Mechanisms of sexual polyploidization and inheritance in triploid citrus populations

PhD THESIS PRESENTED BY

JOSÉ CUENCA IBÁÑEZ

SUPERVISORS:

DR. LUIS NAVARRO LUCAS

DR. PATRICK OLLITRAULT

TUTOR:

DR. Mª JOSÉ DÍEZ

DECEMBER, 2013

Gracias a…

Durante estos años, muchas han sido las personas que han contribuido de una forma o de otra en el desarrollo de esta tesis, por lo que gran parte de ella les pertenece. A todos ellos, debo y quiero expresarles mi gratitud.

En primer lugar, quiero dar las gracias al IVIA y a su ex-director, Dr. Florentino Juste, por haberme dado la oportunidad de desarrollar mi trabajo en este centro. Quiero agradecer al Dr. Luis Navarro su confianza en mí para trabajar en su grupo (hace ahora 13 años) y realizar posteriormente esta tesis doctoral bajo su dirección. Durante todos estos años, su consejo y guía han sido imprescindibles para mi formación como profesional. Quiero agradecer de forma especial al Dr. Patrick Ollitrault, codirector de esta tesis, su apoyo y confianza en mi trabajo desde el principio. Su gran capacidad para orientar las investigaciones ha sido definitiva para el desarrollo de esta tesis, la cual no se habría podido concebir sin su inestimable participación. Quiero dar las gracias también a mi tutora, Dra. Mª José Díez, por estar siempre dispuesta a ayudar en cualquier tarea necesaria para el desarrollo de la tesis.

Quiero dar las gracias a mis compañeros del Laboratorio de Cultivo de Tejidos. A Pepe Juárez, su apoyo incondicional y confianza en mi; siempre he recibido de él buenos consejos y me ha enseñado todo lo que puedo saber sobre cultivo in vitro. A Pablo Aleza debo agradecerle, ante todo, su amistad y, por supuesto, su colaboración en todos los trabajos desarrollados en esta tesis. Gracias también a Juana Mª, Carmen, Toni, Vio, Cloti y Marga por ayudarme siempre en todo lo que he necesitado. Quiero agradecer a mis compañeros del laboratorio CIRAD-IVIA, su inestimable ayuda en este trabajo y su paciencia para aguantarme (Frederique, Juan, María, Marta, Ana, Houssem). De forma especial, quiero dar las gracias a Andrés (Dr. García-Lor *et al*) por su amistad, su gran predisposición a ayudarme y colaboración en esta tesis.

Agradecer también a todos aquéllos que en algún momento me han aconsejado y guiado en mis investigaciones (José Guerri, Leandro Peña, Gema Ancillo, José A. Pina, Alejandro Medina). Gracias a todos mis compañeros del centro, tanto a los que trabajan en los laboratorios (Ana, Elsa, Nuria, Berta, Montse, Jesús, Karelia, Vero, Toni, Daniel, Sandra, Pilar, Sawsen) como a los que trabajan en el campo o invernaderos (Vicente Lloris, José Manuel, Octavio (Paco), Fran, Aurelio, José Lloris, Diego, Raquel, Toni, Ramón, Ferran) y, en especial, a Rafa Montalt y Josep Peris (xufa) por ser más amigos que compañeros. Quiero agradecer a todos los que, de alguna forma, han convivido conmigo durante estos años (administración, conserjería, cafetería, limpieza, mantenimiento). Quiero recordar también a todos los que han pasado por el IVIA durante un tiempo más o menos largo (Magda, Mamen, Pedro, Álida, Giovanni, Jorge, Regina, Inma, Diana, Ezequiel, Lucía, Águeda, Marta, Esther) y especialmente a Mari Cruz, Frank, Rosa y José Manuel, con los que he convivido intensamente durante varios años.

Por supuesto, quiero dar las gracias a mi familia, mi padre, y mis hermanos y, en especial, a mi madre, que seguro que está orgullosa de mí.

Por último, quiero dedicarle esta tesis muy especialmente a mi mujer Mª Carmen, por entenderme y apoyarme en todo momento, ya que sin su ayuda y comprensión, esta tesis no se habría podido realizar. Y a mi chiquitín Andreu (y a mi futuro bebé), quiero agradecerle los momentos de alegría, su sonrisa y sus abrazos, que hacen que aprecie lo realmente importante de la vida.

CONTENTS

Abstract

Citrus is the main fruit crop in the world and Spain is the 6th producer and the major exporter for the fresh fruit market. Seedlessness is one of the most important fruit quality traits for this market since consumers do not accept seedy fruits. Recovery of triploid hybrids has become an important breeding strategy to develop new seedless citrus varieties and several of them have been already released from citrus breeding programs worldwide. Despite the undisputable importance of polyploidy in plant species, their genetics are much less well known than those of their diploid counterparts.

Citrus triploid hybrids can be routinely recovered from sexual polyploidization (2*x* × 2*x*) or interploid crosses (2*x* × 4*x* and 4*x* × 2*x*). In 2*x* × 2*x* sexual crosses, spontaneous triploid hybrids arise from the union of an unreduced (2*n*) megagametophyte with haploid pollen. In the case of interploid sexual crosses (2*x* × 4*x* and 4*x* × 2*x*), triploid hybrids result from the fecundation of a diploid gamete arising from the tetraploid parent and a haploid gamete arising from the diploid parent. The genetic and phenotypic structures of triploid populations greatly depend on the parental heterozygosity restitution (HR) in the diploid gamete at each locus, which is mainly affected by the triploid recovery strategy. In 2*x* × 2*x* crosses, HR depends on the underlying mechanism leading to the unreduced gamete formation, which are genetically equivalent to First Division Restitution (FDR) or Second Division Restitution (SDR) mechanisms. Moreover, under each restitution mechanism, HR also depends on the locus-centromere genetic distance. In the case of interploid crosses, parental heterozygosity restitution from tetraploid parents depends on the double reduction frequency. In citrus, the unreduced gamete formation mechanism is still controversial; FDR has been the mechanism proposed for sweet orange, whereas SDR has been proposed for clementine. On the other hand, inferring the allelic configurations of genetic markers is a main challenge in polyploidy crops to infer genotypic and gametic structures with the objective to analyze meiosis and inheritance mechanisms.

According to this scientific context, the objectives of the thesis where: (i) to develop a new approach for allele dosis assignation when using co-dominant markers, (ii) to implement and apply methods for the analysis of 2*n* gametes origin and locate centromeres, and (iii) to take advantage of this knowledge to locate a major gene of resistance to Alternaria Brown Spot (ABS) which is a major constraint for triploid mandarin breeding.

For microsatellite (SSR) markers, we have demonstrated that triploid progeny genotyping can be successfully performed using the microsatellite allele-counting peak ratio (MAC-PR) method. However, SSR analysis remains relatively costly and time consuming compared with actual SNP genotyping methods. Moreover, with the increasing availability of EST databases and whole genome sequences, SNPs have become the most abundant and powerful polymorphic markers that can be selected along the entire genome. In this thesis, a new method based on competitive allele-specific PCR has been developed to assign SNP allele dosage in an accurate, simple, and cost effective way. Combining the MAC-PR and the new developed SNP

genotyping methods offers the possibility to utilize a broad range of molecular markers in genotyping triploid genotypes. Both methods have been used in further works included in this thesis.

SDR has been demonstrated as the mechanism underlying unreduced gamete production in 'Fortune' mandarin by genotyping triploid progenies with SSR markers. In addition, a new method to locate the centromere, based on the best fit between observed heterozygosity restitution within a linkage group and theoretical functions under either partial or no chiasma interference hypotheses has been developed and successfully applied.

To expand the knowledge of the mechanism underlying unreduced gamete formation to other citrus genotypes besides clementines and 'Fortune' mandarin, a maximum likelihood method based on parental heterozygosity restitution of centromeric loci was developed and successfully applied in sixteen mandarin cultivars. The new method developed in the study allows inferring the restitution mechanism both at population level and even at individual level. Maternal origin of 2*n* gametes was confirmed for all triploid hybrids and SDR was proposed as the restitution mechanism for all analyzed progenies.

The information acquired from the mode of heterozygosity restitution in citrus was useful to determine the genetic and phenotypic structures of new triploid populations arising from different breeding strategies. We studied these structures for the resistance to Alternaria brown spot (ABS), a serious fungal disease producing necrotic lesions on fruits and young leaves in susceptible citrus genotypes. In the present work, different approaches were combined taking advantage of the particular genetic structures of 2*n* gametes resulting from SDR to map a genome region linked to ABS resistance in triploid citrus progeny. The monolocus dominant inheritance of the susceptibility, proposed on the basis of diploid population studies, was corroborated in triploid progeny. A 3.3 Mb genomic region linked to ABS resistance was located near the centromere on chromosome III, which includes clusters of resistance genes. SSR and SNP markers were developed for an efficient early selection of ABS resistant hybrids and they are currently used in our breeding program to perform marker assisted selection.

The knowledge obtained in this thesis on the mechanism of sexual polyploidization and inheritance of concrete traits in citrus will allow implementing much more efficient triploid breeding programs on the basis of current and future needs. Indeed, applied outcomes of this PhD are already routinely used in the IVIA triploid breeding program.

Resumen

Los cítricos son el principal cultivo frutal del mundo, siendo España el sexto productor mundial y el primer país exportador para el mercado en fresco. La ausencia de semillas constituye una de las características más importantes relacionadas con la calidad de los frutos, no siendo aceptados por los consumidores los frutos con semillas. La obtención de híbridos triploides se ha convertido en una estrategia importante para desarrollar variedades sin semillas, algunas de las cuales ya están disponibles en el mercado. A pesar de la importancia de las plantas poliploides, la genética de las mismas es mucho menos conocida que en plantas diploides.

Los híbridos triploides de cítricos se pueden obtener de forma rutinaria a partir de mecanismos de poliploidización sexual (2*x* × 2*x*) o mediante cruzamientos interploides (2*x* × 4*x* ó 4*x* × 2*x*). En los cruzamientos 2*x* × 2*x*, se producen híbridos triploides espontáneos que proceden de la unión de un óvulo no reducido (2*n*) con polen haploide. En el caso de cruzamientos interploides, los híbridos triploides resultan de la fecundación de un gameto diploide procedente del parental tetraploide con polen haploide procedente del parental diploide. Las estructuras genéticas y fenotípicas de las poblaciones triploides obtenidas dependen en gran medida de la restitución de heterocigosidad parental (HR) en el gameto diploide para cada locus, lo que está influido principalmente por la estrategia de obtención de estas poblaciones. En cruzamientos 2*x* × 2*x*, HR depende del mecanismo que da lugar a la formación del gameto no reducido; estos mecanismos son genéticamente equivalentes a una restitución en la primera división meiótica (FDR) o una restitución en la segunda división meiótica (SDR). Además, para cada mecanismo de restitución de heterocigosidad, HR depende a su vez de la distancia genética entre el locus considerado y el centrómero. En caso de cruzamientos interploides, la restitución de la heterocigosidad del parental tetraploide depende de la frecuencia de doble reducción. En los cítricos, existe cierta controversia sobre el mecanismo de formación de gametos no reducidos. En el caso de naranjo dulce, el mecanismo propuesto es FDR, mientras que para el clementino se ha propuesto SDR. Por otro lado, la asignación de las configuraciones alélicas de marcadores genéticos en especies poliploides es esencial para inferir las estructuras genotípicas y gaméticas con el objetivo de estudiar los mecanismos de meiosis y herencia. En este contexto, los objetivos científicos de la tesis han sido: (i) desarrollar un nuevo método para asignar dosis alélicas utilizando marcadores codominantes, (ii) implementar y aplicar nuevas estrategias para analizar el origen de los gametos no reducidos y localizar centrómeros y (iii) utilizar el conocimiento generado para localizar un gen mayor de resistencia a la mancha marrón producida por el hongo *Alternaria*, que supone actualmente un inconveniente en los programas de mejora de mandarino.

En el caso de marcadores microsatelites (SSRs), se ha demostrado que el genotipado de poblaciones triploides puede llevarse a cabo mediante el método *microsatellite allele-counting peak ratio* (MAC-PR). Sin embargo, el empleo de marcadores SSR es relativamente costoso y lento, en comparación con los métodos actuales de genotipado de SNP. Además, con la creciente disponibilidad de bases de datos procedentes de ESTs y secuenciación completa de

genomas, los SNP se han convertido en los marcadores polimórficos más abundantes y efectivos que se pueden utilizar. En la presente tesis, se ha desarrollado un método basado en PCR competitiva específica de alelos para asignar las dosis alélicas de SNPs de una forma precisa, simple y efectiva. La combinación de este nuevo método con el MAC-PR ofrece la posibilidad de emplear un gran abanico de marcadores moleculares para el genotipado de poblaciones triploides. Ambos métodos se han empleado en los trabajos descritos a lo largo de esta tesis.

Mediante la utilización de marcadores SSR, se ha demostrado que el mecanismo que da lugar a la formación de gametos no reducidos en el mandarino 'Fortune' es SDR. Además, se ha desarrollado y aplicado un nuevo método para localizar centrómeros comparando los datos de HR observados dentro de un grupo de ligamiento y funciones teóricas bajo modelos de interferencia parcial y no interferencia.

Para ampliar el conocimiento sobre los mecanismos de origen de los gametos no reducidos a otros genotipos, se ha desarrollado y aplicado un método de máxima similitud basado en la HR de marcadores centroméricos. Este nuevo método permite identificar el mecanismo de origen tanto a nivel poblacional como a nivel individual. Como resultado, se ha confirmado que SDR es el mecanismo de origen de los óvulos diploides en todas las poblaciones de híbridos triploides analizadas.

La información obtenida anteriormente sobre el mecanismo de origen de los gametos no reducidos ha sido utilizada para determinar las estructuras genéticas y fenotípicas de nuevas poblaciones triploides obtenidas mediante diferentes estrategias. En este sentido, se han analizado estas estructuras para la resistencia a la mancha marrón (ABS), una enfermedad fúngica importante en los cítricos que produce lesiones necróticas sobre los frutos y hojas jóvenes de los genotipos susceptibles. En la presente tesis, a partir del conocimiento de la estructura genética particular proveniente de SDR como mecanismo de origen de los gametos no reducidos, se han combinado diferentes estrategias para analizar la resistencia a ABS en poblaciones triploides de cítricos. Se ha corroborado en estas poblaciones triploides la herencia monolocus dominante propuesta a partir de estudios en poblaciones diploides. Además, se ha localizado un región genómica de 3.3 Mb ligada a la resistencia a ABS cerca del centrómero en el cromosoma III, en la que se incluyen varios grupos de genes de resistencia. También se han desarrollado marcadores SSR y SNP para realizar una selección temprana de híbridos resistentes, los cuales están siendo utilizados actualmente en el programa de mejora del IVIA para selección asistida por marcadores.

El conocimiento generado en esta tesis sobre los mecanismos de poliploidización sexual y la herencia de caracteres concretos en cítricos permitirá un desarrollo más eficiente de los programas de mejora en base a necesidades actuales y futuras. De hecho, los resultados de aplicación práctica obtenidos en esta tesis están siendo utilizados actualmente.

Resum

Els cítrics són el principal conreu fruiter del món, sent Espanya el sisé productor mundial i el primer exportador de fruita fresca. L'absència de llavors constitueix actualment una de les característiques més importants relacionades amb la qualitat des fruits, no sent acceptables per part dels consumidors els fruits amb llavors. L'obtenció de híbrids triploides s'ha convertit en una estratègia important per al desenvolupament de varietats sense llavors, algunes de les quals estan disponibles al mercat. Però, a pesar de la importància de les espècies poliploides, la seua genètica és menys coneguda que en espècies diploides.

