

4 | Aplicación de la biotecnología en los programas actuales de mejora

M^a José Rubio-Cabetas¹, Belén Picó², Ana Casas³ y M^a Luisa Badenes⁴

¹ Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA)

² Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, UPV

³ Estación Experimental de Aula Dei, CSIC

⁴ Instituto Valenciano de Investigaciones Agrarias (IVIA)

4.1. Introducción

4.2. Análisis de la diversidad genética a escala genómica. Selección de fuentes de variación para los programas de mejora.

4.3. Identificación de genes y regiones genómicas involucradas en la variación de caracteres de interés

4.4. Empleo de la información genética y genómica para el desarrollo eficiente de cultivares mejorados

4.5. Referencias

4.1. Introducción

La disponibilidad de una secuencia genómica bien anotada y de un transcriptoma completo facilita enormemente la localización e identificación de los genes involucrados en los caracteres de interés. Conocer la variación en estas secuencias génicas y genómicas es un paso previo necesario para asociarla a la variabilidad fenotípica que se observa en los cultivos, lo que resulta esencial para su mejora.

Desde hace varios años venimos asistiendo a una revolución de las técnicas de secuenciación, con el desarrollo de plataformas de secuenciación a gran escala, denominadas colectivamente "*Next Generation Sequencing*" (NGS) technologies (Shendure y Ji, 2008; Metzker, 2010). Las más empleadas son las denominadas de segunda generación, como la plataforma Roche 454[®], el sistema Illumina[®], el más utilizado en la actualidad, la tecnología de Applied Biosystems SOLiD[®] o la estrategia Ion Torrent[™] de Life technologies. Además, se están poniendo a punto otras técnicas NGS de tercera generación, como Helicos Heliscope[®], Complete Genomics[®], and Pacific Biosciences SMRT[®], algunas basadas en la secuenciación de moléculas individuales de mayor longitud (Niedringhaus et al., 2011). Todas estas técnicas llevan a cabo millones de reacciones de secuenciación en paralelo y, asociadas al avance en las estrategias bioinformáticas de análisis, han permitido obtener grandes cantidades de secuencia genómica y transcriptómica en numerosos cultivos y especies silvestres relacionadas, a muy bajo costo y en un tiempo récord. Las secuencias generadas suelen depositarse en la base de datos "*Sequence Read Archive*" del NCBI (<http://www.ncbi.nlm.nih.gov/sra/>). Aparte de la información del NCBI, resultan de interés las bases de datos del PLantGDB Phytozome y Ensembl-Plants; <http://www.plantgdb.org>; <http://www.phytozome.net>; <http://plants.ensembl.org/index.html>. Los avances también se describen en revisiones recientes (Pérez de Castro et al., 2012; Gao et al., 2012).

Los cereales, entre los que se incluyen maíz, arroz, trigo y cebada se encuentran entre los principales cultivos por su producción, a nivel mundial y su uso, principalmente en alimentación humana y animal. En los últimos años, la mejora de cada una de estas especies se ha visto beneficiada por los avances en genómica y secuenciación, que han tenido lugar de forma exponencial. Hasta el año 2010, la secuenciación de genomas se basó en la secuenciación por Sanger de grandes fragmentos genómicos contenidos en cromosomas artificiales de bacterias (BACs). Esta fue la metodología utilizada para secuenciar el genoma de arroz (*Oryza sativa* L., 389 Mb), completado en 2005 (IRGSP, 2005), el de maíz (*Zea mays* L., 2300 Mb) en 2009 (Schnable et al., 2009), o el cromosoma 3B de trigo (Paux et al., 2008). El mayor tamaño de los genomas de cebada (*Hordeum vulgare* L., 5100 Mb) y trigo blando, hexaploide (*Triticum aestivum* L., 17000 Mb), con un 80-90% de ADN repetitivo, ha retrasado la secuenciación de estos dos genomas, hasta que han aparecido nuevos métodos de secuenciación. El empleo de técnicas NGS, la reducción de la complejidad de los genomas, como la separación previa de los cromosomas o brazos cromosómicos por citometría de flujo y el diseño de un orden virtual de genes, por

sintenia con otros genomas ya secuenciados (arroz, sorgo y la especie modelo *Brachypodium distachion*), se ha empleado con éxito en cebada (Mayer et al., 2011) y trigo (Hernandez et al., 2012). En noviembre de 2012 se publicó un ensamblaje del espacio génico en cebada integrado en un mapa físico y genético (IBSC, 2012) y la secuenciación parcial de trigo blando, hexaploide (Brenchley et al., 2012). El genoma de estas dos especies es parcial y habrá que esperar a la secuenciación clon a clon de BACs solapantes que corresponden a un "minimum tilling path" para cada cromosoma. Todos los recursos generados a partir de estos proyectos constituyen la base para el desarrollo de nuevas herramientas en mejora.

Un ejemplo del enorme avance que la aplicación de técnicas NGS ha supuesto en hortalizas lo tenemos en la familia de las Cucurbitáceas, en la que hasta 2009 sólo se disponía de algunas colecciones de secuencias de cDNA, habiéndose secuenciado en los últimos 4 años los genomas de las 4 especies cultivadas más importantes, el pepino (*Cucumis sativus* L., 367 Mb) (Huang et al., 2009), uno de los primeros genomas secuenciados combinando Sanger con NGS, el melón (*Cucumis melo* L., 375 Mb) (García-Mas et al., 2012), la sandía (*Citrullus lanatus* L., 425 Mb) (Guo et al., 2013) y el calabacín (*Cucurbita pepo* L.) (Fei et al., 2014). El genoma del calabacín, todavía no publicado pero accesible en <http://cucurbigene.net>, se ha obtenido en el marco de un proyecto en el que se han secuenciado los genomas de las 3 especies cultivadas más importantes del género *Cucurbita*. En el caso de las Cucurbitáceas tenemos varias bases de datos <http://www.cucurbigene.net>, <http://www.melogene.net>, <http://melonomics.net> y <http://icugi.org> que facilitan el acceso a los últimos avances en secuenciación genómica y transcriptómica en los distintos cultivos de esta familia. Otra familia de hortalizas de gran interés en la que se ha avanzado enormemente ha sido la familia de las Solanáceas. La secuenciación del genoma del tomate (*Solanum lycopersicum* L., 950 Mbp) y de la patata (*Solanum tuberosum*, 840 Mb) se inició en la década pasada con la metodología de Sanger y se finalizó con NGS (Tomato Genome Consortium, 2012; Xu et al., 2011). Otros dos cultivos importantes de esta familia, que se encontraban más atrasados en la generación de información genómica han sufrido también un gran avance. El genoma del pimiento, *Capsicum annum* L (2.700Mb) y de la especie relacionada *Capsicum chinense*, se ha publicado muy recientemente (Kim et al., 2014) y, aunque no se dispone del genoma completo de la berenjena (*Solanum melongena* L.), ya existen estudios que describen secuenciación genómica parcial (Barchi et al., 2011).

El primer genoma disponible de una especie leñosa fue el del chopo *Populus trichocarpa* secuenciado con la tecnología Sanger (Tuskan et al., 2006), seguido del de la vid, *Vitis vinífera* (Velasco et al., 2007). En la familia de las Rosáceas se han secuenciado los del manzano, *Malus domestica* (Velasco et al., 2010), la fresa, *Fragaria vesca* (Shulaev et al., 2011) el peral oriental *Pyrus bretschneideri* Rehd (Wu et al., 2013), el peral europeo *Pyrus communis* (Gardiner et al., 2012b) y el melocotonero, *Prunus persica* (Verde et al., 2013). Este último se inició también en

la década pasada con la metodología de Sanger y se finalizó con NGS. Otros frutales secuenciados han sido *Carica papaya* (Ming et al., 2008), *Citrus sinensis* (Xu et al., 2013), y el último publicado de varias especies del género *Citrus* (Wu et al., 2014). El genoma secuenciado más largo entre las leñosas ha sido el de *Pinus taeda* (Neale et al., 2014). Así mismo están disponibles secuencias parciales de cerezo y de almendro (Knoep et al., 2013), aunque no se dispone del genoma completo. Las secuencias de los genomas de Rosáceas se pueden consultar en <http://www.rosaceae.org/>.

La resecuenciación a escala genómica permite ampliar el número de polimorfismos de tipo SNPs (Single nucleotide polymorphisms) y mejorar su distribución. La resecuenciación de genomas completos se inició en especies modelo como *Arabidopsis* y arroz (Huang et al., 2009; Xu et al., 2012). La cada vez mayor disponibilidad de genomas de referencia y el desarrollo de estrategias de secuenciación de fragmentos de mayor longitud y extremos pareados que facilitan el ensamblaje (genotecas de *pair ends* o *mate pairs*) ha extendido la resecuenciación genómica a un mayor número de cultivos, con una cobertura de los genomas cada vez mayor (Lam et al., 2010; Xu et al., 2014). La abundancia de elementos repetitivos en los genomas de plantas dificulta el ensamblaje de secuencias cortas y la identificación de polimorfismos, por lo que en genomas complejos o en los que se dispone de menor información se ha generalizado la resecuenciación genómica asociada al empleo de técnicas de reducción de la complejidad de los mismos antes de la secuenciación. Algunas de estas técnicas son la selección y secuenciación de regiones poco metiladas, la secuenciación de fragmentos amplificados con AFLPs, o de secuencias diana amplificadas en PCRs múltiples o previamente capturadas mediante hibridación, o el empleo de "*Reduced Representation libraries*" (RRL) (Hyten et al., 2010; Deschamps et al., 2012; Hirsch et al., 2014). Estas estrategias persiguen reducir el coste de identificación de los SNPs sin comprometer en exceso su calidad.

Dos de las estrategias más empleadas utilizan enzimas de restricción para fragmentar previamente el genoma y reducir su complejidad. En el caso de la estrategia RAD, "*Restriction Site-Associated DNA*" (Miller et al., 2007; Baird et al., 2008; Wang et al., 2012), se ligan adaptadores con primers de amplificación y secuenciación a los extremos cohesivos de los fragmentos genómicos digeridos y se fragmentan de nuevo, secuenciándose los fragmentos resultantes. El empleo de RAD ha permitido recientemente la identificación de las primeras colecciones de SNPs en algunos cultivos como berenjena o alcachofa (Barchi et al., 2011; Scaglione et al., 2012). El segundo método, más sencillo que el anterior, es la reducción de la complejidad digiriendo previamente con enzimas que no corten frecuentemente en la fracción repetitiva del genoma, por lo que únicamente se generan fragmentos de tamaño secuenciable en determinadas regiones genómicas ricas en secuencias de copia única. La generación de secuencias se combina con la identificación de marcadores y el genotipado de los individuos en la misma reacción, por lo que nos referimos a este método como Genotipado por Secuenciación "*Genotyping by*

Sequencing" (GBS) (Elshire et al., 2011; Poland et al., 2012a). El número de SNPs identificados en este tipo de estudios es tan grande que podemos aplicar criterios muy restrictivos de calidad para su selección.

Las plataformas que compiten en genotipado masivo a gran escala, con decenas o cientos de miles de SNPs, son los chips de Affymetrix, basados en hibridación, y las plataformas Infinium^R de tecnología Illumina, basadas en extensión del cebador (Perkel, 2008; Edwards et al., 2014). Los primeros chips de Infinium se desarrollaron para un conjunto de especies seleccionadas, en las que se ha avanzado más en la identificación de SNPs, como arroz o maíz (Yu et al., 2014; Wu et al., 2014), pero el número de especies para las que se dispone de estos recursos se está incrementando rápidamente (Ganal et al., 2012). Por ejemplo, ya hay chips de este tipo disponibles en tomate, *Brassica napus*, alfalfa o soja (Song et al., 2013; Hirakawa et al., 2013; Raman et al., 2014; Li et al., 2014). En la familia de las Rosáceas se encuentran disponibles los chip de Infinium de 9K para melocotonero, 9K para el manzano, y de 6K para cerezo (Verde et al., 2012; Chagné et al., 2012).

Cuando no se dispone o no es necesario genotipar un número de SNPs tan elevado se pueden diseñar plataformas de rendimiento medio, entre las que destaca la plataforma Golden Gate^R, de tecnología Illumina (Sharpe et al., 2013; Esteras et al., 2013). Otras estrategias de genotipado a media o pequeña escala que suelen ser necesarias para validar conjuntos limitados de SNPs en muestras concretas tras el empleo de plataformas a gran escala son sondas Taqman^R de Applied Biosystems (Shen et al., 2009), *High resolution melting* (HRM) (Seipp et al., 2009), la tecnología KASPar (KBiosciences) (Cuppen, 2007) y la tecnología *Sequenom Mass Array* en distintos formatos (Gabriel et al., 2009).

Los recientes avances en secuenciación y genotipado descritos ya están cambiando las estrategias de mejora, optimizando los procesos de selección y combinación de genotipos que resulten en fenotipos favorables. La integración de esta información y herramientas afecta a las 3 fases del proceso de mejora. La caracterización y selección de fuentes de variación, la identificación de genes o regiones genómicas involucradas en el control de características de interés y el desarrollo de cultivares mejorados.

4.2. Análisis de la diversidad genética a escala genómica.

Selección de fuentes de variación para los programas de mejora

La domesticación de las plantas ha llevado a la obtención de genotipos más adaptados a los sistemas de cultivo pero a su vez ha supuesto una disminución de la variabilidad genética (Gur y Zamir, 2004). Las nuevas tendencias en mejora se orientan a la identificación de las regiones del genoma de especies silvestres que se han perdido durante el proceso y su utilización en los procesos de mejora genética. Existen numerosas evidencias que demuestran la importancia de las especies silvestres ya que representan un reservorio de genes con gran potencial para

la mejora de determinados caracteres (Hajjar y Hodgkin, 2007). En este contexto conocer la diversidad genética de una determinada especie facilita el empleo de la variación natural, base primaria de la mejora de los cultivos (Morrell et al., 2011; Nybom et al., 2014). A continuación se presentan los avances en el análisis de la diversidad genética a escala genómica en los distintos grupos de cultivos.

En cereales, hasta hace pocos años, los microsatélites fueron los marcadores más populares por su alto nivel de polimorfismo, co-dominancia y gran reproducibilidad, unido a su detección semi-automática. Aunque se siguen utilizando (Maccaferri et al., 2011; Kollers et al., 2013ab), se están viendo sustituidos por otros marcadores de mayor rendimiento. Un método de genotipado que se ha extendido en los últimos años en estos cultivos son los DArT (Diversity Arrays Technology), marcadores derivados de hibridación diferencial a representaciones del genoma después de la digestión con enzimas de restricción para reducir su complejidad. Estos marcadores son propiedad de la empresa Diversity Arrays Technology que es quien los genera. Son bialélicos, dominantes, ya que se valora la presencia/ausencia de ese fragmento en la muestra genotipada, e independientes de la secuencia de nucleótidos. Desde el primer trabajo en cebada en 2004, con la construcción de un mapa genético (Wenzl et al., 2004), se han utilizado con objetivos diversos en cebada (Szűcs et al., 2009), trigo blando (Akbari et al., 2006; Semagn et al., 2006) y trigo duro (Mantovani et al., 2008), en muchos casos integrados con otros marcadores. También se han empleado para evaluar la diversidad genética en cebada (Crossa et al., 2007; Zhang et al., 2009) y trigo blando (Crossa et al., 2007; White et al., 2008).

