Gupta, P.K., Roy, J.K. & Prasad, M., 2001. Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science*, 80(4).
- Gupta, P.K., Rustgi, S. & Mir, R.R., 2008. Array-based high-throughput DNA markers for crop improvement. *Heredity*, 101(1), pp.5–18. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18461083>.
- Haldane, J.B. & Waddington, C.H., 1931. Inbreeding and Linkage. *Genetics*, 16(4), pp.357–374. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1201105&tool=pmcentrez&rendertype=abstract> [Accessed July 4, 2014].
- Hampel, M., Locher, F. & Saur, R., 1981. Control of *Sclerotinia sclerotiorum* and *Botrytis cinerea* on rape and sunflower with Vinclozolin. In *Proc. Br. Crop Prot. Conf. Pests Dis.* pp. 521–527.
- Harter, A. V et al., 2004. Origin of extant domesticated sunflowers in eastern North America. *Nature*, 430(6996), pp.201–205. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15241413>.
- He, C., Poysa, V. & Yu, K., 2003. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 106(2), pp.363–73. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12582864> [Accessed November 30, 2014].
- Heesacker, A. et al., 2008. SSRs and INDELS mined from the sunflower EST database: abundance, polymorphisms, and cross-taxa utility. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 117(7), pp.1021–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18633591> [Accessed November 30, 2014].
- Heesacker, A.F. et al., 2009. Karyotypic Evolution of the Common and Silverleaf Sunflower Genomes. *The Plant Genome Journal*, 2(3), p.233. Available at: <https://www.crops.org/publications/tpg/abstracts/2/3/233> [Accessed November 30, 2014].
- Hegedus, D.D. et al., 2008. Brassica napus possesses an expanded set of polygalacturonase inhibitor protein genes that are differentially regulated in response to *Sclerotinia sclerotiorum* infection, wounding and defense hormone treatment. *Planta*, 228(2), pp.241–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18431596> [Accessed July 4, 2014].
- Hegedus, D.D. & Rimmer, S.R., 2005. *Sclerotinia sclerotiorum*: when “to be or not to be” a pathogen? *FEMS microbiology letters*, 251(2), pp.177–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16112822>.
- Heiser et al., 1981. Infrageneric Classification of *Helianthus* (Compositae). *Taxon*, 30, No. 2, pp.pp. 393–403.

- Heiser, C.B. & Smith, D.M., 1954. New chromosome numbers in *Helianthus* and related genera (Compositae). *Proceedings of the Indiana Academy of Science*, 64, pp.250–253. Available at: <http://journals.iupui.edu/index.php/ias/article/view/6174> [Accessed July 4, 2014].
- Henegariu, O. et al., 1997. Multiplex PCR: critical parameters and step-by-step protocol. *BioTechniques*, 23(3), pp.504–11. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9298224> [Accessed November 30, 2014].
- Hervé, D. et al., 2001. QTL analysis of photosynthesis and water status traits in sunflower (*Helianthus annuus* L.) under greenhouse conditions. *Journal of experimental botany*, 52(362), pp.1857–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11520874> [Accessed July 5, 2014].
- Hewezi, T., Petitprez, M. & Gentzbittel, L., 2006. Primary metabolic pathways and signal transduction in sunflower (*Helianthus annuus* L.): comparison of transcriptional profiling in leaves and immature embryos using cDNA microarrays. *Planta*, 223(5), pp.948–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16307285> [Accessed June 9, 2014].
- Higuchi, R. et al., 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Bio/technology (Nature Publishing Company)*, 11(9), pp.1026–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7764001> [Accessed May 29, 2014].
- Holley, R.C., 1986. Effect of Plant Population and Inoculum Density on Incidence of Sclerotinia Wilt of Sunflower. *Phytopathology*, 76(1), p.71. Available at: [http://www.apsnet.org/publications/phytopathology/backissues/Documents/1986Abstracts/Phyto76\\_71.htm](http://www.apsnet.org/publications/phytopathology/backissues/Documents/1986Abstracts/Phyto76_71.htm) [Accessed July 4, 2014].
- Hu, J., Vick, B.A. & Hu, 2003. Target Region Amplification Polymorphism: A Novel Marker Technique for Plant Genotyping. , (September), pp.289–294.
- Hu, X. et al., 2003. Overexpression of a gene encoding hydrogen peroxide-generating oxalate oxidase evokes defense responses in sunflower. *Plant Physiology*, 133(1), pp.170–181. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=196595&tool=pmcentrez&rendertype=abstract> [Accessed July 4, 2014].
- Huang, H.C. & Dueck, J., 1980. Wilt of sunflower from infection by mycelial-germinating sclerotia of *Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 2(2), pp.47–52. Available at: <http://www.tandfonline.com/doi/abs/10.1080/07060668009501437> [Accessed May 23, 2013].
- Huard-Chauveau, C. et al., 2013. An atypical kinase under balancing selection confers broad-spectrum disease resistance in *Arabidopsis*. *PLoS genetics*, 9(9), p.e1003766. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3772041&tool=pmcentrez&rendertype=abstract> [Accessed November 24, 2014].
- Hwang, T.-Y. et al., 2009. High-density integrated linkage map based on SSR markers in soybean. *DNA research: an international journal for rapid publication of reports on genes and genomes*, 16(4), pp.213–25. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2725787&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- Hyten, D.L. et al., 2008. High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 116(7), pp.945–952. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18278477> [Accessed April 8, 2011].
- Ingaramo, J., 2005. Plan Estratégico ASAGIR 2006-2015. In ASAGIR, ed. *Cuadernillo Informativo, Vol. N° 11*. Buenos Aires, p. 24.
- Inglis, G.D. & Boland, G.J., 1990. The microflora of bean and rapeseed petals and the influence of the microflora of bean petals on white mold. *Canadian Journal of Plant Pathology*, 12(2), pp.129–

134. Available at: <http://www.tandfonline.com/doi/abs/10.1080/07060669009501015> [Accessed July 4, 2014].
- Jan, C.C. et al., 1998. Construction of an RFLP linkage map for cultivated sunflower. *TAG Theoretical and Applied Genetics*, 96(1), pp.15–22. Available at: <http://link.springer.com/10.1007/s001220050703>.
- Jansen, R.C. & Stam, P., 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics*, 136(4), pp.1447–55. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1205923&tool=pmcentrez&rendertype=abstract> [Accessed November 24, 2014].
- Jensen, J., 1989. Estimation of recombination parameters between a quantitative trait locus (QTL) and two marker gene loci. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 78(5), pp.613–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24225819> [Accessed December 1, 2014].
- Jermstad, K.D. et al., 2010. Comparative mapping in Pinus: sugar pine (*Pinus lambertiana* Dougl.) and loblolly pine (*Pinus taeda* L.). *Tree Genetics & Genomes*, 7(3), pp.457–468. Available at: <http://link.springer.com/10.1007/s11295-010-0347-1> [Accessed April 6, 2014].
- Jones, A.C. et al., 1999. Optimal temperature selection for mutation detection by denaturing HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis. *Clinical chemistry*, 45(8 Pt 1), pp.1133–40. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10430776> [Accessed November 30, 2014].
- Kallunki, J.A. & Heywood, V.H., 1994. Flowering Plants of the World. *Brittonia*, 46(4), p.316. Available at: <http://link.springer.com/10.2307/2806914> [Accessed July 4, 2014].
- Kane, N.C.C. et al., 2011. Progress towards a reference genome for sunflower. *Botany*, 437(7), pp.429–437. Available at: <http://www.nrcresearchpress.com/doi/abs/10.1139/b11-032> [Accessed May 23, 2013].
- Katinas, L. et al., 2007. Panorama de la familia Asteraceae (= Compositae) en la Republica Argentina. *Bol. Soc. Argent. Bot*, pp.113–129.
- Kesseli, R. V. & Michelmore, R.W., 1997. The Compositae: systematically fascinating but specifically neglected. In A. H. Paterson, ed. *Genome Mapping in Plants*. Georgetown, TX, pp. 179 – 191.
- Kiani, S. et al., 2007. Genetic analysis of plant water status and osmotic adjustment in recombinant inbred lines of sunflower under two water treatments. *Plant Science*, 172(4), pp.773–787. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0168945206003633>.
- Kiani, S.P. et al., 2007. Genetic variability for physiological traits under drought conditions and differential expression of water stress-associated genes in sunflower (*Helianthus annuus* L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 114(2), pp.193–207. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17103138> [Accessed June 6, 2013].
- Kiani, S.P. et al., 2008. QTL analysis of chlorophyll fluorescence parameters in sunflower (*Helianthus annuus* L.) under well-watered and water-stressed conditions. *Plant Science*, 175(4), pp.565–573. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0168945208001684> [Accessed June 6, 2013].
- Kiani, S.P. & Sarrafi, A., 2010. Genetics, Genomics and Breeding of Sunflower J. Hu, G. Seiler, & C. Kole, eds. Available at: <http://www.crcnetbase.com/doi/book/10.1201/b10192>.
- Köhler, H. & Friedt, W., 1999. Genetic Variability as Identified by AP-PCR and Reaction to Mid-Stem Infection of among Interspecific Sunflower ( L.) Hybrid Progenies. *Crop Science*, 39(5), p.1456. Available at: <https://www.crops.org/publications/cs/abstracts/39/5/1456> [Accessed July 4, 2014].
- Kolkman, J.M. et al., 2007. Single nucleotide polymorphisms and linkage disequilibrium in sunflower. *Genetics*, 177(1), pp.457–468. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2013689&tool=pmcentrez&rendertype=abstract>.

- KOSAMBI, D.D., 1943. THE ESTIMATION OF MAP DISTANCES FROM RECOMBINATION VALUES. *Annals of Eugenics*, 12(1), pp.172–175. Available at: <http://onlinelibrary.wiley.com/doi/10.1111/j.1469-1809.1943.tb02321.x/abstract> [Accessed November 30, 2014].
- Kota, R. et al., 2001. Application of denaturing high-performance liquid chromatography for mapping of single nucleotide polymorphisms in barley (*Hordeum vulgare* L.). *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 44(4), pp.523–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11550885> [Accessed July 5, 2014].
- Kota, R. et al., 2008. EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. *Functional & integrative genomics*, 8(3), pp.223–33. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17968603> [Accessed June 18, 2014].
- Kota, R. et al., 2003. Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). *Molecular genetics and genomics*: MGG, 270(1), pp.24–33. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12938038>.
- Krattinger, S.G. et al., 2009. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science (New York, N.Y.)*, 323(5919), pp.1360–3. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19229000> [Accessed November 17, 2014].
- Kräuter, R., Steinmetz, A. & Friedt, W., 1991. Efficient interspecific hybridization in the genus *Helianthus* via “embryo rescue” and characterization of the hybrids. *TAG Theoretical and Applied Genetics*, 82, pp.521–525.
- Kumar, R. et al., 2007. Single feature polymorphism discovery in rice. J. Borevitz, ed. *PloS one*, 2(3), p.e284. Available at: <http://dx.plos.org/10.1371/journal.pone.0000284> [Accessed November 25, 2014].
- Künzel, G., Korzun, L. & Meister, A., 2000. Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics*, 154(1), pp.397–412. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1460903&tool=pmcentrez&rendertype=abstract> [Accessed July 7, 2014].
- Kwok, P.Y., 2001. Methods for genotyping single nucleotide polymorphisms. *Annual review of genomics and human genetics*, 2, pp.235–58. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11701650> [Accessed July 4, 2014].
- Kwon, S.-J., Truco, M.-J. & Hu, J., 2011. LSGermOPA, a custom OPA of 384 EST-derived SNPs for high-throughput lettuce (*Lactuca sativa* L.) germplasm fingerprinting. *Molecular Breeding*, 29(4), pp.887–901. Available at: <http://link.springer.com/10.1007/s11032-011-9623-5> [Accessed July 8, 2014].
- Lacombe, S. et al., 2010. Interfamily transfer of a plant pattern- recognition receptor confers broad-spectrum bacterial resistance. *Nature biotechnology*, 28(4), pp.365–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20231819> [Accessed September 8, 2014].
- Lai, Z. et al., 2005. Extensive chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species. *Genetics*, 171(1), pp.291–303. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1456521&tool=pmcentrez&rendertype=abstract> [Accessed May 29, 2013].
- Lai, Z. et al., 2005. Identification and mapping of SNPs from ESTs in sunflower. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 111(8), pp.1532–44. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2442914&tool=pmcentrez&rendertype=abstract> [Accessed May 17, 2011].
- Lai, Z. et al., 2006. Microarray analysis reveals differential gene expression in hybrid sunflower species. *Molecular ecology*, 15(5), pp.1213–27. Available at:

- <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2536761&tool=pmcentrez&rendertype=abstract> [Accessed June 27, 2014].
- Lam, E., Kato, N. & Lawton, M., 2001. Programmed cell death, mitochondria and the plant hypersensitive response. *Nature*, 411(6839), pp.848–853. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11459068>.
- Lander, E. & Kruglyak, L., 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature genetics*, 11(3), pp.241–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7581446> [Accessed November 28, 2014].
- Lander, E.S. & Botstein, D., 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1), pp.185–99. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1203601&tool=pmcentrez&rendertype=abstract> [Accessed July 7, 2014].
- Langar, K. et al., 2003. Combined mapping of DALP and AFLP markers in cultivated sunflower using F9 recombinant inbred lines. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 106(6), pp.1068–1074. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12671755>.
- Lapègue, S. et al., 2014. Development of SNP-genotyping arrays in two shellfish species. *Molecular ecology resources*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24447767> [Accessed April 6, 2014].
- Lauter, N. et al., 2008. Quantitative Genetic Dissection of Shoot Architecture Traits in Maize: Towards a Functional Genomics Approach. *The Plant Genome Journal*, 1(2), p.99. Available at: <https://www.crops.org/publications/tpg/abstracts/1/2/99> [Accessed June 10, 2014].
- Lee, H.H. et al., 2004. Crystal structure of T-protein of the glycine cleavage system. Cofactor binding, insights into H- protein recognition, and molecular basis for understanding nonketotic hyperglycinemia. *The Journal of biological chemistry*, 279(48), pp.50514–23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15355973> [Accessed July 7, 2014].
- Leon, A.J.J., Andrade, F.H.F.H. & Lee, M., 2000. Genetic Mapping of Factors Affecting Quantitative Variation for Flowering in Sunflower. *Crop Science*, 40(2), pp.404–407. Available at: <https://www.crops.org/publications/cs/abstracts/40/2/404> [Accessed May 28, 2013].
- Lepoittevin, C. et al., 2010. In vitro vs in silico detected SNPs for the development of a genotyping array: what can we learn from a non-model species? *PLoS one*, 5(6), p.e11034. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2882948&tool=pmcentrez&rendertype=abstract> [Accessed March 28, 2014].
- Lexer, C. et al., 2005. Genetics of species differences in the wild annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics*, 169(4), pp.2225–2239. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1449618&tool=pmcentrez&rendertype=abstract> [Accessed May 29, 2013].
- Libert, M.A., 1837. *Plante cryptogamicae arduennae (Exsiccati) no. 326*,
- Lijavetzky, D. et al., 2007. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BMC genomics*, 8, p.424. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18021442>.
- Lin, C.H. et al., 2009. Medium- to high-throughput SNP genotyping using VeraCode microbeads. *Methods in molecular biology (Clifton, N.J.)*, 496, pp.129–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18839109> [Accessed April 1, 2012].
- Liu, A. & Burke, J.M., 2006. Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics*, 173(1), pp.321–330. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1461453&tool=pmcentrez&rendertype=abstract>.
- Liu, J.H., Lee-Tamon, S.H. & Reid, D.M., 1997. Differential and wound-inducible expression of 1-aminocyclopropane-1-carboxylate oxidase genes in sunflower seedlings. *Plant molecular*