Els híbrids triploides de cítrics es poden obtindre de forma rutinària mitjançant mecanismes de poliploidització sexual (2*x* × 2*x*) o mitjançant creuaments interploides (2*x* × 4*x* ó 4*x* × 2*x*). Als creuaments 2*x* × 2*x* es produeixen híbrids triploides espontanis que procedeixen de la unió d'un òvul no reduït (2*n*) amb pol·len haploide. En el cas de creuaments interploides, els híbrids triploides resulten de la fecundació d'una gàmeta diploide procedent del parental tetraploide amb pol·len haploide procedent del parental diploide. Les estructures genètiques i fenotípiques de les poblacions triploides obtingudes depenen en gran mesura de la restitució d'heterozigosis parental (HR) en la gàmeta diploide per a cada locus, el que està influït principalment per l'estratègia d'obtenció de les poblacions triploides. Als creuaments 2*x* × 2*x*, HR depèn del mecanisme que origina la formació de la gàmeta no reduïda; aquestos mecanismes són genèticament equivalents a una restitució a la primera divisió meiòtica (FDR) o una restitució a la segona divisió meiòtica (SDR). A més, per a cada mecanisme de restitució d'heterozigosis, la HR per un locus depèn de la distància genètica entre el locus considerat i el centròmer. En cas de creuaments interploides, la restitució d'heterozigosis del parental tetraploide depèn de la freqüència de doble reducció. En el cas del cítrics, existeix certa controvèrsia sobre el mecanisme d'origen de les gàmetes no reduïdes. En cas de taronger dolç, el mecanisme que s'ha proposat es FDR, mentre que per al clementí s'ha proposat SDR. D'un altra banda, l'assignació de les configuracions al·lèliques de marcadors genètics en espècies poliploides és essencial per a inferir les estructures genotípiques i gamètiques amb l'objecte d'estudiar els mecanismes de meiosi i herència. En aquest context, els objectius científics de la present tesi doctoral són: (i) desenvolupar un nou mètode per a assignar dosis al·lèliques utilitzant marcadors moleculars codominants, (ii) implementar i aplicar noves estratègies per a analitzar l'origen de les gàmetes no reduïdes i localitzar centròmers i (iii) utilitzar el coneixement generat per a localitzar un gen major de resistència a la taca marró produïda pel fong *Alternaria*, que suposa actualment un inconvenient per als programes de millora de mandarí.

En el cas de marcadors microsatèl·lits (SSR), s'ha demostrat que el genotipatge de poblacions triploides pot realitzar-se mitjançant el mètode *microsatellite allele-counting peak ratio* (MAC-PR). No obstant això, l'ús de marcadors SSR és relativament costós i lent en comparació als actuals mètodes de genotipatge de SNPs. A més, amb la creixent disponibilitat de bases de dades procedents d'ESTs i seqüenciació completa de genomes, els SNP s'han convertit en els marcadors polimòrfics més abundants i efectius que es poden utilitzar. En la present tesi, s'ha

desenvolupat un mètode basat en PCR competitiva específica d'al·lels per a assignar les dosis al·lèliques de SNP d'un forma senzilla, precisa i efectiva. La combinació d'aquest nou mètode amb el MAC-PR ofereix la possibilitat d'utilitzar un gran nombre de marcadors moleculars per al genotipatge de poblacions triploides. Ambdós mètodes s'han emprat en els treballs descrits al llarg de la present tesi.

Mitjançant la utilització de marcadors SSR, s'ha demostrat que el mecanisme d'origen de les gàmetes no reduïdes en el mandarí 'Fortune' és SDR. A més, s'ha desenvolupat i aplicat un nou mètode per a localitzar centròmers comparant les dades de HR observades dins d'un mateix grup de lligament amb funcions teòriques baix models d'interferència parcial i no interferència.

Per a ampliar el coneixement sobre els mecanismes d'origen de les gàmetes no reduïdes a altres genotips, s'ha desenvolupat i aplicar un mètode de màxima similitud basat en la HR de marcadors centromèrics. Aquest nou mètode permet la identificació del mecanisme d'origen tant a nivell poblacional com a nivell individual. Com a resultat, s'ha confirmat que SDR és el mecanisme d'origen del òvuls diploides en totes les poblacions d'híbrids triploides analitzats.

La informació obtinguda sobre aquest mecanisme d'origen ha sigut posteriorment utilitzada per a determinar les estructures genètiques i fenotípiques de noves poblacions triploides obtingudes mitjançant diferents estratègies. En aquest sentit, s'han analitzat aquestes estructures per a la resistència a la taca marró (ABS), una infermetat fúngica importat en els cítrics que produeix lesions necròtiques als fruits i fulles joves dels genotips susceptibles. En la present tesi, a partir del coneixement de l'estructura particular provinent de SDR com a mecanisme d'origen de les gàmetes no reduïdes, s'han combinat diferents estratègies per a analitzar la resistència a ABS en poblacions triploides de cítrics. S'ha confirmat en aquestes poblacions l'herència monolocus dominant proposada a partir d'estudis sobre poblacions diploides. A més, s'ha localitzat una regió genòmica lligada a la resistència a ABS prop del centròmer del cromosoma III, on s'han trobat diversos grups de gens de resistència. També s'han desenvolupat marcadors SSR i SNPs per a realitzar una selecció primerenca dels híbrids resistents, els quals estan sent utilitzats actualment al programa de millora del IVIA per a selecció assistida per marcadors.

El coneixement generat a aquesta tesi sobre el mecanismes de poliploidització sexual i l'herència de caràcters concrets en cítrics permetrà un dessenvolupament més eficient dels programes de millora en base a les necessitats actuals i futures. De fet, els resultats amb aplicació més pràctica obtinguts en aquesta tesi doctoral estan sent utilitzats actualment.

Introduction

Contents

1. ECONOMIC IMPORTANCE OF CITRUS

Citrus is the main fruit crop in Spain and worldwide. Total citrus production was over 128 million tons with 9.2 million hectares of cultivated surface in 2011. Main producer countries are Brazil, China, USA, India, Mexico and Spain (FAO, 2013).

From the agronomic point of view, four varietal groups can be differentiated: oranges, mandarins, lemons/limes and grapefruits. World production is led by oranges, with 70 million tons (60%) followed by mandarins, with 24.5 million tons (21%), lemons/limes, with 14 million tons (12%) and grapefruits, with 8 million tons (7%) (FAO, 2013, Figure 1).

About one third of citrus fruit production is processed, mainly for orange juice production. Main countries in juice production are United States and Brazil, which produce more than 85% of the world market. Relating to mandarins, more than 90% of world production is destined to fresh consumption (USDA, 2013). Mediterranean countries mainly destine their citrus production for the fresh fruit market, exporting about 3.3 million tons of mandarins in 2011 (FAO, 2013). Clementine mandarins and related fruit from Spain and Morocco dominate the easy-peelers category (Ladaniya, 2008).

Spain produced 6.3 million tons in 2011, with a cultivated surface of over 330.000 ha, mainly producing oranges (51%) and mandarins (36%) (Figure 1). Spain exports 57% of citrus fruit production. The Valencia region is the first Spanish citrus producer with 3.5 million tons and a cultivated surface of over 170.000 ha, which represents 56% of total Spanish citrus production (MAGRAMA, 2013). Valencia region's production (Figure 1) is led by mandarins (48%) and oranges (46%). Within the mandarin group, 70% of production is based on clementines and within the orange group, 81% of production is based on navel oranges (GVA, 2013).

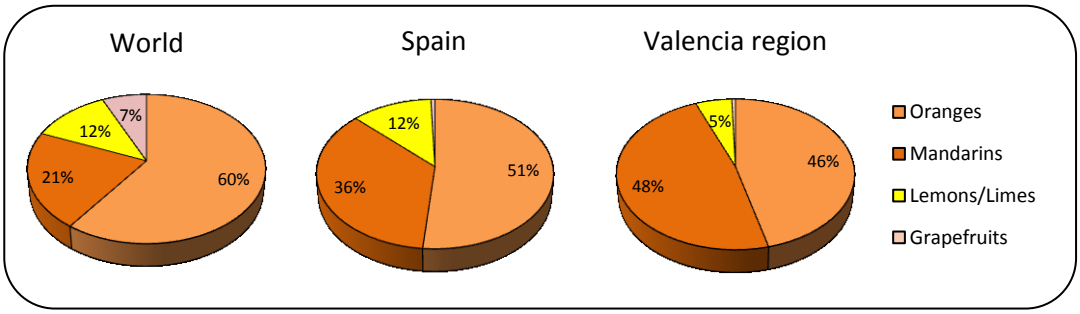

Figure 1. Proportion of the citrus production by group, worldwide, in Spain and in the Valencia region (FAO, 2013; GVA, 2013; MAGRAMA, 2013).

2. TAXONOMY, ORIGIN AND DIVERSITY OF CITRUS AND RELATED GENERA

2.1. Taxonomy

Citrus and related genera are classified within the order *Geraniales*, suborder *Geraniineae* and family *Rutaceae*. This family comprises six subfamilies, with the *Aurantioideae* one containing *Citrus* and related genera. The subfamily *Aurantioideae* is divided into two tribes, *Clauseneae* y *Citreae*. Tribe *Citreae* is also subdivided into three sub tribes, with the *Citrinae* one comprising the true citrus sub tribe group, which includes *Eremocitrus*, *Microcitrus*, *Clymenia*, *Fortunella*, *Poncirus* and *Citrus* genera. The commonly grown citrus cultivars and rootstocks belong to these last three genera.

Fortunella is a genus with several species known as kumquats, all being small trees with a later flowering time than *Citrus* species, relatively cold tolerant and resistant to *citrus canker* and *Phytophthora spp.* They bear small fruits with sweet tasting rind. This genus includes four species: *Fortunella margarita* (Lour.) Swing., *F. japonica* (Thunb.) Swing., *F. polyandra* (Ridl.) Tan. and *F. hindsii* (Champ.) Swing. Kumquats have been cultivated in China for long time and are recently being used as parents in citrus breeding programs.

Poncirus includes only the *Poncirus trifoliata* (L.) Raf species. Its use is exclusively as a rootstock in some producing areas and as parent in rootstock breeding programs, due to its resistance to *Citrus Tristeza Virus*, citrus nematode (*Tylenchulus semipenetrans*), *Phytophthora parasitica* and *P. citrophthora* and its cold tolerance. Its derived hybrids with sweet oranges, [`Carrizo´ and `Troyer´ citranges (*C. sinensis* × *P. trifoliata*)] are the main rootstocks used in Spain; a cross derived from grapefruit and *Poncirus* (*C. paradisi* × *P. trifoliata)* known as Citrumelo ‘CPB-4475’ is also used as a rootstock in some citrus producing Spanish areas.

Botanical classification within the *Citrus* genus has been a challenge for a long time. The most widely accepted taxonomic systems today are those of Swingle (Swingle, 1943; Swingle and Reece, 1967) and Tanaka (1977), who recognized 16 and 162 species, respectively.

Swingle divided the *Citrus* genus into two subgenus, *Papeda* and *Eucitrus*, mainly distinguishable by the fruit edibility. *Papeda* fruits contain essential oil agregates which causes sour and unpleasant taste into their juice, whereas *Eucitrus* fruits contain juice free of essential oils with sweet or acid taste. *Papeda* subgenus includes six species: *Citrus ichangensis* Swing., *C. latipes* (Swing.) Tan., *C. hystrix* D.C., *C. micrantha* Wester, *C. celebica* Koord. and *C. macroptera* Montr. *Eucitrus* genus includes ten species: *C. medica* L., (citron), *C. aurantium* L. (sour orange), *C. limon* (L.) Burn. f. (lemon), *C. aurantifolia* (Christm.) Swing. (lime), *C. grandis* (L.) Osb. (pummelo or shadock), *C. sinensis* (L.) Osb. (sweet orange), *C. reticulata* Blanco (mandarin), *C. paradisi* Macf. (grapefruit), *C. indica* Tan. and *C. tachibana* (Mak.) Tan., with the last two species being of less commercial importance.

Tanaka divided the *Citrus* genus into two subgenera, *Archicitrus* and *Metacitrus*, including 162 species. However, in relation to the most commercially important species, differences with the Swingle classification mostly affect lemons, limes and mandarins. Tanaka divides limes into three species, *C. aurantifolia* (Christm.) Swing. (Mexican, Galego, Key or West Indian limes), *C. latifolia* Tan. (Bearss, Tahití or Persian limes) and *C. limettioides* Tan. (Palestine sweet lime).

The mandarin group is the main germplasm for citrus breeding. Its classification is particularly controversial. Swingle defines only one species (*C. reticulata* Blanco), (Swingle and Reece, 1967), whereas Tanaka includes also *C. deliciosa* Ten. (Mediterranean or Willowleaf mandarin), *C. unshiu* Marc. (satsumas), *C. clementina* Hort. ex Tan. (clementines), *C. tangerina* Hort. ex Tan. (Dancy mandarin), *C. nobilis* Lour. (King mandarin) and other less important species (Tanaka, 1977). From an agronomic point of view, Tanaka's classification is better adapted to the characteristics of the different agronomic groups, and it is widely used to manage germplasm collections (Krueger and Navarro, 2007). Recent molecular studies (García-Lor *et al.*, 2013b) on mandarins reveal that the genomes of most 'mandarin-like' genotypes are complex admixtures of five parental mandarin groups and even include contributions from the other ancestral populations. In the same study, a mitochondrial analysis with indel markers revealed four mitotypes in which mandarin and 'mandarin-like' genotypes were represented.

2.2. Origin and distribution of cultivated citrus

Scora (1975) and Barrett and Rhodes (1976) suggested that there are only three basic true species within the subgenus Citrus: C. medica L. (citron), C. reticulata Blanco (mandarin) and C. maxima (L.) Osb. (pummelo or shaddock). Recently, this thesis has gained support from various biochemical and molecular studies (Herrero *et al.*, 1996; Federici *et al.*, 1998; Nicolosi *et al.*, 2000; Gulsen and Roose, 2001; Barkley *et al.*, 2006; Li *et al.*, 2010; Penjor *et al.*, 2013), suggesting that current citrus groups arise from only four ancestors: citron, mandarin, pummelo and *C. micrantha* Wester (a 'Papeda' wild citrus). Other cultivated species within *Citrus* derive from hybridizations between these true species or closely related genera followed, mainly, by natural mutations (Gmitter *et al.*, 1992; Davies and Albrigo, 1994; Nicolosi *et al.*, 2000; García-Lor *et al.*, 2012; Ollitrault *et al.*, 2012b; Uzun and Yesiloglu, 2012; García-Lor *et al.*, 2013b; Figure 2). In addition, other groups of cultivars originated from hybridizations between ancestral or secondary species are of commercial importance, as tangors (sweet orange × mandarin hybrids) and tangelos (mandarin × grapefruit hybrids).

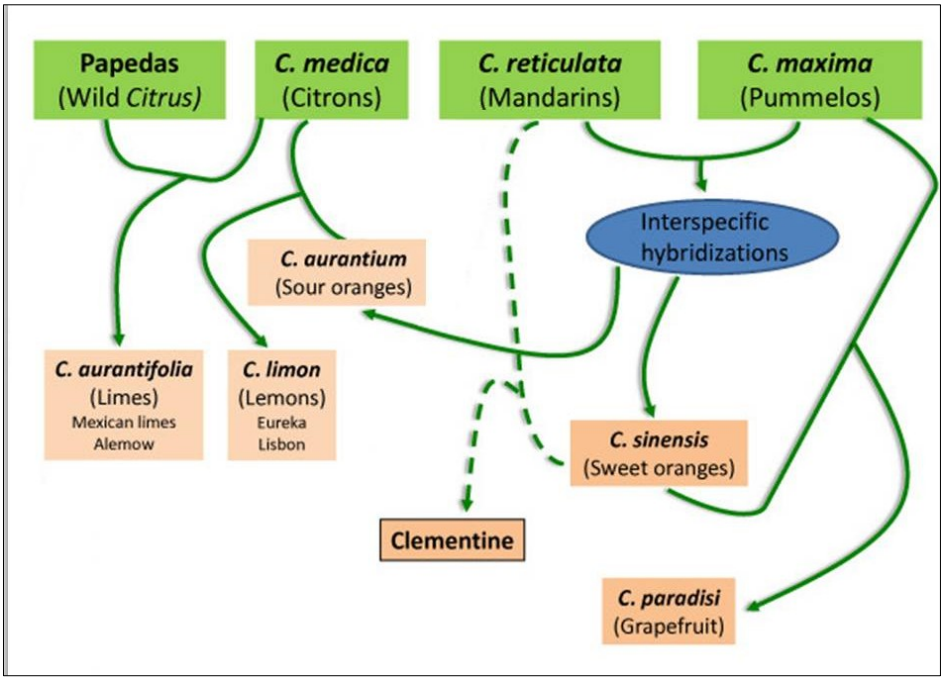

Figure 2. Schematic representation of the origin of the main citrus groups. Modified from Ollitrault *et al.*, (2012b).

All citrus species and related genera, excluding grapefruit, are thought to have been originated in tropical and subtropical regions in Asia and Malaysian archipelago, from where they were extended to the other continents (Webber, 1967; Chapot, 1975; Ollitrault and Navarro, 2012; Figure 3).

Citron was the first citrus genotype reported in Spain in the 7th century. It was cultivated in the Andalousia region, probably introduced from Italy by the Romans (Zaragoza, 2007).

Sour orange is a natural hybrid of a mandarin and a pummelo (Li *et al.*, 2010). Sour orange was expanded in the 10th century by Arab marketers from India to Iraq, Syria, Palestine, Egypt and North Africa, and from there introduced in Spain, Sicilia and Sardinia. Many studies have suggested that lemon is likely to be a natural hybrid of a citron and a lime (Scora, 1975; Barrett and Rhodes, 1976) or a hybrid of citron and sour orange (Nicolosi *et al.*, 2000; Gulsen and Roose, 2001). Lemon is thought to have been introduced in Spain at the same time or little later than sour orange. At that time, all these species are reported to be cultivated in south regions in Spain (Zaragoza, 2007).