Los SNPs o polimorfismos en un único nucleótido son los marcadores más abundantes entre individuos de cualquier especie. Distintos proyectos de secuenciación de colecciones de cDNA o ESTs, han facilitado el diseño de nuevas colecciones de SNPs y la construcción de plataformas de genotipado de baja-media escala, entre 384 y 1536 SNPs en cebada (Close et al., 2009; Muñoz-Amatriáin et al., 2011), trigo blando (Chao et al., 2010) o trigo duro (Trebbi et al., 2011). Poco después, utilizando la información derivada de secuenciación de transcriptoma (RNAseq) o de nuevos proyectos de secuenciación genómica, se han diseñado otras plataformas de genotipado capaces de interrogar un mayor número de SNPs como los chips de Infinium de 9k en cebada (Comadran et al., 2012) o trigo (Cavanagh et al., 2013), 44k en arroz (Zhao et al., 2011), 50k en maíz (Ganal et al., 2011), hasta el nuevo chip de 90k para trigo (Wang et al., 2014). Igual que se ha mencionado para los DArTs, estas plataformas de genotipado se han utilizado para construir mapas genéticos, analizar la diversidad existente en distintos conjuntos de materiales (ej. arroz: Thomson et al., 2012; cebada: Muñoz-Amatriáin et al., 2014; maíz: Wu et al., 2014; trigo: Würschum et al., 2013). En los estudios anteriores se plantea el estudio de la variación en dos fases, primero se identifican los polimorfismos y posteriormente se genotipan los mismos en las poblaciones. Con el desarrollo del Genotipado por secuenciación (GBS), el descubrimiento de los polimorfismos y el

genotipado de los mismos puede realizarse al mismo tiempo. Como se ha comentado en la introducción, el método se basa en secuenciar una fracción del genoma, que se ha seleccionado utilizando métodos de reducción de la complejidad del genoma mediante la digestión con uno (Elshire et al., 2011) o dos enzimas de restricción (Poland et al., 2012a), con la posterior secuenciación de fragmentos cortos multiplexando muestras y utilizando los nuevos sistemas de secuenciación NGS.

Esta tecnología se ha utilizado en arroz, cebada, maíz, trigo blando y trigo duro (Elshire et al., 2011; Poland et al., 2012 a b; Saintenac et al., 2013; Spindel et al., 2013; van Poecke et al., 2013). El número de marcadores obtenidos es muy superior al de los otros sistemas de genotipado. Permite identificar dos tipos de polimorfismos: SNPs, cambios de una base en las secuencias de ADN (co-dominantes) y variantes de presencia/ausencia (dominantes). Para dar una idea de los números obtenidos, con la combinación de enzimas PstI+MseI, Poland et al., (2012a) pudieron localizar 34.000 SNPs y 240.000 variantes de presencia/ausencia en una población de cebada, mientras que Saintenac et al. (2013) consiguieron mapear 33.800 SNPs y 167.000 marcadores de presencia/ausencia en trigo blando. Recientemente se ha utilizado GBS para evaluar la diversidad en sorgo, en un panel de 971 accesiones con 265.000 SNPs (Morris et al., 2013) y se ha publicado otro trabajo en maíz, donde se han analizado 2.815 líneas puras, muchas de ellas conservadas en el banco de germoplasma americano, con 681.000 SNPs (Romay et al., 2013). El poder de detección de estos análisis es muy superior al de otros estudios con menor número de marcadores.

Otra forma de analizar la diversidad genética a escala genómica es mediante re-secuenciación del genoma completo en especies como arroz o maíz, o de una parte del genoma como sería en el caso de cebada o trigo. El menor tamaño del genoma de arroz ha facilitado el análisis de la variación secuenciando a baja cobertura un elevado número de entradas. En un primer estudio se trabajó con 517 cultivares autóctonos chinos, de los que 373 pertenecían a la subespecie *indica* (Huang et al., 2010); en otro trabajo posterior, se añadieron 100 cultivares autóctonos chinos de la subespecie *japonica* y otras 330 variedades de arroz de distintos orígenes (Huang et al., 2012). Estos dos trabajos, resultado de la secuenciación de 950 accesiones de arroz dieron lugar a 4,1 millones de SNPs, con los que se analizó la estructura de población del arroz a nivel mundial, con cinco grupos bien definidos. Un estudio reciente con tres cultivares de arroz que presentan distinta respuesta a sequía y salinidad ha permitido identificar 1,78 millones de SNPs y 154.000 polimorfismos de tipo inserción/delección, entre el genoma de referencia, construido con la variedad Nipponbare y esos tres cultivares. A partir de esos resultados se han podido identificar SNPs dentro de dominios funcionales de genes relacionados con la tolerancia a estrés (Jain et al., 2014).

Los estudios en maíz, han puesto de manifiesto que el genoma de maíz no es estático sino que está en movimiento. Un primer estudio con 6 líneas puras de maíz detectó más de 1 millón de SNPs y 30.000 indels entre la secuencia de esas

líneas y la del genoma de referencia, obtenida con la línea B73 (Lai et al., 2010). Se identificaron cientos de genes completos que presentaban variación tipo presencia/ausencia en las líneas analizadas. Al menos 157 genes no estaban presentes en la referencia B73. Otros dos trabajos posteriores identificaron 55 millones de SNPs y un alto porcentaje de variantes estructurales, de la secuenciación de 103 líneas (Chia et al., 2012), así como el elevado número de cambios genéticos que han tenido lugar a lo largo de la mejora de maíz, a partir de 278 líneas puras (Jiao et al., 2012).

En el caso de otras especies con menos información genómica, se ha recurrido a la secuenciación de sólo una fracción del genoma. Este sistema se empleó para generar un mapa haplotipo del genoma de maíz, secuenciando solo regiones de baja complejidad (Gore et al., 2009), y para capturar sólo el exoma, aquella parte del genoma que codifica para ARN mensajero, permitiendo detectar variaciones en el número de copias en un buen número de genes entre dos líneas recombinantes de maíz (Liu et al., 2012). En cebada se ha desarrollado un sistema de captura de exoma en solución, para aislar y secuenciar 61.6 Mb que es la fracción de ADN que corresponde a secuencia codificante (Mascher et al., 2013). El método, comercializado por la empresa Roche, ha sido utilizado con distintos genotipos de cebada demostrando su utilidad para descubrir nuevas variantes alélicas en zonas codificantes. Un sistema equivalente se utilizó anteriormente en trigo blando, permitiendo obtener más de 500.000 SNPs entre 8 cultivares (Winfield et al., 2012).

Para muchas especies hortícolas el paso de identificación de polimorfismos de forma masiva ya se ha llevado a cabo y existen amplias colecciones de SNPs que se han implementado en plataformas de genotipado para el análisis de la diversidad genética de la especie. Para asegurar el éxito del ensayo de genotipado es esencial optimizar la selección de los SNPs. Su distribución uniforme en el genoma debe asegurarse, bien utilizándolos en la construcción de mapas genéticos, bien situándolos directamente, en el caso de existir un genoma de referencia en la especie. Dada la variabilidad de las poblaciones que se genotipan, es necesario seleccionar SNPs que no presenten polimorfismos adicionales en zonas circundantes, que puedan alterar el genotipado de algunos individuos. Así mismo, resulta de interés incluir SNPs funcionales, es decir mutaciones que puedan estar relacionadas con cambios funcionales en genes de interés, lo que facilita la posterior asociación entre los genotipos y los correspondientes fenotipos en el germoplasma analizado. Una vez establecida la plataforma puede evaluarse su calidad, antes de su uso masivo, ensayándola con un panel de individuos que represente lo mejor posible la variación de la especie.

Como resultado de este tipo de estudios, no sólo se puede conocer, con mayor o menor detalle, la distribución de la variación en la especie y las relaciones entre individuos a escala genómica, sino también pueden obtenerse huellas moleculares específicas de genotipos concretos y estudiar la estructura genética de las poblaciones y la distribución del desequilibrio de ligamiento, factores clave para la se-

lección de fuentes de variación, para la elección de los parentales de poblaciones de mapado, y para llevar a cabo estudios de asociación con colecciones de germoplasma.

Los chips de Infinium empiezan a estar disponibles en hortícolas. Uno de los primeros se desarrolló en tomate, incluyendo más de 7.000 SNPs (Sim et al., 2012a) y se ha utilizado para genotipar distintas colecciones. En el estudio de Sim et al., 2012b, con más de 400 entradas cultivadas, se validan un 97% de los marcadores y se observa una estructura poblacional relacionada con el proceso de mejora, que separa las entradas de procesado de las de consumo en fresco y las procedentes de procesos de mejora modernos de los cultivares antiguos. Así mismo, se distinguen los tipos de tomate cereza y los pertenecientes a la especie silvestre *S. pimpinellifolium*. En este estudio de tomate se describen regiones génicas con mayor diversidad genética en los tipos mejorados, resultado de procesos de introgresión de las especies silvestres durante los programas de mejora. La identificación de las introgresiones de las fuentes de variación (genotipos silvestres, razas locales, etc.) en los cultivares mejorados es una aplicación de gran interés de los estudios de diversidad a escala genómica, ya que facilita la identificación de los genes involucrados en los caracteres mejorados y proporciona dianas de interés para la mejora. Utilizando el mismo chip de tomate, Hirakawa et al. (2013) identificaron mutaciones en 200 genes candidatos que pueden estar relacionadas con variantes fenotípicas en distintos caracteres de interés. Los chips de Affimetrix también se están aplicando a nuevos cultivos. Hill et al., (2013) empleando el Pepper GeneChip de Affimetrix para analizar 40 líneas diversas de *C. annuum* y genotipos representativos de otras 3 especies del género (*C. frutescens*, *C. chinense* y *C. pubescens*), detectaron 33.401 variantes en 13.323 unigenes e identificaron regiones con diversidad genética diferencial en genotipos picantes y no picantes. En muchos otros cultivos, donde todavía no hay disponibles chips de Infinium, ya se han realizado trabajos para el estudio de la diversidad genética y la estructura de las poblaciones empleando otras plataformas de genotipado a media escala, como Golden Gate, lo que en ocasiones se lleva a cabo como un paso previo para la selección de genotipos representativos de la variación y su uso en amplios experimentos de resecuenciación o GBS (Blanca et al., 2012; Esteras et al., 2012, 2013).

Los SNP disponibles derivados de la resecuenciación de genotipos concretos pueden no ser siempre de utilidad para detectar variación si se amplía o reduce el pool genético (Sonah et al., 2013). En este caso necesitaríamos diseñar experimentos de resecuenciación, incluyendo un mayor número de genotipos representativos de la variación a estudiar. Además, para que una plataforma de genotipado sea de utilidad en un amplio espectro de estudios suelen incluirse SNPs con alelos frecuentes, lo que lleva a un sesgo en el estudio de los alelos raros (Perkel et al., 2008). Como hemos comentado, el avance de las técnicas de secuenciación está permitiendo utilizar directamente la resecuenciación genómica para la identificación de polimorfismos a gran escala. Los estudios actuales suelen incluir genotipos silves-

tres y cultivados. La comparación de sus patrones de variación hace posible la identificación de las regiones seleccionadas durante los procesos de domesticación y o posterior mejora de una forma más precisa que las plataformas de genotipado (Tang et al., 2010; Gross y Olsen, 2010).

El avance es muy rápido, y ya comienzan a publicarse los primeros estudios de resecuenciación genómica en hortalizas, generando mapas de variación a escala genómica. Por ejemplo Qi et al., (2013) describen la resecuenciación genómica de una colección nuclear de 115 entradas de pepino, generando una colección de más de 3,6 millones de SNPs, más de 300.000 indels y casi 600 variantes de presencia/ausencia (PAV). Un aspecto esencial de este estudio es la selección previa de las entradas analizadas. Las 115 entradas secuenciadas se seleccionaron a partir de una colección de 3.342 entradas de orígenes diversos a nivel mundial y se estima que representan un 80% de la variación existente en la colección global (Lv et al., 2012). Además, se resecuencia una variedad de un genotipo silvestre de *Cucumis sativus* var. *hardwickii*. La comparación de los resultados obtenidos con los resultados previos en cereales, parece indicar que las hortalizas aprovechadas por sus frutos han sufrido un mayor estrechamiento de su base genética durante la domesticación. En este estudio se identifican más de 100 regiones que han sufrido un sesgo durante el proceso de domesticación, una de las cuales contiene el gen involucrado en la pérdida del amargor del fruto, un suceso esencial para la domesticación de las cucurbitáceas.

En general, como ya hemos comentado, los últimos genomas que se han publicado ya incluyen, aparte de genoma de referencia, varios genomas de entradas diversas, que en muchos casos incluyen especies distintas, consideradas ancestros silvestres o de utilidad en la mejora, como es el caso de la sandía (Guo et al., 2013) o el pimiento (Kim et al., 2014). En este último la secuenciación de *Capsicum annum* y especies relacionadas ha permitido identificar genes relacionados con el carácter picante de esta hortaliza.

En especies no modelo son estrategias como RAD o GBS, basadas en la reducción previa de la complejidad genómica, las que permiten llevar a cabo en un sólo paso la identificación de los polimorfismos y el genotipado de las poblaciones (Barchi et al., 2011). En estos casos, no es necesario hacer una selección previa de los SNPs y es posible añadir nuevos genotipos, o incluso especies, sin necesidad de llevar a cabo una identificación previa de SNPs específicos de los mismos.

Los trabajos citados se orientan al estudio de la diversidad natural. Sin embargo, las estrategias de genotipado y secuenciación descritas también pueden emplearse para el estudio de la variación inducida en colecciones de mutantes. Así, la identificación mediante secuenciación de las variantes inducidas es una alternativa a la digestión enzimática, empleada habitualmente en los estudios de plataformas de TILLING, que ya ha sido validada en algunas colecciones de mutantes (Gilchrist et al., 2013). Recientemente, se han resecuenciado dos genotipos de la variedad de

tomate Microtom, identificándose más de 1 millón de SNPs entre esta variedad y el genoma de referencia. Este recurso será de utilidad para el análisis de las colecciones de mutantes derivadas de esta variedad (Kobayashi et al., 2014).