- biology*, 34(6), pp.923–33. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9290644> [Accessed July 7, 2014].
- Liu, L.-W. et al., 2007. Assessment of genetic purity of tomato (*Lycopersicon esculentum* L.) hybrid using molecular markers. *Scientia Horticulturae*, 115(1), pp.7–12. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0304423807002701> [Accessed November 25, 2014].
- Liu, Z. et al., 2012. Molecular mapping of the Pl(16) downy mildew resistance gene from HA-R4 to facilitate marker-assisted selection in sunflower. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 125(1), pp.121–31. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22350177> [Accessed June 6, 2013].
- Livak, K.J. et al., 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR methods and applications*, 4(6), pp.357–62. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7580930> [Accessed July 5, 2014].
- Liwang, L. et al., 2004. Application of Molecular Marker in Variety Identification and Purity Testing in Vegetable Crops. *Fen zi zhi wu yu zhong*, 2(4), pp.563–568. Available at: <http://europemc.org/abstract/CBA/468927> [Accessed November 30, 2014].
- Lorang, J.M., Sweat, T.A. & Wolpert, T.J., 2007. Plant disease susceptibility conferred by a “resistance” gene. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37), pp.14861–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1976202&tool=pmcentrez&rendertype=abstract> [Accessed November 24, 2014].
- Lumsden, D.R., 1969. Sclerotinia sclerotiorum infection of bean and the production of cellulase. *Phytopathology*, 59, pp.653–657.
- Lumsden, R.D., 1973. Histopathology of Sclerotinia sclerotiorum Infection of Bean. *Phytopathology*, 63(6), p.708. Available at: [http://www.apsnet.org/publications/phytopathology/backissues/Documents/1973Abstracts/Phyto63\\_708.htm](http://www.apsnet.org/publications/phytopathology/backissues/Documents/1973Abstracts/Phyto63_708.htm) [Accessed July 4, 2014].
- Maggio, A. et al., 2002. Does proline accumulation play an active role in stress-induced growth reduction? *The Plant journal: for cell and molecular biology*, 31(6), pp.699–712. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12220262> [Accessed December 1, 2014].
- Makridakis, N.M. & Reichardt, J.K., 2001. Multiplex automated primer extension analysis: simultaneous genotyping of several polymorphisms. *BioTechniques*, 31(6), pp.1374–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11768667> [Accessed July 5, 2014].
- Mandel, J.R. et al., 2013. Association Mapping and the Genomic Consequences of Selection in Sunflower N. M. Springer, ed. *PLoS Genetics*, 9(3), p.e1003378. Available at: <http://dx.plos.org/10.1371/journal.pgen.1003378> [Accessed March 23, 2013].
- Maringolo, C., 2007. *Regiones Cromosómicas Asociadas a Resistencia a Podredumbre Húmeda del Capítulo de Girasol (Sclerotinia sclerotiorum (Lib.) de Bary). Posgrado en Producción Vegetal*. Universidad Nacional de Mar del Plata. Facultad de Ciencias Agrarias, Unidad Integrada Balcarce Argentina.
- Masi, P., Zeuli, P.L.S. & Donini, P., 2003. Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.). *Molecular Breeding*, 11(4), pp.303–313. Available at: <http://link.springer.com/article/10.1023/A%3A1023443109985> [Accessed November 30, 2014].
- McCouch, S.R., 2002. Development and Mapping of 2240 New SSR Markers for Rice (*Oryza sativa* L.). *DNA Research*, 9(6), pp.199–207. Available at: <http://dnaresearch.oxfordjournals.org/content/9/6/199.abstract> [Accessed November 30, 2014].

- McCouch, S.R. et al., 2010. Development of genome-wide SNP assays for rice. *Breeding Science*, 60(5), pp.524–535. Available at: <http://joi.jlc.jst.go.jp/JST.JSTAGE/jsbbs/60.524?from=CrossRef> [Accessed May 29, 2014].
- McCouch, S.R. et al., 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology*, 35(1-2), pp.89–99. Available at: <http://link.springer.com/article/10.1023/A%3A1005711431474> [Accessed November 30, 2014].
- McCouch, S.R. & Doerge, R.W., 1995. QTL mapping in rice. *Trends in genetics*: *TIG*, 11(12), pp.482–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8533164> [Accessed December 1, 2014].
- McDonald, B.A. & Linde, C., 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual review of phytopathology*, 40, pp.349–79. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12147764> [Accessed July 10, 2014].
- McMullen, M.D. et al., 2009. Genetic properties of the maize nested association mapping population. *Science (New York, N.Y.)*, 325(5941), pp.737–40. Available at: <http://www.sciencemag.org/content/325/5941/737> [Accessed May 23, 2014].
- Mei, M. et al., 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 108(2), pp.280–91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14513220> [Accessed November 30, 2014].
- Mestries, E. et al., 1998. Analyses of quantitative trait loci associated with resistance to *Sclerotinia sclerotiorum* in sunflowers (*Helianthus annuus* L.) using molecular markers. *Molecular Breeding*, 4(3), pp.215–226.
- Meyers, B.C., Tingey, S. V & Morgante, M., 2001. Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. *Genome research*, 11(10), pp.1660–76. Available at: <http://genome.cshlp.org/content/11/10/1660> [Accessed November 27, 2014].
- Micic, Z., Hahn, V., Bauer, E., Melchinger, A.E., et al., 2005. Identification and validation of QTL for *Sclerotinia* midstalk rot resistance in sunflower by selective genotyping. *Tag Theoretical And Applied Genetics Theoretische Und Angewandte Genetik*, 111(2), pp.233–242. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15947909>.
- Micic, Z., Hahn, V., Bauer, E., Schön, C.C., et al., 2005. QTL mapping of resistance to *Sclerotinia* midstalk rot in RIL of sunflower population NDBLOSSel x CM625. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 110(8), pp.1490–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15841360> [Accessed May 23, 2013].
- Micic, Z. et al., 2004. QTL mapping of *Sclerotinia* midstalk-rot resistance in sunflower. *Tag Theoretical And Applied Genetics Theoretische Und Angewandte Genetik*, 109(7), pp.1474–1484. Available at: <http://www.hubmed.org/display.cgi?uids=15480534>.
- Mitchell, S.E. et al., 1997. Application of Multiplex PCR and Fluorescence-Based, Semi-Automated Allele Sizing Technology for Genotyping Plant Genetic Resources. *Crop Science*, 37(2), p.617. Available at: <https://www.crops.org/publications/cs/abstracts/37/2/CS0370020617?access=0&view=pdf> [Accessed November 30, 2014].
- Mogg, R. et al., 2002. Characterization of the flanking regions of *Zea mays* microsatellites reveals a large number of useful sequence polymorphisms. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 105(4), pp.532–543. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12582502> [Accessed November 30, 2014].
- Mohan, M. et al., 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding*, 3(2), pp.87–103. Available at: <http://link.springer.com/article/10.1023/A%3A1009651919792> [Accessed December 1, 2014].

- Molinero-Ruiz, M. & Dominguez, J., 2005. REACTION OF FIELD POPULATIONS OF SUNFLOWER DOWNY MILDEW (*Plasmopara halstedii*) TO METALAXYL AND MEFENOXAM/REACCIÓN DE LAS. *Helia*, (42), pp.65-74. Available at: <http://www.degruyter.com/view/j/helia.2005.28.issue-42/helia.2005.28.42.65/helia.2005.28.42.65.xml> [Accessed November 30, 2014].
- Mongkolporn, O., Dokmaihom, Y. & Kanchana-udomkan, C. Pakdeeveraporn, P., 2004. Genetic purity test of F1 hybrid Capsicum using molecular analysis. *Hort Science Biotech*, 79(3), pp.449-451.
- Monna, L. et al., 2006. Genome-wide searching of single-nucleotide polymorphisms among eight distantly and closely related rice cultivars (*Oryza sativa* L.) and a wild accession (*Oryza rufipogon* Griff.). *DNA research: an international journal for rapid publication of reports on genes and genomes*, 13(2), pp.43-51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16766512> [Accessed July 4, 2014].
- Moorthy, K.K. et al., 2011. Identification of informative EST-SSR markers capable of distinguishing popular Indian rice varieties and their utilization in seed genetic purity assessments. *Seed Science and Technology*, 39(2), pp.282-292. Available at: <http://www.ingentaconnect.com/content/ista/sst/2011/00000039/00000002/art00002> [Accessed November 30, 2014].
- Moragues, M. et al., 2010. Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 120(8), pp.1525-34. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20157694> [Accessed June 3, 2014].
- Morales, M. et al., 2004. Single-nucleotide polymorphisms detected in expressed sequence tags of melon (*Cucumis melo* L.). *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 47(2), pp.352-60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15060588> [Accessed July 5, 2014].
- Moreno, M. V. et al., 2013. Genetic characterization of sunflower breeding resources from Argentina: assessing diversity in key open-pollinated and composite populations. *Plant Genetic Resources*, 11(03), pp.238-249. Available at: [http://www.journals.cambridge.org/abstract\\_S1479262113000075](http://www.journals.cambridge.org/abstract_S1479262113000075) [Accessed June 19, 2014].
- Morgante, M., De Paoli, E. & Radovic, S., 2007. Transposable elements and the plant pan-genomes. *Current opinion in plant biology*, 10(2), pp.149-55. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17300983> [Accessed October 17, 2014].
- Moschini, R.C. et al., 2002. Predicción de la incidencia de la podredumbre blanda del capítulo de girasol en base a factores ambientales y genéticos. *Rev. Arg. Agrometeorología*, 2(1), pp.51-57.
- Mukhtar, M.S. et al., 2011. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science (New York, N.Y.)*, 333(6042), pp.596-601. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3170753&tool=pmcentrez&rendertype=abstract> [Accessed July 14, 2014].
- Mutschler, M.A. et al., 1996. QTL analysis of pest resistance in the wild tomato *Lycopersicon pennellii*: QTLs controlling acylsugar level and composition. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 92(6), pp.709-18. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24166395> [Accessed December 1, 2014].
- Nandini, R. & Chikkadevaiah, C., 2005. DNA fingerprinting of sunflower genotypes (*Helianthus annuus* L.). *Helia*, 28(42), pp.9-18. Available at: <http://www.doiserbia.nb.rs/Article.aspx?ID=1018-18060542009N>.
- Naresh, V. et al., 2009. EST-SSR marker-based assay for the genetic purity assessment of safflower hybrids. *Euphytica*, 170(3), pp.347-353. Available at: <http://link.springer.com/10.1007/s10681-009-9995-3> [Accessed November 30, 2014].
- Narvel, J.M. et al., 2000. Development of multiplex sets of simple sequence repeat DNA markers covering the soybean genome. *Molecular Breeding*, 6(2), pp.175-183. Available at:

- <http://link.springer.com/article/10.1023/A%3A1009637119947> [Accessed November 30, 2014].
- Natali, L. et al., 2006. Distribution of Ty3-gypsy- and Ty1-copia-like DNA sequences in the genus *Helianthus* and other Asteraceae. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 49(1), pp.64–72. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16462902>.
- Natali, L. et al., 1993. Nuclear DNA changes within *Helianthus annuus* L.: changes within single progenies and their relationships with plant development. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 85(5), pp.506–12. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24195922> [Accessed November 27, 2014].
- Natali, L. et al., 2013. The repetitive component of the sunflower genome as shown by different procedures for assembling next generation sequencing reads. *BMC genomics*, 14(1), p.686. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3852528&tool=pmcentrez&rendertype=abstract> [Accessed November 25, 2014].
- Neumann, P. et al., 2006. Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. *Genetics*, 173(2), pp.1047–56. Available at: <http://www.genetics.org/content/173/2/1047> [Accessed November 27, 2014].
- Nimchuk, Z. et al., 2003. Recognition and response in the plant immune system. *Annual review of genetics*, 37, pp.579–609. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14616074> [Accessed October 6, 2014].
- Nordborg, M. et al., 2002. The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature genetics*, 30(2), pp.190–193. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11780140> [Accessed June 14, 2011].
- Oefner, P.J., 2000. Allelic discrimination by denaturing high-performance liquid chromatography. *Journal of chromatography. B, Biomedical sciences and applications*, 739(2), pp.345–55. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10755378> [Accessed July 5, 2014].
- Oleykowski, C.A. et al., 1998. Mutation detection using a novel plant endonuclease. *Nucleic acids research*, 26(20), pp.4597–602. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=147896&tool=pmcentrez&rendertype=abstract> [Accessed July 5, 2014].
- Van Ooijen, J.W. & Voorrips, V.E., 2001. JoinMap 3.0, Software for the calculation of genetic linkage maps. *Plant Research International, Wageningen, the Netherlands*.
- Oraguzie et al., 2007. *Association mapping in plants.*, Springer-Verlag GmbH. Available at: <http://www.cabdirect.org/abstracts/20083279521.html> [Accessed June 16, 2011].
- Özdemir, N., Horn, R. & Friedt, W., 2004. Construction and characterization of a BAC library for sunflower (*Helianthus annuus* L.). *Euphytica*, 138(2), pp.177–183. Available at: <http://link.springer.com/10.1023/B:EUPH.0000046803.00050.ec> [Accessed July 4, 2014].
- Paglia, G.P., Olivieri, A.M. & Morgante, M., 1998. Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). *Molecular & general genetics*: MGG, 258(5), pp.466–78. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9669328> [Accessed July 7, 2014].
- Paniego, N. et al., 2012. Genomics, mapping and marker assisted selection strategies for disease resistance. In *18th International Sunflower Conference*. Mar del Plata, Argentina.
- Paniego, N. et al., 2002. Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.). *Genome*, 43(1428), pp.34–43.
- Paniego, N. et al., 2007. Sunflower. In *Oilseed. Genome Mapping and Molecular Breeding in Plants*. pp. 153–177.
- Pavy, N. et al., 2008. Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce

- and black spruce. *BMC genomics*, 9, p.21. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2246113&tool=pmcentrez&rendertype=abstract> [Accessed April 6, 2014].
- Pease, A.C. et al., 1994. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), pp.5022–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=43922&tool=pmcentrez&rendertype=abstract> [Accessed June 29, 2014].
- Pedraza, M.V. V & Pereyra, V.R.R., 2004. Infection courts and length of susceptible period related to sunflower head rot (*Sclerotinia sclerotiorum*) resistance. *Helia*, 27(40), pp.171–182. Available at: <http://www.doiserbia.nb.rs/Article.aspx?ID=1018-18060440171P> [Accessed June 15, 2012].
- Peerbolte, R.P., Paleman, J. & Peerbolte, 1998. The CARTISOL sunflower RFLP map (146 loci) extended with 291 RFLP markers. *Proceedings of the 18th sunflower research forum, Fargo*, pp.11–12.
- Pegadaraju, V. et al., 2013. De novo sequencing of sunflower genome for SNP discovery using RAD (Restriction site Associated DNA) approach. *BMC genomics*, 14(1), p.556. Available at: <http://www.biomedcentral.com/1471-2164/14/556> [Accessed November 17, 2014].
- Peluffo, L., 2007. A parallel analysis of metabolic and transcript profiles reveals key points in control of *Sclerotinia sclerotiorum* infection in sunflower. XIII International Congress on Molecular Plant-Microbe Interactions.
- Peluffo, L., 2010. *Characterization of the defense mechanisms against Sclerotinia sclerotiorum, causal agent of Sunflower Head Rot, through metabolic and transcriptional profiles analyses*, PhD thesis: University of Buenos Aires, Faculty of Exact and Natural Sciences.
- Peluffo, L. et al., 2010. Metabolic profiles of sunflower genotypes with contrasting response to *Sclerotinia sclerotiorum* infection. *Phytochemistry*, 71(1), pp.70–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19853265>.
- Perchepied, L. et al., 2010. Nitric oxide participates in the complex interplay of defense-related signaling pathways controlling disease resistance to *Sclerotinia sclerotiorum* in *Arabidopsis thaliana*. *Molecular plant-microbe interactions*: MPMI, 23(7), pp.846–60. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20521948> [Accessed November 24, 2014].
- Pereyra, V.R. & Valetti, O.E., 1993. *Producción de girasol, manual para productores del Pp, sudeste Bonaerense*. Ed. INTA Balcarce,
- Pereyra, V.R. & Bazzalo, M.E., 1988. Podredumbre del capítulo de girasol. Jornada de Actualización técnica para profesionales. 1988, INTA, Balcarce.
- Pereyra, V.R., Escande, A.R. & Pereyra, V.R., 1994. *Manual de reconocimiento de enfermedades del Girasol en la Argentina*. E. del I. CERBAS, ed., Balcarce, Provincia de Buenos Aires, Argentina.
- Perez Fernandez, J. & Corro Molas, A., 2003. *Efectos de los Causado, fungicidas curasemillas y fechas de siembra sobre Mildiu del girasol EEA, por Plasmopora halstedii (Fasrl.)*. Boletín de divulgación técnica N° 77. EEA Anguil. Cultivos de cosecha gruesa Actualizaciones.,
- Poland, J.A. et al., 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. T. Yin, ed. *PloS one*, 7(2), p.e32253. Available at: <http://dx.plos.org/10.1371/journal.pone.0032253> [Accessed July 29, 2014].
- Poland, J.A. et al., 2009. Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science*, 14(1), pp.21–29. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19062327>.
- Prats, E. et al., 2007. Antifungal activity of a new phenolic compound from capitulum of a head rot-resistant sunflower genotype. *Journal of chemical ecology*, 33(12), pp.2245–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18034282> [Accessed July 4, 2014].
- Purdy, L.H., 1979. *Sclerotinia sclerotiorum*: History, Diseases and Symptomatology, Host Range, Geographic Distribution, and Impact. *Phytopathology*, 69(8), p.875. Available at: 161

- [http://www.apsnet.org/publications/phytopathology/backissues/Documents/1979Abstracts/Phyto69\\_875.htm](http://www.apsnet.org/publications/phytopathology/backissues/Documents/1979Abstracts/Phyto69_875.htm) [Accessed July 4, 2014].
- Putt, E.D., 1997. Early history of sunflower. In Scheiter, ed. *Sunflower Production and Technology*. Madison, WI: American Society Agronomy, pp. 1–19.
- Radwan, O. et al., 2008. Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. *Molecular genetics and genomics*: MGG, 280(2), pp.111–125. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18553106> [Accessed June 14, 2011].
- Rafalski, A., 2002. Applications of single nucleotide polymorphisms in crop genetics. *Current opinion in plant biology*, 5(2), pp.94–100. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11856602>.
- Rafalski, J.A., 2010. Association genetics in crop improvement. *Current opinion in plant biology*, 13(2), pp.174–180. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20089441> [Accessed July 28, 2010].
- Ramsay, L. et al., 2000. A simple sequence repeat-based linkage map of barley. *Genetics*, 156(4), pp.1997–2005. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1461369&tool=pmcentrez&rendertype=abstract> [Accessed July 7, 2014].
- Ramsay, L. et al., 1999. Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *The Plant journal*: for cell and molecular biology, 17(4), pp.415–25. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10205898> [Accessed July 7, 2014].
- Randall, L.P., Coldham, N.G. & Woodward, M.J., 2005. Detection of mutations in Salmonella enterica gyrA, gyrB, parC and parE genes by denaturing high performance liquid chromatography (DHPLC) using standard HPLC instrumentation. *The Journal of antimicrobial chemotherapy*, 56(4), pp.619–23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16141278> [Accessed July 4, 2014].
- Raven, P.H. & Axelrod, D.I., 1974. Angiosperm Biogeography and Past Continental Movements. *Annals of the Missouri Botanical Garden*, 61(3), p.539. Available at: <http://www.jstor.org/stable/2395021?origin=crossref> [Accessed July 4, 2014].
- Regente, M. & de la Canal, L., 2003. A cDNA encoding a putative lipid transfer protein expressed in sunflower seeds. *Journal of plant physiology*, 160(2), pp.201–3. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12685036> [Accessed July 4, 2014].
- Regente, M.C. et al., 1997. A sunflower leaf antifungal peptide active against Sclerotinia sclerotiorum. *Physiologia Plantarum*, 100(1), pp.178–182. Available at: <http://doi.wiley.com/10.1111/j.1399-3054.1997.tb03470.x> [Accessed July 4, 2014].
- Regente, M.C. & de la Canal, L., 2000. Purification, characterization and antifungal properties of a lipid-transfer protein from sunflower (*Helianthus annuus*) seeds. *Physiologia Plantarum*, 110(2), pp.158–163. Available at: <http://www.blackwell-synergy.com/links/doi/10.1034/j.1399-3054.2000.110203.x> [Accessed July 4, 2014].
- Ren, B., Zhou, J.-M. & Komiyama, M., 2004. Straightforward detection of SNPs in double-stranded DNA by using exonuclease III/nuclease S1/PNA system. *Nucleic acids research*, 32(4), p.e42. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=390314&tool=pmcentrez&rendertype=abstract> [Accessed July 5, 2014].
- Di Rienzo, J.A. et al., 2014. InfoStat versión 2014. Available at: <http://www.infostat.com.ar>.
- Rieseberg et al., 1993. Genomic map of a diploid hybrid species. *HEREDITY-LONDON*. Available at: <http://www3.botany.ubc.ca/rieseberglab/loren>.

- Rieseberg, L.H., 2001. Chromosomal rearrangements and speciation. *Trends in ecology & evolution*, 16(7), pp.351–358. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11403867> [Accessed November 27, 2014].
- Rieseberg, L.H., 1998. Genetic Mapping as a Tool for Studying Speciation. *Molecular Systematics of Plants II*, pp.459–487. Available at: [http://link.springer.com/chapter/10.1007/978-1-4615-5419-6\\_16](http://link.springer.com/chapter/10.1007/978-1-4615-5419-6_16) [Accessed May 13, 2014].
- Rieseberg, L.H. et al., 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science (New York, N.Y.)*, 301(5637), pp.1211–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12907807> [Accessed October 21, 2014].
- Rieseberg, L.H., Van Fossen, C. & Desrochers, A.M., 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature*, 375(6529), pp.313–316. Available at: <http://www.nature.com/doifinder/10.1038/375313a0> [Accessed November 18, 2014].
- Roberts, E.H., 1988. Temperature and seed germination. *Symposia of the Society for Experimental Biology*, 42, pp.109–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3077854> [Accessed December 1, 2014].
- Röder, M.S. et al., 1998. A microsatellite map of wheat. *Genetics*, 149(4), pp.2007–23. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1460256&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- Rodríguez, M.A. et al., 2004. Histopathology of *Sclerotinia sclerotiorum* attack on flower parts of *Helianthus annuus* heads in tolerant and susceptible varieties. *Mycopathologia*, 157(3), pp.291–302. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15180158> [Accessed July 4, 2014].
- Ronicke, S. et al., 2004. Interspecific hybrids of sunflower as a source of *Sclerotinia* resistance. *Plant Breeding*, 123(2), pp.152–157. Available at: <http://doi.wiley.com/10.1046/j.1439-0523.2003.00925.x> [Accessed July 4, 2014].
- Rönicke, S. et al., 2005. Quantitative Trait Loci Analysis of Resistance to *Sclerotinia sclerotiorum* in Sunflower. *Phytopathology*, 95(7), pp.834–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18943017> [Accessed May 21, 2013].
- Ronicke, S., Hahn, V. & Friedt, W., 2005. Resistance to *Sclerotinia sclerotiorum* of “high oleic” sunflower inbred lines. *Plant Breeding*, 124(4), pp.376–381. Available at: <http://doi.wiley.com/10.1111/j.1439-0523.2005.01119.x>.
- Roorkiwal, M. et al., 2013. Single Nucleotide Polymorphism Genotyping for Breeding and Genetics Applications in Chickpea and Pigeonpea using the BeadXpress Platform. *The Plant Genome*, 6(2). Available at: <https://www.crops.org/publications/tpg/abstracts/6/2/plantgenome2013.05.0017> [Accessed March 21, 2014].
- Rostoks, N., Mudie, S., et al., 2005. Genome- wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Molecular genetics and genomics*: MGG, 274(5), pp.515–27. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16244872> [Accessed May 30, 2014].
- Rostoks, N. et al., 2006. Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proceedings of the National Academy of Sciences of the United States of America*, 103(49), pp.18656–61. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1693718&tool=pmcentrez&rendertype=abstract> [Accessed July 3, 2014].
- Rostoks, N., Borevitz, J.O., et al., 2005. Single -feature polymorphism discovery in the barley transcriptome. *Genome biology*, 6(6), p.R54. Available at: <http://genomebiology.com/2005/6/6/R54> [Accessed November 10, 2014].
- Roux, F. et al., 2014. Resistance to phytopathogens e tutti quanti: placing plant quantitative disease resistance on the map. *Molecular plant pathology*, 15(5), pp.427–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24796392> [Accessed November 5, 2014].

- Sala, C.A. et al., 2012. Sunflower: Improving Crop Productivity and Abiotic Stress Tolerance. In *Improving Crop Resistance to Abiotic Stress, Volume 1 & Volume 2* (eds N. Tuteja, S. S. Gill, A. F. Tiburcio and R. Tuteja), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany. doi: 10.1002/9783527632930.ch47.
- Sambrook, J., Fritsch, E.F. & Maniatis, T., 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor laboratory press, Cold Spring Harbor.
- Sandve, G.K. et al., 2010. The Genomic HyperBrowser: inferential genomics at the sequence level. *Genome biology*, 11(12), p.R121. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3046481&tool=pmcentrez&rendertype=abstract> [Accessed July 5, 2014].
- Santini, S. et al., 2002. Ty1 /copia- and Ty3 /gypsy-like DNA sequences in Helianthus species. *Chromosoma*, 111(3), pp.192–200. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12355209>.
- Schnabl, H. et al., 2002. Biotechnological methods applied to produce Sclerotinia sclerotiorum resistant sunflower. *Helia*, 25(36), pp.191–197. Available at: <http://www.doiserbia.nb.rs/Article.aspx?ID=1018-18060236191S> [Accessed July 4, 2014].
- Schneider, A.A. & Miller, J.F., 1981. Description of Sunflower Growth Stages1. *Crop Science*, 21(6), p.901. Available at: <https://dl.sciencesocieties.org/publications/cs/abstracts/21/6/CS0210060901>.
- Schuster, I. & Cruz, C., 2004. *Estatística genômica aplicada a populações derivadas de cruzamentos controlados*. Ediciones Universidade Federal de Viçosa, Viçosa, MG, ed.,
- Sebastian, R.L. et al., 2000. An integrated AFLP and RFLP Brassica oleracea linkage map from two morphologically distinct doubled-haploid mapping populations. *TAG Theoretical and Applied Genetics*, 100(1), pp.75–81. Available at: <http://link.springer.com/10.1007/s001220050011> [Accessed November 30, 2014].
- Seeb, L.W. et al., 2011. Single nucleotide polymorphisms across a species' range: implications for conservation studies of Pacific salmon. *Molecular ecology resources*, 11 Suppl 1, pp.195–217. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21429175> [Accessed July 4, 2014].
- Seiler, G.J., 2010. Sunflower breeding on resistance to diseases. In *International Symposium "Sunflower Breeding on Resistance to Diseases"*. Krasnodart, Russia, p. 36.
- Seiler, G.J., Rieseberg, L.H. & Seiler, G.J., 1997. Systematics, Origin, and Germplasm Resources of the Wild and Domesticated Sunflower. In A. A. Schneider, ed. *Sunflower Technology and Production*. Madison: Amer. Soc. Agronomy, pp. 21–65.
- Sharopova, N. et al., Development and mapping of SSR markers for maize. *Plant molecular biology*, 48(5-6), pp.463–81. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12004892> [Accessed November 30, 2014].
- Shen, R. et al., 2005. High -throughput SNP genotyping on universal bead arrays. *Mutation research*, 573(1-2), pp.70–82. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15829238>.
- Shirasawa, K. et al., 2010. SNP discovery and linkage map construction in cultivated tomato. *DNA research*: an international journal for rapid publication of reports on genes and genomes, 17(6), pp.381–91. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2993540&tool=pmcentrez&rendertype=abstract> [Accessed July 5, 2014].
- Smirnoff, N. & Cumbes, Q.J., 1989. Hydroxyl radical scavenging activity of compatible solutes. *Phytochemistry*, 28(4), pp.1057–1060. Available at: <http://www.sciencedirect.com/science/article/pii/0031942289801827> [Accessed December 1, 2014].
- Smith, B.D. & Yarnell, R.A., 2009. Initial formation of an indigenous crop complex in eastern North America at 3800 B.P. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16), pp.6561–6. Available at:

- <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2666091&tool=pmcentrez&rendertype=abstract> [Accessed May 30, 2013].
- Sobrinho, B., Brión, M. & Carracedo, A., 2005. SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic science international*, 154(2-3), pp.181–94. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16182964> [Accessed June 1, 2014].
- Somers, D.J. et al., 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 46(3), pp.431–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12834059> [Accessed July 5, 2014].
- Song, B.-H. et al., 2009. Multilocus patterns of nucleotide diversity, population structure and linkage disequilibrium in *Boechera stricta*, a wild relative of *Arabidopsis*. *Genetics*, 181(3), pp.1021–33. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2651039&tool=pmcentrez&rendertype=abstract> [Accessed June 19, 2014].
- Song, Q. et al., 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. T. Zhang, ed. *PLoS one*, 8(1), p.e54985. Available at: <http://dx.plos.org/10.1371/journal.pone.0054985> [Accessed November 9, 2014].
- Southern, E.M., 1996. High-density gridding: techniques and applications. *Current opinion in biotechnology*, 7(1), pp.85–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8742380> [Accessed July 5, 2014].
- Staton, S.E. et al., 2012. The sunflower (*Helianthus annuus* L.) genome reflects a recent history of biased accumulation of transposable elements. *The Plant journal: for cell and molecular biology*, 72(1), pp.142–153. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22691070>.
- Stevens, P., 2010. Angiosperm phylogeny Website. Available from <http://www.Mobot.Org/mobot/research/apweb/>.
- Studer, A. et al., 2011. Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nature genetics*, 43(11), pp.1160–3. Available at: <http://dx.doi.org/10.1038/ng.942> [Accessed October 20, 2014].
- Syed, N.H. et al., 2006. A detailed linkage map of lettuce based on SSAP, AFLP and NBS markers. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 112(3), pp.517–27. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16341837> [Accessed November 30, 2014].
- Syvänen, A. C., 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nature reviews. Genetics*, 2(12), pp.930–942. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11733746>.
- Talia, P. et al., 2010. Cytological characterization of sunflower by in situ hybridization using homologous rDNA sequences and a BAC clone containing highly represented repetitive retrotransposon-like sequences. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 53(3), pp.172–179. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20237595>.
- Talia, P. et al., 2010. Genetic mapping of EST-SSRs, SSR and InDels to improve saturation of genomic regions in a previously developed sunflower map. *Electronic Journal of Biotechnology*, 13.
- Tang, S. et al., 2002. Simple sequence repeat map of the sunflower genome. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 105(8), pp.1124–1136. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12582890> [Accessed June 3, 2011].
- Tang, S., Kishore, V.K. & Knapp, S.J., 2003. PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 107(1), pp.6–19. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12835928> [Accessed March 11, 2013].

- Tanksley, S.D. & Jones, R.A., 1981. Application of ADH Isozymes in testing the genetic purity of F1 hybrids of tomato. *Hortscience*, 16(1), pp.179–181.
- Temnykh, S. et al., 2000. Mapping and genome organization of microsatellite sequences in rice ( *Oryza sativa* L.). *TAG Theoretical and Applied Genetics*, 100(5), pp.697–712. Available at: <http://link.springer.com/10.1007/s001220051342> [Accessed July 7, 2014].
- Thomson, M.J. et al., 2010. Development and application of 96- and 384-plex single nucleotide polymorphism ( SNP ) marker sets for diversity analysis , mapping and marker-assisted selection in rice. , (March), pp.22–26.
- Thomson, M.J. et al., 2011. High- throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. *Molecular Breeding*, 29(4), pp.875–886. Available at: <http://link.springer.com/10.1007/s11032-011-9663-x> [Accessed August 10, 2013].
- Thuault, M.C. & Tourvieille De Labrouhe, D., 1988. Etudes du pouvoir pathogène de huit isolats de *Sclerotinia* appartenant aux espèces *S. sclerotiorum*, *S. minor* et *S. trifoliorum* sur le tournesol. *Int Tech CETIOM*, 103, pp.21–27.
- Timms, L. et al., 2006. Analyses of synteny between *Arabidopsis thaliana* and species in the Asteraceae reveal a complex network of small syntenic segments and major chromosomal rearrangements. *Genetics*, 173(4), pp.2227–35. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1569713&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- Troglia, C., 2003. *Factores ambientales que afectan la supervivencia y producción del inóculo de Sclerotinia sclerotiorum y umbral mínimo de inóculo para la podredumbre húmeda del capítulo de girasol*. Tesis de Maestría Universidad Nacional de Mar del Plata, Facultad de Ciencias Agrarias, Unidad Integrada Balcarce Argentina.
- Troglia, C., Maringolo, C. & Escande, A., 2005. Caracterización fenotípica de podredumbre húmeda del capítulo por *Sclerotinia sclerotiorum* en dos poblaciones de girasol. XIII Congreso Latinoamericano de Fitopatología.
- Tsuchihashi, Z. & Dracopoli, N.C., 2002. Progress in high throughput SNP genotyping methods. *The pharmacogenomics journal*, 2(2), pp.103–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12049172> [Accessed July 5, 2014].
- Turkington, T.K., 1993. Use of Petal Infestation to Forecast *Sclerotinia* Stem Rot of Canola: The Influence of Inoculum Variation over the Flowering Period and Canopy Density. *Phytopathology*, 83(6), p.682. Available at: [http://www.apsnet.org/publications/phytopathology/backissues/Documents/1993Abstracts/Phyto\\_83\\_682.htm](http://www.apsnet.org/publications/phytopathology/backissues/Documents/1993Abstracts/Phyto_83_682.htm) [Accessed July 4, 2014].
- Ungerer, M.C., Strakosh, S.C. & Zhen, Y., 2006. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Current biology*: CB, 16(20), pp.R872–3. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17055967> [Accessed November 30, 2014].
- Urdangarín, M.C. et al., 2000. A defensin gene expressed in sunflower inflorescence. *Plant Physiology and Biochemistry*, 38(3), pp.253–258. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0981942800007373> [Accessed July 4, 2014].
- Van, B. & Miller, J.F.F., 2004. Methods of inoculation of sunflower heads with *Sclerotinia sclerotiorum*. *Helia*, 27(41), pp.137–142. Available at: <http://www.doiserbia.nb.rs/Article.aspx?ID=1018-18060441137V> [Accessed November 27, 2014].
- Van, K. et al., 2004. Discovery of single nucleotide polymorphisms in soybean using primers designed from ESTs. *Euphytica*, 139(2), pp.147–157. Available at: <http://www.springerlink.com/index/10.1007/s10681-004-2561-0>.

- Van, K. et al., 2005. Discovery of SNPs in soybean genotypes frequently used as the parents of mapping populations in the United States and Korea. *The Journal of heredity*, 96(5), pp.529–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15994422> [Accessed July 19, 2010].
- Vanderplank, J.E., 1969. *Disease Resistance in Plants.*, New York: Academic Press. Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.166.3905.593> [Accessed November 24, 2014].
- Varshney, R.K. et al., 2007. A high density barley microsatellite consensus map with 775 SSR loci. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 114(6), pp.1091–103. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17345060> [Accessed November 30, 2014].
- Vear, F. et al., 2007. Improvement of *Sclerotinia sclerotiorum* head rot resistance in sunflower by recurrent selection of a restorer population. *Helia*, 30(46), pp.1–12. Available at: <http://www.doiserbia.nbs.bg.ac.yu/Article.aspx?ID=1018-18060746001V> [Accessed May 23, 2013].
- Vear, F. et al., 2008. Inheritance of quantitative resistance to downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus* L.). *Euphytica*, 164(2), pp.433–444. Available at: <http://link.springer.com/10.1007/s10681-008-9698-1> [Accessed May 23, 2013].
- Vear, F. & Grezes-Besset, B., 2010. Progress in breeding sunflowers for resistance to *Sclerotinia*. In *International Symposium "Sunflower Breeding on Resistance to Diseases."* Krasnodart, Russia, pp. 31–36.
- Vear, F. & Tourvieille de Labrouhe, D., 1988. Heredity of resistance to *Sclerotinia sclerotiorum* in sunflowers. II. - Study of capitulum resistance to natural and artificial ascospore infections. *Agronomie*, 8(6), pp.503–508. Available at: <http://www.agronomy-journal.org/10.1051/agro:19880604> [Accessed May 23, 2013].
- VEAR, F. & TOURVIEILLE DE LABROUHE, D., 1984. Recurrent selection for resistance to *Sclerotinia sclerotiorum* in sunflowers using artificial infections. *Agronomie*, 4(8), pp.789–794. Available at: [http://www.researchgate.net/publication/248849934\\_Recurrent\\_selection\\_for\\_resistance\\_to\\_Sclerotinia\\_sclerotiorum\\_in\\_sunflowers\\_using\\_artificial\\_infections](http://www.researchgate.net/publication/248849934_Recurrent_selection_for_resistance_to_Sclerotinia_sclerotiorum_in_sunflowers_using_artificial_infections) [Accessed December 1, 2014].
- Vetter, M.M. et al., 2012. Flagellin perception varies quantitatively in *Arabidopsis thaliana* and its relatives. *Molecular biology and evolution*, 29(6), pp.1655–67. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22319159> [Accessed November 24, 2014].
- Vicient, C.M., Kalendar, R. & Schulman, A.H., 2005. Variability, recombination, and mosaic evolution of the barley BARE-1 retrotransposon. *Journal of molecular evolution*, 61(3), pp.275–91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16034651> [Accessed November 27, 2014].
- De Vienne, D., 2003. *Introduction In: Molecular markers in plant genetics and biotechnology.* D. Vienne, ed.,
- Virányi, F., 2008. Research progress in sunflower diseases and their management. In *17th International Sunflower Conference.* Córdoba, España.
- Vitte, C. & Bennetzen, J.L., 2006. Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 103(47), pp.17638–43. Available at: <http://www.pnas.org/content/103/47/17638> [Accessed November 27, 2014].
- Vogl, C. & Xu, S., 2000. Multipoint mapping of viability and segregation distorting loci using molecular markers. *Genetics*, 155(3), pp.1439–47. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1461139&tool=pmcentrez&rendertype=abstract> [Accessed July 8, 2014].

- Voorrips, R.E., 2002. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *Journal of Heredity*, 93(1), pp.77–78. Available at: <http://jhered.oupjournals.org/cgi/doi/10.1093/jhered/93.1.77>.
- Wagner, T. et al., 1999. Denaturing high-performance liquid chromatography detects reliably BRCA1 and BRCA2 mutations. *Genomics*, 62(3), pp.369–376. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10644434>.
- Wang, J. et al., 2005. Methods for genome-wide analysis of gene expression changes in polyploids. *Methods Enzymology*. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0076687905950301> [Accessed June 15, 2011].
- Wang, N. et al., 2008. A functional genomics resource for Brassica napus: development of an EMS mutagenized population and discovery of FAE1 point mutations by TILLING. *The New phytologist*, 180(4), pp.751–65. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18811617> [Accessed July 5, 2014].
- Wang, S., Basten, C.J. & Zeng, Z.B., 2011. Windows QTL Cartographer 2.5. *The Journal of Infectious Diseases*, 204 Suppl(3), pp.198–199. Available at: <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>.
- Wang, X.F., Knoblauch, R. & Leist, N., 2005. Varietal discrimination of tomato (*Lycopersicon esculentum*L.) by ultrathin-layer isoelectric focusing of seed protein. *Seed Science Technology*, 28, pp.521–526.
- Weising, K. et al., 1995. *DNA Fingerprinting in Plants and Fungi*, CRC Press.
- West, M.A.L. et al., 2006. High-density haplotyping with microarray-based expression and single feature polymorphism markers in Arabidopsis. *Genome Research*, 16(6), pp.787–795. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1473188&tool=pmcentrez&rendertype=abstract> [Accessed October 29, 2014].
- Wills, D.M. & Burke, J.M., 2007. Quantitative Trait Locus Analysis of the Early Domestication of Sunflower. *Genetics*, 176(4), pp.2589–2599. Available at: <http://www.genetics.org/cgi/doi/10.1534/genetics.107.075333>.
- Winter, P. et al., 2000. A linkage map of the chickpea (*Cicer arietinum* L.) genome based on recombinant inbred lines from a *C. arietinum* × *C. reticulatum* cross: localization of resistance genes for fusarium wilt races 4 and 5. *TAG Theoretical and Applied Genetics*, 101(7), pp.1155–1163. Available at: <http://link.springer.com/10.1007/s001220051592> [Accessed June 21, 2014].
- Wu, M.S. et al., 2006. Wu, M.S., Jia, X.H., Tian, L., Lv, B.C., 2006. Rapid and reliable purity identification of F1 hybrids of maize (*Zea mays* L.) using SSR markers. *Mol. Plant Breeding* 4, 381–384. *Molecular Plant Breeding*, 4, pp.381–384.
- Xiao, H. et al., 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science (New York, N.Y.)*, 319(5869), pp.1527–30. Available at: <http://www.sciencemag.org/content/319/5869/1527> [Accessed October 16, 2014].
- Xiao, S. et al., 2001. Broad-spectrum mildew resistance in Arabidopsis thaliana mediated by RPW8. *Science (New York, N.Y.)*, 291(5501), pp.118–20. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11141561> [Accessed November 24, 2014].
- Xiao, W. & Oefner, P.J., 2001. Denaturing high-performance liquid chromatography: A review. *Human mutation*, 17(6), pp.439–474. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11385705>.
- Yang, X. et al., 2010. Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 121(3), pp.417–431. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20349034>.