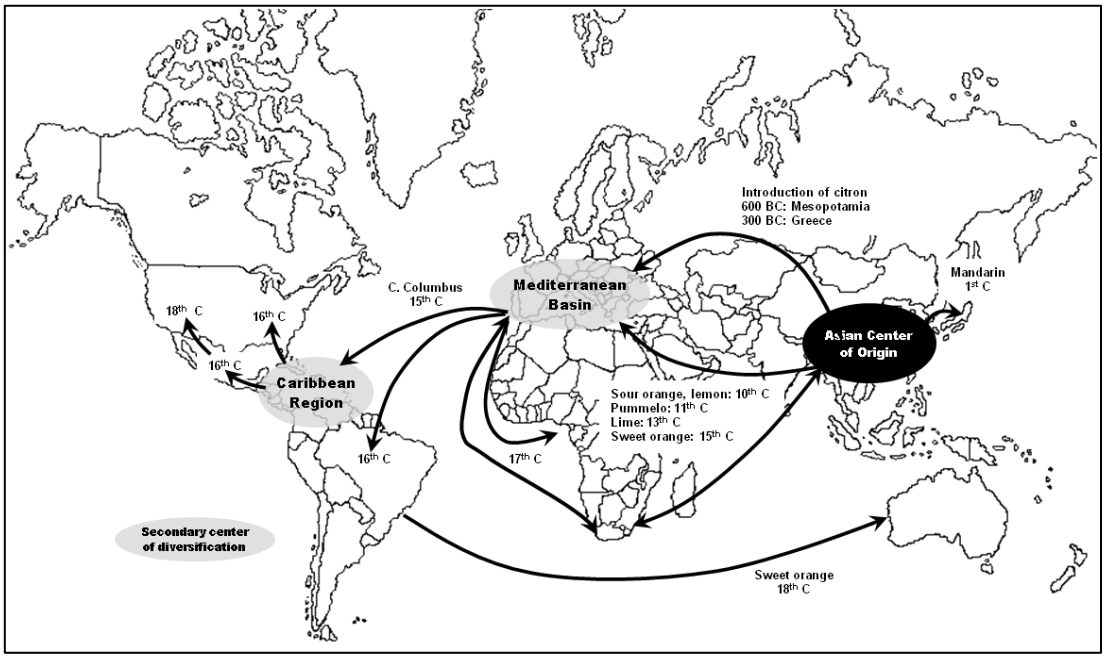

Figure 3. Origin and worldwide distribution of citrus. From Ollitrault and Navarro (2012).

Nicolosi *et al.* (2000) proposed that sweet orange was derived from a single cross based on equal proportions of alleles from mandarin and pummelo. However, Barkley *et al*. (2006) and Luro *et al.* (2008) suggested one or more backcrosses with mandarin and, therefore, its genetic background is mainly derived from mandarin and a small proportion from pummelo. Sweet orange was introduced in Spain four or five centuries later than sour orange, about half 15th century, by Genovese through their commercial relations with Orient. However, it is known that Portuguese greatly contributed to their diffusion around the Iberian Peninsula and Italy, by importing seeds of better and sweeter varieties (Zaragoza, 2007).

Citrus were distributed from the Mediterranean area by three routes: by the Arabs to Africa between eleventh and thirteenth centuries; Christopher Columbus introduced them to Haiti in 1493 and the British and the Dutch introduced them in Capetown in 1654. Due to discovery of America and its gradual conquest, occurred the establishment of citrus in Mexico (1518), Brazil (1540), Florida (1565), Peru (1609) and Texas (1890). The Portuguese introduced citrus in South Africa in 1654. The first settlers brought oranges, lemons and limes from Brazil to Australia in 1769.

Mandarin was the latest citrus species introduced in Spain, by the 19th century. It was extended around north areas in the Valencia region since 1845 (Zaragoza, 2007). The highly diverse group of mandarin includes numerous species, most of which derive from natural intergeneric and intespecific crosses, while others, which are commercially important, derive from man-made crosses (Nicolosi, 2007).

Grapefruit was probably originated from a pummelo × sweet orange spontaneous hybridization in the Barbados Island. It was imported to Spain from California and Florida and first reported to be cultivated in Valencia in 1929 (Herrero, 1929).

2.3. The origin of intraspecific diversity

Most of the cultivars of orange, grapefruit, and lemon are believed to have originated from nucellar seedlings or budsports. Barrett and Rhodes (1976) reported variations in orange, lemon, grapefruit and lime based on mutations occurred on one ancestor tree. Consequently, the amount of genetic diversity within these groups is relatively low, in spite of phenotypic differences among varieties. Conversely, mandarins, pummelos, and citrons have higher levels of genetic diversity since many of the cultivars have arisen through sexual hybridization.

Mandarin germplasm was classified as *C. reticulata* Blanco by Swingle and Reece (1967) and Mabberley (1997). On the contrary, Webber (1967) classified mandarin genotypes into four different groups: king, satsuma, mandarin, and tangerine. Tanaka (1977) divided mandarins into five groups that included 36 species, based on morphological differences in the tree, leaves, flowers, and fruits. The mandarin group has high amount of cultivars, some of them originated from hybridization and the others derived from mutation. So, in the mandarins from hybrid origin there is a clear genetic variation. On the other hand, low level of diversity observed in the cultivars occurred by mutation, such as satsuma and clementine groups (Bretó *et al.*, 2001; Barkley *et al.*, 2006; Uzun and Yesiloglu, 2012). In the mandarin group germplasm, there are many hybrid accessions derived from mandarin × mandarin, mandarin × grapefruit (as tangelo), mandarin × orange (as tangor) or mandarin × tangelo. Clementine was classified by Tanaka as a *Citrus* species (*C. clementina* Hort. ex Tan.); currently, this species is one of the most important cultivated mandarins, especially in the Mediterranean countries.

Sweet orange originated as a natural hybrid between mandarin and pummelo (Barrett and Rhodes, 1976; Luro *et al.*, 1994; Fang *et al.*, 1997; Uzun and Yesiloglu, 2012). Recently, García-Lor *et al.* (2012) proposed that both parents of sweet orange have an interspecific origin. Orange cultivars are classified into four groups: common, low acidity, pigmented and navel oranges. It is considered that sweet oranges varieties originated by mutation from one ancestor tree. So, despite of differences in morphological characters, low level of genetic diversity is observed. Nevertheless, pigmented varieties, which accumulate anthocyanin, have a gene encoding a transcription factor (*Ruby* gene) and can be differentiated from other orange groups (Butelli *et al.*, 2012).

Most lemons have highly similar morphological and biochemical characters, and they are reported to have originated by mutation from a single parental lemon tree (Gulsen and Roose, 2001; Luro *et al.*, 2008; Uzun and Yesiloglu, 2012).

Grapefruits have a nucellar and mutation origin and genetic variation among common them is very low. High level of similarity was found in grapefruit cultivars in various studies (Scora, 1988; Gmitter *et al.*, 1992; Fang *et al.*, 1997; Moore *et al.*, 2000; Nicolosi *et al.*, 2000).

3. NEEDS IN CITRUS SCION BREEDING

The modern citrus industry is based on grafted plant, with the scion cultivar budded on a rootstock (Khan and Kender, 2007). It allows producing disease free planting material without the juvenile phase when coupled with efficient certification schemes (Navarro *et al.*, 1975). Moreover, the rootstock is a major component for (i) resistance to soil pathogens such as *Phytophthora* spp and nematodes, (ii) resistance to the *Citrus tristeza virus* and (iii) for adaptation to adverse environmental conditions (salinity, water deficit, calcareous soils, cold…). The rootstocks also modulate some characteristics of the scion such as fruit quality, productivity or maturing time (Khan and Kender, 2007). However, these last characteristics are mainly determined by the scion genotype, as well as tolerance to diseases affecting the crown of the tree.

3.1. General objectives of scion breeding

In relation with specific market demands and environmental conditions (biotic and abiotic constraints), the main goal of breeding may vary between the production areas. However, some general trends can be outlined. For juice processing, cultivars with a high percentage of pigmented juice and sugar content in their fruits are desirable (Ollitrault *et al.*, 2008). For the fresh fruit market, expanding the harvesting period with high quality seedless fruits is currently the main objective of scion breeding (Navarro *et al.*, 2005; Recupero *et al.*, 2005; Roose and Williams, 2007; Aleza *et al.*, 2010a; Cuenca *et al.*, 2010). Biological mechanisms leading to seedlessness and breeding strategies targeting this trait are detailed in paragraph 4.

The fresh fruit market requests pomological qualities (easy peeling, seedlessness, external appearance) and organoleptic qualities (aroma, taste, acidity, sugar). Since the definition of organoleptic quality varies with the consumer, citrus breeders must therefore endeavor to develop a wide range of varieties likely to meet these diverse needs. Nutritional quality based on vitamin C, carotenoid and polyphenol contents are now considered as breeding criteria in some projects (Alquézar *et al.*, 2009; Sdiri *et al.*, 2012).

Resistance to some diseases that cause considerable damage in orchards is also an important objective in breeding programs. These include Huanglongbing (*ex* citrus greening) caused by the bacterium *Candidatus* Liberibacter asiaticus in Asia, South Africa and recently in Brazil and Florida; Citrus canker, caused by the bacterium *Xanthomonas citri* pv. *citri*, in most tropical and subtropical areas; cercosporiosis, caused by the fungus *Cercospora aurantia,* in

Africa; Mal secco, caused by the fungus *Phoma tracheiphila*, in the Mediterranean Basin for lemon, citrus variegated chlorosis, caused by the bacterium *Xylella fastidiosa* and Sudden Death in Brazil (Ollitrault and Navarro, 2012). Alternaria Brown Spot, caused by the fungus *Alternaria alternata* pv. *citri*, is a major problem for some susceptible mandarin cultivars such as 'Fortune'. In Spain, this cultivar has been widely cultivated, but now replaced due to its high disease susceptibility (Vicent *et al.*, 2004; Navarro *et al.*, 2005). More details on this disease and the potential for resistant cultivar breeding are given in next paragraph 3.2. Ranges of varietal susceptibility have been established for several important diseases and tolerant or resistant parents are selected in some breeding projects. However, in the case of Huanglongbing, no exploitable resistance source has been identified. In this sense, approaches involving genetically-modified organisms are being currently exploited (Alquézar *et al*., 2012).

3.2. Special focus on breeding for Alternaria Brown Spot resistance

Alternaria Brown Spot (ABS) is a serious fungal disease, which induces necrotic lesions on fruits and young leaves, defoliation and fruit drop in susceptible citrus genotypes. ABS is caused by the tangerine pathotype of *Alternaria alternata* (Fr.) Keissl. (Akimitsu *et al.*, 2003).

The disease was first observed in Australia in 1903 on 'Emperor' mandarin (Pegg, 1966), and was further detected in citrus-growing regions in America, the Mediterranean Basin, South Africa, Iran and China (Goes *et al.*, 2001; Timmer *et al.*, 2003; Golmohammadi *et al.*, 2006; Wang *et al.*, 2010). In Spain, the disease was first detected in 1998 (Vicent *et al.*, 2000) and it is currently widespread in most citrus-growing areas.

Due to the environmental flexibility of the fungus, the disease causes severe epidemics in humid areas, as well as in semi-arid regions (Timmer *et al.*, 2003). The pathogen sporulates on affected tissues and conidia are disseminated by air currents and rain splash. Warm temperatures and prolonged wetness on the tree are required for infection.

Currently, ABS control is strongly based on the application of fungicides. Sprays must be scheduled to protect susceptible organs during the critical periods for infection. Depending on the climate of the region and the susceptibility of the cultivar, between 4 and 10 fungicide sprays per year are needed to produce symptomless quality fruit for the fresh market (Swart *et al.*, 1998; Bhatia *et al.*, 2003; Peres and Timmer, 2006; Vicent *et al.*, 2007). Systematic application of fungicides for ABS control over years may create environmental problems and public health concerns (Vicent *et al.*, 2009). Moreover, and despite of this high number of sprays, disease control is not always satisfactory and the production of susceptible cultivars such as 'Fortune' or 'Nova', among others, has declined significantly during recent years in Spain, which are being removed or substituted by resistant cultivars (Navarro *et al.*, 2005).

Due to constraints of the citrus reproductive system (particularly apomixis of many genotypes, specific qualities or capacity to produce 2*n* gametes (see paragraph 5) several ABS-

susceptible monoembryonic cultivars are being used as parents in many mandarin breeding programs, both at diploid and triploid level (Mourao Fo *et al.*, 1996; McCollum, 2007; Aleza *et al.*, 2010a, 2010b; Cuenca *et al.*, 2010; Grosser *et al.*, 2010; Aleza *et al.*, 2012c, 2012d; Froelicher *et al.*, 2012; Navarro *et al.*, 2012). For example, 'Fortune' mandarin has been the most used female parent in 2*x* × 2*x* triploid breeding programs (Recupero *et al.*, 2005; Navarro *et al.*, 2012). It can produce several fold spontaneous triploid genotypes than other cultivars used as female parents, such as 'Fina' or 'Moncada' (Aleza *et al.*, 2010b) and quality transmitted to its progenies have been demonstrated by the newly released varieties (Recupero *et al.*, 2005; Aleza *et al.*, 2010a; Cuenca *et al.*, 2010). These programs have also shown that resistant triploid hybrids can be obtained when using 'Fortune' as female parents. 'Murcott' and 'Ponkan' are widely used parents in many diploid breeding programs carried out in Japan or Brazil (JinPing *et al.*, 2009; Schinor *et al.*, 2012). Their fruits present good quality and late ripening and its segregations are expected to be of major interest in those regions. However, they are also ABS susceptible, resulting in the segregation between resistant and susceptible hybrids in their progeny.

Other ABS-susceptible cultivars, such as 'Minneola', 'Nova', 'Fairchild', 'Fremont', 'Page', 'Orlando', 'Pixie', or 'Daisy' are also used as parents in diploid and triploid breeding programs (Aleza *et al.*, 2012c; Williams, 2012).

Knowledge about inheritance and segregations of ABS-resistance for different hybridizations at diploid and triploid level, and for different strategies to produce triploid hybrids would greatly improve the choice of breeding parents and the managing of future hybridizations.

The tangerine pathotype of *A. alternata*, which causes ABS in citrus, has a gene cluster (*ACTT*) located in a small (<2.0 Mb) conditionally dispensable chromosome, responsible of ACT-toxin I biosynthesis (Ajiro *et al.*, 2010). Another selective toxin, named ACT-toxin II (or ACTG-toxin), was also identified by Kono *et al.* (1986); however, ACT-toxin I (ACT-toxin, hereafter) is at least tenfold more toxic to citrus. This host-specific toxin (HST) is released during the germination of conidia, rapidly affecting the integrity of plasma membrane of susceptible host cells (Kohmoto *et al.*, 1993).

The mode of action of ACT-toxin is still ambiguous; however, a rapid loss of electrolytes from leaf tissues and invaginations in the plasma membranes of cells affected by the toxin indicates that its primary action site is likely to be the plasma membrane (Akimitsu *et al.*, 2003). Moreover, there are indirect evidences suggesting the presence of toxin receptors in the plasma membrane of susceptible citrus genotypes (Tsuge *et al.*, 2012). The damage to host cells caused by these toxins is extremely rapid, and the time required for cell death is extremely short (Maekawa *et al.*, 1984; Kohmoto *et al.*, 1993; Otani *et al.*, 1995). Due to the rapid effect of the ACT-toxin, the incubation period is very short and lesions are visible just one or two days after infection was initiated (Canihos *et al.*, 1999). However, although HSTs are highly toxic to host plant cells and cause cell death, the establishment of infection by the pathogen is probably

caused by a genotype-specific suppression of defence responses rather than the induction of cell death (Tsuge *et al.*, 2012). In addition, recent studies indicated that the mitigation of reactive oxygen species (ROS) produced by the host plants is essential for pathogenicity (Yang and Chung, 2012).

Several studies have been carried out to determine the resistance or susceptibility to ABS in citrus genotypes (Hutton and Mayers, 1988; Goes *et al*., 2001; Vicent *et al*., 2004; Dalkilic *et al*., 2005; Reis *et al*., 2007; de Souza *et al*., 2009). Although there are some discrepancies among them, resistance is clear for clementine (*C. clementina*), 'Willowleaf' (*C. deliciosa*) and satsuma (*C. unshiu*) mandarins. Susceptibility has also been well established for 'Dancy' and 'Fortune' mandarins, 'Orlando', 'Minneola' and 'Nova' tangelos and 'Murcott' tangor. Other cultivars such as 'Ellendale' tangor, and some sweet oranges and grapefruits have been characterized as sensitive or resistant depending on the authors.

