



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



**CSIC**  
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



**IBMCP**  
Instituto de Biología Molecular y Celular de Plantas

*Desarrollo de una genoteca de líneas de introgresión entre  
*Solanum lycopersicum* y *Solanum pimpinellifolium*  
utilizando herramientas genómicas de alto  
rendimiento y detección de QTLs  
implicados en calidad de fruto*

---

*Tesis presentada por:*

**Walter Barrantes Santamaría**

**Para optar al título de Doctor en Biotecnología**

**Directores: Dr. Antonio José Monforte, Dr. Antonio Granell**

**2014**

El **Dr. Antonio José Monforte Gilabert**, científico titular del Consejo Superior de Investigaciones Científicas y el **Dr. Antonio Granell Richart**, Profesor de investigación del Consejo Superior de Investigaciones Científicas, ambos pertenecientes al Instituto de Biología Molecular y Celular de Plantas.

**CERTIFICAN** que el Ingeniero Agrónomo Walter Barrantes Santamaría ha realizado bajo su dirección en el Instituto de Biología Molecular y Celular de Plantas el trabajo que lleva como título “ *Desarrollo de una genoteca de líneas de introgresión entre *Solanum lycopersicum* y *Solanum pimpinellifolium* utilizando herramientas de genómicas de alto rendimiento, y detección de QTLs implicados en calidad de fruto*”, y autorizan su presentación para optar al grado de Doctor en Biotecnología por la Universidad Politécnica de Valencia.

Y para que conste, expiden y firman el presente certificado en Valencia, a 23 de Mayo del 2014.

**Dr. Antonio José Monforte Gilabert**

**Dr. Antonio Granell Richart**

## **Agradecimientos**

Durante mi estancia en Valencia han sido muchas las personas con las que de una u otra forma he compartido momentos especiales, ya sean estos de trabajo o solaz esparcimiento. Todos estos momentos los recordaré con especial cariño, pero especialmente los segundos, que hicieron que mi estancia y la de mi familia fuese mucho más agradable.

Quiero empezar mis agradecimientos con las dos personas cuya participación en el desarrollo de esta tesis fue trascendental: los Drs. Antonio Monforte y Antonio Granell, quienes aparte de su acertada guía durante todas y cada una de las etapas de mi doctorado también, me brindaron su amistad y la de sus estimables familias. A Toni G. gracias por darme la oportunidad de venir al IBMCP y poder vivir esta experiencia por demás maravillosa. Y especialmente, quiero agradecer a Tony M. la paciencia, comprensión y apoyo en todo momento y decirle que esta tesis no habría sido posible sin su irremplazable participación.

Esta tesis fue auspiciada por la Universidad de Costa Rica, quiero agradecer a los funcionarios que en su momento confiaron en mí para que realizara este Doctorado, a los compañeros de la Estación Agrícola Fabio Baudrit, y de la Facultad de Agronomía por su voto de confianza. Agradezco a la Oficina de Asuntos Internacionales (OAICE), en especial a la Dra. Ana Sittenfeld, Yamileth Damazio y Fátima Acosta por el apoyo y comprensión que recibimos mi familia y yo durante nuestra "aventura" por tierras chilenas. Gracias también a la Dra. Julieta Carranza y a Laura Agüero y a todo el personal de OAICE por estar siempre pendiente de nosotros en Valencia.

Al CSIC por cofinanciar parte de mi beca y por brindarme la oportunidad de recibir parte de mi capacitación en uno de sus prestigiosos centros de investigación.

Un especial agradecimiento al Dr. Rafael Fernández, responsable en gran medida de la realización de esta tesis, por su ayuda en el desarrollo de la genoteca y, por estar siempre dispuesto a responder ante mis consultas, algunas veces, con mucho más detalle del esperado. Al Dr. Carlos Romero por su oportuna ayuda en la interpretación de los resultados del capítulo 3, a Asun Fernández que gracias a su guía pude iniciar con el trabajo de genotipado de la plantas, al Dr. José María Bellés por sus consejos y ayuda a lo largo de todo este tiempo.

Igualmente quiero agradecer a todas las personas que de una u otra forma colaboraron en el diseño, instalación, mantenimiento y evaluación de los ensayos de campo en Alginet, Orihuela y Málaga. En Alginet un agradecimiento especial al personal de los invernaderos de la Cooperativa Agrícola (Coagri), también a Teresa León, Erika, Silvia, Teresa Caballero y Soledad Casal por su colaboración en la evaluación de los frutos. En Orihuela un agradecimiento muy especial a Juan José Ruiz, Santiago García-Martínez, Aranzazu Alonso, Fernando Rubio, Adrián Grau, José Joaquín García, Javier Vives y Alberto Lara por toda la colaboración brindada. En Málaga a Gloria López Casado, Antonia Núñez Martín, Luis Rodríguez Caso, Rafael Gómez Cabrera y personal de campo, a todos ellos muchas gracias ya que sin su ayuda esta tesis no se hubiera podido realizar.

También quiero agradecer a mis profesores del COMAV por sus enseñanzas, a los compañeros del IBMCP que de una u otra forma me han ayudado. En especial al Dr. Eugenio Grau del servicio de secuenciación por su apoyo y amistad. Al personal de los invernaderos del IBMCP-UPV especialmente a "Rafa" por su imprescindible ayuda durante el ensayo de campo en Alginet, a María Victoria, Ferrán y David por todo el apoyo logístico recibido en el invernadero a lo largo de estos cuatro años.

A mis compañeros del laboratorio 0.08 con los cuales compartí largas horas de trabajo y uno que otro momento de diversión, Gerardo, Aurora D, Aurora M, Mohamed, Gabriela. En especial a especial a José Añó por su valiosa ayuda en la cosecha de frutos y extracción de semillas y, a Soledad Casal por toda la ayuda brindada durante mi paso por el laboratorio. A todos ellos muchas gracias porque han hecho mi estancia más agradable.

Finalmente quiero agradecer a mi familia y amigos, en especial a mi esposa Kristel y mi hija Ximena porque sin su apoyo no lo hubiera logrado. A mi madre, hermanos, cuñado, sobrinos y primos que realizaron el esfuerzo de sacar parte de su tiempo para venir a compartirlo con nosotros a Valencia. Muchas gracias a todos, este logro también es de ustedes.

## INDICE

<b>RESUMEN</b> .....	<b>11</b>
<b>ABSTRACT</b> .....	<b>13</b>
<b>RESUM</b> .....	<b>15</b>
<b>CAPITULO 1: Introducción General</b> .....	<b>17</b>
1.1 Descripción botánica del tomate.....	17
1.2 Taxonomía.....	17
1.3 Importancia económica del tomate.....	18
1.4 Centro de origen e historia del tomate .....	19
1.5 Recursos fitogenéticos .....	20
1.5.1 Bancos de Germoplasma.....	20
1.5.2 Especies silvestres como fuentes de variabilidad genética aprovechable en mejora.....	20
1.5.3 Sistemas de cruzamiento en tomate.....	22
1.5.4 <i>S. pimpinellifolium</i> como parental donante de genes.....	23
1.6 Uso de Marcadores genéticos para conocer la variabilidad natural de tomate .....	24
1.7 Métodos de genotipado de SNPs.....	26
1.7.1 Métodos de genotipado de SNPs simples .....	26
1.7.2 Genotipado de SNPs de alto rendimiento .....	29
1.7.3 Genotipado basado en secuenciación masiva.....	30
1.8 Genómica del tomate .....	31
1.9 Mapas genéticos del tomate .....	32
1.10 Mapeo de Quantitative Trait Locus .....	34
1.11 Poblaciones para mapeo de QTLs en tomate .....	35
1.12 Usos y ventajas de las ILs en la detección de QTLs.....	37
1.13 QTLs definidos en tomate y regiones cromosómicas relacionadas.....	38
1.14 QTLs para caracteres calidad de fruto en tomate .....	39
1.14.1 QTLs para caracteres de peso, tamaño y forma del fruto.....	39
1.14.2 QTLs para Sólidos totales .....	40
1.14.3 QTLs para acidez y pH .....	41
1.14.4 QTLs para color de fruto.....	41
1.14.5 QTLs para contenido nutricional y compuestos volátiles .....	42
1.14.6 QTLs para vida pos-cosecha .....	43
<b>OBJETIVO GENERAL</b> .....	<b>45</b>
Objetivos específicos.....	45
<b>CAPITULO 2: Highly Efficient Genomics-Assisted Development of a Library of Introgression Lines of <i>Solanum pimpinellifolium</i></b> .....	<b>47</b>
Introduction.....	47
Materials and methods.....	50
1. Plant materials.....	50
2. Genomic characterization of TO-937 .....	51
3. Breeding scheme .....	51

4. Selection differential .....	52
5. Genotyping methods.....	53
5.1 DNA extraction .....	53
5.2 Genotyping.....	53
Results and discussion.....	54
1. Genomic characterization of TO-937 .....	54
2. Development and selection in early backcross populations.....	54
3. Developing HRM assays for selected SNPs from GoldenGate Illumina genotyping array.....	56
4. Development of single-introgression lines .....	56
5. Definitive IL collection characterized by HRM assays .....	57
6. Characterization of the IL collection by high-throughput SNP genotyping .....	58
7. The SP introgression line genomic library as a tool for improving genome assembly .....	62
8. Integration of high-throughput genotyping strategies in IL development .....	62

### **CAPÍTULO 3: Mapeo fino de una región con segregación distorsionada en el extremo distal del brazo largo del cromosoma 4..... 65**

Introducción .....	65
1. Segregación distorsionada (SD).....	66
2. Causas de la SD: bases fisiológicas, estructurales o moleculares .....	66
3. SD en tomate.....	67
4. Causas de la SD en tomate.....	68
Materiales y Métodos .....	70
1. Material vegetal y extracción de ADN .....	70
2. Definición y mapeo de la región con SD .....	70
3. Variabilidad alélica en la región con SD en especies relacionadas de <i>Solanum</i> .....	72
4. Identificación y diseño de marcadores moleculares (SNPs) para las diferentes etapas de mapeo de la región SD.....	72
Resultados .....	73
1. Mapeo inicial de la región con SD .....	73
2. Mapeo fino de la región con SD.....	73
2.1 Desarrollo de nuevos marcadores.....	74
2.2 Generación e identificación de recombinantes.....	77
2.3 Pruebas de progenie a los recombinantes seleccionados .....	79
3. Análisis de la variabilidad en la región con SD en otros silvestres .....	82
4. Genes anotados en la región con SD .....	82
Discusión .....	86

### **CAPITULO 4: Identificación de QTLs implicados en caracteres de calidad de fruto de tomate en la genoteca de líneas de introgresión de TO-937..... 91**

Introducción .....	91
Materiales y Métodos .....	93
1. Material vegetal.....	93
2. Localidades evaluadas y diseño del experimento.....	93
3. Caracteres evaluados .....	94
4. Análisis estadístico .....	96
Resultados.....	97
1. Caracterización del fruto del parental recurrente MM y su variabilidad en las tres localidades.....	97

2. Evaluación de ILs y mapeo de QTLs .....	98
2.1 Caracteres descriptivos .....	98
2.1.1 Precocidad (PREC) .....	98
2.1.2 Vigor (VIG) .....	98
2.1.3 Hombros verdes (HV) .....	98
2.1.4 Densidad de tricomas (TRIC).....	98
2.1.5 Estructura del dosel (DOS).....	100
2.1.6 Color violáceo en ramas y hojas (VIOL) .....	100
2.2 Caracteres cuantitativos de calidad de fruto.....	100
2.2.1 Caracteres relacionados con las dimensiones del fruto .....	101
2.2.1.1 Peso del fruto (FW) .....	102
2.2.1.2 Perímetro (PER) y Área (AR) .....	103
2.2.1.3 Diámetro máximo del fruto (FDL).....	103
2.2.1.4 Longitud máxima del fruto (FLL).....	106
2.2.2 Caracteres de forma del fruto .....	107
2.2.2.1 Índice de forma de fruto (FS).....	107
2.2.2.2 Bloque distal (DBLK), Bloque proximal (PBLK) e Índice de forma de fruto triangular (TRI) .....	109
2.2.2.3 Forma de fruto circular (CIR) .....	109
2.2.2.4 Altura de hombro en el fruto (PSH) .....	109
2.2.3 Caracteres asociados a medidas analíticas relacionadas con el sabor .....	110
2.2.3.1 Contenido en Sólidos Solubles (SSC) .....	110
2.2.3.2 Acidez inicial (pH) .....	112
2.2.3.3 Acidez valorable (TA).....	112
2.2.4 Color externo (EC).....	114
2.2.5 Color interno (IC) .....	117
Discusión .....	120
1. Control genético de los caracteres de calidad en la genoteca de ILs .....	120
2. Disección genética de caracteres de calidad con la genoteca de ILs de TO-937 .....	121
<b>CAPITULO 5: Discusión General.....</b>	<b>127</b>
<b>CONCLUSIONES.....</b>	<b>137</b>
<b>REFERENCIAS.....</b>	<b>139</b>
<b>ANEXOS .....</b>	<b>159</b>
Anexos (CAPÍTULO 2) .....	161
Anexos (CAPÍTULO 3) .....	171
Anexos (CAPÍTULO 4) .....	179



## RESUMEN

Las especies silvestres de tomate contienen una gran cantidad de variabilidad genética que no está siendo aprovechada por los mejoradores debido a las dificultades de trabajar con cruzamientos interespecíficos. Una de las estrategias adecuadas para el manejo de variabilidad silvestre en cultivares modernos es el desarrollo de líneas de introgresión (ILs) que contengan una única introgresión de un genotipo silvestre o exótico en el fondo genético de un cultivar élite. El desarrollo de estas líneas es un trabajo tedioso, lo que explica que el número de colecciones de ILs disponibles sea limitado. En esta tesis se describe el desarrollo de una genoteca de ILs que incorpora la variabilidad genética de la accesión TO-937 de *Solanum pimpinellifolium* (especie silvestre) en el fondo genético del cultivar "Moneymaker" de *Solanum lycopersicum* (parental recurrente). Para aumentar la eficiencia en el desarrollo de la genoteca de ILs se incluyeron varias rondas de genotipado masivo de alto rendimiento en etapas tempranas del plan de mejora que permitió una selección adecuada de las plantas en cada generación teniendo en cuenta: la composición del genoma del parental donante y recurrente, el número de introgresiones y una cobertura del parental donante adecuada. Esto aumentó la eficiencia del proceso, permitiendo reducir el número de generaciones necesarias para obtener la genoteca completa de ILs. La población final, compuesta por 53 ILs, fue caracterizada con una micromatriz de 8 K SNPs desarrollada por la iniciativa SOLCAPs. Así, la genoteca de ILs tiene cobertura del 94% del parental donante (debido a la pérdida de dos regiones proximales en los cromosomas 1 y 8). Cada IL contiene una introgresión homocigota del parental donante con un tamaño promedio de 25 Mb (rango 0.7 – 75Mb), algunas ILs tienen introgresiones adicionales (con un promedio de 0.7 introgresiones extras por IL) en general de tamaño pequeño (rango 0.1 – 60 Mb), localizándose las de mayor tamaño en regiones centroméricas. La resolución genética de la genoteca viene definida por 71 bins con un tamaño promedio de 10 Mb (rango 0.9 – 65 Mb).

Durante el proceso de desarrollo de la genoteca se identificó una región en el cromosoma 4 que mostraba segregación distorsionada en las generaciones intermedias favoreciendo los alelos de TO-937 en detrimento de los alelos de "Moneymaker". Dado el tipo de segregación observada y su localización cromosómica, esta distorsión debe ser producida por el locus "Gamete Eliminator (*Ge*), descrito previamente en *S. pimpinellifolium*. Por medio de un mapeo fino se logró acotar la región a 84 Kb, en esta región se encuentran anotados 10 genes (algunos con función desconocida) y un retrotransposon. Estos resultados nos facilitará el poder determinar en un futuro el gen o genes responsables de la distorsión.

La genoteca de ILs fue caracterizada agronómicamente en tres localidades (Alginet, Orihuela y Málaga) para caracteres relacionados con desarrollo de la planta y, fundamentalmente, calidad de fruto. Se definieron un total de 75 Quantitative Trait Locus (QTLs) con efectos consistentes en al menos dos localidades, de los cuales 8 QTLs correspondieron a caracteres descriptivos (precocidad, vigor, frutos con hombros verdes, pilosidad en tallo y hojas, estructura del dosel y coloración violácea en tallos y ramas) y 67 QTLs a caracteres relacionados con calidad de fruto: caracteres relacionados con las dimensiones de fruto, forma de fruto, caracteres organolépticos, color externo e interno del fruto. El efecto de la mayoría de QTLs definidos estuvo en concordancia con lo esperado según el fenotipo del parental donante, sin embargo, se encontraron alelos en TO-937 que presentaron segregación transgresiva en caracteres de peso del fruto, forma, grados brix y color. Adicionalmente se demostró la existencia de variabilidad genética novedosa en TO-937 al definirse 15 QTLs de interés, que no habían sido descritos antes en otras poblaciones desarrolladas a partir de otras entradas de *S. pimpinellifolium* o incluso de otras especies. Estos resultados demuestran que la genoteca desarrollada en esta tesis es una poderosa herramienta para el estudio de diversos aspectos de la biología del tomate y también para su utilización en planes de mejora

## ABSTRACT

Wild species of tomato contain a large amount of genetic variability that is not being exploited by breeders due to the difficulties of working with interspecific crosses. One of the appropriate strategies for the management of wild variability in modern cultivars is developing introgression lines (ILs) containing a single introgression of a wild or exotic genotype into the genetic background of an elite cultivar. The IL development is a very tedious task, which explains why the number of available IL collections is still limited. One of the major objectives of this thesis, is the development of a genomic library of ILs that incorporates variability from the accession TO-937 of *Solanum pimpinellifolium* (wild species) in the genetic background of cultivar "Moneymaker" *Solanum lycopersicum* (recurrent parent). To increase the efficiency in the development of the IL library several rounds of high-throughput genotyping was performed in the early stages of the breeding scheme which allowed an appropriate selection of plants taking into account: the donor and recurrent parent genome composition, the number of introgressions and adequate coverage of donor genome. This increased the efficiency of the process, allowing to reduce the number of generations needed to obtain the full IL library. The final population comprised 53 ILs that were characterized with a microarray of 8 K SNPs developed by the SOLCAPs initiative. Thus, the IL genomic library has 94% coverage of the donor parent (due to the loss of two proximal regions on chromosome 1 and 8), each IL containing a homozygous introgression of the donor parent with an average size of 25 Mb (range 0.7 - 75 Mb). Some ILs have additional introgressions (with an average of 0.7 extra introgression per IL) generally of small size (range 0.1 - 60 Mb). The larger introgressions were located in centromeric regions. The genetic resolution of the genomic library is defined by 71 bins with an average size of 10 Mb (range 0.9 - 65 Mb).

During the process of IL development a region on chromosome 4 showing distorted segregation favoring the alleles of TO-937 in detriment of "Moneymaker" alleles was identified. Given the kind of observed segregation and its chromosomal location, this distortion must be produced by the locus "Gamete Eliminator (*Ge*), previously described in *S. pimpinellifolium*. Through a fine mapping, the gene was delimited to a region of 84 Kb at the distal part of chromosome 4. Ten annotated genes (some with unknown function) and a retrotransposon were found in this region. These results provide the basis to determine in the future the gene (s) responsible for the distortion.

The Genomic library of ILs was agronomically characterized in three locations (Alginet, Orihuela and Málaga) for characters related to plant development and, mainly, fruit quality. A total of 75 QTLs with consistent effect in at least two locations were defined. Eight QTLs were involved in descriptive characters (earliness, vigor, fruit with green shoulders, hairiness in stem and leaves, structure of canopy and purplish coloration of stems and branches) and 67 QTLs were related to fruit quality t characters: dimensions of fruit, fruit shape, organoleptic characteristics, external and internal fruit color. The allelic effects of most QTLs were in accordance with the phenotype of the parental donor, although, in some cases, alleles from TO-937 induced transgressive segregation in characters such as fruit weight, fruit shape soluble solids content and fruit color. In addition, TO-937 showed unique genetic variability as 15 of the defined QTLs had not been described before in other populations developed from *S. pimpinellifolium*. These results demonstrate that the current IL library developed in this thesis is a powerful tool for the study of various aspects of tomato biology and also for use in breeding programs.

## RESUM

Les espècies silvestres de tomaca contenen una gran quantitat de variabilitat genètica que no està sent aprofitada pels milloradors degut a les dificultats de treballar amb creuaments interespecífics. Una de les estratègies adequades per al maneig de la variabilitat silvestre en cultivars moderns és el desenvolupament de línies d'introgresió (IL) que continguin una única introgresió d'un genotip silvestre o exòtic en el fons genètic d'un cultivar d'èlit. El desenvolupament d'aquestes línies és un treball tediós, la qual cosa explica que el nombre de col·leccions de ILs disponibles siga limitat. "En esta tesi es planteja el desenvolupar una genoteca de ILs que incorpore variabilitat la genètica de l'accessió TO-937 de *Solanum pimpinellifolium* (espècie silvestre) en el fons genètic del cultivar "Moneymaker" *Solanum lycopersicum* (parental recurrent)." "Per tal d'augmentar l'eficiència en el desenvolupament de la genoteca de ILs es van incloure diverses rondes de genotipat massiu d'alt rendiment en etapes primerenques del pla de millora, el que va permetre seleccionar adequadament les plantes més adients en cada generació tenint en compte: la composició del genoma del parental donant i recurrent, el nombre d'introgresions i l'obtenir una cobertura del parental donant adequada. Açò va augmentar l'eficiència del procés, permetent reduir el nombre de generacions necessàries per a obtenir la genoteca completa de ILs. La població final està composta per 53 ILs i va ser caracteritzada amb una micromatriu de 8 K SNPs desenvolupada per la iniciativa SOLCAPs. Així, la genoteca de ILs té cobertura del 94% del parental donant (a causa de la pèrdua de dos regions proximals en el cromosoma 1 i 8), cada IL conté una introgresió homozigota del parental donant amb una grandària mitjana de 25 Mb (rang 0.7 - 75Mb), algunes ILs tenen introgresions addicionals (amb una mitjana de 0.7 introgresions de més per IL) en general de grandària xicoteta (rang 0.1 - 60 Mb), i localitzant-se les més gran en regions centromèriques. La resolució genètica de la genoteca és definida per 71 bins amb una grandària mitjana de 10 Mb (rang 0.9 - 65 Mb).

Durant el procés de desenrotllament de la genoteca es va identificar una regió en el cromosoma 4 que mostrava segregació distorsionada en les generacions intermèdies tot i afavorint els al·lels de TO-937 en detriment dels de "Moneymaker". Donat el tipus de segregació observada i la seua localització cromosòmica, esta distorsió ha de ser produïda pel locus "Gamete Eliminator (Ge)", descrit prèviament en *S. pimpinellifolium*. Per mitjà d'un mapeig fi es va aconseguir fitar la regió a 84 Kb, on es troben anotats 10 gens (alguns amb funció desconeguda) i un retrotransposó. Aquests resultats proporcionen les bases per a determinar en un futur el gen (s) responsable de la distorsió.

La genoteca de ILs va ser caracteritzada agronòmicament en tres localitats (Alginet, Orihuela i Màlaga) per a caràcters relacionats amb el desenvolupament de la planta i, fonamentalment, en la qualitat de fruit. Es van definir un total de 75 QTLs amb efectes consistents en almenys dos localitats, dels quals 8 QTLs van correspondre a caràcters descriptius (precocitat, vigor, fruits amb muscles verds, pilositat en tija i fulls, estructura del dosser i coloració violàcia en tiges i branques)" i 67 QTLs a caràcters relacionats amb qualitat de fruit: caràcters relacionats amb les dimensions de fruit, forma de fruit, caràcters organolèptics, color extern i intern del fruit. L'efecte de la majoria de QTLs definits va estar en concordança amb allò que permet predir el fenotip del parental donant, no obstant, es van trobar al·lels en TO-937 que van presentar segregació transgressiva per a caràcters de pes del fruit, forma, graus brix i color. Addicionalment es va demostrar la variabilitat genètica nova que es troba a TO-937 en definir-se 15 QTLs que no havien sigut descrits abans en altres poblacions desenvolupades a partir de altres accions de *S. pimpinellifolium* inclòs d'altres espècies relacionades. Estos resultats demostren que la genoteca desenvolupada en el marc d'aquesta tesi és una ferramenta poderosa per a l'estudi de diversos aspectes de la biologia de la tomaca i també per a la seua utilització en plans de millora."

## **CAPITULO 1: Introducción General**

### **1.1 Descripción botánica del tomate**

El tomate es una especie dicotiledónea, de porte arbustivo, aunque puede desarrollarse en forma rastrera, semirrecta o erecta. Según su hábito de crecimiento las plantas de tomate se clasifican en determinadas o indeterminadas, con órganos aéreos que están cubiertos de tricomas, los cuales pueden ser glandulares o no glandulares. Los tallos son gruesos y angulosos, de color verde, y la planta presenta tallos secundarios que salen de las axilas de las hojas. El sistema radicular está compuesto por una raíz principal corta y débil, y raíces secundarias que son numerosas y potentes. Las hojas son compuestas anchas, planas con 7-11 folíolos peciolados, lobulados y borde dentado. Las flores son hermafroditas, simétricas, regulares e hipóginas. La corola está formada por 5 o más pétalos de color amarillo, lanceolados y fusionados en la base. Los sépalos son más pequeños que los pétalos. La flor contiene cinco estambres fusionados a la corola por sus filamentos, las anteras largas de color amarillo, están unidas lateralmente, formando un cono estaminal que envuelve el gineceo. El pistilo está formado por un ovario bi- o plurilocular, de estilo delgado y el estigma cubierto por una secreción de naturaleza lipídica. El fruto es una baya con gran diversidad de tamaños y formas, generalmente de color rojo, bilocular o multilocular, de gruesa placenta con numerosas semillas recubiertas de una sustancia mucilaginosa (Peralta et al. 2008).

### **1.2 Taxonomía**

El tomate cultivado pertenece a la familia de las solanáceas, la cual comprende 98 géneros y cerca de 2800 especies que crecen en una gran diversidad de hábitats, desde zonas áridas hasta la alta montaña. Esto ha contribuido, en buena medida, a la importante variabilidad genética presente entre las especies de este grupo (Olmstead y Bohs 2007). La familia Solanáceas, además del tomate, incluye otras especies de interés agronómico como el pimiento, la patata, la berenjena y el tabaco, constituyéndose por lo tanto en una de las familias de mayor importancia alimenticia y económica del mundo.

La taxonomía del tomate ha sufrido cambios a través del tiempo, originalmente el tomate fue clasificado como *Solanum lycopersicum* por Linneo, después fue reclasificado por Miller (1754) como *Lycopersicon esculentum*. En la actualidad el tomate se clasifica como *Solanum lycopersicum* L. (Child 1990; Peralta et al. 2008) (**Tabla1-1**). La taxonomía aceptada para esta especie es la siguiente (Esquinas-Alcázar y Nuez 1995).

Clase: Dicotiledóneas  
 Orden: Solanales  
 Suborden: Solanineae  
 Familia: Solanaceae  
 Subfamilia: Solanoideae  
 Tribu: Solaneae  
 Género: *Solanum* secc. *Lycopersicon*

Nombre dentro del género <i>Solanum</i>	Equivalente del género <i>Lycopersicon</i>
<i>S. lycopersicoides</i> Dunal	<i>L. lycopersicoides</i> Dunal A. Child ex J.M.h.Shaw
<i>S. sitiens</i> I.M. Johnst	<i>L. sitiens</i> I.M. Johnst
<i>S. juglandifolium</i> Dunal	<i>L. juglandifolium</i> Dunal
<i>S. ochranthum</i> Dunal	<i>L. ochranthum</i> Dunal
<i>S. pennellii</i> Correll	<i>L. pennellii</i> Correll D'Arcy
<i>S. habrochaites</i> S.Knapp D.M Spooner	<i>L. hirsutum</i> Dunal
<i>S. chilense</i> Dunal Reiche	<i>L. chilense</i> Dunal
<i>S. huaylasense</i> Peralta	Parte de <i>L. peruvianum</i> L. Miller
<i>S. peruvianum</i> L.	<i>L. peruvianum</i> L. Miller
<i>S. orneliomuelleri</i> J.F.Macbr.L.	Parte de <i>L. peruvianum</i> L. Miller
<i>S. arcanum</i> Peralta	Parte de <i>L. peruvianum</i> L. Miller
<i>S. chmielewskii</i> C.M. Rick, Kesicki, Fobes & M. Holle D.M. Spooner, G.J. Anderson & R.K. Jansen	<i>L. chmielewskii</i> C.M. Rick, Kesicki, Fobes & M. Holle
<i>S. neorickii</i> C.M. Rick, Kesicki, Fobes & M.Holle D.M. Spooner, G.J. Anderson & R.K. Jansen	<i>L. parviflorum</i> C.M. Rick, Kesicki, Fobes & M. Holle
<i>S. pimpinellifolium</i> L.	<i>L. pimpinellifolium</i> L. Miller
<i>S. lycopersicum</i> L.	<i>L. esculentum</i> Miller
<i>S. cheesmaniae</i> L. Riley Fosberg	<i>L. cheesmaniae</i> L. Riley
<i>S. galapagense</i> S. Darwin & Peralta	Parte de <i>L. cheesmaniae</i> L. Riley

**Tabla 1-1 Lista de especies del género *Solanum*, sección *Lycopersicoides*, sección *Juglandifolia* y sección *Lycopersicon*. Se proporciona la correspondencia con la nomenclatura previa. (Tomado de Peralta et al. 2008).**

### 1.3 Importancia económica del tomate

El tomate es uno de los más importantes vegetales en el mundo, con una producción anual en aumento, siendo en el 2011 mayor a las 159 millones de toneladas. China es el mayor productor seguido por India y Estados Unidos. España ocupa el decimo puesto mundial con una producción en el 2011 de 3.8 millones de toneladas. En cuanto al volumen de las exportaciones, España se ubica en el tercer lugar a nivel mundial, exportando tanto tomate fresco como en conserva (FAOSTAT:<http://faostat.fao.org/site/>).

El tomate ha experimentado una gran diversificación varietal después de su domesticación, debido en parte a la selección condicionada a las diferentes formas de consumirlo (directamente crudo, en pastas, salsas, zumos, sopas, enlatados pelados enteros o picados) y a la adaptación a los diferentes sistemas de cultivo que van desde cultivo en campo en países con clima tropical y templado hasta sistemas con ambiente controlado (invernaderos) en países con clima frío.

#### **1.4 Centro de origen e historia del tomate**

El origen de las especies de tomate se localiza en la región andina comprendida entre el sur de Colombia y el norte de Chile, incluyendo las Islas Galápagos. Según Esquinas-Alcazar y Nuez (1995) su domesticación ocurrió en México a partir de *S. lycopersicum*, y los primeros cultivares se crearon a partir de la variante “cerasiforme”, que crece de manera silvestre en regiones tropicales y subtropicales de América. Recientemente, Blanca et al. (2012) indican que hubo un proceso de pre domesticación que ocurrió en la región andina, que se completo en Mesoamérica antes de la llegada de los españoles.

Con la conquista de América, el tomate fue llevado a Europa y otras partes del mundo a inicios del siglo XVI (Peralta y Spooner 2005; Paran et al. 2007). La aceptación en el viejo mundo fue desigual; mientras que en España, Portugal e Italia pasó a formar parte de la gastronomía popular, en el resto de Europa fue utilizado solo como planta ornamental, al ser su fruto considerado como venenoso, hasta alrededor del año 1840 que fue catalogado como vegetal nutricional (Gould 1983).

Algunas de los caracteres más importantes que fueron seleccionados durante el proceso de domesticación y mejora varietal del tomate fueron la apariencia y calidad del fruto, así como una arquitectura de planta que favoreciera su cosecha (Paran et al. 2007). La calidad del fruto en el tomate cultivado es una característica apreciada por los productores a la hora de elegir un cultivar. Hoy en día los mercados están interesados en diferenciar los productos por atributos relacionados con caracteres organolépticos (Causse et al. 2002; Powell et al. 2012). Actualmente, una parte importante de la mejora genética del tomate está dirigida hacia caracteres de calidad de fruto. El mayor interés tanto en fruto fresco como para industrias, incluye características relacionadas con tamaño y forma del fruto, cantidad de sólidos totales, acidez, color, firmeza, textura, contenido nutricional, sabor, maduración entre otros.

## **1.5 Recursos fitogenéticos**

El tomate cultivado posee una variabilidad genética limitada, en gran parte debida a varios cuellos de botella entre ellos las selecciones naturales y artificiales producidas durante la domesticación y desarrollo de los cultivares modernos (Ladizinsky 1985), y especialmente por el cuello de botella que se produjo al importarlo a Europa. Así, el reducido número de tomates introducidos por los españoles tras el descubrimiento constituye la base genética de todos los cultivares modernos. Se estima que solo un 5% de la variabilidad total del género *Solanum* secc. *Lycopersicon* está presente en *S. lycopersicum* (Miller y Tanksley 1990).

### **1.5.1 Bancos de Germoplasma**

Se estima que en bancos de germoplasma en todo el mundo existen más de 62.800 accesiones entre silvestres y cultivadas. Entre los bancos más importantes de Solanaceae podemos citar el Asian Vegetable Research and Development Center (AVRDC) en Tainan, Taiwán, el United States Department of Agriculture (USDA), Plant Genetic Resources Unit de Geneva (PGRU), NY, USA, y el CM Rick Tomato Genetics Resource Center (TGRC), en la Universidad de California en Davis, USA. El TGRC (<http://tgrc.ucdavis.edu>) contiene la colección más grande de tomate silvestre del mundo.

También se mantienen colecciones de germoplasma de tomate en Holanda (IVT), Rusia (VIR), Japón (NIAS), Perú (DHUNA) y Cuba (INIFAT). Poblaciones de mutantes fenotípicamente caracterizadas y catalogadas se pueden rastrear en los sitios web: <http://tgrc.ucdavis.edu>; <http://www.zamir.sgn.cornell.edu>. El Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) conserva 4513 accesiones de tomate y especies relacionadas (<http://www.comav.upv.es>).

### **1.5.2 Especies silvestres como fuentes de variabilidad genética aprovechable en mejora**

Como se ha mencionado anteriormente, la base genética de las variedades comerciales de tomate es estrecha, de ahí que los programas modernos de mejora genética se planteen como objetivo volver la mirada hacia los parientes silvestres, y tratar de aumentar la diversidad genética disponible para los mejoradores, identificando caracteres de interés agronómico dentro del germoplasma exótico (Tanksley et al. 1996; Fulton et al. 2002; Frary et al. 2004; Labate et al. 2007). Un ejemplo lo constituye el aumento del rendimiento, con ejemplos de éxito en arroz (Septiningshi et al. 2003), trigo (Huang et al. 2003), avena (Pillen et al. 2003),

soya (Concibido et al. 2003), garbanzo (Singh y Ocampo 1997), tomate (Bernacchi et al. 1998; Gur and Zamir 2004) pimiento (Rao et al. 2003).

A pesar de que las especies silvestres de tomate actualmente son una fuente de genes de interés agrícola para programas de mejora, aún están muy poco aprovechadas. Para ello hay que identificar primero fuentes de variabilidad en el carácter de interés, y después desarrollar un programa para introducir esa variabilidad en un programa de mejora (Francia et al. 2005). En la **Tabla 1-2** se presenta un resumen de las principales características de las especies silvestres de tomate, con las características de interés que han sido o podrían ser consideradas en eventuales programas de mejora del cultivo.

<b>Especie</b>	<b>Característica de interés</b>
<i>S. lycopersicum var. cerasiforme</i> L.	Tolerancia humedad, resistencia a hongos y enfermedades de suelo.
<i>S. cheesmaniae</i> L. Riley Fosberg y <i>S. galapagense</i> S. Darwin & peralta	Tolerancia a salinidad, "jointless genes" y espesor del pericarpio.
<i>S. pimpinellifolium</i> L.	Color, características de calidad y resistencia a enfermedades
<i>S. chmielewskii</i> C.M. Rick, Kesicki, Fobes & M. Holle D.M. Spooner, G.J. Anderson & R.K. Jansen	Alto contenido de azúcares
<i>S. neorickii</i> C.M. Rick, Kesicki, Fobes & M.Holle D.M. Spooner, G.J. Anderson & R.K. Jansen	Resistencia a bacterias
<i>S. habrochaites</i> S.Knapp D.M Spooner	Tolerancia a frío y heladas, resistencia a insectos y enfermedades
<i>S. pennellii</i> Correll	Resistencia a sequia
<i>S. chilense</i> Dunal Reiche	Resistencia a sequia y enfermedades
<i>S. peruvianum</i> L.; <i>S. huaylasense</i> Peralta & S. Knaap; <i>S. arcanum</i> Peralta; <i>S. corneliomuelleri</i> J.F. Macbr.	Resistencia a virus, hongos y bacterias

**Tabla 1-2 Características agronómicas de interés de las especies silvestre de tomate.** (Tomado de Prohens et al. 2008).

Entre las características agronómicas más interesantes que muestran las especies silvestres del género *Solanum* secc. *Lycopersicon* se encuentran la resistencia a diferentes tipos de estrés. Tradicionalmente, los factores que provocan estrés en las plantas se han dividido en bióticos: plagas, enfermedades, malas hierbas entre otros; y abióticos: sequía, salinidad, bajas y altas

temperaturas, viento, problemas de suelo, etc. (Cuartero et al. 1995). Como ejemplo cabe citar la gran tolerancia a la salinidad en especies como *S. cheesmaniae*. Otras como *S. pennellii*, florecen en hábitats extremadamente secos del Perú Occidental, *S. habrochaites* puede tolerar el frío y las heladas, mientras que variedades de la forma silvestre de *S. lycopersicum* L. “cerasiforme” poseen tolerancia a la humedad (Esquinas-Alcázar 1981; Esquinas-Alcázar y Nuez 1995).

### 1.5.3 Sistemas de cruzamiento en tomate

Los sistemas de cruzamiento en el género *Solanum* varían desde la auto-compatibilidad (AC) en especies autógamas (por ejemplo *S. lycopersicum*, *S. pimpinellifolium*, *S. galapagense*) hasta la auto-incompatibilidad (AI) de las especies alógamas estrictas (por ejemplo *S. chilense*, *S. lycopersicoides*, *S. sitiens*). Existen además especies facultativas que presentan ambos tipos de reproducción. Estas últimas se dividen en dos grupos: el primero que incluye a *S. pimpinellifolium* y *S. chmielewskii* son predominantemente auto-compatibles, aunque sus estructuras florales favorecen el entrecruzamiento y, un segundo grupo formado por *S. pennellii* y *S. peruvianum* (Bedinger et al. 2011) que son especies predominantemente auto-incompatibles con algunas poblaciones auto-compatibles.

Cuando se pretende realizar cruzamientos entre especies de tomate, existe como regla general (aunque con excepciones) que si se utiliza como parental femenino una especie auto-compatible es posible obtener descendencia con prácticamente todas las especies. En cambio si el parental femenino es una especie auto-incompatible es imposible obtener descendencia sea cual sea el parental masculino; a este fenómeno también se le denomina incompatibilidad unilateral (IU) (Bedinger et al. 2011). El desarrollo de poblaciones entre *S. lycopersicum* y otras especies silvestres de tomate ha sido posible en muchas ocasiones al utilizarse *S. lycopersicum* como parental femenino, evitando la IU.

Algunas especies del género *Solanum* tiene un tipo de AI gametofítica, en la cual el polen es abortado a nivel de estilo aun cuando la germinación del polen se haya iniciado (Franklin-Tong y Franklin 2003). La AI es controlada por un simple locus (locus *S*) el cual produce ribonucleasas (S-RNasas) a nivel de pistilo que ejercen actividad sobre el polen rechazándolo cuando proviene de la misma flor (McClure et al. 1989), y fue localizado en el cromosoma 1 en *S. peruvianum* (Tanksley y Loaiza-Figueroa 1985). Un segundo gen HT-B, que produce una proteína HT de función desconocida, rica en asparagina, es al parecer también necesario para la respuesta de las especies AI (McClure et al. 1999).

### 1.5.4 *S. pimpinellifolium* como parental donante de genes

En esta revisión, daremos mayor relevancia a la especie silvestre *S. pimpinellifolium* que utilizaremos en el desarrollo de la genoteca de ILs (**Capítulo II**). Esta especie posee características de interés desde el punto de vista de la mejora y su proximidad filogenética con *S. lycopersicum* facilita su utilización en programas de cruzamiento. Dentro de los caracteres interesantes de *S. pimpinellifolium* destacan su fruto color rojo y comestible y que no presenta problemas de cruzamiento con la especie cultivada. El principal problema es su pequeño tamaño de fruto (entre 1 y 2 gramos). Además puede ser fuente de genes de resistencias a enfermedades. En la tabla 1-3 se indican algunos genes de resistencia introgresados en cultivares comerciales a partir de la especie *S. pimpinellifolium*.

<b>Enfermedad</b>	<b>Agente causal</b>	<b>Gen(s) de resistencia</b>
<b>Hongos</b>		
Marchitamiento por <i>Verticillium</i>	<i>Verticillium dahliae</i>	<i>Ve</i>
Marchitamiento por <i>Fusarium</i>	<i>Fusarium oxysporum</i> <i>f.sp.lycopersici</i> --patotipo 0 --patotipo 1	<i>I</i> <i>I-2</i>
Mancha gris de la hoja	<i>Stemphyllium solani.</i>	<i>Sm</i>
Moho de la hoja	Fulvia fulva ( <i>Cladosporium fulvum</i> )	<i>Cf</i> (1 al 24)
Tizón tardío	<i>Phytophthora infestan</i>	<i>Ph-1 Ph-2 Ph-3</i>
<b>Virus</b>		
Virus del enrollamiento de las hojas	Tomato leaf curl virus (TLCV)	<i>Tlc</i>
<b>Bacterias</b>		
Mancha bacterial	<i>Pseudomonas syringae pv. tomato</i>	<i>Pto</i>
<b>Nematodos</b>		
Nematodos del quiste	<i>Globodera rostochiensis</i>	<i>Hero</i>

**Tabla 1-3 Genes de resistencia a enfermedades** causadas por hongos, virus, bacterias y nematodos introgresados a cultivares comerciales a partir de *S. pimpinellifolium*. (Tomado de Prohens et al. 2008).

Adicionalmente, se han realizado una gran cantidad de estudios de mapeo utilizando poblaciones derivadas de cruces entre *S. lycopersicum* y *S. pimpinellifolium*, donde se han identificado un gran número de loci implicados en caracteres relacionados con la calidad de

fruto como tamaño, forma, color externo e interno, contenido en sólidos solubles, acidez, compuestos orgánicos y aromáticos entre otros (Foolad et al. 2007, Labate et al. 2007, Grandillo et al. 2011 y Grandillo et al. 2013). La accesión de *S. pimpinellifolium* que utilizaremos en este trabajo es la TO-937, proviene del Banco de Germoplasma de tomate de la Estación Experimental La Mayora (EELM) colectada originalmente en Perú, en la provincia de Lambayeque a 50 m de altitud (Cuartero et al. 1984). Una de las características interesantes de TO-937 es su resistencia a la araña roja (*Tetranychus urticae*). La resistencia a la araña roja en esta accesión está asociada a la presencia y densidad de tricomas glandulares de tipo IV y a la producción por parte de estos de acilsacarosas compuestos responsables de la antixenosis y antibiosis a *T. urticae* (Cano 2006). Los resultados más relevantes indican que la resistencia a la araña roja está gobernada por un gen mayor y dominante, aunque otros estudios indican que puede estar gobernada por varios genes menores (Fernández-Muñoz et al. 2000; Fernández-Muñoz et al. 2003).

En algunos genotipos con tricomas glandulares tipo IV y producción de acilazucars, derivados del cruce entre el cultivar comercial "MoneyMaker" y *S. pimpinellifolium* acc. TO-937, hay una disminución del tiempo de supervivencia de *Bemisia tabaci* (mosca blanca) durante el periodo de desarrollo desde el estadio de huevo. El principal interés radica en que *Bemisia tabaci* es el insecto asociado a la transmisión del virus TYLCD, que provoca importantes pérdidas económicas en el cultivo comercial de tomate en todo el mundo (Cano 2006). Otros caracteres que hacen de TO-937 una accesión interesante para mejora genética son su alto contenidos en el fruto de ácido ascórbico, azucars, ácidos orgánicos y carotenoides (Capel et al. 2011; Lima-Silva et al. 2012).

## **1.6 Uso de Marcadores genéticos para conocer la variabilidad natural de tomate**

Los marcadores genéticos se definen como características morfológicas o moleculares (proteína, ADN), cuyo polimorfismo puede ser detectado entre individuos, bajo control genético simple y son utilizados para realizar análisis genéticos (O'Brien 1990).

Hoy en día los marcadores genéticos están siendo ampliamente utilizados en programas de mejora vegetal. Desde este punto de vista, la utilidad de un marcador genético depende, entre otras, de las siguientes características: el grado de polimorfismo, su modo de herencia, preferentemente codominante, la ausencia de efectos epistáticos, que sea neutro o independiente de los efectos ambientales (Nuez et al. 2000).

Los marcadores morfológicos corresponden generalmente con rasgos fenotípicos fácilmente observables debidos a genes con efectos mayores que presentan herencia simple y expresión estable en diferentes ambientes (Nuez et al. 2000).

Los primeros marcadores moleculares fueron las isoenzimas. Estos marcadores fueron populares en los años 70, gracias al avance tecnológico de los geles de almidón y la tinción histoquímica de las proteínas, permitiendo visualizar variaciones en enzimas mediante la catalización de reacciones bioquímicas específicas (Becerra y Paredes 2000). Posteriormente, con los avances en la biología molecular se han desarrollado una amplia gama de marcadores basados en el polimorfismo de las secuencias de ADN. Estos marcadores son más abundantes y estables pues su expresión no es modificada por condiciones ambientales y/o el estado de desarrollo de la planta. (Winter y Kahl 1995).

Dentro de los marcadores moleculares podemos citar los RFLP (Restriction Fragment Length Polymorphisms) se basan en variaciones en el tamaño de fragmentos de ADN generados tras digestión con enzimas de restricción debidos generalmente a cambios en la secuencia de los sitios de restricción o inserciones/deleciones entre sitios de restricción (Botstein et al. 1980; Tanksley et al. 1989). Los RFLPs fueron, con mucho, los marcadores más utilizados durante las décadas de los 80's y 90's, permitiendo por primera vez la construcción de mapas saturados en tomate (Tanksley et al. 1992) y melón (Oliver et al. 2001) entre otras especies.

Con el establecimiento del procedimiento de PCR a finales de los años 80's, se desarrollaron marcadores derivados de esta tecnología. Tienen la ventaja de requerir pocas cantidades de ADN y no necesariamente de alta calidad. Estos marcadores se pueden clasificar en dirigidos o no dirigidos, dependiendo de que se conozca o no la secuencia de ADN contra la que van dirigidos. Los RAPDs (Random Amplified Polymorphic DNAs) y los AFLPs (Amplified Fragment Length Polymorphisms) se ubican dentro del primer grupo, siendo ambos de naturaleza dominante. Dentro de los marcadores diseñados a partir de secuencias de ADN conocidas están los SCARs (Sequence Characterized Amplified Regions), CAPS (Cleaved Amplified Polymorphic Sequences), SSRs (Simple Sequence Repeats) también llamados microsatélites (Hamada et al. 1982). Estos se encuentran distribuidos en forma abundante y uniforme a lo largo de todo el genoma de plantas, de ahí su enorme utilidad con fines diversos en la mejora vegetal (Jarne y Lagoda 1996; Holton 2001). Igualmente son altamente reproducibles entre laboratorios, lo que los hace muy útiles para integrar distintos mapas genéticos de una misma especie, así como para estudios de diversidad genética en colecciones de germoplasma (Powell et al. 1996; Jones et al. 1997; Gupta et al. 1999). Los

SSRs han sido los marcadores de referencia en genética vegetal desde finales de los 90's hasta la llegada de los marcadores SNPs (Single Nucleotide Polymorphisms). Los marcadores SNPs reflejan polimorfismos que afectan a un único nucleótido de una secuencia de ADN. Son marcadores bialélicos, codominantes, altamente reproducibles, son los más abundantes y se encuentran distribuidos aleatoriamente por el genoma de cualquier especie (Gupta et al. 2001; Mammadov et al. 2012).

### **1.7 Métodos de genotipado de SNPs**

Actualmente, hay una gran cantidad de técnicas de genotipado de SNPs que permiten analizar miles de muestras en forma rápida y eficiente (Ragoussis 2009). El principal factor que limita su utilización es la identificación previa de SNPs polimórficos y para ello es necesario determinar la secuencia de todo o parte del genoma(s). Sin embargo, este impedimento es cada vez menor, dada la mayor facilidad que existe actualmente para obtener el tipo de información necesaria mediante las técnicas de secuenciación masiva a partir de ADN o RNA de las especies a analizar.

Dentro de los métodos para genotipar SNP están los basados en la digestión con enzimas de restricción como es el caso de los RFLP; otros están basados en la extensión de un nucleótido simple, en la ligación de un oligonucleótido, o en secuenciación. Otros métodos que incluyen la utilización de pirosecuenciación y espectrometría de masa que permiten automatizar el análisis de SNPs con alto rendimiento (Liew et al. 2004; Pui-Yan 2001). Finalmente se han desarrollado micromatrices de genotipado, sistemas basados en "micromatrices", como los desarrollados por las compañías Affymetrix e Illumina, que permiten detectar desde cientos hasta millones de SNPs en un ensayo (Barnes et al. 2005).

En esta revisión mencionaremos algunos métodos para detección de SNPs, dando especial importancia al método High Resolution Melting (HRM), que será utilizado como parte de los materiales y métodos de esta Tesis.

#### **1.7.1 Métodos de genotipado de SNPs simples**

A) Métodos basados en digestión. El primer intento para identificar SNPs fue buscando alteraciones de sitios de restricción en todo el genoma humano (Botstein et al. 1980). Los RFLPs y CAPS descritos anteriormente están basados en detectar SNPs localizados en dianas de enzimas de restricción: la diferencia en una base entre alelos genera sitios de restricción que pueden ser usados como marcadores. Dentro de este método de detección de SNPs

también se incluye aquellos que utilizan la endonucleasa *Cel I*, que digiere regiones genómicas puntuales con apareamiento desigual. La especificidad de la *Cel I*, es utilizada para detectar mutaciones a través de la metodología TILLING (Targeting Induced Local Lesions in Genomes) (Angaji 2009; Comai et al. 2004).

B) Métodos basados en el cambio de conformación en la molécula de ADN causado por el polimorfismo. El método de SSCP (polimorfismo causado en la conformación del ADN de cadena simple) detecta mutaciones puntuales a través de cambios en la movilidad electroforética de ADN de cadena simple. (Gupta et al. 2001; Bernat et al. 2002; Nikolausz et al. 2003).

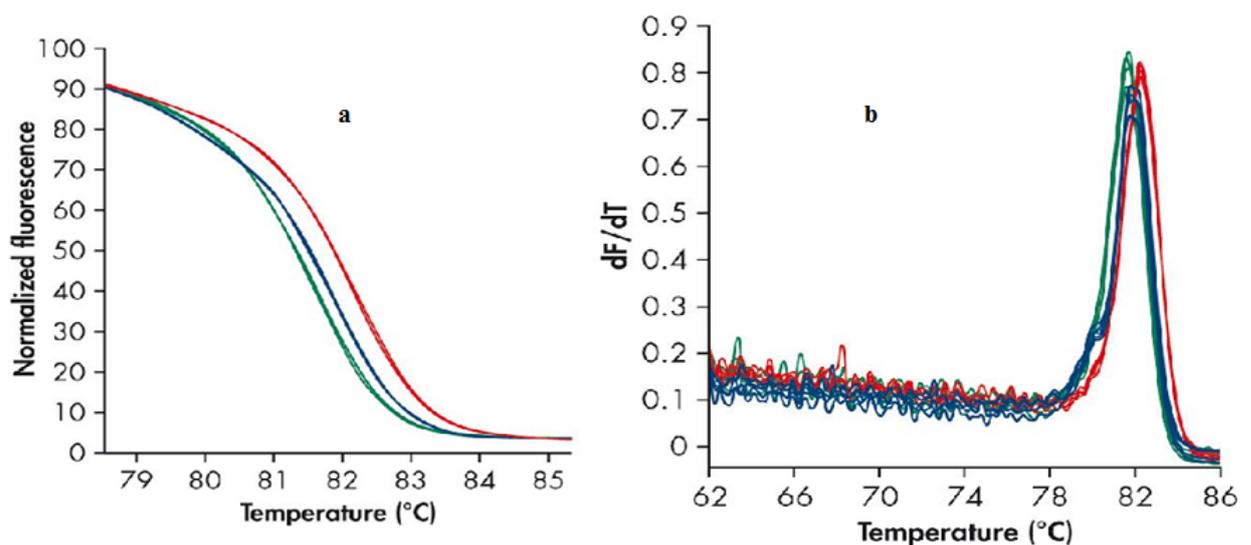
C) Otros métodos son: PCR-alelo específico (Allele Specific Primers) dentro de estos están el método de genotipado por fluorescencia KASP que conlleva cambios en la lectura de fluorescencia de productos obtenidos por PCR con tres oligonucleótidos: dos alelo específicos y uno común (LGC Genomics <http://www.lgcgenomics.com/genotyping/kasp-genotyping-chemistry/>). El método TaqMan consiste en la hibridación de sondas alelo específico y lectura de fluorescencia (Applied Biosystems: <http://www.appliedbiosystems.com>).

D) High Resolution Melting (HRM). Es un método post-PCR rápido, simple, altamente sensible y más barato que otras metodologías (Wittwer 2009). Consiste en una reacción de PCR con ligeras modificaciones en sus condiciones. La modificación más importante es la adición de una molécula intercalante (Dye) que emite fluorescencia cuando se une al ADN de doble cadena. El LC-Green fue el primer Dye utilizado, pero en la actualidad hay muchas opciones (Syto 9, Resolight, EvaGreen) que requieren ligeros cambios en las condiciones de PCR (Vossen et al. 2009).

Este método consiste en detectar el polimorfismo a partir de la generación y análisis de las curvas de fusión (melting curve) de amplicones generados por PCR. El método está basado en el efecto hipercrómico asociado a la intercalación en el ADN de doble hebra del Dye. Durante la fusión, el Dye intercalado en la doble hélice del ADN emite la mayor parte de la fluorescencia, la cual, es continuamente registrada por un sistema óptico muy sensible capaz de detectar un gran número de datos de fluorescencia en pequeñas variaciones de la temperatura (Reed et al. 2007; Vandersteen et al. 2007; Lee et al. 2008; Wu et al. 2008). La fluorescencia es pues alta al inicio de la fusión cuando todo el Dye está intercalado (ADN con cadena doble), pero a medida que aumenta la temperatura el Dye es liberado pasando a la solución y la fluorescencia decrece (**Figura 1-1a**).

La temperatura de fusión o “melting temp” ( $T_m$ ) se define como la temperatura en la cual el 50% de ADN esta en forma de cadena sencilla y la otra mitad como cadena doble. Para obtener la  $T_m$  se deriva matemáticamente la curva de fusión estándar (**Figura 1-1b**). Esta curva es específica para cada producto de PCR, y varía con el contenido de C-G, la longitud del amplicón y la secuencia de bases. De tal forma que, es posible que variaciones de una sola base en la secuencia provoca cambios en la cinética de fusión y por consiguiente, cambios de forma en la curva de fusión. También, si entre dos fragmentos hay un SNP, este se verá reflejado en una distinta  $T_m$ , por tanto, basándonos en la  $T_m$  y en la forma de la curva derivada, es posible detectar la presencia o ausencia de los SNPs, dependiendo del cambio y el contexto de la secuencia amplificadas es posible genotipar plantas con un alto grado de precisión, por medio de la comparación de las curvas de fusión y las  $T_m$  obtenidas para muestras problema utilizando los controles adecuados (Liew et al. 2004; Smith et al. 2010; Han et al. 2012).

El genotipado de SNPs por HRM se ha utilizado con éxito en prácticamente todas las especies, incluyendo a la Manzana (*Mallus domestica* Borkh.) (Chagne et al. 2008), lupino blanco (*Lupinus albus* L.) (Croxford et al. 2008), almendro (*Prunus dulcis* (Mill) D.A. Webb) (Wu et al. 2009), césped (*Lolium perenne* L.) (Studer et al. 2009) y patata (*Solanum tuberosum* L.) (De Koeyer et al. 2010).



**Figura 1-1. Método de genotipado de SNPs High Resolution Melting (HRM).** (a) Curva de fusión. La fluorescencia disminuya a medida que aumenta la temperatura consecuencia de la liberación de la molécula intercalante y (b) curva derivada de la fusión. Mostrando como valor máximo el correspondiente al punto de inflexión y que representa la  $T_m$  donde 50% de las moléculas de ADN son de doble cadena ( $dF/dT$ : derivada de la fluorescencia/derivada de la temperatura) (Fig. Tomada de <http://www.appliedbiosystems.com>)

### 1.7.2 Genotipado de SNPs de alto rendimiento

En los últimos años, ha habido enormes avances en las técnicas de genotipado, incluyendo micromatrices que permiten interrogar millones de SNPs. Además, los costes de genotipado han bajado considerablemente haciéndolos cada día más asequibles.

Actualmente se disponen de distintas plataformas que permiten adaptarse a las necesidades de los distintos proyectos, y que fundamentalmente responden al número de variantes y muestras que se quieren analizar. Las plataformas más populares permiten analizar desde 10 a más de 1 millón de SNP, en una sola reacción, y por tanto, economizan al máximo tiempo y coste.

Dentro de las plataformas para genotipado de alto rendimiento destacan por su mayor utilización el Affymetrix GeneChip, Illumina-Infinium BeadArray y MassArray Platform (Sequenom). A continuación describiremos estas tres tecnologías, pero para una explicación detallada de cada plataforma de genotipado se puede consultar la revisión de Ragoussis (2009).

A) La tecnología de Affymetrix GeneChip está basada en la discriminación alélica por hibridación directa de ADN genómico. Los oligonucleótidos se proporcionan ya sintetizados directamente en la superficie de la micromatriz, permitiendo la presentación de una elevadísima cantidad de sondas en una sola micromatriz; cada sonda contiene una longitud de 25 nucleótidos. Para cada SNP a interrogar se representan dos oligos: uno de complementariedad perfecta (perfect match probe) y otro de complementariedad imperfecta (mismatch probe). Así se pueden detectar y eliminar hibridaciones inespecíficas (Ragoussis 2009). Las principales ventajas de esta tecnología son: su alta reproducibilidad, la posibilidad de estudiar genomas completos y su sencillez de interpretación. Dentro de las desventajas se menciona el alto coste que aun la hace poco accesible para la mayor parte de proyectos de mejora (Aguado 2007).

B) Micromatrices de genotipado Illumina. La empresa Illumina ha desarrollado tecnologías que permiten la detección de un número variable de SNPs dependiendo de las necesidades de los investigadores. En todos los casos hay un primer paso de amplificación de todo el genoma seguido de un segundo paso de hibridación. Los ensayos Infinium que comenzaron con el formato de 10.000 a 100.000 ensayos multiplex, para aumentar varios millones de SNPs. Otra modalidad es la plataforma GoldenGate de Illumina para micromatrices de densidad media que permite genotipar en una reacción de 96 a 384 muestras entre 48 y 1.536 SNPs. El

sistema GoldenGate utiliza dos oligos alelo-específicos (OAE), para discriminar entre las variantes alélicas de un determinado locus (Fan et al. 2006), localizándose el nucleótido variable del SNP en el extremo 3'. Los cebadores y la amplificación están diseñados para que se pueda realizar de forma simultánea, empleando 3 cebadores comunes a todos los loci analizados. Dos de los cebadores empleados están marcados con dos fluorocromos y cada uno de ellos es complementario de una de las secuencias OAE. El tercer cebador es complementario a una secuencia común añadida a todos los OAE. Tras la amplificación, los productos de PCR se hibridan a un "array de beads" (cuentas de sílice de 3 micras situadas en una placa de micropocillos), cada una de las cuales está recubierta por oligos complementarios a un "IllumiCode" que especifica un determinado locus. La fluorescencia de cada uno de los beads permite reconocer la variante alélica que se ha amplificado en cada SNP (Ragoussis 2009).

C) Plataforma MassArray (Sequenon). Esta técnica se basa en la realización de una PCR múltiplex seguida de una reacción de extensión que utiliza como cebador un único oligo para detectar el SNP. Los productos de la extensión se analizan usando un espectrofotómetro de masas (MALDI-TOF MS), que es capaz de detectar entre genotipos homocigotos y heterocigotos a partir de las masas moleculares de los productos de extensión (Ragoussis 2009).

### **1.7.3 Genotipado basado en secuenciación masiva**

Con el surgimiento de la nueva generación de métodos de secuenciación masiva, se ha revolucionado los aportes de la genómica y transcriptómica a la biología. Además, el genoma de muchas especies de plantas se ha podido secuenciar, algo casi imposible desde el punto de vista económico con la tecnología Sanger.

Estas nuevas tecnologías de secuenciación también pueden ser utilizadas para descubrir, validar y evaluar marcadores genéticos en las poblaciones de interés agronómico. Hay descritas varias aproximaciones, que suelen tener en común un primer paso de reducción de la complejidad, para asegurar suficientes lecturas de secuencia de ADN en los mismos loci y en todos los genotipos bajo estudio. Esta reducción de la complejidad se consigue mediante digestión de ADN genómico con una o más enzimas de restricción; la selección de los fragmentos de restricción, el etiquetado de los fragmentos para la posterior identificación del material del que proceden y finalmente, la secuenciación de los fragmentos finales. Estos nuevos métodos son aplicables tanto a especies de las que se dispone de datos de

secuenciación del genoma como a especies no modelo cuyo genoma no está secuenciado. Los métodos de genotipado basados en secuenciación masiva más utilizados son: La librería de representación reducida (siglas en ingles (RRLs) y reducción de la complejidad de secuencias polimórficas (siglas en ingles CRoPS), sitios de restricción asociados a secuencias de ADN (siglas en ingles RAD-seq) y genotipado con baja cobertura de secuenciación, como el “Genotype-by-Sequence” (GBS) (Daver et al. 2011). En todos ellos se usan “barcodes” (secuencias cortas concretas añadidas a los oligonucleótidos adaptadores), que permiten distinguir secuencias procedentes de individuos concretos a partir de una única reacción de secuenciación masiva donde se han mezclado ADNAs procesados de individuos distintos, abaratando los costes. Los tres métodos permiten muestrear y secuenciar un pequeño subconjunto de regiones de todo el genoma. Siguiendo diferentes modificaciones, básicamente, el ADN es digerido con una enzima concreta, generalmente *SbfI*, los fragmentos resultantes son ligados a adaptadores que llevan extremos cohesivos con secuencias universales para la amplificación y secuencias cortas concretas (barcodes) que permitirán distinguir cada muestra individualmente. Los fragmentos ligados se rompen y se amplifican usando las secuencias universales para seleccionar únicamente fragmentos cortos flanqueados por los adaptadores que son secuenciados (Baird et al. 2008; Davey et al. 2011). El GBS consiste en la digestión con una enzima de restricción que no corte en la fracción repetitiva del genoma (como *ApeKI*), ligación con adaptadores con secuencias “barcodes” y adaptadores comunes sin ellas, de manera que no todos los fragmentos serán secuenciados eficientemente con Illumina debido a que no tendrán o los dos adaptadores con “barcodes” necesarios o serán demasiado largos, permitiendo un número suficiente de fragmentos cortos con los adaptadores correctos que podrán ser secuenciados para descubrir varios miles de SNPs en un experimento (Elshire et al. 2011).

## **1.8 Genómica del tomate**

El genoma haploide del tomate tiene un tamaño aproximado de 950 Mb, está organizado en 12 cromosomas y contiene unos 35.000 genes (The Tomato Genome Consortium 2012). Las secuencias codificantes se concentran en las regiones de eucromáticas que corresponden a menos del 25% del total del genoma. Actualmente un buen número de genes han sido descritos y asignados a localizaciones específicas en el mapa genético.

Dentro de la familia de las Solanáceas el tomate es considerado la especie modelo para los análisis genéticos. Ofrece ventajas como son un genoma de tamaño reducido con respecto a otros cultivos, tiempos de generación cortos, autógama y posibilidad de utilizar tecnología

rutinaria de transformación. Por ello ha sido posible obtener gran cantidad de información de tomate, así como recursos genéticos, citogenéticas y fisiológicos (<http://solgenomics.net/>). Entre las herramientas genéticas se encuentran cantidad de mutantes caracterizados (<http://www.zamir.sgn.cornell.edu>) y numerosos recursos fitogenéticos y entradas de especies silvestres (<http://tgrc.ucdavis.edu/>). También se han desarrollado micromatrices con genes de copia única para uso de la comunidad científica, diversas librerías BACs de la variedad Heinz 1706, etc. (<http://www.genome.arizona.edu/fpc/tomato>).

Existen varias bases de datos donde se recopila diversa información sobre distintos aspectos del tomate, que pueden ser consultadas mediante internet. La base de datos SGN (Solanaceae Genome Network; <http://www.solgenomics.net/>) contiene herramientas para el estudio de todas las solanáceas incluye todas las secuencias de genes y ESTs, marcadores moleculares y herramientas para su análisis. El SGN, contiene más de 20 mapas genéticos y físicos de tomate, patata, pimiento y tabaco con miles de marcadores. Los marcadores de esta base de datos van desde AFLP, CAPS, PCR, RFLP, SNP, SSR y dCAPS (Bombarely et al. 2011). Por su parte, existen aproximadamente 300.000 secuencias ESTs, disponibles en bases de datos públicos como el DNA-Banco de datos de Japón (<http://www.ddbj.nig.ac.jp/>); Solgenomics Network (<http://www.solgenomics.net/>) y MiBASE <http://www.pgb.kazusa.or.jp/mibase/> (Shirasawa et al. 2010).