El estudio de la diversidad genética es también esencial para la mejora de frutales, siendo el incremento de la variabilidad un paso básico en muchos programas de mejora. Existen numerosos estudios orientados a la búsqueda de nuevas variantes en especies cultivadas o silvestres de los distintos géneros y especies de interés. Un ejemplo muy estudiado es el manzano, donde en 20 poblaciones de *Malus sieversii* procedentes de Xinjiang, China, se encontró más variabilidad dentro de cada población que entre las poblaciones, por lo que se concluyó que *M. sieversii* es una gran fuente de diversidad genética (Richards et al., 2009). En el género *Prunus* también existen ejemplos del uso de especies silvestres con el objetivo de crear líneas de introgresión que contengan regiones del genoma ventajosas dentro de un fondo genético comercial (Donoso et al., 2012) o que contengan genes de resistencia (Esmenjaud y Dirlewanger 2007; Quilot et al., 2004; Marandel et al., 2009).

Los marcadores tipo SSRs han sido los más utilizados para el estudio de la diversidad en colecciones de germoplasma de todas las especies frutales y especies relacionadas. Desde los primeros trabajos en el género *Prunus*, melocotonero, cerezo y albaricoquero (Aranzana, et al., 2003, Dirlewanger et al., 2002, Hormaza, 2002,) y en el género *Malus* y *Pyrus* (Silfverberg-Dilworth et al., 2006, Yamamoto et al., 2002) han sido varias las colecciones a nivel local y nacional que se han estudiado en prácticamente todos los frutales de hueso (Bouhadida et al., 2011) y pepita (Pina et al., 2014). Los últimos trabajos en frutales han incluido la diversidad con varias especies silvestres relacionadas con manzano, almendro, melocotonero, albaricoquero (Gross et al., 2012; Zeinalabedinia et al., 2008; Fernández i Martí et al., 2014; He et al., 2007). En otros géneros que incluyen especies de gran importancia económica también han se ha estudiado la variabilidad de especie relacionadas de *Citrus* (Noelle et al., 2006), *Vitis* (Wan et al., 2008, 2013) y *Olea* (Belaj et al., 2010). En otras especies frutales los trabajos más recientes basados en marcadores moleculares para estudios de germoplasma se han dado en frutales menores como en los géneros *Eriobotrya* (Gisbert et al., 2009) y *Dyospyrus* (Naval et al., 2010).

Los estudios basados en marcadores tipo SSRs han demostrado la baja diversidad genética presente en variedades comerciales de melocotonero en Europa y USA (Vives, 2012), debido a que se ha utilizado una limitada base genética en todos los programas de mejora tanto públicos como privados. Esta especie es un claro ejemplo en el que el progreso en mejora necesita incorporar nuevas fuentes de diversidad. Un estudio de una colección de germoplasma compuesta por 97 accesiones de diferentes orígenes geográficos demostró la clara ganancia aportada por la misma respecto al material manejado en los programas (Badenes et al., 2014).

El desarrollo de marcadores SSR derivados de genotecas EST también ha incrementado la disponibilidad de marcadores y su aplicación en estudios de diversidad

genética en *Citrus* (Luro et al., 2008), *Vitis* (Myles et al., 2011), *Rubus* (Lewers et al., 2008), *Malus* (Gasic et al., 2009), *Prunus* (Lazzari et al., 2008). La secuenciación parcial del ADN cloroplástico también ha servido para establecer relaciones genéticas en el género *Prunus* (Bielsa et al., 2014) o dilucidar las relaciones filogenéticas del género *Citrus* (Bausher et al., 2006).

A partir de la mejora de las técnicas de secuenciación se han empezado a utilizar marcadores tipo SNPs. En el marco del Consorcio Internacional de SNPs de Melocotonero (IPSC, del inglés, *International Peach SNP Consortium*), Verde et al. (2012) desarrollaron un chip de Infinium de aproximadamente 9K SNPs. Este chip se ha empleado para el genotipado de numerosas colecciones de variedades de melocotonero, cerezo, almendro y algunos híbridos, así como poblaciones de mapeo de *Prunus*, tanto intra como interespecificas (Vives, 2012, Sánchez et al., 2014; Romeu et al., 2014). Los chips 9K de Illumina II array se ha utilizado para evaluar el nivel de diversidad genética en el género *Malus* y ha demostrado un alto nivel de transferibilidad de este género al género *Pyrus*. En manzano, el consorcio internacional RosBREED SNP Consortium (IRSC) ha desarrollado un chip 9K basado en la tecnología Illumina con el fin de evaluar la diversidad alélica con fines de mejora (Chagné et al., 2012). Igualmente en vid también se han desarrollado SNPs que se han aplicado al genotipado de colecciones (Myles et al., 2011).

4.3. Identificación de genes y regiones genómicas involucradas en la variación de caracteres de interés

Uno de los principales objetivos de los avances genómicos en los cultivos de interés agrícola es la conexión entre el genotipo y el fenotipo. La mejora molecular llevada a cabo en las décadas pasadas se ha orientado a explorar la base genética de la variación, utilizando mapas genéticos construidos con un número limitado de marcadores (RFLPS, RAPDS, AFLPs, SSRs). El empleo de los mismos ha resultado exitoso para la localización y clonado de genes involucrados en características cualitativas, en general de alta heredabilidad y fenotipado sencillo. Sin embargo, muchos de los caracteres de interés agronómico presentan variación cuantitativa, debido a un control poligénico, de varios genes con efectos menores e interacciones epistáticas entre ellos, y a la influencia del ambiente en su expresión. El empleo de estos mapas poco densos no permitía obtener información precisa sobre el número y la localización de los QTLs (*Quantitative trait loci*) involucrados en el control de estas características complejas.

Tradicionalmente, las poblaciones empleadas para la construcción de mapas y la localización de genes o regiones cromosómicas asociadas con caracteres de interés han sido poblaciones biparentales. Hay muchos trabajos de este tipo, en los que sólo se pueden evaluar dos alelos a la vez, conforme al diseño del cruzamiento estudiado. Por este motivo, se han impuesto en los últimos años los métodos de mapeo por asociación, o mapeo por desequilibrio de ligamiento (LD). El LD es una

estimación de la asociación no aleatoria de marcadores a lo largo del genoma, no necesariamente en el mismo cromosoma. El desequilibrio de ligamiento decrece con la distancia genética, pero puede estar afectado por múltiples factores, entre ellos el sistema reproductivo (las especies autógamas suelen tener mayor nivel de LD), la estructura poblacional, la existencia de fenómenos de epistasia o la ocurrencia de cuellos de botella durante la domesticación y los procesos de mejora. La intensidad del desequilibrio no es uniforme a lo largo del genoma y difiere bastante entre cultivos, reduciéndose en algunos en pocas kb y conservándose en otros durante varias Mb.

En los estudios de asociación no es necesario construir una población específica para cada carácter. Suelen utilizarse paneles de genotipos ya disponibles, bien caracterizados fenotípicamente en los que se pueda asociar un determinado fenotipo con una variante genética. Al disponer de métodos de genotipado de alto rendimiento, se pueden hacer análisis a lo largo de todo el genoma (*Genome-wide association study*, GWAS) e identificar regiones que pueden estar asociadas con variación en el carácter analizado. Las diferentes plataformas de genotipado mencionadas en el apartado anterior se han utilizado para estudios de asociación genética en varios cultivos, tanto los DArTs (Comadrán et al., 2009; Kulwal et al., 2012; Roy et al., 2010; Yu et al., 2011, 2012; Ziems et al., 2014) como los chip de Infinium (Famoso et al., 2011; Comadrán et al., 2012).

Una revisión de los artículos publicados en *Molecular Breeding*, en el periodo 2011-2014 (Cuadro 4.1) indica que: la mayoría de los artículos de arroz son en poblaciones biparentales y se dirigen a identificar resistencias a bacterias (*Xanthomonas oryzae*), hongos (*Magnaphorte oryzae*), distintos virus y plagas de insectos, y sólo unos pocos trabajos se centran en la búsqueda de QTLs para caracteres agronómicos. En cebada hay un número equivalente de artículos describiendo QTLs para caracteres agronómicos o de calidad maltera, y otros que buscan identificar resistencias a distintos hongos o virus utilizando poblaciones biparentales y también paneles de asociación. En el caso de maíz, la mayoría de los trabajos se centran en la búsqueda de QTLs para caracteres agronómicos (altura de planta, longitud de la mazorca, número de granos, rendimiento) en cruzamientos biparentales o el análisis de varias poblaciones a la vez. En este cultivo hay pocos trabajos de resistencia a enfermedades, entre los que destacan la identificación de QTL para resistencia a la acumulación de aflatoxina, la toxina producida por infección de *Aspergillus flavus* (Willcox et al., 2013). Aparecen trabajos en trigo dirigidos a identificar QTL para caracteres morfológicos y distintos componentes de calidad. Por último, hay muchos artículos describiendo la búsqueda de resistencias a enfermedades fúngicas como carbón, oidio, roya amarilla (*Puccinia striiformis*), roya del tallo (*P. graminis*), roya de la hoja (*P. triticina*), *Septoria*, y a distintos virus. Se debe destacar la identificación de marcadores ligados a resistencia a la raza Ug99, muy virulenta, de roya del tallo, en un estudio con 6 poblaciones biparentales de trigo (Lopez-Vera et al., 2014).

Cuadro 4.1.—Ejemplo de algunos QTLs localizados recientemente en mapas de cereales por especie y carácter cartografiado

Carácter	Especie	Referencia
Crecimiento radicular	arroz	Uga et al. 2013a
Floración y altura de la planta	arroz	Liu et al. 2013a
Rendimiento en sequía	arroz	Venuprasad et al., 2012
Resistencia a podredumbre bacteriana	arroz	Mizobuchi et al., 2013
Resistencia a virus	arroz	Li et al., 2013
Calidad maltera	cebada	Islamovic et al., 2014
Floración	cebada	Comadran et al., 2012
Resistencia a nemátodos	cebada	Galal et al., 2014
Resistencia a <i>Rhynchosporium</i>	cebada	Hofmann et al., 2013
Resistencia a roya (<i>P.hordei</i>)	cebada	Ziems et al., 2014
Tolerancia a frío	cebada	Fisk et al., 2013
Altura de planta	maíz	Teng et al., 2013
Altura de planta	maíz	Peiffer et al., 2014
Caracteres agronómicos	maíz	Xue et al., 2013
Contenido y composición aceite	maíz	Li et al., 2013b
Respuesta al. fotoperíodo	maíz	Hung et al., 2012
Resistencia a virus	maíz	Zambrano et al., 2014
Tolerancia a aluminio	maíz	Maron et al., 2013
Tamaño del grano	maíz	Liu et al., 2014
Tolerancia a sequía	trigo duro	Maccaferri et al., 2011
Resistencia a roya del tallo	trigo duro	Letta et al., 2013
Contenido en carotenoides	trigo blando	Ravel et al., 2013
Resistencia a roya del tallo	trigo blando	Lopez-Vera et al., 2014
Resistencia a plagas de insectos	trigo blando	Joukhadar et al., 2013
Resistencia a <i>Fusarium</i>	trigo blando	Kollers et al., 2013b
Rendimiento, peso hectolitrico	trigo blando	Bordes et al., 2014
Rendimiento	trigo blando	Lopes et al., 2013
Contenido proteína en grano	trigo blando	Bogard et al., 2013

Además de poblaciones biparentales, ha habido grandes avances en mapeo por asociación. No es posible describir todas las iniciativas en este sentido, pero como ejemplo se destacan las siguientes. Para contrarrestar el efecto de la sequía sobre el rendimiento del maíz, Xue et al., (2013) utilizaron 350 líneas puras genotipadas con SNPs, derivados de genes relacionados con sequía, y otros 56.000 SNPs no preseleccionados. En este trabajo se identificaron 33 genes asociados con distintos caracteres agronómicos. Entre ellos uno, que corresponde a un factor de transcripción, presentó asociaciones con la posición de la mazorca, el peso de grano y el tiempo a floración, masculina y femenina. En trigo, se han identificado asociaciones importantes para peso hectolítrico, fecha de floración y rendimiento, a partir de tres poblaciones desarrolladas con material élite de distintos programas de mejora, y genotipadas con 2500 marcadores DArT (Bordes et al., 2014). En cebada, se está utilizando mapeo por asociación para identificar nuevas resistencias a enfermedades fúngicas dentro de programas de mejora. Así se hizo para buscar resistencia estable a la roya, (*P. hordei*) dentro de material élite australiano (360 líneas genotipadas con 3244 marcadores DArT), donde se ha verificado la posición de genes de resistencia conocidos y se han detectado otros nuevos (Ziems et al., 2014). De forma similar, se han buscado QTL de resistencia a infección por *Fusarium* (Huang et al., 2013) o a la misma raza Ug99 de roya del tallo (Zhou et al., 2014), para su incorporación en programas de mejora en Estados Unidos.

En el caso específico del maíz, especie alógama, con bajo desequilibrio de ligamiento (2 kb), se necesitaría emplear un elevado número de marcadores para encontrar asociaciones entre marcadores y fenotipos con esa metodología. Una alternativa que se ha utilizado ha sido la obtención de una población de líneas emparentadas NAM, (Nested Association Mapping, McMullen et al., 2009). La población se construyó a partir de 25 líneas puras, elegidas maximizando la diversidad entre ellas, por cruzamiento con un mismo parental. Se obtuvieron 25 familias de líneas puras recombinantes (RIL), cada una de 200 individuos. La población NAM, de 5000 líneas ha sido genotipada y fenotipada extensivamente y es un recurso excelente para estimar la variación en caracteres muy distintos como fecha de floración (Buckler et al., 2009; Hung et al., 2012), tamaño y disposición de la hoja (Tian et al., 2011); resistencia a patógenos (Kump et al., 2011; Poland et al., 2011), o la composición química del grano. En todos esos trabajos se ha llegado a identificar genes importantes, con variación alélica clara y útiles en mejora. En cebada se está desarrollando una población similar, HEB-25, resultado del cruzamiento de la variedad élite Barke con 25 donantes de la especie silvestre, *H. vulgare ssp. spontaneum* (Schnaithmann et al., 2014).

Otros recursos genéticos especiales, para facilitar la localización e identificación de genes, son las poblaciones MAGIC, (Multiparent Advanced Generation Intercross, Cavanagh et al., 2008) desarrolladas a partir de cruzamientos entre múltiples parentales. Este tipo de poblaciones se están utilizando en arroz y trigo (Bandillo et al., 2013; Rebetzke et al., 2014). Los primeros resultados, de la población MAGIC

han permitido identificar genes conocidos y QTLs nuevos para tolerancia a estrés abiótico, resistencia a enfermedades y caracteres de calidad.