ACT-toxins were found to be structurally analogous metabolites to those produced by Japanese pear (AK-toxins) and strawberry pathotypes (AF-toxins), that are esters of 9,10-epoxy-8-hydroxy-9-methyl-decatrienoic acid (EDA; Figure 4); (Nakashima *et al*., 1985; Nakatsuka *et al*., 1986; Kohmoto *et al*., 1993). Resistance to pear and strawberry pathotypes, and also for the apple pathotype, is controlled by a single recessive allele (Kohmoto *et al*., 1993; Tsuge *et al*., 2012). Inheritance of ABS resistance in citrus has been described as a monogenic control (Dalkilic *et al*., 2005; Gulsen *et al*., 2010), also controlled by a single recessive allele. Therefore, resistant citrus cultivars are considered to be recessive homozygous for this locus, whereas susceptible cultivars could be heterozygous or dominant homozygous.

However, all these resistance/susceptibility evaluations have been conducted on diploid genotypes, and data on triploid segregations are still unknown. Moreover, there is no information about genetic location of ABS resistance locus or markers strongly linked to it. This information should be very useful in order to perform marker-assisted selection in very early stages of a breeding program, both at diploid and triploid level, and to identify candidate resistance genes.

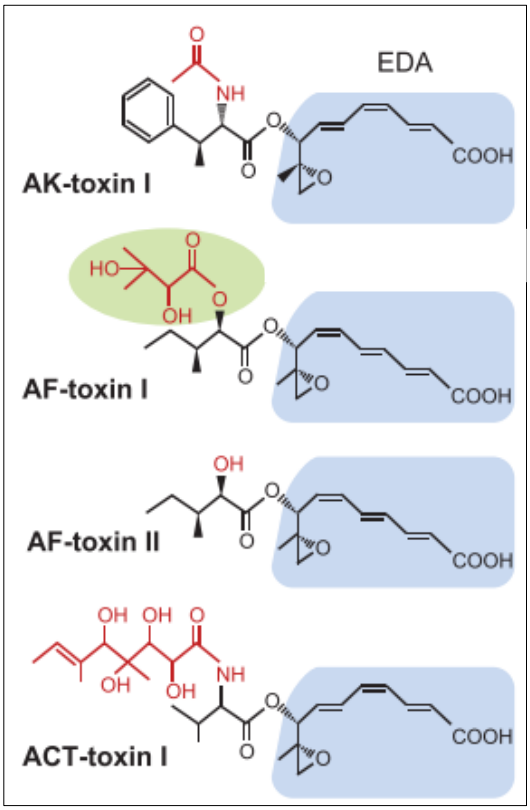

Figure 4. Chemical structures of HSTs produced by *Alternaria alternata*. AK-toxin I produced by the pear pathotype, AF-toxins (I and II) produced by the strawberry pathotype and ACT-toxin I produced by the tangerine pathotype, sharing the EDA: 9,10-epoxy-8-hydroxy-9-methyl-decatrienoic acid. From Tsuge *et al.* (2012).

4. BREEDING FOR SEEDLESS MANDARINS

Seedlessness is one of the most important economic traits relating to fruit quality for fresh-fruit marketing, and it is also desirable for juice industry because of the unfavourable aromatic compounds associated with the presence of seeds in the fruit (Ollitrault *et al.*, 2008). The presence of a large number of seeds in citrus fruits greatly decreases consumer acceptability, even in fruits with high organoleptic quality (Navarro *et al.*, 2004). Therefore, presence of seeds is considered as an obstacle in releasing newly selected high quality mandarins (Vardi *et al.*, 1996). Development of seedless cultivars has become a major goal in many citrus breeding programs around the world (Navarro *et al.*, 2004; Recupero *et al.*, 2005; Ollitrault *et al.*, 2007b; Roose and Williams, 2007; Aleza *et al.*, 2010a; Cuenca *et al.*, 2010).

A citrus cultivar is considered to be seedless if it is able to produce normal fruits that contain no seeds, aborted seeds, or a significantly reduced number of seeds. Strong sterility or self-incompatibility coupled with parthenocarpy is necessary for stable seedless fruit production (Ollitrault *et al.*, 2008). Parthenocarpy is the ability of a cultivar to produce fruits without fertilization. Therefore, parthenocarpy, coupled with a lack of cross- or self-fecundation, can yield seedless fruit (Vardi *et al.*, 2008). Some level of parthenocarpy seems to be widely present in citrus germplasm, resulting in the production of seedless fruit, even in the absence of pollination (Chao, 2005). For this reason, the cultivation of self-incompatible parthenocarpic genotypes like clementines in isolated blocks results in seedless fruit production (Vardi *et al.*, 2008). This implies that, for diversification programs, male and female sterility must be selected, particularly for areas such as the Mediterranean Basin, where the main seedless easy peeler cultivar is the self-incompatible clementine (Ollitrault *et al.*, 2008).

From a genetic point of view, seedlessness in citrus can be achieved at diploid level essentially by managing sterility or self-incompatibility, genetic transformation or mutation (spontaneous or induced) of diploid (seedy) elite genotypes. Other widely used option is ploidy manipulation, to produce triploid hybrids.

4.1. Seedlessness at the diploid level

4.1.1. Managing sterility and self-incompatibility

Various levels of male sterility at diploid level were reported in citrus. Chromosome aberrations (Iwamasa, 1966), such as asynapsis, reciprocal translocation, inversion or failure of the spindle formation are the most widespread factors of pollen sterility (Nakamura, 1943; Raghuvanshi, 1962; Iwamasa and Iwasaki, 1963; Iwamasa, 1966). Other mechanisms of male sterility not caused by chromosome aberrations are also known, such as anther abortion or early degeneration of pollen mother cells (Osawa, 1912; Frost, 1948; Iwamasa, 1966). To

develop new seedless cultivars efficiently, genetic analysis of male sterility has been conducted and particularly the analysis of anther abortion observed in 'satsuma' mandarins or 'Encore' mandarin. This male sterility is due to nucleo-cytoplasmic interaction and is probably controlled by more than two major nuclear genes (Iwamasa, 1966; Nakano *et al.*, 2001; Yamamoto *et al.*, 2001). Seedless cultivars with aborted anthers have been released in Japan (Nishiura, 1964; Matsumoto *et al.*, 1991)). A recent study reveals that dysfunctional mitochondria seemed to cause male sterile phenotype in a cybrid pummelo (Zheng *et al.*, 2012).

Female sterility is also a very important trait closely related to seedlessness and is a heritable characteristic (Yamamoto *et al.*, 2001). Female sterility results from degeneration of embryo sac (Osawa, 1912) or abortion of zygote (Nesumi *et al.*, 2001) due to chromosome aberrations.

The gametophytic self-incompatibility found in citrus is a genetically controlled phenomenon preventing seed set in self-pollinated plants producing functional gametes. In self-incompatible accessions, no pollen tubes are found in the ovaries (Ton and Krezdorn, 1967). Various accessions including almost all of pummelo, some mandarins and several natural or artificial hybrids are self-incompatible. However, many of them produce seedy fruits because of their female fertility and cross-pollination (Miwa, 1951; Mustard *et al.*, 1956; Krezdorn and Robinson, 1958).

4.1.2. Spontaneous mutations and induced mutagenesis

Spontaneous mutations have played an important role in the development of new cultivars, characterized by better fruit colour, longer time periods to fruit maturity, and seedlessness. Despite of the very low frequency of spontaneous mutation, the majority of clementines, satsumas and oranges cultivars arose from spontaneous mutations (Vardi *et al.*, 2008).

Induced mutagenesis can improve characters from the original cultivar, maintaining its genetic background (Hearn, 1984; Spiegel-Roy, 1990; Spiegel-Roy and Vardi, 1992; Vardi *et al.*, 1996; Shu, 2009). Gamma irradiation (^{60}Co) and chemical mutagens are used to induce mutations on seeds or buds, producing several times more mutants than the spontaneous mutation (Zhang *et al.*, 1988). Traits induced by mutagenesis include plant size, blooming time and fruit ripening, fruit colour, alterations in self-compatibility and resistance to pathogens (Predieri, 2001). 'Star Ruby' was the first citrus commercial cultivar obtained by irradiating seeds of 'Hudson' grapefruit; later, 'Rio Red' grapefruit was obtained by irradiation of 'Ruby Red' grapefruit (Hensz, 1971).

Seedlessness is the main goal in the induced mutagenesis of diploid elite cultivars. Examples of recently released seedless cultivars, obtained by induced mutation of diploid mandarin budwood are: 'Nulessin' (from 'Clemenules'), 'Mor' (from 'Murcott'), 'Orri' (from 'Orah'), 'Tango' (from 'Nadorcott'), 'DaisySL' (from 'Daisy'), 'FairchildSL' (from 'Fairchild'),

'NovaSL' (from 'Nova'), 'Murta' (from 'Murcott'), 'Moncalina' (from 'Moncada') and 'Nero' (from 'Clemenules').

4.1.3. Genetic transformation

Traditional breeding programs have been hampered by the long citrus juvenile phase, polyembryony, self-incompatibility and limited knowledge on trait inheritance. Genetic engineering provides a fast and effective way to introgress specific genes in plants aiming to improve targeted traits. Inducing parthenocarpy by genetic engineering and seed-ablated strategy by expressing the cytotoxin gene (*Barnase*) has been experimented for molecular breeding of seedless citrus. (Li *et al.*, 2002; Li *et al.*, 2003) reported the generation of 'Ponkan' and 'Valencia' sweet orange transgenic plants, through *Agrobacterium*-mediated transformation of embryogenic calluses with a chimeric ribonuclease gene (*barnase*) under the control of an anther tapetum-specific promoter (pTA29). Similar approach was developed by (Koltunow *et al.*, 2000) in Mexican limes for decreasing seed set. However, the commercialization of these products is limited by consumers' reluctance to buy genetically modified foods (Vardi *et al.*, 2008).

4.2. Ploidy manipulation for seedlessness. Recovery of triploid hybrids

4.2.1. Origin of polyploid citrus

Polyploidy is a major component of angiosperms evolution (Grant, 1981; Soltis and Soltis, 1993; Wendel and Doyle, 2005). Many plant species result from autopolyploidization or allopolyploidization events and polyploidization should be considered as the most common of the sympatric speciation mechanism (Otto and Whitton, 2000). For a long time, chromosome doubling (autopolyploidization) was considered by most authors (Stebbins, 1971) as the major mechanism leading to polyploidy. However, Harlan and DeWet (1975) argued that spontaneous chromosome doubling must be relatively rare, while polyploidization arising from unreduced gametes (allopolyploidization) seems much more frequent. These conclusions are now assumed by numerous plant evolutionists (Bretagnolle and Thompson, 1995; Ramsey and Schemske, 1998; Ramsey, 2007).

Diploidy is the general rule in *Citrus* and its related genera, with a basic chromosome number x=9 (Krug, 1943). However, some triploid and tetraploid genotypes have been early detected in citrus germplasm or seedlings (Longley, 1925; Lapin, 1937; Iwasaki, 1943). Longley (1925) was the first to formally identify a tetraploid wild form, the 'Hong Kong' kumquat. Other examples of the few natural polyploids found in the germplasm of *Aurantioideae* are the triploid 'Tahiti' lime, tetraploid strains of *Poncirus trifoliata*, allotetraploid *Clausena excavata*, tetraploid *Clausena harmandiana* and hexaploid *Glycosmis pentaphylla* (Ollitrault *et al.*, 2008).

In citrus germplasm, apomictic (nucellar polyembryony) and non-apomictic genotypes are found (Cameron and Frost, 1968). Spontaneous autotetraploidization seems to occur frequently in apomictic citrus genotypes by chromosome doubling of nucellar tissue (Aleza *et al.*, 2011). In non-apomictic genotypes, doubled-diploid plants are not found in the citrus germplasm, but can be artificially produced by treatment of micro-grafted shoot-tips or embryogenic callus with colchicine and oryzalin to achieve chromosome doubling (Tachikawa *et al.*, 1961; Barret, 1974; Gmitter and Ling, 1991; Gmitter *et al.*, 1991; Aleza *et al.*, 2009b). Stable tetraploid plants of different mandarins have been developed for use as male and female parents in interploid hybridizations (Aleza *et al.*, 2012c, 2012d).

Somatic hybridization of two diploid genotypes also allows obtaining tetraploid plants (Grosser *et al.*, 2010; Grosser and Gmitter, 2011). Tetraploid genotypes arising from chromosome doubling, chemical treatments or somatic hybridization are widely used as male parents in breeding programs (Aleza *et al.*, 2012c). Also, tetraploid genotypes are currently interesting to be used as rootstocks (Saleh *et al.*, 2008; Allario *et al.*, 2011; Dambier *et al.*, 2011; Grosser and Gmitter, 2011).

Spontaneous triploidization in citrus has been reported to arise from unreduced (2*n*) gametes (Soost, 1987; Iwamasa *et al.*, 1988; Luro *et al.*, 2004), probably due to the abortion of the second meiotic division in the megaspore (Esen *et al.*, 1979; Luro *et al.*, 2004), whereas non-reduced pollen seems to be extremely rare in citrus (Esen *et al.*, 1979; Luro *et al.*, 2004). It appears that most of the spontaneous triploids arising from diploid parents are found in small and abnormal seeds (Esen and Soost, 1971, 1973; Geraci *et al.*, 1975; Aleza *et al.*, 2010b), which are unlikely to germinate without implementation of an embryo rescue procedure. Unreduced gametes due to meiotic restitution and/or other variant meiotic events and resultant polyploidy have been documented in many plant species (Dewitte *et al.*, 2012; Singh *et al.*, 2012).

4.2.2. Methods for triploid breeding

Obtaining triploid hybrids has become an important breeding strategy to develop new seedless citrus commercial varieties (Starrantino, 1992; Navarro *et al.*, 2004; Ollitrault *et al.*, 2008). Indeed, many triploid varieties have been released from citrus breeding programs worldwide, including 'Oroblanco' and 'Melogold' grapefruits (*C. grandis* × *C. paradisi*) (Cameron and Bernett, 1978; Soost, 1987), and 'Winola' (Spiegel-Roy and Vardi, 1992). More recently, 'Shasta Gold®', 'Tahoe Gold®' and 'Yosemite Gold®' triploid mandarins were released by the USA program in California (Roose *et al.,* 2002); 'Tacle', 'Clara', 'Mandared', 'Mandalate' and other triploid mandarins were released by the Italian program (Recupero *et al.*, 2005) and 'Garbí' (Aleza *et al.*, 2010a) and 'Safor' mandarins (Cuenca *et al.*, 2010) were released by the Spanish program.

Triploid genotypes have generally been considered to be an evolutionary dead-end because they have very low fertility and tend to produce aneuploid gametes, due to problems of chromosome pairing during meiosis (Otto and Whitton, 2000; Ollitrault *et al.*, 2008). Cytogenetic studies have shown that during meiosis of citrus triploid hybrids, trivalent, bivalent and univalent associations are formed (Cameron and Frost, 1968), producing sterile gametes. Moreover, the abortion of megasporogenesis during the period between the first division of embryo-sac and the fecundated egg-cell is common (Fatta Del Bosco *et al.*, 1992). For this reason, citrus triploid plants are generally sterile, although they can occasionally produce fruits with very few seeds.

Triploid hybrids have been produced by sexual hybridizations of 2*x* × 2*x*, 2*x* × 4*x* and 4*x* × 2*x*, endosperm culture (Wang and Chang, 1978; Gmitter *et al.*, 1990) or by diploid + haploid somatic hybridizations (Kobayashi *et al.*, 1995; Ollitrault *et al.*, 1996b). In citrus germplasm, apomictic (polyembryonic) and non-apomictic (monoembryonic) genotypes can be found (Frost and Soost, 1968). The majority of citrus genotypes are apomictic, with the exception of all citrons (*C. medica*), pummelos (*C. grandis*), clementines *(C. clementina*) and some mandarins. One of the major limitations of any sexual hybridization strategy in citrus is that it is extremely difficult to use apomictic cultivars as female parents to produce large progeny and monoembryonic genotypes have to be used as female parents. Concerning polyembryonic cultivars, (Wakana *et al.*, 1981) showed that triploid zygotic embryos should be found with diploid nucellar embryos in small seeds. However, the practical possibility to select these triploid embryos is limited greatly by the polyembryony (Geraci *et al.*, 1977; Wakana *et al.*, 1981). To avoid this problem, Geraci *et al.* (1977) proposed a very early rescue of zygotic embryos from immature fruits, but it seems that selection of triploids from polyembryonic seedlings has not found real applications in citrus breeding.