En el año 2004, se inicio el proyecto de secuenciación del genoma de tomate por un consorcio internacional en donde están involucrados Korea, China, Inglaterra, India, Holanda, Francia, Japón, España, Italia y Estados Unidos. El objetivo inicial fue la secuenciación BAC a BAC de la región eucromatida lográndose secuenciar más de 1200 BACs. En 2009 se inicio la secuenciación del genoma completo habiéndose logrado un ensamblaje de alta calidad. En el 2012 se publicó por parte de The Tomato Genome Consortium los resultados finales de la secuenciación del cultivar "Heinz 1706". Las anotaciones del genoma del tomate, generadas por el proyecto de secuenciación son realizadas por el International Tomato Annotation Group (ITAG) y están disponibles en estado de pre-publicación en el sitio SGN ([http://solgenomics.net/gbrowse/bin/gbrowse/ITAG2.3\\_genomic/](http://solgenomics.net/gbrowse/bin/gbrowse/ITAG2.3_genomic/)).

## **1.9 Mapas genéticos del tomate**

Los mapas genéticos son representaciones gráficas del orden que ocupan los marcadores genéticos en un grupo de ligamiento. Se basa en el estudio de la segregación de marcadores genéticos en una población segregante adecuada (F<sub>2</sub>, BC, DH, RIL), los genes o marcadores

próximos en el cromosoma se transmiten juntos en la descendencia con más frecuencia que los genes o marcadores que están más alejados. Debido a la recombinación entre cromátidas homólogas durante la meiosis, en una población segregante se generarán una mezcla de genotipos recombinantes y no recombinantes entre dos loci cualesquiera. La frecuencia de genotipos recombinantes entre loci se emplea para calcular la frecuencia de recombinación, la cual a su vez es utilizada para inferir la distancia genética entre esos loci (Campa et al. 2012). Cuantos más recombinantes aparecen entre dos loci, mayor es la distancia genética, que se mide en centiMorgans (cM) (un centiMorgan equivale a una frecuencia de recombinación del 1%). La frecuencia de recombinación no es aditiva a lo largo del cromosoma debido a la presencia de doble recombinación, por lo que esta variable subestima el número real de recombinantes. Se han desarrollado funciones de mapeo a partir de algoritmos derivando funciones que estiman con menor sesgo el número real de recombinaciones a partir de las frecuencias de recombinación. Las más comunes son la de Haldane (1919), que asume que no hay interferencia entre eventos de recombinaciones cercanas y la de Kosambi (1944) que asume que hay cierta interferencia.

El primer mapa genético de alta densidad en tomate se publicó en 1992 e incluyó 1030 marcadores de diferente naturaleza (morfológicos, isoenzimas y RFLPs). Este mapa está basado en una población F<sub>2</sub> de 62 plantas derivadas del cruce entre *S. lycopersicum* (VF 36) y *S. pennellii* (LA0716) (Tanksley et al. 1992). Posteriormente, en el año 2000 se desarrolló otra población de 80 plantas con las mismas especies en la que se mapearon más de 2000 marcadores, en su mayoría microsatélites y COS (Conserved Ortholog Set) (Fulton et al. 2002). Además de los marcadores COS, este mapa contiene 149 marcadores SSRs que fueron identificados en ESTs.

Otro de los mapas disponibles está basado en el cruce entre *S. lycopersicum* cv. E6203 x *S. pimpinellifolium* LA1589 (Grandillo y Tanksley, 1996; Tanksley et al. 1996; Doganlar et al. 2002). Contiene 140 marcadores la mayoría de los cuales son marcadores RFLPs. También se encuentra disponible un mapa originado a partir del cruce de *S. lycopersicum* cv. E6203 x *S. habrochaites* LA 1777 (Bernacchi y Tanksley 1997) y que contiene un conjunto de marcadores localizados en el mapa EXPEN 1992 (Tanksley et al. 1992).

Finalmente, contamos en nuestro laboratorio con un mapa genético, elaborado con marcadores SNP, y desarrollado a partir de una población de líneas recombinantes consanguíneas (RILs) obtenidas mediante el cruce entre la variedad “MoneyMaker” de *S.*

*lycopersicum* y la accesión silvestre TO-937 de *S. pimpinellifolium* con más de 4.000 SNPs (Capel et al. 2011).

### 1.10 Mapeo de Quantitative Trait Loci

Algunos caracteres de importancia agronómica tales como rendimiento, calidad, tolerancia o resistencia a estrés bióticos y abióticos suelen presentar una distribución continua o cuantitativa y poseen una herencia compleja controlada por muchos loci o genes (Ferreira et al. 2012). Este tipo de caracteres suelen presentar baja heredabilidad y, por tanto, baja respuesta a la selección, además, tener interacciones con el ambiente y requieren de un fenotipado caro, laborioso y a menudo en condiciones definidas (Liu et al. 2006; Collard et al. 2005). A los genes que controlan estos caracteres se les ha denominado “locus implicado en un carácter cuantitativo” o Quantitative Trait Locus (QTL, Thoday 1960). Mediante el mapeo de QTLs se puede obtener, la posición y el efecto de los QTL, a partir de una población segregante adecuada (F<sub>2</sub>, BC, DH, RIL), y la integración de la variación de marcadores genéticos, generando un mapa genético. Ambos datos se analizan utilizando alguno de los métodos de análisis de QTLs, como el mapeo de intervalos o el mapeo de intervalos compuesto. Una vez obtenida la estimación de la posición y efecto de los QTLs, pueden identificarse los más interesantes para la mejora, y definirlos con marcadores moleculares que flanquean su posición. Los QTLs pueden ser seleccionados por selección asistida por marcadores (MAS, Marker Assisted Selection). Dentro de las ventajas que se atribuyen a MAS están a) el ahorro en tiempo, al sustituir ensayos de campo complejos que deben ser realizados en una estación y región específica por ensayos moleculares; b) aumentar la respuesta a la selección, ya que la heredabilidad de los marcadores es máxima; c) el poder seleccionar los genotipos en estado de semillero; d) el poder piramidar genes; e) el disminuir el arrastre de genes indeseables (particularmente importantes cuando se está introgressando un gen a partir de una especie silvestre); f) el permitir realizar pruebas de caracteres específicos donde la evaluación fenotípica no es factible, como por ejemplo ensayos con patógenos bajo cuarentena (Collard y Mackill 2008). No obstante, la eficacia de MAS depende de varios factores como son: **a)** el modo de herencia del carácter de interés, el grado de asociación entre el marcador empleado y el gen de interés, **b)** el número de individuos a analizar y el genotipo al que se transfiere el gen de interés (Francia et al. 2005).

Aunque el uso de marcadores genéticos para identificar QTLs responsables del control de caracteres fue ya propuesto por Sax (1923), su uso no se generalizó hasta la década de los 80, cuando Paterson et al. (1988) realizó el primer análisis exhaustivo y sistemático. En este

estudio se emplearon marcadores RFLPs para identificar y caracterizar QTLs de calidad de fruto que incluían tamaño, pH y contenido de sólidos solubles. Demostrándose por primera vez que los caracteres cuantitativos podían ser transformados a factores mendelianos discretos (Foolad 2007). Desde entonces se han venido realizando muchos estudios de este tipo en numerosas especies.

A pesar de la proliferación de trabajos de mapeo de QTLs que se ha observado en los últimos años su utilización en mejora de plantas todavía es muy limitado (Francia et al. 2005). Estas limitaciones se pueden deber entre otras a: que el QTL diana sea un falso positivo, la posición poco precisa del QTL identificado, deficiencias en el análisis de QTLs que conducen a una sobre o sub estimación del número de QTLs, la posibilidad de perder el QTL en el proceso de MAS por dobles recombinaciones, la dificultad para evaluar de forma precisa los efectos epispásticos y las interacciones del QTL con el ambiente y, finalmente a cambios en el fondo genético que cambian los efectos de los QTLs detectados en generaciones tempranas (Collard y Mackill 2008).

### **1.11 Poblaciones para mapeo de QTLs en tomate**

Debido a la limitada base genética dentro del tomate cultivado, la mayoría de las poblaciones desarrolladas para el mapeo de QTLs se han basado en cruces interespecíficos con parientes silvestres relacionados (Grandillo et al. 2011; Grandillo et al. 2013). La gran mayoría de las especies silvestres de tomate han sido usadas para mapeo de QTLs y genes, usándose con más frecuencia las especies *S. pennellii*, *S. pimpinellifolium* y *S. habrochaites* y con menor frecuencia (debido a problemas de cruzabilidad, desarrollo de híbrido F<sub>1</sub> y baja fertilidad) las especies *S. chilense* y *S. peruvianum*. La cercanía filogenética entre *S. pimpinellifolium* y el tomate cultivado ha ocasionado que sus poblaciones derivadas hayan sido los más utilizadas en experimentos de mapeo, dado la facilidad de cruzarlas y el fácil manejo de poblaciones segregantes (Chen y Foolad 1999).

Las primeras poblaciones recombinantes que se utilizaron en los primeros estudios de mapeo de caracteres cuantitativos fueron poblaciones tempranas tipo F<sub>2</sub>, BC<sub>1</sub>. Estas poblaciones recombinantes son fáciles de desarrollar y presentan alto desequilibrio de ligamiento, pero presentan varias desventajas entre ellas: **1)** limitaciones en las evaluaciones; por ejemplo en poblaciones F<sub>2</sub> y BC<sub>1</sub> las evaluaciones son realizadas sobre plantas individuales las cuales pueden no ser replicadas; **2)** presencia de excesiva variación genética que podría afectar negativamente la precisión en la detección y mapeo de genes y QTLs; **3)** inestabilidad los

efectos de los QTLs debido cambios en su constitución genética de generación en generación; 4) no son poblaciones de utilización inmediata en mejora genética (Foolad 2009). Para evitar estos problemas se han desarrollado poblaciones recombinantes más estables como son las líneas consanguíneas recombinantes (RILs), líneas de retrocruce consanguíneas (BILs) y las líneas de introgresión (ILs). Las poblaciones RILs presentan muchas ventajas para el mapeo genético que incluyen la posibilidad de repetir los experimentos en diferentes años y localidades y la posibilidad de evaluar múltiples caracteres en diferentes ambientes. Tienen la desventaja de que se requieren mucho tiempo para desarrollarlas particularmente cuando las especies participantes presentan problemas de esterilidad o auto-incompatibilidad. En contraposición a las RILs, las BILs son mejores para el mapeo de QTLs para estudiar cruzamientos inter-específicos al tener una contribución menor del genoma donante silvestre proveyendo por lo tanto un fondo genético más uniforme. Las BILs con una cobertura adecuada del parental donante pueden ser muy precisas en la identificación y caracterización de QTLs, sus altos niveles de homocigosis en comparación a las primeras poblaciones segregantes mencionadas les permite tener más precisión en la estimación y localización de QTLs (Foolad 2009).

Otra estrategia es el desarrollo de poblaciones de retrocruce avanzado para el estudio y transferencia de QTLs a partir de especies silvestres a líneas elite de mejora (AB-QTL) desarrollada por Tanksley y Nelson (1996) aplicada con éxito por primera vez en tomate y luego en otros cultivos. Consiste en combinar el análisis de QTL con el desarrollo avanzado de poblaciones. El análisis de QTLs se realiza en una población de retrocruzamiento avanzado ( $BC_2$  o  $BC_3$ ), de manera que los posibles efectos deletéreos del donante silvestre son minimizados, el fondo genético es en una alta proporción elite y el fenotipo de la población analizada es más similar al fenotipo de la línea elite. Una vez identificados los alelos de un QTL de interés, solamente se necesitan pocas generaciones de selección asistida para desarrollar líneas de introgresión (ILs) que contengan únicamente un fragmento cromosómico del donante conteniendo la región donde se ha mapeado el QTL y el resto del genoma sería isogénico al cultivar elite. Los efectos de los QTLs pueden verificarse en estas ILs, que serían el material de pre-mejora para su integración en planes de mejora. (Grandillo y Tanksley 2005). Una expansión de esta estrategia es el desarrollo de una genoteca completa de líneas de introgresión (ILs). Una genoteca de líneas de introgresión se define como un conjunto de líneas en las que se ha introgresado, en cada una de ellas, un único segmento cromosómico de una especie relacionada o parental donante de interés, en un fondo genético común (Eshed y Zamir 1994). Idealmente, el conjunto de introgresiones representaría el genoma completo del

parental donante. Para la obtención de ILs, se diseña un programa de retrocruzamientos durante 5-6 generaciones a partir de un híbrido F1, sobre el parental que actúa como recurrente, hasta obtener líneas idénticas al parental recurrente excepto en la región cromosómica donde se ubica el segmento introgresado, utilizando marcadores moleculares para seleccionar y monitorear el diseño de las líneas de introgresión.

Wehrhahn y Allard (1965) realizaron un trabajo pionero donde propusieron el análisis de líneas de introgresión para la disección de caracteres cuantitativos. La primera colección de ILs que cubre todo el genoma fue la desarrollada en tomate por Eshed y Zamir (1994). En los últimos 20 años, se han generado ILs y BILs en distintos cultivos entre ellos: tomate (Eshed y Zamir 1994; Fulton et al. 1997; Bernacchi et al. 1998; Chetelat y Meglic 2000; Monforte y Tanksley 2000), lechuga (Jeuken y Lindhout 2004), arroz (Liu et al. 1998), avena (Matus et al. 2003), soya (Concibido et al. 2003), *Brassica oleraceae* (Ramsay et al. 1996), trigo (Liu et al. 2006), cebada (Korff et al. 2004), melón (Eduardo et al. 2005; ver revisión hecha por Grandillo 2014).

### **1.12 Usos y ventajas de las ILs en la detección de QTLs**

Las líneas de introgresión son un recurso genético valioso para la identificación, caracterización y confirmación de la presencia de QTL de importancia agronómica (Eshed y Zamir 1995; Fridman et al. 2004; Schauer et al. 2006; Liu et al. 2006; Mei et al. 2006; Uauy et al. 2006; Hanson et al. 2007; Finkers et al. 2007; Szalme et al. 2007; Torjek et al. 2008; Falke et al. 2009; Di Matteo et al. 2010; Fu et al. 2010). El desarrollo de ILs proporciona ventajas para la mejora comercial al permitir estimar efectos de los QTLs en un fondo genético adecuado (normalmente una variedad elite) además, se pueden realizar estudios de interacciones QTL-ambiente, también, al tener un bajo porcentaje de alelos silvestres se facilita la transferencia de un rasgo interesante a una variedad elite. (Ali et al. 2010; Xu et al. 2010; Schmalenbach et al. 2011). Las ILs presentan especialmente ventajas sobre otras poblaciones segregantes a la hora de piramidar caracteres deseables. A este respecto Gur y Zamir (2004) evaluaron líneas de introgresión del cruce entre *S. pennellii* (parental donante) y variedad comercial de tomate M82, ellos diseñaron una línea (IL789) usando tres ILs independientes con introgresiones en los cromosomas 7, 8 y 9 y que mostraban un efecto sobre el contenido de sólidos solubles y el rendimiento. La línea IL798 mostró niveles superiores a los parentales para los caracteres estudiados principalmente cuando estaba en condición heterocigota.

Otras de las ventajas de las poblaciones de ILs es la posibilidad de visualizar un QTL como un factor mendeliano, esto es transformar la genética cuantitativa en genética mendeliana (Liu et al. 2006; Torjek et al. 2008; Ali et al. 2010; Xu et al. 2010). Otras aplicaciones incluyen el estudio de interacciones entre QTLs mediante cruzamientos entre ILs (Ali et al. 2010), el mapeo fino de QTLs (Eshed y Zamir 1995; Monforte y Tanksley 2000b; Ku et al. 2001; Monforte et al. 2001) y la clonación de QTLs (Frary et al. 2000; Fridman et al. 2000, Liu et al. 2002), o como materiales de pre-mejora.

### **1.13 QTLs definidos en tomate y regiones cromosómicas relacionadas**

Hay varias revisiones publicadas sobre la detección de QTLs en tomate (Foolad et al. 2007; Labate et al. 2007; Grandillo et al. 2011 y Grandillo et al. 2013), incluyendo caracteres de resistencia a enfermedades (hongos, bacterias, virus, nematodos e insectos) y tolerancia a estrés abióticos, relacionados con calidad de fruto (grados Brix, calidad organoléptica, peso, forma, firmeza, color, caracteres de maduración, etc.), y rendimiento, morfología de hojas, ramificación, número de frutos, etc. Estos QTLs han sido identificados en poblaciones obtenidas de cruces interespecíficos de *S. lycopersicum* con distintas especies silvestres del género *Solanum* (*S. pimpinellifolium*, *S. habrochaites*, *S. pennellii*, *S. neorickii*, *S. chmielewskii*, *S. peruvianum*, *S. cheesmaniae*, *S. galapagense*, *S. lycopersicum* "cerasiforme"; *S. arcanum*), destacando *S. pennellii* como la especie donde se han descrito más QTLs en el orden de millares, en cientos de caracteres de importancia agronómica (Lippman et al. 2007; Alseikh et al. 2013).

Estudios recientes han revelado la existencia de ciertas regiones cromosómicas en donde se concentran un elevado número de QTLs relacionados con la calidad del fruto. Por ejemplo la región distal del cromosoma 1 que mide cerca de 25 cM de *S. arcanum* LA1708 controla 15 de los 35 caracteres estudiados por Fulton et al. (1997a, b) o la región en posición distal del brazo largo del cromosoma 6 en la que 10 QTLs ligados a la producción total, firmeza o color del fruto se encuentran agrupados y ligados al marcador TG484. Otro ejemplo lo constituye la región de 22 cM en un extremo del cromosoma 9 que controla doce de los caracteres estudiados en *S. neorickii* LA2133 (Fulton et al. 2000). A su vez Causse et al. (2002) menciona que cerca del 66% de los QTLs de calidad de fruto mapeados en una población de RIL, estaban localizados en tan solo el 16% del longitud del mapa utilizado.

En arroz también se han realizado muchos estudios de QTLs, algunos de los cuales se encuentran publicados en la base de datos de cereales Gramene (<http://www.gramene.org/>).

En ella recogen 8646 QTLs relacionados con caracteres de desarrollo, morfológicos y de producción. Swamy y Sarla (2008), en un revisión de diez estudios de mapeo de QTLs en arroz, enumeran 137 QTLs implicados en el control de diez caracteres relacionados con la productividad, situados su mayoría en ocho regiones de los cromosomas 1, 2, 5, 8 y 9.

#### **1.14 QTLs para caracteres calidad de fruto en tomate**

La calidad de los frutos de tomate es un concepto muy amplio, que comprende parámetros externos e internos. Uno de los aspectos ampliamente usados para determinar la calidad de los frutos es el "flavor", término inglés que incluye la percepción de textura y temperatura en boca, el aroma y el sabor, siendo este último definido fundamentalmente por el contenido de azúcares solubles, ácidos grasos y compuestos volátiles (Saliba-Colombani et al. 2001). La relación entre azúcares y ácidos determina la intensidad del sabor de los frutos, de manera que altos contenidos de azúcares y altos niveles de ácidos son indispensables para alcanzar mejor sabor en el fruto (Grierson y Kader 1986).

Los atributos de la calidad el tomate, van a depender en gran medida del uso que se le dé al mismo, según sea este para la industria del procesado o para el mercado en fresco. Causse et al. (2003) establece relaciones entre los atributos de calidad (propiedades externas y de calidad organoléptica) en los frutos de tomate y la preferencia de los consumidores.

##### **1.14.1 QTLs para caracteres de peso, tamaño y forma del fruto**

Los QTLs relacionados con el peso del fruto (FW) han sido bien estudiados en tomate. Existen QTLs definidos en todos los cromosomas para este carácter en la población de ILS desarrollada a partir de *S. pennellii* acc. LA716 (Eshed y Zamir 1994). Lippman y Tanksley (2001) encontraron tres QTLs mayores para FW localizados en los cromosomas 1, 3 y 11 que explican el 67% de la variación fenotípica de este carácter. La generación de líneas de introgresión de alta resolución ha mostrado que todas las variedades de tomate cultivado llevan alelos *fw2.2* para peso alto de fruto, los alelos silvestres del gen presentes en frutos pequeños reducen el peso en más del 30% (Alpert y Tanksley 1996; Frary et al. 2000). Igualmente Causse et al. (2004) encontró QTLs relacionados con peso del fruto, número de lóbulos y caracteres de calidad en el extremo distal del cromosoma 2, sin embargo, los alelos favorables para peso de fruto y número de lóbulos eran desfavorables para los caracteres de calidad. Anteriormente, Stevens (1986) había encontrado una correlación negativa entre el carácter contenido de sólidos solubles (SCC) y los caracteres tamaño de fruto y rendimiento.

En cuanto a QTLs de tamaño del fruto (longitud (FLL) y diámetro (FDL)), estudios de mapeo molecular ha revelado la presencia de QTLs en los 12 cromosomas del tomate algunos con grandes efectos (Grandillo et al. 1999). La forma del tomate (FS) es un carácter que presenta mucha variabilidad dentro del tomate cultivado, esta variabilidad es explicada en gran parte por la variación en mutaciones en cuatro genes *SUN*, *OVATE*, *LC* y *FAS* (Rodríguez et al. 2011), *SUN* (Xiao et al. 2008) se localiza en el brazo corto del cromosoma 7 codifica para una proteína de la familia de dominio IQ67; *OVATE* (Liu et al. 2002) en el brazo largo del cromosoma 2, *FAS* (Cong et al. 2008) en el brazo largo del cromosoma 11 y *LC* (Muños et al. 2011), en el brazo largo del cromosoma 2, estos dos últimos controlando el número de lóculos y la forma de fruto aplanada. Aunque también se han detectado QTLs para FS de efecto menor en los cromosomas 1, 2, 8, 9, 10, 11 y 12 (Grandillo et al. 1999; Gonzalo y van der Knaap 2008).

#### **1.14.2 QTLs para Sólidos totales**

Es un carácter importante tanto para el tomate para consumo en fresco como para industria, y representan entre el 4-7.5% del peso fresco del fruto. Los sólidos totales se dividen en sólidos solubles (un 75% son principalmente glucosa y fructuosa) y, sólidos insolubles (un 25% que incluyen proteínas, celulosa, hemicelulosa, pectina y polisacáridos). Los sólidos insolubles están implicados en la viscosidad del fruto carácter que determina la calidad del jugo, salsa, pasta y sopa de tomate (Nguyen y Schwartz 1999). El rendimiento y la calidad de ciertos productos procesados está relacionado por el contenido de azúcar en los frutos, por ejemplo un incremento de un 1% en el contenido de sólidos representa un 20% de incremento en el rendimiento (Berry y Uddin 1991). Altos contenidos de azúcares también incrementan el gusto y el sabor en el tomate fresco (Jones y Scott 1984). Por estas razones el incremento en el contenido de sólidos solubles (SSC) en el fruto ha sido el objetivo de números programas de mejora del cultivo. Se estima que SSC en los cultivares de tomate comercial se encuentra entre 4.6% (tomate fresco) y 6.3% (tomate industria) del peso fresco (Foolad 2007). Sin embargo, se han identificado accesiones de tomate silvestre que incluyen las especies *S. pimpinellifolium*, *S. chmielewskii* y *S. cheesmaniae* con concentraciones entre 9-15% de SS (Hewitt y Garvey 1987). A pesar de la existencia de variabilidad genética en este carácter, el éxito en la mejora ha sido limitado debido principalmente a una correlación negativa entre rendimiento y SSC (Georgelis et al. 2004). Existen muchos estudios de identificación de QTLs para SSC usando diferentes tipos de poblaciones interespecíficas, estos se han asociado a diferentes regiones en los 12 cromosomas del tomate (revisión hecha por Foolad 2007). Uno

de los más estudiados es el QTL *Brix 9-2-5* incrementa el contenido de azúcar en un 25%, usando mapeo de ultra alta resolución se ha identificado el gen *lin5* que codifica para un invertasa apoplástica que modula la partición de azúcares en el fruto (Fridman et al. 2000).

### 1.14.3 QTLs para acidez y pH

La acidez influencia la capacidad de almacenamiento del tomate procesado, valores de pH bajos reducen el riesgo de contaminación por patógenos por ejemplo *Bacillus coagulans* es inhibida a pH 4.1. Por esta razón, los objetivos de los programas de mejora en lo referente a estos atributos son pH por debajo de 4.5 y valores de ácido cítrico arriba de 3.5g/100g de peso fresco (Lamberth et al. 1966). Causse et al. (2004) usando IL de *S. pennellii* (LA716) detectó 15 QTLs para acidez (cromosomas 2-5, 7-12) y 11 QTLs para pH (cromosomas 2, 4, 5, 8-12). Sin embargo, la mayoría de los QTLs para pH no mostraron estabilidad a diferencia de los de acidez que fueron detectados en evaluaciones realizadas en un segundo ensayo.

### 1.14.4 QTLs para color de fruto

El color rojo es un atributo de calidad importante tanto para el tomate para consumo fresco como para industria. El proceso de maduración del tomate está bien caracterizado a través de la evolución del color en la superficie del fruto (Hertog et al. 2007). Principalmente, hay degradación de clorofila y acumulación de carotenoides (licopeno) durante la maduración (Brandt et al. 2006). La cantidad del carotenoide-licopeno, el cual causa la coloración roja del fruto (Seymour et al. 2013), está bien caracterizada por el parámetro  $a^*$  y el  $\beta$ -caroteno que da color naranja a los frutos se puede medir a través del parámetro  $b^*$  en el sistema de medición del color CIELab (Sacks y Francis 2001). También, el interés en el color rojo de los frutos se ha incrementado dados los recientes descubrimientos de los beneficios que para la salud representan, con efectos relacionados con su poder antioxidante, disminución en la incidencia de cánceres y de enfermedades coronarias, cataratas y degeneración muscular (Giovannucci 1999; Willcox et al. 2003). El aumento del contenido de licopeno en el fruto de tomate se ha convertido en un importante objetivo de los programas de mejora del cultivo. A finales de los 90, se identificaron dos mutaciones espontáneas dentro del tomate comercial que contribuyen al aumento del nivel de licopeno *hp1* (cromosoma 2) y *hp2* (cromosoma 1) las cuales se han introgresado en cultivares comerciales (Yen et al. 1997; van Tuine et al. 1997). Sin embargo, al tener un efecto pleiotrópico adverso en procesos relacionados con la viabilidad de las semillas, el rendimiento, SSC, sensibilidad a varios patógenos, entre otras, el uso de estos genes para desarrollo de nuevas variedades comerciales no ha sido exitoso (Jarret et al. 1984).

En contraste, el gen *old-gold-crimson* (*og<sup>9</sup>*) localizado en el cromosoma 6, incrementan el licopeno a expensas del  $\beta$ -caroteno (Butler 1962) ha sido incorporado en cultivares modernos, aumentando en un 25% el contenido de licopeno. Otra mutación importante en mejora genética es la *u* (Uniform ripening), localizada en el cromosoma 10, cuyo alelo silvestre causa una coloración verde en los hombros del fruto producto de un incremento en la síntesis de cloroplastos, lo que redundaría en frutos con mayores contenidos de licopeno, ácido ascórbico y SSC (Powell et al. 2012). La mutación *u* se introdujo en los años 50 y está ampliamente distribuida en las variedades modernas, tiene efectos negativos sobre la calidad ya que disminuye el contenido de los compuestos expuestos anteriormente. QTLs relacionadas tanto con el color externo como interno se han asociado a los 12 cromosomas del tomate (Foolad 1999). Algunas ILs derivadas de *S. pennellii* muestran variaciones en el color, frutos de la IL6-3 y IL12-2 son de color naranja oscuro. Ronen et al. (2000) mapearon y clonaron un gen dominante que da coloración naranja en la IL6-3 (llamado *Beta*) que codifica para una licopeno  $\beta$ -ciclase con actividad en la ruta de biosíntesis de carotenoides. En el caso de la IL12-2 un simple gen dominante (*Delta*) es el responsable de la acumulación de carotenos a expensas del licopeno Ronen et al. (1999).

#### **1.14.5 QTLs para contenido nutricional y compuestos volátiles**

A pesar que el tomate no es considerado como un alimento altamente nutritivo, debido a que es de las pocas verduras/frutas que se consume en cantidades relativamente altas, sí que contribuye significativamente a las necesidades diarias de algunos nutrientes como son la vitaminas A y C (Ikeda et al. 2013).

El carotenoides  $\beta$ -caroteno responsable del color amarillo-naranja en los tomates, es también la principal fuente de vitamina A. La identificación y utilización de los genes involucrados en el contenido de vitamina tiene impacto económico y nutritivo significativo. Otro carácter importante que define la calidad el tomate es el sabor. Es un carácter muy complejo, con numerosos componentes genéticos y no-genéticos, no todos bien conocido o entendidos (Foolad 2007). Aspectos como el aroma, la textura, la apariencia, la temperatura, la palatabilidad influyen la percepción que tenemos del sabor. Sin embargo, la proporción entre los valores de SSC y la acidez están directamente correlacionadas con el sabor (Jones y Scott 1983). Un gran número de compuestos aromáticos (cerca de 400) han sido asociados al sabor del tomate, aunque de muchos de ellos no se conoce el rol que desempeñan. Sin embargo, de todos ellos solo 16 compuestos volátiles revisten importancia en la definición del sabor (Buttery 1993; Causse et al. 2002; Tandon et al. 2003; Tieman et al. 2012; Klee y

Tieman 2013). Otros factores que intervienen en el sabor del tomate son los índices de cosecha, cuando los tomates son cosechados en estados tardíos de cosecha son más dulces. Así mismo, condiciones de estrés durante el cultivo pueden favorecer o desfavorecer la calidad y el sabor de los frutos, por ejemplo condiciones de alta salinidad en ciertos estados de crecimiento pueden mejorar el sabor del tomate aunque esto puede causar una reducción en el tamaño del fruto (Mizrahi et al. 1988).

#### **1.14.6 QTLs para vida pos-cosecha**

El estado de madurez es uno de los factores que podría afectar positiva o negativamente atributos de calidad como la firmeza y la vida pos-cosecha del fruto fresco. Uno de los enfoques en mejora de la calidad en tomate ha sido el desarrollo de cultivares con larga vida pos-cosecha. Durante los últimos años se han identificado y caracterizado el rol del etileno y la enzima polygalacturonasa (PG) en los procesos críticos relacionados con la maduración y ablandamiento del tomate (Moore et al. 2002). Estudios fisiológicos y genéticos en los procesos relacionados con la maduración han permitido identificar y caracterización los genes mutantes Never ripe (*Nr*), Non-ripening (*nor*) (cromosoma 10) y ripening inhibitor (*rin*) (cromosoma 5). Los tres mutantes provocan una disminución drástica actividad de la enzima PG durante la maduración (Tigchelaar et al. 1978), que alargan la vida pos-cosecha, pero también reducen atributos de calidad. Otro gen mutante relacionado con la maduración es el Alcobaca (*alc*) descrito por Mutschler (1984) encontrado en una variedad local y que parece ser un alelo de *nor*. La utilización de estos genes para producir variedades con mayor vida pos-cosecha, se ha limitado a la producción de híbridos debido a que solo en condición heterocigota son capaces de retardar la maduración, alargar la vida pos-cosecha e incrementar la firmeza. En condición homocigota, los frutos presentan alteraciones demasiado importantes en la maduración que afectan negativamente su calidad (Foolad 2007). La firmeza del fruto de tomate no solo afecta las preferencias del consumidor sino también otros atributos como resistencia a enfermedades o la vida pos-cosecha. Mejorando la textura del fruto se fomenta una dieta saludable, se simplifica la logística de la cadena alimenticia y se reduce el desperdicio por cosecha (Barrett et al. 2010). La firmeza está influenciada por factores que incluyen la estructura de la pared celular, turgencia (Saladié et al. 2007), propiedades de la cutícula (Chaib et al. 2007) por lo tanto es un carácter muy complejo que involucra muchos genes y rutas metabólicas (Brummell y Harpster 2001).



## **OBJETIVO GENERAL**

Estudiar la variabilidad genética presente en la accesión TO-937 de *Solanum pimpinellifolium* para la disección genética de caracteres relacionados con diversos aspectos de la biología del tomate y facilitar su incorporación a los programas de mejora genética del cultivo.

### **Objetivos específicos**

- 1) Desarrollar una genoteca de líneas de introgresión que contengan representado el genoma de la accesión TO-937 (*Solanum pimpinellifolium*) en el fondo genético del cultivar "Moneymaker" (*Solanum lycopersicum*), apoyándonos en herramientas de genotipado de alto rendimiento y lograr de forma más eficiente una mayor calidad de la colección final.
- 2) Estudiar las características de un locus responsable de la segregación distorsionada observada en descendencias de este cruzamiento durante el curso de la generación de la genoteca y definir genes candidatos.
- 3) Identificar introgresiones de TO-937 que porten genes implicados en caracteres relacionados con la calidad de fruto, mediante análisis de QTLs en ensayos agronómicos realizados con la genoteca de líneas de introgresión cultivada en varias localizaciones.



## **CAPITULO 2: Highly Efficient Genomics-Assisted Development of a Library of Introgression Lines of *Solanum pimpinellifolium***

### **Introduction**

The cultivated tomato (*Solanum lycopersicum* L.) shows a wide range of variation in a number of fruit characteristics such as shape and size (Rodriguez et al. 2011) as well as volatile (Tieman et al. 2012; Rambla et al. 2014) and flavonoid (Slimestad and Verheul 2009) composition; however, the genetic variability available for certain breeding purposes is limited (Miller and Tanksley 1990). Genetic erosion of this crop has resulted from repeated genetic bottlenecks originating during its domestication in Mesoamerica and continued with its cultivation and initial stages of cultivar development in Europe and North America, and has been reduced further still with modern breeding practices (Rick 1986). The genus *Solanum* includes several wild species (Spooner et al. 2005), a number of which may be easily hybridized with *S. lycopersicum*, making it possible to transfer genetic diversity between species. In fact, wild species closely related to *S. lycopersicum* were already used in early tomato breeding programs in the 1940s to find disease-resistance genes (Alexander et al. 1942). Currently, wild *Solanum* species have a significant role in tomato research and breeding. Specifically, *S. pimpinellifolium* L., the species most closely related to *S. lycopersicum*, has been extensively used as a source of genetic variability for disease resistance to tomato yellow leaf curl virus (TYLCV) (Chaque et al. 1997), *Fusarium oxysporum* f.sp. *Lycopersici* (Bournival et al. 1989), *Cladosporium fulvum* (Stevens and Rick 1986), *Botrytis cinerea* (Ignatova et al. 2000), *Pseudomonas syringae* (Pitblado et al. 1984), *Phytophthora infestans* (Chunwongse et al. 2002), *Globodera rostochiensis*, and *Globodera pallida* (Ellis and Maxon-Smith 1971), as well as for fruit quality traits (Tanksley et al. 1996; Bernacchi et al. 1998). Additionally, several *S. pimpinellifolium* accessions have been identified as potential sources of abiotic stress tolerance (Monforte et al. 1996; Foolad 2005). However, the potential of wild species as a source of valuable alleles for the improvement of cultivar germplasm is far from being fully exploited (Grandillo et al. 2011), and this is especially true for traits under complex genetic control. Over the last two decades, the birth of molecular marker technology, the development of mapping populations, and the use of marker-assisted selection (MAS) have facilitated the identification, mapping, and transfer to tomato of many disease-resistance genes and quantitative trait locus (QTL) from wild species (Foolad and Sharma 2005; Labate et al. 2007). Tanksley and Nelson (1996) introduced the

advanced backcross quantitative trait locus (AB-QTL) strategy in order to integrate QTL mapping and breeding population development and to facilitate the introduction of new genetic variability from wild species into elite germplasm. Studies in tomato and other crops have shown that this method efficiently discovers favorable alleles in wild germplasm resources (Xiao et al. 1998; Moncada et al. 2001; Septiningsih et al. 2003; Blair and Izquierdo 2012). The ultimate aim of the AB-QTL approach is to make unadapted germplasm accessible to the breeder by developing introgression lines that carry the favorable QTL alleles. An expansion of this idea is the construction of genetic libraries of introgression lines (IL), each IL in the collection harboring a single, marker-defined chromosome segment from a donor genotype within a uniform genetic background, usually an elite cultivar, with the introgression collection representing ideally the full genome of the donor (Zamir 2001). Each IL typically contains more than 95% of the recurrent cultivated parent genome, thereby avoiding possible sterility problems or deleterious combinations and undesirable linkage drag effects that may occur in wide crosses involving wild species, and has the additional advantage that agronomic management and plant phenotypic characteristics are very similar to those typically found in the crop species, which allows the estimation of the QTLs effects in an appropriate genetic background. QTL mapping in such populations is simplified as phenotypic variation between each IL and the recipient cultivar can be directly associated with the introgressed segment. The first IL population was developed in wheat by Wehrhahn and Allard (1965) who demonstrated that individual QTL effects could be efficiently estimated using backcross inbred lines (BILs) with a low proportion of the donor parent. Most IL populations have been developed during the last fifteen years by means of marker-assisted selection (MAS) involving different marker systems and breeding approaches. In tomato, ILs have been developed from *Solanum pennellii* LA0716 (Eshed and Zamir 1994), *S. habrochaites* LA1777 (Monforte and Tanksley 2000a), *S. lycopersicoides* LA 2951 (Chetelat and Meglic 2000; Canady et al. 2005), *S. habrochaites* LA0407 (Francis et al. 2001), and *S. habrochaites* LYC4 (Finkers et al. 2007). In addition, a reduced number of ILs have been developed for the *S. pimpinellifolium* accessions LA 1589 (Tanksley et al. 1996, Bernachi et al. 1997) and LA2093 (Kinkade and Foolad 2013). These ILs collections have proved to be extremely helpful in the identification of QTLs (Eshed and Zamir 1995; Rousseaux et al. 2005), the verification of QTL effects (Tanksley et al. 1996), to study QTL x environmental, QTL x genetic background, and QTL x QTL interactions (Monforte et al. 2001), QTL fine mapping (Eshed and Zamir 1996; Monforte and Tanksley 2000b; Ku et al. 2000; Ashrafi et al. 2012), and the introduction of new genetic variability from wild species into elite germplasm

(Tanksley and McCouch 1997; Zamir 2001; Gur and Zamir 2004). IL analysis is also a powerful tool for genomics research, as it facilitates the study of the genetics of metabolism (Schauer et al. 2006), transcriptomes and the correlation with metabolism (Lee et al. 2012), enzyme activity (Steinhauser et al. 2011), and QTL cloning (Frery et al. 2000; Fridman et al. 2000; Liu et al. 2002). Apart from tomato, IL populations have been developed for barley (von Korff et al. 2004), cabbage (Ramsay et al. 1996), lettuce (Jeuken and Lindhout 2004), melon (Eduardo et al. 2005), soybean (Concibido et al. 2003), rice (Tian et al. 2006), wheat (Pestova et al. 2006; Liu et al. 2006), and *Arabidopsis thaliana* (Koumproglou et al. 2002) among other species. However, the number of IL collections is still limited even compared to the number of sequenced crop plant genomes. For example, genome resequencing is feasible in a relatively short period of time (Xu et al. 2012) and large resequencing projects are under way in major crops, such as tomato (<http://solgenomics.net/organism/sol100/view>; <http://www.tomatogenome.net/>). However, the availability of appropriate mapping populations to identify genes of agronomic interest as well as QTLs from exotic germplasm (such as IL libraries) is still a key limiting factor to exploiting these genetic resources. Currently, the increased availability of genomics technologies may also help to speed up the development of such populations.

In tomato, current genetic and genomic resources include thousands of molecular markers based on single-nucleotide polymorphisms (SNPs), various molecular linkage maps, BAC libraries and physical maps, genome sequences, multiple permanent mapping populations, germplasm banks containing important collections of wild species, mutant collections, websites providing information about resources for tomato, as well as methodologies and bioinformatics tools (Grandillo et al. 2011; Sato et al. 2012; Sim et al. 2012).

In the current report, we take advantage of the availability of these diverse genomics tools to increase the efficiency and speed of a breeding program aimed at developing a complete IL collection from the *S. pimpinellifolium* accession TO-937 in the genetic background of *S. lycopersicum* cv. "Moneymaker". Several accessions of this wild species, which is closely related to the cultivated tomato (Rick, 1979), have been used to detect QTLs involved in fruit quality traits (Tanksley et al. 1996; Doganlar et al. 2002; Kinkade and Foolad 2013), abiotic tolerance (Monforte et al. 1996; Villalta et al. 2008), and pest resistance (Salinas et al. 2013). Furthermore, the *S. pimpinellifolium* accession LA1589 (utilized for constructing map populations by Tanksley et al. 1996 and Doganlar et al. 2002) has recently been resequenced (<http://solgenomics.net/organism/Solanumpimpinellifolium/genome>). However, an IL

collection covering the whole *S. pimpinellifolium* genome is not currently available. In the current report, we selected the *S. pimpinellifolium* accession TO-937 (hereafter referred to as *SP*), a red-fruited, self-compatible accession, obtained after at least four selfings from the accession PE-2 which was originally collected in Lambayeque, Peru (Cuartero et al. 1984) and maintained in the tomato germplasm bank of the Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (IHSM). TO-937 has interesting features that made it suitable for founding the IL population. It was found to be resistant to pests, such as the two-spotted spider mite *Tetranychus urticae* (Fernández-Muñoz et al. 2000), due to the presence on its leaves and stems of type IV glandular trichomes (Fernández-Muñoz et al. 2003) which produce acylsucroses (Alba et al. 2009) that are not present in the cultivated species or in several other *S. pimpinellifolium* accessions, such as LA1589. Additional interesting traits belonging to TO-937 include high contents of metabolites related to fruit quality, such as ascorbic acid (Lima-Silva et al. 2012), sugars, organic acids, and carotenoids (Capel et al. 2011), as well as aroma volatile compounds (Rambla et al. 2014). Furthermore, populations derived from this accession were used to map the uniform green (*ug*) mutation (Powell et al. 2012). High-throughput genotyping analysis in early backcross populations should allow researchers to obtain a set of plants with a minimum number of introgressions from which single-introgression lines may be developed in a short period of time. This would significantly reduce the time and number of generations necessary to obtain a complete IL library. Finally, the characterization of the definitive IL library with a high-throughput SNP genotyping assay, such as the 8K SNP Illumina Infinium array (Sim et al. 2012), would thoroughly verify the extension of the introgressions and the isogeny of the genetic background.

## **Materials and methods**

### **1. Plant materials**

The recurrent *S. lycopersicum* parent selected for the present study was the indeterminate growing, fresh-market tomato cultivar 'Moneymaker' (hereafter referred to as *SL*), which produces medium-sized, round, red fruits, and which has been used extensively in the literature for fruit development and ripening, physiological, and phytopathological studies, and from which there are available genetic resources such as EMS mutant collections and TILLING platforms (e.g., <http://www-urgv.versailles.inra.fr/tilling/tomato.htm>) and collections of nearly-isogenic lines containing disease-resistance genes (Laterrot 1996). The donor parent TO-937 was obtained from the IHSM germplasm bank.

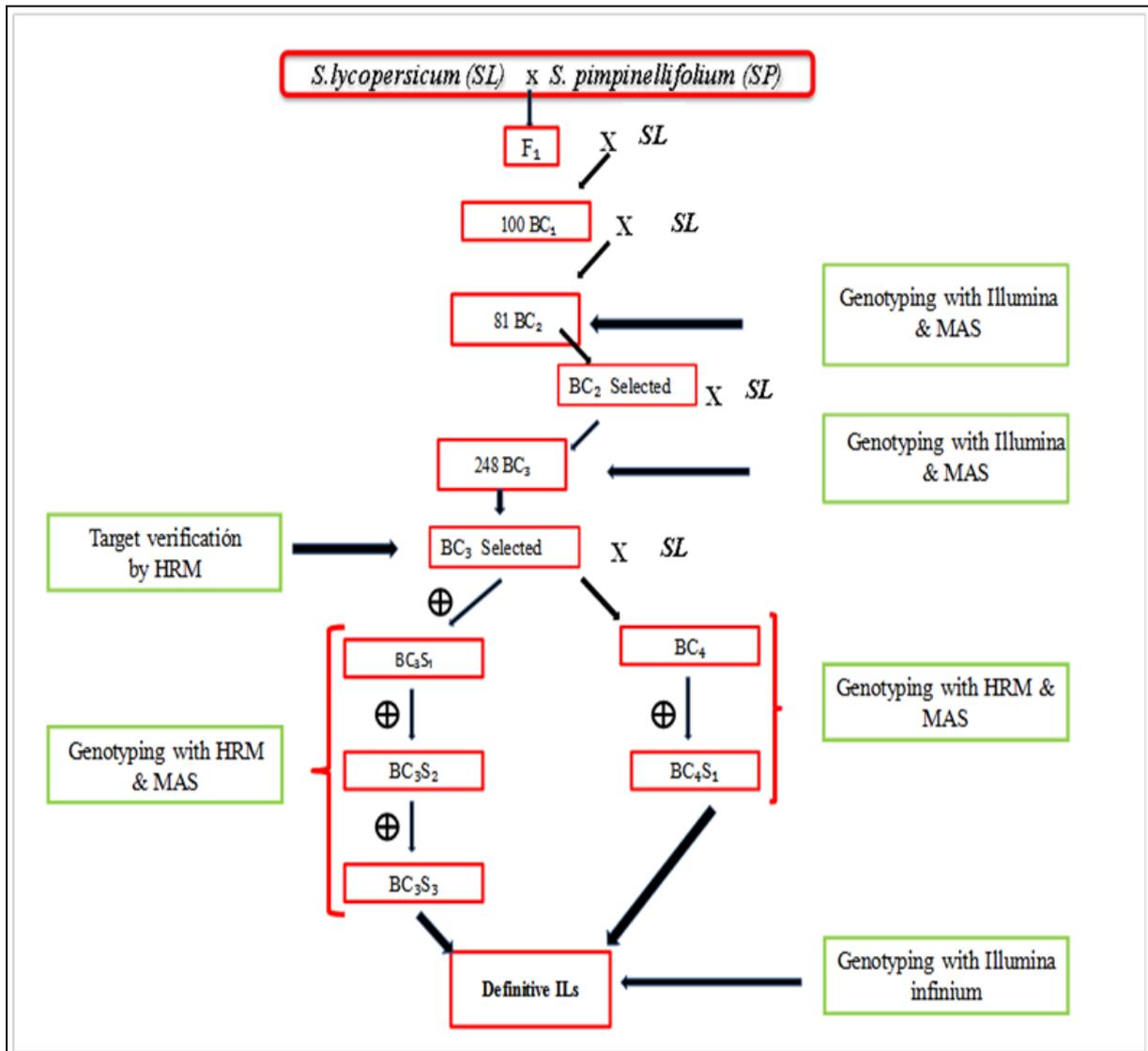
## 2. Genomic characterization of TO-937

TO-937 was genotyped with the 8K SOLCAP SNP Infinium array (see below) and the results were merged with all *S. pimpinellifolium* genotype data from Sim et al. (2012) and Blanca et al. (2012) in order to locate TO-937 in the genetic variability spectrum of this species. Nei et al. (1983) genetic distances and neighbor-Joining (NJ) tree were calculated with Powermarker 3.5 (Liu and Muse 2005), and the tree was drawn using MEGA 6.0 (Tamura et al. 2013).

## 3. Breeding scheme

The breeding scheme used to develop the IL library is depicted in **Fig. 2-1**. A single F<sub>1</sub> plant from the cross between SL and SP was backcrossed with the recurrent parent generating 100 BC<sub>1</sub> plants which were backcrossed following the Single Seed Descendence (SSD) method, thus producing 81 BC<sub>2</sub> plants. These BC<sub>2</sub> plants were genotyped with an array of 712 SNPs (see below) in order to select the most suitable ones for our breeding purposes according to the following criteria: (1) in the whole collection the entire genome should be represented at least twice with overlapping donor segments, and (2) each of the selected plants should contain the lowest possible proportion of the donor genome. A total of 16 BC<sub>2</sub> plants meeting the above criteria were selected and backcrossed to obtain 248 BC<sub>3</sub> plants, which were genotyped at seedling stage with an array of 96 SNPs (see below) and subjected to another round of selection using the criteria outlined above. A total of 76 BC<sub>3</sub> plants were selected and divided into two groups depending on the number of introgressions: an elite group of 41 plants (each carrying three or fewer introgressions) and a backup set with 35 plants (each carrying three or more introgressions). In order to verify the target introgressions, the selected BC<sub>3</sub> plants were genotyped with 96 SNP markers by High Resolution Melting (HRM, see below) and further backcrossed and selfed to obtain BC<sub>4</sub> and BC<sub>3</sub>S<sub>1</sub> families, respectively.

Depending on the number of introgressions, BC<sub>4</sub> or BC<sub>3</sub>S<sub>1</sub> seedlings were screened with the appropriate SNPs by HRM in order to select plants with single (target) introgressions. Given the number of introgressions in a genome (N), the number of progeny that need to be screened to find at least a single plant with a 95% probability can be calculated from the binomial distribution of the allele segregation as  $N = \frac{\log(0.05)}{\log(1-0.25^n)}$  for BC<sub>3</sub>S<sub>1</sub> families and  $\frac{\log(0.05)}{\log(1-0.5^n)}$  for BC<sub>4</sub>.



**Fig. 2-1. Marker-assisted selection (MAS) program to develop a *Solanum pimpinellifolium* acc. TO-937 (SP) introgression line library in the *S. lycopersicum* cv. ‘Moneymaker’ (SL) genetic background.** From the initial cross, the strategy combined high-throughput genotyping in initial populations, MAS, backcrossing, selfing, and within-family MAS by high resolution melting (HRM). The definitive IL collection was characterized using the 8K SOLCAP Illumina Infinium genotyping array.

#### 4. Selection differential

The selection differential for the reduction of *SP*-allele frequency in the BC<sub>2</sub> population was calculated according to Falconer (1989):  $S = 100 \times \{f(\rho) - f(\varsigma)\} / f(\rho)$ , where  $f(\rho)$  and  $f(\varsigma)$  represent the average frequency of *SP* alleles in the whole population and in the subset of selected plants, respectively.

## 5. Genotyping methods

### 5.1 DNA extraction

Genomic DNA was extracted from young leaves, following the Doyle and Doyle (1990) method with slight modifications. For the tissue crushing, two steel balls (3-4 mm) and 370  $\mu$ l of extraction buffer were added to each microtube containing 100 mg of fresh tissue and then ground in a Geno/Grinder 2000 (SPEX Sample Prep, New Jersey, USA) for 5 minutes at 1,400 rpm. The extracted DNA was dissolved in Milli-Q water and the final concentration was adjusted to 100-200 ng/ $\mu$ l.

### 5.2 Genotyping

A genetic map including 665 SNP markers from a 712 GoldenGate Illumina array was developed for a 169-recombinant-inbred-lines (RILs) population from a cross between the same parents (Alba et al. 2009; Fernandez-del-Carmen et al. 2011; details to be published elsewhere). This set of SNPs was used to characterize the BC<sub>2</sub> and BC<sub>3</sub> populations (**Annex 2-1**). Briefly, the map had an average of 59 SNP markers per chromosome (range 44 – 106), with an average marker density of 2 cM/SNP. BC<sub>2</sub> samples were genotyped with a 712-SNP GoldenGate Illumina array by the Centre for Genomic Regulation (CRG) in Barcelona, Spain (<http://pasteur.crg.es>). A 96-SNP IlluminaVeracode array was designed to genotype the BC<sub>3</sub> population by the CRG service.

Genotyping in subsequent generations was performed by high resolution melting (HRM, Gundry et al. 2003). SNP markers from the 8K SNP SOLCAP Tomato Infinium array (Sim et al. 2012) located physically adjacent to the previously used markers were selected and adapted for HRM analysis. Briefly, primers were designed from SNP flanking sequences using Primer Express Software v. 2.0 (Applied Biosystems), and their specificity was checked by BLAST searchers (implemented through the SGN website, <http://solgenomics.net/tools/blast/>) against the tomato genome v. SL 2.40. HRM-based genotyping was tested for each primer pair using selected BC<sub>3</sub> plants in order to verify its agreement with the Veracode-based genotyping. The IL collection was finally characterized by the 8K SNP SOLCAP Tomato Infinium array by the “Unidad de Genotipado y Diagnóstico Genético” service of the “Fundación Investigación Clínica de Valencia-Instituto de Investigación Sanitaria” (INCLIVA) in Valencia, Spain. The introgressions were finally defined based on the physical map according to solgenomics.net and a new genetic map based on the RIL population was

made with SOLCAP markers (Fernandez-del-Carmen et al. 2011, details to be published elsewhere).

## **Results and discussion**

### **1. Genomic characterization of TO-937**

The NJ dendrogram of *S. pimpinellifolium* accessions (**Annex 2-2**) showed a clear separation of accessions based on their geographical origin (Ecuador or Peru) as previously observed by Blanca et al. (2012). TO-937 was grouped with other Peruvian accessions, and was genetically distinct from LA 1589, the accession that was resequenced (<http://solgenomics.net/>), and also from LA2093 (Kinkade and Foolad 2013), demonstrating that TO-937 includes genetic variability that has not been previously exploited.

### **2. Development and selection in early backcross populations**

A set of 81 BC<sub>2</sub> plants was genotyped with the 712-SNP Illumina GoldenGate array, from which a total of 665 SNPs segregated. On average, the BC<sub>2</sub> plants contained 73.5% of the *SL* genome (range 75.5% - 63.9%) and 8.57 introgressions. A subset of 16 BC<sub>2</sub> plants that fulfilled our selection criteria was selected, representing the *SP* genome at least twice. The plants of the selected population had 82.4% of the *SL* genome (range 88.4% - 68.2%) and 6.25 introgressions from *SP* on average (**Annex 2-3**). The selection strategy based on high-throughput genotyping allowed approximately 10 % of the recurrent parent genome to be recovered while decreasing the average number of introgressions by a third as compared to the initial population (**Table 2-1**). The selection differential (S) for reduction of the *SP* allele frequency in the BC<sub>2</sub> population was 33%.

The 16 selected BC<sub>2</sub> plants were backcrossed to generate 248 BC<sub>3</sub> seedlings (16 plants /family). A total of 96 SNP markers were selected from the previous 665-SNP map set in order to genotype these seedlings. The 96 SNP markers were distributed over 8 SNPs on average per chromosome (range 6-11), with an average marker density of 9.5 cM/SNP, ranging from 6.3 cM/SNP for chromosome 12 to 11 cM/SNP for chromosome 7 with a maximum gap of 28.28 cM in chromosome 5 (**Annex 2-4**). The 248 BC<sub>3</sub> seedlings showed an average of 92% of the *SL* genome (range 96.2% - 84.2%) and 3.7 introgressions. Following the previous selection criteria, a total of 76 seedlings of the 248 were selected, which resulted in an average of 92.0% of the *SL* genome (range 95.6% - 88.2%) and 2.9 introgressions (**Anex 2-5**). This selected 76-member sub-population was divided into two groups: an elite group

Population	Genotyping assay	Percentage of <i>SL</i> genome (%)		Average number of <i>SP</i> introgressions
		Average	Range	
Unselected BC <sub>2</sub>	665 SNPs Illumina GoldenGate Array	73.5	(75.9 – 63.9)	8.57
Selected BC <sub>2</sub>		82.4	(88.4 – 68.2)	6.25
Unselected BC <sub>3</sub>	96 SNP Illumina Veracode Array	92	(96.2 – 84.2)	3.7
Selected BC <sub>3</sub>		93	(95.6 – 88.2)	2.9
Elite BC <sub>3</sub>		93.8	(99.4 – 87.3)	2.3
Definitive ILs	HRM (126 SNPs)	96	(99.5 – 92.3)	1
Definitive ILs	8K Illumina Infinium Array	95.7	(99.7 – 91.8)	1.9

**Table 2-1 Average and range of percentages of *Solanum lycopersicum* (*SL*) genome composition and average number of introgressions from *S. pimpinellifolium* (*SP*) in the unselected and selected BC<sub>2</sub>, BC<sub>3</sub> populations and final introgression line (IL) populations estimated by several genotyping methods.**

composed of 41 plants whose genome harbored three or fewer introgressions representing the *SP* genome, and a second set with the remaining 35 plants as a backup for the first group, all with more than two introgressions. The selection performed on the BC<sub>3</sub> seedlings did not reduce the percentage of the *SL* genome (it actually increased slightly from 92.0% to 93.8%). However, the average number of introgressions decreased by one, indicating that this selection step had a major impact on the design of the genetic structure of the population. Thus, the average number of introgressions was 2.3, where 10, 11, and 20 plants presented one, two, and three introgressions, respectively, which made it feasible to fix single-introgression lines for most of the genome with just one additional step of marker-assisted selection. On the other hand, the average number of introgressions in the whole BC<sub>3</sub> was 3.7, making it necessary to screen a large number of seedlings or to produce additional backcross generations to diminish the number of introgressions in order to fix the target introgression.

Depending on the number of introgressions, plants with three or fewer introgressions were selfed, whereas plants with four or more introgressions were backcrossed to *SL* in order to lower the number of target introgressions. Therefore, the two rounds of selection based on high-throughput genotyping performed in early backcross populations allowed us to obtain a

number of genotypes from which it was feasible to extract single-introgression ILs in just a single generation.

### **3. Developing HRM assays for selected SNPs from GoldenGate Illumina genotyping array**

In order to obtain single-introgression lines from BC<sub>3</sub>S<sub>1</sub> or BC<sub>4</sub> seedlings, each family had to be genotyped with specific SNPs depending on the segregating introgressions. For this task, the high-throughput genotyping technologies used for the early stages of our selection scheme were not cost-effective, and more flexible single-marker technologies were needed. We chose HRM SNP genotyping (Gundry et al. 2003) for subsequent genotyping. A total of 126 SNP markers uniformly distributed across the tomato genome and representing an average of 10 SNPs per chromosome (selected from Sim et al. 2012) were adapted for HRM analysis (**Annex 2-6**). The accuracy of these new markers was checked by genotyping selected BC<sub>3</sub> plants, to find a perfect concordance between the 96-SNP Veracode array and the HRM genotyping.

### **4. Development of single-introgression lines**

Unfortunately, we could not obtain seeds from the entire 41-plant elite group, so we needed to use some of the backup plants, which required additional generations in some cases. The first four ILs were obtained directly from the BC<sub>3</sub>S<sub>1</sub> generation by selfing the BC<sub>3</sub> families that included only one non-target introgression. The rest of the population was completed as follows: 38 BC<sub>3</sub>S<sub>2</sub> ILs, 9 BC<sub>4</sub>S<sub>1</sub> ILs, and 2 BC<sub>3</sub>S<sub>3</sub> ILs. In summary, four ILs were developed in the fifth generation, 47 ILs in the sixth generation, and two ILs in the seventh generation, i.e., overall a lower number of generations were needed than in other similar works, such as the seven generations needed by Schmalenback et al. (2011) in oats, 8-10 generations by Eshed and Zamir (1994) to develop *S. pennellii* ILs, 7-10 generations used by Xu et al. (2010) in rice, or nine generations needed by Finkers et al. (2007) in tomato.

No major problems in marker segregation, fertility, or plant vigor were observed in most families except in two cases. Plants with homozygous introgressions in the proximal region of chromosome 3 (between markers *solcap\_snp\_sl\_63048* and *solcap\_snp\_sl\_26315*) showed low vigor, low fruit setting, and very low seed germination. Also, we observed a strong distorted segregation (DS) towards the *SP* alleles on the distal region of chromosome 4, represented around markers *solcap\_snp\_sl\_47165* and *solcap\_snp\_sl\_47590*. Rick (1966,

1970) reported the first DS in this chromosome 4 region, suggesting that the DS was due to a “gamete eliminator” locus (*Ge*) with three allelic forms (*Ge<sup>P</sup>*, *Ge<sup>c</sup>*, and *Ge<sup>n</sup>*), gametes (both male and female) carrying the *Ge<sup>c</sup>* allele being eliminated in the *Ge<sup>P</sup>/Ge<sup>c</sup>* heterozygous. Given that the DS was always towards *SP*, TO-937 should carry the *Ge<sup>P</sup>* allele and ‘Moneymaker’ the *Ge<sup>c</sup>* allele. Additional cases of DS in tomato have been reported for different genomic regions, usually related to self-incompatibility or unilateral incongruity (Bernacchi y Tanksley 1997; Fulton et al. 1997; Chetelat et al. 2000), although, in the current case, the parent genotypes did not show any of the previous phenomena neither in selfing nor intercrossing.

## 5. Definitive IL collection characterized by HRM assays

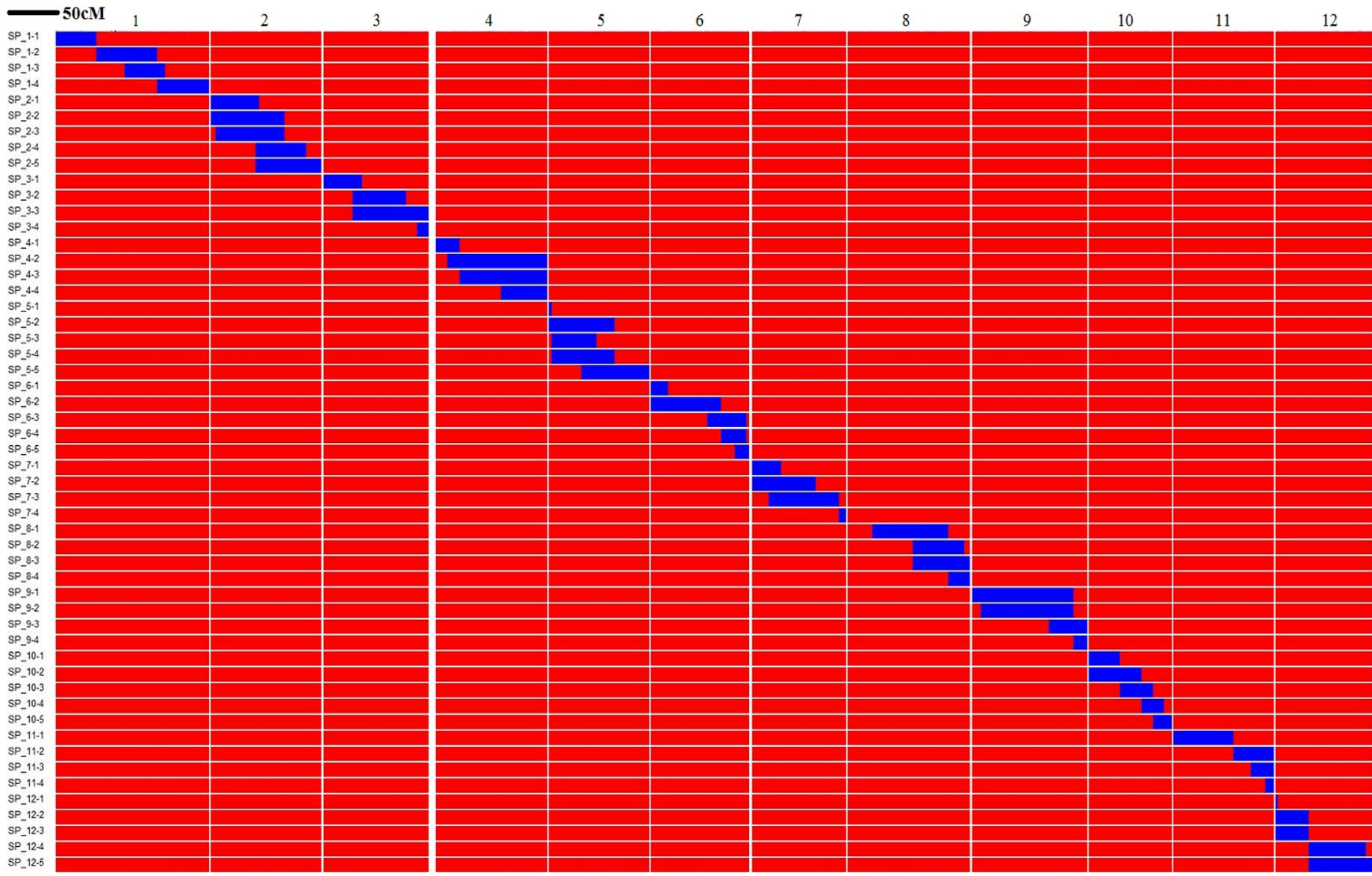
A total of 53 ILs were constructed, each containing an average of 3.7% of the *SP* genome (range 0.5 -7.7%), and altogether covering 98.8% of the donor parent genome. The ILs were named according to the *SP* suffix, a number corresponding to the chromosome on which the introgression was located, and another number to classify the ILs covering different regions of the same chromosome from the top of the chromosome to the bottom (for example, *SP*\_1-1 has the first introgression from the top of chromosome 1). An average of 4.6 ILs/chromosome (range 4-6) was obtained and a total of 58 bins were defined with an average of 4.8 bins/chromosome. Only a small distal region on chromosome 8 was missing. Lack of 100% coverage is common in introgression-line-library development (Monforte and Tanksley 2000a; Tian et al. 2006; Liu et al. 2006; Finkers et al. 2007). Some reasons for this may be self-incompatibility, sterility/lethality (Tian et al. 2006; Keurentjes et al. 2007; Finkers et al. 2007), segregation distortion towards the recurrent genotype (Korff et al. 2004), absence of MAS during at least part of the backcross generations (Tian et al. 2006), inadequate initial population size, or a small number of markers that does not cover the entire genome (Liu et al. 2006; Schmalenbach et al. 2011). In our case, the reason for the loss of coverage was probably the absence of markers in the distal region of chromosome 8 in the early populations, which prevented the selection of *SP* alleles in that region. The average size of the introgression was 51cM (range 3-110 cM) and the entire library has 2X coverage of the *SP* genome (2,385 cM). According to the HRM genotyping, all ILs had a unique introgression, corresponding to the target introgression (**Fig. 2-2**). This result contrasted with other IL collections characterized using similar numbers of markers, where 1 to 4 introgressions per IL were found (Tian et al. 2006; Keurentjes et al. 2007; Finkers et al. 2007).