- Yu et al., 2003. Towards a saturated molecular genetic linkage map for cultivated sunflower. *Crop Science*, 43(1), pp.367–387. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Towards+a+Saturated+Molecular+Genetic+Linkage+Map+for+Cultivated+Sunflower#0> [Accessed June 15, 2011].
- Yu, B. et al., 2005. Denaturing high performance liquid chromatography: high throughput mutation screening in familial hypertrophic cardiomyopathy and SNP genotyping in motor neurone disease. *Journal of clinical pathology*, 58(5), pp.479–85. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1770671&tool=pmcentrez&rendertype=abstract> [Accessed November 30, 2014].
- Yu, J. et al., 2008. Genetic design and statistical power of nested association mapping in maize. *Genetics*, 178(1), pp.539–551. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2206100&tool=pmcentrez&rendertype=abstract>.
- Yu, J. et al., 2002. Towards a Saturated Molecular Genetic Linkage Map for Cultivated Sunflower. , pp.367–387.
- Yue, B. et al., 2008. Identifying quantitative trait loci for resistance to Sclerotinia head rot in two USDA sunflower germplasms. *Phytopathology*, 98(8), pp.926–31. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18943211>.
- Zeng, Z.B., 1994. Precision mapping of quantitative trait loci. *Genetics*, 136(4), pp.1457–1468. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1205924&tool=pmcentrez&rendertype=abstract>.
- Zeng, Z.B., 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America*, 90(23), pp.10972–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=47903&tool=pmcentrez&rendertype=abstract> [Accessed November 24, 2014].
- Zhang, D.Q. et al., 2002. Segregation of AFLP markers in a (*Populus tomentosa* x *P. bolleana*) x *P. Tomentosa* Carr. BC1 family. *For Stud China*, 4, pp.21–16.
- Zhang, F., Wan, X.-Q. & Pan, G.-T., 2006. QTL mapping of Fusarium moniliforme ear rot resistance in maize. 1. Map construction with microsatellite and AFLP markers. *Journal of applied genetics*, 47(1), pp.9–15. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16424603> [Accessed November 30, 2014].
- Zhao et al., 2007. Analysis of gene expression profiles in response to Sclerotinia sclerotiorum in Brassica napus. *Planta*, 227(1), pp.13–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17665211> [Accessed December 4, 2010].
- Zhu, C. et al., 2008. Status and Prospects of Association Mapping in Plants. *The Plant Genome Journal*, 1(1), p.5. Available at: <https://www.crops.org/publications/tpg/abstracts/1/1/5>.
- Zilic, S. et al., 2010. Characterization of sunflower seed and kernel proteins. *Helia*, 33(52), pp.103–113. Available at: <http://www.doiserbia.nb.rs/Article.aspx?ID=1018-18061052103Z>.
- Zimmermann, G. et al., 2006. The multigene family encoding germin-like proteins of barley. Regulation and function in Basal host resistance. *Plant physiology*, 142(1), pp.181–92. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1557593&tool=pmcentrez&rendertype=abstract> [Accessed May 30, 2014].
- Zipfel, C. et al., 2004. Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature*, 428(6984), pp.764–7. Available at: <http://dx.doi.org/10.1038/nature02485> [Accessed November 5, 2014].

## BIBLIOGRAFÍAS

- Zubrzycki, J. et al., 2012. Biparental QTL and Association Mapping for Sclerotinia Head Rot Resistance in Cultivated Sunflower. In *18th International Sunflower Conference*. Available at: [http://www.asagir.org.ar/asagir2008/buscar\\_congreso.asp](http://www.asagir.org.ar/asagir2008/buscar_congreso.asp).
- Zubrzycki, J.E., 2009. *MAPEO DE REGIONES GENÓMICAS DE GIRASOL ASOCIADAS A CARACTERÍSTICAS DE INTERÉS AGRONÓMICO*. Universidad Nacional de Misiones. Facultad de Ciencias Exactas, Químicas y Naturales.

# APÉNDICE

TABLA I. Listado de marcadores SSRs seleccionados para el control de pureza. La descripción de cada uno de ellos es presentada.									
Nombre del marcador	Nº de acceso en el Genebank	Motivo	Fluoróforo	alelo RHA266	alelo PAC2	Tamaño de alelo esperado	Oligonucleótido sentido	Oligonucleótido antisentido	GL
HA1402	BV727899	ATT	FAM	168	153	204	GTGTGAATGTGAGTGTGAATCC	GTCACAGTTGCATATCCATCC	17
HA1590	BV727994	GT	FAM	68	66	71	GTGGGAGAGAAGATAGGAG	GGGGATTATGCAACCTTGCC	-
HA1752	BV728001	ATC	FAM	139	143	147	CCTTCTCCTCAAATCCCG	CGTCATTGTCATTGAGAAGATCG	-
HA1770	BV727917	GA	HEX	179	181	186	GCCGGTCACCCATTCCCCCTTAC	GGCGGAAGCCAGGCGTTGTAGCG	3
HA2147	BV728115	GA	FAM	84	86	90	CCCCATAAACATTTACACCC	CCTGAGAAATCGGTAGATAGTG	-
HA2190	BV728130	GA	HEX	188	192	189	CAAACCCTAATCGCCCAATTGC	GGAAGCTTGTGAATCTGGAGGG	3
HA2605	BV728347	GA	HEX	76	89	151	GCAGATTACAATATGCACAC	GTGTTGAAAGTAGTTGGTTGC	8
HA2989	BV728090	ATCCTC- ATC ATC	HEX	127	142	145	GCCTGTGTCCTCCTCAGAGTC	GTTATGTGGGCTGGCTGCACC	14
HA3204	BV728181		NED	116	119	121	GCCCTTCAATCCTACCATTAA C	GGATATTGAGTTGGTTGTTGGG	10
HA3312	BV728075	GA	FAM	110	100	115	TAACTGCAACAGACATGGAA	ATAACGATTGCACAACACAA	6
HA3555	BV728375	CAT	FAM	102	104	110	GATATCTCTCATAAGTGCCG	GGTCTTGTGATGACGAAGC	12
HA3683	BV728278	GA GT	FAM	186	182	189	TGCCCTTATTCTCCTAATTG	GTGGATCTTGACAGTGTTT	16
HA3971	BV728318	GA	HEX	183	178	179	GGATCTCTCCTGGTCCTAGT	GATCAATATCTTCCATGTTAT	17
HA4021	BV728320	GA	HEX	194	185	188	GTCTGTCTGTACAGAATTCTC	GAAGCAGAAGGTTAGGGCTATG	-
HA4057	BV728333	GT	HEX	206	201	172	AAACCCTTCCGACTTATCTC	TAAAGAGAGAGCCCAACAAG	3
HA4105	BV728354	GA	NED	113	119	118	GGGACTAGTTTGTAAATATCTCTC	GTTTGAGAGTGGTGATAGGTTATG	3
HA541	BV727938	GT	HEX	213	229	220	CAACACACGCTTTGTTGGGTG	CGATAAGGAGGGAGGGGAAGATAG	3
HA806	G67410	GA-GT	HEX	192	186	191	GATGTTCCCTCCTGCAC	GGTTGGATAATGGGGCAGC	-
HA911	BV727888	GA	FAM	186	179	179	CAAAGTTCACCTCGTTTTTC	AAGTGGGAAGGTCTACGAGT	8
ORS1041	BV006423	CT	FAM	281	183	279	AAACAAACCTTAATGGGGTCGTA	ATATTGGCTGGTTGATGCTGAT	7
ORS1097	BV006475	CT	FAM	164	160	161	GACCAAGTGGACTGACACGAG	CGGTGGTGGCTTAGATTGTATAGT	17
ORS1146	BV006511	AG	FAM	344	378	379	GGCTCATCACTTGATCTATTGT	TGAAGACACCATCTCCAATGC	11
ORS1242	BV006598	CT	HEX	265	255	269	GCAATCGTTTTCACTCTTCCATTC	TGGTCGTAGAATTGTCGGTCAT	15
ORS1265	BV006617	CT	HEX	229	227	222	GGGTTTAGCAAATAATAGGCACA	ACCCTTGGAGTTTAGGGATCA	9

APÉNDICE

**TABLA I. Listado de marcadores SSRs seleccionados para el control de pureza.** La descripción de cada uno de ellos es presentada.

ORS166	BV012434	CA	FAM	339	327	344	CAGCCACATGCCCTCTGAC	TGTTAAGAACCGCGACAACCTGC	-
ORS297	BV006634	GT	HEX	226	222	225	GTGTCTGCACGAACTGTGGT	TGCAAAGCTCACACTAACCTG	17
ORS342	BV005933	GT	HEX	340	330	342	TGTTTCATCAGGTTTGTCTCCA	CACCAGCATAGCCATTCAAA	2
ORS460	BV005998	GA	FAM	302	300	302	ACTCGGCTACCACCTCACAC	GCCCTTTGACCCTAACCAAG	7
ORS483	BV006013	GT	FAM	267	269	271	CCGAACAACAATCTCCACAA	GGTTTAGGTGTGCATCACA	6
ORS510	BV006030	GT	HEX	258	249	255	CATCGCGTCCCTCTCTCTAA	CCAACCATCACAGCAATCAG	9
ORS533	BV006044	CT	FAM	162	160	157	TGGTGGAGGTCACTATTGGA	AGGAAAGAAGGAAGCCGAGA	5
ORS613	BV006091	AG	HEX	226	230	228	GTAACCCTAGGTCAATTTGCAG	ATCTCCGAAAACATTCTCG	-
ORS687	BV006138	CT	HEX	166	164	168	ACCGTTACTATTGGTTATTTCATT	GGGGTTTGTGTCTGTTTTG	15
ORS727	BV006163	AG	HEX	154	196	192	GGTGGCAAGTGGTGGTTAGA	AAAGCTGGTCATCTCAAGGGTA	17
ORS767	BV006189	GT	FAM	369	371	368	ACAGGAGCTTTCGGTCTTCA	GGCACAGTCTCGTTCATATCC	12
ORS788	BV006202	GA	FAM	264	262	263	CTGGATAAAGATGGGATAAAGAGAG	GGACCCACCAAGATTTGTTTT	16
ORS844	BV006252	AC	FAM	304	302	307	ACGATGCAAAGAATACTGCAC	CATGTTAATAGGTTTTAATTCTAGGG	-
ORS878	BV006281	AC	FAM	191	200	203	TGCAAGGTATCCATATCCACAA	TATACGCACCGAAAGAAAGTC	10
ORS959	BV006356	GT	HEX	247	236	240	CCGTAAGTATAAACCGCCTATT	CGTCCTCTCGCATCAATCTTAT	1

**TABLA II. Genotipos obtenidos para las 93 RILS ensayadas mediante la técnica dHPLC.** Los alelos derivados de PAC2 fueron identificados como "1", los de RHA266 como alelo "2", los no genotipados como "0" y las indeterminaciones como "\*".

RIL	CAM	LIM	PGIC	SCR1	RIL	CAM	LIM	PGIC	SCR1
C33	2	1	2	2	C125	2	1	2	2
C34	1	2	2	1	C126	2	1	1	2
C35	2	2	2	2	C127	1	1	1	2
C36	2	1	2	1	C128	2	1	1	2
C38	2	1	2	1	C129	2	2	1	2
C39	1	1	2	2	C130	1	2	1	1
C40	1	1	2	2	C133	2	2	2	1
C42	2	1	2	2	C134	1	1	1	1
C47	1	2	2	2	C135	2	2	2	1
C54	2	1	1	1	C136	2	1	2	2
C55	2	*	2	2	C137	2	*	1	1
C56	1	1	1	2	C138	1	2	2	2
C60	2	2	1	1	C139	1	1	1	1
C64	2	*	2	1	C141	1	1	1	1

APÉNDICE

**TABLA II. Genotipos obtenidos para las 93 RILS ensayadas mediante la técnica dHPLC.** Los alelos derivados de PAC2 fueron identificados como "1", los de RHA266 como alelo "2", los no genotipados como "0" y las indeterminaciones como "\*".

C65	2	1	1	2	C143	2	2	1	1
C66	2	1	0	2	C144	1	1	2	2
C67	2	2	2	1	C146	1	1	2	2
C69	1	1	1	2	C148	1	*	*	1
C71	2	2	2	1	C149	2	2	1	1
C75	2	2	1	2	C150	2	2	1	1
C76	2	2	2	2	C153	1	1	2	2
C77	2	1	1	2	C154	2	2	2	2
C78	1	2	2	2	LR1	2	1	1	1
C79	2	1	1	1	LR2	2	1	2	1
C81	1	1	1	1	LR5	1	1	2	2
C82	0	0	0	0	LR7	1	1	2	2
C83	2	2	2	1	LR11	1	1	1	2
C84	1	2	1	2	LR16	1	1	1	1
C88	2	1	1	1	LR18	1	2	2	1
C89	1	1	2	1	LR19	2	2	2	1
C90	2	1	1	2	LR20	2	2	2	1
C91	2	1	1	1	LR25	2	1	2	1
C92	1	2	2	*	LR29	1	2	2	1
C94	1	2	1	1	LR32	2	2	2	1
C99	0	0	0	0	LR34	2	1	*	2
C100	2	1	1	1	LR35	1	2	2	2
C104	1	2	1	1	LR39	1	2	1	*
C105	1	2	1	1	LR41	2	1	2	2
C107	2	2	2	2	LR42	1	2	1	2
C108	1	2	2	1	LR45	2	2	2	1
C114	2	2	1	1	LR46	2	2	1	2
C115	1	1	2	1	LR54	2	1	2	2
C116	2	1	2	2	LR59	1	2	1	2
C120	0	0	0	0	LR63	2	2	2	2
C121	2	2	1	1	LR64	1	2	1	2
C123	2	2	2	2	LR67	2	1	1	1
C124	1	1	2	1					