Sexual polyploidization (2x × 2x crosses): In this case, spontaneous triploid hybrids arise from the union of a 2*n* megagametophyte with haploid pollen (Cameron and Frost, 1968; Esen and Soost, 1971, 1973; Geraci *et al.*, 1975; Ollitrault *et al.*, 1996a; Luro *et al.*, 2004; Chen *et al.*, 2008a; Aleza *et al.*, 2010b). The recovery of triploid citrus hybrids arising from unreduced (2*n*) megagametophytes produced by diploid plants was described in the 1970s (Esen and Soost, 1971, 1973). Cytogenetic studies (Esen and Soost, 1971) showed that triploid embryos are associated with pentaploid endosperm, indicating that triploid hybrids result from the fertilization of unreduced ovules by normal haploid pollen. Moreover, this ratio between the ploidy level of embryos and endosperm is responsible for seed size reduction. The frequency of such events is generally low (Cameron and Frost, 1968; Esen and Soost, 1971; Geraci *et al.*, 1975) and extensive breeding programs based on this type of hybridization require very effective methodologies for embryo rescue and ploidy evaluation of large progenies. The mechanisms that lead to unreduced gamete formation and their genetic consequences are explained at point 5.1 of this introduction.

Interploid crosses (2*x* × 4*x* and 4*x* × 2*x*): Interploid crosses have been used by breeders to obtain triploid hybrids. Triploid plants recovered from interploid crosses result from the fecundation of a diploid gamete arising from the tetraploid parent and a haploid gamete arising from the other parent. Taking advantage of the spontaneous chromosome doubling in apomictic genotypes, most of the early triploid breeding projects were based in the use of these spontaneous tetraploid genotypes as male parents (2*x* × 4*x* hybridization; Cameron and Frost, 1968; Esen *et al.*, 1978; Oiyama *et al.*, 1981; Starrantino and Recupero, 1981; Aleza *et al.*, 2012c). More recently, the recovery of tetraploid monoembryonic genotypes (Aleza *et al.*, 2009b) allowed developing the 4*x* × 2*x* strategy at large scale (Aleza *et al.*, 2012d).

Segregation and recombination of tetraploids is complex (Wu *et al.*, 2001a, 2001b). Indeed, it depends on (i) preferential pairing between homeologous chromosomes that defines the proportion of bivalent and multivalent formation (Wu *et al.*, 2001b) and (ii) double reduction frequency in case of tetravalent formation (Mather, 1936). More details on tetraploid meiosis and resulting structure of diploid gametes are given in paragraph 5.2.

Involvement of somatic hybridization in triploid breeding: Somatic hybridization via protoplast fusion is an approach that has been used in the creation of triploids and novel tetraploid germplasm for improving rootstock and scion varieties (Raza *et al.*, 2003; Ollitrault *et al.*, 2007a). The first somatic hybrid of citrus was produced by Ohgawara *et al.* (1985) and numerous inter and intrageneric somatic hybrids have been produced until now (Grosser and Gmitter, 1990; Louzada *et al.*, 1992; Louzada *et al.*, 1993; Ollitrault *et al.*, 1995; Grosser *et al.*, 2000; Mendes *et al.*, 2001; Grosser *et al.*, 2010; Dambier *et al.*, 2011; Grosser and Gmitter, 2011). Somatic hybridization could be used to produce quality tetraploid breeding parents that can be used in a conventional interploid breeding program to generate seedless triploids or to produce them directly by haploid + diploid protoplast fusion (Ollitrault *et al.*, 1996b; Guo *et al.*, 2000). However, a successful somatic hybridization program requires a complete and long-term effort, and mainly, this last approach have failed to have a major impact on crop development because of difficulties in protoplast isolation, culture and plant regeneration in many elite crop genotypes, small availability of haploid parents and the elevated ploidy levels resulting when somatic hybrid plants could be produced (Grosser *et al.*, 2000; Grosser and Gmitter, 2011).

Endosperm culture: This technique could be a tool to overcome the barriers of sexual hybridization that results from nucellar embryony and can theoretically be applied to all germplasm with female fertility. Successful regeneration of triploid plantlets has been reported by Wang and Chang (1978) and Gmitter *et al.* (1990). However, the step of shoot or embryos regeneration from endosperm calli appears a critical step (Jaskani *et al.*, 1996) that practically limits greatly the application of this technique for breeding purpose. Indeed, it appears difficult to obtain large recombining populations to apply efficient field selection.

5. MECHANISMS OF SEXUAL POLYPLOIDIZATION AND GENETIC STRUCTURES OF TRIPLOID POPULATIONS

5.1. Mechanisms of unreduced gamete formation (2*x* × 2*x* crosses)

Although polyploids can originate by an increase of chromosome number either during somatic growth or during meiosis, the major route is now considered to be via the formation of unreduced (2*n*) gametes (Harlan and DeWet, 1975). The formation of 2*n* gametes occurs in several plant species and can give rise to triploids that can serve as a bridge to the establishment of an even set of chromosomes in subsequent generations (Henry *et al.*, 2005).

Unreduced gametes most commonly arise through meiotic defects and the so-called meiotic nuclear restitution was described for the first time by Rosenberg (1927). Meiosis is a specialized cell division that is essential for sexual reproduction. It involves a single round of DNA replication followed by two rounds of chromosome division to produce cells with half the chromosome number of the mother cell. During meiotic prophase I, the meiosis-specific events of pairing and recombination between homologous chromosomes occur. These processes are important not only for generating genetic variability in the offspring but also for establishing the attachments between chromosomes required for the subsequent divisions (Brownfield and Kohler, 2011). In the first meiotic division, the homologous chromosomes are separated in what is referred to as a reductional division. Second meiotic division resembles mitosis in that it involves the separation of sister chromatids and is referred to as an equational division.

Mechanisms underlying meiotic restitution are still unclear, although more attention has been placed on this complex meiotic process in recent years (Cai and Xu, 2007). Several meiotic aberrations related to spindle formation, spindle function and cytokinesis lead to unreduced gamete formation in plants. Up to seven major mechanisms of 2*n* gamete formation have been cytogenetically characterized: premeiotic doubling; first-division restitution (FDR); chromosome replication during the meiotic interphase; second-division restitution (SDR); postmeiotic doubling; indeterminate meiotic restitution and apospory (Peloquin *et al.*, 1989; Lim *et al.*, 2001; Dewitte *et al.*, 2012). However, FDR and SDR are considered the predominant mechanisms of 2*n* gamete formation (Bretagnolle and Thompson, 1995; Tavoletti *et al.*, 1996; Cai and Xu, 2007). Failure of the first (FDR) or second division (SDR) leads to the formation of restitution nuclei with unreduced chromosomes. If the first meiotic division fails at anaphase I, all the chromosomes stay on the equatorial plate to form one restitution nucleus with the same chromosome number as the mother cell at the first division, which usually undergoes normal second meiotic division. This failure is characterized by an equational division of the entire chromosome complement (as in mitosis) and the formation of two nuclei with unreduced chromosome number. In case of SDR, the first meiotic division proceeds normally, but the second division fails at anaphase II, resulting in two nuclei with unreduced chromosomes (Cai and Xu, 2007). Therefore, a FDR 2*n* gamete contains non-sister chromatids, while a SDR 2*n* gamete contains two sister chromatids [(Tang and Luo, 2002); Figure 5]

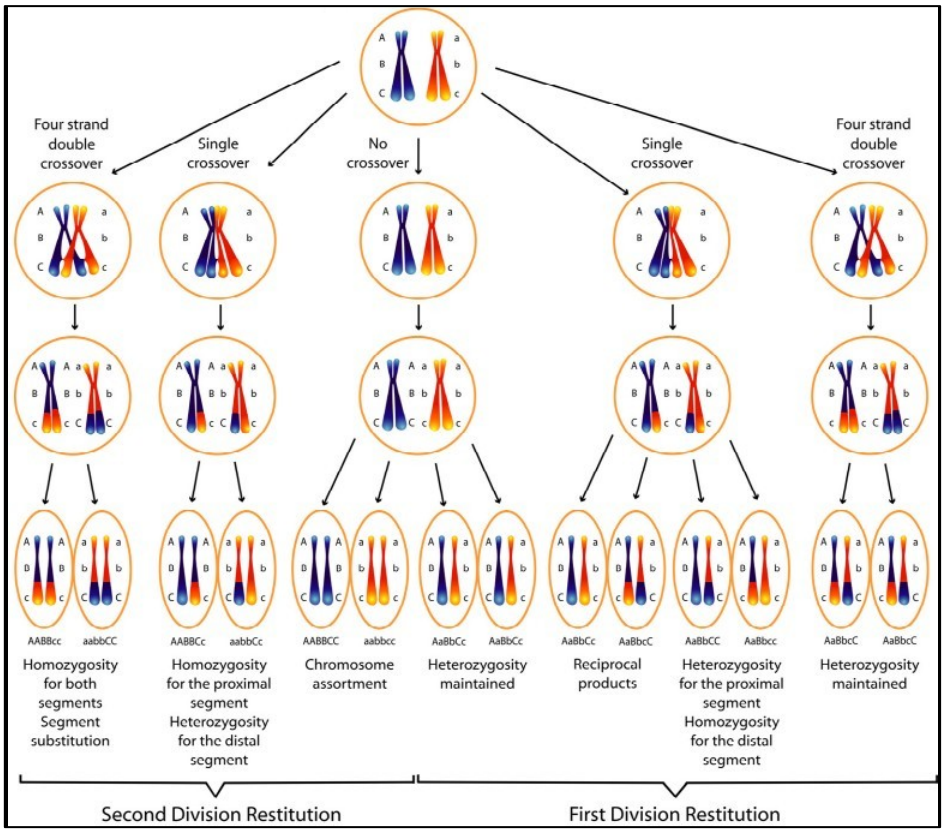

Figure 5. Genetic constitution of unreduced gametes resuting from FDR and SDR mechanisms. From Barba-Gonzalez *et al.* (2008).

From the cytogenetic point of view, several mechanisms lead to FDR or SDR-unreduced gamete formation. Semi-heterotypic division (Rosenberg, 1927), mitotised-meiosis (Stebbins, 1932; 1950), pseudohomoeotypic division (Gustafsson, 1935), aberrant cytokinesis (Ramanna, 1974), parallel spindle (Mok and Peloquin, 1975) and fused spindle (Ramanna, 1979) produce FDR 2*n*-gametes (Ramanna, 1979; Peloquin *et al.*, 1989; Vorsa and Rowland, 1997). In the same way, mechanisms genetically equivalent to SDR, known as premature cytokinesis 1 and 2 (Mok and Peloquin, 1975) have been reported (Veilleux, 1985; Bretagnolle and Thompson, 1995).

In addition to that, other mechanisms have been described leading to unreduced gametes genetically different than those coming from SDR or FDR. Lim *et al.* (2001) described the Indeterminate meiotic restitution (IMR) as a novel type of meiotic nuclear restitution mechanism, in which numerically disproportionate numbers of chromosomes occur due to a restitution mechanism which cannot be categorized as either FDR or SDR. In this case, although the 2*n*-gamete can have the euploid number, each set has not contributed the same number (Dewitte *et al.*, 2012). Premeiotic (chromosome doubling before meiosis) and postmeiotic (chromosome doubling after meiosis) restitution and cytomixis have also been proposed as possible mechanisms for the production of 2*n* gametes (Bastiaanssen *et al.*, 1998).

The identification of the mechanisms behind the formation of 2*n* gametes is complex. However, the use of cytological techniques (genomic in situ hybridization-GISH or fluorescent in situ hybridization-FISH) or marker analysis on polyploid progeny may provide accurate or additional information on these mechanisms (Lim *et al.*, 2001; Crespel and Gudin, 2003; Dewitte *et al.*, 2012). Molecular cytological approaches have been successfully used in the case of allopolyploids, where the constituent genomes can be clearly discriminated. This includes the unequivocal identification of not only genomes and individual alien chromosomes but also recombinant segments in the sexual polyploid progenies. Through DNA in situ hybridization, genomes of allopolyploids can be more critically assigned and intergenomic translocations and recombination events can be detected (Takahashi *et al.*, 1997; Karlov *et al.*, 1999; Lim *et al.*, 2001; Ramanna and Jacobsen, 2003; Barba-Gonzalez *et al.*, 2005). However, although both GISH and FISH are powerful tools for parental genome analysis, few studies in citrus have been reported, partly because citrus chromosomes are very small and indistinguishable (Barba-Gonzalez *et al.*, 2005; Jaskani *et al.*, 2007). In contrast, molecular marker analysis has proved as a very useful tool to estimate the heterozygosity transmission through the diploid gametes to polyploid progenies and, therefore, to identify the mechanism underlying unreduced gamete formation (Barone *et al.*, 1995; Vorsa and Rowland, 1997; Bastiaanssen *et al.*, 1998; Barcaccia *et al.*, 2003; Luro *et al.*, 2004; Chen *et al.*, 2008a; Hayashi *et al.*, 2009).

Although production of 2*n* gametes is highly affected by environmental factors (De Storme and Geelen, 2013), several heritability assays show that the rate of 2*n* gamete production is under a strong genetic control and is often determined by a few loci (Ramsey and Schemske, 1998; Barcaccia *et al.*, 2003; de Storme and Geelen, 2011). Mok and Peloquin (1975) and Qu *et al.* (1995) suggested that four recessive genes determined the 2*n* gamete formation in potatoes. Zhang *et al.* (2000; 2007) indicated that a single major recessive gene determined the 2*n* gamete formation in Chinese cabbage, and it is tightly linked to a centromere. Recent research on *Arabidopsis thaliana* has led to many advances in elucidating the molecular mechanisms underlying unreduced gamete formation, as well as the first genes in which mutations result in the production of viable 2*n* gametes (Brownfield and Kohler, 2011). D'Erfurth *et al.* (2008) were the first to successfully isolate and characterize the *Arabidopsis thaliana Parallel Spindle1* (*AtPS1)* gene involved in 2*n* gamete production due to an abnormal orientation of spindles at meiosis II. Mutations in *Arabidopsis DYAD/SWITCH1* and maize *AGO104* and *AM1* induce a complete loss of meiosis I, and consequently convert the meiotic cell cycle into a mitotic one (D'Erfurth *et al.*, 2009). De Storme and Geelen (2011) observed and characterized another unreduced gamete producer mutant called *jason*. Mutants of two other genes (*OSD1/GIG1* and *TAM/CYCA2;1)* were shown to omit the second meiotic division in both male and female sporogenesis at high frequency resulting in the formation of highly homozygous 2*n* pollen and egg cells (d'Erfurth *et al.*, 2009).

5.1.1. Heterozygosity restitution in the unreduced gametes and centromere mapping

The genetic structures of 2*n* gametes depend on their origin mechanisms and are very different for SDR and FDR. Such origin particularly affects the transmission of the parental heterozygosity in relation to the distance to the centromere. With FDR, the non-sister chromatids transmit the parental heterozygosity from the centromere to the first crossover point. With SDR the two sister chromatids are homozygous between the centromere and the first crossover point. It results in that the SDR 2*n* gametes have lower levels of heterozygosity than the FDR ones (Bretagnolle and Thompson, 1995).

The heterozygosity transmitted through 2*n* gametes can be estimated by analysis of the progeny with molecular markers. Several reports on *Solanum spp.* indicate that FDR gametes transmit 70-80% of the parental heterozygosity, whereas SDR gametes transmit about 30-40% (Douches and Quiros, 1988; Werner and Peloquin, 1990, 1991; Barone *et al.*, 1995). Such values were also reported in other crops: roses, ryegrass, *Begonia* and *Vaccinium darrowi* (Vorsa and Rowland, 1997; Crespel and Gudin, 2003; Chen, 2007; Dewitte *et al.*, 2012). In general, FDR is more than twice as effective in transmitting heterozygosity as SDR (Peloquin *et al.*, 2008).

For both FDR and SDR mechanisms, the rate of parental heterozygosity transmission (H) is a function of the rate of single crossing over (Co) between the centromere and the considered locus (Zhao and Speed, 1996). From the centromere to the telomere, it varies between 0 and 100% for SDR (H=Co) and from 100 to 50% for FDR (H=1-1/2Co), with the hypothesis that single chiasma frequency between the centromere and the locus vary between 0 for centromeric locus and 100% for telomeric ones (Park *et al.*, 2007). Such model assumes a complete chromosome arm interference (a first crossing over in chromosome arm totally prohibits the occurrence of a second one). Such assumption has been verified from half tetrad analysis (HTA) of some plants such as potatoes (Park *et al.*, 2007) and animals such as salmon (Lindner *et al.*, 2000). However, numerous cases of incomplete chromosome interference have been described both in plants and animals (Broman *et al.*, 2002; Copenhaver *et al.*, 2002; Esch and Weber, 2002; Saintenac *et al.*, 2009; Giraut *et al.*, 2011). Under no interference, H varies between 1 and 2/3 for FDR and from 0 to 2/3 for SDR, from the centromere to the telomere (Figure 6).