## 6. Characterization of the IL collection by high-throughput SNP genotyping

In order to improve the resolution and accuracy, the IL collection, together with the parental genotypes ‘Moneymaker’ and TO-937, were genotyped with the 8K SNP SOLCAP Illumina Infinium array (Sim et al. 2012). A total of 4,452 polymorphic markers between the parents were informative in the IL collection, giving an overall polymorphism rate of 59%. An average of 371 SNPs per chromosome were recorded (ranging from 192 SNPs on chromosome 12 to 695 SNPs on chromosome 11), with an average density of 6 SNP/Mb (ranging from 3 to 13 SNPs/Mb on chromosome 12 and 11, respectively).

The high-throughput genotyping assay showed that each of the 53 ILs harbored an average of 4.2% of the *SP* genome, ranging from 0.3% (SP\_5-1) to 8.2% (SP\_3-3), corresponding to an average physical target introgression size of 25 Mb (ranging from 0.7 – 75 Mb). The introgressions altogether covered a total of 94% of the *SP*, and the lack of complete coverage was due to the absence of the aforementioned distal region on chromosome 8 (1.8 Mb, solcap\_snp\_sl\_56370 to solcap\_snp\_sl\_56703) as well as a small region on chromosome 1 (0.7 Mb, solcap\_snp\_sl\_15058 to solcap\_snp\_sl\_24799).

Thus, the general genome composition of the IL genomes estimated by high-throughput genotyping was nearly identical to the estimates based on HRM genotyping. On the other hand, high-throughput genotyping allowed a higher number of bins to be defined: as many as 71 bins (13 bins more than with HRM analysis), with an average *SP* genome size of 10 Mb (range 0.9 – 65 Mb) (**Table 2-1**), as the higher marker density allowed the recombination points to be defined with much greater precision. For example, five bins were defined in chromosome 10 based on HRM genotyping, but seven bins were finally identified after high-throughput genotyping. Furthermore, 37 new non-target introgressions were detected, 18 of which were heterozygous, with the remaining 19 homozygous. However, in more than 90% of the cases, the non-target introgression size was less than 2 Mb, and only six non-target introgressions were larger, with an average size of 54 Mb, located on chromosomes 5, 7, and 9. All these newly detected large introgressions were located in centromeric regions with low recombination that were not covered by the initial genetic map.

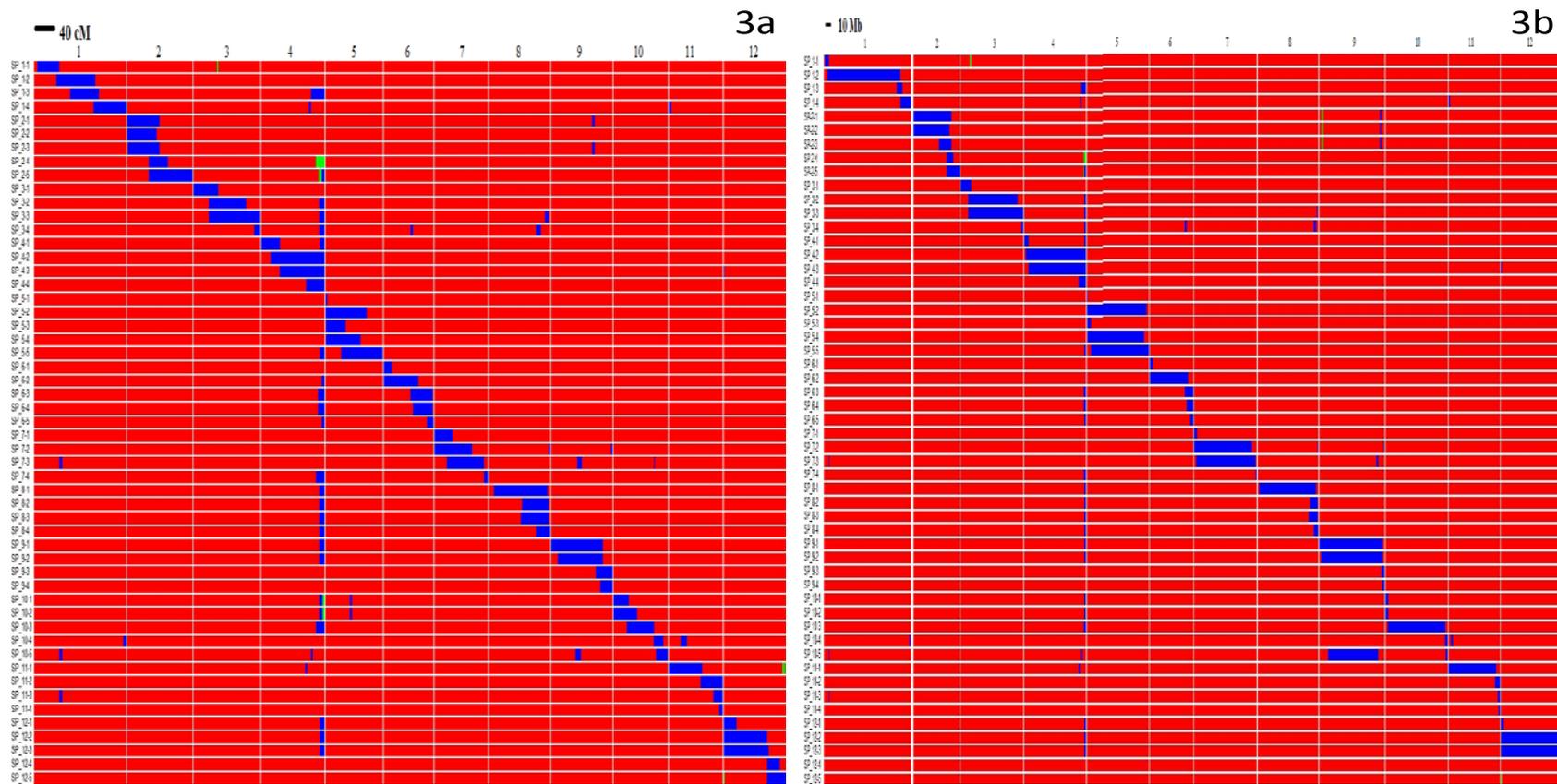


**Figure 2-2. Graphical genotypes of *Solanum pimpinellifolium* TO-937 ILs** in the genetic background of *S. lycopersicum* based on 126 single-nucleotide polymorphism markers. The top rows indicate chromosome number and the left column denotes ILs name. Blue segments represent the homozygous introgressions from TO-937 and the red areas indicate homozygous regions from recurrent parent cv. ‘MoneyMaker’.

Additionally, a non-target introgression located on the distal portion of chromosome 4 was found in 51% of the ILs. This region is linked to the previously observed segregation distortion locus. A recombination between the markers used in the HRM screenings and the distortion locus most likely occurred during the screenings, making it appear that that region had *SL* alleles. The high density of the SOLCAP SNP Infinium array allowed us to uncover the small region on the distal part of chromosome 4 that most likely harbors the distortion locus and which was not detected when using medium resolution maps.

Finally, a last MAS step was performed to remove 14 non-target heterozygous introgressions in the SP\_4-2, SP\_4-4, SP\_5-1, SP\_5-5, SP\_6-5, SP\_7-3, SP\_8-1, and SP\_11-3 introgression lines, resulting in the IL library whose final graphic genotype is shown in (**Figure 2-3a,b**). Comparing genetic and physically based graphical genotypes (**Figures 2-3a and 2-3b**, respectively), the differential recombination rate across the genome became apparent (Sato et al. 2012; Sim et al. 2012). Thus, the non-target introgression on chromosome 9 observed in IL SP\_10-5 represent around 1/10 of the genetic map of chromosome 9, whereas it represents around 1/3 of the physical map. On the other hand, the introgressions located in euchromatic regions (for example, non-target introgression on chromosome 1 observed in ILs SP\_7-3, SP\_10-5, and SP\_11-3) seem relatively smaller in the physical map than in the genetic map.

The presence of non-target introgressions was likely due to double recombination events between the markers used for population development, which went undetected until higher density genotyping technologies were applied. This effect was also observed by other authors after their IL collections were characterized with high-throughput genotyping. Thus, Eichten et al. (2011) and Pea et al. (2013) developed maize ILs without controlling the genetic background to finally find, after high-throughput genotyping, that the number of introgressions were still relatively large (as many as 15 introgressions per IL), thereby demonstrating the need to monitor the genetic background during the introgression procedure in order to ensure appropriate isogenity among ILs. On the other hand, monitoring genetic backgrounds with genetic markers at medium density, even though it certainly reduces the



**Figure 2-3. Graphical genotypes based on genetic distance (figure 2-3a) and physical distance (figure 2-3b) of *Solanum pimpinellifolium* acc. TO-937 introgression lines in the genetic background of *S. lycopersicum*. The top rows indicate chromosome number and the left column denotes ILs name. Blue segments represent homozygous introgressions from TO-937, green areas heterozygous, and red areas indicate homozygous regions from recurrent parent cv. 'Moneymaker'. The 53 ILs have been characterized using the 8K Illumina Infinium SOLCAP array (Sim et al. 2012).**

number of non-target introgressions, does not prevent a certain level of non-isogenity. For example, Xu et al. (2010) developed a library of 128 rice introgression lines, finding 117 non-target introgressions, with 50% of the population having two or more non-target introgressions that were not detected previously, using markers with an average density of 1.6 Mb/marker. Similarly, Schmalenbach et al. (2011) found that 76% of their barley ILs contained one to four additional previously undetected introgressions after high-throughput genotyping. In the current report, the number and size of non-target introgressions was significantly lower than those reported. We consider the larger % of isogenicity in our population to be a consequence of one or more of the following factors: 1) the high density of the genotyping array used in the BC<sub>2</sub> population allowed plants with large intact introgressions to be selected and thereby diminish the probability of dragging small introgressions as a consequence of double recombination among markers, 2) the high selection pressure applied at the BC<sub>2</sub> and BC<sub>3</sub> seedling stage, which allowed most of the ILs to be extracted from the BC<sub>3</sub> level, reduced the number of possible recombinations, and, therefore, maintained the integrity of target and non-target introgressions, and 3) the utilization of two consecutive high-throughput genotypings in BC<sub>2</sub> and BC<sub>3</sub>, which ensured a high-quality control of the background genome.

## **7. The SP introgression line genomic library as a tool for improving genome assembly**

The graphic genotypes and the bin positions of 4,452 polymorphic SNPs were compatible with the physical map positions reported in the 8K SNP Tomato Illumina Infinium array (Sim et al. 2012), except for three regions on chromosome 3 corresponding to chromosome segments 3\_2 (positions SL2.40ch3:2500006 to SL2.40ch3:5671056), 3\_3 (positions SL2.40ch3:6655807 to SL2.40ch3:6930668), and 3\_12 (positions SL2.40ch3:50138804 to SL2.40ch3:54324034). According to the bin mapping of the SNPs included in these chromosome segments, using the ILs SP\_3-1, SP\_3-2, SP\_3-3, chromosome segments 3-2, 3-3 and 3-12 should be located in a different order along tomato chromosome 3 (**Annex 2-7**).

## **8. Integration of high-throughput genotyping strategies in IL development**

The results presented in the current report confirm the previous suggestions by Sato et al. (2011) and Pea et al. (2013) that the genome composition of intermediate materials should be monitored during IL development in order to maximize the isogenicity of the background genome. We found the number of markers used to be most important for this aim in addition

to the genetic and physical location of those markers. In our case, the markers were initially selected from a genetic map, so markers separated by several Mb in the centromeric regions co-segregated in the genetic map, meaning that large centromeric regions were not covered by markers in subsequent generations. This may be why the IL SP\_10-5 carried a large centromeric region from chromosome 9. On the other hand, the selection of markers based solely on their physical position would also introduce a bias in the design of SNP arrays as centromeric regions would be over-represented, whereas SNP density in euchromatic regions would be lower than average. The over-representation of SNPs in centromeric regions would not be cost-effective; in fact, using a relative low number of such markers would be sufficient to ensure the integrity through the generations because of the low recombination ratio in those regions. The lower density on euchromatic regions would increase the probability of undetected double recombinants between adjacent markers, and, therefore, diminish the genetic background isogenicity. The integration of physical and genetic maps would allow a most effective genotyping assay design, selecting markers evenly distributed along the genetic map which would guarantee a low probability of selecting double-recombinants among adjacent markers, as well as markers in telomeric and centromeric regions, in order to ensure the presence of chromosomal distal regions and to avoid the drag of large centromeric regions. High-throughput genotyping should be done in early generations in order to avoid double recombinants and to guarantee the integrity of the target introgression fragments. Also, additional rounds of high-throughput genotyping are desirable in order to control the genetic background more efficiently. This strategy would not only increase the quality of the introgression line library but also reduce the time for its development, i.e., in the current report, most ILs were developed from a BC<sub>3</sub>S<sub>1</sub> generation (a total of five generations), whereas other populations needed a larger number of generations, such as the 10 generations needed by Eshed and Zamir (1994), nine by Xu et al. (2010), seven by Schmalenbach et al. (2011).

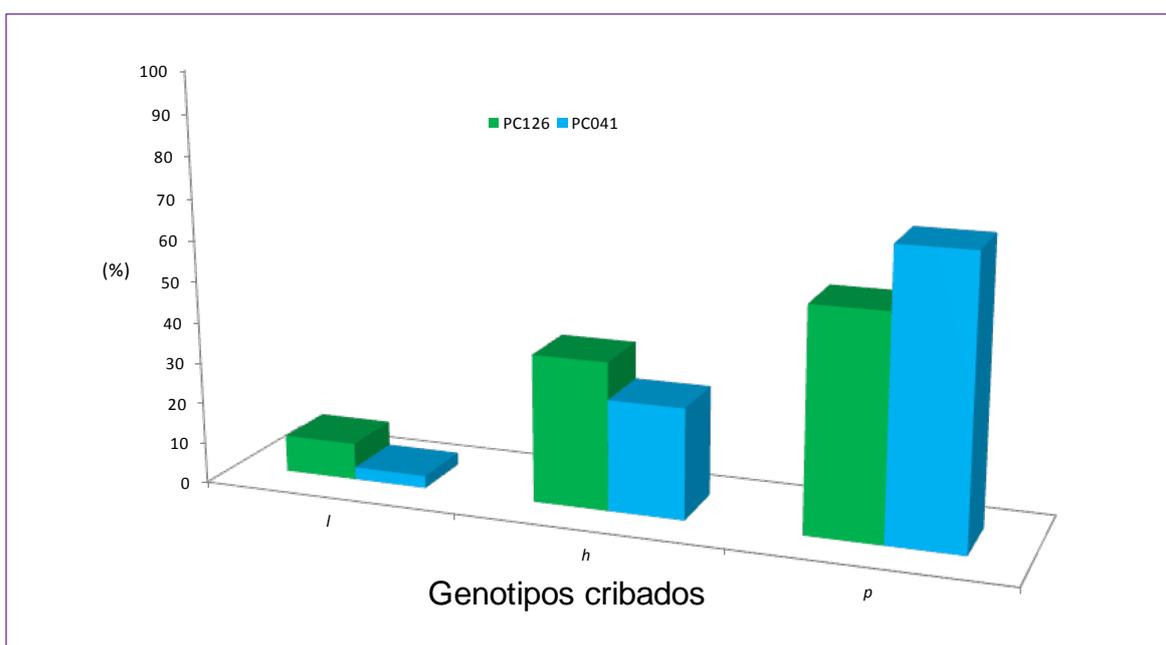
*S. pimpinellifolium* germplasm contains genetic variability that obviously cannot be represented in a single IL collection like the current one (**Annex 2-2**, Blanca et al. 2012). The fact that an important number of *S. pimpinellifolium* accessions are currently genotyped with the SOLCAP Infinium array (Blanca et al. 2012; Sim et al. 2012) will facilitate the development of ILs from other accessions, taking advantage of the genomics tools available for this species.

In summary, in the current report we have shown that the implementation of high-throughput genotyping technologies in IL development schemes significantly reduces the number of generations and increases the genetic background isogenity among ILs, which increases the accuracy of QTL mapping in these populations. The cost of high-throughput genotyping has been decreasing gradually over the last few years, so the task of IL development has become more affordable, and we should therefore expect a larger number of IL collections to be obtained in the coming years.

## CAPÍTULO 3: Mapeo fino de una región con segregación distorsionada en el extremo distal del brazo largo del cromosoma 4

### Introducción

Durante el proceso de desarrollo de las líneas de introgresión, en la etapa de cribado de plántulas de la familia 11M47 (población  $BC_3S_1$ ) que presentaban una introgresión no-diana en el extremo distal del cromosoma 4, nos encontramos con un patrón de segregación en los marcadores de esta región que favorecía los alelos de *S. pimpinellifolium* acc. TO-937 en detrimento de alelos de *S. lycopersicum* "MoneyMaker" (MM) (**Fig. 3-1**), lo que impidió eliminar la introgresión no-diana en dicha familia. Esta misma situación se presentó en otras familias que también portaban la introgresión no-diana en la misma región del cromosoma 4. Tras genotipado final realizado a la genoteca completa de ILs con la micromatriz de 8 K SNP, se pudo cuantificar el alcance de este fenómeno y, cerca del 50% de las ILs contienen el fragmento de *S. pimpinellifolium* de dicha región (**ver anexo 3-1**). Esta situación nos motivó a estudiar este fenómeno con más detalle, con el objetivo final de averiguar la base genética molecular de este fenómeno. En este capítulo describimos el mapeo fino de la región con segregación distorsionada (SD) del cromosoma 4 y proponemos posibles genes candidatos.



**Fig. 3-1. Porcentajes de genotipos encontrados en el cribado con HRM de una población de 200 plantas  $BC_3S_1$ , con una introgresión no-diana en la región distal del cromosoma 4.** Los marcadores SNPs usados fueron PC126 y PC041. Donde *l* representa el genotipo homocigoto de MM, *h* el genotipo heterocigoto y *p* el genotipo homocigoto de TO-937.

## **1. Segregación distorsionada (SD)**

La segregación distorsionada (SD) ocurre cuando la frecuencia de genotipos/ alelos observados se desvía significativamente de las frecuencias mendelianas esperadas para una población segregante. La SD fue descrita por primera vez en plantas de maíz por Mangelsdorf et al. (1926). Posteriormente, se han descrito ejemplos de SD en un gran número de especies (Bintou et al 2012): tomate (Rick 1966), arroz (Nakagahra 1972), trigo (Endo 1990), cebada (Graner et al. 1991), algodón (Rooney et al. 1991), sorgo (Pereira et al. 1994), alfalfa (Echt et al. 1994), cítricos (Luro et al. 1995), café (Ky et al. 2000), garbanzo (Castro et al. 2011). También se han descrito casos en *Drosophila*, etc. confirmando que la SD es un fenómeno común a todos los seres vivos (Taylor et al. 2003).

La SD suele detectarse durante el proceso de construcción de poblaciones de mapeo (García-Dorado y Gallego 1992; Lorieux et al. 1995a; Lorieux et al. 1995b). Por lo general, los loci implicados en SD se detectan al observar un grupo de marcadores ligados que muestran esta SD debido a ligamiento con el gen causal, de manera que los marcadores más cercanos muestran mayor SD que los más alejados (Xu et al. 1997; Li et al. 1997; Song et al. 2006). Al igual que ha ocurrido para otros caracteres/procesos con base genética, la utilización de marcadores ha sido una herramienta muy útil en el mapeo de regiones con SD. La SD se ha detectado en todo tipo de población recombinante, como F<sub>2</sub>, retrocruces (BC), dobles haploides (DH), líneas consanguíneas recombinantes (RILs) y líneas de introgresión (ILs). En poblaciones DH, genes letales en condición homocigota recesiva muestran un alto nivel de SD (Song et al. 2006). Por ejemplo trabajando en café, Lashermes et al. (2001) encontró 40% más de marcadores que mostraban SD cuando se utilizaban líneas DH que cuando se trataba de cruces de los mismos clones. Resultados similares encontró Song et al. (2005) en algodón y Yamagishi et al. (2010) en arroz. Investigaciones realizadas en sorgo, tomate, maíz, repollo y algodón entre otros, han demostrado que la SD suele ser más acentuadas cuanto más distantes sean las especies cruzadas (Xu et al. 1997; Lu et al. 2002; Li et al. 2007).

## **2. Causas de la SD: bases fisiológicas, estructurales o moleculares**

La SD de un carácter o marcador puede deberse a una multitud de factores que pueden tener origen genético, ambiental o ser producto de la interacción entre ambos (Xu et al. 1997; Vancentovic, 2008; Buchi and Mergeai 2012). Se han descrito muchos factores genéticos causante de SD (Buchi and Mergeai 2012) que van desde diferencias estructurales entre los cromosomas que podrían conducir a una amplia gama de anomalías citogenéticas hasta

incompatibilidades alélicas concretas. Dentro de los factores ambientales, Wang et al. (2004) menciona la presión de selección que se ejerce sobre alelos menos adaptados a condiciones ambientales específicas o la selección debido a la endogamia. Por su parte, Livingstone et al. (1999) menciona la deriva genética producida durante la gametogénesis como factor ambiental que puede conducir a SD. Por todo ello es difícil separar los efectos genéticos y los factores ambientales, siendo a menudo la SD una consecuencia de la interacción de ambos (Xu et al. 1997; Vancentovic 2008).

Las alteraciones que conducen a la SD puede producirse antes o después del desarrollo de cigoto y en consecuencia se podrían clasificar en precigóticas o postcigóticas. Antes de la formación del cigoto, la SD estaría ocasionada por alteraciones en las proporciones de los alelos en los gametos y consecuentemente produciendo cambios en las proporciones genotípicas de cigotos, como por ejemplo alteraciones que resultan en la letalidad o aborto del polen (Rick 1966), la desigualdad de la capacidad de los tubos polínicos en la polinización (Mangesldorf et al. 1926), la fertilización preferencial (Endo 1990). El aborto del polen y la competencia de los tubos polínicos son ejemplos de selección de gametos masculinos, y suelen ser más frecuentes en gametos masculinos que en femeninos (Xu et al. 1997; Song et al. 2006). Sin embargo, la SD puede ocurrir en ambos gametos (Rick 1966). Después de la fecundación la SD vendría asociada a alteraciones en las proporciones genotípicas del cigoto (Lorieux et al. 1995; Song et al. 2006), resultantes de que los genotipos segregados sean por ejemplo incompatibles con el endospermo, o produzcan híbridos inviables o débiles, o semilla estéril (Talluri 2012) o los cigotos sean selectivamente eliminados (Rick 1963).

### **3. SD en tomate**

En múltiples poblaciones de mapeo desarrolladas en tomate se ha observado la presencia de regiones genómicas en donde la recombinación se suprime y/o las frecuencias alélicas observadas son diferentes a las esperadas. La frecuencia de estos fenómenos aumenta cuanto más distantes filogenéticamente sean las especies (Anderson et al. 2010). En las últimas décadas, muchos autores han observado SD trabajando con poblaciones de tomate que implicaban diferentes especies silvestres relacionadas. Por ejemplo: Pelham (1968) y Fulton et al. (1997) encontraron un número excesivo de individuos heterocigotos asociados a marcadores del cromosoma 9 de la especie *S. arcanum*, atribuyendo este comportamiento a la acción de un gen promotor de gametos (*Gp*), con una herencia preferencial de los alelos de *S. arcanum* frente a los alelos de *S. lycopersicum*, sin embargo, no encontraron evidencia concluyente sobre la relación entre la distorsión y *Gp*. Resultados similares fueron reportados

por Chetelat et al. (2000) en cruzamientos de tomate con *S. lycopersicoides*, aunque en este estudio se observó SD en detrimento de heterocigotos en los cromosomas 2 y 5 y hacia heterocigóticos en el cromosoma 6 y 9. Otro ejemplo lo constituye la población F<sub>2</sub> derivada del cruce entre *S. lycopersicum* x *S. pennellii* (de Vicente y Tanksley 1993) ellos encuentran un 80% de marcadores con SD, la mayoría favoreciendo alelos silvestres. Así mismo, Bernacchi y Tanksley (1997) reportaron un 15% de marcadores con SD, en una población BC<sub>1</sub> desarrollada a partir de las especies *S. lycopersicum* x *S. habrochaites*, las desviaciones de las frecuencias alélicas se producían hacia los dos parentales, dependiendo de la región cromosómica.

Especial atención merece el fenómeno de segregación distorsionada descrito por Rick (1966, 1970), este autor encontró un sesgo drástico en la segregación en la región proximal a la heterocromatina del cromosoma 4 al analizar diferentes cruzamientos en 43 genotipos de tomate comercial y 11 genotipos de *S. pimpinellifolium*. Rick definió que la causa de la SD era un gen al que denominó "gen eliminador de gametos" (*Ge*), y propuso que el gen presentaba tres formas alélicas (*Ge<sup>p</sup>*, *Ge<sup>c</sup>* y *Ge<sup>n</sup>*). En plantas heterocigóticas (*Ge<sup>p</sup>Ge<sup>c</sup>*), el alelo *Ge<sup>p</sup>* (presente mayoritariamente en accesiones de *S. pimpinellifolium*) es seleccionado preferentemente sobre *Ge<sup>c</sup>* (predominante en *S. lycopersicum*), ocurriendo por igual tanto en los gametos masculinos y femeninos, con una penetración cercana al 95%, generando una fuerte SD en la descendencia. Por otro lado el alelo *Ge<sup>n</sup>* es considerado neutro y no produce SD con ninguna otra combinación alélica. La literatura menciona otros genes con efectos similares al gen "eliminador de gametos" en otras especies vegetales (**Anexo 3-2**).

#### **4. Causas de la SD en tomate**

La SD podría deberse a mecanismos moleculares similares (al menos desde el punto de vista de lógica molecular) a las utilizadas por las plantas para impedir cruzamientos no deseados, relacionados con barreras de cruzamiento desarrolladas entre tomate y las especies relacionadas. La principal barrera precigótica que influencia el éxito o fracaso en los cruzamientos en tomate son las interacciones polen-estilo (Bedinger et al. 2011), existiendo diferencias en los mecanismos de rechazo del polen según el sistema de cruzamiento que tengan las especies (Covey et al. 2010; Li et al. 2010). En el **Anexo 3-3** se mencionan algunos de los genes implicados en crecimiento del polen y guía hacia el ovulo.

En el tomate y en las *Solanáceas* en general los sistemas de cruzamiento van desde especies en auto-compatible (AC) hasta auto-incompatibles (AI) existiendo también especies

facultativas y que presentan incompatibilidad unilateral (IU). Los mecanismos que regulan la AI están bien estudiados (**Anexo 3-4**), se sabe que son controlados por el locus *S* que codifica una S-ribonucleasa (S-RNasa) expresada en el pistilo (McClure et al. 1998) y el locus-S F-Box (SLF) proteínas expresadas en el polen (Sijacic et al. 2004), de manera que la interacción entre las proteínas presente en estas estructura reproductivas son las que definen si ocurre o no la fecundación. De los mecanismos que regulan la IU se han logrado mapear dos factores del polen que intervienen directamente en la incompatibilidad unilateral: el factor *ui1.1* localizado en el locus *S* del cromosoma 1 y el factor *ui6.1* localizado en el brazo corto del cromosoma 6 ambos en *S. pennellii* (Li et al. 2010). Igualmente se han mapeado tres QTLs asociados con la IU a nivel del pistilo. Un QTL mayor que coincide (igual *ui1.1*) con el locus *S* en el cromosoma 1 y dos QTLs con efectos menores localizados en el cromosoma 3 y 12 respectivamente (Bernacchi y Tanksley 1997). Recientemente, Li et al. (2010) propone al gen *CUL1* (cuya actividad es proteger el crecimiento de los tubos polínicos contra la acción de las S-RNasas del pistilo), presente en especies AI, como el responsable que el polen proveniente de estas especies se puede cruzar con especies AC. Las especies auto-compatibles que han sufrido la pérdida de la actividad de este gen (por una delección en el intrón 7) no les permite funcionar como donadores de polen, debido a que este es destruido por las S-RNasas de plantas con genotipos con el locus *S* activo. La especie *S. pimpinellifolium* presenta ambos alelos lo que explicaría que sea facultativa (accesiones alógamas y autógamas) y, que además presenten IU. Esta evidencia aunada a la gran variabilidad genética de esta especie (Rick 1970) hace sugerir que la mutación del gen *CUL1* se haya originado en esta especie.

En cuanto a loci involucrados en barreras poscigótica, se han utilizado ILs para detectar QTLs relacionados con la viabilidad del polen y la esterilidad semillas de híbridos interespecíficos. Por ejemplo Eshed y Zamir (1995) usando ILs de *S. pennellii* detectaron un mínimo de 10 QTLs relacionados con la esterilidad del polen de los híbridos. Con esta misma genoteca los autores Moyle y Nakazato (2008) describen 4 QTLs relacionados con fallos en la germinación de semillas. En otro estudio Moyle y Graham (2005) usando líneas introgresión de *S. habrochaites* asociaron 12 y 5 QTLs relacionados con esterilidad de polen y semillas de los híbridos respectivamente. Según Quiros (1991) la presencia dentro de los híbridos de tomate de esterilidad, segregación distorsionada y reducción de las recombinaciones también pueden deberse a diferencias estructurales en los cromosomas. Las irregularidades sinápticas que se observan en los híbridos son más numerosas y graves cuanto más alejada filogenéticamente estén las especies cruzadas (Anderson et al. 2010).

## **Materiales y Métodos**

### **1. Material vegetal y extracción de ADN**

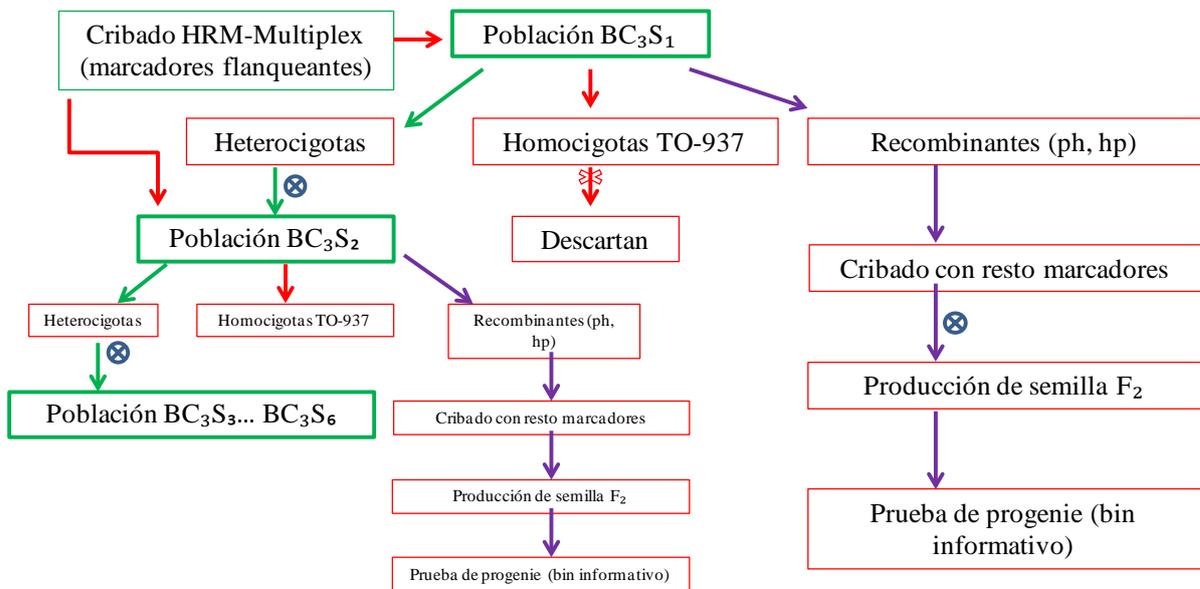
El material vegetal utilizado inicialmente fueron 200 plantas pertenecientes a una población de retrocruce avanzado ( $B_3CS_1$ ), que formaba parte de la colección pre-ILs procedente del cruce entre *S. lycopersicum* cv "MoneyMaker" x *S. pimpinellifolium* acc. TO-937 y que fueron utilizadas durante el desarrollo de la colección de líneas de introgresión. En esta población original, se seleccionaron genotipos heterocigotos en la región que presentaba la distorsión y se cultivaron bajo condiciones de invernadero, dejándose en autofecundación espontánea para generar plantas recombinantes para las siguientes etapas. Esta operación se repitió en sucesivas ocasiones como se muestra en la **Figura 3-2**, de tal forma que, la población final de mapeo estuvo compuesta por cerca de 2000 plantas distribuidas en un ámbito de generación desde la  $BC_3S_1$ , hasta la  $B_3CS_6$ .

Adicionalmente, se seleccionaron cinco accesiones de *S. pimpinellifolium* (PE\_7, LA0722, LA1589, BGGV006507, BGGV007168) y una accesión de *S. galapagense* (LA0530) provenientes del banco de germoplasma del COMAV-UPV y de la Estación Experimental La Mayora (**ver apartado 3**). Estas accesiones fueron cruzadas con el cv. "MoneyMaker" para generar poblaciones  $F_1$  y posteriormente por autofecundación poblaciones  $F_2$  con el objetivo de estudiar la variabilidad alélica en el locus SD.

El ADN genómico fue aislado de hojas primarias jóvenes siguiendo el método Doyle y Doyle (1990) (descrito en el Cap. II).

### **2. Definición y mapeo de la región con SD**

La SD fue investigada en varias poblaciones derivadas del mismo cruzamiento:  $BC_2$ ,  $BC_3$  seleccionadas y RILs, contrastando las segregaciones en marcadores ligados al loci SD con las esperadas en cada generación. Adicionalmente, el análisis de la SD en la población de RILs se utilizó para realizar un primer mapeo del locus aprovechando que se disponía de un mapa de alta densidad (>4000 SNPs, Fernández-del-Carmen et al. 2011) en esta población con SNPs realizados en nuestro grupo y por la plataforma SOLCAP (Slim et al. 2012).



**Fig. 3-2. Esquema utilizado para generar, genotipar y evaluar las poblaciones utilizadas para estudiar la región con SD del cromosoma cuatro.** (ph y hp: son recombinantes con genotipo TO-937 (*p*) en un marcador flanqueante y heterocigoto en el otro (*h*), respectivamente).

Aproximadamente 2000 plantas conformaron la población final de mapeo sobre la cual se realizó la búsqueda de recombinantes. Esta población fue cribada por medio de la técnica de genotipado High Resolution Melting (HRM) (Gundry et al. 2003), empleada en el desarrollo de las ILs y explicada ampliamente en el Capítulo II. El cribado de los individuos se realizó por medio de marcadores SNPs (**Anexo 3-5**) localizados en la región flanqueante de la zona que presentaba la distorsión, las plantas que presentaban eventos de recombinación en la región bajo estudio se genotiparon con marcadores localizados dentro del intervalo inicial para definir la localización del evento de recombinación (**Figura 3-2**). Los recombinantes informativos (los que tenían recombinaciones en diferentes intervalos entre los marcadores desarrollados) se autofecundaron para recolectar la semilla F<sub>2</sub>. La determinación del genotipo para el locus SD en las plantas F<sub>2</sub> seleccionadas se realizó mediante pruebas de progenie, al menos 50 plántulas de la descendencia de cada planta fueron genotipadas por HRM con un marcador localizado en la región heterocigota del recombinante (**Figura 3-2**) y la segregación esperada (1:2:1) se contrastó con un test Chi-cuadrado ( $X^2$ ), en el caso de que se observara SD significativa, el locus SD se localizaría en la región heterocigótica segregante, y en el caso de no observarse SD, se localizaría en la región homocigótica.

Finalmente, se realizó una búsqueda en la región genómica definida tras las pruebas de progenie de los recombinantes informativos de genes candidatos en la versión del genoma de tomate depositada en <http://solgenomics.net/> de acuerdo con la anotación realizada por The International Tomato Annotation Group (ITAG).

### **3. Variabilidad alélica en la región con SD en especies relacionadas de *Solanum***

En un estadio intermedio de este objetivo se definió la localización del gen responsable de la SD entre los marcadores *solcap\_snp\_sl\_3920* y *solcap\_snp\_sl\_3977* (separados por 229 Kb) en el cromosoma 4. Se integraron los datos de genotipado obtenidos por Sim et al. (2012), Blanca et al. (2012) y propios de nuestro grupo de investigación con la misma micromatriz 8K SNP Infinium de Illumina desarrollado por la iniciativa SOLCAP. Se calculó la distancia de Nei (1983) entre los genotipos y el dendograma "Neighbour-Joining" con el programa Powermarker (Liu y Muse 2005) que se dibujó con MEGA 6 (Tamura et al. 2013).

Se seleccionaron accesiones representativas de cada uno de los diferentes grupos definidos en el dendograma para investigar el efecto de la diversidad genética en este locus sobre la SD. Las accesiones seleccionadas se cruzaron con MM para generar poblaciones F<sub>2</sub> en las que se determinó (tal como se explica en el apartado 2) si se observaba SD estudiando la segregación de marcadores ligados al locus determinados previamente, contrastando con la segregación esperada con una prueba Chi-cuadrado ( $\chi^2$ ).

### **4. Identificación y diseño de marcadores moleculares (SNPs) para las diferentes etapas de mapeo de la región SD**

Los primeros marcadores SNPs que se usaron provenían de un mapa de alta densidad (> 4.000 SNPs), en la población de RILs a partir del mismo cruzamiento. A pesar de la alta densidad del mapa, fue necesario desarrollar nuevos marcadores SNPs en zonas no cubiertas por los marcadores anteriores. Para ello se utilizaron secuencias de TO-937 obtenidas por RNA-Seq mediante secuenciación masiva con SOLID (Applied Biosystems), a partir de una preparación de RNA de una mezcla de tejidos y órganos en diferentes estadios de desarrollo (resultados no publicados). Se utilizó el programa CLCBIO (CLC Genomics Workbench <http://www.clcbio.com/>) para identificar los nuevos SNPs, por medio de la comparación de las secuencias de RNA-Seq de TO-937 con la secuencia del genoma de referencia de tomate (The Tomato Genome Consortium 2012) en la región con SD del cromosoma 4.

## Resultados

### 1. Mapeo inicial de la región con SD

La existencia de una región con segregación distorsionada fue detectada por primera vez durante el proceso de desarrollo de la genoteca de líneas de introgresión concretamente realizando el cribado de una introgresión heterocigota no-diana del cromosoma 4 en una familia de retrocruce avanzado ( $BC_3S_1$ ). La SD se investigó también en otras poblaciones del mismo cruzamiento: retrocruces  $BC_2$  (81 plantas),  $BC_3$  (121 plantas) y población de RILs (182 plantas). Los resultados que se presentan en la (**Tabla 3-1**) demuestran que en todos estos casos las segregaciones observadas se diferencian estadísticamente de las esperadas. Dado que las poblaciones de retrocruzamiento se desarrollaron siempre con polen de MM, la SD estaría afectando a los gametos femeninos, sin descartar que también estuviera afectando a los masculinos. Con esta primera aproximación, la definición del tamaño y posición de la región con SD no fue muy precisa dado la carencia de marcadores en la región. Con la disponibilidad de más marcadores provenientes del diseño de la micromatriz de SOLCAP<sup>®</sup> Sim et al. (2012), se diseñaron oligonucleótidos a partir de estos marcadores para usarlos en el genotipado por medio de HRM (individuos recombinantes en el intervalo entre los marcadores PC026 y PC041) de la población  $BC_3S_1$  (**Tabla 3-2**). Pudimos observar que a medida que el cribado avanzaba hacia el extremo distal del cromosoma 4 los alelos de *S. lycopersicum* disminuían drásticamente y como consecuencia los genotipos homocigotos MM (*l*) desaparecen por completo y solo se encuentran plantas heterocigotas (*h*) u homocigotas *S. pimpinellifolium* (*p*) para esa región. Paralelamente, gracias a un genotipado de alto rendimiento en la población de RILs con la micromatriz de SOLCAP (**Fig 3-3**), se observó una región del cromosoma 4 donde el genotipo de todas las RILs (182 plantas) era homocigóticos *p*. Combinando los resultados de las RILs y de la población  $BC_3S_1$  se definió la región con SD, delimitada en el extremo proximal por el marcador `solcap_snp_sl_47697` posición física de 62.8 Mb y en el extremo distal por el marcador `solcap_snp_sl_47749` posición física de 63.2 Mb, lo que representa un tamaño físico de 380 Kb (**Fig. 3-3**). Nos propusimos a buscar en esa región recombinantes para mejorar la resolución de la posición de la SD.

### 2. Mapeo fino de la región con SD

Una vez se delimitó la región que presentaba segregación distorsionada, se procedió al mapeo fino de la misma. Esta etapa se realizó en tres pasos: primero se buscaron marcadores SNPs

para saturar adecuadamente la región a mapear, segundo se buscaron los recombinantes por medio de HRM-múltiple, y finalmente se realizó una prueba de progenie sobre una F<sub>2</sub> de los recombinantes informativos con objeto de verificar si contenían o no los alelos de que provocan la SD

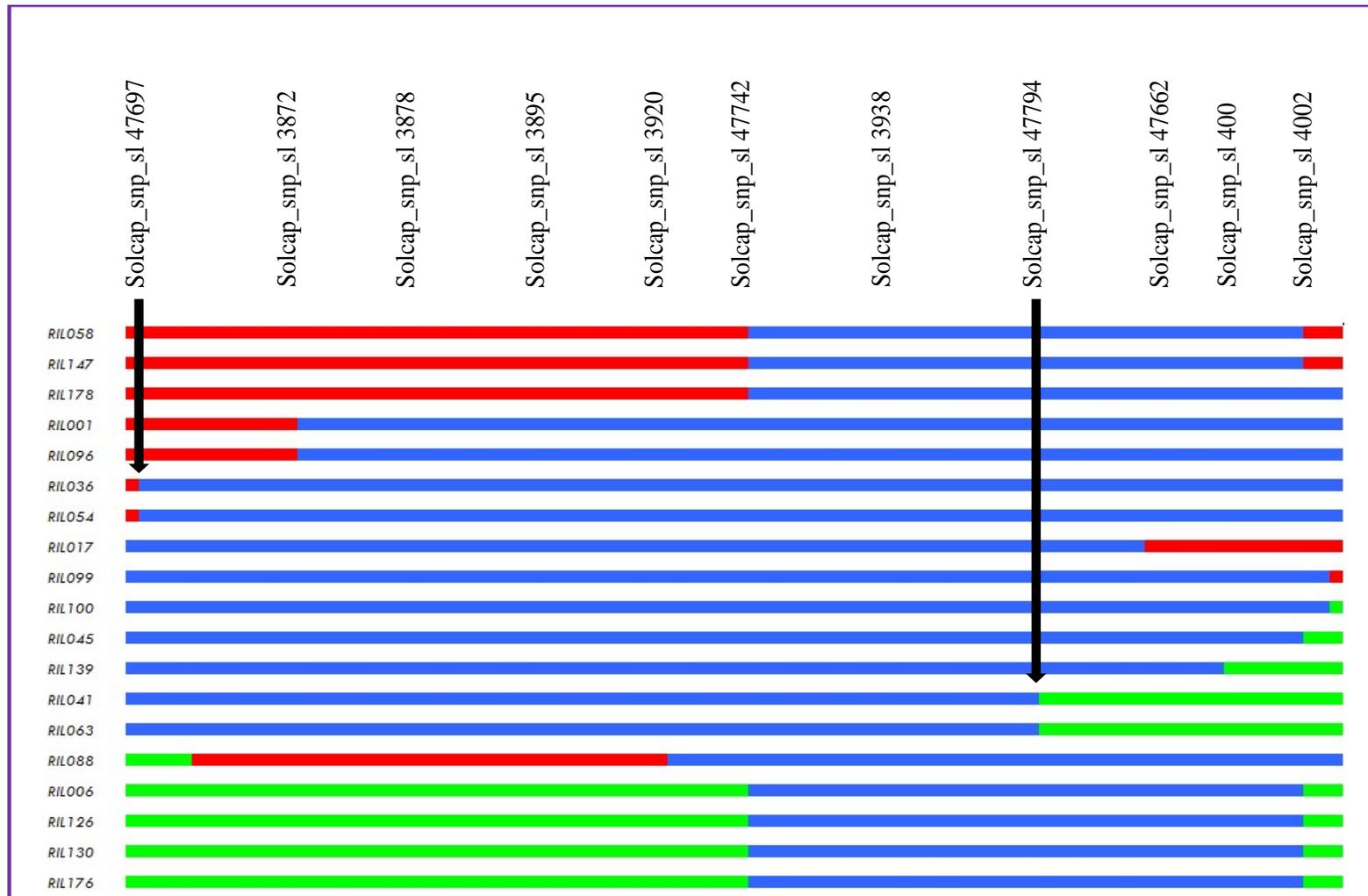
Población	Segregación (genotipos) Esperados (%)			Segregación (genotipos) observados (%)			X <sup>2</sup>	Resultado
	<i>l</i>	<i>h</i>	<i>p</i>	<i>l</i>	<i>h</i>	<i>p</i>		
RILs	42	16	42	11	6	83	223**	SD
BC <sub>2</sub>	75	25	0	33	66	0	64**	SD
BC <sub>3</sub>	50	50	0	40	60	0	5*	SD
BC <sub>3</sub> S <sub>1</sub>	25	50	25	6	31	63	165**	SD

**Tabla 3-1. Segregaciones esperadas y observadas y prueba de Chi-cuadrado en el cromosoma 4 de diferentes poblaciones desarrolladas a partir de MM x TO-937.** Los porcentajes corresponden a genotipado con HRM con los marcadores PC026 y PC041 de 182 RILs, 81 BC<sub>2</sub>, 121 BC<sub>3</sub>, 200 BC<sub>3</sub>S<sub>1</sub>. La letra *l*: genotipo homocigoto MM, *h* genotipo heterocigoto y *p* genotipo homocigoto TO-937, SD: Segregación distorsionada (\*p<0.05, \*\* p<0.0001).

## 2.1 Desarrollo de nuevos marcadores

Con el objetivo de saturar con marcadores la región SD del cromosoma 4 (**Fig. 3-4**), se desarrollaron un total de 33 marcadores nuevos (**Anexo 3-2**). Un primer grupo 23 marcadores se diseñaron a partir del mapa de alta densidad de la población de RILs, obteniéndose una densidad de 0.1 cM entre marcadores (**Fig. 3-4**). A pesar de este significativo avance, quedaba un fragmento importante en el centro de la región con SD delimitado por los marcadores solcap\_snp\_sl\_47742 en el extremo proximal y solcap\_snp\_sl\_47749 en el extremo distal con un tamaño físico de 160 Kb sin cobertura.





**Figura 3-3. Resumen del genotipado gráfico de la región con SD del cromosoma 4 en la población de RILs**, realizado con la “micromatriz de SOLCAP”. En la parte superior marcadores SOLCAP de la región mapeada y en lado izquierdo el nombre de las RILs. Las líneas negras indican el intervalo seleccionado para la búsqueda de recombinantes. Las barras de color rojo representan fragmentos homocigotos de MM (*S. lycopersicum*), las de color verde fragmentos heterocigotos y las azules fragmento homocigotos de TO-937 (*S. pimpinellifolium*)

Para lograr dar cobertura a esta región, se utilizaron secuencias de RNA-Seq del parental silvestre TO-937 obtenidas en nuestro grupo mediante secuenciación con SOLiD™. Con la ayuda del programa CLCBIO, se alinearon las secuencias de RNA-Seq de TO-937 de la región con SD contra las secuencias de la misma región del parental de referencia (Heinz 1706). Se lograron descubrir 10 nuevos marcadores SNPs, que presentaban un número de lecturas adecuadas, lo que permitió reducir la brecha entre SNPs en esta región a 16 Kb (**Fig. 3-4**).

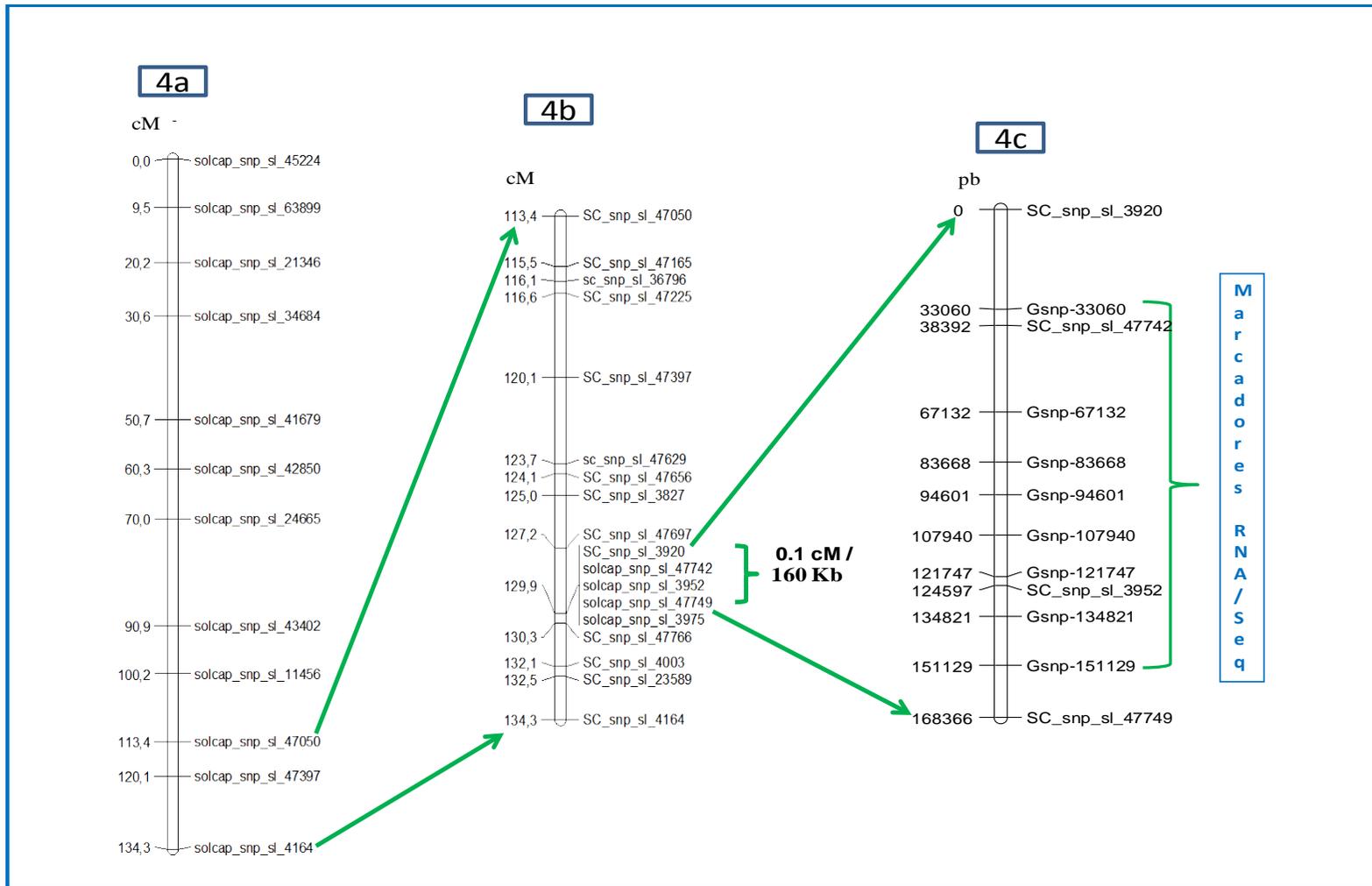
El mapa final de trabajo comprendido entre los marcadores Solcap\_snp\_sl\_47697 y Solcap\_snp\_sl\_47749, tiene una densidad promedio de 31 Kb entre marcadores, con un rango que va desde 2.8Kb hasta 211Kb (Tabla 3-3).

## 2.2 Generación e identificación de recombinantes

Para la búsqueda de recombinantes se utilizó una adaptación de la técnica de genotipado High Resolution Melting (HRM). La modificación consistió en hacer una PCR-múltiple en una primera etapa, interrogando los marcadores flanqueantes de la región con SD (Solcap\_snp\_sl\_47697 y Solcap\_snp\_sl\_47749), en la segunda etapa del HRM se generaron las respectivas curvas de fusión (HRM-multiplex) (**Ver anexo 3-6**). La técnica de HRM-multiplex permitió identificar con precisión las cuatro posibles combinaciones de genotipos en la región en estudio.

Los resultados del cribado con HRM de la PFM fue de un 75% de plantas con genotipo homocigoto (TO-937) "**pp**" para ambos marcadores, un 23% de plantas con genotipo heterocigoto para ambos marcadores "**hh**"; y 2% restante fueron los recombinantes, con genotipo heterocigoto para un marcador y homocigoto TO-937 para el otro y viceversa "**ph**" y "**hp**". Las plantas con genotipo heterocigoto para ambos marcadores fueron seleccionadas para continuar generando recombinantes (**Fig. 3-2**).

A partir de los 46 recombinantes en la región diana de 380 Kb, se definieron 10 bins, con un tamaño promedio de 30 Kb. Un 69% de los recombinantes presentaron el punto de recombinación en los dos primeros bins (**Tabla 3-3**), entre ellos el 54% presentó el punto de recombinación entre los marcadores Solcap\_snp\_sl\_47697 y Solcap\_snp\_sl\_3920 y un 15 % entre los marcadores Solcap\_snp\_sl\_3920 y Gsnp-33060. Los 46 recombinantes se plantaron



**Figura 3-4. Saturación con marcadores SNPs de la región distal del cromosoma 4 con SD. (4a)** Mapa del cromosoma cuatro con distancias genéticas, **(4b)** saturación de la región diana inicial con marcadores desarrollados a partir del "mocromatriz" de SOLCAP". **(4c)** saturación con marcadores diseñados a partir de las secuencias de RNA\_Seq en la región final de 160 Kb.

en invernadero para obtener semilla de autofecundación. De éstos, 17 recombinantes produjeron semilla en cantidades adecuadas para las pruebas de progenie. Como algunos pertenecían al mismo número de bin, al final solo seis recombinantes fueron informativos y fueron usados en las pruebas de progenie y consecuentemente en el mapeo fino (**Tabla 3-3**).

Marcador (SNP)	Posición Física (Pb)	Tamaño región heterocigota (pb)	211850	244910	250242	278982	336447	346671	16898	362979	101596	185264	Distancia física/entre marcadores (pb)	N° recomb/Kb	
		Tamaño región homocigota (pb)	185264	152204	146872	118132	60667	50443	362979	34135	295518	211850			
		N° Bin	1	2	3	4	5	6	7	8	9	10			
		N° recomb/Bin	25	7	1	2	1	1	3	4	1	1			
<i>solcap_snp_sl_47697</i>	62877923	Región diana	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>p</i>	<i>p</i>	23848	nd	
<i>solcap_snp_sl_3920</i>	63089773		<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>p</i>	<i>h</i>	211850	0,12	
<i>Gsnp-33060</i>	63122833		<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>p</i>	<i>h</i>	33060	0,21	
<i>solcap_snp_sl_47742</i>	63128165		<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>p</i>	<i>h</i>	5332	0,19	
<i>Gsnp-67132</i>	63156905		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>p</i>	<i>h</i>	28740	0,07	
<i>Gsnp-83668</i>	63173441		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>p</i>	<i>h</i>	16536	0,06	
<i>Gsnp-94601</i>	63184374		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>h</i>	10933	0,00	
<i>Gsnp-107940</i>	63197713		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>h</i>	13339	0,00	
<i>Gsnp-121747</i>	63211520		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>h</i>	13807	0,00	
<i>solcap_snp_sl_3952</i>	63214370		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>h</i>	2850	0,35	
<i>Gsnp-134821</i>	63224594		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>h</i>	10224	0,10	
<i>Gsnp-151129</i>	63240902		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>h</i>	16308	0,25	
<i>solcap_snp_sl_47749</i>	63258139		<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>h</i>	<i>h</i>	<i>h</i>	17237	nd	
<i>solcap_snp_sl_3975</i>	63275037		Semilla viable	si	si	si*	si*	si*	no	no	si*	no	no		

**Tabla 3-3. Características de los recombinantes en la región de 380 Kb del cromosoma cuatro con segregación distorsionada.** Se definieron 10 bins con los 46 recombinantes obtenidos. En la parte superior se muestra el tamaño físico de las regiones heterocigotas (*h* color verde) y homocigotas TO-937 (*p* color celeste), el numero de recombinantes obtenidos/bin. A la izquierda los marcadores diseñados con su posición física, a la derecha la distancia en pb entre marcadores y el N° de recombinantes/Kb, en la parte inferior la producción de semilla viable por grupo de recombinantes.

### 2.3 Pruebas de progenie a los recombinantes seleccionados

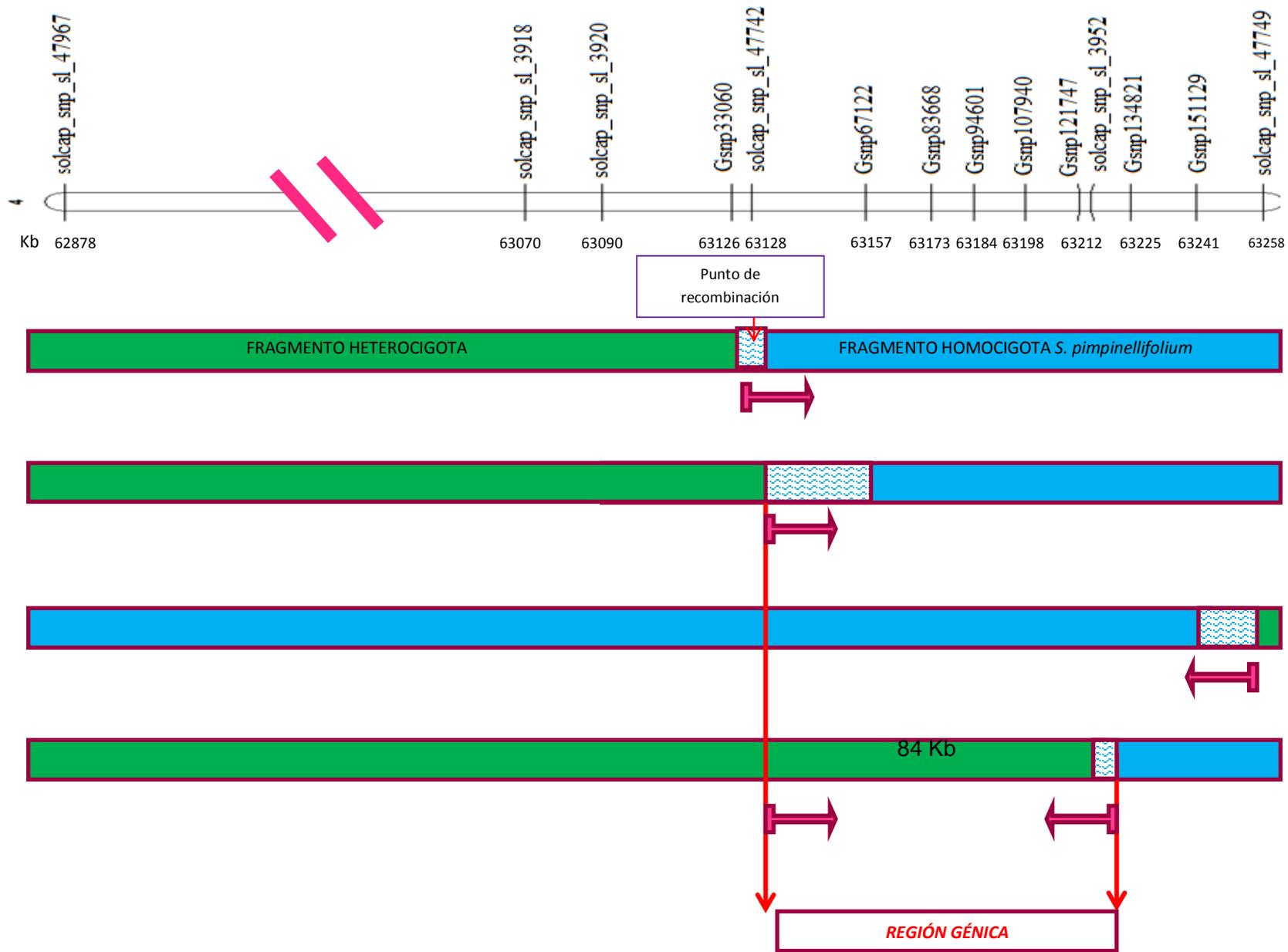
La selección de los recombinantes para las pruebas de progenie se basó: 1) en la disponibilidad de semilla F<sub>2</sub> en cantidades adecuadas y 2) en la posición donde se localizaba el punto de recombinación. Como regla general si las segregaciones obtenidas en las progenies de cada recombinante no difieren estadísticamente de las esperadas (1:2:1) significa que el alelo responsable del fenómeno de SD está fijado en el fragmento homocigoto de TO-937, por el contrario si las segregaciones difieren según la prueba de Chi-cuadrado significa que el alelo sigue segregando dentro de población (condición heterocigota en la región SD). El mapeo fino de la región con SD, se inició con el recombinante dentro del bin 3, cuya

posición de recombinación está entre los marcadores Gsnp-33060 y Solcap\_snp\_sl\_47742. La prueba de progenie no encontró diferencias significativas entre las segregaciones observadas y esperadas ( $p > 0.1$ ) (**Tabla 3-4**). Por consiguiente, el locus responsable de la SD tendría su nuevo límite proximal en el marcador Gsnp-33060 (**Fig. 3-6**). En la prueba de progenie realizada con recombinante del bin 4, la segregación no fue distorsionada (**Tabla 3-2**), por consiguiente, el nuevo límite proximal de la región con SD es el marcador solcap\_snp\_sl\_47742.

Pedigrí	Nº bin	Población	Segregación observada (%)			$\chi^2$	Resultado	Valor de P
			l	h	p			
11M 47_12	3	BC <sub>3</sub> S <sub>1</sub>	20	51	29	2	NSD	P > 0.1
12M 47_56	4	BC <sub>3</sub> S <sub>2</sub>	27	39	34	3	NSD	P > 0.1
13M 47_23	6	BC <sub>3</sub> S <sub>4</sub>	25	50	25	0	NSD	P > 0.1
13M 47_54	5	BC <sub>3</sub> S <sub>6</sub>	0	71	29	21	SD	P < 0.005

**Tabla 3-4. Prueba de progenie realizada con los recombinantes 3, 4, 6 y 5** para acotar la región SD del cromosoma 4. Genotipos: *l*: homocigoto MM, *h*: heterocigoto y *p*: homocigoto TO-937. NSD: ausencia de distorsión en la progenie y SD: progenie con distorsión.

Para confirmar el límite distal, establecido previamente con la población de RILs, se utilizaron los recombinantes correspondientes al bin 6. Esta familia no mostró SD, confirmándose que el límite distal lo delimita el marcador solcap\_snp\_sl\_47749, y definiendo por tanto una región de 130 Kb. Finalmente, la descendencia del recombinante del bin 5, presentó SD, lo que permitió cambiar el límite distal de la región, quedando finalmente acotada la región SD en 84 Kb y quedando flanqueada por el marcador solcap\_snp\_sl\_3952 en el extremo proximal y solcap\_snp\_sl\_47742 en el extremo distal (**Fig. 3-5**).



**Figura 3-5. Mapeo fino de la región SD.** Se muestra el tamaño final del fragmento del extremo distal del brazo largo del cromosoma 4 con segregación distorsionada flanqueado por los marcadores *solcap\_snp\_sl\_3952* y *solcap\_snp\_sl\_47742*. En la parte superior están los marcadores y su posición física (Kb) dentro del fragmento estudiado. Las barras de color verde representan regiones del cromosoma heterocigotos, las de color celeste regiones homocigotos con genotipo TO-937, los rectángulos punteados son las regiones donde ocurre la recombinación. En la derecha las letras indican el tipo de segregación (N normal y A anormal) y el número de bin al que pertenece el recombinante.

### 3. Análisis de la variabilidad en la región con SD en otros silvestres

El árbol de filogenia desarrollado con accesiones de *S. lycopersicum* y especies silvestres, muestra que la variabilidad genética en *S. lycopersicum* en la región con SD es baja (**Fig. 3-6**). Por el contrario, las accesiones de *S. pimpinellifolium* tienen un rango de variabilidad mayor, llegándose a formar hasta doce grupos o “clústeres” en esta región. Para realizar un primer estudio sobre la variabilidad alélica en la región SD, se escogieron accesiones representativas de seis clúster representativos del entorno en el que se localiza TO-937.