APÉNDICE

TABLA III. Condiciones ambientales para el período de inoculación a lo largo de las campañas							
DÍA	MES	AÑO	TEMP MAX	TEMP MIN	TEMP MEDIA	HUM REL	PRECIP (mm)
6	1	2010	24,9	13,8	19,4	48	
7	1	2010	29,2	9,9	19,6	47	
8	1	2010	29,9	13,6	21,8	57	
9	1	2010	34,8	16,4	25,6	53	
10	1	2010	26,3	16,8	21,6	49	
11	1	2010	28,8	8,8	18,8	52	27,5
12	1	2010	23,9	15,4	19,2	52	
13	1	2010	26,5	11,8	19,2	48	
14	1	2010	30,7	12,2	21,5	52	
15	1	2010	31,5	12,5	22	53	
16	1	2010	29,1	15,2	22,2	70	
17	1	2010	31,7	17,1	24,4	59	
18	1	2010	29,7	13,1	21,4	66	4
19	1	2010	29,4	18,2	23,8	79	0,8
20	1	2010	31,9	11,5	21,7	60	
21	1	2010	33,2	17,2	25,2	66	
22	1	2010	27,4	12	19,7	62	
23	1	2010	32,4	14	23,2	72	4
24	1	2010	36	20	28	59	11,5
25	1	2010	26,9	16,2	21,6	66	
6	2	2012	34,8	19	26,9	59	
7	2	2012	35,2	13,6	24,4	51	0,4
8	2	2012	25,3	15,9	20,6	64	
9	2	2012	23,4	10,5	17	59	
10	2	2012	29,3	8	18,7	63	
11	2	2012	33	9,5	21,3	53	
12	2	2012	37,2	14,6	25,9	55	
13	2	2012	31,1	18	24,6	68	
14	2	2012	34,8	17,1	26	54	
15	2	2012	37,1	19,6	28,4	47	0,8
16	2	2012	33,2	19,8	26,5	63	0,4
17	2	2012	30,2	15,6	22,9	70	65,5
18	2	2012	25,8	21,3	23,6	89	20,5
19	2	2012	29,3	18,8	24,1	63	
20	2	2012	31,5	17,4	24,5	67	5,3
8	2	2013	33	15,2	24,1	66	
9	2	2013	32,2	15,9	24,1	63	
10	2	2013	32,5	19,6	26,1	75	17,8
11	2	2013	33,3	19,3	26,3	65	
12	2	2013	30	17,7	23,9	57	
13	2	2013	30,2	13,2	21,7	64	

APÉNDICE

14	2	2013	32,3	10,7	21,5	59	
15	2	2013	34	20	27	56	
16	2	2013	34,2	20	27,1	67	
17	2	2013	26,5	15,2	20,9	69	
18	2	2013	24,8	9	16,9	60	
19	2	2013	28	9,6	18,8	52	
20	2	2013	26,8	11,4	19,1	62	
7	2	2014	29,7	15,5	22,6	65	
8	2	2014	29,3	12,9	21,1	78	0,5
9	2	2014	30,3	13,2	21,8	59	
10	2	2014	29,9	12,2	21,1	57	4,5
11	2	2014	26,9	17,7	22,3	89	9
12	2	2014	24,4	15,3	19,9	85	
13	2	2014	23,8	9,7	16,8	70	
14	2	2014	22,9	7,5	15,2	70	
15	2	2014	24,3	10,1	17,2	69	
16	2	2014	25,5	15,7	20,6	71	2,5
17	2	2014	26,6	17,9	22,3	82	
18	2	2014	24,6	19,6	22,1	94	15
19	2	2014	28,8	18	23,4	89	12
PROMEDIO 2009-10			29,7	14,2	21,9	58,5	9,56
PROMEDIO 2011-12			31,4	15,9	23,7	61,7	15,5
PROMEDIO 2012-13			30,6	15,1	22,9	62,7	17,8
PROMEDIO 2013-14			26,7	14,3	20,5	75,2	7,3

RIL	INCIDENCIA					SEVERIDAD					INTENSIDAD					PERÍODO DE INCUBACIÓN				
	PROMEDIO	09-10	11-12	12-13	13-14	PROMEDIO	09-10	11-12	12-13	13-14	PROMEDIO	09-10	11-12	12-13	13-14	PROMEDIO	09-10	11-12	12-13	13-14
C122	0,82 A				0,85 A	0,44 A				0,40 A	0,24 A				0,23 B	16,32 B				16,85 A
LR18	0,80 A	0,94 A	0,41 A	0,89 A	0,77 A	0,36 A	0,24 B	0,11 B	0,64 B	0,46 A	0,28 A	0,20 A	0,03 B	0,50 A	0,40 B	17,25 B	20,05 A	21,82 C	19,57 B	15,80 A
WA03-80	0,79 A	0,68 A	0,12 B	1,00 A	0,91 A	0,30 A	0,17 B	0,26 B	0,24 B	0,33 A	0,13 B	0,04 B	0,05 B	0,20 B	0,16 C	19,03 B	22,61 B	16,98 B	20,61 B	19,55 B
C34	0,73 A	1,00 A	0,35 A	0,76 A	0,71 A	0,54 A	0,18 B	0,76 A	0,74 B	0,60 A	0,31 A	0,11 B	0,08 B	0,54 A	0,52 A	17,11 B	19,40 A	22,06 C	16,07 B	15,70 A
C130	0,72 A	0,75 A	0,33 A	0,95 A		0,34 A	0,11 B	0,67 A	0,45 B		0,17 B	0,05 B	0,18 A	0,30 A		18,40 B	18,40 A	19,27 B	19,62 B	
C142	0,69 A			0,93 A	0,66 A	0,41 A			0,70 B	0,20 B	0,19 B			0,45 A	0,06 C	18,53 B			19,03 B	18,70 A





APÉNDICE

**TABLA IV. Medias ajustadas para todas las variables analizadas.** La asignación de letras diferentes por medio del test DGC, indica pertenencia a grupos cuyas medias son significativamente diferentes.

LR41	0,28	A	0,23	B	0,00	B	0,76	A	0,44	A	0,26	A	0,28	B		0,45	B	0,13	B	0,10	B	0,09	B	0,02	B	0,28	B	0,03	C	18,68	B	21,08	A		18,86	B	21,25	B		
LR46	0,27	A	0,61	A	0,04	B	0,36	B	0,34	B	0,30	A	0,08	B	0,29	B	0,60	B	0,28	A	0,07	B	0,02	B	0,03	B	0,15	B	0,06	C	18,56	B	21,30	A	17,45	B	18,54	B	21,15	B
C91	0,27	A			0,00	B	0,75	A	0,24	B	0,06	B				0,21	B	0,09	B	-0,02	B			0,03	B	0,11	B	-0,01	C	22,31	B				21,37	B	24,40	B		
LR19	0,27	A	0,22	B	0,03	B	0,56	B	0,64	A	0,44	A	0,22	B	0,80	A	0,54	B	0,41	A	0,12	B	0,02	B	0,01	B	0,24	B	0,25	B	19,59	B	24,81	B	17,67	B	20,44	B	18,85	A
LR20	0,25	A			0,00	B	0,62	B			0,23	A				0,42	B			0,01	B			0,04	B	0,14	B			23,21	B					22,98	C			
C121	0,25	A	0,00	B	0,06	B	0,58	B	0,37	B	0,38	A		1,09	A	0,46	B	0,18	B	0,11	B	-0,01	B	0,04	B	0,35	A	0,02	C	19,17	B		21,41	C	17,40	B	22,90	B		
C47	0,25	A	0,26	B	0,02	B	0,26	B	0,51	A	0,36	A		0,27	B	0,57	B	0,33	A	0,08	B			0,03	B	0,18	B	0,09	C	20,23	B	22,93	B	17,86	B	19,62	B	22,35	B	
C129	0,23	A					0,27	B	0,39	A							0,38	A	0,05	B							0,08	C	21,24	B							22,30	B		
WA03-73	0,23	A	0,13	B	0,00	B	0,62	B	0,49	A	0,24	A	0,15	B		0,35	B	0,23	B	0,07	B	0,01	B	0,03	B	0,12	B	0,09	C	17,36	B	18,76	A			20,96	B	17,05	A	
WA03-90	0,22	B	0,50	A	0,02	B	0,45	B	0,20	B	0,12	A	0,03	B	0,03	B	0,25	B	0,21	B	0,00	B	0,00	B	-0,02	B	0,04	B	0,02	C	20,54	B	27,32	B	22,48	C	23,69	C	16,50	A
C104	0,21	B	0,45	A	0,00	B	0,19	B	0,50	A	0,23	A	0,12	B		0,49	B	0,10	B	-0,01	B	0,03	B	-0,01	B	0,04	B	0,03	C	22,34	B	24,50	B			22,22	B	24,00	B	
WA03-133	0,21	B	0,33	A					0,39	A	0,23	B								0,14	B	0,05	B							16,29	B	18,30	A							
C99	0,20	B	0,08	B	0,18	A	0,54	B	0,33	B	0,39	A	0,11	B	0,61	A	0,57	B	0,29	A	0,12	B	0,01	B	0,08	B	0,31	A	0,10	C	18,22	B	20,44	A	17,26	B	17,62	B	20,87	B
WA03-50	0,20	B	0,39	A	0,00	B	0,45	B			0,30	A	0,43	A		0,40	B			0,01	B	0,04	B	-0,01	B	0,10	B			21,32	B	23,77	B			21,48	B			
C43	0,19	B	0,30	A	0,06	B	0,57	B	0,14	B	0,33	A	0,16	B	0,24	B	0,45	B	0,47	A	0,07	B	0,01	B	0,05	B	0,18	B	0,04	C	19,79	B	26,12	B	18,29	B	20,03	B	17,50	A
LR32	0,19	B			0,05	B			0,22	B	0,19	A			-0,10	B			0,25	B	0,07	B			-0,01	B		0,08	C	22,08	B			22,22	C			22,80	B	
LR67	0,18	B	0,32	A			0,43	B	0,34	B	0,26	A	0,38	A		0,38	B	0,17	B	-0,01	B	0,04	B			0,09	B	0,03	C	20,13	B	20,81	A			21,59	B	22,75	B	
C62	0,18	B			0,02	B	0,38	B			0,08	B		0,16	B	0,12	B			0,03	B			0,00	B	0,05	B			21,05	B			17,22	B	23,67	C			
C81	0,15	B	0,21	B	0,00	B	0,60	B	0,22	B	0,16	A				0,42	B	0,02	B	0,05	B			0,01	B	0,24	B	0,00	C	20,63	B	25,41	B			19,53	B	23,15	B	
C33	0,15	B			0,02	B	0,41	B	0,16	B	0,17	A		0,04	B	0,75	B	0,09	B	0,05	B			-0,01	B	0,30	A	0,00	C	19,82	B		24,20	C	16,23	B	21,25	B		
LR7	0,14	B	0,14	B	0,00	B	0,65	A			0,28	A	0,08	B		0,50	B			0,05	B	0,01	B	0,03	B	0,15	B			19,85	B	20,89	A			27,46	C			
C41	0,13	B	0,19	B					0,30	A	0,15	B								0,12	B	0,03	B							18,27	B	21,26	A							
C125	0,11	B	0,17	B	0,05	B	0,56	B	0,02	B	0,23	A		0,34	B	0,41	B	-0,01	B	0,02	B			0,02	B	0,15	B	-0,02	C	21,81	B	24,36	B	21,33	C	20,88	B	28,00	C	
C82	0,11	B	0,10	B	0,01	B	0,48	B	0,14	B	0,20	A		0,07	B	0,42	B	0,20	B	0,00	B			-0,01	B	0,10	B	0,01	C	22,56	B	25,92	B	23,06	C	21,08	B	24,80	B	
C101	0,10	B					0,21	B			0,01	B				0,21	B			-0,11	C					0,06	B			26,05	B					25,78	C			
C138	0,10	B	0,14	B	0,00	B	0,32	B			0,40	A				0,60	B			0,08	B	-0,01	B	0,02	B	0,20	B			18,30	B	24,91	B			13,52	A			
C39	0,10	B	0,07	B	0,00	B	0,45	B	0,22	B	0,22	A				0,53	B	0,14	B	0,06	B	0,01	B	0,01	B	0,25	B	0,01	C	20,75	B	23,06	B			18,11	B	23,90	B	
LR5	0,09	B	0,10	B	0,00	B	0,46	B	0,12	B	0,13	A	0,15	B		0,26	B	0,16	B	-0,02	B	0,03	B	0,03	B	0,04	B	0,03	C	20,44	B	25,20	B			28,93	C	20,30	B	
C147	0,07	B	0,12	B	0,00	B	0,13	B	0,10	B	0,04	B	0,07	B		0,10	B	-0,01	B	0,00	B	0,01	B	0,02	B	0,03	B	-0,01	C	21,91	B	23,27	B			23,04	C	24,50	B	
C40	0,05	B	0,00	B	0,00	B			0,24	B	0,40	A					0,39	A	0,07	B	0,00	B	0,00	B			0,03	C	22,67	B							23,85	B		
C65	0,04	B					0,26	B	0,00	B	-0,19	B				0,03	B			-0,04	B					0,01	B	0,00	C	24,29	B					23,93	C			

APÉNDICE

TABLA V. Resumen de las anotaciones ontológicas para los genes candidatos SNPs corroborados por el análisis <i>in silico</i> de sus secuencias				
	Annotation	BLAST		
Feature Name	Go Term Annotation	Match accession	Name	Norm Score
HeAn_C_10493.1	metabolic process ; catalytic activity ; transition metal ion binding ; oxidation-reduction process ; oxidoreductase activity ; zinc ion binding ; transferase activity, transferring acyl groups other than amino-acyl groups ; binding ;	XP_002300061	predicted protein [Populus trichocarpa]	974.0
		EEE84866	predicted protein [Populus trichocarpa]	974.0
		BAJ34603	unnamed protein product [Thellungiella halophila]	969.0
		EFH39540	oxidoreductase [Arabidopsis lyrata subsp. lyrata]	956.0
		XP_002863281	oxidoreductase [Arabidopsis lyrata subsp. lyrata]	956.0
		AAK76531	unknown protein [Arabidopsis thaliana]	955.0
		AAL06521	AT4g13010/F25G13_100 [Arabidopsis thaliana]	955.0
		AAL08252	AT4g13010/F25G13_100 [Arabidopsis thaliana]	955.0
		AAM51229	unknown protein [Arabidopsis thaliana]	955.0
		NP_193037	oxidoreductase, zinc-binding dehydrogenase family protein [Arabidopsis thaliana]	955.0
HeAn_C_11352.1		AAD25161	proline transporter 2 [Solanum lycopersicum]	845.0
		AAL91181	proline transporter 2 [Arabidopsis thaliana]	837.0
		AAP49512	At3g55740 [Arabidopsis thaliana]	837.0
		EFH57713	hypothetical protein ARALYDRAFT_482636 [Arabidopsis lyrata subsp. lyrata]	834.0
		XP_002881454	hypothetical protein ARALYDRAFT_482636 [Arabidopsis lyrata subsp. lyrata]	834.0
		EFH56089	hypothetical protein ARALYDRAFT_483024 [Arabidopsis lyrata subsp. lyrata]	832.0
		XP_002879830	hypothetical protein ARALYDRAFT_483024 [Arabidopsis lyrata subsp. lyrata]	832.0
		EFH52589	hypothetical protein ARALYDRAFT_486008 [Arabidopsis lyrata subsp. lyrata]	831.0
		XP_002876330	hypothetical protein ARALYDRAFT_486008 [Arabidopsis lyrata subsp. lyrata]	831.0
		AAM63275	putative proline transporter [Arabidopsis thaliana]	828.0
HeAn_R_534.1	carbon fixation ; glyceraldehyde-3-phosphate dehydrogenase activity ; plastid ; nucleotide binding ; hexose metabolic process ; oxidation-reduction process ; oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor ; glucose metabolic process ; NAD binding ; glyceraldehyde-3-phosphate dehydrogenase activity ; binding ;	XP_002307806	predicted protein [Populus trichocarpa]	1818.0
		EEE94802	predicted protein [Populus trichocarpa]	1818.0
		AAA84543	glyceraldehyde-3-phosphate dehydrogenase B subunit [Pisum sativum]	1814.0
		P12859	RecName: Full=Glyceraldehyde-3-phosphate dehydrogenase B, chloroplastic; AltName: Full=NADP-dependent glyceraldehydephosphate dehydrogenase subunit B; Flags: Precursor	1814.0
		CBI40744	unnamed protein product [Vitis vinifera]	1814.0
		XP_002300632	predicted protein [Populus trichocarpa]	1787.0
		EEE79905	predicted protein [Populus trichocarpa]	1787.0
		BAJ34314	unnamed protein product [Thellungiella halophila]	1765.0