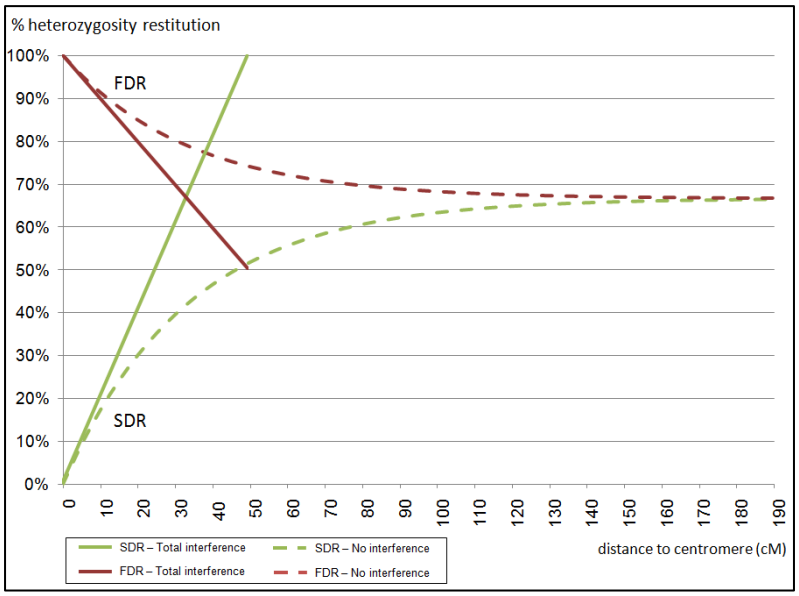

Figure 6. Theoretical curves of heterozygosity restitution as function of genetic distance to the centromere for FDR and SDR without chiasma interference and total interference models.

Conversely, the relation between the heterozygosity restitution at a locus and the distance of this locus to the centromere allow mapping centromere in linkage groups. The centromere is a specialized domain in most eukaryotic chromosomes that ensures delivery of one copy of each chromosome to each daughter cell during cell division by the mechanisms of kinetochore nucleation, spindle attachment, and sister chromatid cohesion.

Half-tetrad analysis (HTA) has been used in several plant and animal species for centromere mapping. HTA is an approach comparable to tetrad analysis, although based on only two chromatids from a single meiosis. These two chromatids remain together due to omission of the first (FDR) or the second (SDR) meiotic division, resulting in numerically unreduced gametes (Douches and Quiros, 1988; Kauffman *et al.*, 1995; Lindner *et al.*, 2000; Okagaki *et al.*, 2008). HTA could be used without a previous genetic map of the markers and can be applied with a predefined order of markers (Tavoletti *et al.*, 1996) or without any previous information about marker position (Da *et al.*, 1995). However, many models of HTA are based on the hypothesis of complete interference, which not always occurs. Indeed, pattern of heterozygosity restitution along the chromosome may vary between the different models of chiasma interference assumed (Zhao and Speed, 1998a). Therefore, a method which allows comparing a large range of partial interference model functions is suitable.

Once centromere is mapped, molecular markers close to it could be identified (Mendiburu and Peloquin, 1979) and, therefore, the mode of 2*n* gamete formation could be determined through analysis of the segregating marker in the progeny.

5.1.2. Mechanisms of unreduced gamete formation in citrus

Cytogenetic studies showed that triploid embryos are associated with pentaploid endosperm, indicating that triploid hybrids result from the fertilization of unreduced ovules by normal haploid pollen (Esen and Soost, 1971). Esen *et al.* (1979) proposed that 2*n* eggs result from the abortion of the second meiotic division in the megaspore (SDR) in citrus. This hypothesis has been corroborated for clementine (*Citrus clementina* Hort. ex Tan.) by molecular marker analysis (Luro *et al.*, 2004). However, Chen *et al.*, (2008a) proposed that 2*n* eggs of sweet orange (*C. sinensis* (L.) Osb.) resulted from first meiotic division restitution (FDR). It is, therefore, important to shed light on the mechanism underlying 2*n* gamete formation in a wide range of citrus genotypes used as female parents in 2*x* × 2*x* triploid breeding programs.

5.2. Diploid gametes from tetraploid parents (interploid crosses)

There is a basic distinction between autopolyploids and allopolyploids, both having multiple sets of chromosomes. Autopolyploids result from a variation of ploidy within a single species, so chromosomes are of the same type and have the same origin. Allopolyploids refer to the association of two differentiated genomes, through the process of hybridization and subsequent chromosome doubling, so both the type and the origin of chromosome are different (Gallais, 2003). Allopolyploids are generally considered to be much more prevalent in nature than are autopolyploids but, as detected from a growing number of genetic analyses, autopolyploids should be more common than previously assumed (Soltis and Soltis, 2000). An immediate consequence of polyploidy is the change in gametic and segregation frequencies (Comai, 2003).

In allopolyploids, identical or at least fully homologous genomes occur in pairs, but different pairs of a genome have a strong pairing barrier (Sybenga, 1996). Because only homologous chromosomes pair, allopolyploids strictly exhibit bivalent formation (two chromosomes pair) at meiosis and undergo disomic inheritance for each locus (Gallais, 2003).

For autopolyploids, all the chromosomes are homologous and have equal opportunities to pair at meiosis. Since pairing can start at different chromosomal sites, homologous chromosomes may switch partners, leading to multivalent formation (more than two chromosomes pair) and a type of inheritance called polysomic (Jackson and Jackson, 1996; Sybenga, 1996; Hauber *et al.*, 1999). Heterozygosity transmission from an autotetraploid to its diploid gamete depends on the double reduction frequency, defined as the probability of two sister chromatids occurring in the same gamete (Marsden *et al.*, 1987). Although segregation patterns depend on the frequency of crossing over between the centromere and a given loci, and therefore, they are expected to vary with the position of the locus on the chromosome, little empirical information concerning the frequency of double reduction is available (Gallais, 2003). This frequency assumes maximum values of 0 (random chromosome segregation), 1/7 (with

pure random chromatid segregation), and 1/6 (with maximum equational segregation; (Muller, 1914; Mather, 1936). The average rate of parental heterozygosity restitution is 2/3 in case of random segregation, while it will be only 5/9 in case of maximum chomatid equational segregation.

It appears that many tetraploids have not systematic disomic or tetrasomic inheritance, but display intermediary inheritance (Danzmann and Bogart, 1983; Marsden *et al.*, 1987; Jackson and Jackson, 1996). Indeed, allo- and autotetraploids (with disomic and tetrasomic inheritance, respectively) are the extremes of a range (Figure 7). In cases where parents are divergent but have retained enough homology to prevent exclusive preferential pairing, inheritance patterns intermediate between di- and tetrasomic can be expected (Stift *et al.*, 2008). For example, in citrus Kamiri *et al.* (2011) showed that a somatic hybrid between lemon and mandarin had an intermediate inheritance with preferential tetrasomic tendency, while an intergeneric somatic hybrid between *Poncirus* and mandarin displayed an intermediary inheritance with preferential disomic tendency (Kamiri *et al.*, 2012).

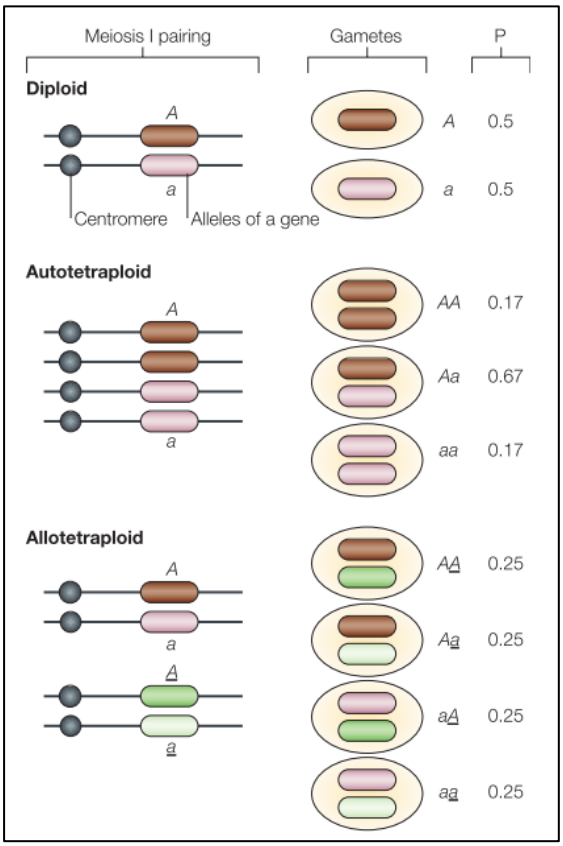

Figure 7. Meiotic arrangements and gametic output in diploids, autotetraploids and allotetraploids. A diploid heterozygote Aa produces two types of gamete in equal proportion (P). An autotetraploid with genotype AAaa produces, under the model of random chromosome segregation three types of gamete. An allotetraploid with genotype AaAa produces four types of gamete. From Comai (2003).

6. MOLECULAR RESOURCES FOR CITRUS GENETICS AND BREEDING

6.1. Molecular markers and genetic maps in citrus

Molecular markers are specific fragments of DNA that can be identified within the whole genome and are transmitted by the standard laws of inheritance from one generation to the next. The development of molecular markers based on DNA sequences has provided a useful tool for enlarging the knowledge of inheritance of economically important characters. A wide variety of DNA-based markers have been developed, both PCR-based and non-PCR-based (Agarwal *et al.*, 2008).

Since isozymes markers (Torres *et al*., 1978; 1982; Roose, 1988), several kinds of nuclear markers have been used for citrus genetic studies such as Random Amplified Polymorphic DNA [RAPDs; (Luro *et al*., 1994)], Sequence Characterized Amplified Regions [SCARs; (Nicolosi *et al*., 2000)], Restriction Fragment Length Polymorphism [RFLPs; (Federici *et al*., 1998)], Intersimple sequence repeat [ISSRs, (Fang *et al*., 1997)], Amplified Fragment Length Polymorphism (AFLPs; (Liang *et al*., 2007; Pang *et al*., 2007) and Cleaved Amplified Polymorphic Sequences (CAPs) from ESTs (Lotfy *et al*., 2003). Single-stranded conformational polymorphism (SSCP) analysis has been used for cytoplasm characterization (Olivares-Fuster *et al*., 2007). In the last 10 years, a limited number of Simple Sequence Repeat (SSRs) have been derived from genomic libraries (Kijas *et al*., 1997; Barkley *et al*., 2006; Froelicher *et al*., 2008): 56 SSRs were obtained from the Genbank citrus EST data (Chen *et al*., 2006) and more than 200 SSR markers have been developed (Luro *et al*., 2008) from the 1,600 microsatellite sequences from 37,000 ESTs characterized by Terol *et al*. (2007). The same group identified more than 7,600 SSRs from BAC end sequencing (Terol *et al*., 2008) that have been used to develop SSR markers allowing direct anchoring of the genetic and physical maps (Ollitrault *et al*., 2012a). In addition to genetic mapping, SSRs have been used for the analysis of genetic diversity (Luro *et al*., 2001; Barkley *et al*., 2006; García-Lor *et al.*, 2013b), characterization of somatic hybrids, discrimination between zygotic and nucellar seedlings (Ruiz *et al*., 2000; Ruiz and Asins, 2003), control of the origin of plants obtained by induced gynogenesis (Froelicher *et al.*, 2007), molecular characterization of triploid cultivars (Aleza *et al.*, 2010a; Cuenca *et al.*, 2010) and the analysis of the origin of unreduced gametes (Luro *et al.*, 2004; Chen *et al.*, 2008a).

More recently, the availability of large set of sequencing data has opened the way for insertion-deletion (InDel) and single nucleotide polymorphism (SNP) marker development (García-Lor *et al.*, 2012; Ollitrault *et al.*, 2012, 2012b; García-Lor *et al.*, 2013b).

A very important outbreak for efficient use of molecular markers in genetics and breeding is the availability of genetically mapped codominant markers anchored with the physical sequence. Such citrus reference map was established recently for clementine (Ollitrault *et al.*, 2012a). Despite the heterogeneous dispersion of markers, this medium density reference map

(961 markers for 1084.1 cM) constitutes a good framework for further marker-trait association studies, and it has been used to enable the chromosome assembly of the reference whole genome citrus sequence (Wu *et al.*, 2013).

6.2. Specific challenge for polyploid species genotyping

Despite the undisputable importance of polyploid plant species, the genetics of these plant species are less well known than those of their diploid counterparts, which is especially evident in research fields as DNA analysis (Esselink *et al.*, 2004). Indeed, the estimation of molecular marker allele copy number has long been considered a challenge for polyploid species with polysomic inheritance, while it is essential to assign the allelic configuration for different types of heterozygotes for accurate population genetic studies. In segregating polyploid progenies, the population genetic structure can provide relevant information about the underlying meiosis mechanisms that take place in the formation of these progenies, which also greatly affect character segregation (Hutten *et al.*, 1993; Tai and DeJong, 1997; Douches and Maas, 1998; Barcaccia *et al.*, 2003; Brownfield and Kohler, 2011). Moreover, allelic dosage can affect gene expression and phenotype. Therefore, the determination of allelic dosages is particularly important for marker/trait association studies (De Jong *et al.*, 2003; Sjoling *et al.*, 2005). However, there is a major problem to define which allele(s) occur in more than one copy when the number of displayed DNA alleles in a sample is less than the possible maximum number for that ploidy level in species with polysomic inheritance, such as SNP markers, which are mostly biallelic. For a triploid plant with two detected alleles, two heterozygous configurations are possible (2:1 or 1:2); for a tetraploid plant, three different allelic configurations are possible (3:1, 2:2 or 1:3). With higher ploidy levels, the number of possible allelic configurations becomes even larger (Esselink *et al.*, 2004).

Several techniques have been used to estimate allele dosage in polyploids, such as the useful MAC-PR method (Esselink *et al.*, 2004) for SSR markers; however, SSR analysis remains relatively costly and time consuming compared with present SNP genotyping methods. Moreover, with the increasing availability of EST databases and whole genome sequences, SNPs have become the most abundant and powerful polymorphic markers that can be selected along the entire genome (Edwards and Batley, 2010).

Techniques for SNP genotyping include allele-specific primer extension (Kwok, 2001), temperature-switch PCR (Tabone *et al.*, 2009), array methodologies (Ishikawa *et al.*, 2005), and targeted pyrosequencing (Ahmadian *et al.*, 2000). The application of SNP markers has been limited primarily to diploid organisms, while the application of these markers to polyploid organisms for allele dosage estimation remains limited. The usefulness of SNPlex™ (Berard *et al.*, 2009) and Illumina Golden Gate™ assays (Akhunov *et al.*, 2009) for polyploid wheat genotyping has been demonstrated. However, these techniques are more suitable for

genotyping large numbers of samples over numerous markers than for performing small scale analysis.

It is, therefore, important to develop alternative methods that offer a wider spectrum of genotyping possibilities to infer SNP allelic configurations in polyploid plants, particularly in citrus, for small- to larger-scale projects.

6.3. Whole-genome sequencing

Genomics has provided new tools for crop improvement, helping to identify and select candidate genes responsible of agronomic characters of interest, and allowing the development of fast methods to incorporate these characters into crop plants (Terol *et al.*, 2007).

Citrus, with a basic chromosome number of 9, has a relatively small genome size. Haploid genomes of *C. sinensis* and *C. clementina* are, respectively, 380 Mb and 370 Mb of size (Arumuganathan and Earle, 1991).

BAC libraries of sweet orange, clementine, and satsuma have been established in the last few years in Spain, Japan, and the USA. The Spanish Citrus Genomic Consortium has constructed three BAC libraries from the clementine mandarin (EcoR I, Hind III, and MboI) containing a total of 57,000 clones with an average insert size of 120 kb (19x coverage). 28,000 BAC clones were end-sequenced and these sequences analyzed (Terol *et al.*, 2008). The Citrus Genome Analysis Team from Japan is engaged in the construction of a physical map of citrus by HICF (High-Information-Content Fingerprinting) analysis of a BAC library from the 'satsuma' mandarin consisting of 37,000 clones, with 13.3x coverage of the citrus genome. A BAC library of 'Ridge Pineapple' sweet orange was produced in USA (USDA-ARS, Ft. Pierce, Fl, USA) containing 18,432 clones (BamHI/Mbo I) with an average insert size of 145 kb, or an estimated 7x coverage. A total of 16,727 clones from this library have been fingerprinted and assembled into 472 contigs (http://phymap.ucdavis.edu:8080/citrus/).