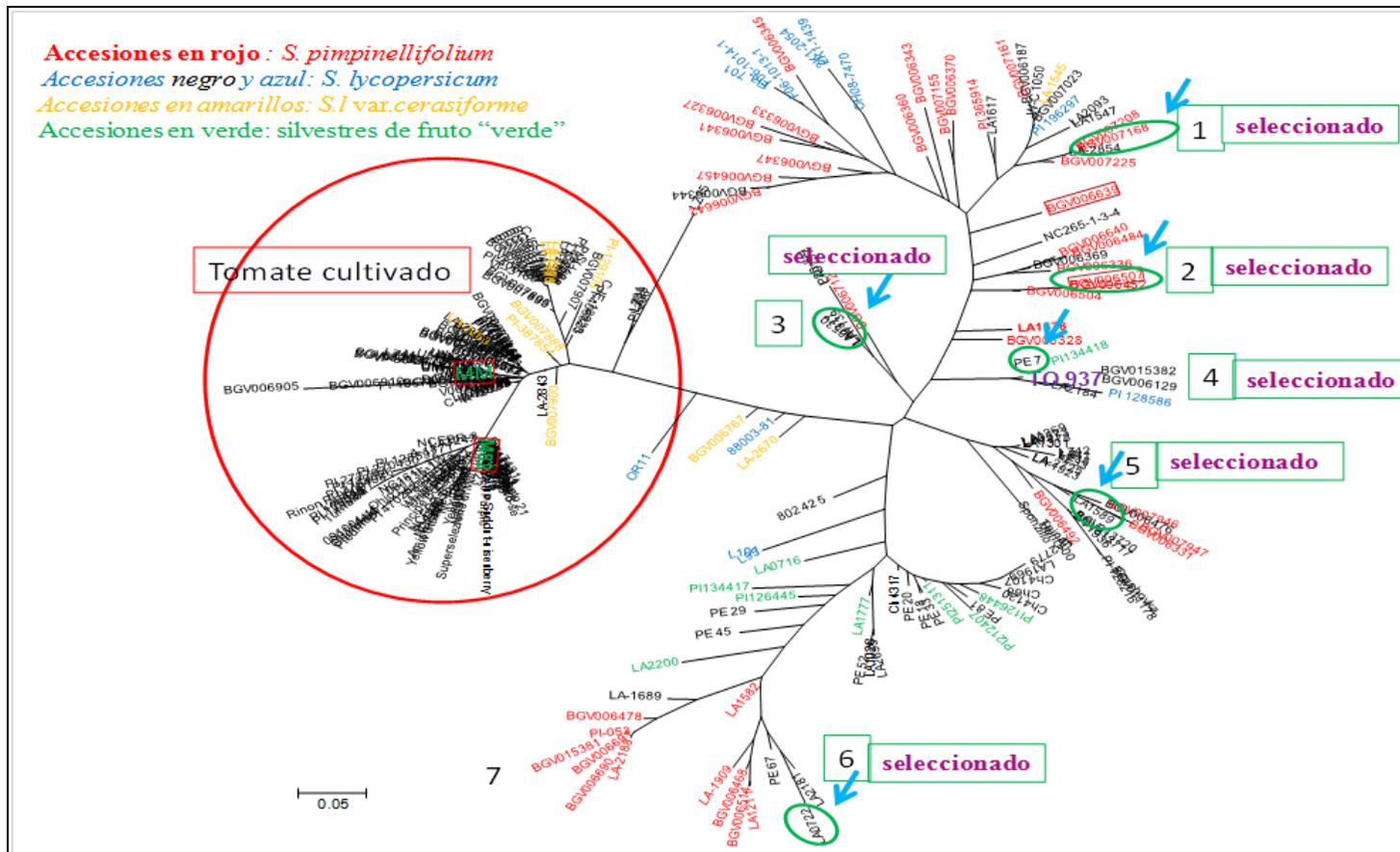
Los resultados de las pruebas de progenie realizadas a las poblaciones F<sub>2</sub> generadas a partir de los cruces entre MM y las 6 accesiones silvestres, indican que solamente, la accesión PE\_7 de *S. pimpinellifolium*, localizada según el árbol filogenético desarrollado a partir de los datos de variabilidad en la región SD en el mismo clúster que TO-937, presentó segregación distorsionada en la región SD (**Tabla 3-5**).

### 4. Genes anotados en la región con SD

La región diana de 84Kb delimitada en el extremo proximal por el marcador *Solcap\_snp\_sl\_47742* (pos. fis. 63128165 pb) y el marcador *Solcap\_snp\_sl\_3952* (pos. fis. 63212165pb (coordenadas SL2.40ch04:63128165...63212165; <http://solgenomics.net/gbrowse/bin/gbrowse/ITAG2.3>) contiene 10 genes anotados (**Fig. 3-7**).

Según la anotación realizada por el ITAG la región contiene: un gen RLK (Receptor like Kinasa), dos genes HSP90 en tándem (anotados como HSP90 y endoplasmina), dos Ciclinas B en tándem, una proteasa de procesamiento vacuolar y cuatro genes que codifican proteínas pequeñas de función desconocida.

El gen *Solyc04g081590*, anotado como RLK por el ITAG es probablemente el ortólogo en tomate de *Clavata 1* (CLV1) ya que cuando realizamos un BLAST con AtCLV1 contra las



**Figura 3-6. Dendrograma filogenético generado a partir de la variabilidad genética de SNPs localizados en región SD (tamaño 229 Kb).** Genotipado realizado con la "micromatriz 8K SNP Infinium de Illumina desarrollado por la iniciativa SOLCAP". A partir de los datos publicados por Sim et al. (2012), Blanca et al. (2012) y propios de nuestro grupo de investigación, se calculó la distancia de Nei (1983) entre los genotipos y el dendrograma "Neighbour-Joining" con el programa Powermarker (Liu y Muse 2005) se dibujó con MEGA 6 (Tamura et al. 2013).

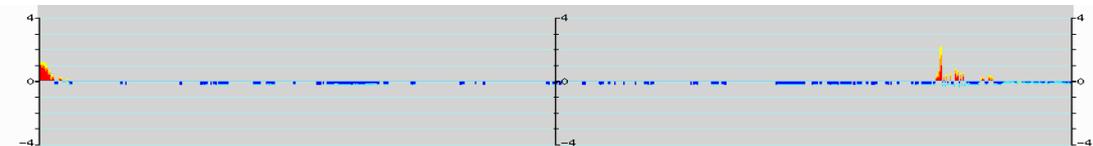
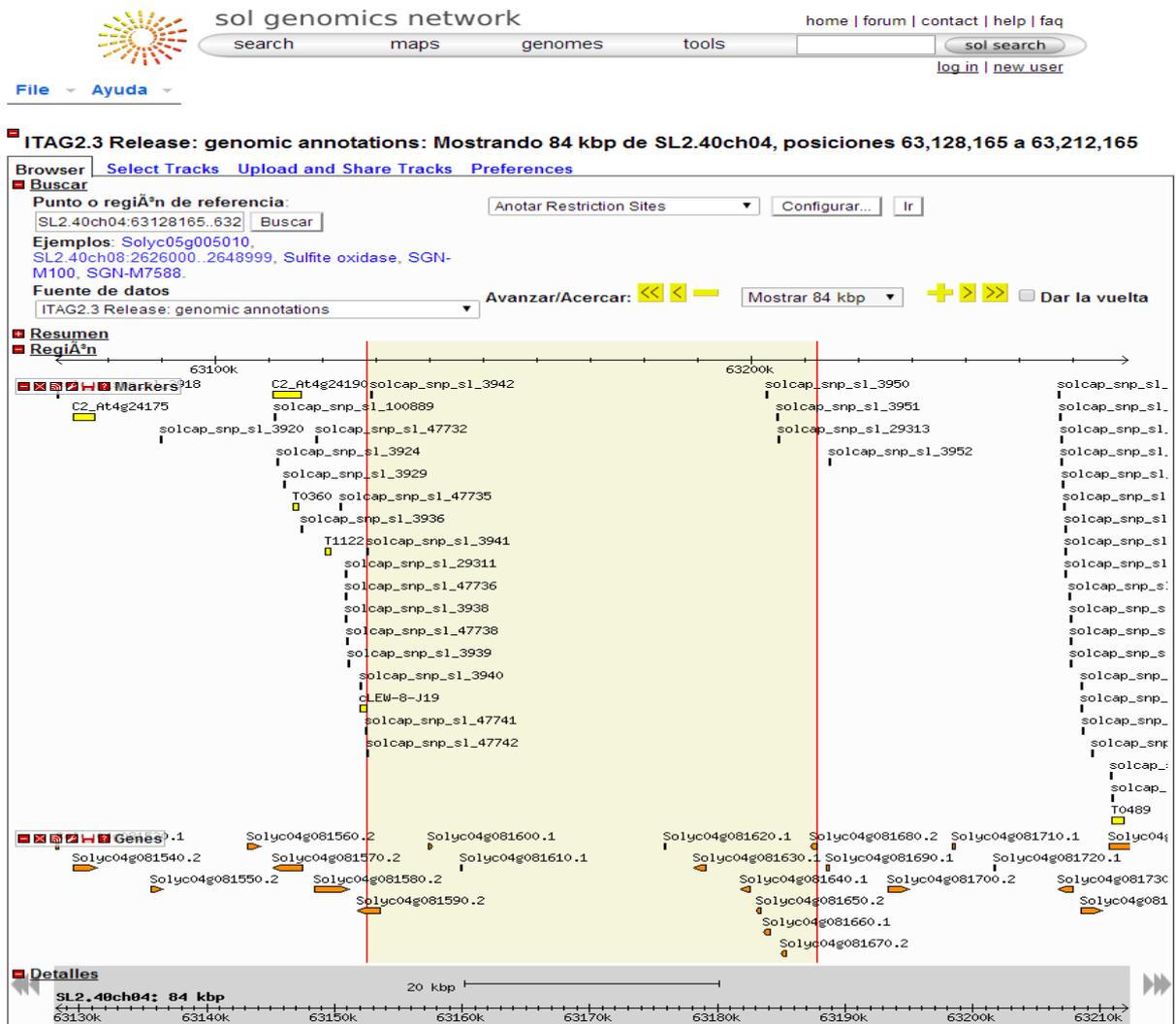
Pedigrí (cruce inicial)		Segregación observada (%) Población F <sub>2</sub>			X <sup>2</sup>	Resultado	Valor de P
Madre	Padre	<i>l</i>	<i>h</i>	<i>p</i>			
<i>S.l</i> (cv.MM)	<i>S.p</i> (acc. PE-7)	0	37	63	41	SD	P<0.005
	<i>S.p</i> (acc. LA0722)	26	48	26	0	NSD	P > 0.1
	<i>S.g</i> (acc. LA0530)	24	62	14	4	NSD	P > 0.1
	<i>S.p</i> (acc. LA1589)	19	49	32	2	NSD	P > 0.1
	<i>S.p</i> (acc. BGGV006507)	13	50	37	2	NSD	P > 0.1
	<i>S.p</i> (acc.BGGV007168)	19	56	25	3	NSD	P > 0.1

**Tabla 3-5. Resultados de las pruebas de progenie realizadas a las poblaciones F<sub>2</sub> de seis accesiones de silvestres seleccionadas para estudiar si presentaban SD en la región SD. *S.l* (*Solanum lycopersicum*), *S.p* (*Solanum pimpinellifolium*), *S.g* (*Solanum galapagense*). *l*: genotipo homocigoto MM, *h*: genotipo heterocigoto y *p*: genotipo homocigoto TO-937. NSD: ausencia de distorsión en la progenie y SD: progenie con distorsión.**

secuencias de tomate esta es la que da mayor identidad. Al igual que la proteína CLV1 de *Arabidopsis*, SICLV1 posee tres dominios claramente definidos y conservados (**ver anexo 3-4**), un primer dominio tipo LRR (rico en leucinas) que reconocería una señal de tipo peptídico, un segundo dominio también rico en L denominado RNI (de la familia de Cyclin A/CDK2-associated p19, familia Skp2) y un dominio kinasa COOH terminal.

En esta región se encuentra el gen *Solyc04g081640* cuyo producto es una HSP90 similar al gen adyacente *Solyc04g081630*, pero dado que la HSP es residente en el retículo endoplasmático el ITAG lo anota como "endoplasmina". Esta endoplasmina es seguramente el ortólogo de *SHEPHERD* (*SHD*) también llamada *GRP94* (Ishiguro et al. 2002). En *Arabidopsis* *SHD* es necesario para el plegamiento correcto y formación de complejos de las proteínas CLV, sin cuya actividad las CLV no son funcionales (Ishiguro et al. 2002). Sin embargo una revisión más detallada de *Solyc04g081630* y *Solyc04g081640* indican que se trata

de versiones truncadas en ambos casos. Curiosamente una versión completa de la endoplasmina se encuentra dos genes "aguas arriba" de CLV1 en *Solyc04g081570*, cuya secuencia muestra que codifica una HSP90/GRP con señal de retención en el RE.



**Figura 3-7. Genes anotados por ITAG en la región con SD del cromosoma 4 (SL2.40ch04:63128165...63212165).** En la parte superior de la figura se muestra la escala en el genoma, debajo marcadores moleculares (la mayoría de la plataforma SOLCAP), seguido de los genes con la anotación ITAG y por último datos de expresión incluidos en la base de datos SGN ([http://solgenomics.net/gbrowse/bin/gbrowse/ITAG2.3\\_genomic/](http://solgenomics.net/gbrowse/bin/gbrowse/ITAG2.3_genomic/)).

Por su parte las dos ciclinas B localizadas en esta región pertenecen a las denominadas ciclinas mitóticas, que son necesarias para que la célula entre y salga de la fase de mitosis con niveles de expresión más elevados en las células que están en proliferación. Sin embargo, estas dos ciclinas B son realmente dos fragmentos de una sola secuencia anotada como LOC101266716 "cyclin-B2-4-like" (*Solanum lycopersicum*) por NCBI y que correspondería a una ciclina competa. Funcionan en interacción con la enzima quinasa dependiente de ciclina B (CdKs) de la cual hay una copia aguas abajo al final del cromosoma 4. Adicionalmente, en la parte central de la región SD se localiza un posible retrotransposon (Ty3/Gypsy LTR) en Heinz-1706 y que pasó por desapercibido para ITAG.

## Discusión

La cercanía filogenética entre las especies *S. lycopersicum* y *S. pimpinellifolium*, presupondría la no existencia de barreras a la cruzabilidad entre ambas especies. Sin embargo, hemos podido poner de manifiesto que no todas las regiones del genoma se transfieren a la descendencia con la misma eficiencia. Más concretamente hemos podido definir que en una región con SD del cromosoma 4, los alelos de TO-937 se transmiten preferentemente frente a los equivalentes del tomate cultivado MM. Adicionalmente, hemos comprobado que la región donde se localiza el locus SD, muestra variabilidad genética dentro de *S. pimpinellifolium* (**Figura 3-6**). Únicamente las accesiones TO-937 y PE\_7 (ambas muy próximas genéticamente según la variabilidad de esta región) presentan SD en las poblaciones derivadas tras el cruzamiento con MM (**Tabla 3-3**).

Con la evidencia recopilada en las diferentes pruebas realizadas a las distintas poblaciones usadas para estudiar la SD en este capítulo, podemos hacer una primera descripción de las características del locus de SD estudiado en el presente trabajo: **1)** el locus se encuentra en la región distal del brazo largo del cromosoma 4, dentro de un intervalo de 84 Kb flanqueado por los marcadores Solcap\_snp\_sl\_47742 en el extremo proximal solcap\_snp\_sl\_3952 en el extremo distal, **2)** la distorsión favorece los alelos del parental silvestre (*p*), **3)** no es causada

por alelos letales, ya que las plantas son fértiles independientemente de la combinación alélica en este locus, produciendo polen viable y semillas fértiles, **4**) la segregación observada en la descendencia de una autofecundación (por ejemplo un población  $BC_3S_1$ ) es prácticamente 0:1:3 ( $l:h:p$ ), si se produjera la letalidad de los alelos “ $l$ ” únicamente en un tipo de gameto (masculino o femenino), la descendencia debería de ser 50% de heterocigotos ( $h$ ) y 50% de homocigotos  $p$ , por consiguiente, gametos de ambos sexos pueden estar afectados, **5**) la penetración no es del 100%, dada la presencia genotipos heterocigotos ( $h$ ) en la región con SD en las diferentes poblaciones estudiadas.

Si consideramos las características antes mencionadas que describen la región con SD y las propias de los parentales involucrados en este cruce (autocompatibles), tendríamos que descartar las hipótesis de incompatibilidad (Locus  $S$ ) e incompatibilidad unilateral (gen  $CUL1$ ,  $HT$ ) como responsables del sesgo en las segregaciones observadas.

Por otro lado, quien si describe un fenómeno muy similar al nuestro es Rick (1966,1970) quien descubrió un locus de SD que atribuyó a la presencia de un gen “eliminador de gametos” ( $Ge$ ) con tres alelos (definidos en la introducción), que provocaba distorsiones en las poblaciones derivadas que guardan mucha semejanza con las descritas en el presente capítulo como son: ocurre en un entorno heterocigoto, se favorecen los alelos silvestres, afecta tanto gametos femeninos como masculinos y su penetración es menor al 100%. Además, mapeó  $Ge$  en el cromosoma 4, aunque en posiciones centroméricas, no distales como en el presente trabajo. Otros grupos de investigación también han encontrado un locus para SD en la misma región que nosotros (D. Francis, comunicación personal), por lo que, dadas las características de la segregación y la coincidencia cromosómica, pensamos que el locus SD descrito en el presente capítulo corresponde al locus  $Ge$  descrito por Rick (1966) y que las diferencias en la localización cromosómica se deben a que actualmente la precisión de los mapas genéticos es muy superior que en el momento en que se describió el locus  $Ge$ .

En relación a los genes que anotados en la región con SD cuya función involucra procesos relaciones con gametogénesis o división celular, tenemos miembros de la familia CLV que, regulan la proliferación celular durante la formación del fruto de la misma forma que ocurre con los meristemos (Durbak y Tax 2011). En el ovario estos meristemos son los encargados de generar tejido placentario y los óvulos (Ferrandiz et al. 1999). Es posible plantearnos como hipótesis que alteraciones en proteínas CLV podría afectar el desarrollo del óvulo y del polen. Siendo posible que CLV1/TO-937 proporcione una ventaja adaptativa frente a la de CLV1/MM en el óvulo/polen en el entorno del esporofito heterocigoto. Sin embargo, al no

detectarse variación en la secuencia proteínica en CLV1/TO-937, el mecanismo causal debería deberse a alteraciones en el patrón de expresión de los dos alelos. Estas alteraciones en el patrón de expresión han sido descritas en múltiples diferentes mutaciones de interés en tomate, incluidas algunas con papel importante durante la domesticación por ejemplo *fw2.2* (Frary et al. 2000).

Otro gene candidato mencionado es *Solyc04g081630* que corresponde a una endoplasmina, cuyo ortólogo en *Arabidopsis* es de *SHD*. *SHD* es necesario para que el meristemo se mantenga sin diferenciar, además tiene una función en el desarrollo de los gametofitos (Ishiguro et al. 2002), siendo requerido para la elongación del tubo polínico o/y la penetración del estilo. Puede que la SHD/TO-937 presente un patrón de expresión optimizado para el contexto de TO-937 y que tenga ventajas en el contexto del esporofito heterocigoto con respecto a SHD/MM (mutación heterocrónica). Sin embargo, no sabemos si SHD tiene la misma función en el desarrollo de gametofitos femeninos.

En relación a los demás genes localizados en la región final de 84 Kb, las dos ciclinas B, que en realidad correspondería a una ciclina completa, están implicadas con procesos de división celular su alteración podría generar el fenómeno de SD. La ciclina B se une a la enzima quinasa dependiente de ciclina (CdK1). Liu et al. (2008) mostraron el rol de la quinasa dependientes de ciclina (CDKs) en la gametogénesis. Se sabe que la degradación de unas proteínas que inhiben las CDKs es regulada por proteólisis mediada por ubiquitinación y que en ello participan las proteínas ubiquitina ligasa tipo RING para luego ser degradada. Trabajando con dobles mutantes para ubiquitin ligasas de tipo RING, se encontró que causaban degeneración de óvulos, polen inviable (50-70%) y células germinativas de polen sin división. Sin embargo, las secuencias de los inhibidores de quinasas (CDK-inhibidores) y de las ubiquitinas ligasa tipo Ring) anotadas en el trabajo de Liu et al. (2008), no tienen homologías con la región con SD del cromosomas cuatro (SGN <http://solgenomics.net/tools/blast/>). Por otra parte, las secuencias de las proteínas correspondientes a la ciclina-B en TO-937 es idéntica a las de MM por lo que de nuevo la hipótesis que se podría plantear es que una alteración en los niveles o perfil de expresión de la ciclina-B podría causar desajustes en la división celular de los gametofitos que ocasionaría el que finalmente los gametofitos o las semillas portadoras de alelos de TO-937 tuvieran alguna ventaja sobre los de MM en un contexto heterocigoto.

Tampoco podemos descartar que los pequeños péptidos codificados por *Solyc04g081610* y *Solyc04g081620*, sin homología en las bases de datos, puedan tener algún papel clave en el

proceso de direccionamiento del tubo polínico hacia el óvulo como se ha descrito para los péptidos tipo LURE (ver tabla anexo 3-4). Aunque no se observa la presencia de presentan los residuos de Cys conservadas típicos de las proteínas de este tipo descritas hasta el momento (ver anexo 3-4) no podemos descartar que se trate de un nuevo grupo de péptidos señalizadores cuya diferente expresión o afinidad por el receptor (normalmente del tipo RLKs) determine el fenotipo observado. En este sentido no podemos descartar el que el RLK que hemos anotado como CLV1 pueda ser receptora de esta señal. Aunque CLV3 codifica el péptido con el que normalmente CLV1 interacciona y está localizado en otro cromosoma podría ocurrir que CLV1 reconociera estos nuevos péptidos. La capacidad de las RLKs de interaccionar con más de una molécula señalizadora se ha descrito previamente en plantas (Montoya et al. 2002)

Finalmente, la presencia en la parte central de la región de un posible retrotransposon Ty3/Gypsy LTR en Heinz-1706 puede alterar de forma importante la expresión de los genes flanqueantes. Desde anularla o incrementarla (Lacombe et al. 2010; Butelli et al. 2012), hasta insertar nuevos genes o producir reordenamientos en las regiones donde se insertan (Xiao et al. 2008). La domesticación de plantas ilustra diferentes casos en los que la activación / represión o modificación heterocronica de la expresión de genes fue clave para la adquisición de nuevas características que fueron seleccionadas por los primeros agricultores. Ese es el caso de "teosinte branch" en maíz (Studer et al. 2011) y de *fw2.2* (Frary et al. 2000). El análisis de expresión combinado con una disección de la región para las diferentes accesiones que difieren en SD nos permitirá averiguar si existe una asociación entre la presencia de diferentes versiones del transposon, los niveles de expresión de los genes flanqueantes y la SD.

Alternativamente es posible que este elemento proporcione inestabilidad a esta zona o alta variabilidad entre las accesiones con posibilidades de que diferentes especies o accesiones de la misma especie presenten alteraciones en cuanto al tamaño del elemento retrotransposon insertado como el ejemplo que describe Goulet et al. (2012). Es posible que ello implique alteraciones en la viabilidad o capacidad para recombinar cuando se presenta en heterocigosis y eso se refleje a nivel de SD. Es razonable pensar que si las alteraciones estructurales causadas por las inserciones o reorganizaciones de elementos retrotransponibles sean las que ocasionen la SD, estas también ocurran en otras posiciones del genoma donde se detectan retrotransposones (un blast de este elemento contra el genoma del tomate permitió identificar

333 inserciones). Será interesante ver si estas inserciones puedan estar asociadas a regiones con SD en otras poblaciones.

En resumen, a pesar de los errores de anotación encontrados en los genes dentro de región con SD, la información aquí recabada liga a estos genes y sus productos con procesos gametofíticos, los cuales aparentemente a través de una acción complementaria, interactúan favoreciendo genotipos en donde están presentes los alelos de TO-937. Para ahondar en el conocimiento del mecanismo que produce SD descrito en este capítulo se plantean realizar las siguientes actividades: 1) obtener más recombinantes en esa región que permitan disminuir el listado de genes recombinantes, 2) estudiar la estructura de la región genómica de TO-937, para lo que sería necesario generar genotecas de BACs, 3) analizar con más detalle posibles alteraciones en el desarrollo de los gametofitos masculinos / femeninos en el entorno del esporofito heterocigoto para la región, 3) analizar el desarrollo del tubo polínico de las diferentes combinaciones genotípicas posibles, 4) hacer estudios de expresión en las estructuras reproductivas, 5) hacer un rastreo para determinar la presencia/ausencia o variaciones en el tamaño del retrotransposon /inserción en accesiones de tomate, especialmente en *S. pimpinellifolium*, y verificar su relación con las SD. Esperamos que estos estudios nos permitan definir cuál es el gen responsable de la SD y su mecanismo de acción en un futuro.

## **CAPITULO 4: Identificación de QTLs implicados en caracteres de calidad de fruto de tomate en la genoteca de líneas de introgresión de TO-937**

### **Introducción**

La teoría del mapeo de QTLs fue descrita por primera vez por Sax (1923) al observar asociación entre el peso de semillas (carácter complejo) con el color de la testa (carácter monogénico) en *Phaseolus vulgaris*. Con el desarrollo de la tecnología de marcadores moleculares, a finales de los 80s empezaron a publicarse trabajos de identificación de QTLs en mapas de ligamiento. Por ejemplo, Paterson et al (1988), utilizando poblaciones desarrolladas tras el cruzamiento de una accesión de *S. chmielewskii* con una variedad de tomate, demostró con éxito que la variación de algunos marcadores moleculares explicaban una parte sustancial de la variación fenotípica en caracteres cuantitativos como SSC, rendimiento, pH etc. Acuñando el término mapeo de QTLs al estudio para localizar los genes responsables de caracteres de herencia compleja en un mapa genético y explorar sus efectos e interacciones. Durante los últimos 25 años se han publicado números estudios que describen el mapeo de QTLs en tomate y otras especies hortícolas. Entre los QTLs mapeados incluyen tanto los relacionados con caracteres de importancia agronómica como son resistencia a enfermedades e insectos, tolerancia a estrés abiótico y otros relacionados con caracteres de calidad de fruto (revisiones hechas por Foolad et al. 2007; Labate et al. 2007; Grandillo et al. 2011; Grandillo et al. 2013).

El primer estudio de mapeo de QTLs en tomate lo realizó Weller et al. (1988) en una población F<sub>2</sub> generada a partir del cruce entre *S. lycopersicum* x *S. pimpinellifolium* (accesión CIAS27) y a partir de mediados de la década de los 90s el número de trabajos de mapeo de QTLs en tomate aumentó considerablemente. Para llevar a cabo estos estudios se utilizaron diferentes tipos de poblaciones: poblaciones basales segregantes (F<sub>2</sub> - BC<sub>1</sub>) (Grandillo y Tanksley, 1996; Monforte et al. 1996; Foolad et al. 1998), líneas recombinantes consanguíneas (RILs) (Graham et al. 2004; Villalta et al. 2008), poblaciones avanzadas de retrocruce (Tanksley y Nelson, 1996; Tanksley et al. 1996; Bernacchi et al. 1998, entre otras), líneas de retrocruce consanguíneas (Backcross Inbred Lines o BILs; Doganlar et al. 2002) y líneas de introgresión (ILs, Eshed and Zamir 1995).

Dadas las ventajas que tienen las genotecas de ILs para la disección genética de caracteres complejos e introducir variabilidad exótica en los cultivares modernos, se han desarrollado

varias colecciones de ILs a partir de diversas especies silvestres de tomate: *Solanum pennellii* LA0716 (Eshed and Zamir 1994), *S. habrochaites* LA1777 (Monforte and Tanksley 2000a), *S. lycopersicoides* LA2951 (Chetelat and Meglic 2000), *S. habrochaites* LA0407 (Francis et al. 2001), *S. lycopersicoides* (Canady et al. 2005) y *S. habrochaites* LYC4 (Finkers et al. 2007). Entre éstas, la colección de *S. pennellii* es sin duda la más utilizada, y ha permitido la identificación de más de 3069 QTLs implicados en centenares de caracteres de importancia agronómica (Alseekh et al. 2013).

En esta tesis, hemos desarrollado una genoteca completa de ILs utilizando como parental donante de genes la accesión TO-937 de *S. pimpinellifolium* sobre el fondo genético del cultivar “Moneymaker” de *S. lycopersicum*. La accesión TO-937, presenta resistencia a *Tetranychus urticae* (Fernández-Muñoz et al. 2000; Fernández-Muñoz et al. 2003; Alba et al. 2009) y, en estudios realizados en una población de RILs desarrollada a partir de este cruzamiento, también se encontraron alelos que pueden ser fuente de variabilidad genética para la mejora de caracteres relacionados con la calidad del fruto como son: contenido de ácido ascórbico (Lima-Silva et al. 2012), azúcares, carotenoides (Capel et al. 2011) y compuestos volátiles (Rambla et al, en preparación). Lo que hace de esta accesión un material de interés para su inclusión en programas de mejora genética del cultivo. Dentro de los aportes más significativos del mapeo de QTLs en otras poblaciones de ILs generadas a partir de parientes silvestres están: desarrollo de nuevas variedades (Gur y Zamir 2004), mapeo fino de QTLs (Eshed y Zamir 1995; Monforte y Tanksley 2000b; Ku et al. 2001; Monforte et al. 2001), clonaje de QTLs (Frary et al. 2000; Fridman et al. 2000; Liu et al. 2002), estudios de metabolismo genético (Schauer et al. 2006), transcriptoma y su correlación con metabolitos (Lee et al. 2012), actividad enzimática (Steinhauser et al. 2011).

El objetivo del presente capítulo es evaluar en la genoteca de ILs varios aspectos de la calidad del fruto, con objeto de mapear QTLs asociados a esos caracteres. La definición de esos QTLs resultará de interés tanto para el estudio básico de la genética de los caracteres asociados como para su posible integración en futuros planes de mejora.

## **Materiales y Métodos**

### **1. Material vegetal**

El material vegetal evaluado en este ensayo estuvo compuesto por plantas de tomate de la variedad de tomate para consumo fresco "MoneyMaker" (MM, parental recurrente de la población IL), y 53 líneas de introgresión (ILs) cuyo obtención ha sido descrita en el Capítulo II de esta tesis. Semillas de cada línea y del cultivar MM se pusieron a germinar y transcurridas tres semanas se seleccionaron 10 plántulas de cada línea de introgresión y 60 plántulas del parental MM, las cuales se llevaron al campo experimental definitivo para ser cultivadas.

### **2. Localidades evaluadas y diseño del experimento**

Los ensayos se realizaron durante los meses de marzo a agosto del 2013 replicados en tres localizaciones de España: Cooperativa Agrícola Alginet (Coagri) localizada en Alginet provincia de Valencia coordenadas (Latitud Norte 39° 15' 32" Longitud Oeste 0° 28' 05"), Escuela Politécnica Superior de Orihuela (Universidad Miguel Hernández), provincia de Alicante, coordenadas Latitud Norte 38° 04' 05" Longitud Oeste 0° 58' 51" y en el Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (CSIC), Algarrobo-Costa, Provincia de Málaga, con coordenadas son Latitud Norte 36° 46' 22" Longitud Oeste 4°25' 25". Las ILs se cultivaron en invernaderos de malla transparente, usando suelo como sustrato. El manejo agronómico, las prácticas culturales y las evaluaciones del cultivo fueron los mismos en los tres emplazamientos. El diseño experimental de plantación fue de bloques al azar con 8 repeticiones para cada IL y 60 repeticiones para el parental MM. El diseño se dividió en dos partes: una parte con un diseño de bloques completamente aleatorizados, comprendiendo 7 bloques; la unidad experimental fue una planta de cada IL en cada bloque y 6 replicas de MM. En el octavo bloque las unidades experimentales fueron parcelas de tres plantas de cada IL, plantándose una única parcela por IL, se plantaron siguiendo un orden ascendente desde la SP\_1-1 hasta la SP\_12-5, intercalándose con parcelas de MM. El objetivo del diseño de este último bloque fue facilitar la evaluación de caracteres descriptivos al poder observarse en un grupo de tres plantas, en lugar de planta única como en los primeros siete bloques.

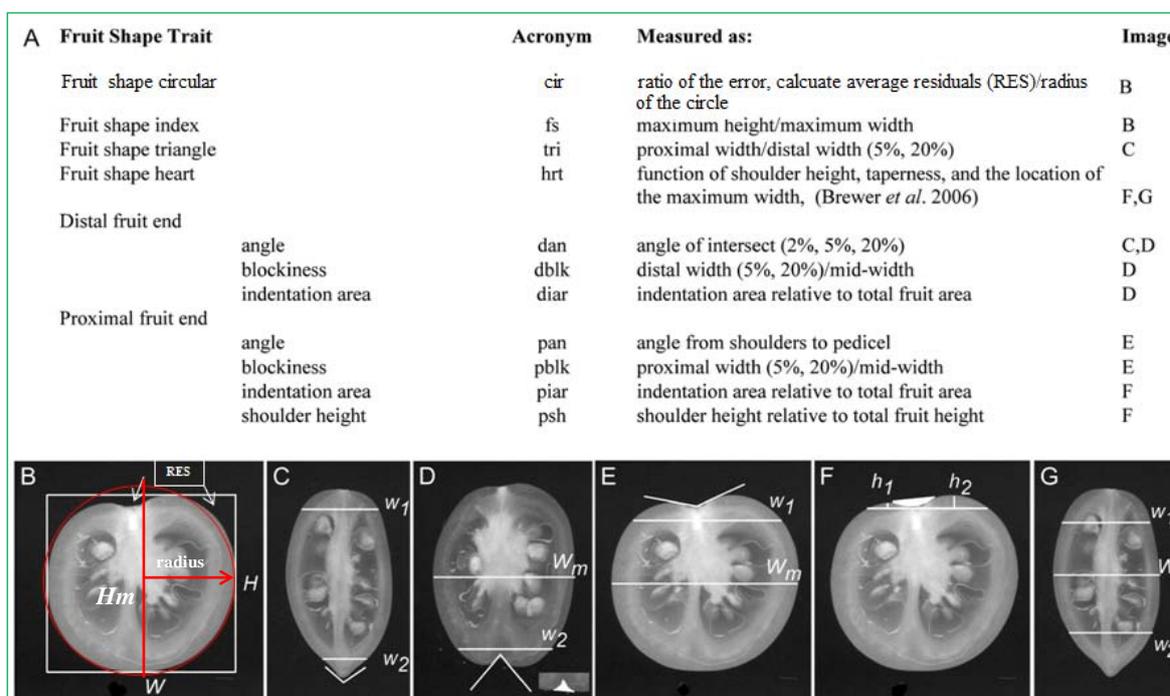
### 3. Caracteres evaluados

Durante las diferentes fases del cultivo se realizaron evaluaciones descriptivas relacionadas con el desarrollo de las plantas en los bloques compuesto por parcelas de tres plantas por IL. Estas evaluaciones se llevaron a cabo con una periodicidad semanal y sólo en Alginet y Orihuela; se evaluaron los siguientes caracteres: precocidad (PREC), definida como número de días transcurridos desde que la planta fue trasplantada (ddt) hasta alcanzar el primer fruto maduro; vigor de la planta (VIG) que se midió (al momento del cuaje del primer fruto: altura y porte), y comparó con MM utilizando una escala subjetiva (menor que MM (1), similar a MM (2), mayor que MM (3)); densidad de tricomas en hojas y ramas (TRIC), determinada según una escala de 0 = ausencia, 1 = baja, 2 = intermedia, 3 = alta; estructura del abierta o compacta del dosel (DOS); ausencia/presencia hombros verdes en frutos (HV); y presencia/ausencia de coloración violácea (posiblemente por acumulación de antocianinas) en ramas y follaje (VIOL).

Para realizar las evaluaciones de las diferentes características del fruto, se cosecharon 4 frutos maduros (rojo ligero) de cada planta de los bloques del 1-7 y en el bloque 8 se recolectó una muestra compuesta también de cuatro frutos, pero proveniente de las tres plantas de la IL representada en el bloque.

El peso promedio del fruto (FW) se determinó a partir de los valores obtenidos con una balanza digital con los cuatro frutos. El color externo (EC) se analizó con un colorímetro Konica-Minolta modelo CR-400 en tres puntos distribuidos alrededor de la región ecuatorial de cada fruto. Se utilizó el espacio de color CIELAB compuesto por los parámetros:  $L^*$ ,  $a^*$ ,  $b^*$ ,  $C^*$ ,  $H^*$ . El parámetro  $L^*$  (luminosidad), es una medida de la cantidad de luz que refleja el fruto, y varía entre  $L^*=0$  (negro) y  $L^*=100$  (blanco); los parámetros  $a^*$  y  $b^*$  se utilizan para calcular  $C^* = (a^2 + b^2)^{1/2}$  que se denomina cromacidad e indica la saturación de color con valores entre 0 a 100 ; y  $H^*$  es la tonalidad o posición del color, calculándose como  $H = \arctan\left(\frac{b}{a}\right)$ , se expresa por tanto en grados (de 0 a 360), donde el color rojo corresponde a un valor de  $H = 0^\circ$ , el amarillo a un valor  $H = 90^\circ$ , el verde un valor de  $H = 180^\circ$  y el azul un valor  $H = 270^\circ$  (**Anexo 4-1**). Para obtener las imágenes digitales de los frutos, los frutos se cortaron en dos mitades longitudinales y una de ellas se colocó sobre un escáner calibrado a 300 dpi por cm para su digitalización. Las imágenes se guardaron en formato *jpeg* para proceder a su análisis posterior utilizando el paquete informático "Tomato Analyzer 3.0" TA ([http://oarc.ohio-state.edu/vanderknaap/tomato\\_analyzer.php](http://oarc.ohio-state.edu/vanderknaap/tomato_analyzer.php)). Las mediciones realizadas a

los frutos con el programa TA fueron 1) Dimensión de fruto: Perímetro (PER), área de la sección longitudinal (AR), longitud máxima (FLL), diámetro máximo (FDL), 2) Forma de fruto: índice de forma de fruto (FS), bloque distal (DBLK), bloque proximal (PBLK), forma de fruto circular (CIR), forma de fruto triangular (TRI), altura de hombros (PSH), los acrónimos usados para la forma del fruto y la forma de cómo se calcula cada índice corresponden con las definidas por Brewer et al. (2006) (**Fig. 4-1**) Color interno (IC) del pericarpio del fruto, en el mismo espacio de color (CIELAB). Para medir el IC, se modificaron las imágenes con Photoshop CS5 12.0.4x32, dejando sólo un fragmento de tejido de pericarpio sobre el cual el programa "Tomato Analyzer" realizó las lecturas, previa normalización con la carta de color, escaneada para tal efecto con cada escáner de cada localidad.



**Figura 4-1. Definición de los caracteres usados para medir la forma de los frutos en la genoteca de ILs y, con sus respectivos acrónimos.** Modificado de Brewer et al. (2006). (*W* ancho máximo, *W<sub>m</sub>* ancho medio, *W<sub>1</sub>* ancho proximal, *W<sub>2</sub>* ancho distal, *H* altura máxima, *H<sub>m</sub>* altura media, *h<sub>1</sub>* y *h<sub>2</sub>* altura de cada hombro).

Para determinación de los sólidos solubles (SSC), pH inicial (pH-i) y la acidez titulable (TA) se utilizó una muestra compuesta de tejido de pericarpio triturando los cuatro frutos que se almacenó a -20°C hasta su utilización. Las muestras, previamente descongeladas a temperatura ambiente, se centrifugaron durante 10 min a 3,500 rpm para obtener el

sobrenadante. La determinación de SSC se realizó colocando una alícuota del sobrenadante de cada muestra sobre la superficie correspondiente de un refractómetro digital marca ATAGO modelo PAL-1 (rango de 0- 53° Brix). La acidez titulable se midió en una alícuota de 1 ml de sobrenadante que se diluyó con 10 ml de agua destilada. Para la medición se utilizó un analizador electrónico PH-Matic 23 marca CRISON y se tituló con NaOH 0.1N, la acidez se expresó como porcentaje de ácido cítrico equivalente.

#### 4. Análisis estadístico

Los datos obtenidos en las tres localidades se sometieron a un análisis de varianza (ANOVA) de dos vías para obtener las estimaciones de los efectos Genético (G), Localidad (L) y la Interacción G x L, sobre el fenotipo observado ( $ec^1$ ). Luego, se realizó un ANOVA de una vía en cada localidad ( $ec^2$ ) para calcular la heredabilidad de los caracteres ( $ec^3$ ). Para los análisis estadísticos se utilizó el programa JPM 5.1.

- ( $ec^1$ ) Modelo matemático ANOVA dos vías:

$$Y_{ijk} = \mu + G_i + L_j + (G \times L)_{ij} + \varepsilon_{ijk}$$

Donde:

$\mu$  es la media general

$G_i$  es el efecto del genotipo

$L_j$  es el efecto de la localidad

$(G \times L)_{ij}$  es la interacción genotipo X localidad

$\varepsilon_{ijk}$  el error

- ( $ec^2$ ) Modelo matemático ANOVA una vía:  $Y_{ij} = \mu + G_i + \varepsilon_{ij}$

Donde:

$\mu$  es la media general

$G_i$  es el efecto del genotipo

$\varepsilon_{ij}$  el error

- ( $ec^3$ ) Cálculo de la heredabilidad

$$h^2 = \frac{V_g}{V_t}$$

Donde:

$V_g$  es la varianza genética, calculada como la varianza entre genotipos a partir del ANOVA y

$V_t$  es la varianza total

Para detectar diferencias significativas entre las medias de cada IL con MM, se utilizó la prueba de medias de Dunnett's ( $p < 0.05$ ). Si la diferencia entre la media de una IL con la

media de MM resultaba significativa, se interpretaba que existía un QTL para el carácter en la introgresión de *SP*. Para determinar si una IL mostraba el efecto de más de un QTLs en un carácter determinado (QTLs ligados), se realizaron pruebas de contrastes entre las medias de las ILs cuyas introgresiones solapaban. Si se encontraban diferencias significativas entre las medias se asumió que existían al menos dos QTLs y fueron mapeados por separado. Finalmente se hizo un análisis de correlaciones de Pearson entre los caracteres en cada localidad.

## **Resultados**

### **1. Caracterización del fruto del parental recurrente MM y su variabilidad en las tres localidades**

Las plantas de MM mostraron una precocidad promedio en las tres localidades de 75ddt, siendo más precoz en Málaga (69 ddt), seguido por Orihuela (73 ddt) y Alginet (82 ddt), ausencia de hombros verdes en los frutos, poca presencia o ausencia de pilosidad y coloración violácea en ramas y tallos y estructura del dosel abierta. Las características generales del fruto del cultivar "Moneymaker" utilizado como control, se presentan en la **Tabla 4-1**. El peso promedio de fruto fue de 98.16 gr, no presentando diferencias estadísticamente significativas para este carácter entre las localidades ( $p > 0.7$ ). Con respecto a los caracteres perímetro, área, diámetro y longitud del fruto se encontraron diferencias estadísticamente significativas ( $p < 0.01$ ) siendo Málaga la localidad que presentó los valores más altos.

Los caracteres de forma de fruto estudiados presentaron diferencias estadísticamente significativas ( $p < 0.01$ ) entre localidades con excepción del índice de forma (FS). Los valores de SSC de MM fueron cercanos a 4.5° brix, con diferencias significativas entre localidades. Con respecto a los valores de acidez, los frutos fueron siempre ligeramente ácidos, también con diferencias significativas entre localidades. El color interno y externo estuvo en valores dentro del espectro del rojo, aunque con diferencias significativas en sus componentes entre localidades. En general en las localidades de Alginet y Orihuela se obtuvieron valores más altos en la mayoría de los parámetros asociados a caracteres de calidad de fruto. Con respecto al color externo, la localidad de Málaga mostró los frutos menos luminosos, con menor saturación y tono.

## 2. Evaluación de ILs y mapeo de QTLs

Los caracteres estudiados los podemos dividir en dos grandes grupos: las evaluaciones descriptivas realizadas durante el ciclo de cultivo (apartado 2.1) y caracteres cuantitativos de frutos maduros obtenidos por métodos analíticos o análisis de imagen (apartado 2.2). En ambos casos enfatizamos aquellos caracteres que mostraron diferencias consistentes entre ILs y MM en al menos dos de las localizaciones de ensayo.

### 2.1 Caracteres descriptivos

Las mediciones de los caracteres descriptivos fueron realizadas en el bloque en el que cada réplica estaba compuesta por tres plantas de la misma IL en Alginet y Orihuela. El propósito fue tener una primera aproximación de algunos caracteres cualitativos interesantes y contrastantes entre cada grupo de ILs, así como su localización dentro de cada cromosoma.

#### 2.1.1 Precocidad (PREC)

La SP\_2-5 fue la IL que mostró más precocidad en ambas localidades (**Tabla 4-2**), necesitó en promedio 69 ddt para formar el primer fruto maduro, 8 días menos que MM. Por otra parte la SP\_11-4 fue la más tardía requiriendo entre 84 y 90 ddt para producir el primer fruto maduro, 8 días más que MM. Para este carácter se definieron dos QTLs *prec2.1* y *prec11.1* (**Fig. 4-7**).

#### 2.1.2 Vigor (VIG)

Las ILs que mostraron mayor vigor con respecto a MM fueron la SP\_2-4 y SP\_2-5. Por su parte, la SP\_3-1 fue la menos vigorosa en todas las localidades (**Fig. 4-2a**). Para este carácter se definieron dos QTLs *vig2.1* y *vig 3.1* (**Fig. 4-7**).

#### 2.1.3 Hombros verdes (HV)

Se observaron hombros de color verde oscuro en los frutos en las líneas SP\_10-1 y SP\_10-2 (**Fig. 4-2b**). Para este carácter se definió el QTL *hv 10.1* (**Fig. 4-7**).

#### 2.1.4 Densidad de tricomas (TRIC)

La IL SP\_10-3, mostró una alta densidad de tricomas en el tallo y ramas laterales (**Fig. 4-2c**). Para este carácter se definió el QTL *tric10.1* (**Fig. 4-7**).

Carácter		Localidad del ensayo					
		Alginet		Orihuela		Málaga	
		Promedio	Desv. Est.	Promedio	Desv. Est.	Promedio	Desv. Est.
Tamaño	Peso (ns)	96.18	12.63	98.79	19.50	99.58	12.38
	Perímetro (*)	19.10	0.91	19.39	1.34	20.42	1.07
	Área (*)	24.58	2.19	24.88	3.08	27.13	2.34
	Diámetro máx. (*)	5.89	0.31	5.89	0.42	6.20	0.32
	Longitud máx. (*)	5.18	0.22	5.25	0.28	5.43	0.30
Forma	FS (ns)	0.88	0.02	0.89	0.03	0.88	0.04
	PBLK (*)	0.75	0.02	0.73	0.03	0.76	0.04
	DBLK (*)	0.59	0.02	0.61	0.03	0.61	0.02
	TRI (*)	1.28	0.06	1.21	0.08	1.25	0.07
	CIR (*)	0.05	0.01	0.06	0.01	0.06	0.01
	PSH (*)	0.02	0.01	0.05	0.01	0.06	0.01
Químicos	Brix (*)	4.49	0.33	4.61	0.57	4.26	0.36
	pH (*)	3.70	0.15	4.91	0.28	4.65	0.06
	Acidez (*)	0.32	0.03	0.34	0.03	0.21	0.05
Color externo	L* (*)	39.04	1.84	43.75	1.73	37.33	2.07
	a* (*)	25.48	2.87	23.16	1.65	18.70	2.38
	b* (*)	24.73	3.03	24.93	2.74	21.76	2.91
	C* (*)	35.59	3.75	34.16	3.16	29.20	2.78
	H* (*)	44.07	2.84	47.02	3.34	49.04	4.24
Color interno	L* (*)	36.81	3.27	40.11	3.62	41.65	1.11
	a* (*)	28.18	2.61	18.80	2.48	25.51	1.59
	b* (*)	27.74	1.98	26.07	3.38	28.70	1.45
	C* (*)	39.63	2.79	32.20	2.90	38.61	1.53
	H* (*)	44.65	2.51	54.22	4.72	48.33	2.70

**Tabla 4-1. Valores promedio y desviaciones estándar por ensayo de los caracteres medidos en la variedad de tomate "Moneymaker", utilizada como control en los ensayos de fenotipado de las líneas de introgresión. (\*p<0.01) (ns diferencias no significativas entre localidades).**

### 2.1.5 Estructura del dosel (DOS)

La SP\_4-1 mostró una estructura compacta de las ramas (dosel cerrado). Definiéndose para este carácter el QTL *dos4.1* (Fig. 4-7).

### 2.1.6 Color violáceo en ramas y hojas (VIOL)

Las ILs con introgresiones en el cromosoma 5 (SP\_5-2 y SP\_5-3), mostraron un color violáceo en ramas y hojas (Fig. 4-2d), probablemente debido a un alto contenido en antocianos. Definiéndose para este carácter el QTL *viol5.1* (Fig. 4-7).

Carácter descriptivo evaluado		Línea de introgresión	QTLs definidos
Precocidad	Mayor	SP_2-5	<i>prec2.1</i>
	Menor	SP_11-4	<i>prec11.1</i>
Vigor	Mayor	SP_2-4, SP_2-5	<i>vig2.1</i>
	Menor	SP_3-1	<i>vig3.1</i>
Presencia de hombros verdes en fruto		SP_10-1, SP_10-2	<i>hv10.1</i>
Desnidad de tricomas en tallos y ramas		SP_10-3	<i>tric10.1</i>
Estructura de dosel cerrada		SP_4-1	<i>dos4.1</i>
Coloración violácea en tallo y ramas		SP_5-2, SP_5-3	<i>viol5.1</i>

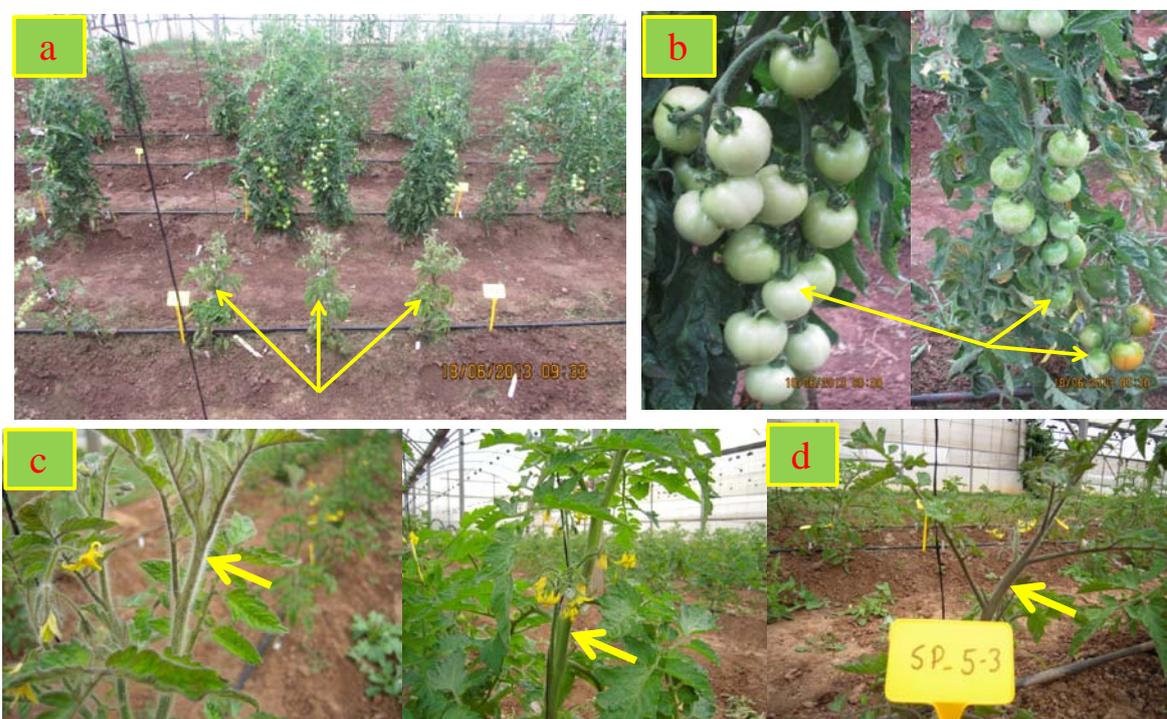
**Tabla 4-2. Caracteres descriptivos evaluados y lista de QTLs definidos en la colección de líneas de introgresión** (precocidad *prec*, vigor *vig*, hombros verdes en fruto *hv*, tricomas ramas y tallos *tric*, dosel cerrado *dos*, color violáceo en ramas y tallos *viol*). Las evaluaciones se realizaron visualmente en el bloque ocho (cada repetición estuvo formada por tres plantas de la misma IL) en las localidades de Alginet y Orihuela. Solo se presentan los datos consistentes en las dos localidades.

## 2.2 Caracteres cuantitativos de calidad de fruto

Estos caracteres los agrupamos en cinco categorías: dimensiones del fruto, forma de fruto, caracteres relacionados con el sabor (sólidos solubles y medidas de acidez), color externo y color interno.

### 2.2.1 Caracteres relacionados con las dimensiones del fruto

En general, los caracteres peso, perímetro, área, diámetro máximo y altura máxima del fruto mostraron un importante efecto del genotipo (superior al 30%). El efecto de la localidad también fue significativo en todas las variables, explicando un 23% de la varianza, valor considerado como intermedio. También se observó una interacción G x L relativamente baja (5%) para los caracteres peso, diámetro y longitud máxima ( $p < 0.001$ ), no siendo significativa para perímetro y área del fruto (**Tabla 4-3**). La heredabilidad  $h^2$  de este tipo de caracteres fue en general alta ( $h^2 > 0.4$ ) para todos los caracteres en los tres localidades. De todas las localizaciones, fue Alginet donde se obtuvieron los valores más altos (en promedio  $h^2 = 0.61$ )



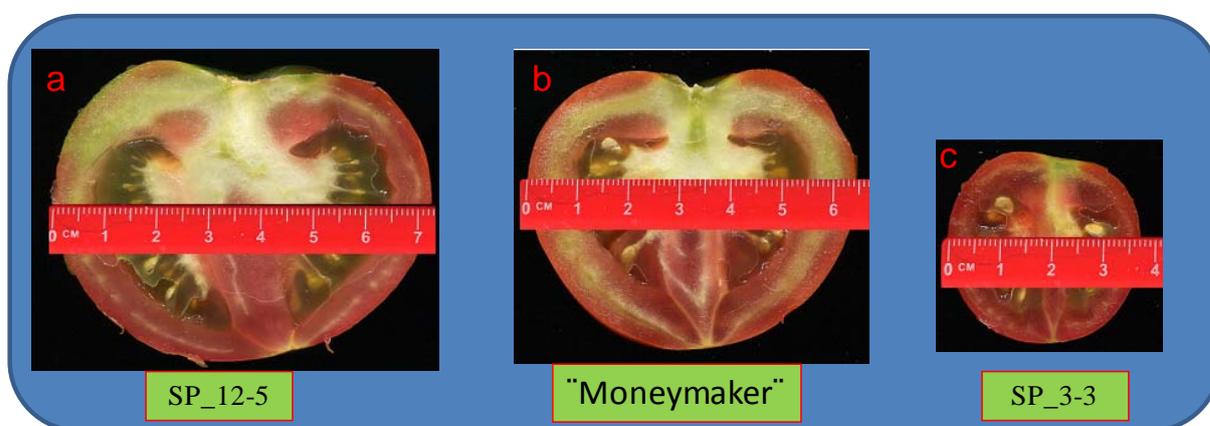
**Figura 4-2. Caracteres descriptivos evaluados en el ensayo de fenotipado de las ILs.** (a) Plantas de la SP\_3-1 mostrando el vigor más pobre de toda la genoteca (b) a la derecha frutos de las SP\_10-1 y SP\_10-2 con color verde oscuro en los hombros a la izquierda frutos de MM (c) a la izquierda la SP\_10-3 mostrando una densidad alta de tricomas en ramas a la derecha una planta de MM con pocos tricomas y tallo de color verde (d) coloración violácea en los tallos y ramas de la SP\_5-3.

seguido de Orihuela ( $h^2 = 0.5$ ) y Málaga ( $h^2 = 0.43$ ) (**Tabla 4-4**). Las correlaciones entre los diferentes caracteres relacionados con la dimensión de fruto fueron significativos ( $p < 0.005$ ), con valores positivos y altos en las tres localidades (**Anexo 4-2**). Con respecto a las

correlaciones con otros caracteres, se observó correlación negativa significativa con FS y positiva con CIR y PSH, mientras que para los caracteres relacionados con el sabor y el color del fruto fueron bajas o no significativas ( $p < 0.005$ ).

### 2.2.1.1 Peso del fruto (FW)

El peso promedio de fruto en la colección de ILs fue de 86.23 gr (11.3% menor que MM). El rango de variación de ese carácter en la población varió entre 47.81 g (SP\_3-3) y 117.86 g (SP\_12-5) (**Fig. 4-3**). En cuanto a las localidades, en Málaga se observó el mayor FW (100.7 g) seguido de Orihuela (81.47 g) y Alginet (76.61 g). La heredabilidad fue relativamente alta, siendo el promedio de  $h^2 = 0.53$ .



**Figura 4-3. Efecto contrastante de QTLs FW con respecto al control.** (a) Fruto de la SP\_12-5 que presentó los valores máximos de FW, (b) fruto del parental MM, (c) fruto de la SP\_3-3 que mostró los valores más bajos en los valores de FW.

Prácticamente todas las ILs mostraron diferencias significativas en FW al menos en una localidad ( $p < 0.05$ ) (**Anexo 4-3**). En general las introgresiones actuaron disminuyendo el FW, aunque algunas líneas con introgresiones en los cromosomas 5, 6, 10, 11 y 12 aumentaron el FW al menos en alguna localidad. Tres ILs (SP\_1-2, SP\_2-5, SP\_3-3) presentaron diferencias significativas en las tres localidades, disminuyendo el FW entre un 28% y 61%, dependiendo de la localidad. Por otra parte, 14 ILs disminuyeron FW en dos localidades (SP\_1-3, SP\_1-4, SP\_2-1, SP\_2-3, SP\_2-4, SP\_3-2, SP\_4-1, SP\_4-2, SP\_4-3, SP\_4-4, SP\_7-3, SP\_10-2, SP\_11-2). En cambio la SP\_12-5 aumentó el peso FW del fruto entre 21% y 31% respecto a MM. Adicionalmente, 19 ILs mostraron diferencias significativas, pero únicamente en una localidad.

Atendiendo a las diferencias significativas entre las ILs y MM, que eran consistentes en al menos dos localidades, se definieron 11 QTLs para FW (*fw1.1*, *fw1.2*, *fw2.1*, *fw2.2*, *fw3.1*, *fw3.2*, *fw4.1*, *fw7.1*, *fw10.1*, *fw11.1* y *fw12.1*) (Fig. 4-6), de los cuales diez disminuyen el carácter (con un valor mínimo del 27% en el QTL *fw3.1*) y solamente el *fw12.1* aumenta el peso de fruto en promedio un 20% con respecto al control (Tabla 4-5). En los cromosomas 5, 6, 8 y 9 no se detectaron QTLs relacionados con este carácter de forma consistente.

### 2.2.1.2 Perímetro (PER) y Área (AR)

El perímetro promedio de los frutos en la genoteca de ILs fue de 18.71cm, el cual representa un 4.73% menos del valor promedio de MM (19.63 cm). Por su parte, el área promedio de la sección longitudinal de los frutos de las ILs fue de 23.3cm<sup>2</sup> que representa un 8.74% menor a MM (Tabla 4-1). El rango varió entre 15.4 cm de perímetro y 16.08 cm<sup>2</sup> de área (SP\_3-3) y hasta 20.9 cm de perímetro y 28.56 cm<sup>2</sup> de área (SP\_12-5) (Fig. 4-3).

Las ILs SP\_1-2, SP\_2-4, SP\_2-5, SP\_3-3, presentaron diferencias significativas en las tres localidades para ambos caracteres ( $p < 0.05$ ), mientras que 8 y 9 ILs mostraron diferencias estadísticas significativas para PER y AR respectivamente en dos localidades. Adicionalmente, 16 y 18 ILs mostraron diferencias significativas en PER y AR, respectivamente en una única localidad (Anexo 4-3).

Se pudieron definir 10 QTLs para PER (*per1.1*, *per1.2*, *per2.1*, *per2.2*, *per3.1*, *per3.2*, *per4.1*, *per7.1*, *per10.1* y *per11.1*) y 9 para AR (*ar1.1*, *ar1.2*, *ar2.1*, *ar2.2*, *ar3.1*, *ar3.2*, *ar4.1*, *ar10.1* y *ar11.1*) (Fig. 4-7), en ambos casos, los QTLs disminuyen el valor del carácter, siendo las ILs que contienen los QTLs *per3.1* y *ar3.1* las que muestran los menores valores de perímetro de fruto y área del fruto (Tabla 4-5). La posición de los QTLs de los caracteres PER y AR coincide con otros definidos en variables relacionadas con las dimensiones del fruto (Fig. 4-7).

### 2.2.1.3 Diámetro máximo del fruto (FDL)

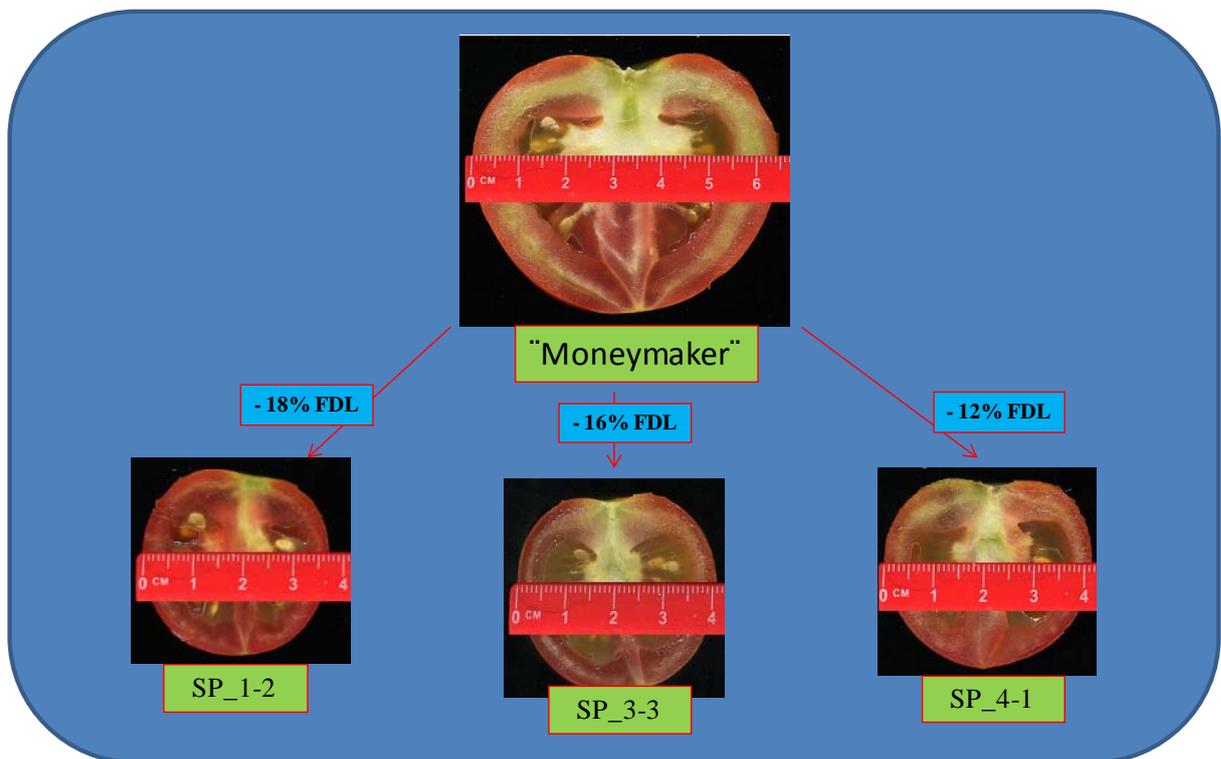
En su conjunto las ILs registraron un diámetro máximo de fruto promedio de 5.71cm con un rango entre 4.67cm en la SP\_3-3 y 6.38 cm en la SP\_12-5. Las ILs que mostraron diámetros de fruto significativamente menores al control en las tres localidades ( $p < 0.05$ ) (Fig. 4-4) fueron: SP\_1-2, SP\_2-5, SP\_3-3 y SP\_2-4. Seis ILs tuvieron valores significativamente menores al control en dos localidades (SP\_2-1, SP\_3-2, SP\_4-2, SP\_4-3, SP\_10-2 y SP\_11-2) y 15 ILs presentaron diferencias significativamente en una única localidad Dentro de estas

Carácter		Efecto (%)		
		Genotipo	Localidad	Genotipo x Localidad
Tamaño	Peso	35**	19**	5**
	Perímetro	31**	25**	Ns
	Área	31**	25**	Ns
	Diámetro máx.	30**	26**	5*
	Longitud máx.	35**	20**	5*
Forma	FS	23**	9**	9**
	PBLK	8**	15**	10**
	DBLK	15**	10**	9**
	TRI	10**	15**	Ns
	CIR	26**	15**	Ns
	PSH	6**	57**	4**
Químicos	Brix	29**	9**	9**
	pH	7**	58**	3**
	Acidez	12**	44**	5**
Color externo	L*	6**	59**	4**
	a*	20**	37**	8**
	b*	15**	24**	8**
	C*	16**	38**	7**
	H*	22**	7**	12**
Color interno	L*	11**	26**	12**
	a*	7**	50**	7**
	b*	14**	1**	13**
	C*	10**	37**	9**
	H*	7**	42**	7**

**Tabla 4-3. Efectos Genético, Localidad e Interacción (G x L), sobre el valor fenotípico de los caracteres utilizados con la genoteca de ILs, en los tres sitios estudiados. \*\* p<0.001), \* p<0.05 y ns, estadísticamente no significativo.**

Carácter		h <sup>2</sup> Localidad		
		Alginet	Orihuela	Málaga
Tamaño	Peso	0.62	0.52	0.45
	Perímetro	0.60	0.49	0.40
	Área	0.61	0.50	0.42
	Diámetro máx.	0.61	0.50	0.41
	Longitud máx.	0.62	0.49	0.49
Forma	FS	0.48	0.36	0.29
	PBLK	0.29	0.2	0.23
	DBLK	0.37	0.3	0.23
	TRI	0.34	0.15 <sup>ns</sup>	0.26
	CIR	0.47	0.49	0.35
	PSH	0.30	0.35	0.30
Químicos	Brix	0.51	0.48	0.43
	pH	0.25	0.40	0.2
	Acidez	0.37	0.34	0.46
Color externo	L*	0.37	0.26	0.23
	a*	0.51	0.55	0.32
	b *	0.41	0.29	0.23
	C *	0.48	0.46	0.25
	H*	0.42	0.37	0.33
Color interno	L*	0.44	0.49	0.24
	a*	0.39	0.45	0.29
	b *	0.38	0.31	0.23
	C *	0.40	0.36	0.24
	H*	0.36	0.42	0.29

**Tabla 4-4. Valores h<sup>2</sup> de los caracteres de calidad de fruto evaluados en la genoteca de ILs.** Todos los valores fueron estadísticamente significativos en las tres localidades estudiadas (p<0.0001), con excepción del carácter TRI (ns) no significativo.