APÉNDICE

TABLA V. Resumen de las anotaciones ontológicas para los genes candidatos SNPs corroborados por el análisis <i>in silico</i> de sus secuencias				
		AAK62594	At1g42970/F13A11_3 [Arabidopsis thaliana]	1756.0
		AAK64065	putative glyceraldehyde-3-phosphate dehydrogenase [Arabidopsis thaliana]	1756.0
		CBI24757	unnamed protein product [Vitis vinifera]	557.0
		XP_002323710	predicted protein [Populus trichocarpa]	525.0
		EEF03843	predicted protein [Populus trichocarpa]	525.0
		AAC23406	hypothetical protein [Arabidopsis thaliana]	419.0
		AAL49852	unknown protein [Arabidopsis thaliana]	419.0
		AAM20216	unknown protein [Arabidopsis thaliana]	419.0
HeAn_C_11434.1	intracellular ; kinase activity ;	O80568	RecName: Full=Inositol-tetrakisphosphate 1-kinase 4; AltName: Full=Inositol 1,3,4-trisphosphate 5/6-kinase 4; Short=Atltk-4; Short=Inositol-triphosphate 5/6-kinase 4; Short=Ins(1,3,4)P(3) 5/6-kinase 4	419.0
		EFH58203	inositol 1,3,4-trisphosphate 5/6-kinase family protein [Arabidopsis lyrata subsp. lyrata]	412.0
		XP_002881944	inositol 1,3,4-trisphosphate 5/6-kinase family protein [Arabidopsis lyrata subsp. lyrata]	412.0
		NP_001063689	Os09g0518700 [Oryza sativa Japonica Group]	358.0
		XP_002326118	predicted protein [Populus trichocarpa]	787.0
		AAD00295	auxin-binding protein ABP19 [Prunus persica]	746.0
		Q9ZRA4	RecName: Full=Auxin-binding protein ABP19a; Flags: Precursor	746.0
		AAB51240	auxin-binding protein [Prunus persica]	743.0
		ACA03785	putative auxin-binding protein 2 [Prunus salicina]	743.0
GLP4	transition metal ion binding ; signal transducer activity ; hormone-mediated signaling pathway ; extracellular region ; external encapsulating structure ; apoplast ; manganese ion binding ; nutrient reservoir activity ;	O04011	RecName: Full=Auxin-binding protein ABP20; Flags: Precursor	743.0
		ACG69486	germin-like protein 10 [Glycine max]	737.0
		ADD71879	germin-like protein subfamily 3 member 4 precursor [Arachis hypogaea]	728.0
		ACA03784	putative auxin-binding protein 3 [Prunus salicina]	728.0
		AAK62573	AT5g20630/T1M15_30 [Arabidopsis thaliana]	725.0
		AAP59427	phospholipid hydroperoxide glutathione peroxidase [Solanum lycopersicum]	727.0
		AAL40914	phospholipid hydroperoxide glutathione peroxidase [Momordica charantia]	702.0
HeAn_C_11835.1	metabolic process ; response to stress ; peroxidase activity ; oxidation-reduction process ; response to oxidative stress ; glutathione peroxidase activity ;	EFH38550	ATGPX2 [Arabidopsis lyrata subsp. lyrata]	695.0
		XP_002862292	ATGPX2 [Arabidopsis lyrata subsp. lyrata]	695.0
		EFH57440	ATGPX2 [Arabidopsis lyrata subsp. lyrata]	694.0
		XP_002881181	ATGPX2 [Arabidopsis lyrata subsp. lyrata]	694.0

APÉNDICE

TABLA V. Resumen de las anotaciones ontológicas para los genes candidatos SNPs corroborados por el análisis <i>in silico</i> de sus secuencias				
		XP_002310444	glutathione peroxidase [Populus trichocarpa]	689.0
		EEE90894	glutathione peroxidase [Populus trichocarpa]	689.0
		AAD24836	putative glutathione peroxidase [Arabidopsis thaliana]	688.0
		AAL25600	At2g31570/T9H9.9 [Arabidopsis thaliana]	688.0
		ACD56611	putative photosystem II protein [Gossypioides kirkii]	715.0
		Q9SMB4	RecName: Full=Photosystem II 22 kDa protein, chloroplastic; AltName: Full=CP22; Flags: Precursor	709.0
		AAG48610	photosystem II 22 kDa protein precursor [Solanum sogarandinum]	708.0
		Q9FPP4	RecName: Full=Photosystem II 22 kDa protein, chloroplastic; AltName: Full=CP22; Flags: Precursor	708.0
HeAn_C_11928.1	photosystem ; cellular metabolic process ; intrinsic to membrane ; chloroplast thylakoid ;	CBI19017	unnamed protein product [Vitis vinifera]	708.0
		AAA63649	22 kDa component of photosystem II [Solanum lycopersicum]	706.0
		P54773	RecName: Full=Photosystem II 22 kDa protein, chloroplastic; AltName: Full=CP22; Flags: Precursor	706.0
		ABF72923	chloroplast chlorophyll a/b-binding protein [Solanum nigrum]	703.0
		ABF72924	chloroplast chlorophyll a/b-binding protein [Solanum nigrum]	703.0
		XP_002300987	hypothetical protein POPTRDRAFT_816277 [Populus trichocarpa]	693.0
		EEE86435	predicted protein [Populus trichocarpa]	1862.0
		XP_002305924	predicted protein [Populus trichocarpa]	1862.0
		XP_002327592	predicted protein [Populus trichocarpa]	1840.0
		CBI28628	unnamed protein product [Vitis vinifera]	1832.0
		AAN15672	putative protein kinase gb AAD22129 [Arabidopsis thaliana]	1719.0
HeAn_C_12025.1	catalytic activity ; metabolic process ; oxidation-reduction process ; oxidoreductase activity ;	AAO85405	putative amine oxidase 2 [Arabidopsis thaliana]	1719.0
		NP_176759	ATPAO4 (ARABIDOPSIS THALIANA POLYAMINE OXIDASE 4); amine oxidase/ polyamine oxidase [Arabidopsis thaliana]	1719.0
		Q8H191	RecName: Full=Probable polyamine oxidase 4; Short=AtPAO4; AltName: Full=Amine oxidase 2	1719.0
		AAK43885	putative protein kinase gb AAD22129 [Arabidopsis thaliana]	1716.0
		EFH63242	ATPAO4 [Arabidopsis lyrata subsp. lyrata]	1705.0
		XP_002300824	homocysteine s-methyltransferase [Populus trichocarpa]	1408.0
HeAn_C_12824.1	one-carbon metabolic process ; S-methyltransferase activity ; homocysteine S-methyltransferase activity ;	EEE80097	homocysteine s-methyltransferase [Populus trichocarpa]	1408.0
		XP_002307602	homocysteine s-methyltransferase [Populus trichocarpa]	1398.0
		EEE94598	homocysteine s-methyltransferase [Populus trichocarpa]	1398.0

APÉNDICE

TABLA V. Resumen de las anotaciones ontológicas para los genes candidatos SNPs corroborados por el análisis <i>in silico</i> de sus secuencias				
		ABF47292	selenocysteine methyltransferase [Camellia sinensis]	1362.0
		AAF23822	homocysteine S-methyltransferase AtHMT-2 [Arabidopsis thaliana]	1292.0
		CBI28727	unnamed protein product [Vitis vinifera]	1289.0
		EFH54766	homocysteine S-methyltransferase AtHMT-2 [Arabidopsis lyrata subsp. lyrata]	1285.0
		XP_002878507	homocysteine S-methyltransferase AtHMT-2 [Arabidopsis lyrata subsp. lyrata]	1285.0
		AAQ22634	At3g63250/F16M2_100 [Arabidopsis thaliana]	1282.0
	primary metabolic process ; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen ; oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor ; lipid metabolic process ; oxidation-reduction process ; oxidoreductase activity, acting on paired donors, with oxidation of a pair of donors resulting in the reduction of molecular oxygen to two molecules of water ;	AAL68982	delta-12 oleate desaturase [Helianthus annuus]	2095.0
		AAX19392	delta-12 oleate desaturase [Helianthus annuus]	2095.0
		AAX19393	delta-12 oleate desaturase [Helianthus annuus]	2095.0
		AAX19394	delta-12 oleate desaturase [Helianthus annuus]	2095.0
		AAX19395	delta-12 oleate desaturase [Helianthus annuus]	2095.0
		AAX19396	delta-12 oleate desaturase [Helianthus annuus]	2095.0
HeAn_C_1638.1		AAK26633	delta-12 fatty acid desaturase FAD2 [Calendula officinalis]	1951.0
		AAF04094	delta-12 oleate desaturase [Vernonia galamensis]	1937.0
		AAF04093	delta-12 oleate desaturase [Vernonia galamensis]	1928.0
		AAL68983	delta-12 oleate desaturase [Helianthus annuus]	1925.0
	cation binding ; phosphate metabolic process ; adenylyl ribonucleotide binding ; plastid ; phosphotransferase activity, paired acceptors ; phosphorylation ; transferase activity, transferring phosphorus-containing groups ; catalytic activity ;	AAA86940	cold stable pyruvate, orthophosphate dikinase [Flaveria brownii]	958.0
		Q39734	RecName: Full=Pyruvate, phosphate dikinase, chloroplastic; AltName: Full=Cold-sensitive pyruvate, orthophosphate dikinase; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor	958.0
		Q42736	RecName: Full=Pyruvate, phosphate dikinase, chloroplastic; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor	955.0
HeAn_C_185.1		AAA86941	cold stable pyruvate, orthophosphate dikinase [Flaveria bidentis]	954.0
		Q39735	RecName: Full=Pyruvate, phosphate dikinase, chloroplastic; AltName: Full=Cold-sensitive pyruvate, orthophosphate dikinase; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor	954.0
		P22221	RecName: Full=Pyruvate, phosphate dikinase, chloroplastic; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor	954.0
		CBI26150	unnamed protein product [Vitis vinifera]	894.0

APÉNDICE

TABLA V. Resumen de las anotaciones ontológicas para los genes candidatos SNPs corroborados por el análisis <i>in silico</i> de sus secuencias				
		EFH44457	hypothetical protein ARALYDRAFT_915248 [Arabidopsis lyrata subsp. lyrata]	879.0
		XP_002868198	hypothetical protein ARALYDRAFT_915248 [Arabidopsis lyrata subsp. lyrata]	879.0
		Q42910	RecName: Full=Pyruvate, phosphate dikinase, chloroplastic; AltName: Full=Pyruvate, orthophosphate dikinase; Flags: Precursor	877.0
		AAK83073	putative cytosolic NADP-malic enzyme [Flaveria pringlei]	2939.0
		AAK83074	putative cytosolic NADP-malic enzyme [Flaveria pringlei]	2920.0
		BAG16517	putative NADP-dependent malic enzyme [Capsicum chinense]	2613.0
		AAB58728	cytosolic NADP-malic enzyme [Solanum lycopersicum]	2599.0
		CBI28009	unnamed protein product [Vitis vinifera]	2562.0
HeAn_C_4316.1	intracellular part ; oxidation-reduction process ; malate metabolic process ; malate dehydrogenase (oxaloacetate-decarboxylating) activity ; malic enzyme activity ; NAD binding ; oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor ; binding ;	XP_002324450	predicted protein [Populus trichocarpa]	2560.0
		EEF03015	predicted protein [Populus trichocarpa]	2560.0
		CBI20568	unnamed protein product [Vitis vinifera]	2559.0
		AAA67087	malate dehydrogenase (NADP+) [Vitis vinifera]	2556.0
		P51615	RecName: Full=NADP-dependent malic enzyme; Short=NADP-ME	2556.0
		ADL36590	BHLH domain class transcription factor [Malus x domestica]	1141.0
		XP_002302066	predicted protein [Populus trichocarpa]	1133.0
		EEE81339	predicted protein [Populus trichocarpa]	1133.0
		XP_002306828	predicted protein [Populus trichocarpa]	1102.0
HeAn_C_4498.1	regulation of transcription, DNA-dependent ; nucleus ;	EEE93824	predicted protein [Populus trichocarpa]	1102.0
		CBI17963	unnamed protein product [Vitis vinifera]	1093.0
		AAL16298	AT4g16430/dl4240w [Arabidopsis thaliana]	985.0
		AAL15257	AT4g16430/dl4240w [Arabidopsis thaliana]	985.0
		AAN73298	At4g16430/dl4240w [Arabidopsis thaliana]	985.0
		AAL55710	putative transcription factor BHLH3 [Arabidopsis thaliana]	985.0
		ABW80752	chloroplast ribulose 1,5-bisphosphate carboxylase/oxygenase activase [Flaveria bidentis]	1563.0
		ADD60245	rubisco activase [Glycine max]	1554.0
		ADD60248	beta-form rubisco activase [Glycine max]	1549.0
HeAn_R_174.2	adenyl ribonucleotide binding ; ATP binding ;	ADD60246	beta-form rubisco activase [Glycine max]	1549.0
		ABI94075	chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase activase small protein isoform [Acer rubrum]	1538.0
		ABI94076	chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase activase large protein isoform [Acer rubrum]	1538.0