A low-coverage (1.2X) shotgun sequence of the *C. sinensis* genome has revealed the difficulties related to high heterozygosity, and lead the International Citrus Genomic Consortium to select a haploid clementine (Aleza *et al.*, 2009a) as the model for whole citrus genome sequencing. The International Consortium for Citrus Genomics has recently released the whole genome sequence of the haploid clementine, assembled in pseudo-chromosome [www.phytozome.net, (Wu *et al*., 2013)]. This reference sequence is currently used as template to organize sequences of the highly heterozygous species such as *C. sinensis*, *C paradisi*, and *C. limon*.

Recently, Xu *et al.* (2013) sequenced and assembled a dihaploid genome of sweet orange and mapped the parental diploid genomic DNA sequence reads to the haploid reference genome to complete the construction of the heterozygous genome map. The availability of the

sweet orange and clementine genome sequence provides a valuable genomic resource for citrus genetics and breeding improvement.

In addition, the complete chloroplast genome sequence of *C. sinensis* was published by Bausher *et al.* (2006). It is 160.129 bp in length and contains 133 genes (89 protein-coding, 4 rRNAs, and 30 distinct tRNAs).

6.4. Marker-trait association studies and marker assisted selection

Quantitative Trait Loci (QTLs) and genetic association studies are used to find candidate genes or genome regions that contribute to a specific phenotypic trait, by testing for a correlation between phenotypic diversity and genetic variation (Lewis and Knight, 2012). Then, the identified DNA markers can be used to infer the presence of allelic variation in the genes underlying these traits and to assist plant breeding (marker-assisted selection – MAS). Selection can be carried out at the seedling stage, allowing more effectively selection of target genotypes and resulting in faster variety development and releasing (Bertrand *et al.*, 2008). This is of particular interest for tree species with long juvenile phase, such citrus.

The detection of major genes and QTLs controlling traits is based on the linkage disequilibrium between closely linked loci. Significant genetic association may be interpreted as either direct association, in which the genotyped molecular marker is the true causal variant conferring phenotypic variation; or indirect association, in which a molecular marker in linkage disequilibrium with the true causal variant is genotyped. Distinguishing between direct and indirect association is challenging and may require resequencing of the candidate region, dense genotyping of all available markers, or functional studies to confirm the role of a putative mutation in the phenotypic trait (Lewis and Knight, 2012). In citrus, only a few characters of agronomic interest have been linked to molecular markers, such as RAPD markers linked to dwarfing (Cheng and Roose, 1995), RAPD markers linked to a gene controlling fruit acidity (Fang *et al.*, 1997), SSR markers linked to CTV resistance from *Poncirus trifoliata* (Gmitter *et al.*, 1996; Yang *et al.*, 2003; Asins *et al.*, 2004; Bernet *et al.*, 2004), AFLP markers linked to nucellar embryony (García *et al.*, 1999; Kepiro and Roose, 2010) and the dominant PCR assay for the anthocyan content of pulp from blood orange due to a transposable element in the 5' extremity of the Ruby gene (Butelli *et al.*, 2012). Other characters of interest have been tagged to QTLs, such as salinity tolerance (Moore *et al.*, 2000) and nematode resistance (Ling *et al.*, 2000). Alternaria brown spot resistance has also been tagged with molecular markers. (Dalkilic *et al.*, 2005) reported two RAPD markers with loose linkage with the locus (15.3 cM and 36.7 cM far from ABS resistance locus in the same side). More recently, (Gulsen *et al.*, 2010) identified two flanking SRAP markers at 3 cM and 13 cM. However, markers tightly (or completely) linked to ABS resistance would greatly improve the selection of resistant genotypes in early development stages and avoid growing susceptible genotypes (false positives) or discard resistant ones (false negatives).

Objectives

Sexual polyploidization through unreduced gamete formation is currently a central approach in citrus triploid breeding programs aiming to develop new seedless cultivars. However, despite the undisputable importance of polyploid plant species, their genetics are less well known than those of their diploid counterparts, which is especially evident in research fields as DNA analysis. Indeed, the estimation of molecular marker allele copy number has long been considered a challenge for polyploid species with polysomic inheritance, while it is essential to assign the allelic configuration for different types of heterozygotes for accurate population genetic studies. For SSR markers, these allelic configurations in polyploids can be assigned using the microsatellite allele counting – peak ratios method (MAC-PR). However, SSR analysis remains relatively costly and time consuming compared with actual SNP genotyping methods; moreover, with the increasing availability of SNP information in citrus, an efficient method for SNP genotyping allowing inferring allele doses in polyploid genotypes appears essential for optimal integration of molecular tools in citrus triploid breeding. Allelic configurations in triploid segregating polyploid progeny, particularly in citrus, should provide useful information in the underlying meiosis mechanisms that take place in the formation of these progeny and other genetic studies.

The genetic structure of triploid progeny arising from unreduced gametes depends on their mechanism of origin, being very different for first-division restitution (FDR) and second-division restitution (SDR). Moreover, for a concrete gene, segregations in triploid populations also depend on the gene-centromere distance. In this context, deep knowledge of the mechanism underlying meiotic nuclear restitution producing unreduced gametes and location of centromeres are crucial to optimize breeding strategies based on sexual hybridization. Moreover, trait segregations in triploid populations still not well understood, due to the complexity of polyploid genomes and eventually non-Mendelian segregations. In the case of citrus, the inheritance of very few characters has been studied on triploid progenies and genetic determinants underlying their segregations have not been characterised in any case.

Alternaria brown spot (ABS) is a major fungal disease which affects susceptible mandarin cultivars and causes a substantial loss of production and fruit quality, limiting the number of usable breeding parents for triploid breeding. Knowledge on segregations for ABS-resistance for different breeding strategies and the identification of linked molecular markers would greatly improve the breeding efficiency for resistant cultivar selection.

The specific objectives of the PhD thesis are the following:

Objective 1: To implement a versatile SNP genotyping method to efficiently assign allelic configuration in polyploid plants

With the increasing availability of SNPs markers in citrus, which have become the most abundant and powerful polymorphic markers that can be selected along the entire genome, a new method to assign SNP allele dosage in an accurate, simple, and cost effective way, usable in small- to large-scale projects will be developed.

Objective 2: To shed light on the mechanism underlying unreduced gamete formation in a wide range of citrus genotypes used as female parents in 2*x* × 2*x* triploid breeding programs

The origin of unreduced megagametophytes in citrus is still controversial, although differences in genetic structures are expected for FDR and SDR origin. Moreover, segregations for concrete loci depend on the locus-centromere distance and are different for each mechanism. In this context, three objectives have been developed:

2.1 To develop and apply a method for centromere mapping which allows comparing a large range of partial interference model functions

2.2 To develop and apply a method for identifying unreduced gamete formation mechanism both at population and at individual level based on marker-centromere distances

2.3 To investigate the origin of unreduced gametes for sixteen mandarin genotypes used as female parents

Objective 3: To gain knowledge about inheritance, genetic and molecular determinism of ABS-resistance

The mode of inheritance for ABS-resistance has been previously studied at diploid level, but segregations on triploid populations still unknown. Moreover, development of molecular markers for an early selection of ABS resistant hybrids would greatly improve the efficiency of breeding programs. In this context, three objectives have been developed:

3.1 To analyze segregations of ABS-resistance for different parents and strategies of triploid breeding

3.2 To locate a chromosome region associated with the ABS-resistance and to identify candidate resistance genes in the located region

3.3 To develop molecular markers for marker-assisted selection in citrus breeding programs for ABS-resistance

The manuscript is structured in four chapters, corresponding to published or in preparation scientific articles:

Chapter 1: Assignment of SNP allelic configuration in polyploids using competitive allele-specific PCR: application to citrus triploid progenies. Cuenca, *et al*. (2013). *Annals of Botany* 111: 731 – 742. Objective 1.

Chapter 2: Multilocus half-tetrad analysis and centromere mapping in citrus: evidence of SDR mechanism for 2*n* megagametophyte production and partial chiasma interference in mandarin cv 'Fortune'. Cuenca *et al.* (2011). *Heredity* 107: 462 – 470. Objective 2.

Chapter 3: Maximum-likelihood method based on parental heterozygosity restitution of centromeric loci identifies SDR as the predominant mechanism leading to 2*n* megagametophytes in *C. reticulata*. Cuenca *et al*. *Annals of Botany*, Submitted. Objective 2.

Chapter 4: Genetically based location from triploid populations and gene ontology of a 3.3-Mb genome region linked to Alternaria brown spot resistance in citrus reveal clusters of resistance genes. Cuenca *et al*. (2013). *PLoS ONE* 8(10): e76755. Objective 3.

Chapter 1

Assignment of SNP allelic configuration in polyploids using competitive allele-specific PCR: application to citrus triploid progenies

José Cuenca, Pablo Aleza, Luis Navarro and Patrick Ollitrault

Annals of Botany (2013) 111: 731 – 742. doi:10.1093/aob/mct032

Abstract

Background: Polyploidy is a major component of eukaryote evolution. Estimation of allele copy numbers for molecular markers has long been considered a challenge for polyploid species, while this process is essential for most genetic research. With the increasing availability and whole genome coverage of SNP markers, it is essential to implement a versatile SNP genotyping method to efficiently assign allelic configuration in polyploids.

Scope: This work evaluates the usefulness of the KASPar method, based on competitive allele-specific PCR, for the assignment of SNP allelic configuration. Citrus was chosen as a model because of its economic importance, the ongoing worldwide polyploidy manipulation projects for cultivar and rootstock breeding, and the increasing availability of SNP markers.

Conclusions: Fifteen SNP markers were successfully designed that produced clear allele signals that were in agreement with previous genotyping results at the diploid level. The analysis of DNA mixes between two haploid lines (Clementine and pummelo) at 13 different ratios revealed very high correlation (average = 0.9796; SD = 0.0094) between the allele ratio and two parameters [theta angle = $\tan^{-1}$ (y/x) and $y' = y/(x+y)$] derived from the two normalised allele signals (x and y) provided by KASPar. Separated cluster analysis and analysis of variance (ANOVA) from mixed DNA simulating triploid and tetraploid hybrids provided 99.71% correct allelic configuration. Moreover, triploid populations arising from 2*n* gametes and interploid crosses were easily genotyped and provided useful genetic information. This work demonstrates that the KASPar SNP genotyping technique is an efficient way to assign heterozygous allelic configurations within polyploid populations. This method is accurate, simple, and cost effective. This genotyping technique has been successfully applied to two citrus triploid populations arising from 2*n* gametes and interploid crosses. Moreover, this method appears to be useful for quantitative studies, such as relative allele-specific expression analysis and bulk segregant analysis.

Introduction

Polyploidy is a major component of eukaryote evolution, particularly in angiosperms (Grant, 1981; Soltis and Soltis, 1993; Wendel and Doyle, 2005). Many plant species result from autopolyploidization or allopolyploidization events, and polyploidization is considered the most common sympatric speciation mechanism (Otto and Whitton, 2000). Despite the undisputable importance of polyploid plant species, the genetics of these plant species are less well known than those of their diploid counterparts. Indeed, the estimation of molecular marker allele copy number has long been considered a challenge for polyploid species with polysomic inheritance, while it is essential to assign the allelic configuration for different types of heterozygotes for accurate population genetic studies. In segregating polyploid progenies, the population genetic structure can provide relevant information about the underlying meiosis mechanisms that take place in the formation of these progenies, which also greatly affect character segregation (Hutten *et al.*, 1993; Tai and DeJong, 1997; Douches and Maas, 1998; Barcaccia *et al.*, 2003; Brownfield and Kohler, 2011). Moreover, allelic dosage can affect gene expression and phenotype. Therefore, the determination of allelic dosages is particularly important for marker/trait association studies (De Jong *et al.*, 2003; Sjoling *et al.*, 2005). When parents of a polyploid progeny share one allele, only the dosage allele estimation allows knowing the alleles transmitted by each parent to heterozygous progeny. Therefore, knowledge of allelic dosage in polyploids appears to be essential for studies using single nucleotide polymorphisms (SNPs) markers, most of which are biallelic.

Several techniques have been used to estimate allele dosage in polyploid genotypes or tissues. When analysing microsatellite markers (SSRs), the microsatellite allele counting – peak ratios method (MAC-PR; (Esselink *et al.*, 2004) is especially useful. However, SSR analysis remains relatively costly and time consuming compared with actual SNP genotyping methods. Moreover, with the increasing availability of EST databases and whole genome sequences, SNPs have become the most abundant and powerful polymorphic markers that can be selected along the entire genome (Edwards and Batley, 2010).

Several SNP genotyping methods have been developed. Some methods are based on electrophoretic separation after PCR amplification, including allele-specific primer extension (Kwok, 2001) and temperature-switch PCR (Tabone *et al.*, 2009). High-throughput genotyping can be obtained using array methodologies (Sapolsky *et al.*, 1999; Ishikawa *et al.*, 2005); other techniques are based on pyrosequencing™ (Ronaghi *et al.*, 1998; Ahmadian *et al.*, 2000). However, the application of SNP markers has been limited primarily to diploid organisms, while the application of these markers to polyploid organisms for allele dosage estimation remains limited. Rickert *et al*. (2002) reported the use of pyrosequencing™ in polyploid potatoes, with some sequence-specific limitations. The usefulness of SNPlex™ (Berard *et al.*, 2009) and Illumina Golden Gate™ assays (Akhunov *et al.*, 2009) for the genotyping of polyploid wheat has been demonstrated. For genotype calling in tetraploid species with SNP analysis using the Illumina GoldenGate™ array, Voorrips *et al.* (2011) developed an algorithm using mixture

models, but they assumed Hardy Weinberg equilibrium within the population, which does not occur in all segregating polyploid progeny. Microarray data (Kirov *et al.*, 2006; Meaburn *et al.*, 2006; Steer *et al.*, 2007) have also been used to estimate allelic frequencies in bulk populations or DNA pools, i.e., to perform genome-wide association scans. Array analysis is more suitable for genotyping large numbers of samples over numerous markers than for performing small scale analysis, as array analysis lacks flexibility in term of the numbers and panels of SNP loci that can be analyzed. Targeted pyrosequencing™ (Gruber *et al.*, 2002; Neve *et al.*, 2002; Wasson *et al.*, 2002; Lavebratt *et al.*, 2004) can be useful for performing allele frequency estimation for a few genes in pooled DNA, but this technique remains relatively costly and time consuming. It is, therefore, important to develop alternative methods that offer a wider spectrum of genotyping possibilities to infer SNP allelic configurations in polyploid plants in small- to larger-scale projects.

The KBiosciences Competitive Allele-Specific PCR SNP genotyping system (KASPar) is a homogeneous fluorescent endpoint genotyping system (Cuppen, 2007), that utilises a unique form of competitive, allele-specific PCR and combines the use of a highly specific 5'-3' exonuclease-deleted Taq DNA polymerase with two competitive, allele-specific, tailed forward primers and one common reverse primer. This system is simple and cost effective compared to other SNP genotyping assays and is well adapted to low-to-medium throughput genotyping projects (Chen *et al.*, 2010). This technology has been successfully applied to the study of humans, animals, and plants (Nijman *et al.*, 2008; Bauer *et al.*, 2009; Cortés *et al.*, 2011; Rosso *et al.*, 2011).

Citrus is mainly diploid. However, many modern breeding projects for the production of seedless mandarins based on the production of triploid hybrids (Ollitrault *et al.*, 2008; Aleza *et al.*, 2010b, 2012c, 2012d) and tetraploid rootstocks are promising (Saleh *et al.*, 2008; Dambier *et al.*, 2011; Grosser and Gmitter, 2011). Triploid populations in citrus can arise from unreduced gametes in crosses between diploid parents or from interploid (diploid × tetraploid or tetraploid × diploid) crosses. Sexual polyploidization resulting from 2*n* megagametophyte production is routinely exploited for triploid citrus breeding (Aleza *et al.*, 2010b). In such crosses, segregation of a marker depends on the parental genetic structure, the relative distance to the centromere, and the mode of restitution (First Division Restitution [FDR] or Second Division Restitution [SDR]). The MAC-PR method has been successfully applied to demonstrate the SDR origin of the 2*n* gametes arising from the 'Fortune' mandarin cultivar and to locate the centromere in one chromosome (Cuenca *et al.*, 2011). For interploid crosses, marker segregations are almost exclusively dependent on the parental genetic conformation and preferential chromosome pairing. SSR markers were also used to analyse the meiotic behaviour of a tetraploid interspecific somatic hybrid of *C. deliciosa* + *C. lemon* (Kamiri *et al.*, 2011), concluding that there was predominant tetrasomic segregation. However, the low availability of SSR markers displaying favourable parental allelic structure that can be used to differentiate male and female contributions to the hybrids limits such studies to just a few areas of the genome. Conversely,

large SNP resources have become available from extensive sequencing projects (Terol *et al.*, 2007; Terol *et al.*, 2008; Gmitter *et al.*, 2012; Ollitrault *et al.*, 2012b).