**Figura 4-4. Diferencias en diámetro máximo de fruto (FDL) entre ILs seleccionadas.** Frutos característicos de las ILs SP\_1-2, SP\_3-3 y SP\_4-1, con disminuciones de un 18%, 16% y 12%, respectivamente, con respecto al control MM.

últimas, cabe destacar SP\_12-4 como la única línea con un diámetro de fruto mayor al control (Anexo 4-3).

Se definieron 8 QTLs que disminuyen el diámetro máximo del fruto (*fdl1.1*, *fdl2.1*, *fdl2.2*, *fdl3.1*, *fdl3.2*, *fdl4.1*, *fdl10.1* y *fdl11.1*) siendo las plantas que contienen los QTLs *fdl1.1*, *fdl3.1* y *fdl4.1* las que dieron frutos con el menor diámetro con respecto al control, la posición de estos QTLs coincide con otros definidos en variables relacionadas con las dimensiones del fruto (Tabla 4-5, Fig. 4-7).

#### 2.2.1.4 Longitud máxima del fruto (FLL)

El valor promedio de FLL de los frutos en la genoteca completa fue de 5.05 cm (4.5% menor que MM), con un rango que varió entre 4.27 (SP\_3-3) y 5.55 cm (SP\_12-5). Cuatro ILs presentaron diferencias significativas en las tres localidades (SP\_1-2, SP\_2-5, SP\_3-3 y SP\_3-2), ocho ILs en dos localidades (SP\_1-3, SP\_1-4, SP\_4-1, SP\_4-2, SP\_7-2, SP\_7-3, SP\_10-2 y SP\_11-2) y 18 IL fueron significativas en una única localidad ( $p < 0.05$ ) (Anexo 4-3). Se definieron 9 QTLs que disminuyen FLL (*fu1.1*, *fu1.2*, *fu2.1*, *fu3.1*, *fu3.2*, *fu4.1*, *fu7.1*, *fu10.1*

y *fl11.1*), la posición de los QTLs coincide con otros definidos en variables relacionadas con las dimensiones del fruto (**Fig. 4-7**). Las plantas que contienen el QTL *fl1.1*, *fl2.1* *fl3.1* mostraron los valores más bajos (**Tabla 4-5, Fig. 4-5**).

En resumen, se encontró que la colección de ILs presentaba de una buena fuente de variabilidad para los caracteres de dimensión de fruto estudiados. La mayoría de los cromosomas mostraron efectos sobre diferentes componentes relacionados con la dimensión del fruto, generalmente disminuyendo el tamaño comparado con MM. Los QTLs con los efectos más importantes se asociaron en los cromosomas 1, 2, 3, y 4, en los cromosomas 5, 6, 8 y 9 no se detectaron QTLs relacionados con estos caracteres, y el cromosoma 12 fue el único donde se localizó un QTL que aumenta el peso con respecto a MM.

## 2.2.2 Caracteres de forma del fruto

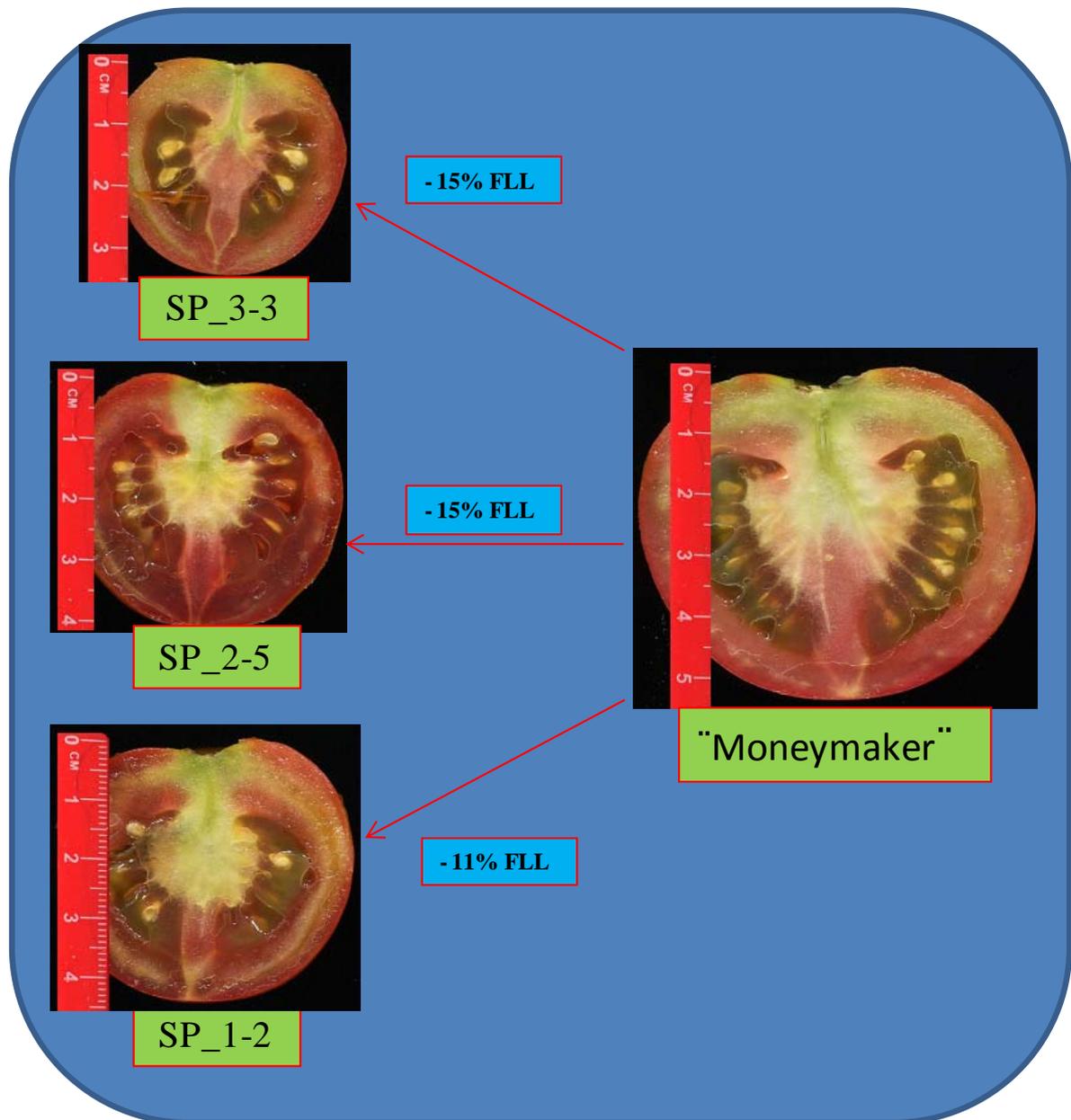
El efecto promedio del genotipo en los caracteres de forma de fruto en las tres localidades fue de aproximadamente 16%, observándose un mayor efecto del genotipo en el índice forma de fruto circular (CIR), con un 26%, seguido del índice de forma de fruto (FS) con un 23%, bloque distal (DBLK) con un 15%, forma de fruto triangular (TRI) con 10%, bloque proximal (PBLK) con un 8% y altura de hombros (PSH) con un 6%. Para todos estos caracteres el efecto de localidad y la interacción GxL fue bajo o nulo con la excepción del carácter altura de hombros (PSH), para el que el efecto de localidad fue alto (57%) ( $p < 0.0001$ ) (**Tabla 4-3**). Los valores de  $h^2$  fueron significativos ( $p < 0.0001$ ) para todas los caracteres menos para TRI, siendo altos ( $h^2 > 0.4$ ) para FS y CIR e intermedios ( $0.20 < h^2 < 0.40$ ) para DBLK, PBLK y PSH. También se pudo observar que  $h^2$  varió entre las diferentes localizaciones, observándose valores más altos en Alginet, seguido de Orihuela y los menores en Málaga (**Tabla 4-4**).

Entre los caracteres de forma de fruto se observaron correlaciones significativas altas, siendo las más importantes entre TRI con DBLK, PBLK y CIR en las tres localidades ( $r > 0.8$ ). CIR y FS también están altamente correlacionados ( $-0.9 < r < -0.8$ ). Entre los caracteres de forma y el resto, las correlaciones más importantes se observaron entre FS y CIR con FDL (**Anexo 4-2**).

### 2.2.2.1 Índice de forma de fruto (FS)

El promedio general de FS fue muy similar tanto en la genoteca de ILs como en MM (FS=0.89) (**Tabla 4-1**). El rango de valores estuvo entre 0.82 (SP\_6-2) y 0.93 (SP\_4-2) (FS < 1 frutos achatados). La  $h^2$  puede considerarse media-alta (entre 0.489 y 0.29, dependiendo de localidad) (**Tabla 4-4**). La SP\_6-2 mostró frutos más achatados que MM en las tres

localidades, mientras SP\_10-5 produjo frutos más achatados en dos localidades ( $FS=0.83$ ) y SP\_4-2 más alargados ( $FS=0.93$ ) en dos localidades. Adicionalmente, cuatro ILs mostraron diferente FS en una única localidad (**Anexo 4-3**).



**Figura 4-5. Efectos de QTLs implicados en la longitud máxima del fruto (FLL) representados en 3 ILs.** Frutos característicos de las ILs SP\_3-3, SP\_2-5 y SP\_1-2, con disminuciones de un 15%, 12% y 11%, respectivamente, con respecto al control MM.

Por tanto, se definieron 3 QTLs, dos de ellos (*fs6.1* y *fs10.1*) induciendo frutos más achatados y *fs4.1* (Tabla 4-5 y Fig. 4-7), que induce frutos más alargados que MM, acercándose más a la forma redonda perfecta (FS=1). (Fig. 4-6a).

#### 2.2.2.2 Bloque distal (DBLK), Bloque proximal (PBLK) e Índice de forma de fruto triangular (TRI)

Los valores promedio de estas tres variables en la genoteca de ILs fueron muy similares a los obtenidos por el control MM, en los tres sitios (Tabla 4-1). Los valores de  $h^2$  promedio en las distintas localidades fueron de 0.24 para DBLK, 0.30 para PBLK y 0.25 para TRI (Tabla 4-4), lo que se clasifica como medio-bajo. Se observaron diferencias significativas solamente en una única localidad en las ILs: SP\_5-4, SP\_6-3, SP\_2-4, SP\_2-5, SP\_6-4 para PBLK, en las ILs SP\_5-4, SP\_6-3, SP\_12-2, SP\_12-1 para DBLK, y en las ILs SP\_2-5, SP\_12-2, SP\_12-1, SP\_2-3 y SP\_1-2 para TRI (Anexo 4-3). No fue posible definir QTLs consistentes en más de una localidad para estos caracteres.

#### 2.2.2.3 Forma de fruto circular (CIR)

El valor promedio del CIR en la genoteca de ILs fue 0.06, con un rango de valores de 0.04 para SP\_2-4 (34% más redondos que MM) y 0.08 para SP\_10-5 (frutos 50% menos redondo que MM). La  $h^2$  promedio del carácter fue alta en las tres localidades (Tabla 4-4). Las ILs SP\_10-5 y SP\_6-2 presentaron valores significativamente mayores que MM en las tres localidades y la SP\_2-4 fue significativamente menor que MM en dos localidades. Cinco ILs mostraron diferencias significativas con MM para este carácter, pero sólo en una única localidad (Anexo 4-3).

Se definieron por tanto, tres QTLs consistentes relacionados con CIR (*cir2.1*, *cir6.1*, *cir10.1*) (Fig. 4-7), uno de ellos (*cir2.1*) induce la producción de frutos hasta un 26% más redondos con respecto a MM (ver imagen 2d), mientras que *cir6.1*, *cir10.1* inducen la producción de frutos hasta un 46% menos redondos que MM (Tabla 4-5, Fig. 4-6b).

#### 2.2.2.4 Altura de hombro en el fruto (PSH)

Los valores promedio de PSH del conjunto de la genoteca de ILs fueron similares al control (PSH=0.04). La  $h^2$  promedio del carácter fue de 0.31. Solamente dos ILs (SP\_1-2 y SP\_3-3) presentaron diferencias significativas en las tres localizaciones, mostrando en promedio un 42% del valor de PSH que caracteriza MM. Otras ILs fueron diferentes estadísticamente, pero en una única localidad (SP\_2-4, SP\_3-4, SP\_7-4, SP\_8-1 y SP\_11-1) y con valores de PSH menores que MM. Por el contrario, la SP\_10-5 presentó mayor PSH que MM (entre

32% y 25% más) aunque estas diferencias fueron significativas únicamente en Orihuela (**Anexo 4-3**). Se definieron consecuentemente dos QTLs consistentes relacionados con PSH (*psh1.1* y *psh3.1*). Estos QTLs inducen una reducción de PSH de alrededor del 21% (**Fig. 4-6c**).

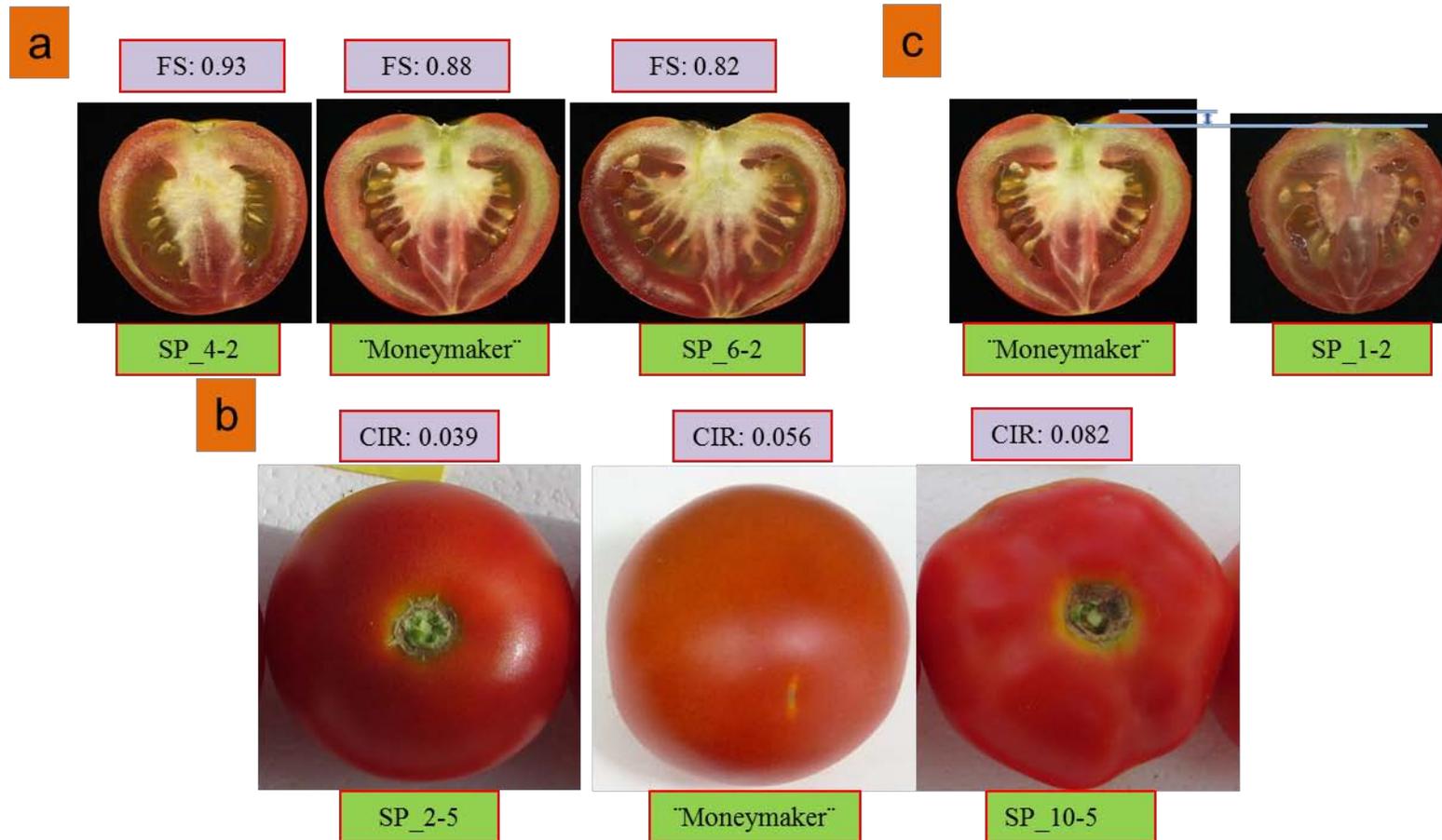
En resumen se definieron ocho QTLs relacionados con caracteres de forma de fruto en seis cromosomas. Los QTLs *fs4.1*, *fs6.1* y *fs10.1* que inducen variaciones en la relación longitud/del fruto, los QTLs *cir2.1*, *cir6.1* y *cir10.1* que alterarían la forma circular y los *psh1.1* y *psh3.1* están involucrados en la altura de hombros. Los caracteres DBLK, PBLK y TRI mostraron poca variabilidad y no permitieron detectar QTLs para ellos. Los QTLs FS y CIR son coincidentes en la misma región de los cromosomas 6 y 10, lo que indica que posiblemente estén reflejando los efectos de los mismos genes, sin embargo, los QTLs para PSH se localizan en cromosomas distintos, indicando un control genético distinto con respecto a FS y CIR.

### **2.2.3 Caracteres asociados a medidas analíticas relacionadas con el sabor**

El efecto genético promedio fue alto para SSC (29%) y bajo para pH y TA (7 y 12% respectivamente). El efecto localidad fue alto para TA y pH (44 % y 58% respectivamente) y bajo para SCC (9%). La interacción (G x L) fue significativa para los tres caracteres ( $p < 0.0001$ ) (**Tabla 4-3**). Los valores de  $h^2$  promedio para SSC fueron altos ( $h^2 > 0.51$ ) en las tres localidades e intermedio para los caracteres pH y TA (**Tabla 4-4**). TA correlaciona positivamente con SSC ( $0.30 < r < 0.44$ ) y negativamente con pH ( $-0.26 > r > -0.40$ ), no existió correlación entre los SSC y el pH. Se presentaron correlaciones bajas y negativas ( $p < 0.005$ ), únicamente en Málaga. SSC y pH correlacionaron negativamente con el color de fruto externo (**Anexo 4-2**).

#### **2.2.3.1 Contenido en Sólidos Solubles (SSC)**

El promedio general de SSC en la genoteca de ILs fue muy similar al presentado por MM (4.48 °brix) (**Tabla 4-1**). El rango de valores varió entre 3.85° (SP\_4-3) y 5.1° brix (SP\_10-6). La SP\_3-3 presentó un contenido SSC estadísticamente mayor al control ( $p < 0.05$ ) en las tres localidades (**Anexo 4-3**). Adicionalmente, seis ILs mostraron valores de SSC



**Figura 4-6. Diferencias entre ILs en varios caracteres relacionados con la forma del fruto.** (a) Efecto del *fsl4.1* y *fsl6.1* sobre la forma del fruto en las SP\_4-2 y SP\_6-2. (b) Efecto los QTLs *cir4.1* y *cir6.1* sobre la variable CIR en las SP\_2-5 y SP\_10-5. (c) Efecto del *psh1.1* sobre la altura de los hombros en la SP\_1-2.

estadísticamente significativos en dos localidades. De estas seis ILs, cuatro tuvieron valores mayores a MM (SP\_2-1, SP\_2-2, SP\_2-5 y SP\_9-1) y dos menores a MM (SP\_4-3 y la SP\_5-2). 10 ILs adicionales fueron significativamente diferentes a MM testigo ( $P < 0.05$ ) en una única localidad sitio, seis de ellas con valores mayores y cuatro con valores menores.

Se definieron de acuerdo con estos resultados, cinco QTLs consistentes relacionados con el contenido de sólidos solubles (*ssc2.1*, *ssc3.1*, *ssc4.1*, *ssc5.1*, *ssc9.1*), de los cuales los QTLs *ssc2.1*, *ssc3.1* y *ssc9.1* aumentaron SSC hasta en un 19% y los QTLs *ssc4.1*, *ssc5.1* disminuyeron SSC hasta en un 20% con respecto a MM (Tabla 4-5, Fig.4-7).

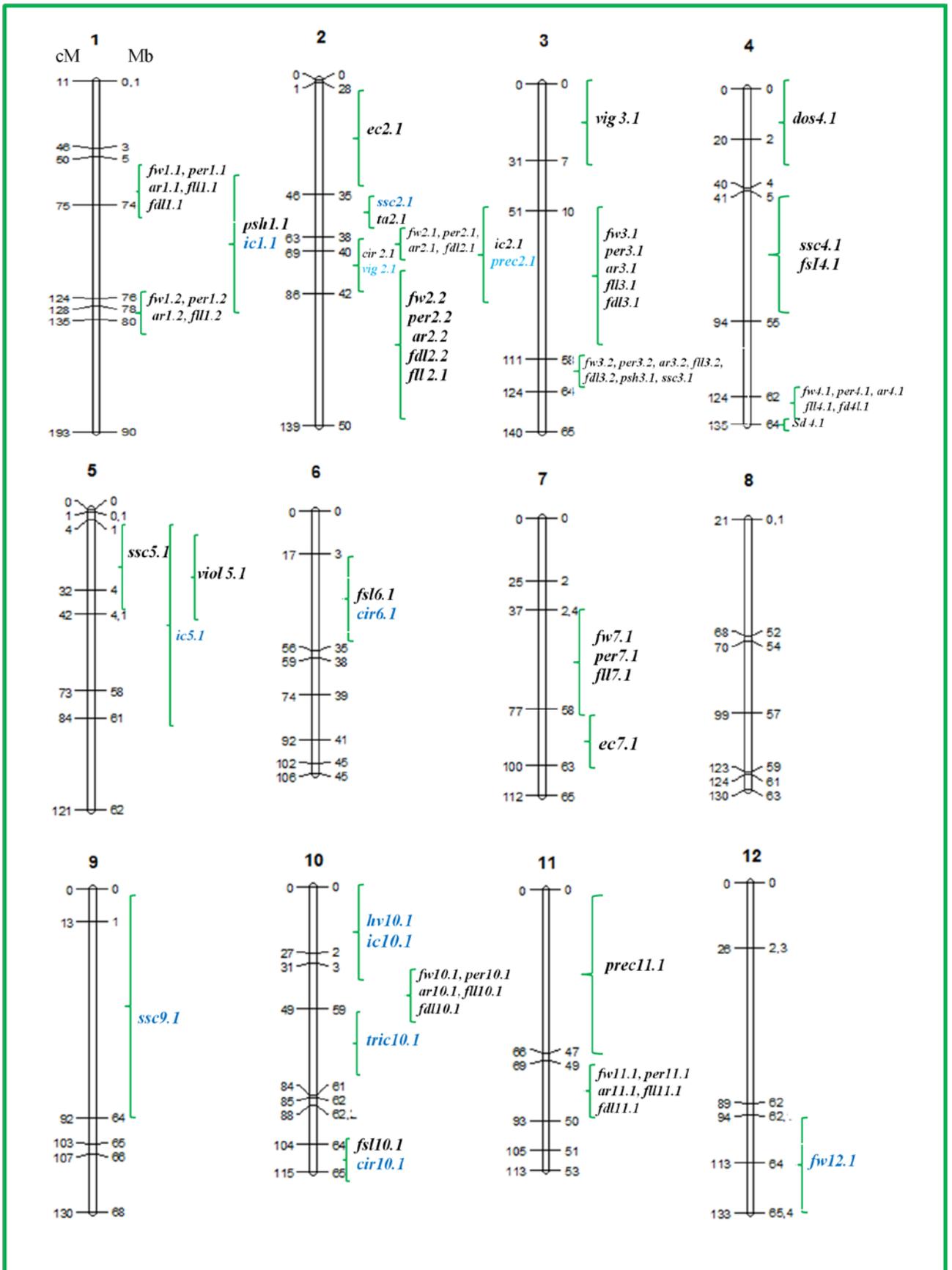
### 2.2.3.2 Acidez inicial (pH)

El promedio general de pH en la genoteca fue de 4.36 con un rango entre 4.22 (SP\_3-2) y 4.56 (SP\_2-4). No se encontraron diferencias estadísticas significativas consistentes entre las ILs y el control (Anexo 4-3).

### 2.2.3.3 Acidez valorable (TA)

El promedio general de TA en la genoteca de ILs fue idéntico a MM (0.29) (Tabla 4-1), el rango de valores se distribuyó entre 0.24 (SP\_2-4) y 0.35 (SP\_3-3). La SP\_2-4 fue la única que mostró diferencias significativas respecto a MM en más de una localidad, disminuyendo cerca del 19% con respecto a MM. 10 ILs adicionales aumentaron o disminuyeron considerablemente los valores de TA en una sola localidad. Dentro de estas cabe destacar la ILs SP\_1-4, SP\_3-2 y SP\_4-1 que en promedio sus frutos mostraron valores de TA un 33% mayor que MM (Anexo 4-3). Se encontró un único QTL consistente en la genoteca de ILs, *ta2.1*, localizado en el cromosoma 2 (Fig.4-7), el cual disminuye TA hasta en un 21% con respecto a MM (Tabla 4-5).

En resumen se definieron 6 QTLs para los caracteres relacionados con características organolépticas: cinco QTLs para SSC (tres aumentan y dos disminuyen), y un QTL que disminuye la TA. Con respecto al pH no se pudieron definir QTLs consistentes



**Figura 4-7. Localización de los QTLs** (precocidad *prec*, vigor *vig*, hombros verdes en fruto *hv*, tricomas ramas y tallos *tric*, dosel cerrado *dos*, color violáceo en ramas y tallos *viol*, peso de fruto *fw*, perímetro de fruto *per*, área de fruto *ar*, diámetro de fruto *fdl*, longitud del fruto *fll*, índice de forma de fruto *fs*, forma de fruto circular *cir*, altura de hombros *psh*, sólidos solubles *ssc*, acidez *ta*, color externo *ec*, color interno *ic*), en el mapa de bins realizado con las líneas de introgresión desarrolladas entre la variedad comercial "Moneymaker" y el silvestre TO-937. Al costado izquierdo de cada cromosoma están las distancias genéticas (cM) y al derecho las distancias físicas (Mb). Los QTLs de color negro disminuyen el valor y los de color azul lo aumentan (Mapa basado en una población de RILs genotipadas con la micromatriz de SOLCAP (Fernandez-del-Carmen et al. 2011)).

#### 2.2.4 Color externo (EC)

El efecto genético observado en los componentes del caracteres del color externo varió entre valores medios ( $a^*$  y  $H^*$ ) y bajos ( $L^*$ ,  $b^*$  y  $C^*$ ). El efecto localidad en la mayoría de ellos fue alto ( $> 24\%$ ) con excepción del componente  $H^*$  ( $7\%$ ). Por su parte las interacciones ( $G \times L$ ) fueron siempre bajas ( $< 12\%$ ) (**Tabla 4-3**). Los valores de heredabilidad se distribuyeron entre altos e intermedios para todos los componentes del carácter color externo. Los valores de la  $h^2$  fueron altos en todos los componentes del carácter en la localidad de Alginet, mientras que en Orihuela y Málaga los valores fueron en su mayoría intermedios (**Tabla 4-4**).

Las correlaciones dentro de los componentes de EC fueron en su mayoría significativas, positivas y altas ( $0.50 < r < 0.97$ ), con la excepción del componente  $H^*$  que se correlacionó negativamente con  $a^*$  ( $r = -0.60$ ). EC correlacionó positivamente con dimensiones de fruto y color interno y negativamente con los caracteres relacionados con el sabor SSC, TA, pH, aunque estas correlaciones fueron de magnitud baja. No se observó correlación con los caracteres de forma de fruto (**Anexo 4-2**).

De todos los componentes de EC únicamente se observaron diferencias significativas entre ILs y MM en más de una localidad para en el componente  $a^*$  (**Anexo 4-3**). El promedio general de  $a^*$  en la genoteca de ILs fue de 25.41, con un rango entre 17.65 (SP\_7-3) y 25.73 (SP\_3-2). La SP\_7-3 mostró valores de  $a^*$  significativamente menor al control en los tres sitios, disminuyendo el color rojo de los frutos en un 20% con respecto a MM. Por su parte, las SP\_2-3 y SP\_2-1 mostraron diferencias significativas con valores menores que el control en dos localidades. Un total de 47 ILs mostraron diferencias significativas con MM en una única localidad para alguno de los componentes de EC: 5 ( $L^*$ ), 17 ( $a^*$ ), 5 ( $b^*$ ), 11 ( $C^*$ ) y 9 ( $H^*$ ) (**Anexo 4-3**).

Carácter	QTL	Cr	Posición del QTL (cM)	Nº Localidades detectado	Efecto de TO-937 respecto a MM (%)*	Posible QTL ortólogo en otros <i>S. pimpinellifolium</i> (**)
Peso de fruto (FW)	<i>fw1.1</i>	1	57-75	3	-21	b, e, i, g
	<i>fw1.2</i>	1	77-128	2	-17	nd
	<i>fw2.1</i>	2	47-69	2	-23	nd
	<i>fw2.2</i>	2	69-138	3	-20	a, b, c, e, g, h
	<i>fw3.1</i>	3	51 - 111	3	-27	f, g
	<i>fw3.2</i>	3	111-124	2	-25	a, c, f, g, h
	<i>fw4.1</i>	4	124 - 135	2	-24	nd
	<i>fw7.1</i>	7	39 - 77	2	-18	a, g
	<i>fw10.1</i>	10	33 - 49	2	-23	nd
	<i>fw11.1</i>	11	69 - 93	2	-17	a, b, e
<i>fw12.1</i>	12	94 - 133	2	+20	nd	
Perímetro del fruto (PER)	<i>per1.1</i>	1	57-75	3	-9	nd
	<i>per1.2</i>	1	77-128	2	-8	nd
	<i>per2.1</i>	2	47-69	2	-9	nd
	<i>per2.2</i>	2	69-138	3	-8	nd
	<i>per3.1</i>	3	51 - 111	3	-11	nd
	<i>per3.2</i>	3	111-124	2	-11	nd
	<i>per4.1</i>	4	124 - 135	2	-6	nd
	<i>per7.1</i>	7	39 - 77	2	-7	nd
	<i>per10.1</i>	10	33 - 49	2	-8	nd
	<i>per11.1</i>	11	69 - 93	2	-8	nd
Área de fruto (AR)	<i>ar1.1</i>	1	57-75	3	-14	nd
	<i>ar1.2</i>	1	77-128	2	-15	nd
	<i>ar2.1</i>	2	47-69	2	-17	nd
	<i>ar2.2</i>	2	69-138	3	-13	nd
	<i>ar3.1</i>	3	51 - 111	3	-19	nd
	<i>ar3.2</i>	3	111-124	2	-18	nd
	<i>ar4.1</i>	4	124 - 135	2	-26	nd

	<i>ar10.1</i>	10	33 - 49	2	-15	nd
	<i>ar11.1</i>	11	69 - 93	2	-14	nd
Diámetro máximo de fruto (FDL)	<i>fdl1.1</i>	1	57-75	3	-18	b, e, g
	<i>fdl2.1</i>	2	47-69	2	-10	nd
	<i>fdl2.2</i>	2	69-138	3	-6	b, d, e, g, h
	<i>fdl3.1</i>	3	51 - 111	3	-7	b, g, h
	<i>fdl3.2</i>	3	111-124	2	-15	g
	<i>fdl4.1</i>	4	124 - 135	2	-12	nd
	<i>fdl10.1</i>	10	33 - 49	2	-7	nd
	<i>fdl11.1</i>	11	69 - 93	2	-6	b, g
Longitud máximo de fruto (FLL)	<i>fl1.1</i>	1	57-75	3	-6	b, g
	<i>fl1.2</i>	1	77-128	2	-9	b
	<i>fl2.1</i>	2	69-138	3	-15	b
	<i>fl3.1</i>	3	51 - 111	3	-11	b, d
	<i>fl3.2</i>	3	111-124	2	-8	nd
	<i>fl4.1</i>	4	124 - 135	2	-10	nd
	<i>fl7.1</i>	7	39 - 77	2	-9	nd
	<i>fl10.1</i>	10	33 - 49	2	-9	nd
	<i>fl11.1</i>	11	69 - 93	2	-9	b
Forma circular fruto (CIR)	<i>cir2.1</i>	2	69 - 86	2	-26	a, c, e, f
	<i>cir6.1</i>	6	17 - 56	3	+46	a
	<i>cir10.1</i>	10	104 - 115	3	+42	a, c, g
Índice de forma de fruto (FS)	<i>fs4.1</i>	4	41 - 94	2	+6	nd
	<i>fs6.1</i>	6	17 - 56	3	-7	a
	<i>fs10.1</i>	10	104 - 115	2	-6	c, g
Altura de hombros (PSH)	<i>psh1.1</i>	1	50 - 135	3	-43	nd
	<i>psh3.1</i>	3	111 - 124	3	-42	i
Grados Brix (SSC)	<i>ssc2.1</i>	2	46 - 63	2	+14	f, g
	<i>ssc3.1</i>	3	111 - 124	3	+15	g
	<i>ssc4.1</i>	4	41 - 94	2	-13	nd
	<i>ssc5.1</i>	5	4 - 32	2	-10	f
	<i>ssc9.1</i>	9	0 - 92	2	+13	d, e, g

Acidez (TA)	<i>ta2.1</i>	2	46 - 63	2	-17	h
Color externo (EC)	<i>ec2.1</i>	2	1 - 46	2	-10	c, e
	<i>ec7.1</i>	7	77 - 100	2	-20	nd
Color interno (IC)	<i>ic1.1</i>	1	50 - 128	2	+2	c, e
	<i>ic2.1</i>	2	63 - 86	2	-11	nd
	<i>ic5.1</i>	5	4 - 84	2	+8	nd
	<i>ic10.1</i>	10	0 - 49	2	+18	nd

**Tabla 4-5. Lista de QTLs para calidad de fruto** (peso de fruto *fw*, perímetro de fruto *per*, área de fruto *ar*, diámetro de fruto *fdl*, longitud del fruto *fll*, índice de forma de fruto *fs*, forma de fruto circular *cir*, altura de hombros *psh*, sólidos solubles *ssc*, acidez *ta*, color externo *ec*, color interno *ic*), detectados en la genoteca de líneas de introgresión, su ubicación dentro del mapa de bins, número de localidades donde se detecto, efecto de TO-937 sobre el carácter con respecto al control "Moneymaker" (\*efectos calculados en ILs que presentaron diferencias estadísticas en más de una localidad  $p < 0.05$ ) y posibles ortólogos (\*\* en otras poblaciones generadas a partir de accesiones de *S. pimpinellifolium*): (a) van der Knaap y Tanksley (2003), (b) Lippman y Tanksley (2001), (c) Bernacchi et al. (1998), (d) Doganlar et al. (2002), (e) Grandillo y Tanksley (1996), (f) Tanksley et al. (1996), (g) Chen et al. (1999), (h) Saliba-Colombani et al. (2001) detectados en *S. lycopersicum* var *ceraciforme*, (i) van der Knaap et al. (2002).

En resumen, se definieron dos QTLs consistentes en más de una localidad que están relacionados con alguno de los parámetros que definen el EC: concretamente, *ec2.1* y *ec7.1* que disminuyen el valor del componente  $a^*$  con respecto a MM (**Tabla 4-5, Fig. 4-7 y Fig. 4-8a**).

### 2.2.5 Color interno (IC)

Los resultados de los parámetros relacionados con el IC se encuentran en las **Tablas 4-1- a 4-4 y Anexo 4-2**). Según se puede observar, el efecto genético promedio sobre los componentes de IC fue bajo en todos los casos (9.8% de media), con un rango de variación que oscila entre 7% para los componentes  $a^*$  y  $C^*$  y 14% para  $b^*$ . Sin embargo, el efecto de la localidad fue alto para casi todos ( $> 26\%$ ), con la excepción de  $b^*$  (1%). La interacción (G x L) para todos los componentes de IC fue baja ( $< 13\%$ ) (**Tabla 4-3**). Los valores de  $h^2$  fueron altos en Orihuela (0.41), y medios para Alginet (0.39) y Málaga (0.26) (**Tabla 4-4**). Se encontraron correlaciones altas y positivas entre la los componentes  $C^*$  con  $a^*$  ( $r=0.91$ ) y  $C^*$  con  $b^*$  ( $r=0.83$ ) y negativas entre  $H^*$  con  $a^*$  ( $r= -0.82$ ). No se encontraron correlaciones consistentes entre IC y los demás caracteres (**Anexo 4-2**).

Los valores promedio de cada una de los componentes de IC en la genoteca no difieren mucho del control (**Tabla 4-1, Anexo 4-4**). El componente luminosidad (**L\***) presentó un promedio en la genoteca de ILs de 39.2 con un rango entre 35.28 (SP\_7-4) y 42.56 (SP\_9-1). La SP\_5-2 presentó un aumento en la luminosidad de un 8%, mientras que en SP\_2-5 disminuyó la luminosidad en un 11% respecto al control. Estos resultados se reprodujeron en más de una localización.

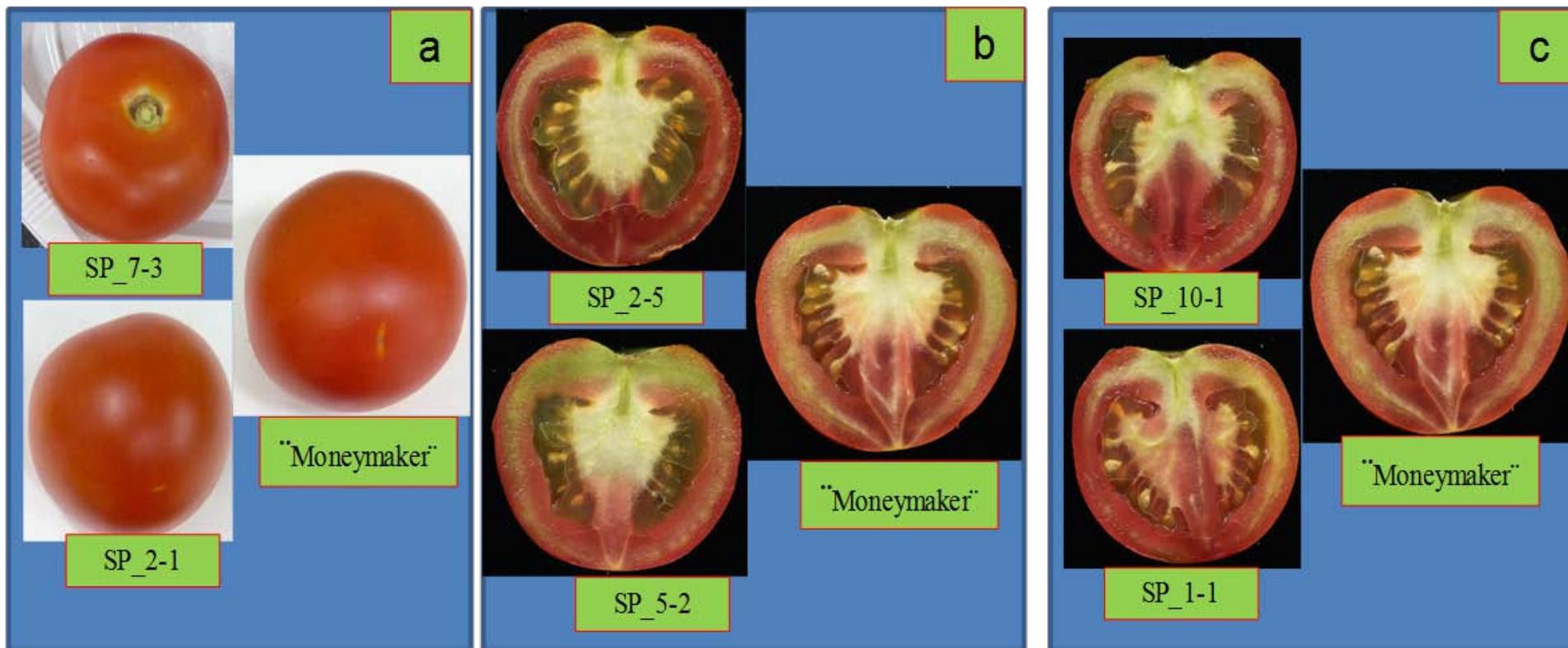
Con respecto al componente **a\***, el promedio general fue de 24.72 con un rango de 21.44 (SP\_2-4) y 28.37 (SP\_4-4). La IL SP\_10-1 mostró efecto significativo en dos localidades, aumentando su valor en un 18% con respecto al control MM y por tanto la componente del color rojo interno.

Por otra parte, el promedio general del componente **H\*** en la genoteca de ILs fue 48.42 con un rango de variación que osciló entre 45.46 (SP\_4-4) y 53.84 (SP\_2-4). La IL SP\_1-2 presentó diferencias significativas en dos localidades, aumentando el valor de **H\*** (tono) de los frutos un 7% con respecto a MM.

Finalmente, un total de 50 ILs mostraron diferencias significativas en una única localidad, desglosado de la siguiente forma: 9 (**L\***), 10 (**a\***), 12 (**b\***), 11 (**C\***) y 8 (**H\***) (**Anexo 4-3**).

Se definieron con estos resultados cuatro QTLs (**Fig. 4-7**) relacionados con el color interno del fruto: **ic1.1** que aumenta el tono (**H\***), **ic2.1** y **ic5.1** que disminuye y aumentan la luminosidad (**L\***) respectivamente y **ic10.1** que aumenta la componente de color rojo (**a\***) interno de los frutos con respecto a MM (**Fig. 4-8b-c**).

En resumen, se ha logrado mapear un total de 75 QTLs con efecto consistente (confirmado en dos o más localizaciones), distribuidos entre caracteres descriptivos y cuantitativos relacionados con calidad de fruto. La cantidad de QTL definidos por carácter varió entre uno para TA hasta 11 QTLs para FW (**Tabla 4-5**). También se encontraron diferencias en el número de QTLs identificados por cromosoma que varió entre un QTL en el cromosoma 9 y 16 QTLs en el cromosoma 2, no detectándose ningún QTL en el cromosoma 8.



**Figura 4-8. Efectos de QTLs de color externo e interno.** (a) Imágenes de frutos pertenecientes a la SP\_7-3 y SP\_2-1 que contienen los QTLs *ec2.1* y *ec7.1* que disminuyen  $a^*$  (determina el color rojo) entre el 10% y 20% respectivamente con respecto a MM). (b) Imágenes de frutos pertenecientes a la SP\_2-5 (contiene *ic2.1* baja  $L^*$  (luminosidad) en un 11%) y SP\_5-2 (contiene *ic5.1* sube  $L^*$  en un 8%). (c) Imágenes de frutos pertenecientes a la SP\_10-1 (contiene *ic10.1* sube  $a^*$  (color rojo) en un 18%) y SP\_1-1 (contiene *ic1.1* sube  $H^*$  (saturación del color) en un 2%)

## Discusión

### 1. Control genético de los caracteres de calidad en la genoteca de ILs

El efecto genético fue muy alto para los caracteres relacionados con la morfología del fruto, mientras que para el resto, en general, se observa un menor efecto del genotipo, principalmente en los caracteres relacionados con el sabor y de color. Concomitantemente, las interacciones G x L más bajas las presentaron los caracteres relacionados con las dimensiones del fruto, seguidos por los de forma del fruto y caracteres asociados con el sabor; estos resultados concuerdan con Chen et al. (1999) que menciona como poco comunes las interacciones de los caracteres FW y SSC con el ambiente, siendo menores las interacciones en FW.

Las correlaciones altas y positivas entre FW y los demás caracteres de dimensión de fruto (PER, AR, FDL y FLL), eran esperables, ya que estos caracteres están relacionados directa o indirectamente con FW. Resultados similares han sido descritos en poblaciones desarrolladas a partir de *S. pimpinellifolium* (Grandillo y Tanksley 1996; Chen et al. 1999; Lippman y Tanksley 2001; Saliba-Colombani et al. 2001; Doganlar et al. 2002). La correlación negativa observada entre FW y FS en el presente estudio, también descrita previamente por (Lippman y Tanksley 2001), puede explicarse debido a que ambos caracteres están influenciados por el mismo componente (FDL), esto es, frutos con mayor diámetro son más achatados (menor FS), pero más pesados (FW). Es decir, la variación en FS observada en esta población no se debe a variación en el crecimiento longitudinal que se observa en los mutantes *OVATE* (Ku et al. 2000) o *SUN* (van der Knaap y Tanksley 2001), sino a variación en el eje ecuatorial.

Interesantemente, no se observaron correlaciones importantes entre FW y SSC, al contrario con lo observado más comúnmente en trabajos previos (Tanksley et al. 1996; Bernacchi et al. 1998; Chen et al. 1999; Doganlar et al. 2002,) donde se han descrito correlaciones negativas entre FW y SSC, que generalmente se interpretan como efectos pleiotrópicos de la disminución de FW en el aumento de SSC. Los resultados del presente trabajo indican que debería ser posible utilizar los QTLs de SSC encontrados en la presente colección para aumentar SSC sin efectos negativos sobre el tamaño del fruto, al menos usando como fondo genético MM.

Las correlaciones altas observadas entre DBLK, PBLK y TRI son obvias debido a que las mediciones de los frutos que se usan para calcular estos índices son muy similares, Asimismo

la alta correlación negativa entre FS y CIR indican que la forma del fruto está más influenciada por el diámetro y no por la longitud del fruto (confirmando lo mencionado anteriormente sobre la influencia del diámetro en el tamaño del fruto de algunas ILs).

Los niveles de SSC se correlacionaron negativamente con los parámetros del EC, lo que explicaría que a medida que los frutos alcanzan niveles de color externo mayores (fruto sobremaduro), los SSC tienden a disminuir. Esto dependerá de la etapa de maduración en que se realicen las determinaciones de SSC y de la forma como se mida el color. Por ejemplo Grandillo y Tanksley (1996), midiendo el color en forma visual, encontraron una correlación positiva entre SSC y color de fruto. Las correlaciones entre EC y IC fueron predominantemente positivas, resultados similares fueron encontrados por Doganlar et al. (2002) en una población de BILs hechas a partir de *S. pimpinellifolium*.

## **2. Disección genética de caracteres de calidad con la genoteca de ILs de TO-937**

En los últimos 25 años se han descrito en tomate una enorme cantidad de QTLs implicados en distintos aspectos de la calidad de fruto de tomate (Labate et al. 2005; Foolad et al. 2007; Grandillo et al. 2011; Grandillo et al. 2013; Monforte et al. 2014). En prácticamente todos los trabajos se han encontrado alelos de especies silvestres que tenían efectos contrarios a lo esperado atendiendo al fenotipo silvestre donante, como aumento de tamaño (Tanksley et al. 1996; Causse et al. 2004), color del fruto (Monforte et al. 2001), productividad (Gur y Zamir 2004), entre otros. En el presente trabajo también hemos encontrado QTLs con efectos contrarios al esperado en varios caracteres, como FW, FS, PSH, CIR, SSC, EC e IC.

Pero en general, se han encontrado alelos con efectos en las dos direcciones, mostrando la complejidad genética implicada en la variabilidad de estos caracteres (**Anexo 4-4**). Por ejemplo, en los QTL relacionados con las dimensiones de los frutos como *fw1.1*, *fw1.2*, *fw2.1*, *fw2.2*, *fw3.1*, *fw3.2* *fw4.1*, entre otros, los alelos de TO-937 disminuyeron los valores con respecto a MM, lo cual es esperable dadas las grandes diferencias en tamaño y peso entre ambos genotipos. Por el contrario en QTLs como SSC e IC, los alelos de TO-937 muestran efectos positivos o favorables, coincidiendo con los valores más altos en estos caracteres de TO-937, comparado con MM. Sin embargo, en estos mismos caracteres algunos los efectos observados son distintos a los esperados, como por ejemplo el QTL *fw12.1* en donde los alelos de *S* TO-937 aumentan FW en cerca del 20% con respecto al control. De forma similar los QTLs *cir6.1* y *cir10.1* cuyos efectos son responsables en más de un 40% de la forma no

circular de los frutos o los QTLs *ssc4.1* y *ssc5.1* que disminuyen SSC en 13 y 10% respectivamente.

Para el carácter peso de fruto, se definieron once QTLs en los cromosomas 1, 2, 3, 4, 7, 10, 11 y 12 (también fueron detectados QTLs en los cromosomas 5, 6, 8 y 9 pero sólo en una única localidad). Estos resultados coinciden con otros estudios realizados con poblaciones desarrolladas a partir de *S. pimpinellifolium*, en donde se han descrito QTLs para FW en los todos los cromosomas en una o más localidades (Grandillo y Tanksley 1996; Tanksley et al. 1996; Bernacchi et al. 1998; Chen et al. 1999; Lippman y Tanksley 2001; Doganlar et al. 2002; van der Knaap et al. 2002). Aunque como analizaremos más adelante algunos de estos QTLs no fueron definidos en las mismas regiones genómicas.

El control genético del tamaño y morfología del fruto es probablemente uno de los aspectos más conocidos del fruto del tomate (Tanksley 2004, Monforte et al. 2014). Entre los QTL de FW definidos en el presente trabajo, *fw2.2* y *fw3.2* coinciden en su posición con los QTLs ya caracterizados molecularmente: *fw2.2* (Frary et al. 2000) y *fw3.2* (Zhang et al. 2012, Chakrabarti et al. 2013), respectivamente, que corresponden a dos QTLs claves en la domesticación y diversificación de tomate. En general, *fw2.2* se ha definido como el QTL de mayor efecto, sin embargo, en nuestro caso QTLs como por ejemplo *fw1.1*, *fw2.1*, *fw3.1*, *fw3.2*, *fw10.1* presentan un efecto similar y/o mayor que el posible ortólogo de *fw2.2*, lo que apoyaría la hipótesis de que la domesticación fue un proceso más complejo e implicando a un mayor número de genes de lo que se podía deducir tras el clonaje de *fw2.2* (Rodríguez et al. 2011).

En nuestro estudio, un QTL en el extremo distal del cromosoma 12 (*fw12.1*) aumentó el peso del fruto en un 20%. El único estudio previo realizado en poblaciones desarrolladas a partir de *S. pimpinellifolium*, en el que se define un QTL de peso de fruto en el cromosoma 12 lo realizó Chen et al. (1999), sin embargo, mostró un efecto contrario en FW, y además, al comparar las posiciones cromosómicas de los marcadores que flanquean el QTL en los respectivos mapas (Tomato-EXPEN 1992 y Tomato-Kazusa y SolCAP) la localización del los QTLs es distinta. Lo mismo podemos decir del QTL descrito por Saliba-Colombani et al. (2001) en una poblaciones desarrolladas a partir de una accesión de *S. lycopersicum* var. *cerasiforme*. Por otra parte, Causse et al. (2004) definió el QTL *fw12.1* en la IL12\_4-1 de *S. pennellii* que aumenta el peso del fruto en 19.6% con respecto al control (M-82). Otro posible homólogo lo describe Fulton et al. (2000) en una población desarrollada a partir de la accesión LA2133 de *S. parviflorum*, sin embargo, en ambos casos los efectos del QTL no

fueron estables en distintos experimentos. En el presente trabajo *fw12.1* ha mostrado una gran estabilidad, reforzada al encontrar otros QTLs relacionados con dimensiones del fruto (FW, PER, AR, FDL y FLL) mapeados en la misma región cromosómica en dos ILs que se solapan (SP\_12-4 y SP\_12-5).

Otros QTLs para FW poco comunes en otros estudios en poblaciones desarrolladas a partir de *S. pimpinellifolium* han sido definidos en el presente: *fw1.1*, *fw2.1*, *fw4.1* *fw10.1*. Sin embargo, todos estos QTLs ya fueron descritos anteriormente en otras poblaciones desarrolladas a partir de cruzamientos con otras especies silvestres por ejemplo: *fw1.1*, *fw2.1*, *fw4.1* *fw10.1* probablemente sean alélicos a los definidos por Eshed y Zamir (1994) en ILs desarrolladas a partir de *S. pennellii* acc. LA716. Por su parte, Monforte et al. (2001) identificó un QTL para FW en la misma región que *fw4.1* en líneas de introgresión a partir de accesiones de *S. pennellii*, *S. arcanum* y *S. habrochaites*.

Los QTLs que definidos en este trabajo para PER y AR pueden adscribirse como componentes de FW, lo que queda de manifiesto al observar la alta correlación entre estos caracteres y FW y la posición idéntica en el genoma, dada la correspondencia entre la variación del peso y el tamaño de los frutos con las dimensiones longitud, diámetro y profundidad.

En cuanto al FDL, en la literatura se menciona la presencia de QTLs en todos los cromosomas con excepción de los cromosomas 5 y 10 cuando se utilizan poblaciones desarrolladas a partir de *S. pimpinellifolium* (Foolad 2007). En el presente estudio se detectaron QTLs en todos los cromosomas con excepción del cromosoma seis, de estos *fdl1.1*, *fdl2.1*, *fdl2.2*, *fdl3.1*, *fdl3.2*, *fdl4.1*, *fdl10.1* y *fdl11.1* se detectaron en más de una localidad. En cuanto a la FLL, estudios previos mencionan la presencia de QTLs en los cromosomas 1, 2, 3, 4, 6, 7, 8, 9, 11 y 12. En el presente estudio se detectaron QTLs para estos caracteres en la mayoría de los cromosomas mencionados, además se detectó un nuevo QTL no descrito previamente en cromosoma 10 que disminuye la altura del fruto en un 9% con respecto MM. Los QTLs para los que no hemos encontrado ortólogos claros en la literatura en poblaciones desarrolladas a partir de *S. pimpinellifolium* fueron *fl4.1*, *fl7.1*, *fl10.1* (altura máxima) y *fdl4.1*, *fdl10.1* (diámetro máximo). Sin embargo, en estas regiones se han encontrado efectos en el carácter en la población de ILs desarrolladas a partir del silvestre *S. pennellii* acc. LA716 (Eshed y Zamir 1994).

En cuanto a FS, se han descrito más de 20 QTLs en estudios realizados en poblaciones desarrolladas a partir de *S. pimpinellifolium* y otros silvestres como donantes de genes. (Bernacchi et al. 1998; Tanksley et al. 1996; Doganlar et al. 2002; Chen et al. 1999; Grandillo y Tanksley 1996; van der Knaap et al. 2002, Barrero y Tanksley 2004, Rodríguez et al. 2011, Monforte et al. 2014). En el presente estudio se definen ocho QTLs en los cromosomas 1, 2, 3, 4, 6 y 10 con efectos sobre la forma del fruto. El QTL *cir2.1*, responsable de un 26% de la variación en la forma redonda de los frutos, ha sido detectado en estudios anteriores con poblaciones desarrolladas a partir de *S. pimpinellifolium* (van der Knaap y Tanksley 2003; Bernacchi et al. 1998; Grandillo y Tanksley 1996; Tanksley et al. 1996). La posición física de *cir2.1* coincide con la posición del gen *OVATE*, y probablemente se deba a un alelo de este gen (Liu et al. 2002; Rodríguez et al. 2011). Por su parte los QTLs *cir6.1* y *cir10.1* se asocian a frutos menos redondos (> 40%) que MM, que coinciden con los QTLs *fs6.1* y *fs10.1* que inducen frutos entre 6 y 7% mas achatados que MM. Un QTL similar en el cromosomas 6 fue descrito por van der Knaap y Tanksley (2003) en una población F<sub>2</sub> derivada de la accesión LA1589 de *S. pimpinellifolium*, que probablemente sea ortólogo con el encontrado en este estudio. Por su parte, Bernacchi et al. (1998) y Chen et al. (1999) han descrito QTLs probablemente ortólogos para *fs10.1* en poblaciones desarrolladas a partir de *S. pimpinellifolium*. Uno de los QTLs identificados como diferenciadores en la forma del fruto entre las especies silvestres y el tomate cultivado es *fs8.1* (Ku et al. 2000), que se encuentra en la región distal del cromosoma 8. Este QTL no ha sido detectado en la colección de ILs de TO-937, una inspección más detallada del genotipo de las ILs del cromosoma 8 nos mostró que la región que debería incluir este QTL no está presente en la genoteca debido a no contar con marcadores para esa región durante proceso de cribado de la IL SP\_8-1 (Fig. 4-3).

Con respecto a los QTLs que midieron la altura de los hombros en frutos, no hay estudios donde se describan ortólogos usando esta nomenclatura. Van der Knaap et al. (2002) define a *fs3.1* en una población derivada de la accesión LA1589 de *S. pimpinellifolium* que mapea en la misma posición que *psh3.1*, lo que podría indicar que ambos QTLs son ortólogos.

Con respecto a QTLs de SSC, estos se han detectado en estudios anteriores en todos los cromosomas, siendo más frecuente su detección en los cromosomas 2, 3, 6 y 9. En el presente estudio se detectaron QTLs significativos para SSC en los cromosomas 2, 3, 4, 5 y 9 en más de una localidad. La gran mayoría de QTLs detectados en plantas con introgresiones de fragmentos de accesiones de *S. pimpinellifolium* aumentan los valores del carácter. Como es el caso del *ssc9.1* que aumenta un 13% SSC, que posiblemente sea ortólogo *ssc9.1* descrito y

aislado por Fridman et al. (2000) y que incrementa en más del 20% en contenido de azúcares. Dos QTLs interesantes son *ssc2.1* y *ssc3.1* que aumentan en 14 y 15% los grados brix respectivamente, sin que muestren efectos pleiotrópicos negativos sobre FW, por lo que serían buenos candidatos para su utilización en mejora genética sin afectar el tamaño de los frutos. Por otra parte, el QTL *ssc4.1* detectado en el presente estudio, tuvo un efecto negativo sobre el contenido de azúcares disminuyéndolo el valor de SSC en un 13% aproximadamente. Un QTLs con el mismo efecto fue descrito previamente por Tanksley et al. (1996) sin embargo, no deben de ser ortólogos, ya que al comparar su ubicación en los mapas respectivos, se ubican en distintas posiciones (**Tabla 4-5**).

En el presente estudio se lograron detectar QTLs para contenido de acidez en los cromosomas 1, 2, 3, 4, 5, 10 y 11 la mayoría aumentando los niveles con respecto a MM, aunque sus efectos no fueron consistentes entre localidades, estos resultados concuerdan con los encontrados por Causse et al. (2004) quienes describen QTLs para acidez en prácticamente todos los cromosomas en líneas de introgresión desarrolladas a partir de *S. pennellii*. Por otra parte, el único estudio sobre detección de QTL para acidez en una población desarrolladas a partir de *S. pimpinellifolium* lo realizó Fulton et al. (2002), logrando detectar 3 QTLs que se localizan en los cromosomas 1, 3 y 8. Sin embargo, en nuestro caso, solamente el QTL *ta2.1* fue consistente en más de una localidad, disminuyendo el contenido de ácido cítrico en un 17% con respecto al control. El QTL *ta2.1* mapea en la misma posición de un posible ortólogo descrito por Saliba-Colombani et al. (2001) en una población desarrolladas a partir de *S. lycopersicum* var *cerasiforme*.

En relación a EC, los dos QTLs encontrados *ec2.1* y *ec7.1*, disminuyen el color rojo con respecto al control en 10 y 20% respectivamente. Bernacchi et al. (1998) y Grandillo y Tanksley (1996) describen un posible ortólogo para el QTL *ec2.1* en una población desarrolladas a partir de la accesión LA1589 de *S. pimpinellifolium*. Adicionalmente el *ec2.1* podría estar relacionado con el locus *hp-1* (mapean en la misma posición), incrementa los niveles de licopeno y Beta-caroteno durante el desarrollo del fruto (Yen et al. 1997). Se definió el QTL *ic1.1* (componente H\*) responsable de que los frutos sean menos rojos con respecto a MM. Bernacchi et al. (1998) y Grandillo y Tanksley (1996), describen un posible ortólogo de *ic1.1* en una población desarrolladas a partir de *S. pimpinellifolium*. La posición física de *ic1.1* es homóloga al locus *hp-2* involucrado en el desarrollo de plásmidos y biosíntesis de carotenoides (van Tuinen et al. 1997). Los frutos que contenían *ic2.1* o *ic5.1*, mostraron valores, para el componente L\*, menores y mayores respectivamente con respecto

al control. De acuerdo a lo mencionado por Sack y Francis (2001), valores bajos de  $L^*$  se asocian a frutos rojos (*ic5.1*) y valores altos de  $L^*$  se asocian a frutos de color amarillo o naranja (*ic2.1*). Por su parte el QTL *ic10.1* se asocia a la variable  $a^*$  define frutos más rojos con respecto al control. No se encontraron ortólogos en estudios anteriores para estos tres últimos QTLs.

Resumiendo, a pesar de la existencia de gran cantidad de accesiones que recogen mucha de la variabilidad genética de *S. pimpinellifolium*, la gran mayoría de trabajos de mapeo de QTLs realizados en esta especie se han hecho en una sola accesión (LA1589) usando diferentes fondos genéticos de tomate comercial (industria y consumo fresco), repitiéndose muchas veces los mismos resultados (Tanksley et al. 1996; Grandillo y Tanksley 1996; Bernacchi et al. 1998; Lippman y Tanksley 2001; Doganlar et al. 2002; van der Knaap et al. 2002; van der Knaap y Tanksley 2003).

En el presente trabajo definimos cerca de 15 QTLs que no han sido descritos antes en poblaciones desarrolladas a partir de *S. pimpinellifolium* (sin considerar los QTLs de perímetro y área no descritos antes en Solanáceas), dentro de estos nuevos QTLs, siete están relacionados con las dimensiones del fruto, dos de FS, uno de SSC, y cuatro de color externo e interno del fruto. Estos nuevos QTLs pueden representar variabilidad alélica en *S. pimpinellifolium* que no ha sido rastreada previamente en poblaciones derivadas de esta especie, aunque no podemos descartar que la detección diferencial pueda ser debida a variaciones en el fondo genético y/o condiciones ambientales entre los distintos experimentos. El número de QTL consistentes que fueron definidos en este experimento son algo menor comparado con los definidos, par los mismos caracteres, con la genoteca de *S. pennellii*, probablemente como consecuencia una menor variación genética en esta población debido a de la mayor proximidad filogenética de TO-937 con MM. Sin embargo, esta proximidad también se puede traducir en un beneficio, debido a que los fenotipos de las ILs tienen menor el lastre genético, lo que facilitaría futuras incorporaciones de estos QTLs a líneas elites en eventuales programas de mejora (Tanksley y Nelson 1996).

## CAPITULO 5: Discusión General

La reducida variabilidad que muestran las variedades comerciales de tomate en relación a sus parientes silvestres (Miller y Tanksley 1990), tiene entre otras causas el tener su sitio de domesticación y diversificación muy lejos de su centro de origen. A partir de los primeros trabajos en pre-mejora (Rick 1986), se supo que las especies silvestres pueden ser una fuente de alelos de interés agronómico (incluyendo alelos con efectos favorables dentro de genotipos silvestres que fenotípicamente son inferiores agronómicamente), que una vez localizados en el genoma por medio del mapeo de QTLs son asociados a marcadores moleculares y pueden ser introducidos en los cultivares modernos mediante Selección Asistida por Marcadores (Collard et al. 2005). Como resultado, los mejoradores pueden hacer selección basándose exclusivamente en el fenotipo, en la composición alélica de marcadores moleculares o combinar ambas estrategias. Esta circunstancia ha fomentado el desarrollo de poblaciones a partir de cruces interespecíficos como herramienta para conocer la diversidad genética de materiales silvestres, paso previo a ser usados en programas de mejora vegetal. Sin embargo, el potencial de las especies silvestres como fuente de alelos valiosos para la mejora genética del tomate todavía está muy lejos de ser totalmente explotado (Grandillo et al. 2011).

El objetivo último de esta tesis es acercar al mejorador la variabilidad genética presente en TO-937 para que pueda ser incluida en planes de mejora y también para el estudio básico de caracteres de interés, relacionados con la calidad del fruto en una primera aproximación, pero que pueden extenderse a muchos más en el futuro. Para lograr este objetivo se desarrolló una genoteca de líneas de introgresión utilizando esta accesión como parental donante. Si bien este tipo de poblaciones ya se han desarrollado con anterioridad en tomate a partir de otras especies silvestres del género *Solanum* (Eshed y Zamir 1994; Monforte y Tanksley 2000a; Chetelat y Meglic 2000; Francis et al. 2001; Canady et al. 2005; Finkers et al. 2007) y, en otros cultivos como avena (von Korff et al. 2004), melón (Eduardo et al. 2005), trigo (Liu et al. 2006), entre otros, el número de genotecas disponibles es muy reducido (considerando que hace 20 años Eshed y Zamir desarrollaron la primera), dado el gran esfuerzo que significa su desarrollo en términos de tiempo y recursos. Mediante la estrategia metodológica desarrollada en esta tesis pretendemos revertir esta tendencia. Dentro de los principales aportes que hace esta tesis podemos destacar una significativa mejora de la eficiencia en el desarrollo de la genoteca (tiempo requerido para obtener las ILs). Sin lugar a duda el genotipado masivo realizado a la totalidad de plantas que formaron las poblaciones BC<sub>2</sub> y BC<sub>3</sub> contribuyó a

optimizar la selección de plantas con las características idóneas, lo que repercutió que en generaciones tempranas ( $BC_3S_1$ ) iniciáramos la extracción de las primeras ILs definitivas. Con esta estrategia metodológica desarrollada, se estima que hasta un tercio de la genoteca se podría obtener en la quinta generación. Sin embargo, debido a problemas surgidos durante el cultivo en el invernadero, no pudieron obtenerse semillas de todas las plantas seleccionadas en las generaciones intermedias, por lo que únicamente un 8% de las ILs pudieron obtenerse en la quinta generación y, el resto se obtuvo entre la sexta y séptima generación. A pesar de estos problemas, el número de generaciones es sensiblemente inferior (entre 2 y 3 generaciones) comparado con la construcción de las últimas genotecas publicadas en *Solanum* (Francis et al. 2001; Canady et al. 2005; Finkers et al. 2007).