Los nuevos avances en secuenciación y genotipado permiten la construcción de mapas genéticos de alta densidad y la construcción de poblaciones más adecuadas para el mapado de características cuantitativas. En el Cuadro 4.2 se muestran QTLs identificados en mapas de especies hortícolas. Son numerosos los ejemplos en los que se orienta la identificación de SNPs para generar marcadores útiles para mapado, resecuenciando los parentales de poblaciones de mapado de

Cuadro 4.2.—Ejemplo de algunos QTLs localizados recientemente en mapas de algunos cultivos hortícolas por especie y carácter cartografiado

Carácter	Especie	Referencia
Patrón de Volátiles del fruto	tomate	Rambla et al., 2014
Contenido en Licopeno del fruto	tomate	Kimkade y Foolad, 2013
Resistencia a <i>Phytophthora infestans</i>	tomate	Chen et al., 2014
Número y tamaño del fruto	pimiento	Dwivedi et al., 2013
Contenido en pigmentos del fruto	pimiento	Brand et al., 2011
Contenido en capsicina	pimiento	Reddy et al., 2014
Contenido en antocianina	berenjena	Barchi et al., 2011 y 2012
Presencia de espinas	berenjena	Frary et al., 2014
Resistencia a hongos y virus.	melón	Díaz et al., 2011
Tipo de maduración	melón	Vegas et al., 2013
Precocidad de floración y parámetros de calidad del fruto	calabacín	Esteras et al., 2013
Contenido en azúcares del fruto	sandía	Ren et al., 2014
Precocidad de floración	pepino	Lu et al., 2014
Resistencia a oidio	pepino	He et al., 2013
Control de la ginoecia	melón amargo	Matsumura et al., 2014
Características de semilla	judía	Yuste-Lisbona et al., 2014
Resistencia a la antracnosis	altramuz azul	Yang et al., 2012
Resistencia a mildiu	lechuga	Den Boer et al., 2013
Color de la semilla y estructura y color de la hoja	lechuga	Kwon et al., 2013
Precocidad de floración y resistencia a enfermedades	colza	Raman et al., 2014

interés. Por ejemplo, Barchi et al., (2011 y 2012), empleando RAD y secuenciación con illumina, ampliaron la colección disponible de microsatélites en berenjena, una de las especies de la familia de las solanáceas de mayor interés agronómico y con menor información genómica, generando la primera colección de marcadores de tipo SNPs que se utilizan para construir varios mapas genéticos, inter- e intra-específicos. Estos mapas se han empleado, entre otras cosas, para el estudio de la base genética del contenido en antocianina, identificando QTLs y genes candidatos. Estudios similares se han llevado a cabo en otras hortalizas (Esteras et al., 2012; Garcia-Mas et al., 2012). Recientemente, empleando un chip de SNPs de Affymetrix se ha construido un mapa de alta densidad en una población de RILs de lechuga, con más de 13.000 marcadores en más de 12.000 unigenes (Truco et al., 2013). El chip de Infinium de tomate se utilizó para generar 3 mapas a partir de poblaciones F_2 derivadas de cruzamientos interespecíficos, 2 de *Solanum lycopersicum* \times *S. pennellii*, y la tercera de *S. lycopersicum* \times *S. pimpinellifolium* (Sim et al., 2012a). Para mapear QTLs a gran escala en *Brassica napus* se ha empleado un chip de infinium, con más de 5.000 SNPs, en el genotipado de una población de líneas diplohaploides. El mapa construido se ha empleado para identificar *loci* asociados a precocidad de floración y resistencia a enfermedades (Raman et al., 2014).

Los mapas, al igual que ocurre en los estudios de diversidad, pueden construirse directamente resecuenciando una población adecuada de mapado (Schneeberger y Weigel, 2011)., aunque hay otros métodos que no requieren de la resecuenciación de la población completa de utilidad sobre todo cuando se buscan marcadores asociados a caracteres de control genético simple y fenotipado sencillo. En estos casos se acelera el proceso resecuenciando sólo individuos informativos con fenotipos contrastantes para el carácter. Así, Yang et al., (2012) describen la identificación de marcadores ligados a menos de 1 cM al gen de resistencia *Lanr1*, que confiere resistencia a la antracnosis causada por *Colletotricum lupini* en *Lupinus angustifolius*, combinando NGS con RAD secuenciando 20 individuos informativos y confirmando el ligamiento en una población extendida. Una alternativa más sencilla es la combinación de técnicas de secuenciación masiva con BSA (*Bulk Segregant Analysis*), sustituyendo el genotipado por la secuenciación de dos pools de plantas. Así el ligamiento entre los marcadores y el gen que controla el carácter se determina mediante el análisis de las frecuencias alélicas en ambos pools. Por ejemplo, Matsumura et al., (2014) combinan la resecuenciación con la técnica RAD de una población F_2 de *Momordica charantia* segregante para la ginoecia/monoeecia, con el BSA de pools ginoicos/monoicos para acelerar la identificación de marcadores ligados a los genes que controlan la ginoecia.

Una vez identificados el gen o genes candidatos relacionados con procesos de interés, una forma de comprobar si están asociados al mismo es determinar su expresión diferencial en distintas condiciones. La abundancia diferencial de distintos transcritos medida por ejemplo con qRT-PCR o microarrays ha sido ampliamente utilizada en décadas anteriores. La secuenciación masiva, concretamente el RNA-

seq, puede también aplicarse a la cuantificación de la expresión génica. Estos estudios de expresión con RNA-seq están sustituyendo a los microarrays (Martin et al., 2013), ya que presentan un rango más amplio de cuantificación y permiten extender el estudio a transcritos y variantes de splicing no conocidas previamente, incluyendo transcritos raros, que son difíciles de identificar con otras estrategias. Esta información permite mejorar las anotaciones de los genomas, siendo las variantes de splicing muy útiles para la identificación de las uniones exón/intrón. Sin embargo, el RNA-seq para estudios de expresión no está exento de limitaciones, como las dificultades de un correcto ensamblaje y anotación del gran número de secuencias cortas que se generan para cada fragmento y el sesgo que se puede producir durante la construcción de las genotecas para secuenciar.

Algunos de los genes que se han identificado como responsables de la variación de QTLs son factores de transcripción (Powell et al., 2012). Una forma de encontrar factores de transcripción que regulan la expresión de otros genes es mediante QTLs de expresión (e QTLs), cuando la variación en una determinada región genómica resulta en cambios en la abundancia de determinados transcritos (Holoway y Li, 2010). La mejora se ha orientado fundamentalmente a la identificación de factores en *cis* (polimorfismos nucleotídicos o presencia o ausencia del gen involucrado en el carácter), pero también resulta imprescindible el estudio de factores en *trans* que puedan regular la expresión de estos genes. Los estudios de expresión basados en RNA-seq pueden contribuir a avanzar en este campo.

A pesar de que el empleo de nuevas poblaciones y el incremento de la densidad de los mapas genéticos han incrementado la eficiencia en la detección de QTLs, con marcadores asociados útiles para la selección asistida, en muchos casos esta estrategia no ha llevado a la identificación de los genes responsables de los mismos (Salvi y Tuberosa, 2005). Como se ha comentado en el caso de los cereales, una alternativa al mapado en poblaciones derivadas del cruce de parentales seleccionados son los estudios de asociación a escala genómica (GWAS) (Rafalski et al., 2011; Duran et al., 2010). Los estudios de asociación explotan la variabilidad natural generada por sucesos de recombinación ocurridos en múltiples generaciones en una población determinada.

Uno de los pasos preliminares para abordar un estudio de asociación es el análisis a escala genómica del desequilibrio de ligamiento, ya comentado anteriormente. Esta información permite optimizar las estrategias de mejora, ya que la existencia de bajos niveles de desequilibrio permite llevar a cabo con éxito estudios de GWA, pero como contrapartida requiere un número muy elevado de marcadores para saturar el genoma e identificar las asociaciones con una resolución óptima. Los elevados niveles de desequilibrio permiten rastrear el genoma con un menor número de marcadores, pero la resolución del mapado será menor. Los marcadores que se encuentren ligados serán útiles en procesos de selección asistida por marcadores, pero será difícil el clonaje posicional de los genes asociados a los mismos.

Un ejemplo de cómo el nivel y la distribución del desequilibrio pueden afectar a los resultados lo tenemos en tomate, una especie con un alto desequilibrio de ligamiento, lo que dificulta el mapado posicional. Ranc et al. (2012) llevaron a cabo con éxito la identificación de asociaciones significativas con el peso del fruto, el número de lóculos y el contenido en sólidos solubles en tomate. Para lograr una mayor resolución en el mapado por asociación emplearon *S. lycopersicum* var. *cerasiforme*, que presenta una estructura genómica mixta entre *S. lycopersicum* y su ancestro silvestre *S. pimpinellifolium*, y en la que se ha encontrado un menor nivel de desequilibrio de ligamiento.

Otros estudios que analizan una mayor diversidad de material vegetal, empleando entradas de *S. lycopersicum*, *S. cerasiforme* y *S. pimpinellifolium*, identifican diferencias en la extensión del desequilibrio de ligamiento en regiones específicas, relacionadas con introgresiones de especies silvestres en especies cultivadas, llevadas a cabo durante el proceso de mejora (Sim et al., 2012b; Xu et al., 2013; Causse et al., 2013). Estos resultados deben ser tenidos en cuenta en los estudios de asociación. Empleando estrategias similares, Kwon et al. (2013) describen un estudio empleando 300 líneas homocigotas de lechuga en el que identifican marcadores asociados a caracteres de color de la semilla, y estructura y color de la hoja, útiles para selección asistida.

En lugar de estudiar la asociación a escala genómica, el estudio puede orientarse a la asociación de mutaciones en genes candidatos. Reddy et al. (2014) llevan a cabo un estudio de asociación para identificar SNPs relacionados con el contenido en capsicina en pimiento. Partiendo de 4 genes candidatos identifican un conjunto de SNPs en el promotor de uno de ellos, relacionados de forma consistente con la variación en el contenido en capsicina de esta hortaliza.

Una de las principales estrategias que se ha empleado en especies arbóreas para encontrar asociación entre genotipo y fenotipo ha sido la elaboración de mapas genéticos con la identificación de QTLs o regiones del genoma ligadas a los caracteres de interés. En especies arbóreas la mayor parte de los mapas realizados conciernen a especies frutales de interés económico como son los frutales de hueso o género *Prunus* y frutales de pepita o género *Malus*. Dada la alta heterozigosidad de estas especies y la herencia cuantitativa o poligénica de muchos de los caracteres de interés, se ha conseguido determinar en la mayoría de los casos las zonas del genoma ligadas, pero no con la suficiente precisión para poder realizar mejora asistida. Sin embargo, estos resultados suponen puntos de partida para que a través de los genomas publicados y los marcadores SNPs disponibles que permiten saturar las regiones diana, obtener marcadores y potenciales genes candidatos para MAS.

Uno de los caracteres que ha despertado más interés en los últimos años es el de las necesidades en frío para la salida de la latencia, así como caracteres fenológicos, como fechas de floración y maduración del fruto, pues comprenden el

mayor factor limitante para la expansión de los frutales en zonas cálidas y a su vez representan la base genética sobre la que actuar en condiciones de cambio climático. Los primeros trabajos se realizaron en *Populus* (Frewen et al., 2000) y posteriormente en especies del género *Prunus*. En melocotonero se localizó un mutante que no entraba en latencia (evergrowing), el mapeo posicional de esta mutación determinó que se debía a una delección en un tandem de genes MADS-box (Bielenberg et al., 2008). Posteriormente, mediante análisis de QTLs se determinó que el mayor QTL asociado a necesidades en frío y fecha de floración se localizaba en el cromosoma 1 de albaricoquero (Okolulu et al., 2009) y melocotonero (Fan et al., 2010), y ambos se co-localizaban con los genes MADS-box. Esta localización se ha corroborado en un mapa posterior de melocotonero, que además incluye otros caracteres fenológicos (Romeu et al., 2014). A partir de la secuencia de individuos extremos para el carácter salida de latencia se identificaron polimorfismos y potenciales genes candidatos en 3 de los más importantes QTLs descritos ligados a este carácter (Zhebentayeva et al., 2014). Similares QTLs, aunque con diferencias en la contribución individual, se han detectado en otras especies de *Prunus* (Campoy et al., 2011; Sánchez-Pérez et al., 2012; Dirlenwanger et al., 2012) y en *Malus* (Celton et al., 2011). Los primeros análisis de QTLs ligados a caracteres de calidad del fruto se realizaron en melocotonero. Posteriormente, se han estudiado el efecto pleiotrópico entre caracteres de calidad y fecha de maduración (Eduardo et al., 2011). En melocotonero se han desarrollado mapas de alta densidad de SNPs obtenidos por secuenciación de los parentales para determinar QTLs de calidad del fruto (Martínez-García et al., 2013). La alta densidad de marcadores ha permitido hacer hipótesis sobre los potenciales genes candidatos. En albaricoquero se han obtenido QTLs para la arquitectura del árbol, fecha de brotación y características del fruto (Socquet-Juglard et al., 2013). En manzano se han realizado análisis de QTLs en 20 poblaciones que raramente comparten parentales. Entre los primeros trabajos se encuentra el de Liebhard et al. (2003) que cartografió numerosos caracteres de calidad del fruto y fenología. Otros trabajos estudiaron caracteres como la firmeza y características organolépticas (Cuadro 4.3). La resistencia a patógenos también es una prioridad en la mejora. Se han identificado algunas de herencia mendeliana que han permitido el establecimiento de marcadores y mejora asistida (Esmenjaud y Dirlwanger, 2007), mientras que en otras de herencia cuantitativa, se han identificado QTLs, lo que comporta limitaciones para la selección asistida por marcadores (Cuadro 4.3). Otra opción es la selección a partir del genoma o GS (genomic selection), una alternativa viable en aquellas especies que cuentan con genomas conocidos y dispongan de SNPs para genotipar. Kumar et al. (2012) utilizaron el chip 8K de manzano para identificar polimorfismos ligados a caracteres del fruto que permitan la mejora asistida.