APÉNDICE

TABLA V. Resumen de las anotaciones ontológicas para los genes candidatos SNPs corroborados por el análisis <i>in silico</i> de sus secuencias				
		CBI30513	unnamed protein product [Vitis vinifera]	1533.0
		ABI94077	chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase activase small protein isoform [Acer rubrum]	1529.0
		ABI94078	chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase activase large protein isoform [Acer rubrum]	1529.0
		ABB20913	ribulose-1,5-bisphosphate carboxylase/oxygenase activase alpha 2 [Gossypium hirsutum]	1526.0
		AAR83891	ubiquitin-conjugating enzyme 8 [Capsicum annuum]	789.0
		ACF72670	ubiquitin carrier-like protein [Cucumis melo var. cantalupensis]	787.0
		AAG40371	AT4g27960 [Arabidopsis thaliana]	785.0
		AAN13102	E2 ubiquitin-conjugating enzyme 9 (UBC9) [Arabidopsis thaliana]	785.0
		XP_002309194	predicted protein [Populus trichocarpa]	785.0
		XP_002325633	predicted protein [Populus trichocarpa]	785.0
HeAn_C_8647.2	adenyl ribonucleotide binding ; ligase activity, forming carbon-nitrogen bonds ; acid-amino acid ligase activity ;	EEE92717	predicted protein [Populus trichocarpa]	785.0
		EEF00015	predicted protein [Populus trichocarpa]	785.0
		AAA32894	ubiquitin conjugating enzyme [Arabidopsis thaliana]	785.0
		P35132	RecName: Full=SUMO-conjugating enzyme UBC9; AltName: Full=UBCAT4B; AltName: Full=Ubiquitin carrier protein 9; AltName: Full=Ubiquitin-conjugating enzyme E2 9; AltName: Full=Ubiquitin-conjugating enzyme E2-17 kDa; AltName: Full=Ubiquitin-protein ligase 9	785.0
		AAD56390	1-deoxy-D-xylulose-5-phosphate synthase [Artemisia annua]	707.0
		ACF21004	1-deoxy-D-xylulose 5-phosphate synthase [Salvia miltiorrhiza]	669.0
		AAD38941	1-D-deoxyxylulose 5-phosphate synthase [Solanum lycopersicum]	667.0
		CAZ66648	1-deoxy-D-xylulose 5-phosphate synthase 1 precursor [Solanum lycopersicum]	667.0
		ADK73609	1-D-deoxyxylulose 5-phosphate synthase [Solanum tuberosum]	665.0
HeAn_R_283.1	thiamine metabolic process ; oxidoreductase activity, acting on the aldehyde or oxo group of donors ; transferase activity, transferring aldehyde or ketonic groups ; terpenoid metabolic process ; metabolic process ; catalytic activity ;	XP_002312717	predicted protein [Populus trichocarpa]	664.0
		EEE90084	predicted protein [Populus trichocarpa]	664.0
		AAS99588	chloroplast 1-deoxy-D-xylulose-5-phosphate synthase [Elaeis guineensis]	663.0
		AAQ84169	1-deoxy-D-xylulose 5-phosphate synthase [Pueraria montana var. lobata]	662.0
		ACF60511	1-D-deoxyxylulose 5-phosphate synthase [Nicotiana tabacum]	662.0

APÉNDICE

TABLA VI. Resultados de la comparación efectuada para los SSRs contra la base de datos Heliagene									
SSR	HIT	FROM-TO QUERY REGION	QUERY COVERAGE (%)	QUERY LENGHT	SCORE	IDENTITY	E-VALUE	GO	INTERPROSCAN
HA1108	Ha412T4I739C0S1	20-135	30,91	372	440	87.93	6,83E-15	BP: response to oxidative stress (GO:0006979)	<i>IPR000889:Glutathione peroxidase; IPR012335; IPR012336:Thioredoxin-like fold;</i>
								BP: oxidation-reduction process (GO:0055114)	
MF: glutathione peroxidase activity (GO:0004602)									
HA3102	Ha412T4I10350C0S4	246-378	34,92	378	657	99.25	8,12E-28	BP: regulation of transcription (GO:0045449)	<i>IPR005539:ELK; IPR005540:KNOX1; IPR005541:KNOX2; IPR012287:Homeodomain-related; IPR016233:Homeobox Pitx/unc30;</i>
								CC: nucleus (GO:0005634)	
								DNA-dependent (GO:0006355)	
	Ha412T4I10350C0S3	246-378	34,92	378	657	99.25	8,12E-28	MF: sequence-specific DNA binding transcription factor activity (GO:0003700)	
								MF: transcription regulator activity (GO:0030528)	
								MF: protein binding (GO:0005515)	
MF: sequence-specific DNA binding (GO:0043565)									
MF: DNA binding (GO:0003677)									
HA3621	Ha412T4I41930C0S1	1-288	57,29	501	1245	93.43	2,25E-62		
	Ha412T4I41930C0S2	48-288	47,90	501	870	86.53	1,19E-40		
HA4057	Ha412T4I4412C0S20	157-264	17,37	616	530	99.07	9,42E-45	BP: tRNA aminoacylation for protein translation (GO:0006418)	<i>IPR014729:Rossmann-like alpha/beta/alpha sandwich fold;</i>
		378-464	13,96	616	435	100		BP: arginyl-tRNA aminoacylation (GO:0006420)	
	Ha412T4I4412C0S8	157-264	17,37	616	530	99.07	2,42E-40	BP: translation (GO:0006412)	
		378-464	13,96	616	435	100		CC: cytoplasm (GO:0005737)	
		MF: aminoacyl-tRNA ligase activity (GO:0004812)							
		MF: arginine-tRNA ligase activity (GO:0004814)							
Ha412T4I4412C0S6	157-264	17,37	616	530	99.07	7,68E-40	MF: ATP binding (GO:0005524)		
378-464	13,96	616	435	100		MF: nucleotide binding (GO:0000166)			
HA503	Ha412T4I14873C0S1	18-336	74,65	426	1135	86.65	3,65E-56		
HA911	Ha412T4I92505C0S1	471-574	16,30	632	420	90.38	7,09E-14	CC: intracellular (GO:0005622)	<i>IPR007087:Zinc finger, C2H2;</i>
								MF: zinc ion binding (GO:0008270)	
ORS1242	Ha412T4I30833C0S1	256-460	44,35	460	851	91.22	1,80E-70		
ORS401	Ha412T4I31424C0S1	46-222	79,28	222	675	89.83	5,48E-29		
ORS437	Ha412T4I22228C0S1	116-404	62,20	463	1390	98.28	1,26E-71		<i>IPR001251:Cellular retinaldehyde-binding/triple function, C-terminal; IPR011074:Phosphatidylinositol transfer protein-like, N-terminal;</i>
ORS687	Ha412T4I16005C0S2	90-464	80,60	464	1775	97.33	3,72E-94		
ORS687	Ha412T4I16005C0S1	364-464	21,55	464	485	98.02	4,15E-18		

APÉNDICE

TABLA VII. Resumen de las características asociadas con los marcadores SSRs evaluados en mapas públicos disponibles mediante la herramienta Cmap. La ubicación está dada por un rango para los QTLs y por único valor en caso de genes

Marcador	GL	Mapa de Referencia	Posición del Marcador (cM)	Característica próxima	Ubicación de la Característica
HA77	1	-	-	-	-
ORS662	1	CMSHA89 x ANN1238_F3_Burke et al. 2005	31,3	Contenido de ácido oléico	3,30-25,50
		NDBLOSsel x CM625_F2_Micic et al. 2004	68,5	Número de flores radiales	25,50-41,50
HA4057	3	-	-	Diámetro del tallo	4,80-39,50
HA4105	3	-	-	Tasa de crecimiento fúngico	68
ORS1041	7	Hopi x ANN1238_Wills & Burke 2007	1	Días a floración	0-5,30
				Hojas sobre el tallo principal	0,7,30
HA3878	7	-	-	-	-
HA1848	7	PAC2 x RHA266_RIL_Kiani et al. 2007b	21,6	Potencial de turgencia	14,5
HA911	8	-	-	-	-
HA503	8	-	-	-	-
HA3621	8	-	-	-	-
ORS894_2	8	CMSHA89 x ANN1238_F3_Burke et al. 2002	61,9	Día de floración	49,30-59
		PAC2 x RHA266_RIL_Alfadil et al. 2007	103,1	Tamaño del achenio	50,70-75
HA1108	10	PAC2 x RHA266_RIL_Alfadil et al. 2007	18,2	Resistencia parcial a necrosis basal del tallo ocasionada por Phoma	109,1
				Resistencia parcial a necrosis de las raíces ocasionada por Phoma	113,1
				Resistencia parcial a necrosis del tallo ocasionada por Phoma	6,9
ORS437	10	Hopi x ANN1238_Wills & Burke 2007	17,1	Número de flores radiales	4-22,90
				Diámetro del capítulo	7,80-17,10
				Tamaño de la hoja	7,80-18,90
				Peso del achenio	9,80-18,90
				Número de ramificaciones	9,80-24,80
				Número de capítulos	11,80-18,90
	RHA280 x RHA801_RIL_Tang et al. 2006b	60,3	Peso del pericarpio	40,10-54,70	
			Peso del grano	40,10-54,70	
			Largo de la semilla	40,10-54,70	
			Ancho de la semilla	40,10-54,70	
			Alto de la semilla	40,10-54,70	
			Peso de la semilla	40,10-54,70	
ORS591	10	ANN1295 x PET1277_BC_Lexer et al. 2005	0	Concentración de aceite	40,10-54,70
				Relación de peso grano/pericarpio	40,10-54,70
				Área de la hoja	0
				Concentración de calcio	1
				Diámetro del tallo	3
				Ancho de la hoja	4
ORS691	10	ANN1295 x PET1277_Brouillette et al. 2007	28,8	Concentración de manganeso	5
		ANN1295 x PET1277_BC_Lexer et al. 2005	28,8	Concentración de potasio	5
				Concentración de carbono en la hoja	26-28
				Ancho de la hoja	26

APÉNDICE

**TABLA VII. Resumen de las características asociadas con los marcadores SSRs evaluados en mapas públicos disponibles mediante la herramienta Cmap. La ubicación está dada por un rango para los QTLs y por único valor en caso de genes**

				Concentración de manganeso	26
				Área foliar	27
				Número de lígulas	28
				Concentración de potasio	28
ORS301	14	-	-	-	-
HA3102	15	PAC2 x RHA266_RIL_Alfadil et al. 2007	26,4	Resistencia parcial a Phoma basal del tallo	26,2
ORS401	15	PAC2 x RHA266_RIL_Alfadil et al. 2007	30,8	Resistencia parcial a Phoma basal del tallo	26,2
HA2940	15	-	-	-	-
ORS1242	15	-	-	-	-
ORS499	15	NDBLOSSel x CM625_RIL_Micic et al. 2005a	67,7	Longitud de la hoja	62
ORS687	15	-	-	-	-
HA2191	16	-	-	-	-

**TABLA VIII. Porcentaje de la varianza explicada por todos los términos incluidos en el modelo de ajuste de medias.**

CAMPAÑA	TÉRMINO DEL MODELO	IEMAX			SEV21			INT 21			PI		
		VAR	%VAR	DE	VAR	%VAR	DE	VAR	%VAR	DE	VAR	%VAR	DE
2009-10	RIL:FECHA.INOC	0	0	0	1,90E-03	16,35	0,04	8,10E-04	27,71	0,03	2,6	10,21	1,61
	RIL	8,49	45,61	2,91	2,30E-03	20,4	0,05	5,50E-05	1,86	0,01	3,51	13,74	1,87
	FECHA.INOC	0,82	4,43	0,91	0	0	0	1,40E-03	49,02	0,04	9,61	37,66	3,1
	REP	0	0	1,80E-07	4,30E-04	3,73	0,02	2,10E-05	0,72	4,60E-03	2,03	7,94	1,42
	Residual	9,31	49,97	3,05	0,01	59,52	0,08	6,10E-04	20,69	0,02	7,76	30,44	2,79
2011-12	RIL:FECHA.INOC	0,42	11,82	0,65	0	0	0	5,00E-04	13,35	0,02	0	0	2,40E-08
	RIL	0,26	7,34	0,51	0,04	43,36	0,19	1,10E-04	2,87	0,01	6,01	45,48	2,45
	FECHA.INOC	1,83	51,83	1,35	0,01	6,51	0,07	3,60E-04	9,65	0,02	0	0	0
	REP	0	0	0	6,00E-05	0,07	0,01	7,80E-05	2,09	0,01	0,13	0,97	0,36
	Residual	1,02	29,01	1,01	0,04	50,06	0,2	2,70E-03	72,04	0,05	7,08	53,55	2,66
2012-13	RIL:FECHA.INOC	5,54	19,87	2,35	0,01	23,04	0,1	0,01	22,18	0,09	0	0	4,00E-08
	RIL	9,76	34,98	3,12	0,01	20,49	0,1	0,01	26,49	0,1	2,98	37,94	1,73
	FECHA.INOC	1,5	5,37	1,22	1,50E-03	3,37	0,04	3,50E-03	9,5	0,06	0,41	5,22	0,64
	REP	0,03	0,1	0,17	1,30E-03	2,77	0,04	1,70E-04	0,45	0,01	0,18	2,29	0,42
	Residual	11,07	39,68	3,33	0,02	50,34	0,15	0,02	41,38	0,12	4,29	54,54	2,07
2013-14	RIL:FECHA.INOC	1,22	7,32	1,11	0	0	1,30E-09	1,10E-04	0,59	0,01	0	0	1,00E-07
	RIL	5,26	31,49	2,29	0,02	50,52	0,13	0,01	59,9	0,1	4,22	52,45	2,05
	FECHA.INOC	3,45	20,67	1,86	6,50E-04	2,08	0,03	1,90E-03	10,74	0,04	0,07	0,92	0,27
	REP	0,19	1,12	0,43	0	0	0	0	0	0	0	0	0
	Residual	6,58	39,41	2,56	0,01	47,4	0,12	0,01	28,77	0,07	3,75	46,63	1,94