Una segunda consecuencia positiva de la estrategia empleada es la calidad de la genoteca de líneas de introgresión, que viene definida por la resolución genética y la isogenicidad del fondo genético. Las tecnologías recientes de genotipado masivo de alta densidad con varios miles de marcadores, o re-secuenciación de genomas permiten determinar la calidad de las genotecas con una resolución impensable hasta hace muy pocos años. La aplicación de estas tecnologías en genotecas construidas anteriormente usando mapas genéticos de media densidad ha mostrado que en general la isogenicidad del fondo genético es menor de la esperada, con más introgresiones “no-diana” que las que originalmente se habían detectado con mapas menos saturados (Xu et al. 2010; Eichten et al. 2011; Schmalenbach et al. 2011; Pea et al. 2013). En nuestro caso el genotipado final con la micromatriz de 8 K SNPs SOLCAP también nos permitió detectar introgresiones no-diana en el 51% de las ILs, aunque en cantidades y tamaño muy inferiores a las encontradas en los estudios antes mencionados. En cuanto a introgresiones no-diana, nuestra genoteca contiene 0.7/IL, es ligeramente menor a las 1.5 detectadas por Schmalenbach et al. (2011), pero mucho menor a las más de 8 y entre 10-19 detectadas por Eichten et al. (2011) y Pea et al. (2013) respectivamente. En cuanto al tamaño de estas introgresiones, en nuestra genoteca el 90% de las introgresiones no-dianas tenían un tamaño promedio de 2 Mb localizadas en regiones eucromáticas donde las frecuencias de recombinación son mayores y el restante 10% de las introgresiones no-dianas tenían en promedio 54 Mb de tamaño y se localizaron en las regiones centroméricas, regiones con pocos eventos de recombinación (The Tomato Genome Consortium 2012).

La presencia de estas introgresiones no-diana obedece a eventos de doble recombinación entre los marcadores utilizados para caracterizar las generaciones intermedias durante el desarrollo de las líneas de introgresión. Así, en un mapa de densidad media (por ejemplo con intervalos

de 20 cM entre marcadores), aunque la probabilidad de producirse una doble recombinación entre dos marcadores es baja (4%), la probabilidad de esto ocurra en al menos uno de los intervalos del genoma es, sin duda, no despreciable. Una vez producida una doble recombinación, esta pasa desapercibida si no se aumenta la densidad de marcadores. En nuestro trabajo, el reducido número de introgresiones no-dianas encontradas después del genotipado masivo, en relación a otras genotecas, obedeció a varias razones: a una alta densidad en la micromatriz de genotipado utilizada en la población BC<sub>2</sub> que permitió seleccionar plantas con secciones grandes de cromosoma del parental donante, disminuyendo la probabilidad del arrastre de pequeñas introgresiones consecuencia de dobles recombinaciones: La caracterización genómica de la generación BC<sub>3</sub> se realizó basándonos en mapas genéticos de mediana resolución, lo que explicaría que se produjera algún evento de doble recombinación que no se detectó hasta el genotipado masivo final. Por otra parte, gracias a la alta presión de selección que se pudo realizar durante el desarrollo de la genoteca, el número de generaciones fue relativamente bajas, lo que también permite que la frecuencia de doble recombinaciones sea más baja que en otros trabajos y por tanto, que la calidad de la presente genoteca de ILs sea sensiblemente mejor con respecto a la isogenicidad del fondo genético comparado con estrategias anteriores (Tian et al. 2006; Keurentjes et al. 2007; Finkers et al. 2007)

Con respecto a la resolución genética de la colección de ILs, el genotipado masivo de alta densidad realizado al final del proceso, permitió definir un total de 71 bins (tamaño promedio de 21 cM y rango entre 1 y 79 cM), 13 más respecto a los bins definidos con el genotipado con HRM basado en un mapa de densidad media, consecuencia de una mayor precisión en la definición de los puntos de recombinación. El aumento en el número de bins es muy importante de cara a la precisión en el mapeo de QTLs objetivo final de la genoteca. El número, tamaño promedio y rango de valores de los bins obtenidos en nuestra genoteca son comparables con los definidos en la genotecas desarrollada por Canady et al. (2005) (68 bins, tamaño promedio de 18 cM y rango máximo de 64 cM) usando *S. lycopersicoides* como parental donante.

Como era de esperarse, durante el desarrollo de la genoteca no se observaron en general problemas de recombinación o compatibilidad entre los genomas, dada la cercanía filogenética que existe entre los genomas fundadores. Sin embargo, las introgresiones representadas en la genoteca cubren el 94% de genoma donante, a causa de la pérdida de dos fragmentos de 47 Kb y 3 Mb en el extremo proximal de los cromosomas 1 y 8

respectivamente. La inspección detallada de estas regiones demostró que se debió a la ausencia de marcadores en esas regiones en los diseños de SNPs utilizados en las generaciones tempranas, por lo que se pudieron producir eventos de recombinación en las etapas iniciales que no fueron detectados y se eliminaron dichos fragmentos de TO-937. Una forma de evitar la pérdida de estas regiones cromosómicas y evitar el lastre de grandes regiones centroméricas es considerar tanto las distancias genéticas como físicas a la hora de elegir los marcadores que se utilizarían para el programa de selección asistida por marcadores. La distribución de los genes y las recombinaciones a lo largo de los cromosomas no es homogénea, la mayoría de los genes se encuentran en las regiones de eucromatina, que es donde se producen la mayor parte de los eventos de recombinación, mientras que en las regiones de heterocromatina, que incluyen centrómeros y telómeros, la densidad de genes es muy baja y la recombinación está suprimida, además los centrómeros pueden representar regiones muy grandes, abarcando gran parte del cromosoma (The Tomato Genome Consortium 2012). Así, mediante la selección de marcadores basándose en distancias genéticas se saturan las regiones de eucromatina y se asegura que el número de dobles recombinantes sea el menor posible. Al no incluirse prácticamente marcadores en las zonas de heterocromatina, si se produjeran dobles recombinantes se podrían arrastrar grandes regiones heterocromáticas no deseadas. El incluir unos pocos marcadores en estas regiones, evitaría este lastre en la construcción de ILs. Por otro lado, la selección de marcadores basado exclusivamente en distancias físicas no sería adecuado, ya que la mayoría de los marcadores caerían en regiones heterocromáticas sin apenas recombinación, mientras que la densidad de marcadores en las regiones eucromáticas no sería suficiente para un control adecuado de los posibles doble recombinantes.

Resumiendo, en el capítulo 2 hemos puesto en práctica una variante al procedimiento utilizado anteriormente para el desarrollo de líneas de introgresión, combinando cruces controlados con la selección basada en la utilización de herramientas genómicas de alto rendimiento. Lo que se ha traducido en mejoras sustanciales si la comparamos con otras genotecas desarrolladas anteriormente. Con esta estrategia se mejoró la eficiencia reduciéndose el número de generaciones necesarias para obtener las ILs, la genoteca posee un alta isogenicidad reflejada en un bajo número de introgresiones no-dianas y en una cobertura cercana al 100% del parental donante, y con una relación coste/beneficio adecuada promovida en primer lugar por el menor número de plantas a cribar (gracias a un mejor selección de genotipos en la BC<sub>2</sub> y BC<sub>3</sub>), y a que en los últimos años el costo del genotipado masivo ha disminuido considerablemente haciéndolos más asequibles.

Otro aspecto interesante que nos encontramos durante el desarrollo de las ILs fue la presencia de una región que mostró segregación distorsionada (SD) en todas las poblaciones (RILs, BC<sub>2</sub>, BC<sub>3</sub>, ILs) analizadas resultado del cruzamiento entre TO-937 y MM. A pesar que el análisis de la SD no fue un objetivo trazado inicialmente en esta tesis, decidimos ahondar más en el fenómeno incorporándolo como parte de la investigación.

El locus responsable de SD se localizó en el extremo distal del brazo largo del cromosoma 4, mostrando unas características muy peculiares: la SD se observa en la descendencia de plantas heterocigóticas para la región, que en caso de autofecundación se observa una segregación 0:1:3 (*l: h: p*), lo que descarta que se deba a mecanismos de autoincompatibilidad, esterilidad de polen, o incompatibilidad unilateral y que afecta tanto gametos femeninos como masculinos. La segregación distorsionada se observa en las poblaciones de retrocruzamiento donde el donante de polen fue MM, confirmando que esta distorsión afecta al gameto femenino. Esta segregación indica que la eliminación de los gametos de MM no es del 100%. En conjunto, estos resultados son compatibles con las segregaciones distorsionadas observadas por Rick (1966 1971), donde definen un gen implicado en la distorsión de segregación que denominan *Gamete Eliminator (Ge)*, que localizan en el cromosoma 4. La localización de este gen en los trabajos de Rick (1966, 1971) está en la región centromérica del cromosoma 4, mientras que según nuestros resultados se encontraría en la región distal de este cromosoma. Estos resultados han sido verificados en otras poblaciones derivadas del cruzamiento entre otras accesiones de *S. pimpinellifolium* y variedades de tomate cultivado (D. Francis comunicación personal). Esta diferencia en la posición cromosómica del locus causante de SD puede deberse a las diferencias de resolución genéticas de los mapas desarrollados a finales de los 60s con los mapas a los que tenemos acceso actualmente.

Mediante mapeo de alta resolución, se consiguió localizar el locus responsable de SD dentro de una región cromosómica de 84 Kb. Dentro de los genes anotados en la región está ***Solyc04g081590*** (ortólogo de *Clavata (CLV)*) que se perfila como un posible candidato. Miembros de la familia CLV regulan la proliferación celular durante la formación del fruto de la misma forma que ocurre con los meristemos (Durbak y Tax 2011). En el ovario estos meristemos son los encargados de generar tejido placentario y los óvulos (Ferrandiz et al. 1999). De acuerdo a estas evidencias es posible plantearnos como hipótesis que alteraciones en proteínas CLV podría afectar el desarrollo del óvulo y del polen. Siendo posible que CLV1/TO-937 proporcione una ventaja adaptativa frente a la de CLV1/MM en el óvulo/polen en el entorno del esporofito heterocigoto. Otro de los posibles genes candidatos es

*Solyc04g081640* anotado como endoplasmina, cuyo ortólogo en *Arabidopsis* es Shepherd. (*SHD*) es necesario para que la identidad del meristemo, además tiene una función en el desarrollo de los gametofitos (Ishiguro et al. 2002), siendo requerido para la elongación del tubo polínico o/y la penetración del estilo. Igualmente la hipótesis podría ser que el alelo de TO-937 de *SHD* presente un patrón de expresión que le proporcione una ventaja adaptativa a los gametos que lo lleven. A su vez, la ciclina-B está implicada en procesos de división celular. La hipótesis que se podría plantear es que una alteración en los niveles o perfil de expresión de la ciclina-B podría causar desajustes en la división celular de los gametofitos, confiriendo a los gametofitos TO-937 alguna ventaja sobre los de MM en un contexto heterocigoto. Finalmente, la presencia de un retrotransposon en la región podría estar afectando la expresión o integridad de algunas de las secuencias que lo flanquean, como ha sido descrito en otros casos en tomate y especies relacionadas (Xiao et al. 2008; Lacombe et al. 2010; Butelli et al. 2012). Sin lugar a duda hará falta realizar ensayos con múltiples abordajes para tratar de aceptar o rechazar cada una de estas hipótesis y lograr esclarecer la participación o no de cada uno de los genes y elementos cromosómicos presentes en la zona donde se localiza la SD.

Consideramos que el avance realizado con el mapeo fino ha sido considerable, pero a sabiendas que la información generada nos es concluyente estamos proponiendo un abordaje integral del problema, realizando estudios que contemplen: 1) generación de más recombinantes en esa región, idealmente hasta que la resolución sea de un único gen, 2) estudiar la estructura de la región a nivel genómico, generando librerías en BACs para realizar un ensamblaje *de novo* 3) analizando con más detalle posibles alteraciones en el desarrollo de los gametofitos masculinos / femeninos en el entorno del esporofito heterocigoto para la región y en cruces con genotipos que contengan los tres posibles alelos que menciona Rick (1966, 1971) ( $Ge^p$ ,  $Ge^c$  y  $Ge^n$ ) 4) haciendo estudios de expresión en las estructuras reproductivas en plantas que presenten las diferentes combinaciones alélicas en la región con SD. Esperamos que estos estudios nos permitan definir a corto o mediano plazo cuál (s) es el mecanismo responsable del fenómeno de SD en la región diana estudiada.

El objetivo último del desarrollo de una población de mapeo de QTLs es su utilización en la exploración y acercamiento al germoplasma cultivado de diversidad genética de especies silvestres. En el caso del tomate, la relación entre la gran cantidad de diversidad genética almacenada en los bancos de germoplasma y las pocas poblaciones de mapeo de QTLs desarrolladas a partir de parientes silvestres es muy desigual. En cuanto a la especie *S.*

*pimpinellifolium* a pesar de su gran variabilidad genética (Rick 1971, 1977), la gran mayoría de trabajos de mapeo de QTLs publicados, se han realizado con la accesión LA1589, (Tanksley et al. 1996; Grandillo y Tanksley 1996; Bernacchi et al. 1998; Lippman y Tanksley 2001; Doganlar et al. 2002; van der Knaap et al. 2002; van der Knaap y Tanksley 2003). Si hablamos de exploración de variabilidad genética de tomate usando genotecas completas de ILs, la gran mayoría de trabajos se han realizado usando una única genoteca: la desarrollada por Eshed y Zamir (1994). Las razones, ya mencionadas anteriormente, versan sobre los largos períodos de tiempo y altos costos que se requiere para su desarrollo. Esperamos que la estrategia desarrollada en esta tesis ayude a cambiar el panorama y, en los próximos años la construcción de nuevas genotecas de ILs repunte de manera significativa, al menos con especies que no presenten problemas serios de compatibilidad.

La genoteca desarrollada en esta tesis nos permitió definir un total de 75 QTLs entre caracteres descriptivos y cuantitativos con efectos consistentes en al menos dos localidades. La gran mayoría de la dirección de los efectos de los QTLs estuvo en concordancia con lo que se esperaba según el fenotipo mostrado por el parental donante. Sin embargo, al igual que en una buena parte de los trabajos de mapeo de QTLs publicados (Tanksley et al. 1996, Monforte et al. 2001; Causse et al. 2004, Gur y Zamir 2004, entre otros), se encontraron alelos en TO-937 que presentaron segregación transgresiva en caracteres de peso del fruto, forma, grados brix y color. Por ejemplo, *fw12.1* aumento el peso de fruto en un 20% con respecto a MM, *fs6.1* y *fs10.1* son responsables de que los frutos sean más achatados en promedio un 6.5% respecto a MM y *ssc4.1* y *ssc5.1* responsables de frutos con niveles de brix de 13 y 10% menos que MM.

Por otro lado, el numero de QTLs consistentes definidos en este experimento no son tan numerosos con los que se han definido, para los mismos caracteres, con la genoteca de *S. pennellii*, muy probablemente como consecuencia de una menor diversidad genética entre TO-937 y MM que entre *S. pennellii* LA716 y el cultivar M82. Sin embargo, esta proximidad también se puede traducir en un beneficio, debido a que los fenotipos de las ILs desarrolladas a partir de TO-937 tienen menor el lastre genético, lo que facilitaría futuras incorporaciones de estos QTLs a líneas elites en eventuales programas de mejora (Tanksley y Nelson 1996).

Dentro de los QTLs más interesantes por su efecto transgresor fue *fw12.1* que aumenta el peso promedio del fruto en un 20% con respecto a MM. Aunque anteriormente Causse et al. (2004) había definido este QTL en la IL12\_4-1 de *S. pennellii* aumentando el peso del fruto en 19.6% con respecto a M-82, había mostrado poca consistencia. En nuestro caso *fw12.1* si

mostró consistencia, que se vió reforzada con la detección de QTLs para perímetro, área, diámetro y longitud de fruto en ILs con introgresiones de TO-937 en esta misma región donde mapea *fw12.1*, haciendo evidente la presencia de genes en esta región de *S. pimpinellifolium* que aumentan las dimensiones de los frutos. Habría que investigar si la presencia de estos QTLs transgresivos podría estar asociado al hábito de crecimiento indeterminado de MM, que implica una mayor área foliar y consecuentemente mayor disponibilidad de recursos, en comparación a otros fondos genéticos que tienen crecimiento determinado.

Dentro de los datos de correlaciones obtenidas, resulta muy interesante que no se observaran correlaciones entre FW y SSC (al menos en dos localidades), todo lo contrario a lo publicado en trabajos previos (Tanksley et al. 1996; Chen et al. 1999; Bernacchi et al. 1998; Doganlar et al. 2002) donde se han descrito correlaciones negativas entre FW y SSC, atribuyendo el aumento SSC a disminuciones de peso de fruto y viceversa. Lo que podría interpretarse como que incorporando algunos alelos de TO-937 podemos aumentar los SSC sin afectar el tamaño del fruto y viceversa, al menos usando MM como fondo genético, lo cual resultaría de mucho valor en la mejora del cultivo. Haría falta comprobar estas correlaciones en otros fondos genéticos. Otro aspecto interesante fue la correlación negativa entre FW y FS, esto es debido a que los frutos cuanto más achatados son más pesados. Es decir, la variación en FS observada en esta población no se debe a variación en el crecimiento longitudinal que se observa en los mutantes *OVATE* (Ku et al. 2000) o *SUN* (van der Knaap y Tanksley 2001), sino a variación en el eje ecuatorial.

En resumen del total de QTLs definidos en el capítulo 4, 15 QTLs no habían sido definidos antes en otras poblaciones desarrolladas a partir de *S. pimpinellifolium* (sin considerar los QTLs de perímetro y área no descritos antes en *Solanáceas*), dentro de los cuales, siete están relacionados con las dimensiones del fruto, dos FS, uno de SSC, y 4 de color externo e interno del fruto. Aunque esta detección diferencial en la genoteca presente pueda deberse en parte a factores como el fondo genético y condiciones experimentales, quedando demostrada una vez más la utilidad de las genotecas de líneas de introgresión en el mapeo de QTLs y por otra parte, la presencia de variabilidad genética aprovechable en TO-937. A este respecto, creemos que las líneas que contienen QTLs como *fw12.1*, *ssc2.1*, *ssc3.1* y *ssc9.1* entre otros interesantes, podrían ser seleccionadas para realizar sub-ILs, que serían usadas en el mapeo fino del QTL. Con el objetivo de acotar lo máximo posible la región que contiene el QTL e identificar el gen (s) responsable del efecto. Seguidamente, se podrían piramidalizar algunos de estos QTLs de cara a ser introgresados en líneas de mejora con deficiencias en estos

caracteres, afín de estudiar su comportamiento en otros fondos genéticos. Por ejemplo se podrían cruzar las líneas que contengan *ssc2.1*, *ssc3.1* y *ssc9.1* potenciando al máximo el carácter en una sola línea que podrá ser luego cruzada con líneas elites en programas de mejora del cultivo. Los materiales y conocimiento obtenido serán pues utilizados en investigaciones de mejora de la calidad del fruto en los laboratorios participantes.



## CONCLUSIONES

- 1) Se ha construido una genoteca de 53 líneas de introgresión de la accesión TO-937 (*Solanum pimpinellifolium*: parental donante) en el cultivar comercial "MoneyMaker" (*Solanum lycopersicum*: parental recurrente). Dicha colección tiene una cobertura del 94%, cada IL tiene una introgresión diana homocigota de TO-937 con un tamaño promedio de 25 Mb (rango 0.7 – 75Mb) y en promedio 0.7 introgresiones no-dianas con un tamaño promedio de 6.8 Mb (rango 0.1 – 60 Mb), y en conjunto representan el genoma de TO-937 en un total de 71 bins con un tamaño promedio de 10 Mb (rango 0.9 – 65 Mb).
- 2) La estrategia utilizada para la generación de la genoteca de ILs combinó cruces controlados con técnicas de genotipado masivo, en poblaciones tempranas y etapas finales del cribado, lo que mejoró ostensiblemente la eficiencia, reduciendo el número de generaciones necesarias para obtener la genoteca completa de ILs, resultó además en ILs con un fondo genético altamente isogénico, reflejado por un bajo número de introgresiones no-dianas y en una cobertura cercana al 100% del parental donante.
- 3) Para disminuir las posibilidades de perder regiones del genoma durante el desarrollo de la genoteca y evitar el arrastre de grandes regiones centroméricas, se debe considerar tanto las distancias genéticas como físicas a la hora de elegir los marcadores a utilizar en el programa de selección asistida por marcadores.
- 4) Se identificó un locus que provoca distorsión en la segregación (SD) que favorece los alelos de TO-937 en detrimento de los alelos de MM y que afecta tanto los gametos femeninos como masculinos, con una penetración inferior al 100% y que se encuentra en el extremo distal del brazo largo del cromosoma 4. Dadas que las características observadas en este locus guarda mucha similitud con el descrito por Rick (1966), por lo que debe de tratarse del fenómeno descubierto por este autor y, que atribuyó a la acción de un gen eliminador de gametos (*Ge*).
- 5) Con el mapeo fino realizado en la región diana que presenta SD se logró definir una región final de 84 Kb en donde se localizan los siguientes genes candidatos: un gen *RLK* (Receptor like Kinasa), un gen *HSP90* (anotados endoplasmina), una Ciclina-B, una proteasa de procesamiento vacuolar, cuatro genes que codifican proteínas pequeñas de función desconocida y un retrotransposon.

6) El mapeo de QTLs realizado con la genoteca de ILs logro definir un total de 75 QTLs consistentes en al menos dos localidades, de los cuales 8 QTLs correspondieron a caracteres descriptivos (precocidad, vigor, frutos con hombros verdes, pilosidad en tallo y hojas, estructura del dosel y coloración violácea en tallos y ramas) y 67 QTLs corresponden a caracteres cuantitativos relacionados con calidad de fruto en las siguientes categorías: caracteres relacionados con las dimensiones de fruto, forma de fruto, caracteres organolépticos, color externo e interno del fruto.

7) Se lograron definir QTLs cuyo efecto es contrario a lo esperado dado el fenotipo del parental donante (QTL transgresivos): *fw12.1* que aumenta el peso del fruto en un 20% con respecto a MM, *cir6.1* y *cir 10.1*, que son responsables de formar frutos 44% menos redondos que MM, *fs 6.1* y *fs10.1* relacionados con los anteriores, son responsables de formar de fruto 6% mas achatadas con respecto a MM, *ssc4.1* y *ssc5.1*, producen frutos con el 10 y 13% menos de grados brix y *ta2.1* produce frutos con 17% menos de contenido de acidez titulable con respecto a MM.

8) Dentro de los QTLs que podrían ser interesantes en un programa de mejora destinado a aumentar el contenido de azucares en sus variedades, están *ssc2.1*, *ssc3.1* y *ssc9.1*, que en conjunto podrían aumentar más del 40% el contenido de SSC de MM, lo que podría representar un importante aumento en los grados brix de las variedades a desarrollar si se confirma la validez del QTL(s) en diferentes fondos genéticos.

9) Con respecto al control genético de la morfología del fruto en este cruzamiento en particular, el aumento en FW y las variaciones en FS no se deben a variación en el crecimiento longitudinal que se observa en los mutantes *OVATE* o *SUN* sino a variación en el eje ecuatorial.

## REFERENCIAS

- Aguado M (2007) DNA microarrays in microbiology. *RCCV* 1:125-134.
- Alba JM, Montserrat M, Fernández-Muñoz R (2009) Resistance to the two-spotted spider mite (*Tetranychus urticae*) by acylsucroses of wild tomato (*Solanum pimpinellifolium*) trichomes studied in a recombinant inbred line population. *Exp Appl Acarol* 47:35-47.
- Alexander L, Lincoln RE, Wright A (1942) A survey of the genus *Lycopersicon* for resistance to the important tomato disease occurring in Ohio and Indiana. *Plant Dis Rep Suppl* 136:51-85.
- Ali LM, Sanchez PL, Yu SB, Lorieux M, Eizenga CG (2010) Chromosome segment substitution lines: A powerful tool for the introgression of valuable genes from *Oriza* wild species into cultivated (*O. sativa*). *Rice* 3:218-234.
- Alpert KB, Tanksley S D (1996) High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: a major fruit weight quantitative trait locus in tomato. *PNAS* 93:15503-15507.
- Amien S, Kliwer I, Márton ML, Debener T, Geiger D, Becker D, Dresselhaus T (2010) Defensin like ZmES4 mediates pollen tube burst in maize via opening of the potassium channel KZM1. *PLoS Biol* 8:e1000388.
- Anderson LK, Covey PA, Larsen LR, Bedinger PA, Stack SM (2010) Structural differences in chromosomes distinguishes species in the tomato clade. *J Cytogen Genome Res* 129:24-34.
- Angaji SA (2009) Single nucleotide polymorphism genotyping and its application on mapping and marker-assisted plant breeding. *Afr J Biotechnol* 8:908-914.
- Artigas J M, Capilla P, Pujol J (2002) Tecnología del color. Universitat de Valencia, Valencia, España.
- Alseekh S, Ofner I, Pleban T, Tripoli P, Di Dato F, Cammareri M, Mohammad M, Grandillo S, Fernie AR, Zamir D (2013) Resolution by recombination: breaking up *Solanum Pennellii* introgressions. *Trend Plant Sci* 18:536-538.
- Ashrafi H, MP Kinkade, H Merk, MR Foolad (2012) Identification of novel QTLs for increased lycopene content and other fruit quality traits in a tomato RIL population. *Mol Breed* 30:549-567.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers *PloS ONE* 3:e3376.
- Barnes M, Freudenberg J, Thompson S, Aronow B, Pavlidis P (2005) Experimental comparison and cross-validation of the Afflymetrix and Illumina gene expression analysis platforms. *Nucl Acid Res* 33:5914-5923
- Barrero LS, Tanksley SD (2004) Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars. *Theor Appl Genet* 109:669-679.
- Barrett DM, Beaulieu JC, Shewfelt R (2010) Color, flavor, texture and nutritional quality of fresh-cut fruits and vegetables: desirable levels, instrumental and sensorial measurement, and the effects of processing. *Food Sci Nut* 50:369-389.
- Becerra VV y Paredes CM (2000) Use of biochemical and molecular markers in genetic diversity studies. *Agrc Tec* 60:270-281.
- Bedinger PA, Chetelat RT, MvClure B, Moyle LC, Rose JKC, Stack SM, Knaap E, Baek YS, Lopez-Casado G, Covet PA, Kumar A, Li W, Nunez R, Cruz-Garcia F, Royer S (2011) Interspecific reproductive barriers in the tomato clade: opportunities to decipher mechanisms of reproductive isolation. *Sex Plant Reprod* 24:171-187.
- Bernacchi D, Beck-Bunn T, Emmatty D, Eshed Y, Inai S, Lopez J, Petiard V, Sayama H, Uhlig J, Zamir D, Tanksley SD (1998) Advanced backcross QTL analysis of tomato. II

- Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *Lycopersicon pimpinellifolium*. *Theor Appl Genet* 97:170-180.
- Bernacchi D, Tanksley SD (1997) An interspecific backcross of *Lycopersicon esculentum* x *L. hirsutum*: linkage analysis and QTL study of sexual compatibility factors and floral traits. *Genetics* 147:861-877.
- Bernat M, Titos E, Claria J (2002) Rapid identification of single nucleotide polymorphisms by fluorescence-based capillary electrophoresis. *Genet Mol Res* 1:72-78.
- Berry SZ, Uddin MR (1991) Breeding tomato for quality and processing attributes. in *Genetic Improvement of Tomato*, G. Kalloo, Ed., Springer, Berlin, Germany of Monographs on. *Theor Appl Genet* 14:197-206.
- Bintou F, Hasedine Diouf H, Mergeai G (2012) Distorsions de ségrégation et amélioration génétique des plantes (synthèse bibliographique). *Biotechnol Agron Soc Environ* 16:499-509.
- Blair MW, Izquierdo P, Astidillo C, Grusak MA (2013) A legume biofortification quandary: variability and genetic control of seed coat micronutrient accumulation in common beans. *Front Plant Sci* 4:275.
- Blanca J, Cañizares J, Cordero L, Pascual L, Diez MJ, Nuez F (2012) Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PLoS ONE* 4:e48198.
- Bombarely A, Menda N, Teclé IY, Buels RM, Strickler S, Fischer-York T, Pujar A, Leto J, Gosselin J, Mueller LA (2011) The Sol Genomics Network (solgenomics.net): growing tomatoes using. *Perl Nucl Acids Res* 39:1149-1155.
- Botstein D, White RL, Skolnick M, David RW (1980) Construction of a genetic linkage map a man using restriction fragment length polymorphisms. *AM J Hum Genet* 32:314-331.
- Bournival BL, Scott JW, Vallejos CE (1989) An isozyme marker for resistance to race 3 of *Fusarium oxysporum* f.sp. *lycopersici* in tomato. *Theor Appl Genet* 78:489-494.
- Brandt S, Pék Z, Barna E, Lugasi A, Helyes L (2006) Lycopene content and color ripening tomatoes as affected by environmental conditions. *J Sci Food Agric* 86:568-572.
- Breto MP, Asins MJ, Carbonell EA (1993) Genetic variability in *Lycopersicon* species and their genetic relationships. *Theor Appl Genet* 86:113-120.
- Brewer MT, Lang L, Fujimura K, Dujmovic N, Gray S, van der Knaap E (2006) Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. *Plant Physiol* 141:15-25.
- Brummell DA, Harpster MH (2001) Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Mol Biol* 47:311-340.
- Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato PG, Martin C (2012) Retrotransposons Control Fruit-Specific, Cold-Dependent Accumulation of Anthocyanins in Blood Oranges. *The Plant Cell* 243:1242-1255.
- Butler L (1962) Muddy fruit, a phenotype of the gene *gf*. *Tomato Genet Coop Rpt* 12:19.
- Buttery RG (1993) Quantitative and sensory aspects of flavor of tomato and other vegetables and fruits. In: Acree TE, Teranishi R (eds) *Flavor Science: Sensible Principles and Techniques*. American Chemical Society, Washington, DC, USA. pp 259-286.
- Cameron DR, Moav RM (1957) Inheritance in *Nicotiana tabacum* XXVII. Pollen killer, an alien genetic locus inducing abortion of microspores not carrying it. *Genetics* 42:326-335.
- Campa A, Perez-Vega E, Giraldez R, Ferreira JJ (2012) Conceptos básicos sobre la elaboración de mapas genéticos de ligamiento en plantas. In: Ferreira J, Ordás A, Perez de la Vega M (eds) *Genética de los caracteres cuantitativos en la mejora vegetal del siglo XXI*, Serida-INIA, Madrid, España, pp 149-176.

- Canady MA, Meglic V, Chetelat RT (2005) A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome* 48:685-697.
- Cano A. 2006. Herencia de los mecanismos de resistencia a araña roja en tomate. In: Tesis Doctoral. Universidad de Málaga. España.
- Capel C, Salinas M, Ruiz-Rubio C, Hernández-Gras F, Lima V, Valpuesta V, Fernández del Carmen A, Rambla JL, Medina A, Fernández-Muñoz R, Boronat A, Botella MA, Granell A, Angosto T, Capel J, Lozano R (2011) A novel *Solanum lycopersicum* x *S. pimpinellifolium* genetic linkage map based on a RIL population displaying locations of QTL for fruit quality traits. XVII Eucarpia Meeting on Tomato Genetics and Breeding, Abstracts Book, P2-1, p. 45, Fuengirola (Málaga), Spain.
- Castro P, Rubio J, Cabrera A, Millan T (2011) A segregation distortion locus located on linkage group 4 of the chickpea genetic map. *Euphytica* 179:515-523.
- Causse M, Buret M, Robini K, Verschave P (2003) Inheritance of nutritional and sensory quality traits in fresh market tomato and relation to consumer preferences. *J Food Sci* 68:2342-2350.
- Causse M, Chaib J, Lecomte L, Buret M, Hospital F (2007) Both additivity and epistasis control the genetic variation for fruit quality traits in tomato. *Theor Appl Genet* 115:429-442.
- Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C (2004) A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. *J Exp Bot* 55:1671-1685.
- Causse M, Saliba-Colombani V, Lecomte L, Duffe P, Rousselle P, Buret M (2002) QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. *J Exp Bot* 53:2089-2098.
- Chagné D, Gasic K, Crowhurst RN, Hanb Y, Asset HC, Bowatte DR, Lawrence TJ, Rikkerink EHA, Gardiner SE, Korban CHS (2008) Development of a set SNP markers present in expressed genes of the apple. *Genomics* 92:353-358.
- Chaib J, Devaux MF, Grotte MG, Robini K, Causse M, Lahaye M, Marty I (2007) Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. *J Exp Bot* 58:1915-1925.
- Chakrabarti M, Zhang NA, Sauvage CH, Munos S, Blanca J, Cañizares J, Diez MJ, Schneider R, Mazourek M, McClead J, Causse M, van der Knaap E (2013) A cytochrome P450 regulates a domestication trait in cultivated tomato. *PNAS* 110:17125-17130.
- Chaque V, Mercier JC, Guenard M, de Courcel A, Vedel F (1997) Identification of RAPD markers linked to a locus involved in quantitative resistance to TYLCV in tomato by bulked segregant analysis. *Theor Appl Genet* 95:671-677.
- Chapman NH, Bonnet J, Grivet L, Lynn J, Graham N, Smith R, Sun G, Walley PG, Poole M, Causse M, King GJ, Baxter C, Seymour GB (2012) High resolution mapping of a fruit firmness-related QTL in tomato reveals epistatic interactions associated with a complex combinatorial locus. *Plant Physiol* 159:1644-1657.
- Chen FQ, Foolad MR (1999) A molecular linkage map of tomato based on an interspecific cross between *Lycopersicon esculentum* and *L. pimpinellifolium* and its comparison with other molecular maps of tomato. *Genome* 42:94-103.
- Chen YF, Matsubayashi Y, Sakagami Y (2000) Peptide growth factor phytosulfokine- $\alpha$  contributes to the pollen population effect. *Plant* 211:752-755.
- Chetelat RT y Devern, JW (1991) Expression of unilateral incompatibility in pollen of *Lycopersicon pennellii* is determined by major loci on chromosomes 1, 6 and 10. *Theor Appl Genet* 82:704-712.
- Chetelat RT, Meglic V (2000) Molecular mapping of chromosome segments introgressed from *Solanum lycopersicoides* into cultivated tomato (*Lycopersicon esculentum*) *Theor Appl Genet* 100:232-241.

- Chil A (1990) A Synopsis of *Solanum* Subgenus *Potatoe* (G. Don) D'Arcy (*Tuberarium* (Dum.) Bitter (s.l.). Feddes Report 101:209-235.
- Chunwongse J, Chunwongse C, Black L, Hanson P (2002) Molecular mapping of the *Ph-3* gene for the blight resistance in tomato. *J Hortic Sci Biotechnol* 77:281-286.
- Collard BC, Jahufer MZ, Brouwer JB, Pang EC (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concept. *Euphytica* 142:169-196.
- Collard BC, Mackill DL (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Phil Trans R Soc B* 363:557-572.
- Comai L, Young K, Till JB (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J* 37:778-786.
- Concibido VC, Vallee BL, Mcclaird P (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor Appl Genet* 106:575-582.
- Cong B, Barrero LS, Tanksley SD (2008) Regulatory change in YABBY like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nat Genet* 40:800-804.
- Covey PA, Konfok, Welch L, Frank E, Kumar A, Knaap E, Nunez R, Lopez-Casado G, Rose JKC, McClure BA, Bedinger PA (2010) Multiple features that distinguish unilateral incongruity and self-incompatibility in the tomato clade. *Plant J* 64:367-378.
- Covey PA, Subbaiah CC, Parsons RL, Pearce G, Lay FT, Anderson MA, Ryan CA, Bedinger PA (2010) A pollen-specific RALF from tomato that regulates pollen tube elongation. *Plant Physiol* 153:703-15.
- Croxford AE, Rogers T, Caligari PDS, Wilkinson MJ (2008) High-resolution melts analysis to identify and map sequence-target site anchor points onto linkage maps: a white lupin *Lupinus albus* map as an exemplar. *New Phytol* 180:594-607.
- Cuartero J, Fernández-Muñoz R, González-Fernández J (1995) Estreses abióticos. In: F Nuez (ed) *El cultivo del tomate*, Mundi-Prensa, Madrid, España, pp 351-384.
- Cuartero J, Nuez F, Díaz A (1984) Catalog of collections of *Lycopersicon* and *L. pennellii* from northwest of Perú. *Tomato Genet Coop Rpt* 34:43-46.
- Davey WJ, Hohenlohe PA, Etter PD (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499-510.
- De Koeber D, Douglass K, Murphy A (2010) Application of high-resolution DNA melting for genotyping and variant scanning of diploid and autotetraploid potato. *Mol Breed* 25:67-70.
- de Vicente, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134:585-596.
- Di Matteo A, Sacco A, Anacleria M, Pessotti M, Delledonne M, Ferrarini A, Frusciante L, Borobe A (2010) The ascorbic acid content of tomato fruits is associated with the expression of genes involved in pectin degradation. *BMC Plant Biol* 10:163.
- Doganlar S, Frary A, Ku HM, Tanksley SD (2002) Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). *Genome* 45:1189-1202.
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.
- Durbak RA, Frans ET (2011) CLAVATA Signaling Pathway Receptors of *Arabidopsis* Regulate Cell Proliferation in Fruit Organ Formation as well as in Meristems. *Genetics* 189:177-194.
- Dvorak J (1980) Homoeology between *Agropyron elongatum* chromosomes and *Triticum aestivum* chromosomes. *Can J Genet Cytol* 22:237-259.
- Dvorak J, Appels R (1986) Investigation of homologous crossing over and sister chromatid exchange in the wheat Nor-2 locus coding for rRNA and Gli-B2 locus coding for gliadins. *Genetics* 113:1037-1056.

- Echt CS, Kidwell KK, Knaap SJ, Osborn TC, Mccoy TJ (1994) Linkage mapping in diploid alfalfa (*Medicago sativa*). *Genome* 37:61-71.
- Eduardo I, Arùs P, Monforte AJ (2005) Development of a genomic library of near isogenic lines (NILs) in melon (*Cucumis melo* L.) from the exotic accession PI161375. *Theor Appl Genet* 112:139-148.
- Eichten SR, Foerster JM, de Leon N, Kai Y, Yeh CT, Liu S, Jeddloh JA, Schnable PS, Kaeppler SM, Springer NM (2011) B73-Mo17 near-isogenic lines demonstrate dispersal structural variation in maize. *Plant Physiol* 156:1679-1690.
- Ellis PR, Maxon-Smith JW (1971) Inheritance of resistance to potato cyst-eelworm (*Heterodera rostochiensis* Woll.) in the genus *Lycopersicon*. *Euphytica* 20:93-101.
- Endo TR (1990) Gametocidal chromosomes and their induction of chromosome mutations in wheat. *Jpn J Genet* 65:135-152.
- Eshed Y, Zamir D (1994) Introgressions from *Lycopersicon pennellii* can improve the solute-solids yield of tomato hybrids. *Theor Appl Genet* 88:891-897.
- Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* 141:1147-1162.
- Eshed Y, Zamir D (1996) Less-than-additive epistatic interactions of quantitative trait loci in tomato. *Genetics* 143:1807-1817
- Esquinas-Alcázar J (1981) Genetic resources of tomatoes and wild relatives. A global report. IBPGR, Rome, Italy.
- Esquinas-Alcázar J, Nuez F (1995) Situación taxonómica, domesticación y difusión del tomate. In: F Nuez (ed) *El cultivo del tomate*, Mundi-Prensa, Madrid, España, pp:13-42.
- Falconer DS (1989) *Introduction to Quantitative Genetics*, 3rd ed. Longman Scientific & Technical, Essex, UK.
- Falke KC, Miedaner T, Frish M (2009) Selection strategies for the development of rye introgression libraries. *Theor Appl Genet* 119:595-603.
- Fan JB, Chee MS, Gunderson KL (2006) Highly parallel genomic assays. *Nat Rev Genet* 7:632-644.
- Ferreira JJ, Ordás A, Pérez de la Vega M (2012) *La genética de los caracteres cuantitativos en la mejora vegetal del siglo XXI*. SERIDA-INIA Madrid, España.
- Fernandez-del-Carmen A, Abad J, Fernández-Muñoz R, Granell A, Monforte AJ (2011) Applications of the SolCap Illumina SNP array in tomato genetics. 8<sup>th</sup> Solanaceae and 2<sup>nd</sup> Cucurbitaceae Genome Joint Conference. Kobe, Japan. 28 November-2 December 2011.
- Fernández-Muñoz R, Domínguez E, Cuartero J (2000) A novel source of resistance to the two-spotted spider mite in *Lycopersicon pimpinellifolium* (Jusl.) Mill.: its genetics as affected by interplot interference. *Euphytica* 111:169-173.
- Fernández-Muñoz R, Salinas M, Alvarez M, Cuartero J (2003) Inheritance of resistance to the two-spotted mite and glandular leaf trichomes in wild tomato *Lycopersicon pimpinellifolium* (Jusl.) Mill. *J Am Soc Hort Sci* 128:188-195.
- Ferrándiz C, Pelaz S, Yanofsky MF (1999) Control of carpel and fruit development in *Arabidopsis*. *Annu Rev Biochem* 68:321-354.
- Finch RA, Miller TE y Bennett MD (1984) "Cuckoo" *Aegilops* addition chromosome in wheat ensures its transmission by causing chromosome breaks in meiospores lacking it. *Chromosoma* 90:84-88.
- Finkers R, Heusden AW, Dekens-Meijer F, Kan JA, Maris P, Lindhout P (2007) The construction of a *Solanum habrochaites* LYC4 introgression line population and the identification of QTLs for resistance to *Botrytis cinerea*. *Theor Appl Genet* 112:1360-1373.

- Foolad MR (1999) Comparison of salt tolerance during seed germination and vegetative growth in tomato by QTL mapping. *Genome* 42:727-734.
- Foolad MR (2005) Breeding for a biotic stress tolerances in tomato. In: Ashraf M, Harris PJC (eds) *Abiotics stresses: plant resistance through breeding and molecular approaches*. Haworth, New York, USA, pp 613-684.
- Foolad MR (2007) Genome mapping and molecular breeding of tomato. *Int J Plant Genomics* doi:10.1155/2007/64358.
- Foolad MR, Chen FQ, Lin GY (1998) RFLP mapping of QTLs conferring cold tolerance during seed germination in an interspecific cross of tomato. *Mol Breed* 4:519-529.
- Foolad MR, Sharma A (2005) Molecular markers as selection tools in tomato breeding. *Acta Hort* 695:115-240.
- Francia E, Tacconi G, Crosatti C, Barabaschi D, Bulgarelli D, Dall'Aglio E, Vale G (2005) Marker assisted selection in crop plants. *Plant Cell Tiss Organ Cult* 82: 317-342.
- Francis DM, Kabelka E, Bell J, Franchino B, St. Clair D (2001) Resistance to bacterial canker in tomato (*Lycopersicon hirsutum* LA407) and its progeny derived from crosses to *L. esculentum*. *Plant Dis* 85:1171-1176
- Franklin-Tong VE, Franklin FCH (2003) The different mechanisms of gametophytic self-incompatibility. *Phil Trans R Soc Lond B* 358:1025-1032.
- Frary A, Fulton TM, Zamir D, Tanksley SD (2004) Advanced backcross QTL analysis of a *Lycopersicon esculentum* X *L. pennellii* cross and identification of possible orthologs in the Solanaceae. *Theor Appl Genet* 108:485-496.
- Frary A, Nesbitt TC, Grandillo S (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289:85-88.
- Fridman E, Carrari F, Liu YS, Zamir D (2004) Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* 305:1786-1789.
- Fridman E, Pleban T, Zamir D (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484pb within an invertase gene. *PNAS* 97:4718-4723.
- Friebe B, Zhang P, Nasuda S, Gill BS (2003) Characterization of a knock-out mutation at the *Gc2* locus in wheat. *Chromosoma* 111:509-517.
- Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD (1997a) QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTL found in other wild species. *Theor Appl Genet* 95:881-94.
- Fulton TM, Nelson JC, Tanksley SD (1997b) Introgresión and DNA marker analysis of *Lycopersicon peruvianum*, a wild relative of the cultivated tomato, into *Lycopersicon esculentum*, followed through three successive backcross generations. *Theor Appl Genet* 95: 895-902.
- Fulton TM, Grandillo S, Beck-Bun T, Fridman E, Frampton A, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD (2000) Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *Lycopersicon parviflorum* cross. *Theor Appl Genet* 100:1025-1042.
- Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. *Plant Cell* 14:1457-1467.
- Fu Q, Zhang P, Tan L, Zhu Za, Ma D, Fu Y, Zhan X, Cai H, Sun Ch (2010) Analysis of QTLs for yield related traits in Yuanjiang common wild rice (*Oriza rufipogon* Griff) *J Genet Genomics* 37:147-157.
- Garcia-Dorado A, Gallego A (1992) On the use of the classical tests for detecting linkage. *J Hered* 83:143-146.

- Georgelis N, Scott JW, Baldwin EA (2004) Relationship of tomato fruit sugar concentration with physical and chemical traits and linkage of RAPD markers. *J Am Soc Hort Sci* 129:839-845.
- Giovannucci E (1999) Tomatoes, tomato-based products, lycopene, and cancer: review of the epidemiologic literature. *J Natl Cancer Inst* 91:317-331.
- Goldraij A, Kondo K, Lee CB, Hancock CN, Sivaguru M, Vazquez-Santana S, Kim S, Phillips TE, Cruz-Garcia F, McClure B (2006) Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. *Nature* 439:805-810.
- Gonzalo MJ, van der Knaap E (2008) A comparative analysis into the genetic bases of morphology in tomato varieties exhibiting elongated fruit shape. *Theor Appl Genet* 116:647-656.
- Gould WA. 1983. *Tomato Production, processing and quality evaluation*, 2ed. AVI Publishing Company, Inc. Westport, CT USA. pp 3-50.
- Goulet CH, Mageroy MH, Nghi BL, Floystad A, Tieman Dm, Klee HJ (2012) Role of an esterase in flavor volatile variation within the tomato clade. *PNAS* 109:19009-19014.
- Graham EB, Frary A, Kang JJ, Jones CM, Gardener RG (2004) A recombinant inbred line mapping population derived from a *Lycopersicon esculentum* × *L. pimpinellifolium* cross, *Tomato Genet Coop Rep* 54:22-25.
- Grandillo S (2014) Introgression libraries with wild relatives of crops. In: Tuberosa R, Graner A, Frison E (eds) *Genomics of Plant Genetics Resources*, Springer Science+Business Media Dordrecht, pp 87-122.
- Grandillo S, Chetelat R, van der Knaap E, Spooner D, Peralta I, Cammareri M, Perez O, Termolino P, Tripodi P, Chiusano MI, Ercolano MR, Frusciantè L, Monti L, Pignone D (2011) *Solanum* sect. *Lycopersicon*. In: Kole C (ed) *Wild Crop Relatives: Genomic and Breeding Resources*, Vol. 5 Vegetables. Springer, Netherlands, pp 129-215.
- Grandillo S, Ku HM, Tanksley SD (1999) Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theor Appl Genet* 99:978-987.
- Grandillo S, Tanksley SD (1996) QTL Analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. *Theor Appl Genet* 92:935-951.
- Grandillo S, Tanksley SD (2005) Advanced backcross QTL analysis: Results and perspectives. In: Tuberosa R, Phillips RL, Gale M (eds) *Proceeding of the International Congress "The wake of the double helix: From the green revolution to the gene revolution"*. Edizioni, Avenue Media, Bologna, Italia, pp 115-132.
- Grandillo S, Termolino P, van der Knaap E (2013) Molecular mapping of Complex Traits in Tomato. In: Leidl BE, Labate JA, Stommel JR, Slade A, Kole Ch (eds) *Genetics, Genomic and Breeding of Tomato*, CRC Press, US, pp 150-228.
- Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. *Theor Appl Genet* 83:250-256.
- Grierson D, Kader AA (1986) Fruit ripening and quality. In: Atherton JG, Rudich J (Eds.) *The tomato crop: A scientific basis for improvement* Chapman and Hall, London: pp 241-280.
- Gundry CN, Vandersteen JG, Reed GH, Pryor RJ, Chen J, Wittwer CT (2003) Amplicon Melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clin Chem* 49:396-406
- Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: a new paradigm for molecular marker technology and AND polymorphism detection with emphasis on their use in plants. *Curr Sci* 80:524-535.
- Gur A, Zamir D (2004) Unused natural variation can lift yield barriers in plant breeding. *PLoS Biol* 2:e245.

- Haldane JBS (1919) The combination of linkage values and the calculation of distance between the loci of linked factors. *J Genet* 8:299-309.
- Hamada H, Petrino MG, Kakunaga T (1982) A novel repeated element with z-dna-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *PNAS* 79:6465-6469.
- Hamilton JP, Sim S, Stoffel K, Van Deynze A, Buell CR (2012) Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. *Plant Gen* 5:17-29.
- Hancock CN, Kent L, McClure BA (2005) The stylar 120 kDa glycoprotein is required for S-specific pollen rejection in *Nicotiana*. *Plant J* 43:716-723.
- Hanson PM, Sitathani K, Sadashiva AT, Yang R, Graham E, Ledesma D (2007) Performance of *Solanum habrochaites* LA 1777 introgression line hybrids for marketable tomato fruit yield in Asia. *Euphytica* 158:167-178.
- Han Y, Khu DM, Monteros MJ (2012) High-resolution melting analysis for SNP genotyping and mapping in tetraploid alfalfa (*Medicago sativa* L.). *Mol Breed* 29:489-501.
- Hertog ML, Lammertyn J, Scheerlink N, Nicolai BM (2007) The impact of biological variation on postharvest behaviour: the case of dynamic temperature conditions. *Postharvest Biol Technol* 43:183-192.
- Hewitt JD, Garvey TC (1987) Wild sources of high soluble solids in tomato. In: Nevins DJ, Jones RA (eds) *Plant Biology. Vol 4: Tomato Biotechnology*, AR Liss, New York, USA, pp 45-54.
- Higashiyama T (2010) Peptide signaling in pollen-pistil interactions. *Plant Cell Physiol* 51:177-189.
- Holton TA (2001) Plant genotyping by analysis of microsatellites. In: Henry RJ (ed) *Plant genotyping: the DNA fingerprinting of plant*. CAB International, NY, USA, pp15-27.
- Huang XQ, Coster H, Ganai MW, Roder MS (2003) Advanced backcross QTL analysis for the identification of quantitative trait loci alleles wild relatives wheat *Triticum aestivum* L. *Theor Appl Genet* 106:1379-1389.
- Hua Z, Kao TH (2006) Identification and characterization of components of a putative petunia S-locus F-box-containing E3 ligase complex involved in S-RNase-based self-incompatibility. *Plant Cell* 18:2531-2553.
- Ignatova SI, Gorshkova NS, Tereshonkova TA (2000) Resistance of tomato F1 hybrids to grey mold. *Acta Physiol Plant* 22:326-328.
- Ikeda H, Hiraga M, Shirasawa K, Nishiyama M, Kanahama K, Kanayama Y (2013) Analysis of a tomato introgression line IL8-3, with increased Brix content. *Sci Hortic* 153:103-108.
- Iltis HH (2000) Homeotic sexual translocations and the origin of maize (*Zea mays*, Poaceae): A new look at an old problem. *Econ Bot* 54:7-42.
- Ishiguro S, Watanabe Y, Ito N, Nonaka H, Tadeka N, Sakai T, Kanya H, Okada K (2002) SHEPHERD is the *Arabidopsis* GRP94 responsible for the formation of functional CLAVATA proteins. *EMBO* 21:898-908.
- Iwano M, Takayama S (2012) Self/non-self discrimination in angiosperm self-incompatibility. *Curr Opin Plant Biol* 15:78-83.
- Jarne P, Lagoda PJ (1996) Microsatellites, from molecules to populations and back. *Trends Ecol Evol* 11:424-429.
- Jarret RL, Sayama H, Tigchelaar EC (1984) Pleiotropic effects associated with the chlorophyll intensifier mutations *high pigment* and *dark green* in tomato. *J Am Soc Hort Sci* 109:873-787.
- Jeuken MJW, Lindhout P (2004) The development of lettuce backcrossing inbred lines (BILs) for exploitation of the *Lactuca saligna* (wild lettuce) germoplasm. *Theor Appl Genet* 109:394-401.

- Jimenez-Duran K, McClure B, Garcia-Campusano F, Rodriguez-Sotres R, Cisneros J, Busot G, Cruz-Garcia F (2013) NaStEP: a proteinase inhibitor essential to self-incompatibility and a positive regulator of HT-B stability in *Nicotiana glauca* pollen tubes. *Plant Physiol* 161:97-107.
- Johnson MA, Preuss D (2003) On your mark, get set, GROW LePRK2-LAT52 interactions regulate pollen tube growth. *Trends Plant Sci* 8:97-99.
- Jones N, Ougham H, Thomas H (1997) Markers and mapping. We are all geneticists now. *New Phytol* 137:165-177.
- Jones RA, Scott SJ (1983) Improvement of tomato flavor by genetically increasing sugar and acid contents. *Euphytica* 32: 845-855.
- Jones RA, Scott SJ (1984) Genetic potential to improve tomato flavor in commercial F<sub>1</sub> hybrids. *J Am Soc Hort Sci* 109:318-321.
- Kabelka E, Franchino B, Francis DM (2002) Two loci from *Lycopersicon hirsutum* LA 407 confer resistance to strains of *Clavibacter michiganensis* subsp. *michiganensis*. *Phytopathology* 92:504-510.
- Kermicle JL, Evans MMS (2010) The *Zea mays* Sexual Compatibility Gene *ga2*: Naturally Occurring Alleles, Their Distribution, and Role in Reproductive Isolation. *J Hered* 101: 737-749.
- Kermicle JL, Evans MMS (2005) Pollen-pistil barriers to crossing in maize and teosinte result from incongruity rather than active rejection. *Sex Plant Reprod* 18:187-194.
- Keurentjes BJ, Bentsink L, Blanco CA, Hanhart CJ, De Vries HB, Effgen S, Vreugdenhil D, Koornneef M (2007) Development of a near-isogenic line population of *Arabidopsis thaliana* and comparison of mapping power with a recombinant inbred line population. *Genetics* 175:891-905.
- Kim S, Govindan JA, Tu ZJ, Greenstein D (2012) SACY-1 DEAD-Box helicase links the somatic control of oocyte meiotic maturation to the sperm-to-oocyte switch and gamete maintenance in *Caenorhabditis elegans*. *Genetics* 192:905-928.
- Kindale PM, Foolad MR (2013) Validation and fine mapping of *lyc12.1*, a QTL for increased tomato fruit lycopene content. *Theor Appl Genet* 126:2163-2175.
- Klee HJ, Tieman DM (2013) Genetic challenges of flavor improvement in tomato. *Trends in genetics* 29:257-262.
- Kondo K, Yamamoto M, Itahashi R, Sato T, Egashira H, Hattori T, Kowayama Y (2002) Insights into the evolution of self-compatibility in *Lycopersicon* from a study of stylar factors. *Plant J* 30:143-53.
- Korff MV, Wang H, Leon J, Pillen K (2004) Development of candidate introgression using exotic barley accession (*Hordeum vulgare* ssp. *Spontaneum*) as donor. *Theor Appl Genet* 109:1736-1745.
- Kosambi DD (1944) The estimation of map distances from recombination values. *Ann Eugenics* 12:172-175.
- Koumproglou R, Wilkes TW, Townson P, Wang XY, Beynon J, Pooni HS, Newbury HJ, Kearsley MJ (2002) STAIRS: a new genetic resource for functional genomics studies of *Arabidopsis*. *Plant J* 31:355-364.
- Kubo K, Entani T, Takara A, Wang N, Fields AM, Hua Z, Toyoda M, Kawashima S, Ando T, Isogai A, Kao TH, Takayama S (2010) Collaborative non-self recognition system in S-RNase-based self-incompatibility. *Science* 330:796-799.
- Ku HM, Doganlar S, Chen KY, Tanksley SD (1999) The genetic basis of pear-shaped tomato fruit. *Theor Appl Genet* 99:844-850.
- Ku HM, Grandillo S, Tanksley SD (2000) *fs8.1*, a major QTL, sets the pattern of tomato carpel shape well before anthesis. *Theor Appl Genet* 101:873-878.

- Ky CL, Barre P, Lorieux M, Trouslot P, Akaffou S, Louarn J, Charrier A, Hamon S, Nairot M (2000) Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.) *Theor Appl Genet* 101:669-676.
- Labate JA, Grandillo S, Fulton T, Muños S, Caicedo AL, Peralta I, Ji Y, Chetelat RT, Scott JW, Gonzalo MJ, Francis D, Yang W, van der Knaap E, Baldo AM, Smith-White B, Mueller LA, Prince JP, Blanchard NE, Storey DB, Stevens MR, Robbins MD, Wang JF, Liedl BE, O'Connell MA, Stommel JR, Aoki K, Iijima Y, Slade AJ, Hurst SR, Loeffler D, Steine MN, Vafeados D, McGuire C, Freeman C, Amen A, Goodstal J, Facciotti D, Van Eck J, Causse M (2007) Tomato. In: Kole C (ed) *Genome mapping and molecular breeding in plants. Vol 5: Vegetables*. Springer, Berlin, Heidelberg, Germany, pp 1-96.
- Lacombe S, Rougon CA, Sherwood E, Peeters N, Dahlbeck D, van Esse P, Smoker M, Rallapalli G, thomma BP, Staskwicz B, Jones J, Zipfel C (2010) Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. *Nat Biotechnol* 28:365-369.
- Ladizinsky G (1985) Founder effect in crop-plant evolution. *Econ Bot* 33:191-199.
- Lambeth VN, Straten EF, Fields ML (1966) *Fruit Quality Attributes of 250 Foreign and Domestic Tomato Accessions*, Agricultural Experiment Station, University of Missouri, Columbia, Mo, USA.
- Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M, Charrier A (2001) Genetic linkage map of *Coffea canephora*: effect of segregation rate in male and female meioses. *Genome* 44:589-596.
- Laterrot H (1996) Twenty-one near isogenic lines in Moneymaker type with different genes for disease resistances. *Tomato Genet Coop Rpt* 46:34.
- Lee C, Kim J, Bae S (2008) Identification of mycobacterium using high resolution amplicon DNA melting analysis. *Clin Chem* 54:160.
- Lee JM, Joung JG, McQuinn R, Chung MY, Fei Z, Tieman D, Klee H, Giovannoni J (2012) Combined transcriptome genetic diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor SIERF6 plays an important role in ripening and carotenoid accumulation. *Plant J* 70:191-204.
- Libault M, Stacey G. (2010) Evolution of FW2.2-like (FWL) and PLAC8 genes in eukaryotes. *Plant Signal Behav* 5:1226-1228.
- Li W, Lin Z, Zhang X (2007) A novel segregation distortion in intraspecific population of Asian cotton (*Gossypium arboreum* L.) detected by molecular markers. *J Genet Genomics* 34:634-640.
- Li W, Royer S, Chetelat RT (2010) Fine mapping of UI6.1, a gametophytic factor controlling pollen-side unilateral incompatibility in interespecific *Solanum* hybrids. *Genetics* 185:1069-1080.
- Li W, Chetelat RT (2014) The role of a pollen-expressed Cullin1 protein in gametophytic self-incompatibility in *Solanum*. *Genetics* 196:439-442.
- Liew M, Pryor R, Palais R, Meadows C, Erali M, Lyon E, Wittwer C (2004) Genotyping of single-nucleotide polymorphisms by High-Resolution Melting of small amplicon. *Clin Chem* 50:1156-1164.
- Lima-Silva V, Rosado A, Amorin-Silva V, Muñoz-Merida A, Pons C, Bombarely A, Trelles O, Fernandez-Muñoz R, Granell A, Valpuerta V, Botella MA (2012) Genetic on genome-wide transcriptomic analyses identify co-regulation of oxidative response and hormone transcript abundance with vitamin C content in tomato fruit. *BMC Genomic* 13:187.
- Lin SY, Sasaki T, Yano M (1998) Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa* L., using backcross inbred lines. *Theor Appl Genet* 96:997-1003.

- Lippman ZB, Semel Y, Zamir D (2007) An integrated view of quantitative trait variation using tomato interspecific introgression lines. *Curr Opin Genet Dev* 17:545-552.
- Lippman ZB, Tanksley SD (2001) "Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom," *Genetics* 158:413-422.
- Liu J, Zhang Y, Qin G, Tsuge T, Sakaguchi N, Luo G, Sun K, Shi D, Aki S, Zheng N, Aoyama T, Oka A, Yang W, Umeda M, Xie Q, Gu H, Qua L (2008) Targeted degradation of the Cyclin-Dependent Kinase inhibitor ICK4/KRP6 by RING-Type E3 ligases is essential for mitotic cell cycle progression during *Arabidopsis* gametogenesis. *Plant Cell* 20:1538-1554.
- Liu JP, van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *PNAS* 99:813302-813306.
- Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinform* 21:2128-2129.
- Liu S, Zhou R, Dong Y (2006) Development, utilization of introgression lines using synthetic wheat as donor. *Theor Appl Genet* 114:1071-1080.
- Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the *Solanaceae*. *Genetics* 152:1183-1202.
- Loefering WQ, Sears ER (1963) Distorted inheritance of stem-rust resistance of timstein wheat caused by a pollen-killing gene. *Can J Genet Cytol* 5:65-72.
- Lorieux M, Goffinet B, Perrier X, León G, Lanaud C (1995a) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations. *Theor Appl Genet* 90:73-80.
- Lorieux M, Perrier X, M, Goffinet B, Lanaud C, León G (1995b) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 2. F2 populations. *Theor Appl Genet* 90:81-89.
- Lu H, Romero-J. Severson, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. *Theor Appl Genet* 105:622-628.
- MacNeil BH, Kerr EA (1984) Chromosomal identity and linkage relationships of *Pto*, a gene for resistance to *Pseudomonas syringae* pv. *tomato* in tomato. *J Plant Pathol* 6:48-53.
- Mangelsdorf PC, Jones DF (1926) The expression of mendelian factors in the gemetophyte of maize. *Genetics* 11:423-455.
- Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S (2012) SNP Markers and Their Impact on Plant Breeding. *Int J Plant Genomics*. doi:10.1155/2012/728398.
- Matteo AD, Sacco A, Anacleria M, Pezzotti M, Delledonne M, Ferrarini A, Frusciante L, Barone A (2010) The ascorbic acid content of tomato fruits is associated with the expression of genes involved in pectin degradation. *MBC Plant Biol* 10:163.
- Matus I, Corey A, Filichkin T (2003) Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor allele in a *Hordeum vulgare* subsp. *vulgare* background. *Genome* 46:1010-1023.
- McClure BA, Haring V, Elbert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE (1989) Style self-incompatibility gene products of *Nicotiana glauca* are ribonucleasas. *Nature* 342:955-957.
- McClure BA, Mou B, Canevascini S, Bernatzky R (1999) A small asparagine-rich protein required for S-allele-specific pollen rejection in *Nicotiana*. *PNAS* 96:13548-13553.
- Mei HW, XU JL, Li JL (2006) QTLs influencing panicle size detected in two reciprocal introgressive line (IL) populations in rice (*O. sativa* L.) *Theor Appl Genet* 112:648-656.
- Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor Appl Genet* 80:437-448.

- Miller P (1754) The gardener`s dictionary, Abridged 4<sup>th</sup> ed. John and James. Rivington, London, UK.
- Mizrahi Y, Taleisnik E, Kagan-Sur V (1988) A saline irrigation regime for improving tomato fruit quality without reducing yield. *J Am Soc Hort Sci* 113:202-205.
- Moncada P, Martinez CP, Borrero J, Chatel M, Gauch Jr H, Guimaraes E, Tohme J, McCouch SR (2001) Quantitative trait loci for yield and yield components in an *Oryza sativa* × *Oryza rufipogon* BC2 F2 population evaluated in an upland environment. *Theor Appl Genet* 102:41-52.
- Monforte AJ, Asins AJ, Carbonell EA (1996) Salt tolerance in *Lycopersicon* species IV. Efficiency of marker-assisted selection for salt tolerance improvement. *Theor Appl Genet* 93:765-772.
- Monforte AJ, Diaz AI, Caño-Delgado A, van der Knaap E (2014) The genetic basic of fruit morphology in horticultural crops: lessons from tomato and melon. *J Exp Bot.* doi:1031093/jxb/eru017.
- Monforte AJ, Friedman E, Zamir D, Tanksley (2001) Comparison of set of allelic QTL\_NILs for chromosome 4 of tomato deductions about natural variation and implications for germoplasm utilization. *Theor Appl Genet* 102:572-590.
- Monforte AJ, Tanksley SD (2000a) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. *Genome* 43:803-813.
- Monforte AJ, Tanksley SD (2000b) Fine mapping of a quantitative trait locus (QTL) from *Lycopersicon hirsutum* chromosome 1 affecting fruit characteristics and agronomic traits: breaking linkage among QTLs affecting different traits and dissection of heterosis for yield. *Theor Appl Genet* 100:471-479.
- Montoya T, Nomura T, Farrar K, Kaneta T, Yokota T, Bishop G J (2002) Cloning the tomato *curl3* gene highlights the putative dual role of the leucine-rich repeat receptor kinase tBRI1/SR160 in plant steroid hormone and peptide hormone signaling. *The Plant Cell* 14:3163-76.
- Moore S, Vrebalov J, Payton P, Giovannoni J (2002) Use of genomics tools to isolate key ripening genes and analyze fruit maturation in tomato. *J Exp Bot* 53:2023-2030.
- Moyle LC, Graham EB (2005) Genetics of hybrid incompatibility between *Lycopersicon esculentum* and *L. hirsutum*. *Genetics* 169:355-373
- Moyle LC, Nakazato T (2008) Comparative genetics of hybrid incompatibility: sterility in two *Solanum* species crosses. *Genetics* 179:1437-1453.
- Muños S, Ranc N, Botton E, Bérard A, Rolland S, Duffe´ P, Carretero Y, Le Paslier MC, Delalande C, Bouzayen M, Brunel D, Causse M (2011) Increase in tomato locule number is controlled by two SNPs located near WUSCHEL. *Plant Physiol* 156:2244-2254.
- Muschiatti J, Dircks L, Vancanneyt G, McCormick S (1994) LAT52 protein is essential for tomato pollen development: pollen expressing antisenseLAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. *Plant J* 6:321-338.
- Mutschler MA (1984) Inheritance and linkage of the Alcobaca ripening mutant in tomato. *J Am Soc Hort Sci* 109:500-503.
- Nakagahra M (1972) Genetic mechanism on the distorter segregation of marker genes belonging to the eleventh linkage group in cultivated rice. *Jpn J Breed* 22:232-238
- Nasuda S, Friebe B y Gill BS (1998) Gametocidal genes induce chromosome breakage in the interphase prior to the first mitotic cell division of the male gametophyte in wheat. *Genetics* 149:1115-1124.
- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data II. Gene frequency data. *J. Mol. Evol* 19:153-170.