La identificación de los genes subyacentes a estos QTLs se ve facilitada por los estudios de expresión de genes candidatos. Se han llevado a cabo estudios de expresión con Microarrays conteniendo unigenes y ESTs, para identificar candidatos de calidad de fruto en *Prunus*. El primer microarray empleado fue el uPEACH1.0,

Cuadro 4.3.—Ejemplo de algunos QTLs localizados recientemente en mapas de algunos cultivos leñosos por especie y carácter cartografiado

Carácter	Especie	Referencia
Necesidades en frío	melocotonero	Fan et al., 2010, Romeu et al., 2014 Zhebentayeva et al., 2014
Necesidades en frío	almendro	Sánchez Pérez et al., 2012
Necesidades en frío	albaricoquero	Olokulu et al., 2009
Fenología de yemas	manzano	Celton et al., 2011
Fenología de yemas	melocotonero, cerezo	Dirlewanger et al., 2012
Latencia de yemas	chopo	Chen et al., 2002
Cuajado de yemas	chopo	Fabbrini et al., 2012
Latencia de yemas	melocotonero	Blaker et al., 2013
Fenología	albaricoquero	Campoy et al., 2011 Soquet-Juglard et al., 2013
Calidad del fruto	almendro	Sánchez Pérez et al., 2010 Font i Forcada et al., 2012 Fernández i Martí et al., 2013
Calidad y características del fruto	melocotonero	Martinez-Garcia et al., 2013 Eduardo et al., 2011 Quilot et al., 2004
Características del fruto	vid	Cabezas et al., 2006
Calidad del fruto y fenología	manzano	Liebhard et al., 2003 Kenis et al., 2008 Potts et al., 2014
Resistencia al fuego bacteriano	manzano	Peil et al., 2007, Gardiner et al., 2012a
Resistencia a <i>Xanthomonas</i>	melocotonero	Yang et al., 2013
Resistencia a oidio	manzano, melocotonero	Calenge y Durel 2006 Lalli et al., 2005
Resistencia a moteado	peral	Pierantoni et al., 2007
Resistencia al virus de la tristeza	naranjo	Asins et al., 2004
Tolerancia a clorosis	vid, <i>Prunus</i>	Bert et al., 2013, Gonzalo et al., 2012

con 4,806 oligonucleotidos (Trainotti et al., 2006) con el cual se identificaron genes involucrados en la fase climatérica, así como en calidad post-cosecha en melocotonero (Ziliotto et al., 2008) y albaricoquero (Manganaris et al., 2011). El siguiente microarray Chill-Peach fue desarrollado para estudiar daños por frío en melocotonero (Ogundiwin et al., 2008), conteniendo 4.468 unigenes de mesocarpo. Este último microarray ha sido utilizado para identificar distintos genes involucrados en la tolerancia a la asfixia radicular en *Prunus* (Rubio-Cabetas et al., 2010). En chocho se han identificado genes diferenciales con respecto a *Arabidopsis* (Christianson et al., 2010) en respuesta a hypoxia, e igualmente mediante la comparación con *Arabidopsis* se han identificado dos genes en respuesta a estreses abióticos en *Prunus* (Almada et al., 2013). Otra técnica que se ha utilizado para identificar genes asociados a caracteres cuantitativos ha sido la hibridación sustractiva por supresión (SSH). Esta técnica, combinada con el empleo de microarrays, permitió determinar los genes relacionados con la salida del reposo invernal en melocotonero (Leida et al., 2010) y la identificación de genes que se inducían ante infecciones de fuego bacteriano (Noreli et al., 2009). En manzano se han identificado genes en respuesta a la sequía por medio de una genoteca de raíz SSH (Kilian et al., 2007). En cítricos Yun et al. (2012) estudiaron la expresión diferencial de genes que intervienen en el mantenimiento de la calidad del fruto en post-cosecha. Por último, los estudios de asociación se han desarrollado en melocotonero con SNPs asociados a fecha de cosecha, contenido en flavonoides, sorbitol, azúcares totales y firmeza (Picañol et al., 2013; Font i Forcada et al., 2013).

4.4. Empleo de la información genética y genómica para el desarrollo eficiente de cultivares mejorados

Una de las primeras aplicaciones de las nuevas tecnologías genómicas con mayor impacto en mejora es el desarrollo de protocolos para selección asistida por marcadores (MAS), que facilitan la introgresión de *loci* de interés en materiales élite mediante retrocruzamiento. Un ejemplo fue la introducción de un QTL para tolerancia a la inmersión (sumergimiento) en arroz (*Sub1*), lo que supuso mejorar los rendimientos en una zona muy amplia de Asia (Septiningsih et al., 2009). La mejora privada, especialmente en maíz, utiliza rutinariamente esta metodología. En otros cereales como cebada y trigo, donde la mejora pública es importante, se han diseñado marcadores funcionales, del propio gen o asociados a genes que controlan caracteres morfológicos, agronómicos, o de resistencia a enfermedades que pueden ser utilizados para diferenciar variedades, o facilitar el seguimiento de determinados alelos en programas de mejora (Cockram et al., 2012; Liu et al., 2012). Además, en el caso específico de trigo, se están implementando protocolos específicos para selección asistida por marcadores para genes de resistencia a hongos (royas, oídio, *Fusarium*, etc.), virus, insectos, calidad de grano (alto contenido de proteínas, contenido en amilosa y amilopectina, firmeza del gluten, color de la semolina, etc.), así como otros marcadores para caracteres agronómicos (genes de

enanismo y de vernalización). Es una herramienta en constante desarrollo, a la que se puede acceder en la dirección <http://maswheat.ucdavis.edu/Index.htm>.

La selección asistida con marcadores funciona bien para caracteres monogénicos, en retrocruzamientos, o en la piramidalización de unos pocos genes, como se ha hecho en arroz (Luo et al., 2013; Steele et al., 2013) o maíz (Yang et al., 2013).

Una vez que se han identificado SNPs específicos para un gen, o asociados a un determinado carácter, para su aplicación en mejora se está empezando a utilizar el sistema KASP (Kompetitive allele specific PCR), desarrollado por KBioscience y ofrecido por la empresa LGCGenomics. Este sistema permite genotipar a pequeña escala, con un diseño específico para cada polimorfismo. Es el sistema de genotipado utilizado por el CIMMYT de forma rutinaria, con un coste menor que los de otras plataformas (Semagn et al., 2014). Además, ha sido el sistema elegido para convertir SNPs en marcadores útiles para mejora en once cultivos, entre los que se encuentran arroz, maíz y trigo, en la iniciativa del Generation Challenge Programme.

En el caso de caracteres cuantitativos, controlados por muchos genes con efecto pequeño, este método es ineficaz. Una metodología nueva en plantas, que puede tener gran impacto en la mejora de cereales, es la selección genómica (GS, *Genome Selection*), (Heffner et al., 2009), combinando información fenotípica y genotípica. En selección genómica se estima el valor de un individuo basado en su información genética, teniendo en cuenta todos los marcadores a la vez y no sólo unos pocos como se hace en MAS. La base de la selección genómica está en la implementación de nuevas plataformas de genotipado, y en los nuevos métodos de secuenciación, capaces de generar un elevado número de marcadores que cubran todo el genoma. En selección genómica se utilizan dos poblaciones distintas, por un lado está la población de entrenamiento ("training"), que es genotipada a escala genómica con un número elevado de marcadores y fenotipada con detalle en ensayos repetidos para los caracteres de interés en el programa. La información derivada de la población de entrenamiento es la que se utiliza para seleccionar candidatos en función de su genotipo en la población de selección. El éxito de la selección genómica no sólo se basa en los avances genómicos, sino que también requiere del desarrollo de técnicas de fenotipado, fundamentalmente para los caracteres cuantitativos. Así, la fenómica (phenomics), tecnología de fenotipado a gran escala, es esencial para establecer estos modelos y llevar a cabo con éxito los procesos de GS (Topppa et al., 2013).

Aunque es una metodología reciente en plantas, ya se está aplicando en la práctica. Los primeros estudios en maíz y trigo utilizan, en la población de entrenamiento, 300-400 líneas genotipadas con 30.000-100.000 SNPs en combinación con unos 1300-1700 DArT (Poland et al., 2012b; Crossa et al., 2013, 2014; Heslot et al., 2013; Lado et al., 2013), a partir de los cuales hay que ajustar los modelos para predecir el potencial genético de distintos individuos, sin tener que evaluarlos fenotípicamente en campo.

Como queda patente en los puntos anteriores, se está avanzando mucho en el establecimiento de la conexión entre genotipo y fenotipo. Como consecuencia de los estudios de ligamiento llevados a cabo en numerosas hortalizas se dispone de muchos marcadores, basados en el propio gen o ligados a genes de resistencia a enfermedades causadas por virus, hongos o bacterias (Foolad y Panthee, 2012). Por ejemplo, en tomate se dispone de marcadores para la resistencia a numerosos virus, hongos, bacterias y nematodos (*Meloidogyne*). A pesar de su disponibilidad sólo algunos parecen utilizarse de forma rutinaria para la selección asistida por marcadores en los programas de mejora. Entre estos destacan la incorporación en material comercial de resistencia a *Fusarium oxysporum* f sp. *lycopersici* (genes *I*), a *Phytophthora infestans* (genes *Ph*), a *Cladosporium fulvum* (genes *Cf*), a *Leveillula taurina* (gen *Lv*), a *Verticillium* (genes *Ve*), a *Pseudomonas syringae* (genes *Pto*) o a *Xanthomonas* (genes *Xv*). Aunque en algunos casos ha sido necesario combinar la selección asistida por marcadores con una selección fenotípica para conseguir mayores niveles de resistencia. La selección más eficiente por marcadores ha sido la resistencia a virus, concretamente al Tobamovirus ToMV (Tomato Mosaic Virus), al geminivirus, TYLCV (Tomato Yellow Leaf Curl Virus) y al Tospovirus TSWV (Tomato Spotted Wilt Virus), permitiendo una selección no sólo más rápida, sino también más económica que el fenotipado por resistencia.

Situaciones similares a la de tomate se encuentran en otras hortalizas. Por ejemplo, en el caso del pimiento donde la selección asistida por resistencia a *Phytophthora* y *Colletotrichum* y a virus, como CMV (Cucumber Mosaic Virus), TMV (Tobacco Mosaic Virus), PepMoV (Pepper Mottle Virus) y TSWV (Tomato Spotted Wilt Virus), es posible, pero también por características de calidad, por ejemplo utilizando marcadores ligados al gen *Pun1*, involucrado en la síntesis de capsinoides, compuestos similares a los capsicinoides, pero que no pican (Garcés-Claver et al., 2007; Tanaka et al., 2014). En el caso del melón, existen marcadores de utilidad probada para resistencias a áfidos (gen *Vat*), al MNSV (Melon Necrotic Spotted Virus, gen *nsv*), a distintas razas de *Fusarium oxysporum* f *melonis* (genes *Fom*), y también a parámetros relacionados con la producción y la calidad, como la determinación sexual (gen α , andromonoecious) o el color de la carne (gen *Wf*) (Fernández-Trujillo et al., 2012; Díaz et al., 2011).

La selección asistida de estas características monogénicas ha sido más frecuente y eficiente para el desarrollo de nuevos cultivares que la de características de genética compleja, fundamentalmente por la falta de estabilidad de los QTLs al cambiar de población y por el arrastre de ligamiento de características desfavorables. Aunque la mejora para caracteres de calidad, tales como el sabor y el dulzor basada en los QTLs disponibles no ha sido demasiado exitosa, sí se han logrado selecciones adecuadas en la mejora de caracteres cuantitativos relacionados con la resistencia a enfermedades. Por ejemplo, en el caso de la resistencia a *Phytophthora*, donde existe la posibilidad de selección de resistencias monogénicas específicas de raza, se han identificado QTLs no específicos de raza en entradas de

Solanum habrochaites. Sin embargo, su uso en selección asistida no se ha llevado a cabo con éxito debido a que arrastran por ligamiento características desfavorables (como madurez tardía y excesivo tamaño de la planta), por lo que se requiere separar estos efectos previamente, por ejemplo mediante el desarrollo de NILs antes de utilizarlo. A pesar del esfuerzo dedicado a identificación de QTLs asociados a la mejora de la resistencia a estreses abióticos, salinidad, sequía o bajas temperaturas, todavía no se ha conseguido una mejora efectiva de estos caracteres basada en marcadores.

En relación con lo comentado anteriormente, el enorme esfuerzo invertido en la identificación de QTLs no ha podido ser suficientemente explotado, debido fundamentalmente a las limitaciones de los QTLs identificados empleando determinados tipos de poblaciones. Por ello, para poder llevar a cabo una mejora asistida por marcadores de características cuantitativas es esencial el empleo de poblaciones adecuadas. Las nuevas técnicas genómicas están permitiendo avanzar en el desarrollo de poblaciones que son a la vez herramientas de estudio de caracteres cuantitativos y herramientas de combinación de variantes genéticas para el desarrollo de nuevos cultivares. Entre estos tipos de poblaciones se encuentran las colecciones de líneas de introgresión, en las que, mediante un retrocruzamiento asistido con marcadores a escala genómica, se introducen en un genoma cultivado, a ser posible de una variedad de élite, introgresiones únicas de un genoma exótico. De esta forma, la colección completa de líneas representa el genoma exótico en su totalidad. La separación de segmentos genómicos individuales, minimiza los fenómenos de epistasia entre QTLs, que dificultan su detección y la introgresión de los mismos en fondos genéticos de élite, valida su expresión en los mismos. El empleo de este tipo de poblaciones está permitiendo mejorar la selección por marcadores de caracteres cuantitativos.

Entre las especies frutales la mejora asistida destaca en el manzano y el melocotonero, tanto por su importancia económica como por la mayor disponibilidad de recursos (mapas de ligamiento, secuencia del genoma y plataformas de SNPs), en ambas especies para resistencia a patógenos y para algunas características de calidad del fruto.

Para la resistencia al moteado en manzano, causado por el hongo *Venturia inaequalis*, se identificó el gen Vf posteriormente renombrado Rvi6, localizado en el GL1 donde se identificaron 3 genes candidatos del tipo RGA con un dominio LLR. Después se identificó un nuevo gen de resistencia como Rvi15 o Vr2 y se cartografió en una zona de 48,7 Kb, determinado por la presencia de 3 genes candidatos del tipo TIR-NBS-LRR (Galli et al., 2010). En la actualidad se han identificado hasta un total de 8 genes de resistencia alternativos al Vf por lo que se requiere la piramidización de genes mayores (Gessler y Pertot, 2012). Se dispone de un total de 14 SSRs organizados en 4 multiplex que permiten el análisis de 9 genes de resistencia mayores (Patocchi et al., 2009). A partir del genotipado con las plataformas de SNPs de RosBreed existen además SNPs disponibles para la mejora asistida de

resistencia al moteado, concretamente para los genes Rvi6 y Rvi2. Los marcadores que permiten seleccionar caracteres de fruto en manzano son la acidez mediante el SNP2 (ss475876558), textura crujiente, Crispness-SNP1 (ss475881704) y conservación post-cosecha (ACO_FEM_cg_4 ACO_FEM_cg_5). En melocotonero se dispone de marcadores ligados al carácter melocotón/nectarina, color de la pulpa y forma del fruto (Picañol et al., 2013), el tipo de fruto basado en polimorfismos en un intrón de la endoPG (Peace et al., 2005), SSRs asociados a alto contenidos en sorbitol, fenoles y azúcares (Font i Forcada et al., 2013), y una inserción en un exón del gen candidato ppa008301m que cosegrega con la fecha de maduración y permite desarrollar marcadores de selección (Pirona et al., 2013).