- Nelson OE (1994) The gametophyte factors of maize. In: Freeling M, Walbot V (eds) *The Maize Handbook*, Springer-Verlag, New York USA, pp.496-503
- Nguyen ML y Schwartz SJ (1999) Lycopene: chemical and biological properties. *Food Technol* 53:38-45.
- Nikolausz M, Chatzinotas A, Palatinszky M, Imfeld G, Martinez P, Kästner M (2003) Single-Nucleotide Primer Extension Assay for Detection and Sequence Typing of "Dehalococcoides" spp. *Appl Environ Microbiol* 74:300-304.
- Nuez F, Carrillo JM, Ron AM (2000) Introducción. In: Nuez F, Carrillo JM (ed) *Los marcadores genéticos en la mejora vegetal*, Universidad Politécnica de Valencia, Valencia, España pp 23-89.
- Nyquist WE (1962) Differential fertilization in the inheritance of stem rust resistance in hybrids involving a common wheat strain derived from *Triticum timopheevi*. *Genetics* 47:1109-1124.
- O'Brien SL (1990) *Genetics maps: locus maps of complex genomes*. Cold Spring Harbor Laboratory Press (ed), New York, USA.
- O'Brien M, Kapfer C, Major G, Laurin M, Bertrand C, Kond, K, Kowyama Y, Matton DP (2002) Molecular analysis of the stylar- expressed *Solanum chacoense* small asparagine-rich protein family related to the HT modifier of gametophytic self-incompatibility in *Nicotiana*. *Plant J* 32:985-996.
- Okuda S, Sutsui T, Shiina K, Keiko S, Sprunck S, Takeuchi H, Hidenori T, Ryoko Y, Ryushiro D K, Yuki H, Akane M, Daichi S, Nao K, Takashi S, Shoko N, Kie I, Kurataka O, Motomichi M, Hisayoshi N, Tsuneyoshi K, Akihiko N, Masahiro K, Thomas D, Narie S, Tetsuya H (2009) Defensin like polypeptide LUREs are pollen tube attractants secreted from synergid cells. *Nature* 458:357-361.
- Olmstead RG, Bohs L (2007) A Summary of molecular systematic research in Solanaceae: 1982-2006. *Acta Hort* 745:255-268.
- Oliver M, Garcia-Mas J, Cardus J, Puedo N, Lopez-Sesé AI, Arroyo M, Gómez-Paniagua H, de Vicente MC (2001) Construction of a reference linkage map for melon. *Genome* 44:836-845.
- Paran I, van der Knaap E (2007) Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J Exp Bot* 58:3841-3852.
- Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map restriction fragment length polymorphisms. *Nature* 335:721-726.
- Pea G, Aung HH, Frascaroli E, Landi P, Pé ME (2013) Extensive genomic characterization of set of near-isogenic lines for heterotic QTL in maize (*Zea mays* L.). *BMC Genomics* 14:61
- Pelham J (1968) Disturbed segregation of genes on chromosome 9: gamete promoter, *Gp*, a new gene. *Rep. Tomato Genet Coop Rep* 18:27-29.
- Peralta IE, Knapp S, Spooner DM (2005) New species of wild tomatoes (*Solanum* Section *Lycopersicon*: Solanaceae) from Northern Peru. *Syst Bot* 30:424-434.
- Peralta IE, Spooner DM, Knapp S (2008) Taxonomy of wild tomatoes and their relatives (*Solanum* sections *Lycopersicoides*, *Juglandifolia*, *Lycopersicon*; Solanaceae). *Syst Bot Monogr* 84:1-186.
- Pereira MG Lee M, Bromel-Cox P, Woodman JW, Doeblen J, Whitkus R (1994) Construction of an RFLP map in sorghum and comparative mapping in maize. *Genome* 37:236-243.
- Pestsova EG, Borner A, Roder MS (2006) Development and QTL assessment of *Triticum aestivum*-*Aegilops tauschii* introgression lines. *Theor Appl Genet* 112:634-647.

- Peterson DG, Price HJ, Johnston S, Stack SM (1996) DNA content of heterochromatin and euchromatin in tomato (*Lycopersicon esculentum*) pachytene chromosomes. *Genome* 39:77-82.
- Pillen K, Zacharias A, León J (2003) Advanced backcross QTL of yield in barley *Hordeum vulgare* L. *Theor Appl Genet* 107:340-352.
- Pitblado RE, MacNeil BH, Kerr EA (1984) Chromosomal identity and linkage relationships of *Pto*, a gene for resistance to *Pseudomonas syringae* pv. *tomato* in tomato. *Can J Plant Pathol* 6:48-53.
- Powell A, Nguyen C, Hill T, Cheng KL, Figueroa R, Aktos H, Ashrafi H, Pons C, Fernandez-Muñoz R, Vicente A, Lopez-Baltazar J, Barry C, Liu Y, Chetelat R, Granell A, Deynze A, Giovannoni J, Bennett (2012) Uniform ripening Encodes a Golden 2-like Transcription Factor Science Regulating Tomato Fruit Chloroplast Development. *Science* 336:1711-1715.
- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trend Plant Sci* 1:215-222.
- Prohens J, Nuez F (2008) Handbook of plant breeding: vegetables II. Springer New York, USA pp 249-326.
- Puerta AR, Ushijima K, Koba T, Sassa H (2009) Identification and functional analysis of pistil self-incompatibility factor HT-B of Petunia. *J Exp Bot* 60:1309-1318.
- Pui-Yan K (2001) Methods for genotyping single nucleotide polymorphisms. *Ann Rev Genomics Hum Genetics* 2:235-258.
- Quiros C (1991) *Lycopersicon* cytogenetics. In: Tsuchiya T, Gupta PK (eds) Chromosome engineering in plants: genetics, breeding, evolutions (part B). Elsevier Science Publishers, Amsterdam, pp 119-137.
- Ragoussis J (2009) Genotyping technologies for genetic research. *Ann Rev Genomics Hum Genet* 10:117-133.
- Rambla JL, Tikunov YM, Monforte AJ, Bovy AG, Granell A (2014) The expanded tomato fruit volatile landscape. *J Exp Bot* (In press).
- Ramsay LD, Jennings DE, Kearsey MJ (1996) The construction of a substitution library of recombinant backcross lines in *Brassica oleraceae* for the precision mapping of quantitative trait loci. *Genome* 39:558-567.
- Rao GU, Ben Chaim A, Borovsky Y y Paran I (2003) Mapping of yield-related QTLs in pepper an interspecific cross of *Capsicum annuum* and *C. frutescens*. *Theor Appl Genet* 106:1457-1466.
- Reed GH, Kent JO, Wittwer CT (2007) High-Resolution DNA melting analysis for single and efficient molecular diagnostics. *Pharmacogenomics* 8:597-608.
- Rick C, Smith MPG (1953) Novel variation in tomato species hybrids. *Am Nat* 87: 359-373.
- Rick CM (1966) Abortion of male and female gametes in the tomato determined by allelic interaction. *Genetics* 53:85-96.
- Rick CM (1971) The Tomato *Ge* Locus Linkage Relations and Geographic Distribution of Alleles. *Genetics* 67:75-85.
- Rick CM (1979) Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*. In: Hawkes JG, Lester RN, Skelding AD (eds) The Biology and Taxonomy of the Solanaceae, Linnean Soc Symposium Series No 7, pp 667-678.
- Rick CM (1986) "Genetic resources in *Lycopersicon*," In: Nevins DJ, Jones RA (eds) Plant Biology. Vol 4: Tomato Biotechnology, AR Liss, New York, USA pp17-26.
- Rodriguez GR, Munoz S, Anderson C, Sim SC, Michael A, Causse M, Gardener M, Francis D, van der Knaap E (2011) Distribution of SUN, OVATE, LC and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiol* 156:275-285.

- Ronen GL, Cohen M, Zamir D, Hirschberg J (1999) Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is downregulated during ripening and is elevated in the mutant *Delta*. *Plant J* 17:341-351.
- Ronen G, Carmel-Goren L, Zamir D, Hirschberg J (2000) An alternative pathway to beta-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *PNAS* 97:11102-11107.
- Rooney WL, Stelly DM (1991) Preferential transmission and somatic elimination of a *Gossypium sturtianum* chromosome in *G. hirsutum*. *J Heredity* 82:151-155.
- Rousseaux MC, Jones CM, Adams D, Chetelat R, Bennett A, Powell A (2005) QTL analysis of fruit antioxidants in tomato using *Lycopersicon pennellii* introgression lines. *Theor Appl Genet* 111:1396-1408.
- Sacks EJ, Francis DM (2001) Genetic and environmental variation for flesh color of tomato fruit in a population of modern breeding lines. *J Am Soc Hort Sci* 126:221-226.
- Saladié M, Matas AJ, Isaacson T, Jenks MA, Mark Goodwin S, Niklas KJ, Xiaolin R, Labavitch JM, Shackel KA, Fernie AR, Lytovchenko A, O'Neill MA, Watkins ChB, Rose JKC (2007) A reevaluation of the key factors that influence tomato fruit softening and integrity. *Plant Physiol* 144:1012-1028.
- Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M (2001) Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theor Appl Genet* 102:259-272.
- Salinas M, Capel C, Alba JM, Mora, B, Cuartero J, Fernández-Muñoz R, Lozano R, Capel J (2013) Genetic mapping of two QTL from the wild tomato *Solanum pimpinellifolium* L. controlling resistance against two-spotted spider mite (*Tetranychus urticae* Koch). *Theor Appl Genet* 126:83-92.
- Sano Y, Chu Y-E, Oka H-I (1979) Genetic studies of speciation in cultivated rice. I. Genic analysis for the F1 sterility between *O. sativa* L. and *O. glaberrima* Steud. *Jap J Genet* 54:121-132.
- Sano Y (1983) A new gene controlling sterility in F1 hybrids of two cultivated rice species. *J Hered* 74:435-439.
- Sano L, Sano R, Eiguchi M, Hirano Y (1994) Gamete eliminator adjacent to the *wx* locus as revealed by pollen analysis in Rice. *J Hered* 85:310-312.
- Sato K, Close T, Bhat P, Muñoz-Amatrian M, Muehlbauer GJ (2011) Single Nucleotide Polymorphism Mapping and Alignment of Recombinant Chromosome Substitution Lines in Barley. *Plant Cell Physiol* 52:728-737.
- Sax K (1923) Association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8:552-560.
- Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotechnol* 24:447-454.
- Schmalenbach I, March TJ, Bringezu T, Waugh R, Pillen K (2011) High-resolution genotyping of wild barley introgression lines and fine-mapping of the threshability locus *thresh-1* using the Illumina GoldenGate assay. *G3* 1:187-196.
- Schopfer CR, Nasrallah ME, Nasrallah JB (1999) The male determinant of self-incompatibility in Brassica. *Science* 286:1697-1700.
- Schwartz D (1950) The analysis of a case of cross-sterility in maize. *PNAS* 36:719-724.
- Scoles GJ y Kibirge-Sebunya IN (1982) Preferential abortion of gametes in wheat induced by an *Agropyron* chromosome. *Can J Genet Cytol* 25:1-6.
- Septiningsih EM, Prasetyono J, Lubis E, Tai TH, Tjubaryat T, Moeljopawiro S, McCouch SR (2003) Identification of quantitative trait loci for yield and yield components in an

- advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theor Appl Genet* 107:1419-1432.
- Seymour B, Chapman NH, Chew B, Rose JKC (2013) Regulation of ripening and opportunities for control in tomato and other fruits. *Plant Biotech J* 11:269-278
- Shirasawa K, Isobe S, Hirakawa H, Asamizu E, Fukuoka H, Just D, Rothan Ch, Sasamoto Sh, Fujishiro T, Kishida Y, Kohara M, Tsuruoka H, Wada T, Nakamura Y, Sato S, Tabata S (2010) SNP discovery and linkage map construction in cultivated tomato. *DNA Res* 17:381-391.
- Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao TH (2004) Identification of the pollen determinant of S-RNase-mediated self-incompatibility. *Nature* 429:302-305.
- Sim S-C, Durstewitz G, Plieske J, Wieseke R, Ganai MW, Van Deynze A, Hamilton JP, Robin Buell C, Causse M, Wijeratne S, Francis DM (2012) Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE* 7:e40563.
- Singh KB y Ocampo B (1997) Exploitation of wild *Cicer* species for yield improvement in chickpea. *Theor Appl Genet* 95:418-423.
- Slimestad S, Verheul M (2009) Review of flavonoids and other phenolics from fruit of different tomato (*Lycopersicon esculentum* Mill) cultivars. *J Sci Food Agri* 89:1255-1270.
- Smith BL, LU CP, Bremer AR (2010) High-resolution melting analysis (HRMA): a highly sensitive inexpensive genotyping alternative for population studies. *Mol Ecol Res* 10:193-196.
- Song X, Sun X, Zhang T (2006) Segregation distortion and its effect on genetic mapping in plants. *Chin J Agric Biotechnol* 3:163-169.
- Song X, Wang K, Guo W, Zhang J Zhang T (2005) A comparison of genetic maps constructed from haploid and BC1 mapping populations from the same crossing between *Gossypium hirsutum* L. and *Gossypium barbadense* L. *Genome* 48:378-390.
- Spooner DM, Peralta IE, Knapp S (2005) Comparison of AFLPs to other markers for phylogenetic inference in wild tomatoes (*Solanum* L. Section *Lycopersicon* (Mill. (Wattst)). *Taxon* 54:43-61.
- Steinhauser MC, Steinhauser D, Gibon Y, Bolger M, Arrivault S, Usadel B, Zamir D, Fernie AR, Stitt M (2011) Identification of enzyme activity quantitative trait loci in a *Solanum lycopersicum* x *Solanum pennellii* introgression line population. *Plant Physiol* 157:998-1014.
- Stevens MA, Rick CM (1986) Genetics and breeding. In: Atherton JG, Rudich J (eds) *The tomato crop: a scientific basic for improvement*. Chapman and Hall, London, UK, pp 35-109.
- Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43:1160-1163.
- Studer B, Jensen L, Fiil A, Asp T (2009) “Blind” mapping of genic DNA sequence polymorphisms in *Lolium perenne* L. by high resolution melting curve analysis. *Mol Breed* 24:191-199.
- Sukamoto T, Ando T, Takahashi K, Omori T, Watanabe H, Kokubun H, Marchesi E, Kao TH (2003) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by loss of pollen function. *Plant Physiol* 131:1903-1912.
- Swamy M, Sarla N (2008) Yield-enhancing quantitative trait loci (QTLs) from wild species. *Biotechnol Adv* 26:106-120.
- Szalme SJ, Hostert BM, Le Deaux JR, Stuber CW, Holland JB (2007) QTL mapping with near-isogenic lines in maize. *Theor Appl Genet* 114:1211-1228.

- Takayama S, Shiba H, Iwano M, Shimosato H, Che F, Kai N, Watanabe M, Suzuki G, Hinata K, Isogai A (2000) The pollen determinant of self-incompatibility in *Brassica campestris*. PNAS 97:1920-1925.
- Takeuchi H, Higashiyama T (2012) A Species-Specific Cluster of Defensin-Like Genes Encodes Diffusible Pollen Tube Attractants in *Arabidopsis*. PLoS Biol 10:e1001449.
- Talluri RS (2012) Interploidy interspecific hybridization in *Fuchsia*. J Genet 91:71-74.
- Tandon KS, Baldwin EA, Scott JW, Shewfelt RL (2003) Linking sensory descriptors to volatile and nonvolatile components of fresh tomato flavor. J Food Sci 68:2366-2371.
- Tang W, Ezcurra I, Muschietti J, McCormick S (2002) A cysteine-rich extracellular protein, LAT52, interacts with the extracellular domain of the pollen receptor kinase LePRK2. Plant Cell 14:2277-2287.
- Tang W, Kelley D, Ezcurra I, Cotter R, McCormick S (2004) LeSTIG1, an extracellular binding partner for the pollen receptor kinases LePRK1 and LePRK2, promotes pollen tube growth in vitro. Plant J 39:343-353
- Tanksley SD (2004) The genetic developmental and molecular bases of fruit size and shape variation in tomato. Plant Cell 16:S181-S189.
- Tanksley SD, Ganai MW, Prince JP, De Vincent MC, Bonierbael MW, Broun P, Fulton JJ (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141-1160.
- Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet 92:213-224.
- Tanksley SD, Loaiza-Figueroa F (1985) Gametophytic self compatibility is controlled by single locus on chromosome 1 in *Lycopersicon peruvianum*. Proc Am Soc Hort Sci 82:5093-5096.
- Tanksley SD, McCouch SR (1997) Seed banks molecular maps: unlocking genetic from the wild. Science 277:1063-1066.
- Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191-203.
- Tanksley SD, Yound ND, Paterson AH, Bonierbael MW (1989) RFLP mapping in plant breeding: new tools for an old science. Nat Biotechnol 7:257-264.
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 30: 2725-2729.
- Taylor DR, Ingvarsson PK (2003) Common features of segregation distortion in plants and animals. Genetica 117:27-35.
- The Tomato Genome Sequence Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635-641.
- Thoday JM (1960) Location of polygenes. Nature 191:368-370.
- Tian F, Li DJ, Fu Q, Zhu ZF, Fu YC, Wang XK, Sun CQ (2006) Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. Theor Appl Genet 112:570-580.
- Tieman D, Bliss P, Mclatyre LM, Blondon-ubeda A, Bies D, Odabasi AZ, Rodriguez GR, van der Knaap E, Taylor MG, Goulet C, Mageroy MH, Snyder DJ, Colguhoun T, Moskowitz H, Clark DG, Sims C, Bartoshuk L, Klee HJ (2012) The chemical interactions underlying tomato flavor preferences. Curr Biol 22:1035-1039.
- Tigchelaar EC, McGlasson W, Buescher R (1978) Genetic regulation of tomato fruit ripening. Hort Sci 13:508-513.

- Torjek O, Meyer RC, Zehnsdorf M, Teltow M, Strompen G, Witucka-Wall H, Blacha A, Altmann, T (2008) Construction and analysis of 2 reciprocal *Arabidopsis* introgression line populations. *J Hered* 99:396-406.
- Uauy C, Distelfeld A, Fahima T, Blechl A y Dubcovsky J (2006) A NAC gene regulation senescence improves grain protein, zinc and iron content in wheat. *Science* 314:1298-1301.
- Vancetovic J (2008) An impact of environment on segregation ratio of qualitative traits in maize. *Genetika* 40:145-156.
- Vandersteen JC, Bayrak-Toydemir P, Palais RA, Wittwer CT (2007) Identifying common genetic variants by high-resolution melting. *Clin Chem* 53:1191-1198.
- van der Hoeven R, Ronning C, Giovannoni J, Martin G, Tanksley S (2002) Deductions about the number, organization, and evolution of genes in the tomato genome base on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell* 14:1441-1456.
- van der Knaap E, Lippman ZB, Tanksley SD (2002) Extremely elongated tomato fruit controlled by four quantitative trait loci with epistatic interactions. *Theor Appl Genet* 104:241-247.
- van der Knaap E, Tanksley SD (2001) Identification and characterization of a novel locus controlling early fruit development in tomato. *Theor Appl Genet* 103:353-358.
- van der Knaap E, Tanksley SD (2003) The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in yellow stuffer tomato. *Theor Appl Genet* 107:139-147.
- van Tuinen A, Cordonnier-Pratt MM, Pratt LH, Verkerk R, Zabel P, Koornneef M (1997) The mapping of phytochrome genes and photomorphogenic mutants of tomato. *Theor Appl Genet* 94:115-122.
- Villalta I, Reina-Sanchéz A, Bolarín MC, Cuartero J, Belver A, Venema K, Carbonell EA, Asins MJ (2008) Genetic analysis of Na<sup>+</sup> and K<sup>+</sup> concentrations in leaf and stem as physiological components of salt tolerance in tomato. *Theor Appl Genet* 116:869-880.
- von Korff M, Wang H, Leo'n J, Pillen K (2004) Development of candidate introgression lines using an exotic barley accession (*Hordeum vulgare* ssp. *spontaneum*) as donor. *Theor Appl Genet* 109:1736-1745.
- Vossen R, Aten E, Roos A, Dunnen JT (2009) High-Resolution melting analysis (HRMA) more than just sequence variant screening. *Hum Mutat* 30:860-866.
- Wang YJ, WU X, Yu D, Zhang Y, Chen S, Gai J (2004) Method of evaluation and adjustment of recombinant inbred line population and its application to the soybean RIL population NJRIKY. *Acta Agron Sinica* 30:413-418.
- Wehrhahn C, Allard W (1965) The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. *Genetics* 51:109-119.
- Weller JI, Soller M, Brady T (1988) Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* x *Lycopersicon pimpinellifolium*) by means of genetic markers. *Genetics* 118:329-339.
- Wentao L, Chetelat RT (2010) A pollen factor linked inter- and intraspecific pollen rejection in tomato. *Science* 230:1827-1830.
- Willcox JK, Catignani G L, Lazarus S (2003) Tomatoes and cardiovascular health. *Crit Rev Food Sci Nutr* 43:1-18.
- Winter P, Kahl G (1995) Molecular marker technologies for plant improvement. *World J Microbiol Biotechnol* 11:438-448.
- Wittwer CT (2009) High-resolution DNA melting analysis: advancements and limitations. *Hum Mutat* 30:857-859.

- Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-resolution genotyping by amplicon melting analysis using LC Green. *Clin Chem* 49:853-860.
- Wu SB, Tavassolian I, Rabiei G, Hunt P, Wirthensohn M, Gibson JP, Ford ChM, Sedgley M (2009) Mapping SNP-anchored genes using High-resolution melting analysis in almond. *Mol Genet Genomics* 282:273-281.
- Wu SB, Wirthensohn MG, Hunt P, Gibson JP, Sedgley M (2008) High Resolution melting analysis of almond SNPs derived from ESTs. *Theor Appl Genet* 118:1-14.
- Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* 319:1527-1530.
- Xiao J, Li J, Grandillo S, Ahn S, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150:899-909.
- Xu J, Zhao Q, Du P, Xu Ch, Wang B, Feng Q, Liu Q, Tang S, Gu M, Han B, Liang G (2010) Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (*Oryza sativa* L.). *Genomics* 11:2-14.
- Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fong L, Huang L, Li J, He W, Zhang G, Zheng X, Zhang F, Li Y, Yu Ch, Kristiansen K, Zhang X, Wang J, Wright M, McCouch S, Nielsen R, Wang J, Wang W (2012) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat Biotechnol* 30:105-111.
- Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F<sub>2</sub>, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol Gen Genet* 253:535-545.
- Yen HC, Shelton BA, Howard LR, Lee S, Vrebalov J, Giovannoni JJ (1997) The tomato high-pigment (*hp*) locus maps to chromosome 2 and influences plastome copy number and fruit quality. *Theor Appl Genet* 95:1069-1079.
- Zamir D (2001) Improving plant breeding with exotic genetic libraries. *Nat Rev Genet* 2:983-989.
- Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. *Bot Gaz* 147:355-358.
- Zhang NA, Brewer MT, van der Knaap E (2012) Fine mapping of *fw3.2* controlling fruit weight in tomato. *Theor Appl Genet* 125:273-284.
- Zhan H y Xu S (2011) Generalized linear mixed model for segregation distortion analysis. *BMC Genetics* 12:97.
- Zhao L, Huang J, Zhao Z, Li Q, Sims, TL, Xue Y (2010) The Skp1-like protein SSK1 is required for cross-pollen compatibility in S-RNase-based self-incompatibility. *Plant J* 62:52-63.

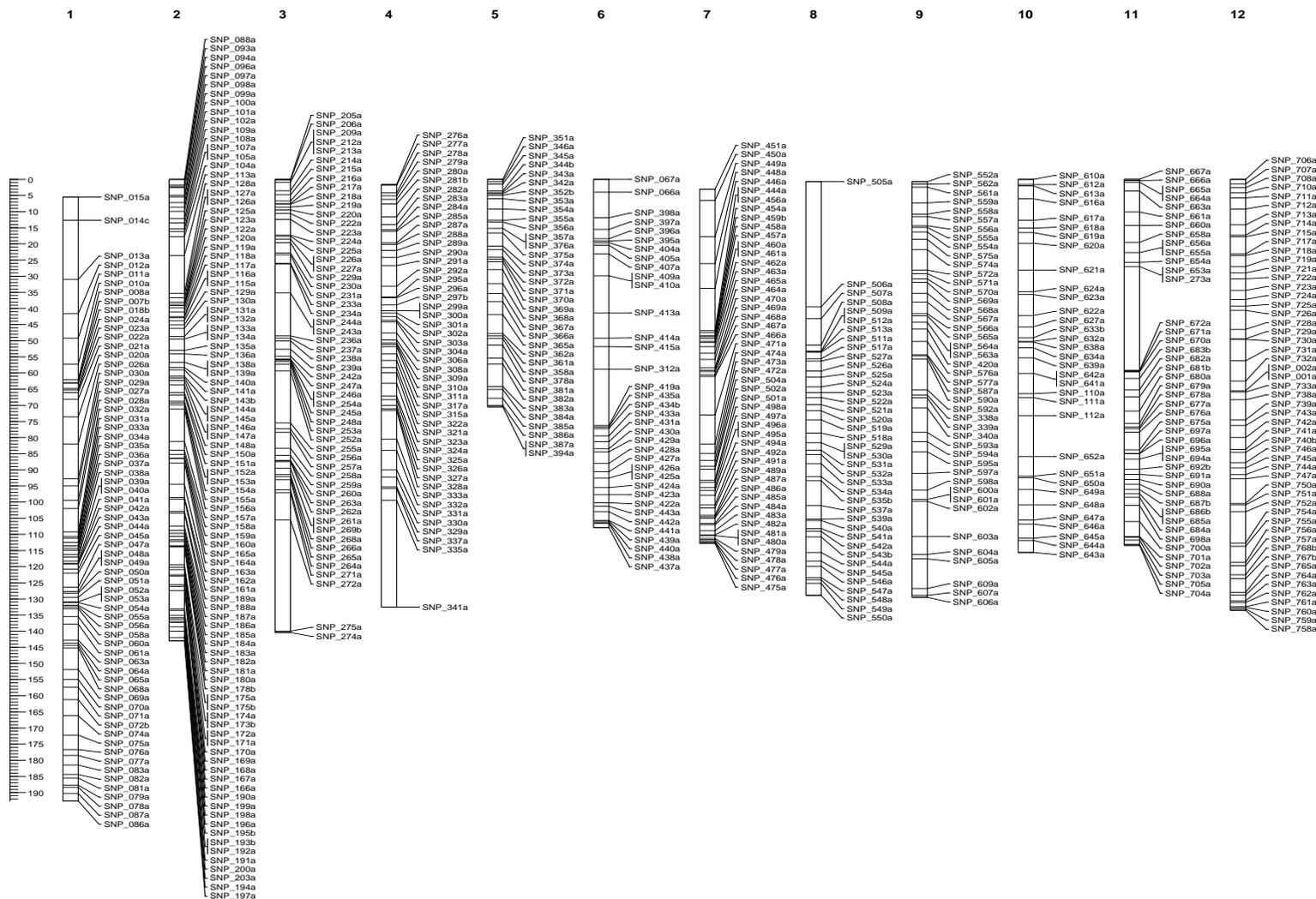


## ANEXOS

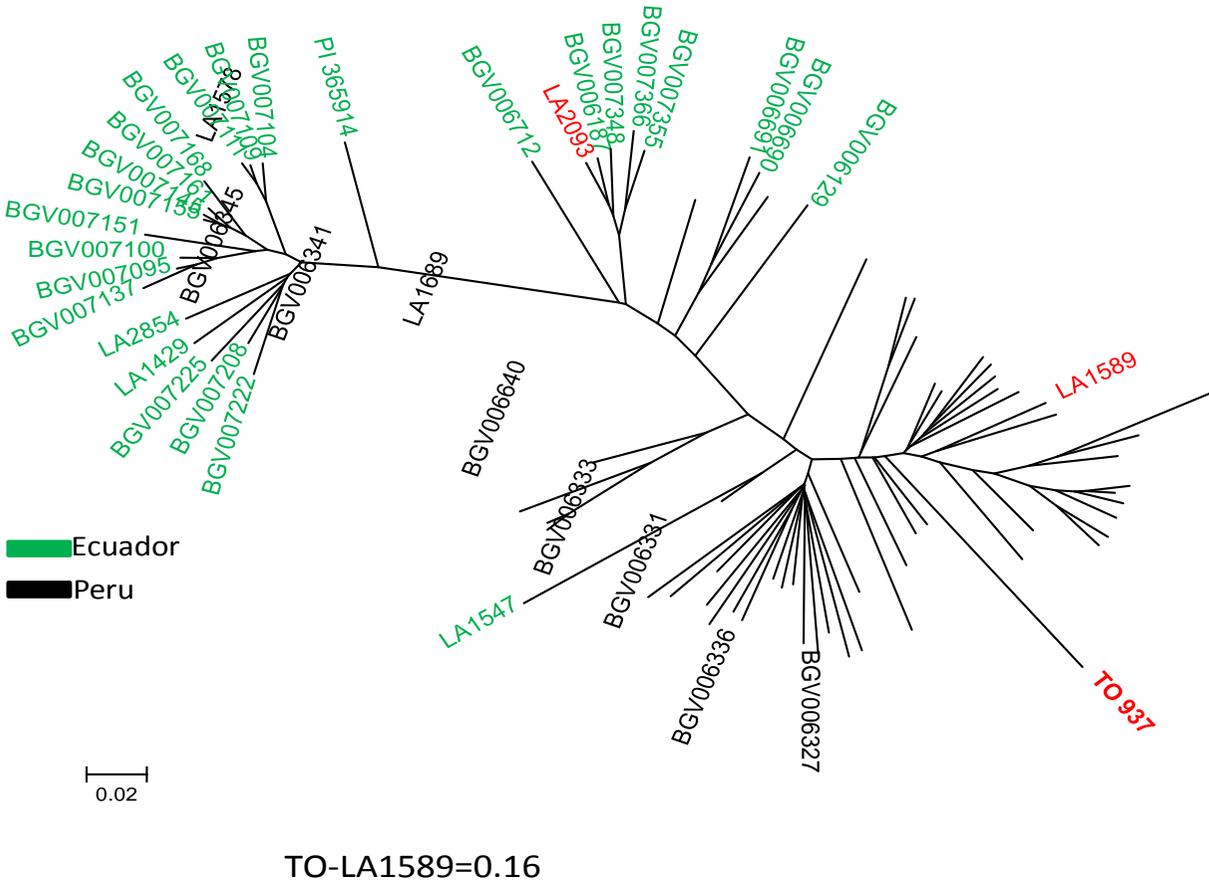


## **1. Anexos (Capítulo 2)**

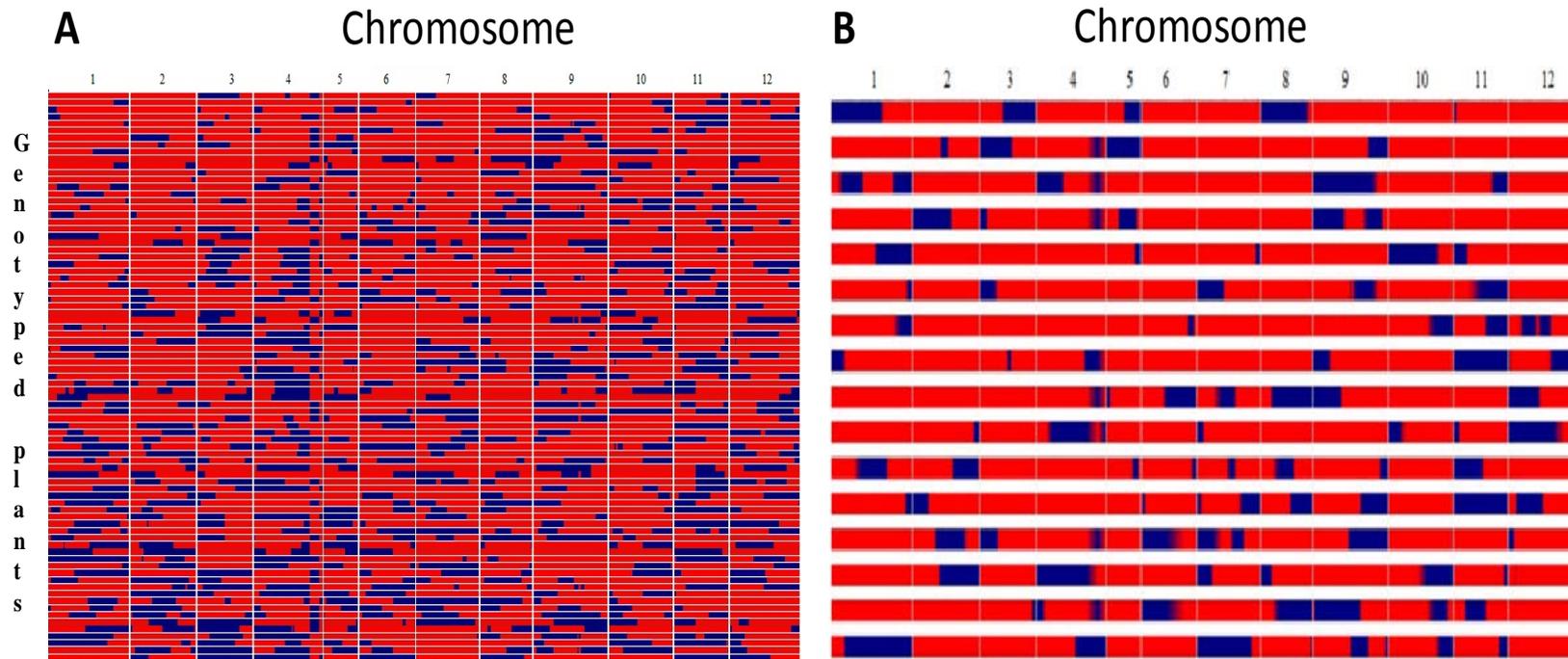
# Annex 2-1. Genetic map including the SNP markers used to genotype the BC<sub>2</sub> population.



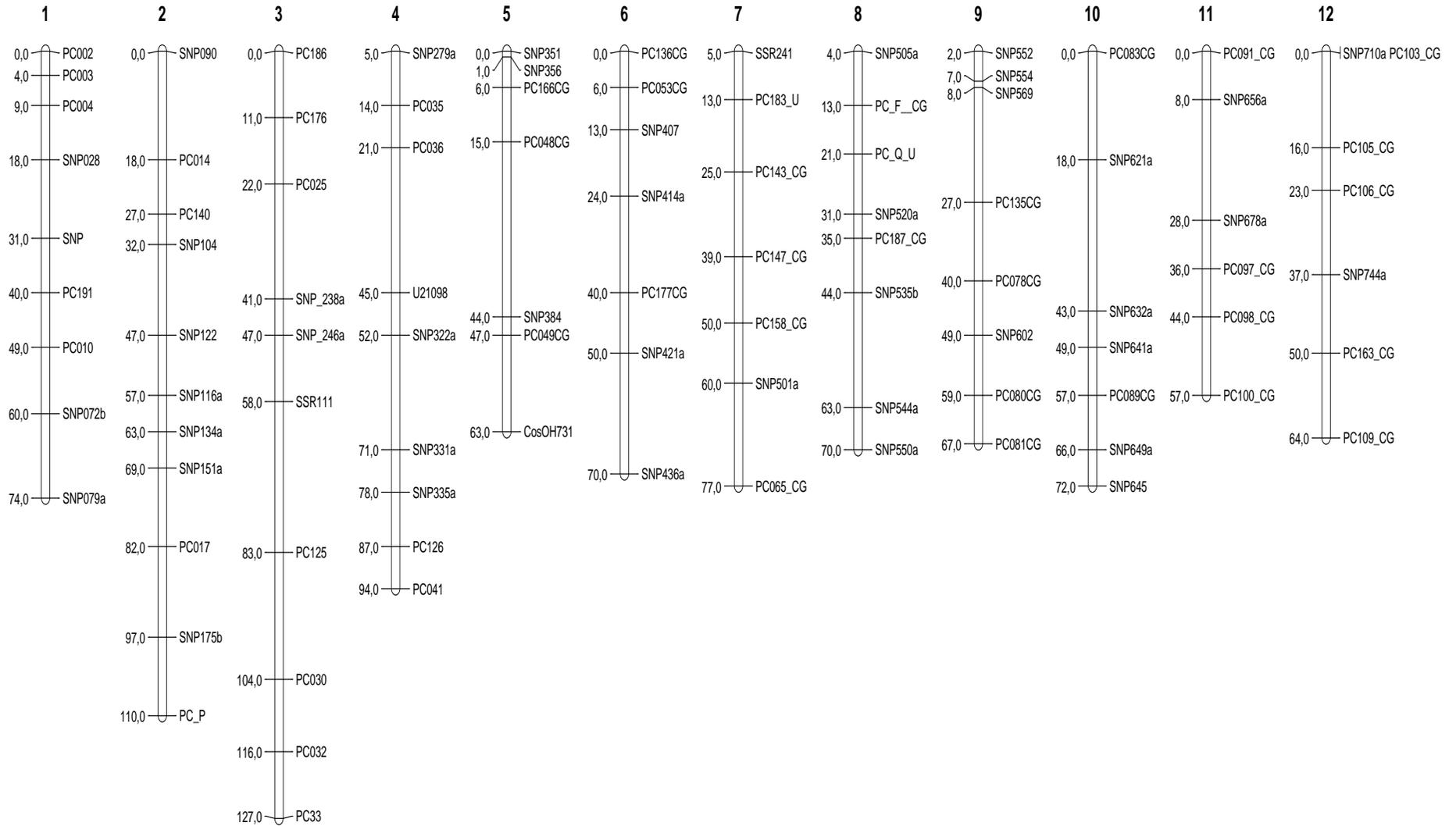
**Annex 2-2. Neighbor-joining dendrogram of *S. pimpinellifolium* accessions calculated from Nei (1983) genetic distance based on 8K SNP SOLCAP Infinium array genotyping.** Data from all accessions except from TO-937 were obtained from Sim et al. (2012) and Blanca et al. (2012). TO-937, LA1589, and LA2093 accessions (the latter two have also been used in other works for several QTL analyses) are highlighted in red. The origin of the accessions is indicated by the colors green for Ecuador, and black for Peru.



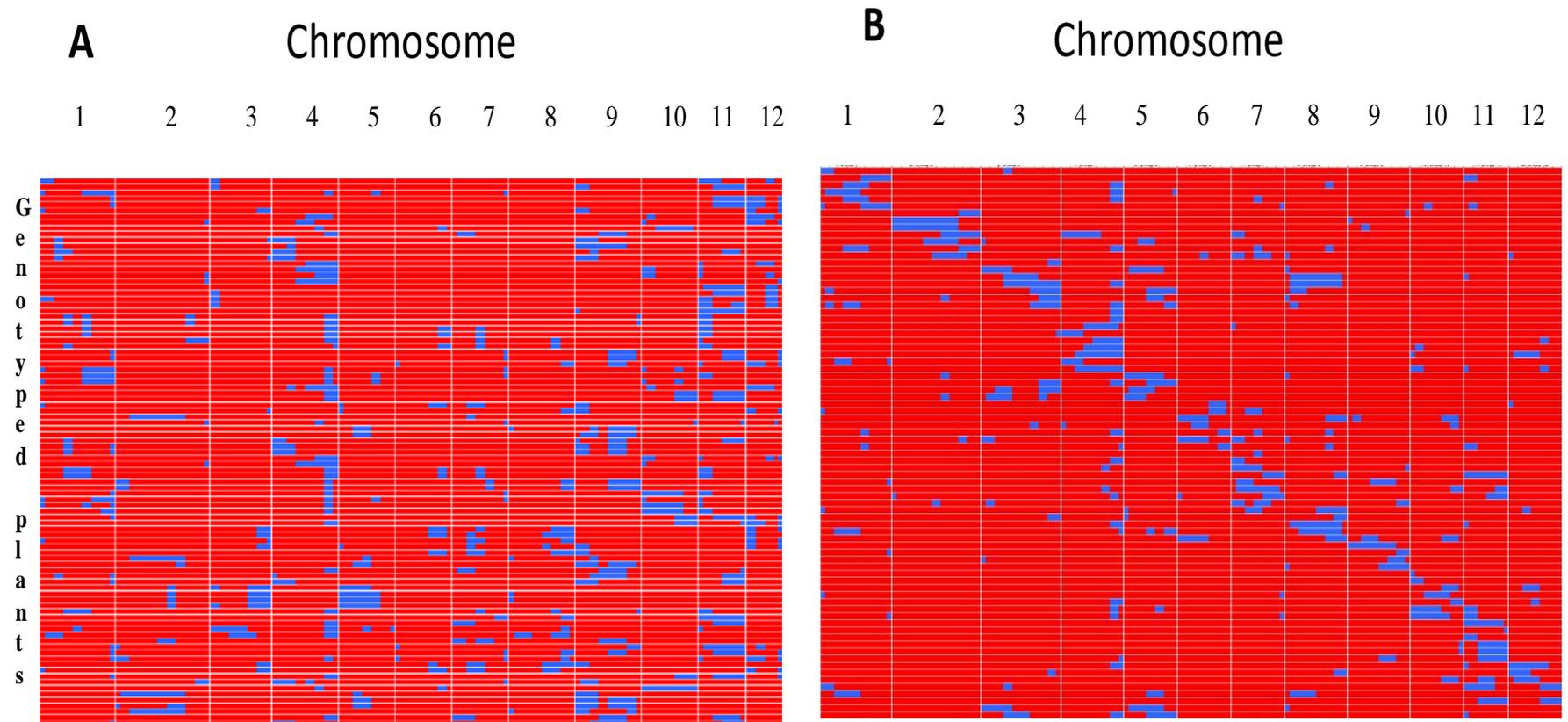
**Annex 2-3. Graphical genotypes of unselected (A) and selected (B) BC2 population that have been characterized with the 665 polymorphic SNP markers indicated in the Anex 2-1. Blue and red areas indicate regions that are heterozygous and homozygous for cv. 'Moneymaker', respectively.**



**Annex 2-4. Genetic map of the SNP markers used to genotype the BC<sub>3</sub> population.**



**Annex 2-5. Graphical genotypes of unselected (A) and selected (B) BC<sub>3</sub> population that have been characterized with 96 SNP markers (Illumina Veracode array). Blue and red regions indicate regions that are heterozygous or homozygous for the recurrent parent cv. 'Moneymaker', respectively.**



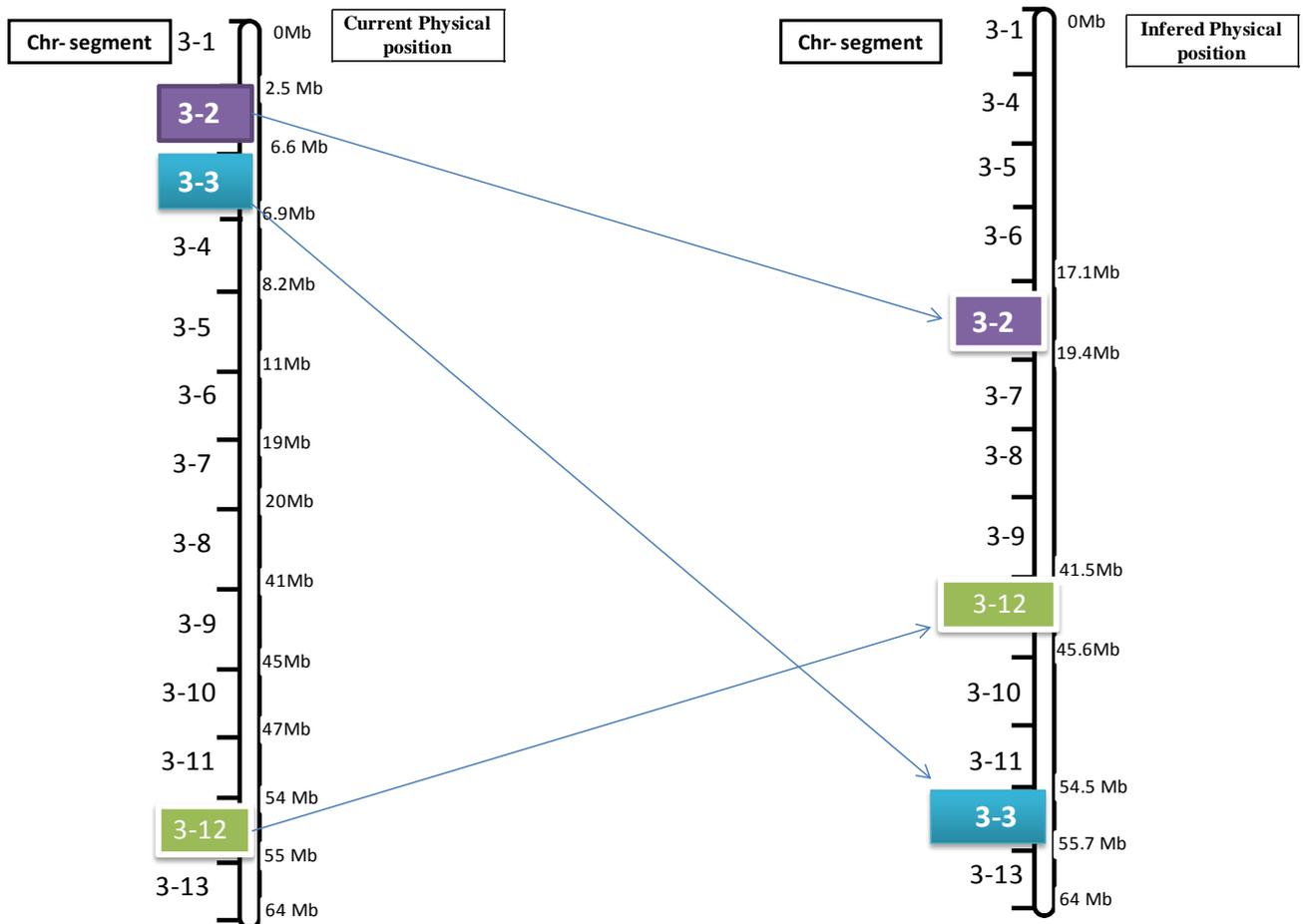
**Annex 2-6. Primers designed to genotype the BC<sub>3</sub> and subsequent generations by High Resolution Melting to develop the introgression lines. The genetic (cM) and physical position (pb) of each locus according to (<http://solgenomics.net/>) is also shown.**

Locus_Name	Chr.	Physical Genetic	Physical position	Forward	Reverse
SC_snp_sl_24799	1	10.94	669736	GCACAGAAAAGAAGCCCAAATC	GAAATGCTGAAGAAAGAGGAAGAAAT
SC_snp_sl_60089	1	39.28	2879206	TGATACATAGGCAAGAACACAAAA	TTGTTGGCACTCGGATATAAT
SC_snp_sl_50373	1	75.24	70221378	TCCCCCTACACTGTTTTACGTTATT	AGGAATGTAGTTTTTGCGCCA
SC_snp_sl_17080	1	101.76	74643317	CAGTTTTTGCCCTGGTTTGG	TGTTTCATCGGATGGTCTATTTC
SC_snp_sl_28168	1	127.29	76600314	CATTACCATTAGACCAACCTTAGGC	CCGACTTCTCAGAGGTTTATATGGAA
SC_snp_sl_43902	1	131.29	80289726	GAGGTTAAGTTGCTGAATAGGCAAT	AGTGAATCTTTCCGCTTTTGG
SC_snp_sl_40076	1	145.22	83277965	CCCCTCTCTTCAACAATTCC	TGCAATTTCCACTACCACATAA
SC_snp_sl_26879	1	159.95	85236107	TGAAATGCTGAAAAAGATCAA	TCAGATCAGTCTTGTATGGACAA
SC_snp_sl_48121	1	187.93	88665777	GGCTTGGACGGTAGAGAATCAC	AAAGCATTTGGAAAGCAACTG
SC_snp_sl_6247	2	0.00	12014192	GGAATATCACCTTTCCGGG	ACCGGAGTTTTTGTGATCG
SC_snp_sl_26050	2	8.58	28962589	TTTCAACCAATCGCCCGA	CGTGAAGATGCACACATATATTTT
SC_snp_sl_33255	2	17.78	32182950	GTTGAGCAGAAAAATGAAAATCCT	GCATGGCTCAAGAGAGGGGA
SC_snp_sl_36373	2	21.14	32707329	TCCAAGTCAGTGTGGTGACG	GTATAACCCAGGCCTGGTGC
SC_snp_sl_35844	2	39.14	35296865	GGCACTGTTCCACAATACCA	CCTACTTCAATGGAGCAACGGA
SC_snp_sl_25414	2	43.79	36057989	ACGTTGAAATGTCTTAACATACCA	ACTGGGAAAAGAACCTCGAG
SC_snp_sl_29602	2	50.44	37554774	GCCGAAAATCGAGAATGC	TGCTGGGTTTTCCGGTTC
SC_snp_sl_29627	2	53.66	40270742	CGGAAAATGTCGATATGTGAA	CATCAGCAGAACAAAATCCGGT
SC_snp_sl_42346	2	99.08	43978187	TTGGAATGGAATTGACTTGAA	TCAATAGACTAGTCCCAAATAGTT
SC_snp_sl_67017	2	120.25	47026891	TGAATAAGCAGTCAGCAACGAG	TTTGTTTGCCGTATCCCTTGA
SC_snp_sl_67551	2	133.10	48395256	GGGTCCCAAATCCTATGGTC	TGCAAGAATCAGAACCAGAGT
SC_snp_sl_63048	3	0.25	71026	TCCAGACACATTTCCACTTTCAA	TGAACTTCAATGTCTTAAATGAAAT
SC_snp_sl_63273	3	10.59	1323026	CCACCTTCTCTGTCAGTGTCA	GAAAAGCGGTTAAGGGAGGC
SC_snp_sl_9929	3	23.44	2111779	GAAGACACACTCCGTTCTATTGA	AGATCAGTCCCTCGCCACC
SC_snp_sl_26315	3	49.20	9416813	TGGACCTGCCTCTGATGACA	CTCTGCATTTCTCCCTGTTC
SC_snp_sl_30015	3	55.94	16383260	AATTCATTACATCCCTCTTAAAC	TTTGTGAATCACTGTCAAATGTGTT
SC_snp_sl_35459	3	78.42	52829430	TACTACGCTTTGCCCGC	TGCAAGTTTTGCTAATTTATGGGTAC
SC_snp_sl_9382	3	98.91	57030473	TGTGAACCAAATCCCAAGA	GGCATAGACACATAAAGTGACGGT
SC_snp_sl_9115	3	119.42	60818362	GAGCTACTTCTGTCATACGTCCTTG	CGTCTCAACACAGATTCAGAAA
SC_snp_sl_34027	3	130.24	62347422	ACACGAGTCCAGTGATATTTTGG	GCTCAAATATCTGAAGGAAGCAATT
SC_snp_sl_33815	3	139.54	64135699	GTCTTTAGCGTACTCCCGTATAGGTT	TTGTCCAGCTCAAACGAATGT
SC_snp_sl_9849	4	4.67	983682	AGCGGCATTTTCAGTAATGG	TGTCTTTGAGTCTGAGCTTCCA
SC_snp_sl_63996	4	16.05	1707761	GCAGTCTGTTCCGAATGTACAAAAG	TTGCTATAAGGCAGCAGAAAATAAAA
SC_snp_sl_64142	4	40.75	2413535	GAGAAAGCTCGGAAAGCTTAGAA	GAAGTGGTTCGCCAAGG
SC_snp_sl_41648	4	49.63	5526423	CAAATAGAAAATCGACACTTTAGC	CAATTTCCCTTAAATCTCCTTCC
SC_snp_sl_42919	4	57.87	51444136	AGAGACTACTGTTTCTTTGGGCT	CTCACCCAAAACAAATGGGC
SC_snp_sl_43439	4	89.74	56793148	TGGCTGATGATTTTCACTGG	GCTTTTAGITTCGGTGAGCTATC
SC_snp_sl_47165	4	115.46	60552747	TGTTGTGTGAGATTCGGGTTTC	TGTCATTGATTCAACTTAAAAGCAA
SC_snp_sl_47457	4	122.36	61853302	AGATAAGAAGTATCGGCTAGGCAAG	TTAGGTGTGAAGTCTGTGTTGTAGA
SC_snp_sl_47590	4	123.25	62310794	TGTGGTCCAAGTGTGCTG	ACTGCCATCGTATCAGGC
SC_snp_sl_69219	5	0.00	108062	CCTTCTCAGGCTTTAAGCGATT	CAACACAATGGGAGGCAGATCT
SC_snp_sl_52785	5	13.01	1664053	GCAAAAAGTCATCATAGGAAAAACAAA	TCAAAAAGTCTCCTCGCGTAG
SC_snp_sl_49009	5	25.48	3012858	CGTTGGATTGGTTGAAATTG	CCAAGTACTACTACGCCAAAACA
SC_snp_sl_29482	5	41.90	4066280	GGATCTCTGCTTTGCTGAATACA	ATAAGCCTTAAATGCTCGGTGAA
SC_snp_sl_5086	5	65.10	7473704	TTTTCCATAGCTCTTAGCTGACCAT	TGTTGGTTGTAAAAGGTGAAAAATG

SC_snp_sl_37626	5	118.18	63550293	AACAAAACCTATTCACGATAGCATCG	GCTCTGTACCATTTCCATTATGTCAC
SC_snp_sl_26826	6	0.00	208038	TGAGGTTTCGAGTCTTTGTCAT	CATTGAGTTTCCAATTAACCTTTTG
SC_snp_sl_35208	6	9.50	1748371	CATTGTGACTTTGTGGGTTTCAA	CCAGACCCAACCTGTGGG
SC_snp_sl_25167	6	13.01	28994438	ACGCGCAGTAGCCAACCA	GTGGCTGGCATCTGTTTTTGT
SC_snp_sl_28367	6	20.00	30372712	TCTATTTCGTGTTCCTATTGTTGCA	CTTGTCTGTTTCCCGTAGTGA
SC_snp_sl_14458	6	48.34	35241488	TGGTTTGCAGTGTATCCGACA	TTAAGAGCATCAAAAAGCCAGAGT
SC_snp_sl_12748	6	65.10	36958528	TTCATTACCTTCAGAACGGATGTC	GATTACTGGCTCCTTTGTCACA
SC_snp_sl_57512	6	83.50	41454775	ACCTTGTCTCATCTGCATTTT	TGATGATCAAACCATTCGAGG
SC_snp_sl_57294	6	92.63	42683784	TTCTTAACGGCCATACCTCGA	TTGGTGGAAATGAAGAGTTGGAG
SC_snp_sl_57093	6	100.16	43961536	TGTTATTTGGACTTTGGTGAAGAGC	TTGTTTCGCCTGAGTTTCTTTG
SC_snp_sl_54214	6	107.92	45759965	CCACATACCTCTACTCTCGAACCTT	CCGACATCTTGTCTATCGAGCTC
SC_snp_sl_68592	7	2.23	538803	GAGCACAAAAGGCGTCAGT	CCAACAAAACTCCATAAAGTGG
SC_snp_sl_68398	7	16.25	1557601	TTTCCCAATTCAGTTTCATC	CATGGGTCAGTCAGCAGAGA
SC_snp_sl_11171	7	20.41	1890867	GGTGTGACTGGTACTGCTATGAGAA	CAATCACCTGGAGTACCTGCAA
SC_snp_sl_68261	7	23.35	2091612	TCCATGAATCTTTCCACAATATCA	TGATTTTTTCATGCACGCTCTG
SC_snp_sl_68166	7	25.17	2354578	TTCAGACATTGATGACTTATTGCC	GTTGCCTGCATTTAGACAAAATA
SC_snp_sl_68143	7	26.06	2439173	AAGCGATGGGATCGAAGGA	CTTCTTCTCCGCATACAGCG
SC_snp_sl_52221	7	47.52	40553217	TTTGCCCGACTGTTTCTAG	TGCATACAGAACTCCAACAGG
SC_snp_sl_53591	7	63.97	58140668	CAAGACAGTGGATGAGAGGAACAC	GCGTATTACTCAGTTATCGTTTGGG
SC_snp_sl_53393	7	73.02	59572686	CTGTTGTGGGTTGGATGAGAAAG	AGACTATTTCTGAAACACCCCTCTGCT
SC_snp_sl_53237	7	79.74	60502871	CAATCAACAAAACCAAGAAGCTACTT	GGAGAGAGCTAGACGAGTCGAAA
SC_snp_sl_12149	7	93.25	62303494	ICTTTTCCACTTGCTCTTAGCCA	GCAGAGGATACCTGTTGIGGG
SC_snp_sl_70595	7	110.37	64536239	AAAGAGCAAGGACGCCCTTCA	GGATGAGATGATAAAGGCTGGTAGAC
SC_snp_sl_56404	8	0.44	269765	GCAAAATCACAACAATGGAAGA	GGTCAACCACGAATGTCCTCG
SC_snp_sl_56631	8	15.75	1403177	CGTCTAAGGAGAAACGTAGAGCTTGT	TATTAGCCCTCAAATTAGACTGCGC
SC_snp_sl_56791	8	28.54	2580721	TTGCCATTAGGGTTTGAATG	AATCCACCACCAAAATCTTTC
SC_snp_sl_5428	8	42.53	3261185	CGTAGGCGTTGGTAAAGGAGG	TGAGGGTCCAGGGTGAACA
SC_snp_sl_51919	8	50.75	3681415	TCTTACCCACCCTCTGTA	GTGAAAATGATGCGTGTGG
SC_snp_sl_46508	8	55.62	49588810	CCTCTTTGGAAGCGGCAA	TTGGAACGTGTTGAGAAGCAGG
SC_snp_sl_48469	8	71.70	54379535	GAAACCGACTACGAGACCATG	CAACAGTCCATCTTTTAGCCATC
SC_snp_sl_29413	8	74.95	54785958	GGATACACCTTTACCTTGTGATG	CATATTTTCTACACCTCTCATGG
SC_snp_sl_64736	8	93.89	57711488	CAAGTACTGGAGATTAGTACTGCTA	GGCAGCAGCAAAGGCTTATT
SC_snp_sl_65262	8	115.60	60436829	TCGATCCTTTCCGGTTTATCA	TCTTATATCTCTTAGCACCAATGCC
SC_snp_sl_65541	8	128.91	62135911	GGCGGGTGTACAATGAGATAGT	TTGTTTCGGTAAGGCAAAGTCC
SC_snp_sl_17481	9	0.66	629911	TGAGAAAACGAAAGCCTCGTC	TGAGAAAACGAAAGCCTCGTC
SC_snp_sl_39886	9	16.27	2052812	ATGTTGAGTGCCAACGTAGCC	CAGTTTTGCTACTGTTTTTCCC
SC_snp_sl_39658	9	31.94	3742541	CGTGCCATTGTGTAACCTGAG	TGCCATATTCATCAGCAAGTG
SC_snp_sl_45076	9	48.45	5836955	TTCCAGCCTCAGGCAATT	TGAATGATAGCACCAGACCAACTAA
SC_snp_sl_53165	9	54.35	22093677	GAGGAGGAATTTGAGAAGGCTAAA	TTGGTGAAGAACTTGATTGAATG

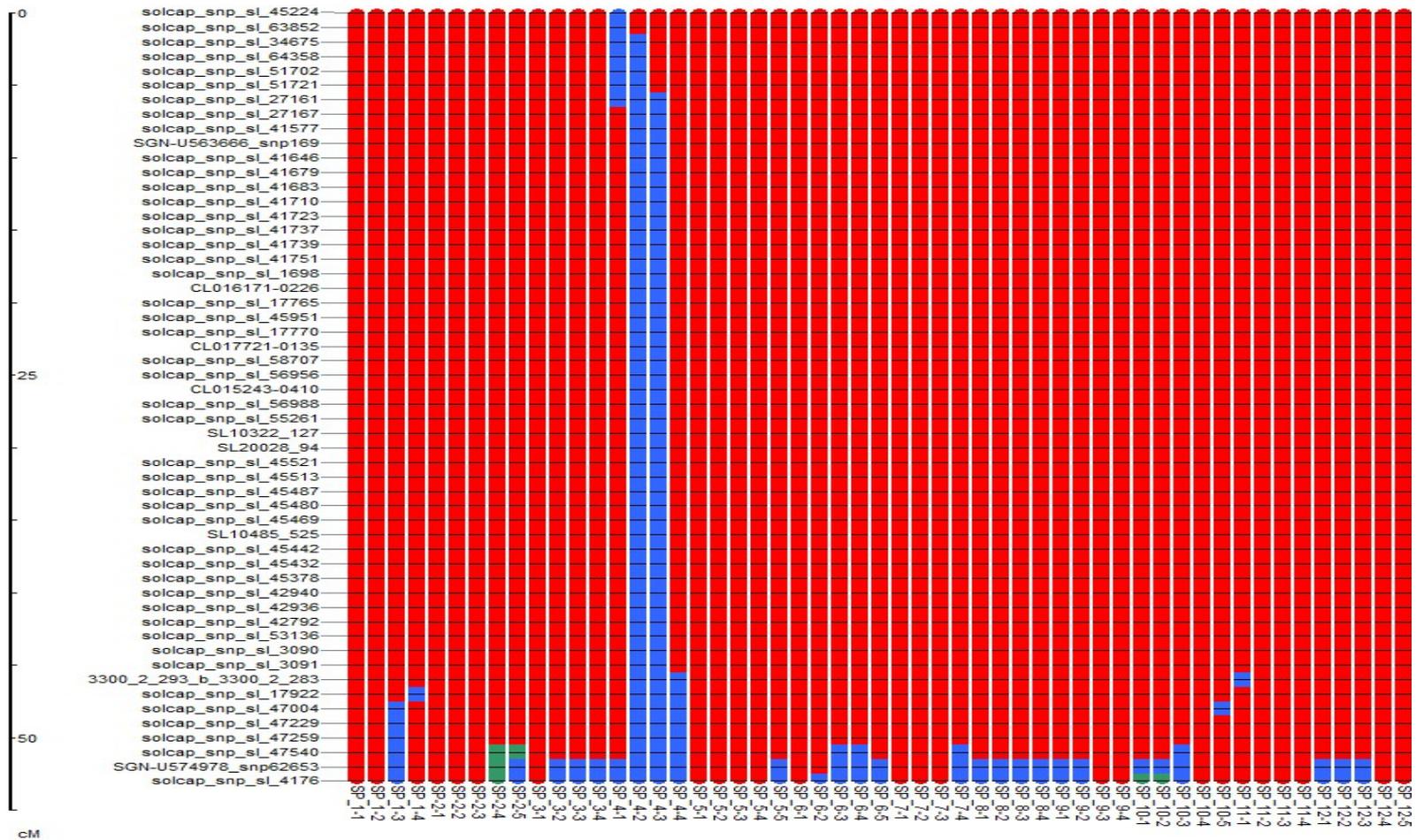
SC_snp_sl_46701	9	67.88	61995687	AGCCCTCCGATACCTTTTCC	ATGGCAATCCCCTGTCCA
SC_snp_sl_58299	9	92.13	64214465	AGGAGGAGTAAATGATTGGTTGTACA	TACCAAACCTCAATTTCAAACCAC
SC_snp_sl_69496	9	97.57	64900266	TCCACTTTGAGTTGATTGATAACAT	TCGTCCAGTACACACCTCCA
SC_snp_sl_11699	9	115.34	66031462	TGGGCTCGCTCATCTACCC	TTGTGTCCCCACTGCGG
SC_snp_sl_69874	9	122.50	66641208	ATATCAGCAGCCGTTCTCTCT	TGGAGAACTGGGCTTACATTTATT
SC_snp_sl_45992	10	0.00	49075	ACCTTTTATAGCCCATTCGGACT	CCAAAGCAGAGCAACCAATAAGT
SC_snp_sl_17839	10	15.08	1732945	TTAATAGCTTACC GGCTACTGC	AAATGCAATGGAAGGAGAGTTGAT
SC_snp_sl_46475	10	28.37	2417069	TAATATGACAATTTGCGACACCA	CAGTGATGAATGAAAAGAATCACA
SC_snp_sl_32616	10	50.22	42608313	GGGTCAACGGTAGAACTCCTCAT	GGCCGGTACAGCAAACATGT
SC_snp_sl_28675	10	60.08	59267524	CGACAGAGCATAAGGAGATAGTTTCC	TGGACCTTAGAGCATCTCAGC
SC_snp_sl_59437	10	84.60	61947683	CACTGATGCACAGAGACAACATCA	TGTGACGTGCCAAAGATTAGGAT
SC_snp_sl_61037	10	96.09	63051047	TGTTATGTTGGGAATGGAGGG	GCAACAAATGGAAAATCTTCTTG
SC_snp_sl_8807	10	111.91	64312177	CGGCGAAAACAGACAACAT	AACCCACATTGATTCCACCA
SC_snp_sl_66678	11	0.00	138728	CAAGTGCCCGACATGCTAT	ACGCTAGTTACCTGCAAAGAAGC
SC_snp_sl_66221	11	22.72	2593750	CAAAGGAAAGATGGAGTCACCTG	GCAGTTTCTCAGACAGCCGTTT
SC_snp_sl_21767	11	28.03	3158673	AGGAGGAGCCGAAGAGCG	AACAACCTCTCATCCATCTTAGCTT
SC_snp_sl_20963	11	34.28	3652933	TGTTCTTTGGGTTTCATTAATGTTCTC	CCACAGCAGCACAACCTTTATCAT
SC_snp_sl_9480	11	46.41	4921098	AGGTGCTTTGCCATACCAAG	TTTTCCATTACTTTCTCTCGTG
SC_snp_sl_10893	11	59.67	45478757	AAAATAGCAGGGGAGGCTGT	TGGGAGTAGAAAAAGGCAGAG
SC_snp_sl_52997	11	76.75	49345350	GGGAACTAATACCTTCAAAGAACAAA	TTGGTTCTTGCAATCTCTCGG
SC_snp_sl_32093	11	84.93	50308149	AATTTTGGCTGAATCTTGGTTATC	ACCAACCTTAGGTATGTTCTAGCAAAA
SC_snp_sl_56161	11	96.09	51293292	CAAGAACAAGTGTCTTTTATTGTCCCT	TGACAGCCCAACCATTGTTACA
SC_snp_sl_28471	11	110.89	52667653	CTACTCGTAACATCAGTCTCCAAAATTT	ATAGAAGCCAGTGCAGACAAAGG
SC_snp_sl_17723	12	3.81	581566	TCAAAAATGATACCAAACCTAGCCTT	CCTTTGTCCACCTTTGAATCGA
SC_snp_sl_41031	12	28.32	2329881	TCTCTTTGCTTTCCCTGTTAATT	GGTGGTTGTCGTTGTATAAGACTTGT
SC_snp_sl_41222	12	37.95	3318670	CAAAGGTGGAATGCTACCGA	TGTGATGAAGGCAAAGGAGGT
SC_snp_sl_1573	12	47.61	4038762	GAGGCTGCTCAGAAAGACCTAAGA	TCAGCAGCAGTTCGAGAAGATTT
SC_snp_sl_52464	12	67.00	44167775	CCTCTCTTTTGGATGCCCTCTGC	CAGCGGATAGGGAAAGATGTGT
SC_snp_sl_42740	12	90.51	48364193	TTGTAGGATTAATGTTCTCTGACCG	CCCAATAGCTTTTCTCAGGC
SC_snp_sl_55530	12	92.76	61834330	GTATCGCCGAAAATGTGACC	CCACTCCGGCAAGCAATC
SC_snp_sl_55712	12	102.36	62766003	ACAAATCCGTTCCCTAAAATCTG	GGCATCCAAAATCCGCTG
SC_snp_sl_32706	12	112.74	63329227	TCTCGGTAATACGCTGCAACACA	CACCTCTTTTCAATCACCAGTCT
SC_snp_sl_31464	12	130.63	64775887	CCCCCGTTATTGTACACGTTTT	GAACGGCGAGGTTGAGACC

**Annex 2-7. Assembly of the chromosome segments of chromosome 3** according to the SL2.40 genome build on the left, and according to bin-mapping with the current introgression lines on the right.