En los programas de mejora de patrones para melocotonero la resistencia a nematodos agalladores *Meloidogyne spp.* es uno de los objetivos prioritarios (Rubio-Cabezas et al., 2012). El gen Ma de resistencia a *M. arenaria* del ciruelo mirabolán (*P. cerasifera*) pertenece al grupo de genes TIR-NBS-LRR (Claverie et al., 2011) y está localizado en el GL7 donde se dispone de dos marcadores genómicos NSCALF2 (dentro del primer intron) y CT3-4N (dentro de la región C-terminal) (Esmenjaud y Dirlewanger 2007, Claverie et al., 2011). Otro gen RMia (R) confiere resistencia a *M. incognita* y *M. arenaria* que co-segregan en otros genotipos, 'Nemared', 'Shalil' y 'Juseitou' en el GL 2. Para este gen se dispone de 5 marcadores SSR específicos para 'Nemared'. Se ha cartografiado en una distancia de 92-kb que comprende 4 genes candidatos para el gen de la familia TNL (Duval et al., 2014). El polimorfismo de los tres marcadores más estrechamente ligados son dos SNP_APP92, SNP_APP91, y un SSR, y discriminan perfectamente las accesiones R y S, sugiriendo que al menos la fuente de R de 'Nemared', 'Nemaguard' y 'Shalil' comparte un ancestro común de resistencia.

Para el virus de la sharka sólo se han identificado fuentes de resistencia en albaricoquero. Aunque se identificó el locus PPVres (Vera et al., 2011), la comparación con la secuencia disponible del melocotonero permitió la obtención de marcadores ligados. En la actualidad se cuenta con 3 marcadores SSR: PGS1.21, PGS1.23 y PGS1.24 que muestran diferencias alélicas ligadas a la resistencia a sharka y que permiten realizar la selección precoz de las descendencias (Soriano et al., 2012). Por medio de mapeo posicional se han identificado genes candidatos (Zuriaga et al., 2013).

La autoincompatibilidad es otro carácter en el que se han invertido muchos esfuerzos y en el que se realiza mejora asistida por marcadores. Este carácter (SI) está controlado por un locus multialélico basado en la actividad de las RNAsas del pistilo. En el género *Prunus* la región codificante de la RNAsa está interrumpida por dos intrones que varían en tamaño, lo que permite identificar diferencias alélicas por secuenciación parcial de dichos intrones y posterior diseño de cebadores. Esta metodología ha permitido diferenciar por medio de PCR convencionales los distintos alelos y realizar mejora asistida para la selección de los alelos de autoincompatibilidad o para determinar la fertilidad de cruces intervarietales en los pro-

gramas de mejora de almendro (Fernandez i Marti et al., 2010); cerezo (Cachi y Wünsch, 2014), albaricoquero (Vilanova et al., 2005), manzano (Matsumoto et al., 2007), peral (Sanzol, 2009) y níspero (Gisbert et al., 2009).

4.5. Referencias

- Akbari M et al. 2006. Diversity arrays technology (DART) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* 113: 1409-1420.
- Almada R et al. 2013. Class 1 non-symbiotic and class 3 truncated hemoglobin-like genes are differentially expressed in stone fruit rootstocks (*Prunus* L.) with different degrees of tolerance to root hypoxia. *Tree Genet. Genomes* 9: 1051-1063.
- Aranzana M et al. 2003. A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. *Theor. Appl. Genet.* 106: 819-825.
- Asins MJ et al. 2004. QTL analysis of citrus tristeza virus-citradia interaction. *Theor. Appl. Genet.* 108: 603-611.
- Badenes ML et al. 2014. A peach germplasm collection for increasing the genetic diversity in European breeding programs. *Acta Hort.* (in press).
- Baird NA et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*: e3376.
- Bandillo N et al. 2013. Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. *Rice* 6: 1-15.
- Barchi L et al. 2011. Identification of SNP and SSR markers in eggplant using RAD tag sequencing. *BMC Genomics* 12: 304.
- Barchi L et al. 2012. A RAD tag derived marker based eggplant linkage map and the location of QTLs determining anthocyanin pigmentation. *PLOS One* 7: e43740.
- Bausher M. et al. 2006. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biol.* 6: 21.
- Belaj A et al. 2010. Genetic diversity and relationships of wild and cultivated olives at regional level in Spain. *Scientia Hort.* 124: 323-330.
- Bert PF et al. 2013. Mapping genetic loci for tolerance to lime-induced iron deficiency chlorosis in grapevine rootstocks (*Vitis* sp.). *Theor. Appl. Genet.* 126: 451-473.
- Bielenberg DG et al. 2008. Sequencing and annotation of the evergrowing locus in peach (*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genomes* 4: 495-507.
- Bielsa B et al. 2014. Detection of SNP and validation of a SFP InDel (deletion) in inverted repeat region of the *Prunus* species chloroplast genome. *Scientia Hort.* 168: 108-112.
- Blaker K. et al. 2013. Identification of QTLs controlling seed dormancy in peach (*Prunus persica*) *Tree Genet. Genomes* 9: 659-670.
- Blanca J et al. 2012. Transcriptome sequencing for SNP discovery across *Cucumis melo*. *BMC Genomics* 13: 280.
- Bogard M et al. 2013. Identifying wheat genomic regions for improving grain protein concentration independently of grain yield using multiple inter-related populations. *Mol. Breed.* 31: 587-599.

- Bordes J et al. 2014. Genome-wide association mapping of three important traits using bread wheat elite breeding populations. *Mol. Breed.* 33: 755-768.
- Bouhadida M et al. 2011. Genetic variability of introduced and local Spanish peach cultivars determined by SSRs markers. *Tree Genet. Genomes* 7: 257-270.
- Brand A et al. 2011. pc8.1, a major QTL for pigment content in pepper fruit is associated with variation in plastid compartment size. *Planta* 235: 579-588.
- Brenchley R et al. 2012. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491: 705-710.
- Buckler ES et al. 2009. The genetic architecture of maize flowering time. *Science* 325: 714-718.
- Cabezas JA et al. 2006. A genetic analysis of seed and berry weight in grapevine. *Genome* 49: 1572-1585.
- Cachi AM, Wünsch A. 2014. S-genotyping of sweet cherry varieties from Spain and S-locus diversity in Europe. *Euphytica* 197: 229-236.
- Calenge F, Durel CE. 2006. Both stable and unstable QTLs for resistance to powdery mildew are detected in apple after four years of field assessments. *Mol. Breed.* 17: 329-339.
- Campoy JA et al. 2011. Inheritance of flowering time in apricot (*Prunus armeniaca* L.) and analysis of linked quantitative trait loci (QTL) using simple sequence repeat (SSR) markers. *Plant Mol. Biol. Rep.* 29: 404-410.
- Causse M et al. 2013. Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genomics* 14: 791.
- Cavanagh CR et al. 2008. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* 11: 215-221.
- Cavanagh CR et al. 2013. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci. USA* 110: 8057-8062.
- Celton JM et al. 2011. Deciphering the genetic determinism of bud phenology in apple progenies: a new insight into chilling and heat requirement effects on flowering dates and positional candidate genes. *New Phytol.* 192: 378-392.
- Chagné D et al. 2012. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS One* 7(2): e31745.
- Chao S et al. 2010. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). *BMC Genomics* 11: 727.
- Chen A L et al. 2014. Reassessment of QTLs for late blight resistance in the tomato accession L3708 using a restriction site associated DNA (RAD) linkage map and highly aggressive isolates of *Phytophthora infestans*. *PLoS one* 9: e96417.
- Chen TH et al. 2002. Molecular genetic analysis of dormancy-related traits in poplars. *Weed Sci.* 50: 232-240.
- Chia JM et al. 2012. Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* 44: 803-807.
- Christianson JA et al. 2010. Comparisons of early transcriptome responses to low-oxygen environments in three dicotyledonous plant species. *Plant Signal. Behav.* 5: 1006-1009.
- Claverie M et al. 2011. The Ma gene for complete-spectrum resistance to *Meloidogyne* species in *Prunus* is a TNL with a huge repeated C-terminal post-LRR region. *Plant Physiol.* 156: 779-792.
- Close TJ et al. 2009. Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics* 10: 582.

- Cockram J et al. 2012. Evaluation of diagnostic molecular markers for DUS phenotypic assessment in the cereal crop, barley (*Hordeum vulgare* ssp. *vulgare* L.). *Theor. Appl. Genet.* 125: 1735-1749.
- Comadran J et al. 2009. Patterns of genetic diversity and linkage disequilibrium in a highly structured *Hordeum vulgare* association-mapping population for the Mediterranean basin. *Theor. Appl. Genet.* 119: 175-187.
- Comadran J et al. 2012. Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44: 1388-1392.
- Crossa J et al. 2007. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 177: 1889-1913.
- Crossa J et al. 2013. Genomic prediction in maize breeding populations with genotyping-by-sequencing. *G3* 3: 1903-1926.
- Crossa J, et al. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity* 112: 48-60.
- Cuppen E. 2007. Genotyping by allele-specific amplification (KASPar). *Cold Spring Harb Protoc.* 4841.
- den Boer E et al. 2013. Fine mapping quantitative resistances to downy mildew in lettuce revealed multiple sub-QTLs with plant stage dependent effects reducing or even promoting the infection. *Theor. Appl. Genet.* 126: 2995-3007.
- Deschamps et al. 2012. Genotyping by sequencing in plants. *Biology* 1: 460-483.
- Diaz A et al. 2011. A consensus linkage map for molecular markers and Quantitative Trait Loci associated with economically important traits in melon (*Cucumis melo* L.). *BMC Plant Biol.* 11: 111.
- Dirlewanger E et al. 2002. Development of microsatellite markers in peach and their use in genetic diversity analysis in peach and sweet cherry. *Theor. Appl. Genet.* 105: 127-138.
- Dirlewanger E et al. 2012. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity* 109: 280-292.
- Donoso JM et al. 2012. Towards the enrichment of the peach gene pool with alleles from wild species: the collection of peach-almond introgression lines. 6th Int. Rosaceae Genomics Conf., pp40.
- Duran C et al. 2010. Future tools for association mapping in crop plants. *Genome* 53: 1017-1023.
- Duval H. et al. 2014. High-resolution mapping of the *RMia* gene for resistance to root-knot nematodes in peach. *Tree Genet. Genomes* 10: 297-306.
- Dwivedi N et al. 2013. QTL mapping for important horticultural traits in pepper (*Capsicum annuum* L.). *J. Plant Biochem. Biotechnol.* (DOI: 10.1007/s13562-013-0247-1)
- Eduardo I et al. 2011. QTL analysis of fruit quality traits in two peach intraspecific populations and importance of maturity date pleiotropic effect. *Tree Genet. Genomes* 7: 323-335.
- Edwards D et al. 2014. New technologies for ultrahigh-throughput genotyping in plant taxonomy. *Methods Mol. Biol.* 1115:151-175.
- Elshire RJ et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6: e19379.
- Esmenjaud D, Dirlewanger E. 2007. Plum. En: Kole CR (ed). *Genome mapping and molecular breeding. Fruits & nuts*, vol 4. Springer, Heidelberg, Alemania, pp 119-136.
- Esteras C et al. 2012. High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping *BMC Genomics* 13: 80.
- Esteras C et al. 2013. SNP genotyping in melons: genetic variation, population structure, and linkage disequilibrium. *Theor. Appl. Genet.* 126: 1285-1303.

- Fabbrini F et al. 2012. Phenotypic plasticity, QTL mapping and genomic characterization of bud set in black poplar. *BMC Plant Biol.* 12: 47.
- Famoso AN et al. 2011. Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS Genet* 7: e1002221.
- Fan S et al. 2010. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytol.* 185: 917-930.
- Fei Z et al. 2014. Genome sequencing of *Cucurbita* species Cucurbitaceae 2014, 12-16 October, Bay Harbor, Michigan, EE UU.
- Fernández i Martí A et al. 2010. The almond *S_r* haplotype shows a double expression despite its comprehensive genetic identity. *Scientia Hort.* 125: 685-691.
- Fernández i Martí A et al. 2013. Genetic analysis for physical nut traits in almond. *Tree Genet. Genomes* 9: 455-465.
- Fernández i Martí A et al. 2014. Genomic analysis of the molecular evolution and the population structure of a worldwide almond pool assessed by microsatellite markers. *Genet. Resour. Crop Evol.* (aceptado, en prensa).
- Fernández-Trujillo JP et al. 2011. Breeding for fruit quality in melon. En: Jenks MA, Bebeli P (eds.). *Breeding for Fruit Quality.* Wiley-Blackwell, Ames, IA, EE UU.
- Fisk SP et al. 2013. FR-H3: a new QTL to assist in the development of fall-sown barley with superior low temperature tolerance. *Theor. Appl. Genet.* 126: 335-347.
- Font i Forcada C et al. 2012. Mapping quantitative trait loci for kernel composition in almond. *BMC Genetics* 13: 47.
- Font i Forcada et al. 2013. Population structure and marker-trait associations for pomological traits in peach and nectarine cultivars. *Tree Genet. Genomes* 9: 331-349.
- Foolad MR, Panthee DR. 2012. Marker-assisted selection in tomato breeding. *Crit. Rev. Plant Sci.* 31: 93-123.
- Frary A et al. 2014. QTL hotspots in eggplant (*Solanum melongena*) detected with a high resolution map and CIM analysis. *Euphytica* 197: 211-228.
- Frewen BE et al. 2000. Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics* 154: 837-845.
- Gabriel S et al. 2009. SNP genotyping using the Sequenom MassARRAY iPLEX platform. En: Haines JL (ed.) *Current Protocols in Human Genetics*, unit 2.12.
- Galal A et al. 2014. Comparative QTL analysis of root lesion nematode resistance in barley. *Theor. Appl. Genet.* 127: 1399-1407.
- Galli P et al. 2010. The Rvi15 (Vr2) apple scab resistance locus contains three TIR-NBS-LRR genes. *Mol. Plant-Microbe Interactions* 23: 608-617.
- Ganal MW et al 2012. Large SNP arrays for genotyping in crop plants. *J. Biosci.* 37: 821-828.
- Ganal MW et al. 2011. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PloS One* 6: e28334.
- Gao Q et al 2012. Recent progress using high-throughput sequencing technologies in plant molecular breeding. *J. Integrat. Plant Biol.* 54: 215-227.
- Garcés-Claver A et al. 2007. Identification, validation and survey of a single nucleotide polymorphism (SNP) associated with pungency in *Capsicum* spp. *Theor. Appl. Genet.* 115: 907-916.
- García-Mas J et al. 2012. The genome of melon (*Cucumis melo* L.) *Proc. Natl. Acad. Sci. USA* 109: 11872-11877.

- Gardiner SE et al. 2012a. Putative resistance gene markers associated with quantitative trait loci for fire blight resistance in *Malus* "Robusta 5_ accessions. *BMC Genomics* 13: 25.
- Gardiner SE et al. 2012b. A draft sequences of European pear (*Pyrus communis* L. 'Barlett) 6th International Rosaceae Genomics Conference 2012, pp 21.
- Gasic K et al. 2009. Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. *Mol. Breed.* 23: 397-411.
- Gessler C, Pertot I. 2012. Vf scab resistance of *Malus*. *Trees* 26: 95-108.
- Gilchrist E J et al. 2013. A mutant *Brassica napus* (canola) population for the identification of new genetic diversity via TILLING and next generation sequencing. *PLoS One* 8: e84303.
- Gisbert et al. 2009. Genetic diversity evaluation of a loquat [*Eriobotrya japonica* (Thun.) Lindl.] germplasm collection by SSR and S-allele fragments. *Euphytica* 168: 121-134.
- Gonzalo MJ et al. 2012. Genetic analysis of iron chlorosis tolerance in *Prunus* rootstocks. *Tree Genet. Genomes* 8: 943-955.
- Gore MA et al. 2009. A first-generation haplotype map of maize. *Science* 326: 1115-1117.
- Gross BL et al. 2012. Identification of interspecific hybrids among domesticated apple and its wild relatives. *Tree Genet. Genomes* 8: 1223-1235.
- Gross BL, Olsen KM. 2010. Genetic perspectives on crop domestication. *Trends Plant Sci.* 15: 529-537.
- Guo S et al. 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* 45: 51-58.
- Gur A, Zamir D. 2004. Unused natural variation can lift yield barriers in plant breeding. *PLoS Biology* 2: e245.
- Hajjar R, Hodgkin T. 2007. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156: 1-13.
- He T et al. 2007. Using SSR markers to determine the population genetic structure of wild apricot (*Prunus armeniaca* L.) in the Ily Valley of West China. *Genet. Resour. Crop Evol.* 54: 563-572.
- He X et al. 2013. QTL mapping of powdery mildew resistance in WI 2757 cucumber (*Cucumis sativus* L.). *Theor. Appl. Genet.* 126: 2149-2161.
- Heffner E L et al. 2009. Genomic selection for crop improvement. *Crop Sci.* 49: 1-12.
- Hernandez P et al. 2012. Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. *Plant J.* 69: 377-386.
- Heslot N et al. 2013. Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. *PLoS One* 8: e74612..
- Hill TA et al. 2013. Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30K unigene pepper gene chip. *PLoS One* 8: e56200.
- Hirakawa H et al. 2013. Genome-wide SNP genotyping to infer the effects on gene functions in tomato. *DNA Res.* 20: 221-233.
- Hirsch CD et al. 2014. Reduced representation approaches to interrogate genome diversity in large repetitive plant genomes. *Brief. Funct. Genomics* (DOI: 10.1093/bfgp/elt051)
- Hofmann K et al. 2013. Fine mapping of the *Rrs1* resistance locus against scald in two large populations derived from Spanish barley landraces. *Theor. Appl. Genet.* 126: 3091-3102.
- Holloway B, Li B. 2010. Expression QTLs: applications for crop improvement. *Mol. Breed.* 26: 381-391.

- Hormaza JI. 2002. Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. *Theor. Appl. Genet.* 104: 321-328.
- Huang S et al. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41: 1275-1281.
- Huang X et al. 2009. High-throughput genotyping by whole-genome resequencing. *Genome Res.* 19: 1068-1076.
- Huang X et al. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42: 961-967.
- Huang X et al. 2012. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44: 32-39.
- Huang Y et al. 2013. Haplotype diversity and population structure in cultivated and wild barley evaluated for *Fusarium* head blight responses. *Theor. Appl. Genet.* 126: 619-636.
- Hung HY et al. 2012. *ZmCCT* and the genetic basis of day-length adaptation underlying the post domestication spread of maize. *Proc. Natl. Acad. Sci. USA* 109: E1913-E1921.
- Hyten D et al. 2010. High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics* 11: 38.
- IBSC Consortium 2012. A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491: 711-716.
- IRGSP. 2005. The map-based sequence of the rice genome. *Nature* 436: 793-800.
- Islamovic E et al. 2014. Quantitative trait loci of barley malting quality trait components in the Stellar/01Ab8219 mapping population. *Mol. Breed.* 34: 59-73.
- Jain M et al. 2014. Genome-wide discovery of DNA polymorphisms in rice cultivars with contrasting drought and salinity stress response and their functional relevance. *Plant Biotechnol. J.* 12: 253-264.
- Jiao Y et al. 2012. Genome-wide genetic changes during modern breeding of maize. *Nat. Genet.* 44: 812-815.
- Joukhadar R et al. 2013. Genome-wide association mapping for five major pest resistances in wheat. *Mol. Breed.* 32: 943-960.
- Kenis K et al. 2008. Identification and stability of QTLs for fruit quality traits in apple. *Tree Genet. Genomes* 4: 647-661.
- Kilian J et al. 2007. The AtGenExpress global stress expression dataset: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J.* 55: 526-542.
- Kim S et al. 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nat. Genet.* 46: 270-278.
- Kinkade MP, Foolad MR. 2013. Validation and fine mapping of lyc12.1, a QTL for increased tomato fruit lycopene content. *Theor. Appl. Genet.* 126: 2163-2175.
- Kobayashi M et al. 2014. Genome-wide analysis of intraspecific DNA polymorphism in 'Micro-Tom', a model cultivar of tomato (*Solanum lycopersicum*). *Plant Cell Physiol.* 55: 445-454.
- Koepke T. et al. 2013. Comparative genomics analysis in Prunoideae to identify biologically relevant polymorphisms. *Plant Biotechnol. J.* 11: 883-893.
- Kollers S et al. 2013a. Genetic architecture of resistance to *Septoria tritici* blotch (*Mycosphaerella graminicola*) in European winter wheat. *Mol. Breed.* 32: 411-423.
- Kollers S et al. 2013b. Whole genome association mapping of *Fusarium* head blight resistance in European winter wheat (*Triticum aestivum* L.). *PLoS One* 8: e57500.

- Kulwal P et al. 2012. Association mapping for pre-harvest sprouting resistance in white winter wheat. *Theor. Appl. Genet.* 125: 793-805.
- Kumar S et al. 2012. Genomic selection for fruit quality traits in apple (*Malus × domestica* Borkh.) *PLoS One* 7: e36674.
- Kump KL et al. 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* 43: 163-168.
- Kwon SJ et al. 2013. Genome-wide association of 10 horticultural traits with expressed sequence tag-derived SNP markers in a collection of lettuce lines. *Crop J.* 1: 25-33.
- Lado B et al. 2013. Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3* 3: 2105-2114.
- Lai J et al. 2010. Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat. Genet.* 42: 1027-1030.
- Lalli DA et al. 2005. Identification and mapping of resistance gene analogs (RGA) in *Prunus*: a resistance map for *Prunus*. *Theor. Appl. Genet.* 111: 1504-1513.
- Lam HM et al. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42: 1053-1059.
- Lazzari B et al. 2008. Version VI of the ESTree db: an improved tool for peach transcriptome analysis. *BMC Bioinformatics* 9: S9.
- Leida C et al. 2010. Identification of genes associated with bud dormancy release in *Prunus persica* by suppression subtractive hybridization. *Tree Physiol.* 30: 655-666.
- Letta T et al. 2013. Searching for novel sources of field resistance to Ug99 and Ethiopian stem rust races in durum wheat via association mapping. *Theor. Appl. Genet.* 126: 1237-1256.
- Lewers K et al. 2008. A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biol.* 8: 69.
- Li A et al. 2013. Identification and fine mapping of *qRBSDV-6^{MH}*, a major QTL for resistance to rice black-streaked dwarf virus disease. *Mol. Breed.* 32: 1-13.
- Li H et al. 2013. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45: 43-50
- Li X et al. 2014. Development of an alfalfa SNP array and its use to evaluate patterns of population structure and linkage disequilibrium *PLoS One.* 9: e84329.
- Liebhart R et al. 2003. Mapping quantitative physiological traits in apple (*Malus × domestica* Borkh.). *Plant Mol. Biol.* 52: 511-526.
- Liu S et al. 2012. Changes in genome content generated via segregation of non-allelic homologs. *Plant J.* 72: 390-399.
- Liu T et al. 2013. Validation and characterization of *Ghd7. 1*, a major QTL with pleiotropic effects on spikelets per panicle, plant height, and heading date in rice (*Oryza sativa* L.). *J. Integrative Plant Biol.* 55: 917-927.
- Liu Y et al. 2012. Functional markers in wheat: current status and future prospects. *Theor. Appl. Genet.* 125: 1-10.
- Liu Y et al. 2014. Genetic analysis and major QTL detection for maize kernel size and weight in multi-environments. *Theor. Appl. Genet.* 127: 1019-1037.
- Lopes M S et al. 2013. QTL for yield and associated traits in the Seri/Babax population grown across several environments in Mexico, in the West Asia, North Africa, and South Asia regions. *Theor. Appl. Genet.* 126: 971-984.

- Lopez-Vera EE et al. 2014. Resistance to stem rust Ug99 in six bread wheat cultivars maps to chromosome 6DS. *Theor. Appl. Genet.* 127: 231-239.
- Lu H et al. 2014. QTL-seq identifies an early flowering QTL located near flowering locus T in cucumber. *Theor. Appl. Genet.* 127: 1-9.
- Luo Y et al. 2013. Marker-assisted breeding of Thai fragrance rice for semi-dwarf phenotype, submergence tolerance and disease resistance to rice blast and bacterial blight. *Mol. Breed.* 32: 709-721.
- Luro F L. et al. 2008. Transferability of the EST-SSRs developed on Nules clementine (*Citrus clementina* Hort ex Tan) to other *Citrus* species and their effectiveness for genetic mapping *BMC Genomics* 9: 287.
- Lv J et al. 2012. Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). *PLoS One* 7: e46919.
- Maccaferri M et al. 2011. Association mapping in durum wheat grown across a broad range of water regimes. *J. Exp. Bot.* 62: 409-438.
- Manganaris G A et al. 2011. Comparative transcript profiling of apricot (*Prunus armeniaca* L.) fruit development and on-tree ripening. *Tree Genet. Genomes* 7: 609-616.
- Mantovani P et al. 2008. An integrated DArT-SSR linkage map of durum wheat. *Mol. Breed.* 22: 629-648.
- Marandel G et al. 2009. Quantitative trait loci metaanalysis of Plum pox virus resistance in apricot (*Prunus armeniaca* L.): new insights on the organization and the identification of genomic resistance factors. *Mol. Plant. Pathol.* 10: 347-360.
- Maron LG et al. 2013. Aluminum tolerance in maize is associated with higher *MATE1* gene copy number. *Proc. Natl. Acad. Sci. USA* 110: 5241-5246.
- Martin LB et al. 2013. Catalyzing plant science research with RNA-seq. *Front. Plant Sci.* 4: 66.
- Martinez-Garcia P et al. 2013. High density SNP mapping and QTL analysis for fruit quality characteristics in peach (*Prunus persica* L.). *Tree Genet. Genomes* 9: 19-36.
- Mascher M et al. 2013. Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. *Plant J.* 76: 494-505.
- Matsumoto S. et al. 2007. Determination and confirmation of S-RNase genotypes of apple pollinators and cultivars. *J. Hort. Sci. Biotechnol.* 82: 323-329.
- Matsumura H et al. 2014. Mapping of the gynoecey in bitter melon (*Momordica charantia*) using RAD-seq analysis. *PLoS One* 9: e87138.
- Mayer KFX et al. 2011. Unlocking the barley genome by chromosomal and comparative genomics. *Plant Cell* 23: 1249-1263.
- McMullen MD et al. 2009. Genetic properties of the maize nested association mapping population. *Science* 325: 737-740.
- Metzker ML. 2010. Sequencing technologies-the next generation. *Nat. Rev. Genet.* 11: 31-46.
- Miller MR et al. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17: 240-248.
- Ming R et al. 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452: 991-996.
- Mizobuchi R et al. 2013. Identification of *qRBS1*, a QTL involved in resistance to bacterial seedling rot in rice. *Theor. Appl. Genet.* 126: 2417-2425.
- Morrell PL et al. 2011. Crop genomics: advances and applications. *Nat. Rev. Genet.* 13: 85-96.

- Morris GP et al. 2013. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci. USA* 110: 453-458.
- Muñoz-Amatriaín M et al. 2014. The USDA barley core collection: genetic diversity, population structure, and potential for genome-wide association studies. *PLoS One* 9: e94688.
- Muñoz-Amatriaín M et al. 2011. An improved consensus linkage map of barley based on flow-sorted chromosomes and single nucleotide polymorphism markers. *Plant Genome* 4: 238-249.
- Myles S et al. 2011. Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. USA*. 108: 93530-93535.
- Naval M et al. 2010. Analysis of genetic diversity among persimmon cultivars using microsatellite markers *Tree Genet. Genomes* 6: 677-687.
- Neale D et al. 2014. Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies *Genome Biol.* 15: R59.
- Niedringhaus TP et al. 2011. Landscape of next-generation sequencing technologies. *An. Chem.* 83: 4327-4341.
- Noelle A et al. 2006. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theor. Appl. Genet.* 112: 1519-1531
- Norelli JL et al. 2009. Rapid transcriptional response of apple to fire blight disease revealed by cDNA suppression subtractive hybridisation analysis. *Tree Genet. Genomes* 5: 27-40.
- Nybom H et al. 2014. DNA fingerprinting in botany: past, present, future. *Invest. Genet.* 5: 1.
- Ogundiwin E et al. 2008. Development of ChillPeach genomic tools and identification of cold-responsive genes in peach fruit. *Plant Mol. Biol.* 68: 379-397.
- Olukolu BA et al. 2009. Genetic linkage mapping for molecular dissection of chilling requirement and bud break in apricot (*Prunus armeniaca* L.). *Genome*, 52: 819-828.
- Patocchi A et al. 2009. Towards improvement of marker assisted selection of apple scab resistant cultivars: *Venturia inaequalis* virulence surveys and standarization of molecular markers associated with resistance genes. *Mol. Breed.* 24: 337-347.
- Paux E et al. 2008. A physical map of the 1-Gigabase bread wheat chromosome 3B. *Science* 322: 101-104.
- Peace CP et al. 2005. Endopolygalacturonase: a candidate gene for freestone and melting flesh in peach. *Mol. Breed.* 16: 21-31.
- Peiffer JA et al. 2014. The genetic architecture of maize height. *Genetics* 196: 1337-1356.
- Peil A et al. 2007. Strong evidence for a fire blight resistance gene of *Malus robusta* located on linkage group 3. *Plant Breed.* 126: 470-475.
- Pérez-de-Castro AM et al. 2012. Application of genomic tools in plant breeding. *Curr. Genomics* 13: 179-195.
- Perkel J 2008. SNP genotyping: six technologies that keyed a revolution. *Nat. Methods* 5: 447-453.
- Picañol R et al. 2013. Combining linkage and association mapping to search for markers linked to the flat fruit character in peach. *Euphytica* 190: 279-288.
- Pierantoni L et al. 2007. Pear scab resistance QTLs via a European pear (*Pyrus communis*) linkage map. *Tree Genet. Genomes* 3: 311-317.
- Pina A et al. 2014. Analysis of the genetic diversity of local apple cultivars from mountainous areas from Aragón (Northeastern Spain). *Scientia Hort.* 174: 1-9.
- Pirona R et al. 2013. Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. *BMC Plant Biol.* 13: 166.