## **2. Anexos (Capítulo 3)**

**Anexo 3-1. Genotipado grafico realizado en el cromosoma 4 de las líneas de introgresión desarrolladas a partir del silvestre *S. pimpinellifolium* acc. TO-937, en el fondo genético de la variedad "Moneymaker" (MM). En el lado izquierdo del grafico se indican los marcadores moleculares utilizados con la escala de su posición en cM y abajo las ILs. Los fragmentos de color rojo corresponden al fondo genético de la variedad cultivada MM, los sectores verdes definen las introgresiones heterocigotas y las azules a introgresiones homocigotas del parental silvestre TO-937.**



### Anexo 3-2- Genes eliminadores de gametos en plantas.

Cultivo	Nombre del gen	Observaciones
<i>S. lycopersicum</i>	Eliminador de gametos "Ge"	<i>Ge</i> es un locus en la región proximal de la heterocromatina del cromosoma 4 con tres alelos ( $Ge^p$ , $Ge^c$ , $Ge^n$ ). En heterocigotos $Ge^p/Ge^c$ , el polen y los óvulos $Ge^c$ son eliminados y $Ge^p$ se transmite en >95% del polen viable y en >85% de los óvulos viables. $Ge^n$ es neutral. Se encuentra en <i>S. lycopersicum</i> y en <i>S. pimpinellifolium</i> (Rick 1966, 1971).
<i>Nicotiana tabacum</i>	Eliminador de polen	En plantas de <i>N. tabacum</i> que llevan un cromosoma particular de <i>N. plumbaginifolia</i> como supernumerario, el polen que no lleva el supernumerario es eliminado; es transmitido a cerca del 95% de la progenie. La transmisión a través del óvulo es del 15% debido a la pérdida del supernumerario en la meiosis (Cameron y Moav 1957).
<i>Triticum</i>	<i>Ki</i>	En híbridos de <i>T. aestivum vulgare</i> var. Chinese Spring × var. Timstein, el polen que lleva el alelo <i>Ki</i> -Timstein es eliminado y <i>Ki</i> -Chinese Spring se transmite a >90% de la progenie. La transmisión es normal a través de los óvulos. Otras accesiones son neutras (Loegering and Sears 1963; ver también Nyquist 1962).
<i>Triticum</i>	<i>Gc2</i>	En plantas de <i>Triticum aestivum</i> que llevan cromosomas particulares de especies de <i>Aegilops</i> , bien como sustituciones o como adiciones (supernumerarios), tanto el polen como el ovulo que no lleva el cromosoma de <i>Aegilops</i> son eliminados por muchas rupturas en el cromosoma y el cromosoma de <i>Aegilops</i> se transmite a 98–100% de la progenie a través del polen o los óvulos. Posiblemente se deba a una desrepresión de elementos transponibles. (Finch et al. 1984; Endo 1990; Nasuda et al. 1998; Friebe et al. 2003).
<i>Oryza</i>	<i>S1, S2</i>	<i>O. glabberima</i> parece que está fijada por un gen polen killer <i>S1</i> (grupo de ligamiento 1, cromosoma 6), que actúa sobre el gen homólogo de <i>O. sativa</i> . Según sea el fondo genético, <i>S1</i> puede también eliminar óvulos. La recombinación se encuentra suprimida cerca de <i>S1</i> en el híbrido, quizás indicando una inversión. Otro locus en el grupo de ligamiento 8 actúa como un eliminador de polen. Finalmente, <i>O. sativa</i> parece que tiene fijado un eliminador de polen en un locus diferente, <i>S2</i> , que elimina el polen que lleva el homólogo de <i>O. glabberima</i> . El resultado es un alto grado de esterilidad del híbrido (Sano et al. 1979; Sano 1983).
<i>Agropyron / Lophopyrum</i>	<i>Sd-1</i>	Un factor de distorsión de la segregación <i>Sd-1</i> en el cromosoma 7 del híbrido <i>A. elongatum</i> y <i>L. ponticum</i> elimina los óvulos que no lo llevan en un fondo de trigo (Dvorak 1980; Scoles and Kibirge-Sebunya 1982; Dvorak y Appels 1986).

**Anexo 3-3 Genes implicados en autoincompatibilidad (siglas en inglés SI) y en incompatibilidad unilateral (siglas en inglés UI) interespecífica en Solanaceas.**

<i>Nombre del locus o loci</i>	<b>Observaciones</b>
<i>S-locus</i>	Gen determinante de la especificidad de la SI por parte del esporofito. El sistema de auto-incompatibilidad en <i>Solanum</i> SI está controlado normalmente por este locus <i>S</i> . Codifica la expresión de proteínas con actividad S-RNAasa en el pistilo, las cuales controlan la especificidad con la que el pistilo rechaza al polen. Se expresa en estilo, estigma y ovario
<i>S-locus F box</i>	La proteína producto de este locus es una F-box, específica para cada locus <i>S</i> (SLF) y que se expresa en el polen y confiere especificidad a la interacción con cada RNasa del pistilo, sea de compatibilidad o de incompatibilidad (Iwano y Takayama 2012). El polen es rechazado solo cuando el haplotipo-S del polen es idéntico a cualquiera de los dos haplotipos-S del pistilo diploide.
<i>SI-modifier genes sporofito</i> <i>HT GENES</i> <i>120 kDa GlyProt</i> <i>Inhibidor proteasa</i> <i>NaStep</i>	Existen genes para proteínas modificadores de la SI, que se requieren para las funciones SI del polen / pistilo pero que no contribuyen a la especificidad del sistema S-per se (McClure et al. 1999; Tsukamoto et al, 2003; Hancock et al. 2005; Hua y Kao 2006; Puerta et al. 2009; Zhao et al. 2010; Jimenez-Duran et al. 2013). Por la parte del pistilo se han identificados una serie de proteínas de los genes modificadores que incluyen: Las proteínas HT ricas en Asn/Asp y de expresión tardía en el estilo, la glicoproteína de 120 kDa arabinogalactano glycoprotein que se acumula en la ECM del estilo y es importada por el tubo polínico y el inhibidor de proteasas <i>NaStep</i> . Casi todos cuales se han implicado directamente en el rechazo del polen especificado por el tipo de haplotipo S mediante experimentos de RNAi o antisense en <i>Nicotiana</i> (McClure et al. 1999; O'Brien et al. 2002; Hancock et al. 2005; Jiménez -Duran et al. 2013). Se han propuesto dos modelos: 1) degradación de S-RNAsas (Kubo et al. 2010) y 2) compartimentalización en el que HT (dos parálogos A y B en <i>Sol</i> ) y factor 4936 actúan estabilizando la SRNAasa en la vacuola y previniendo su actuación (Goldraj et al. 2006)
<i>SI-modifier genes En Polen: Skp1Cullin1 homolog S-Rnase BP</i>	Como los genes SLF expresados en el polen probablemente funcionen como parte de un complejo SCF ubiquitin ligasa, los genes modificadores del polen incluyen al menos los que codifican un homólogo de Skp1 (Zhao et al. 2010), una Cullin1 (Hua y Kao. 2006; Li y Chetelat. 2014) y una proteína de unión a S-RNasa (Hua y Kao 2006)
<i>UI locus</i>	Barrera de polinización inter-específica que actúa a nivel de especie y que por lo tanto los sistemas de reconocimiento y rechazo son inherentemente menos específicos que los SI. Sin embargo, la evidencia indica que algunas barreras de polinización interespecíficas están relacionadas con al SI. La incongruencia/ incompatibilidad unilateral (UI) es una relación interespecífica según la cual las polinizaciones solo son compatibles en una dirección que normalmente siguen la regla SI x SC.
<i>Ui1.1</i> <i>S-locus</i>	Estudios genéticos demuestran que los QTLs de mayor efecto sobre la UI tanto en pistilo como en polen coinciden con el locus <i>S</i> (Chetelat y Deverna 1991; Bernacchi y Tanksley 1997).
<i>Ui6.1</i> <i>Cullin1</i>	El gen <i>ui6.1</i> codifica una proteína tipo Cullin1 que se expresa en el polen y que es similar a la proteína de <i>Petunia</i> que ha sido implicada en la SI (Hua y Kao 2006; Li y Chetelat 2010). El producto de <i>ui6.1</i> solo funciona cuando el gen se expresa en conjunción con <i>ui1.1</i> , un QTL para UI que se localiza en el locus <i>S</i> (Li y Chetelat 2010; Li et al. 2010).
<i>HT</i>	Un QTL adicional que contribuye a la parte de la UI que depende del pistilo (Bernacchi y Tanksley 1997) incluye al locus del gene <i>HT HTA</i> (Covey et al. 2010),

### Anexo 3-4 Ejemplos de genes implicados en crecimiento del polen y guía hacia el ovulo incluidos los que proporcionan fertilización diferencial.

Cultivo	Nombre del gen/proteína	Observaciones
Maíz	<i>Gal-10</i>	Hasta 10 factores gametofíticos se han descrito en maíz que favorecen la fertilización cuando el alelo del polen se corresponde con el del pistilo (Kermicle y Evans, 2005), favoreciendo de esta forma la autopolinización (Nelson 1994).
Maíz	<i>Ga-1</i>	El factor gametofítico 1 ( <i>gal</i> ) de maíz identificado por su efecto en distorsionar la segregación del locus ligado <i>sugary 1</i> en cruces entre maíz dulce ( <i>gal gal sul sul</i> ) y de palomitas ( <i>Gal Gal Sul Sul</i> ) cuando se realizaban retrocruces con palomitas como parental femenino (Mangelsdorf y Jones 1926; Nelson 1994). El alelo <i>Gal-S</i> causa esterilidad no-recíproca; una planta <i>Gal-S/Gal-S</i> no es receptiva al polen <i>gal</i> ; sin embargo, el polen <i>Gal-S</i> induce un llenado de grano completo en la mazorca de plantas homocigotas <i>gal / gal</i> . Un tercer alelo, <i>Gal-m</i> , desde el punto de vista de la compatibilidad, se ha definido como alelo de acción neutral dado que fertiliza todos los genotipos incluyendo a <i>Gal-S/ Gal-S</i> y las plantas homocigotas para este alelo son receptivas al polen con cualquiera de los tres alelos del locus <i>gal</i> (Kermicle y Evans 2010; Schwartz 1950).
Maíz	<i>Phytosulphokine</i>	La germinación y crecimiento del tubo polínico están autoregulados por diferentes péptidos producidos por el polen. Una serie de fitosulfokinas son responsables de ejercer un efecto estimulador sobre la población de polen durante la germinación del tubo polínico (Chen et al. 2000).
Maíz	<i>LAT52</i>	La proteína pequeña y rica en Cys <i>LAT52</i> es necesaria para la hidratación y germinación normal del polen (Muschietti et al. 1994). En estos dos procesos los péptidos actúan como factores “autocrinos” para regular la germinación del tubo polínico (Tang et al. 2002; 2004; Johnson y Preuss 2003).
Tomate	<i>SIPRALF</i>	<i>SIPRALF</i> , es un péptido específico del polen de tomate que no afecta su viabilidad, hidratación o etapas iniciales de la germinación, pero sí que inhibe la elongación del tubo polínico durante una ventana específica del desarrollo. Posee un motivo ISY cerca del N terminal, un motivo GASYY cerca del primer y segundo motivo de Cys conservados, y un motivo YXRGCS que incluye el tercer residuo conservado de Cys (Covey et al. 2010).
Brasicas	<i>SCR/SP11</i>	El péptido <i>SP11</i> interacciona con el <i>SCR</i> de Brassicaceae es el determinante masculino del reconocimiento “self/nonself” durante la polinización (Schopfer et al. 1999 y Takayama 2000).
	<i>LURE peptides</i>	Los péptidos <i>LURE</i> de <i>T. fournieri</i> (Linderniaceae, Lamiales) se expresan específicamente por las células sinérgicas (célula haploides que acompañan al huevo) para atraer los tubos polínicos (Okuda et al. 2009).
Maíz	<i>ZmES4</i>	El péptido de maíz producto de <i>EMBRYO SAC 4</i> ( <i>ZmES4</i> ) se expresa en las células sinérgicas, induce la ruptura del tubo polínico la liberación de las células espermáticas (Amien et al. 2010).
<i>Arabidopsis</i>	<i>AtLURE</i>	El clúster de genes <i>AtLURE1</i> agrupa una serie de genes específicos de especie cuyo productos son unos péptidos que se expresan en las sinérgicas y se secretan al camino que seguirán los tubos polínicos para alcanzar el gametofito femenino. Los péptidos <i>AtLURE1</i> atraen preferentemente los tubos polínicos de la especie y su represión desbarata el mecanismo de rastreo /guía del tubo polínico. La introducción de un único gene <i>AtLURE1</i> desde <i>A. thaliana</i> a plantas de <i>T. fournieri</i> , es suficiente para destruir las barreras de aislamiento en lo que respecta a la guía y penetración del tubo polínico (Takeuchi y Higashiyama 2012).

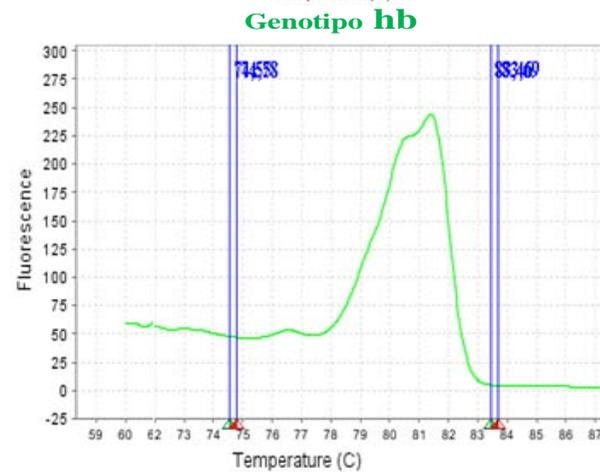
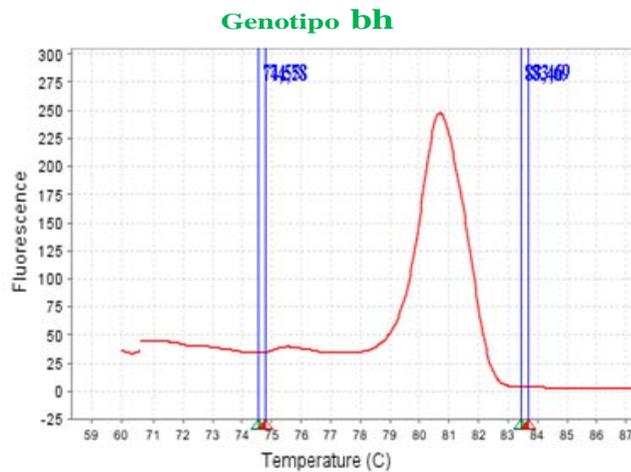
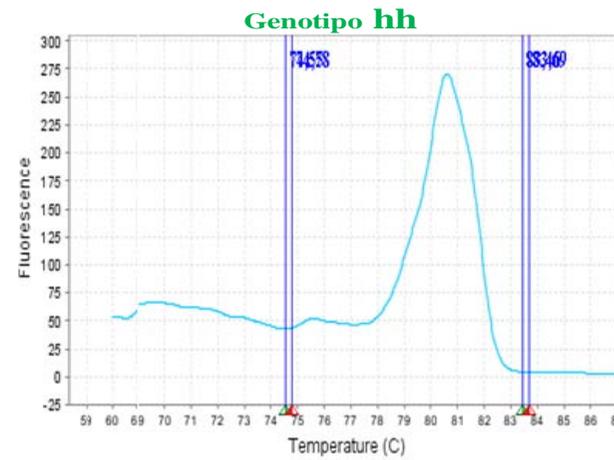
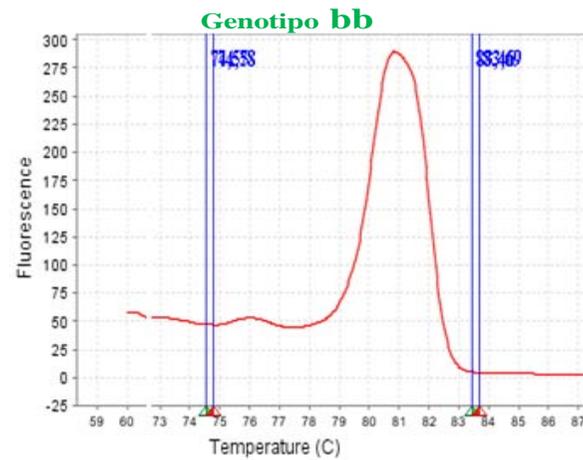
**Anexo 3-5. Marcadores SNP utilizados para el genotipado por medio de HRM de las plantas usadas en el mapeo fino de la región del cromosoma 4 que presenta SD.**

NÚMERO	NOMBRE OLIGO	SECUENCIA	FUENTE
1	SNP 47590R	ACTGCCATCGTGATCAGGC	MAPA/RILS
2	SNP 47590F	TGTGGCTCCAAGTGTGCTG	MAPA/RILS
3	SNP 335aR	TTGCACTCAAAGAAATGCCG	MAPA/RILS
4	SNP 335aF	AGCTGTCTTGCGATGTCAGGT	MAPA/RILS
5	Solcap_snp_sl_47050 F	CCAGTTTCCCTGTCCTTAGCAA	CHIP SOLCAP
6	Solcap_snp_sl_47050 R	CGAGAGATACGAGTGATGAGATGC	CHIP SOLCAP
7	Solcap_snp_sl_47165 F	GCTGACGATGCAAGAGATAGACA	CHIP SOLCAP
8	Solcap_snp_sl_47165 R	GTGGTGGTTTGCCTTGTGAA	CHIP SOLCAP
9	Solcap_snp_sl_3748 F	AACTTCCATGTCATGTCCCA	CHIP SOLCAP
10	Solcap_snp_sl_3748 R	CATCCTCGATTGTTTTCCGTT	CHIP SOLCAP
11	Solcap_snp_sl_47225 F	TTCTAGTGGTGGATGGGTCAGTC	CHIP SOLCAP
12	Solcap_snp_sl_47225 R	AAGAAATCACCGGCGAAGATC	CHIP SOLCAP
13	Solcap_snp_sl_47229 F	TGGCAATGAAATTAGGCAGGT	CHIP SOLCAP
14	Solcap_snp_sl_47229 R	CAAGCCACCAGATTCCTTCC	CHIP SOLCAP
15	Solcap_snp_sl_47397 F	TGGAATACGACTTACTCAGCGAGA	CHIP SOLCAP
16	Solcap_snp_sl_47397 R	GAAGATGCAACAGGCGTGGTA	CHIP SOLCAP
17	Solcap_snp_sl_47413 F	AGGGTGTGGTTAATATGAAGGGTTAT	CHIP SOLCAP
18	Solcap_snp_sl_47413 R	GAAGCCAAGCATCTTGAGGGT	CHIP SOLCAP
19	Solcap_snp_sl_47457 F	CCTTCGAGCTGCACGTCC	CHIP SOLCAP
20	Solcap_snp_sl_47457 R	CCATCACATGCTCGTAGAGCC	CHIP SOLCAP
21	Solcap_snp_sl_47656 F	GCGTATGTAAAGGGAAGTTGTCTCT	CHIP SOLCAP
22	Solcap_snp_sl_47656 R	CGTAGATCCACCAGCAAATTCA	CHIP SOLCAP
23	Solcap_snp_sl_3827 F	CCTGAACATCTGAAAAACATACACG	CHIP SOLCAP
24	Solcap_snp_sl_3827 R	ATTCTAAAAGCTCATCTCGCCTA	CHIP SOLCAP
25	Solcap_snp_sl_47694 F	TGGCGAGCACTAAATCCCA	CHIP SOLCAP
26	Solcap_snp_sl_47694 R	TTTCAGCCAAGCTGTTTGCC	CHIP SOLCAP
27	Solcap_snp_sl_47697 F	GGCAAACAGCTTGGCTGAAA	CHIP SOLCAP
28	Solcap_snp_sl_47697 R	TTGGCCTGAAGACGGTTGA	CHIP SOLCAP
29	Solcap_snp_sl_3920F	GCCATCAACGATCAACAACCTCTAT	CHIP SOLCAP
30	Solcap_snp_sl_3920R	ACGTGGTGGTTGAAGATGCA	CHIP SOLCAP
31	Solcap_snp_sl_47766F	GCAAAACCTTCATCCCCAAA	CHIP SOLCAP
32	Solcap_snp_sl_47766R	GCCACTGCTAAATATGCCCAAT	CHIP SOLCAP
33	Solcap_snp_sl_4003F	TCAAACAGGGTGGTGCTCTG	CHIP SOLCAP
34	Solcap_snp_sl_4003R	AATCACTCGAAAAAATCATCCAA	CHIP SOLCAP
35	Solcap_snp_sl_23589F	CACACTAACCTTTGGCAATCCC	CHIP SOLCAP
36	Solcap_snp_sl_23589R	TTCAAGGCAGCTCAGGCG	CHIP SOLCAP
37	Solcap_snp_sl_4164F	CGGAGAGTAGTTGGAGGAACACC	CHIP SOLCAP
38	Solcap_snp_sl_4164R	GAGAAGCTGGAGAAGGTTGGG	CHIP SOLCAP
39	Solcap_snp_sl_47742F	CACTCCAAAACATCGTCAGCTC	CHIP SOLCAP
40	Solcap_snp_sl_47742R	ACTGTTGCATTGCTGTTGGC	CHIP SOLCAP
41	Solcap_snp_sl_3952F	TCTTCTTCGTGCAACCTTCTG	CHIP SOLCAP
42	Solcap_snp_sl_3952R	CGATATAAAGTTGAAGGAGATCGGTC	CHIP SOLCAP

43	Solcap_snp_sl_47749F	TTCTGGGTGCTGTCGGATG	CHIP SOLCAP
44	Solcap_snp_sl_47749R	TGGCAAGGCCACTCATGA	CHIP SOLCAP
45	Solcap_snp_sl_3975F	TTTTGATGTTTTCCGTCAATGG	CHIP SOLCAP
46	Solcap_snp_sl_3975R	GATGCACGAAACAATGTACGAGA	CHIP SOLCAP
47	Gsnp_23169F	AAGGGCCTTACGGATGAGTT	RNA-sec/TO-397
48	Gsnp_23169R	TGAATTTTAACACTGATTGTGGTG	RNA-sec/TO-397
49	Gsnp_33060F	TTCATTCTGACTGGTGTGGAA	RNA-sec/TO-397
50	Gsnp_33060R	TCATCATGAACTTATCACCCAAA	RNA-sec/TO-397
51	Gsnp_67132F	ACACCCTAGGAACAAAGATG	RNA-sec/TO-397
52	Gsnp_67132R	CGACCTTGCCCAATATCAT	RNA-sec/TO-397
53	Gsnp_83668F	TGATTTTTAATTTTATCCACCTTATTC	RNA-sec/TO-397
54	Gsnp_83668R	AGTGTTAGGCACCACAATCAT	RNA-sec/TO-397
55	Gsnp_94601F	CATGTGCTGCTTTGCTTGAT	RNA-sec/TO-397
56	Gsnp_94601R	CAATGGCAGCAACTGGATTA	RNA-sec/TO-397
57	Gsnp_99848F	CACGACAAATATTCCTTTTCCA	RNA-sec/TO-397
58	Gsnp_99848R	CCGTTTCGAGCAGAAAACGT	RNA-sec/TO-397
59	Gsnp_107940F	TTCCATGCTAACCACCTTTGATG	RNA-sec/TO-397
60	Gsnp_107940R	ACCACCAGGGTTTGAGTGAT	RNA-sec/TO-397
61	Gsnp_121747F	CAAGCTTTGTCTCCTGTAAGCA	RNA-sec/TO-397
62	Gsnp_121747R	GGGAGGGGAAAGGCTTTAG	RNA-sec/TO-397
63	Gsnp_134821F	TTCTTATCGCCTGTCCATTT	RNA-sec/TO-397
64	Gsnp_134821R	TGCACTTGACACCTTTTTTCG	RNA-sec/TO-397
65	Gsnp_151129F	TTGAATTGATGGACAATGAAAA	RNA-sec/TO-397
66	Gsnp_151129R	AATTTAGGACCCACCGAACC	RNA-sec/TO-397

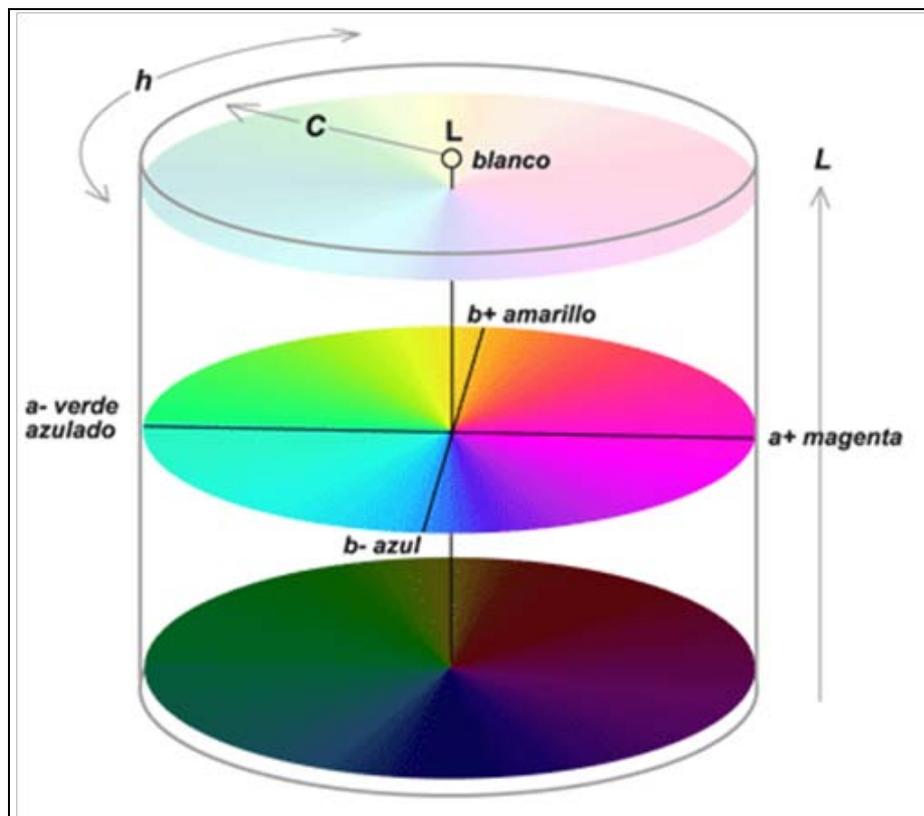
Anexo 3-6. Curvas de fusión utilizadas en la búsqueda de recombinantes, obtenidas a partir de los marcadores: Solcap\_snp\_sl\_47697 localizado en el extremo proximal y Solcap\_snp\_sl\_47749 localizado en el extremo distal de la región con SD del cromosoma cuatro. Genotipo bb significa plantas con la región con SD en condición homocigota para TO-937, bh y hb son plantas con genotipo recombinantes y hh plantas con genotipo heterocigoto para la región con SD.

□



### **3. Anexos (Capítulo 4)**

Anexo 4-1. Diagrama del espacio de color CIELAB utilizado en la medición del color externo e interno de los frutos.



**Anexo 4-2. Coeficiente de correlación ( $p < 0.005$ ,  $-0.13 > r > 0.13$ ) de los caracteres usados en el fenotipado de las líneas de introgresión en las tres localidades estudiadas: Alginet, Orihuela y Malaga. Los nombres de los caracteres son sustituido por sus respectivos acrónimos propuestos por Brewer et al. (2006) (Fig 4-1)**

	Dimensiones del fruto					Forme del fruto						Químicos			Color externo					Color interno				
	FW	PER	AR	FDL	FLL	FS	PBLK	DBLK	TRI	CIR	PSH	BRIX	Ph_i	TA	L*	a*	b*	C*	H*	L*	a*	b*	C*	
PER	0,93																							
AR	0,93	0,99																						
FDL	0,93	0,98	0,98																					
FLL	0,88	0,95	0,96	0,90																				
FS	-0,44	-0,42	-0,39	-0,56	-0,15																			
PBLK	0,18	0,18	0,15	0,21		-0,39																		
DBLK		-0,14	-0,14		-0,23	-0,20	-0,30																	
TRI	0,17	0,20	0,17	0,19	0,16		0,80	-0,80																
CIR	0,45	0,44	0,39	0,55	0,17	-0,92	0,50		0,26															
PSH	0,30	0,31	0,28	0,33	0,22	-0,32	0,32		0,26	0,44														
BRIX								0,14																
Ph_i												0,21												
TA												0,30	-0,32											
L*	0,13	0,15	0,16	0,15	0,16							-0,36	-0,18											
a*	0,18	0,16	0,18	0,15	0,19							-0,26			0,42									
b*	0,17	0,19	0,20	0,18	0,21							-0,41	-0,19		0,97	0,52								
C*	0,20	0,19	0,22	0,19	0,23							-0,38	-0,19		0,81	0,86	0,88							
H*															0,50	-0,55	0,42							
L*						-0,16	0,26		0,18	0,22	0,20		-0,21	0,27	0,35	0,30	0,34	0,37						
a*															0,37		0,24	-0,35	0,14					
b*							0,17			0,05	0,09		-0,17		0,21	0,24	0,19	0,25		0,66	0,52			
C*							0,08								0,36	0,13	0,28	-0,27	0,41	0,91	0,83			
H*	-0,14	-0,14	-0,16		-0,18		0,13					0,13			0,15	-0,23			0,34	0,39	-0,70	0,24	-0,33	

Localidad: Alginet

## Continúa anexo 4-2

	Dimensiones del fruto					Forme del fruto						Químicos			Color externo					Color interno				
	FW	PER	AR	FDL	FLL	FS	PBLK	DBLK	TRI	CIR	PSH	BRX	Ph_i	TA	L*	a*	b*	C*	H*	L*	a*	b*	C*	
PER	0,54																							
AR	0,55	0,98																						
FDL	0,54	0,95	0,96																					
FLL	0,49	0,90	0,92	0,80																				
FS				-0,36	0,27																			
PBLK					-0,24	-0,53																		
DBLK					-0,27	-0,57	0,40																	
TRI						-0,20	0,81	-0,20																
CIR	0,18	0,32	0,27	0,48		-0,79	0,21	0,30																
PSH	0,22	0,40	0,36	0,46	0,20	-0,42	0,20	0,18		0,57														
BRX								0,14																
Ph_i						0,18				-0,24	-0,20													
TA			0,13	0,17				0,16		0,25	0,26	0,45	-0,26											
L*	0,19	0,25	0,24	0,21	0,26								-0,28		-0,15									
a*													-0,22	-0,30										
b*	0,17	0,22	0,22	0,18	0,25			-0,19	0,13				-0,39		-0,17	0,89	0,16							
C*	0,12	0,13	0,13	0,13	0,13				0,12				-0,40	-0,25		0,63	0,73	0,78						
H*	0,10	0,17	0,17	0,12	0,23	0,16		-0,16					0,19	-0,17	0,62	-0,68	0,61							
L*	0,22	0,26	0,27	0,23	0,27						0,17				0,18		0,17			0,21				
a*													-0,15		0,41		0,21	-0,37	-0,39					
b*				-0,15		0,13				-0,16					0,16		0,16							
C*															0,37		0,25	-0,29	-0,21	0,68	0,78			
H*				-0,15		0,14									-0,28			0,29	0,37	-0,85	0,44	-0,21		

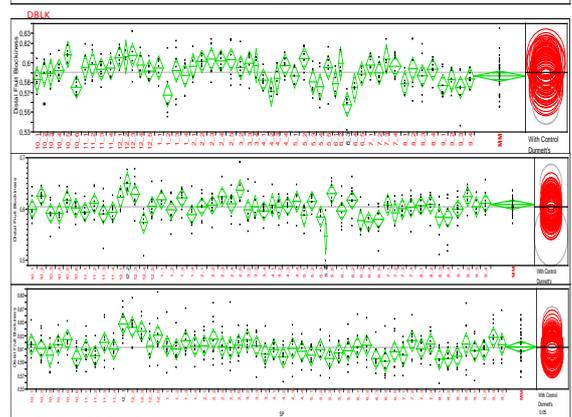
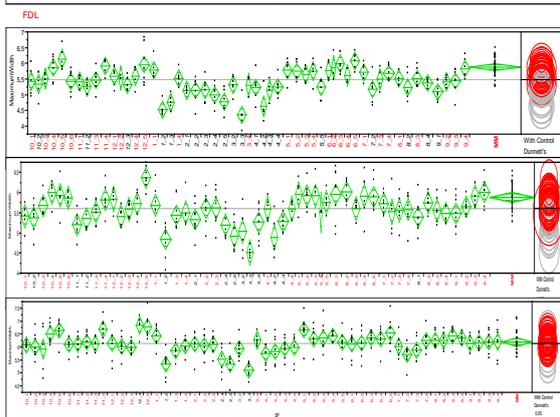
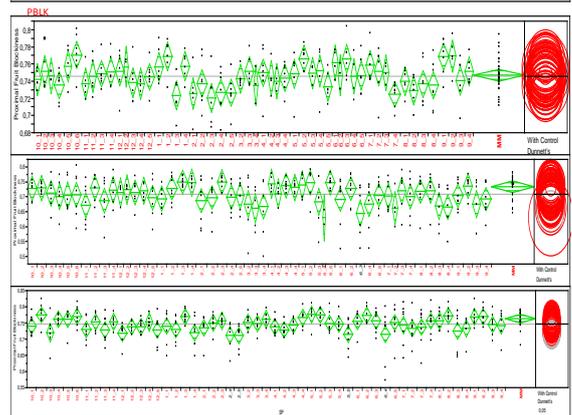
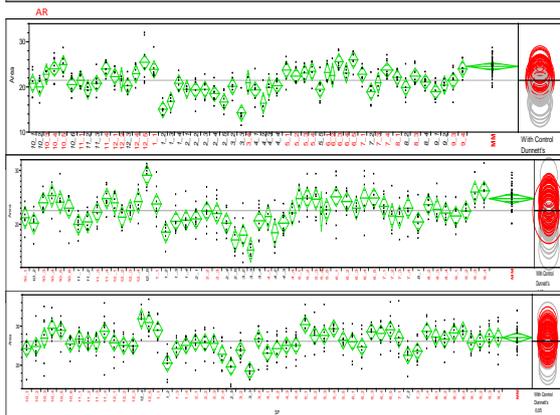
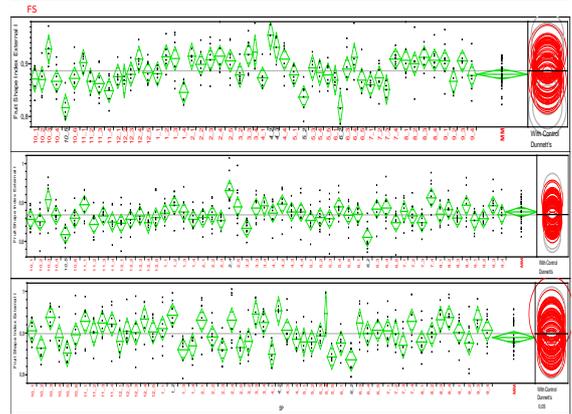
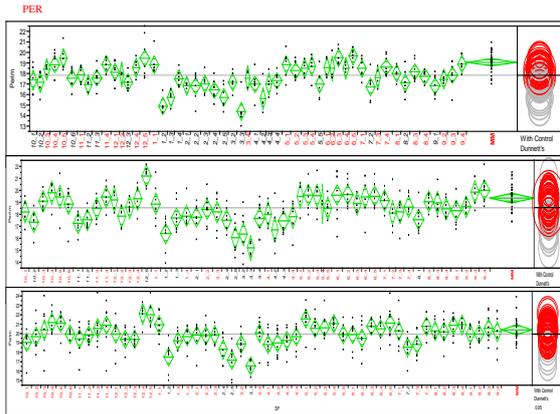
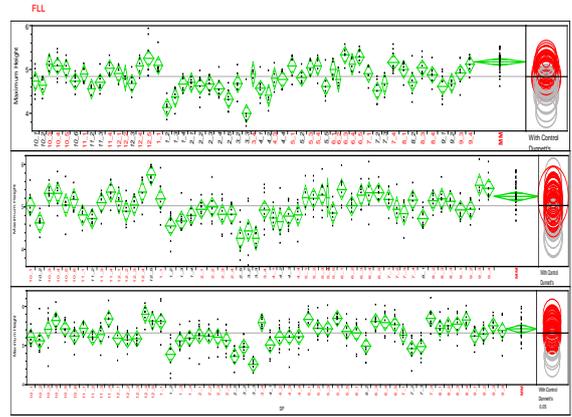
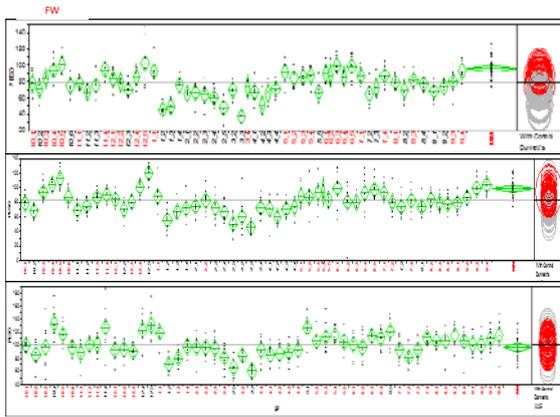
Localidad: Orihuela

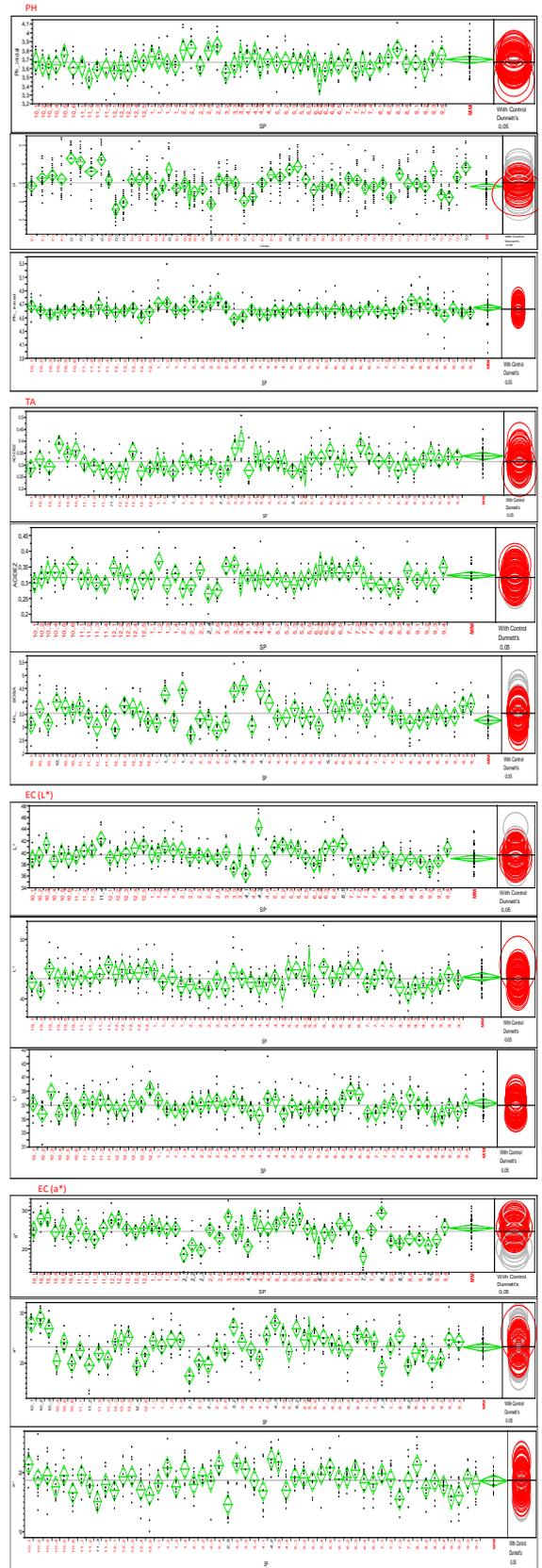
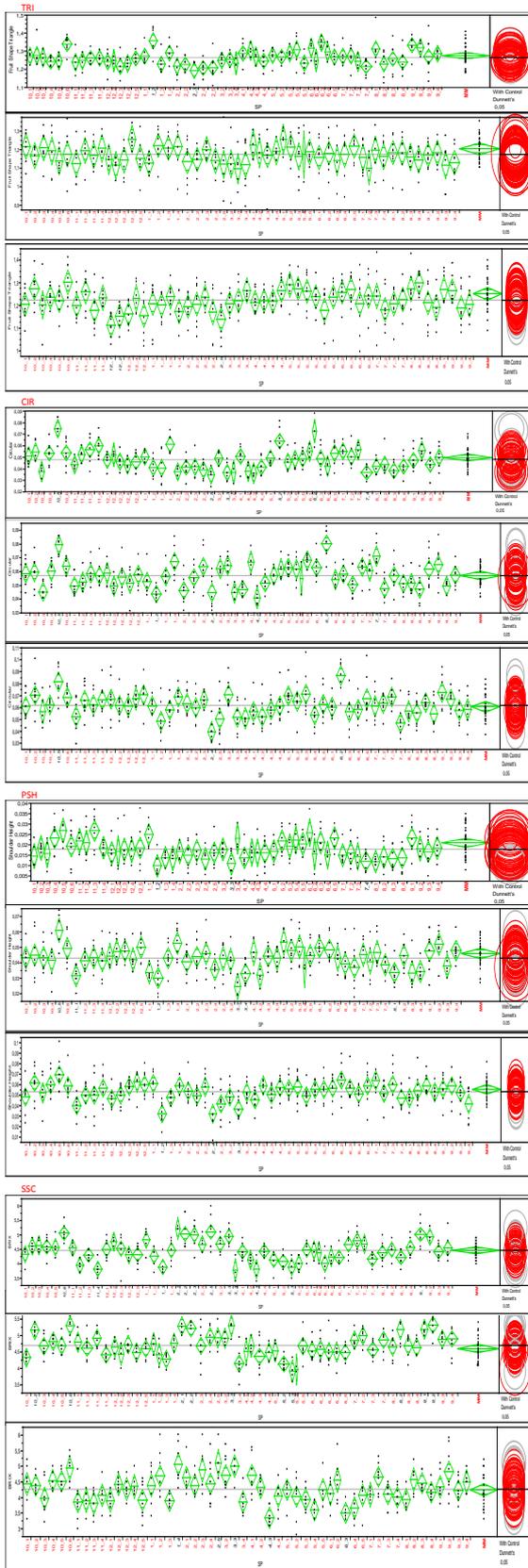
## Continua anexo 4-2

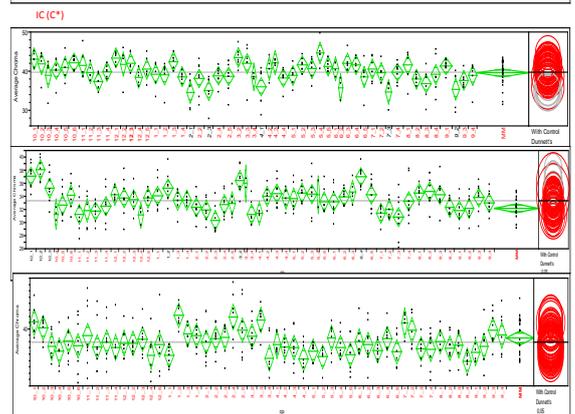
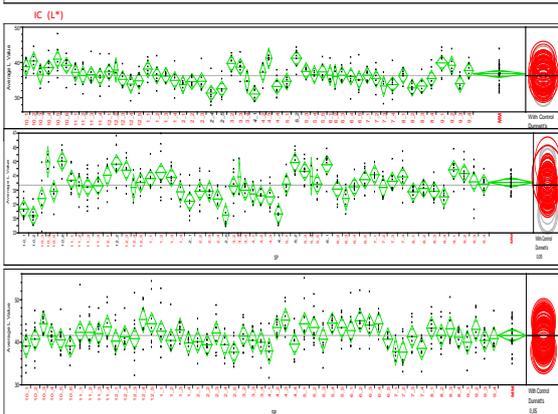
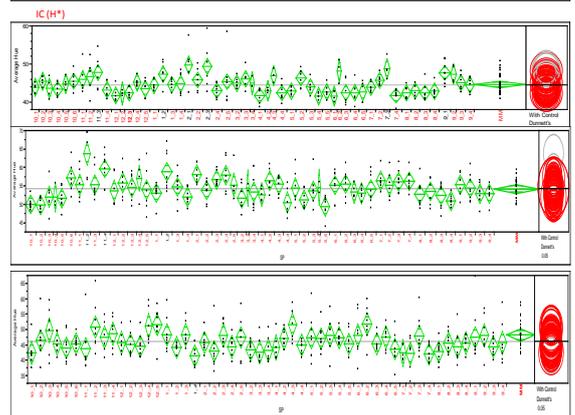
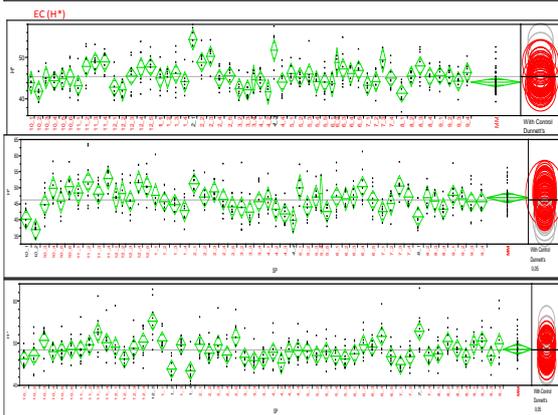
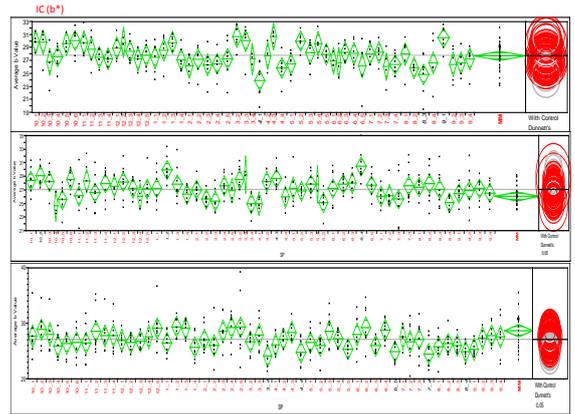
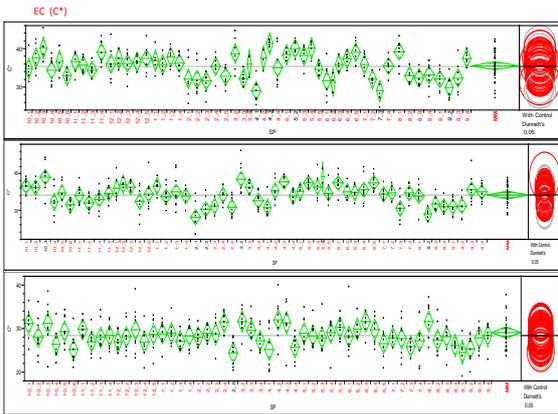
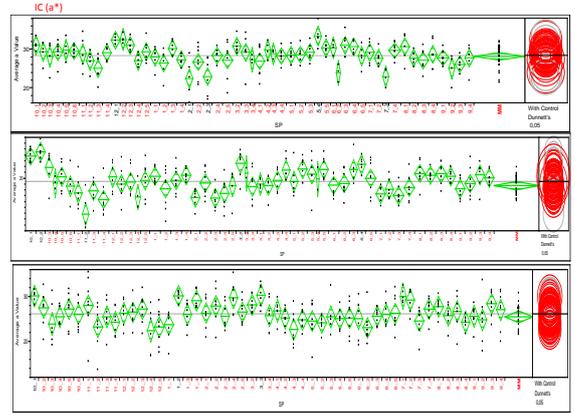
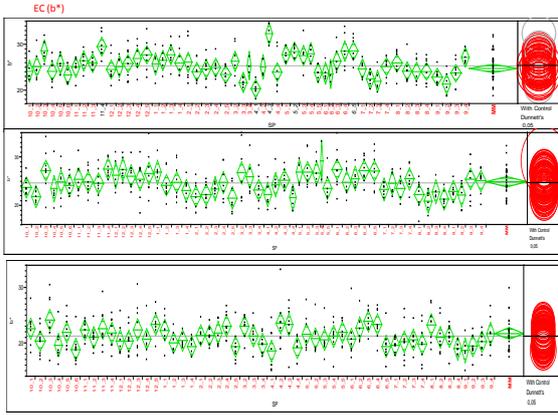
	Dimensiones del fruto					Forme del fruto						Quimicos			Color externo					Color interno					
	FW	PER	AR	FDL	FLL	FS	PBLK	DBLK	TRI	CIR	PSH	BRIX	Ph_i	TA	L*	a*	b*	C*	H*	L*	a*	b*	C*		
PER	0,85																								
AR	0,86	0,99																							
FDL	0,85	0,97	0,97																						
FLL	0,80	0,91	0,94	0,83																					
FS	-0,31	-0,36	-0,31	-0,53																					
PBLK		0,16	0,14	0,17		-0,31																			
DBLK					-0,19	-0,18																			
TRI	0,13	0,15	0,14	0,16		-0,14	0,81	-0,61																	
CIR	0,36	0,43	0,36	0,57		-0,91	0,29	0,14	0,14																
PSH	0,44	0,50	0,45	0,52	0,35	-0,42	0,24		0,19	0,56															
BRIX	-0,18	-0,22	-0,24	-0,20	-0,27																				
Ph_i												0,15													
TA												0,38	-0,40												
L*												-0,17													
a*													-0,14	0,22	0,41										
b*													0,11	-0,14		0,93	0,62								
C*														-0,16		0,78	0,87	0,92							
H*													0,14	-0,18		-0,28	0,46	-0,58	0,28						
L*	0,26	0,26	0,28	0,29	0,25	-0,15						0,17	-0,31	-0,26		0,30		0,30	0,24	0,17					
a*	-0,18	-0,17	-0,16	-0,18	-0,13	0,14						-0,19	-0,19	0,28	0,22		0,21			-0,34	-0,61				
b*	-0,34	-0,32	-0,33	-0,33	-0,32											0,19		0,14		0,15	0,26				
C*	-0,37	-0,35	-0,35	-0,37	-0,33	0,18						-0,19	-0,29	0,16			0,17			-0,17	-0,31	0,73	0,60		
H*												-0,28	-0,21	-0,22	0,20	-0,16	0,15		0,35	0,63	-0,82	0,64	-0,21		

Localidad: Malaga

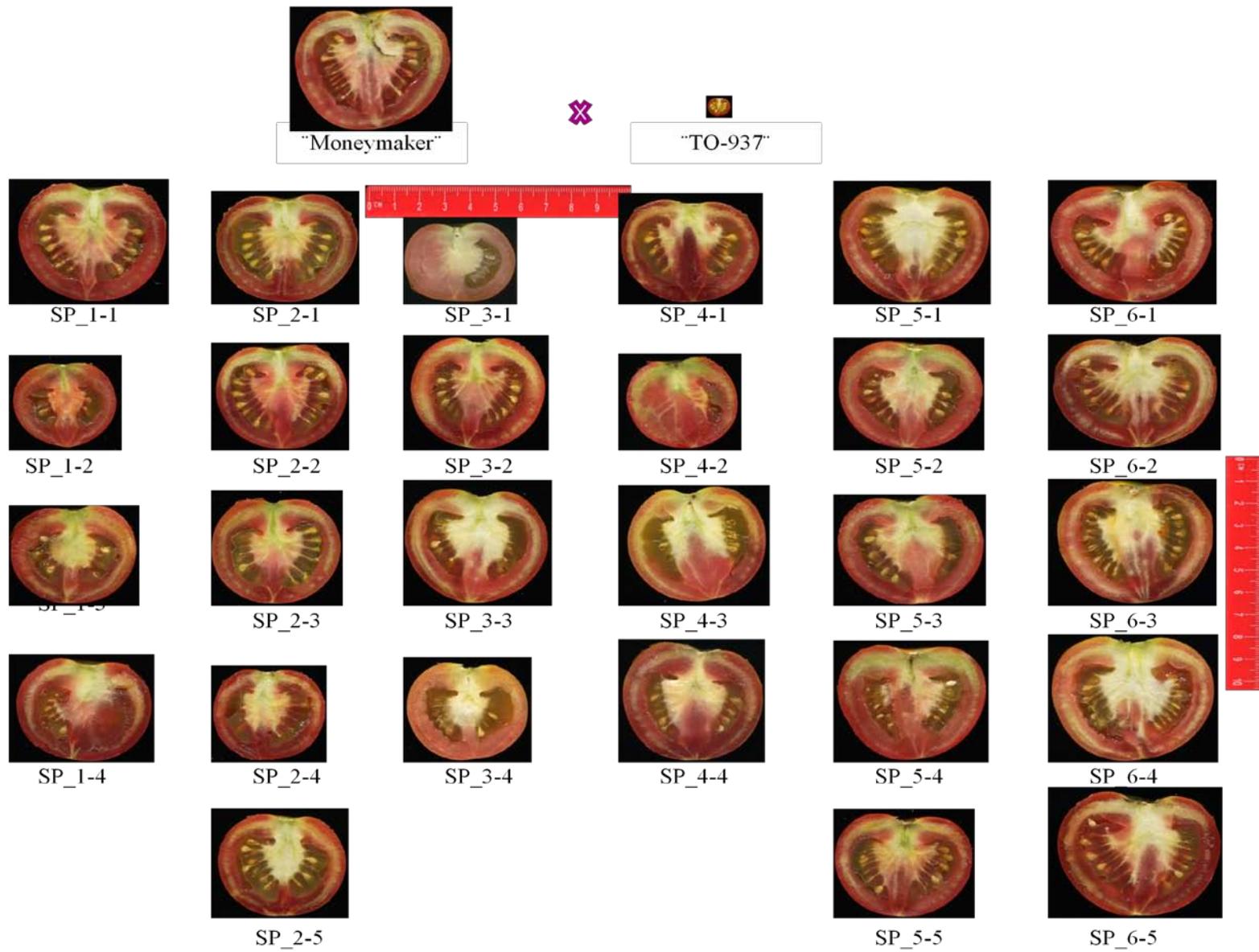
**Anexo 4-3. Análisis de varianza y prueba de Dunnett's realizadas a la genoteca completa de ILs para cada uno de los caracteres evaluados en las tres localidades. Los caracteres evaluados son: 1) Dimensión de fruto:** Peso (FW), Perímetro (PER), área de la sección longitudinal (AR), longitud máxima (FLL), diámetro máximo (FDL), 2) **Forma de fruto:** índice de forma de fruto (FS), bloque distal (DBLK), bloque proximal (PBLK), forma de fruto circular (CIR), forma de fruto triangular (TRI), altura de hombros (PSH), 3) **Químicos:** Sólidos solubles (SCC), pH inicial (pH-i) Acidez (TA), 4) **color externo e interno del fruto** (EC, IC) parámetros de color evaluados L\*, a\*, b\*, C\*, H\*. En el eje horizontal se muestran las SP (ILs) y en el eje vertical los caracteres. El orden de las pruebas / localidad es Alginet, Orihuela y Malaga. SP con letras de color negro presentan diferencias significativas ( $P < 0.05$ ) con respecto a MM.

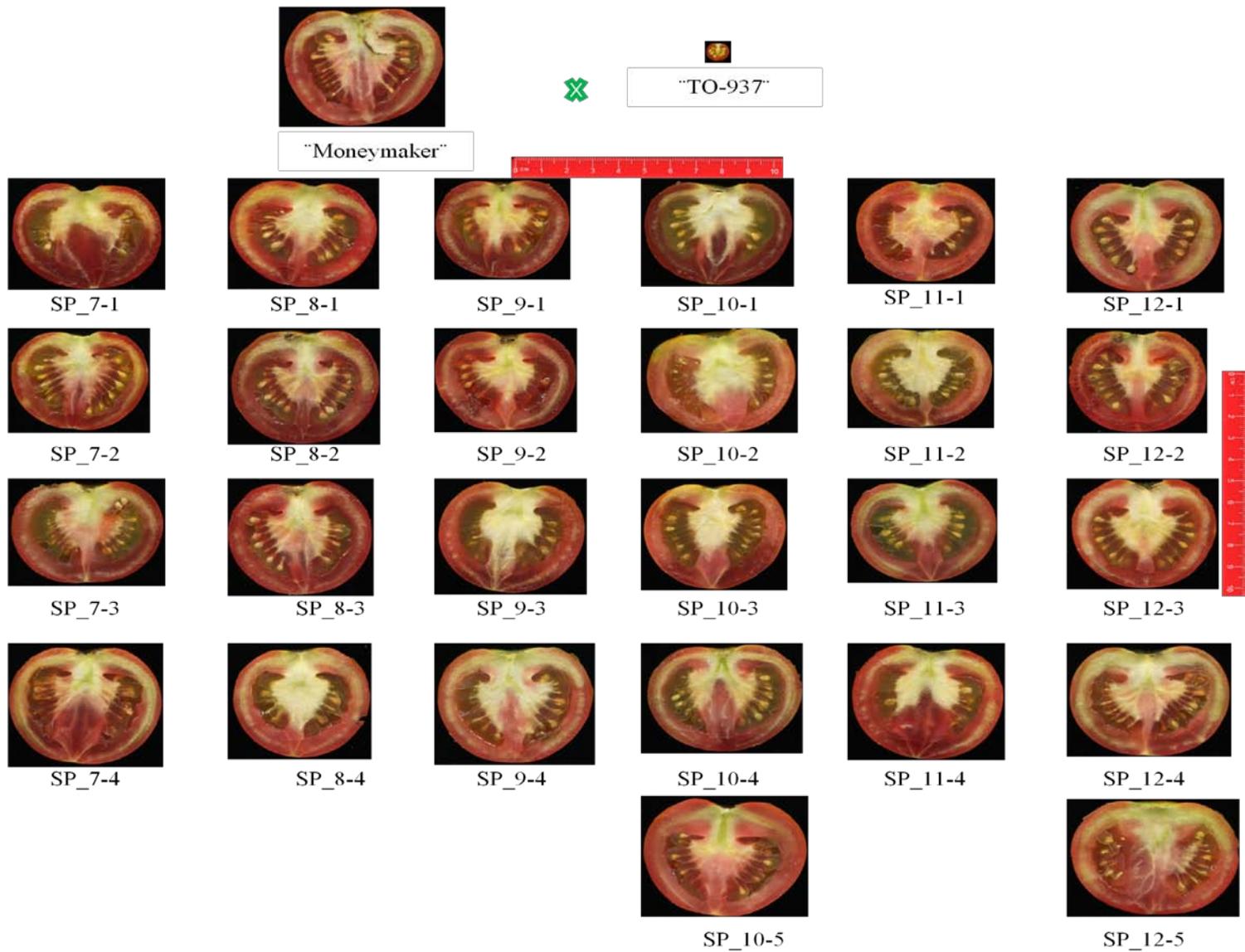






**Anexo 4-4. Imágenes de mitades de fruto de la colección completa de líneas de introgresión desarrollada a partir del cv "MoneyMaker" y acc. TO-937.**





**Anexo 4-5. Representación grafica de la variabilidad fenotípica observada en la librería completas** de líneas de introgresión en las tres localidades (A:Alginet, M:Málaga, O:Orihuela). El nombre de los caracteres se sustituye por el acrónimo respectivo. Los colores indican diferencias entre cada IL y MM (color azul: diferencias negativas y color rojo: diferencia positivas), entre más intenso sea el color mayor o menor las diferencias, en las áreas más claras las diferencias entre las ILs y el control fueron pequeñas en ambos sentidos.

Continúa anexo 4-5

IL/Loc	Dimensiones fruto					Forma de fruto						Quimicos			Color externo					Color interno				
	fv	pr	ar	fl	fd	fsl	pellk	dblk	fsll	cir	psl	BRIX	Ph	ta	L*	a*	b*	C*	H*	L*	a*	b*	C*	H*
1 1A																								
1 1M																								
1 1O																								
1 2A																								
1 2M																								
1 2O																								
1 3A																								
1 3M																								
1 3O																								
1 4A																								
1 4M																								
1 4O																								
2 1A																								
2 1M																								
2 1O																								
2 2A																								
2 2M																								
2 2O																								
2 3A																								
2 3M																								
2 3O																								
2 4A																								
2 4M																								
2 4O																								
2 5A																								
2 5M																								
2 5O																								
3 2A																								
3 2M																								
3 2O																								
3 3A																								
3 3M																								
3 3O																								
3 4A																								
3 4M																								
3 4O																								
4 1A																								
4 1M																								
4 1O																								
4 2A																								
4 2M																								
4 2O																								
4 3A																								
4 3M																								
4 3O																								
4 4A																								
4 4M																								
4 4O																								
5 1A																								
5 1M																								
5 1O																								
5 2A																								
5 2M																								
5 2O																								
5 3A																								
5 3M																								
5 3O																								
5 4A																								
5 4M																								
5 4O																								
5 5A																								
5 5M																								
5 5O																								
6 1A																								
6 1M																								
6 1O																								
6 2A																								
6 2M																								
6 2O																								
6 3A																								
6 3M																								
6 3O																								
6 4A																								
6 4M																								
6 4O																								
6 5A																								
6 5M																								
6 5O																								

Continúa anexo 4-5

IL/Loc	Dimensiones fruto					Forma de fruto						Quimicos			Color externo					Color interno				
	fw	per	ar	fl	fdl	fsl	pdlk	dblk	fsll	cir	psl	BRIX	Ph	ta	L*	a*	b*	C*	H*	L*	a*	b*	C*	H*
7_1A																								
7_1M																								
7_1O																								
7_2A																								
7_2O																								
7_2M																								
7_3A																								
7_3O																								
7_3M																								
7_4A																								
7_4O																								
7_4M																								
8_1A																								
8_1O																								
8_1M																								
8_2A																								
8_2O																								
8_2M																								
8_3A																								
8_3O																								
8_3M																								
8_4A																								
8_4O																								
8_4M																								
9_1A																								
9_1O																								
9_1M																								
9_2A																								
9_2O																								
9_2M																								
9_3A																								
9_3O																								
9_3M																								
9_4A																								
9_4O																								
9_4M																								
10_1A																								
10_1O																								
10_1M																								
10_2A																								
10_2O																								
10_2M																								
10_3A																								
10_3O																								
10_3M																								
10_4A																								
10_4O																								
10_4M																								
10_5A																								
10_5O																								
10_5M																								
11_1A																								
11_1O																								
11_1M																								
11_2A																								
11_2O																								
11_2M																								
11_3A																								
11_3O																								
11_3M																								
11_4A																								
11_4O																								
11_4M																								
12_1A																								
12_1O																								
12_1M																								
12_2A																								
12_2O																								
12_2M																								
12_3A																								
12_3O																								
12_3M																								
12_4A																								
12_4O																								
12_4M																								
12_5A																								
12_5O																								
12_5M																								