- Poland JA et al. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc. Natl. Acad. Sci. USA* 108: 6893-6898.
- Poland JA et al. 2012a. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 7: e32253.
- Poland JA et al. 2012b. Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5: 103-113.
- Potts S et al. 2014. Identification of quantitative trait loci (QTLs) for fruit quality traits in apple. *Plant Mol. Biol. Rep.* 32: 109-116.
- Powell A et al. 2012. Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science* 336: 1711-1715.
- Qi J et al. 2013. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat. Genet.* 45: 1510-1515.
- Quilot B et al. 2004. QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild relative species *P. davidiana*. *Theor. Appl. Genet.* 109: 884-897.
- Rafalski J A et al. 2011. Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13: 174-180.
- Raman H et al. 2014. SNP markers-based map construction and genome-wide linkage analysis in *Brassica napus*. *Plant Biotechnol. J.* (DOI: 10.1111/pbi.12186).
- Rambla JL et al. 2014. The expanded tomato fruit volatile landscape. *J. Exp. Bot.* (DOI: 10.1093/jxb/eru128).
- Ranc N et al. 2012. Genome-wide association mapping in tomato (*Solanum lycopersicum*) is possible using genome admixture of *Solanum lycopersicum* var. *Cerasiforme*. *G3* 2: 853-864.
- Ravel C et al. 2013. Improving the yellow pigment content of bread wheat flour by selecting the three homoeologous copies of *Psy1*. *Mol. Breed.* 31: 87-99.
- Rebetzke GJ et al. 2014. Use of a large multiparent wheat mapping population in genomic dissection of coleoptile and seedling growth. *Plant Biotechnol. J.* 12: 219-230.
- Reddy UK et al. 2014. Identification of gene-specific polymorphisms and association with capsaicin pathway metabolites in *Capsicum annuum* L. collections. *PLoS One* 9: e86393.
- Ren Y et al. 2014. An integrated genetic map based on four mapping populations and quantitative trait loci associated with economically important traits in watermelon (*Citrullus lanatus*). *BMC Plant Biol.* 20: 14:33.
- Richards CM et al. 2009. Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genet. Genomes* 5: 339-347.
- Romay MC et al. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14: R55.
- Romeu JF et al. 2014. Quantitative trait loci affecting reproductive phenology in peach. *BMC Plant Biol.* 14: 52.
- Roy JK et al. 2010. Association mapping of spot blotch resistance in wild barley. *Mol. Breed.* 26: 243-256.
- Rubio-Cabetas MJ. 2012. Present and future trends in peach rootstock breeding worldwide. *Acta Hort.* 962: 81-89.
- Rubio-Cabetas MJ et al. 2010. A microarray analysis revealed and oxidative response genes underlying the differential response to hypoxia of two *Prunus* genotypes. 5th International RGC. 042.
- Saintenac C et al. 2013. Sequence-based mapping of the polyploid wheat genome. *G3* 3: 1105-1114.

- Salvi S, Tuberosa R. 2005. To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* 10: 297-304.
- Sánchez G et al. 2014. The peach volatilome modularity is reflected at the genetic and environmental response levels in a QTL mapping population *BMC Plant Biol.* 14: 137.
- Sánchez-Pérez R et al. 2010. Molecular markers for kernel bitterness in almond. *Tree Genet Genomes* 6: 237-245.
- Sánchez-Pérez R et al. 2012. Inheritance of chilling and heat requirements for flowering in almond and QTL analysis. *Tree Genet. Genomes* 8: 379-389.
- Sanzol J 2009. Genomic characterization of self-incompatibility ribonucleases in European pear cultivars and development of PCR detection for 20 alleles. *Tree Genet. Genomes* 5: 393-405.
- Scaglione D et al. 2012. RAD tag sequencing as a source of SNP markers in *Cynara cardunculus* L. *BMC Genomics* 13: 3.
- Schnable PS et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112-1115.
- Schnaithmann F et al. 2014. A first step toward the development of a barley NAM population and its utilization to detect QTLs conferring leaf rust seedling resistance. *Theor. Appl. Genet.* (on line).
- Schneeberger K, Weigel D. 2011. Fast-forward genetics enabled by new sequencing technologies *Trends Plant Sci.* 16: 282-288.
- Seipp MT et al. 2009. Multiplex amplicon genotyping by high-resolution melting. *J. Biomol. Tech.* 20: 160-164.
- Semagn K et al. 2006. Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. *Genome* 49: 545-555.
- Semagn K et al. 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. *Mol. Breed.* 33: 1-14.
- Septiningsih EM et al. 2009. Development of submergence-tolerant rice cultivars: the *Sub1* locus and beyond. *Ann. Bot.* 103: 151-160.
- Sharpe A et al. 2013. Ancient orphan crop joins modern era: gene-based SNP discovery and mapping in lentil *BMC Genomics* 14: 192.
- Shen GQ et al. 2009. The TaqMan method for SNP genotyping. *Methods Mol. Biol.* 578: 293-306.
- Shendure J, Ji H. 2008. Next-generation DNA sequencing. *Nat Biotechnol.* 26: 1135-1145.
- Shulaev V et al. 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* 43: 109-116.
- Silfverberg-Dilworth et al. 2006. Microsatellite markers spanning the apple (*Malus × domestica* Borkh.) genome. *Tree Genet. Genomes* 2: 202-224.
- Sim SC et al. 2012a. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS One* 7: e40563.
- Sim SC et al. 2012b. High-density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding. *PLoS One* 7: e45520.
- Socquet-Juglard D et al. 2013. Mapping architectural, phenological, and fruit quality QTLs in apricot. *Plant Mol. Biol. Rep.* 31: 387-397.
- Sonah H et al. 2013. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS One* 8: e54603.
- Song Q et al. 2013. Development and evaluation of soy SNP50K, a high-density genotyping array for soybean. *PLoS One* 8: e54985.

- Soriano JM et al. 2012. Identification of SSR markers tightly associated to PPV resistance in apricot. *Mol. Breed.* 30: 1017-1026.
- Spindel J et al. 2013. Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor. Appl. Genet.* 126: 2699-2716.
- Steele KA et al. 2013. QTLs associated with root traits increase yield in upland rice when transferred through marker-assisted selection. *Theor. Appl. Genet.* 126: 101-108.
- Szűcs P et al. 2009. An integrated resource for barley linkage map and malting quality QTL alignment. *Plant Genome* 2: 134-140.
- Tanaka Y et al. 2014. Application of marker-assisted selection in breeding of a new fresh pepper cultivar (*Capsicum annuum*) containing capsinoids, low-pungent capsaicinoid analogs. *Scientia Hort.* 165: 242-245.
- Tang H et al. 2010. Domestication and plant genomes. *Curr. Opin. Plant Biol.* 13: 160-166.
- Teng F et al. 2013. *ZmGA3ox2*, a candidate gene for a major QTL, *qPH3. 1*, for plant height in maize. *Plant J.* 73: 405-416.
- Thomson MJ et al. 2012. High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. *Mol. Breed.* 29: 875-886.
- Tian F et al. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43: 159-162.
- Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485: 635-641.
- Toppa CN et al. 2013. 3D phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. *Proc. Natl. Acad. Sci. USA* 30: E1695-E1704.
- Trainotti L et al. 2006. The use of microarray uPEACH1. to investigate transcriptome changes during transition from preclimacteric to climacteric phase in peach fruit. *Plant Sci.* 170: 606-613.
- Trebbi D et al. 2011. High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theor. Appl. Genet.* 123: 555-569.
- Truco MJ et al. 2013. An ultra-high-density, transcript-based, genetic map of lettuce. *G3* 3: 4617-4631.
- Tuskan GA et al. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596-1604.
- Uga Y et al. 2013. Control of root system architecture by *DEEPER ROOTING 1* increases rice yield under drought conditions. *Nat. Genet.* 45: 1097-1102.
- van Poecke RMP et al. 2013. Sequence-based SNP genotyping in durum wheat. *Plant Biotechnol. J.* 11: 809-817.
- Vegas J et al. 2013. Interaction between QTLs induces an advance in ethylene biosynthesis during melon fruit ripening. *Theor Appl Genet.* 126:1531-1544.
- Velasco R et al. 2010. The genome of the domesticated apple (*Malus × domestica* Borkh.) *Nat. Genet.* 42: 833-839.
- Velasco R et al. 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS One* 2: e1326.
- Venuprasad R et al. 2012. A large-effect QTL for rice grain yield under upland drought stress on chromosome 1. *Mol. Breed.* 30: 535-547.

- Vera EM et al. 2011. Narrowing down the apricot Plum pox virus resistance locus and comparative analysis with the peach genome syntenic region. *Mol. Plant Pathol.* 12: 535-547.
- Verde I et al. 2012. Correction: development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. *PLoS One* 7: e35668.
- Verde I et al. 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet.* 45: 487-496.
- Vilanova S et al. 2005. Identification of self(in)compatibility alleles in apricot by PCR sequence analysis. *J. Amer. Soc. Hort. Sci.* 130: 893-898.
- Vives C et al. 2012. Aplicación de marcadores moleculares a la mejora genética del melocotonero. Ms. Thesis. Instituto Agronómico Mediterráneo, Zaragoza (Spain).
- Wan Y et al. 2008. The eco-geographic distribution of wild grape germplasm in China. *Vitis* 47: 77-80.
- Wan Y et al. 2013. A phylogenetic analysis of the grape genus (*Vitis* L.) reveals broad reticulation and concurrent diversification during neogene and quaternary climate change. *BMC Evolut. Biol.* 13: 141.
- Wang S et al. 2014. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* (online).
- Wang S et al. 2012. 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nat. Methods* 9: 808-810.
- Wenzl P et al. 2004. Diversity arrays technology (DART) for whole-genome profiling of barley. *Proc. Natl. Acad. Sci. USA* 101: 9915-9920.
- White J et al. 2008. The genetic diversity of UK, US and Australian cultivars of *Triticum aestivum* measured by DART markers and considered by genome. *Theor. Appl. Genet.* 116: 439-453.
- Willcox MC et al. 2013. Confirming quantitative trait loci for aflatoxin resistance from Mp313E in different genetic backgrounds. *Mol. Breed.* 32: 15-26.
- Winfield MO et al. 2012. Targeted re-sequencing of the allohexaploid wheat exome. *Plant Biotechnol. J.* 10: 733-742.
- Wu J et al. 2013. The genome of the pear (*Pyrus bretschneideri* Rehd.) *Genome Res.* 23: 396-408.
- Wu X et al. 2014. Fine genetic characterization of elite maize germplasm using high-throughput SNP genotyping. *Theor. Appl. Genet.* 127: 621-631.
- Wu G A et al. 2014. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat. Biotechnol.* (on line).
- Würschum T et al. 2013. Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theor. Appl. Genet.* 126: 1477-1486.
- Xu X et al. 2012. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30: 105-111.
- Xu J et al. 2013. Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species. *Theor. Appl. Genet.* 126:567-581.
- Xu J et al. 2014. Identification of candidate genes for drought tolerance by whole-genome resequencing in maize. *BMC Plant Biol.* 14: 83.
- Xu Q et al. 2013. The draft genome of sweet orange (*Citrus sinensis*) *Nat. Genet.* 45: 59-66.
- Xu X et al. 2011. Genome sequence and analysis of the tuber crop potato. *Nature* 475: 189-195.
- Xue Y et al. 2013. Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. *Theor. Appl. Genet.* 126: 2587-2596.

- Yamamoto T et al. 2002. Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. *Theor. Appl. Genet.* 106: 9-18.
- Yang H et al. 2012. Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius* L. *BMC Genomics* 13: 318.
- Yang L et al. 2013. Marker-assisted selection for pyramiding the *waxy* and *opaque-16* genes in maize using cross and backcross schemes. *Mol. Breed.* 31: 767-775.
- Yang N et al. 2013. Mapping quantitative trait associated with resistance to bacterial spot (*Xanthomonas arboricola* pv *pruni*). *Tree Genet. Genomes* 9: 573-586.
- Yu H et al. 2014. A whole-genome SNP array (RICE6K) for genomic breeding in rice. *Plant Biotechnol. J.* 12: 28-37.
- Yu LX et al. 2011. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theor. Appl. Genet.* 123: 1257-1268.
- Yu LX et al. 2012. Identification of Ug99 stem rust resistance loci in winter wheat germplasm using genome-wide association analysis. *Theor. Appl. Genet.* 125: 749-758.
- Yun Z et al. 2012. Comparative transcriptomics and proteomics analysis of citrus fruit, to improve understanding of the effect of low temperature on maintaining fruit quality during lengthy post-harvest storage. *J. Exp. Bot.* 63: 2873-2893.
- Yuste-Lisbona et al. 2014. Genetic analysis of single-locus and epistatic QTLs for seed traits in an adapted × nuña RIL population of common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 127: 897-912.
- Zambrano JL et al. 2014. Genetic analysis of resistance to six virus diseases in a multiple virus-resistant maize inbred line. *Theor. Appl. Genet.* 127: 867-880.
- Zeinalabedini MK et al. 2008. Comparison of the use of morphological, protein and DNA markers in the genetic characterization of Iranian wild *Prunus* species. *Scientia Hort.* 116: 80-88.
- Zhang LY et al. 2009. Population structure and linkage disequilibrium in barley assessed by DArT markers. *Theor. Appl. Genet.* 119: 43-52.
- Zhao K et al. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Comm.* 2: 467.
- Zhebentayeva TN et al. 2014. Dissection of chilling requirement and bloom date QTLs in peach using a whole genome sequencing of sibling trees from an F₂ mapping population. *Tree Genet. Genomes* 10: 35-51.
- Zhou H et al. 2014. Association mapping of stem rust race TTKSK resistance in US barley breeding germplasm. *Theor. Appl. Genet.* 127: 1293-1304.
- Ziems LA et al. 2014. Association mapping of resistance to *Puccinia hordei* in Australian barley breeding germplasm. *Theor. Appl. Genet.* 127: 1199-1212.
- Ziliotto F et al. 2008. Transcriptome profiling of ripening nectarine (*Prunus persica* L. Batsch) fruit treated with 1-MCP. *J. Exp. Bot.* 59: 2781-2791.
- Zuriaga E et al. 2013. Genomic analysis reveals MATH gene(s) as candidate(s) for Plum pox virus (PPV) resistance in apricot (*Prunus armeniaca* L.). *Mol. Plant Pathol.* 14: 663-677.