The objective of the present work was to evaluate the potential of the KASPar method to assess SNP allelic configurations in polyploid plants. Citrus was chosen as a model system because of its economic importance, the worldwide ongoing polyploidy manipulation projects for cultivar and rootstock breeding, and the increasing availability of SNP markers.

The quantitative value of the KASPar assay was estimated by pooling DNA from two haploid lines at several relative concentrations, simulating, among others, triploid and tetraploid heterozygous progeny. A method was developed for semi-automated polyploid genotype calling and applied for allelic configuration analysis of 170 triploid hybrids from two families arising from both sexual polyploidization and interploid crosses.

Material and methods

Sample preparation

DNA pool preparation:

Genomic DNA from two haploid lines, *Citrus maxima* (Burm.) Merr. (pummelo) cv 'Chandler' and *C. clementina* Hort. ex Tan. [clementine; (Aleza *et al.*, 2009a)] was isolated using a Plant DNAeasy Kit from Qiagen Inc. (Valencia, CA, USA) following the manufacturer's protocol. DNA concentrations were estimated with PicoGreen® and adjusted to 30 ng/µl. DNA from the two haploid lines were pooled at ratios of 9:1, 5:1, 3:1, 2:1, 3:2, 1:1, 2:3, 1:2, 1:3, 1:5, and 1:9. Five samples (replications) of each haploid line and pool were prepared and used to test the accuracy of the technique.

Simulation of triploid and tetraploid hybrid samples by pooling DNA from haploid lines:

Two subsets of haploid DNA pool, one corresponding to 2:1 and 1:2 ratios that simulated heterozygous triploid genotypes and the other corresponding to 3:1, 1:1, and 1:3 ratios that simulated tetraploid heterozygous genotypes, were jointly used with the haploid genotypes to test the capability of the technique to discriminate among different types of heterozygotes within triploid and tetraploid populations.

Natural triploid populations:

2x × 2x crosses: 'Fortune' (*C. clementina* × *C. tangerina*) and 'Willowleaf' (*C. deliciosa* Ten.) diploid mandarins and 39 triploid hybrids segregating from this cross (Aleza *et al.*, 2010b) were selected to test the accuracy of the technique by analysing two replicates of each sample. Moreover, 86 triploid hybrids from 'Clementine' (*C. clementina*) × 'Nadorcott' (*C. reticulata*

Blanco) population (Aleza *et al.*, 2010b) were also analyzed as individual samples to perform genotype calling.

4x × 2x cross: Tetraploid 'Clementine' (*C. clementina* 4*x*) and diploid 'Pink' pummelo (*C. maxima* 2*x*) and 88 triploid hybrids segregating from this cross (Aleza *et al.*, 2012d) were also analyzed. Tetraploid 'Clementine' was obtained by treating buds of the diploid 'Clementine' with colchicines (Aleza *et al.*, 2009b); therefore, this genotype should be duplex (*aabb*) for all heterozygous loci.

SNP selection

SNPs for the analysis of signal-dosage correlation:

To validate the quantitative value that was obtained from the KASPar assay using pooled DNA, seven SNPs differentiating the two haploid lines (*C. maxima* and *C. clementina*) were selected from previous genotyping data obtained on the Illumina GoldenGate® platform (Ollitrault *et al*., 2012b). These SNPs were also used to test the accuracy of the technique in genotyping repetitions of the same sample over the 'Fortune' × 'Willowleaf' triploid population.

SNPs for triploid population analysis:

Three out of seven SNPs used for the previous analysis, and eight other SNP markers, were selected to test the capacity of the technique for differentiating between heterozygous genotypes within two triploid populations, one arising from a 2*x* × 2*x* cross, and the other from a 4*x* × 2*x* cross. These SNPs are heterozygous for 'Clementine' and homozygous or heterozygous with null allele for 'Nadorcott' and 'Pink'.

SNP genotyping

All samples were genotyped for the SNP markers using KASPar technology by KBioscience® (http://www.kbioscience.co.uk/). The KASPar™ Genotyping System is a competitive, allele-specific dual Förster Resonance Energy Transfer (FRET)-based assay for SNP genotyping. It uses two FRET cassettes where fluorometric dye [FAM (6-carboxy-fluorescein) or VIC®] is conjugated to primer but quenched via resonance energy transfer; ROX dye (6-carboxy-X-rhodamine, succinimidyl ester) is used to normalise the data. Sample DNA is amplified with a thermal cycler using allele-specific primers, leading to the separation of fluorometric dye and quencher when the FRET cassette primer is hybridised with DNA. Primers were designed by KBioscience®, based on the SNP locus flanking sequence (approximately 50 nt on each side of the SNP). Two 40-mer allele-specific oligonucleotides and one common 20-mer oligonucleotide were defined for each locus. Detailed information for all SNP markers can be found in Additional File 1. Additional details about this genotyping method can be found in Cuppen (2007).

Data analysis method

Normalised signals from each SNP allele (*x* and *y*) were provided by KBioscience® services, and two-dimensional plot representations were obtained using SNPViewer software (http://www.kbioscience.co.uk/software/SNP%20viewer%20intro.html). From the *x* and *y* normalised values, the theta angle [$\theta=\tan^{-1}(y/x)$; $0º \leq \theta \leq 90º$] and the relative *y* allele signal [$y'=y/(x+y)$; $0 \leq y' \leq 1$] of each sample were calculated. Further analyses were carried out that considered the *y'* parameter, as this parameter was found to provide better clustering and genotype calling of the samples.

Data from all haploid lines and DNA pools with different allele configurations (9:1, 5:1, 3:1, 2:1, 3:2, 1:1, 2:3, 1:2, 1:3, 1:5, and 1:9) were tested for correlations between doses and both the theta angle and *y'* values that were obtained. Cluster analysis (MacQueen, 1966) using the farthest-neighbour method with standardised squared euclidean distances and ANOVA were performed from the normalised allele signals (*x*, *y*) jointly and from the *y'* parameter data for each SNP.

Data from triploid and tetraploid simulated populations were also analyzed separately by cluster analysis and ANOVA. Replications of the same samples were used to test the precision of the technique by genotype calling.

All statistical data were analyzed using Statgraphics® Plus v5.1 software (Rockville, MD, USA).

Results

Marker design and data acquisition

Primers for the KASPar assay were successfully designed by KBioscience® for all 15 of the submitted SNP surrounding sequences. Data acquisition for *x* and *y* allele signals allowed successful allelic calling for 2535 out of 2563 marker/genotype combinations (98.91%). The validity of the genotyping results was verified by comparing the results for 24 diploid varieties with previous data obtained with an Illumina GoldenGate® array. Complete conformity was observed (data not shown).

Analysis of the correlation between relative allele signals and relative allele frequencies in the DNA pools

To confirm the value of the KASPar assay for producing semi-quantitative data, equimolar DNA extracts from two haploid lines (clementine and pummelo) were mixed at 13 different relative concentrations, and five replicates were analyzed for each of seven SNP markers. The correlations between relative allele signals and relative doses were analyzed.

An example of correlation analysis between relative allele dosage and signals is shown in Figure 1.1 for the CiC2840-01 SNP marker. From *x* and *y* signal values (Figure 1.1A), theta angle [$\theta = \tan^{-1}(y/x)$; Figure 1.1B] and the relative *y* allele signal [$y' = y/(x+y)$; Figure 1.1C] of each haploid line sample and DNA pool were calculated. High values of correlation coefficients between both parameters and relative allele dosage in the DNA pools were obtained for all analyzed SNP markers, with an average of 0.9796 and a standard deviation of 0.0094 for the *y'* parameter and an average of 0.9710 and a standard deviation of 0.0176 for angle theta. Correlation values obtained for the *y'* parameter were slightly superior to those obtained by angle theta for six of the seven SNP markers that were analyzed (Table 1.1).

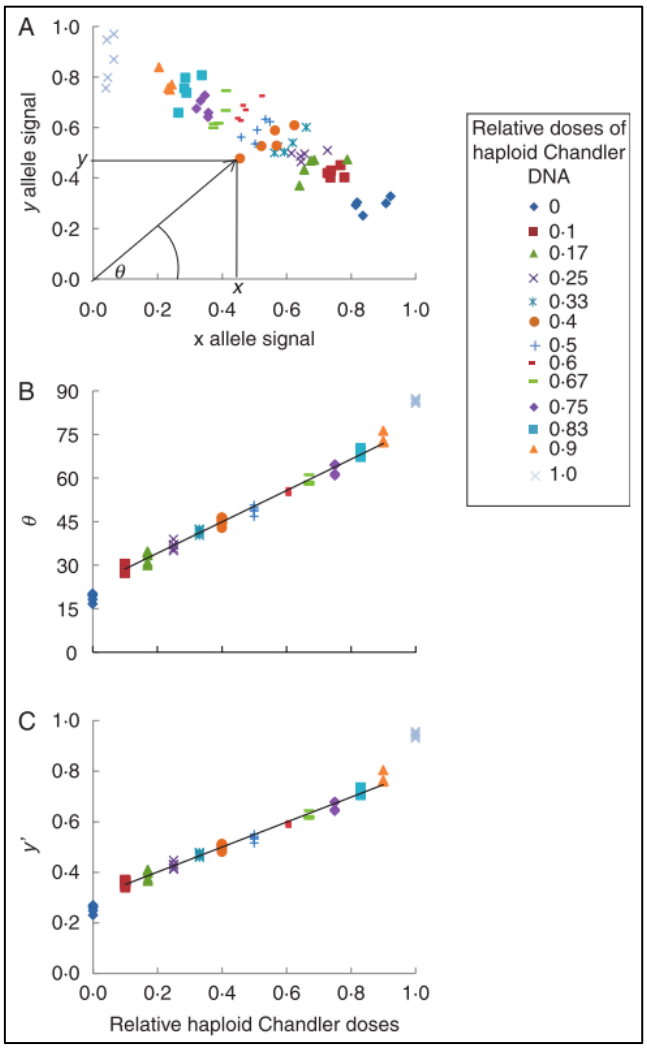

Figure 1.1. Correlation study of allele doses and x, y signals for the CiC2840-01 SNP marker. (A) Plot of normalized x, y allele signals. (B) Correlation between relative haploid Chandler doses and theta angle [$\theta = \tan^{-1}(y/x)$]. (C) Correlation between relative haploid Chandler doses and the y′ parameter [y′=y(x+y)].

Table 1.1. Correlation coefficients between relative allele dosage and allele signals from DNA pools at intermediate proportions for theta angle and *y'* parameter for the seven SNP markers analyzed.

SNP Marker	Correlation coefficient for angle θ	Correlation coefficient for *y'* parameter
CiC2840-01	0,9941	0.9919
CiC5089-06	0,9753	0.9779
CiC5785-01	0,9580	0.9747
DXS-M618	0,9881	0.9923
F3H-M309	0,9788	0.9803
FLS-M400	0,9535	0.9717
TRPA-M593	0,9492	0.9684
AVERAGE	0.9710	0.9796

For some markers (CiC5785-01, F3H-M309, FLS-M400, and TRPA-M593; Figure 1.2), the lineal regression established from the mixed sample did not fit with the signals from the pure sample. For all these markers, the relative signals corresponding to the haploid 'Chandler' allele in the DNA mixes appear to be lower than expected in relation to the relative DNA dosages. This can probably be attributed to PCR allele competition between the 'Clementine' and 'Chandler' allele in the DNA mixes.

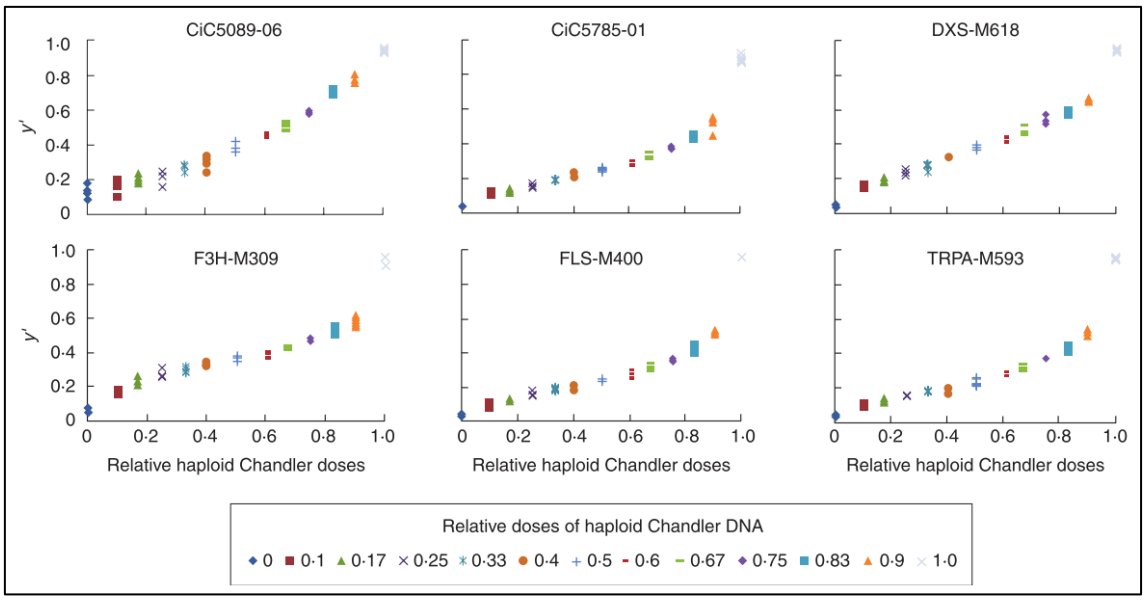

Figure 1.2. Correlation study between the relative haploid Chandler doses and the y′ parameter for six SNP markers.

However, in the DNA mixes, the correlations between allele signals and allele doses remained high for these markers (between 0.9684 and 0.9803) testifying for a very good lineal regression between relative signals of the two alleles and relative allele dosages.

These data indicate that the KASPar technique, using either the *y'* parameter or the theta angle, can be useful for a quantitative analysis of the relative allele frequency in a genotype or DNA pool. Because *y'* produced a slightly higher correlation coefficient, this parameter was employed in further analyses.

Cluster analysis and ANOVA for simulated triploid and tetraploid allele dosage

Separated cluster analyses and ANOVA from relative *y* allele signals (*y'* parameter) in triploid and tetraploid simulated populations were performed. With diallelic markers, for a triploid heterozygous genotype, there are only two allelic configurations to distinguish: *aab* and *abb* (duplex and simplex of *a*-allele). For a heterozygous tetraploid genotype, three allelic configurations may be differentiated: *aaab*, *aabb*, and *abbb* (triplex, duplex and simplex of *a*-allele). With higher ploidy levels, the number of possible allelic configurations becomes even larger (*n-1* configurations for *n* ploidy).

ANOVA analysis (Table 1.2) revealed a complete and correct classification of the average value of the different configurations that were simulated. An example of the *x* and *y* allele signals, the frequency histogram for the *y'* parameter and the LSD intervals for the mean from ANOVA for simulating triploid and tetraploid populations, is provided for the CiC2840-01 SNP marker in Figure 1.3.

Table 1.2. Homogeneous groups formed, and F-values from ANOVA analysis of SNPs from DNA pools simulating triploid and tetraploid populations showing the percentage of correctly classified replications by cluster analysis based on the *y'* parameter.

	Simulating 3n populations						Simulating 4n populations						
SNP Marker	**1:0**	**2:1**	**1:2**	**0:1**	**F value from ANOVA**	**% 3n correctly classified**	**1:0**	**3:1**	**1:1**	**1:3**	**0:1**	**F value from ANOVA**	**% 4n correctly classified**
CiC2840-01	a	b	c	d	2975.34	100%	a	b	c	d	e	1644.48	100%
CiC5089-06	a	b	c	d	1045.50	100%	a	b	c	d	e	673.68	96%
CiC5785-01	a	b	c	d	5110.86	100%	a	b	c	d	e	6084.71	100%
DXS-M618	a	b	c	d	2938.98	100%	a	b	c	d	e	3013.55	100%
F3H-M309	a	b	c	d	2063.52	100%	a	b	c	d	e	1445.66	100%
FLS-M400	a	b	c	d	10046.43	100%	a	b	c	d	e	10459.43	100%
TRPA-M593	a	b	c	d	11140.59	100%	a	b	c	d	e	6666.05	100%
AVERAGE						100%							99.43%

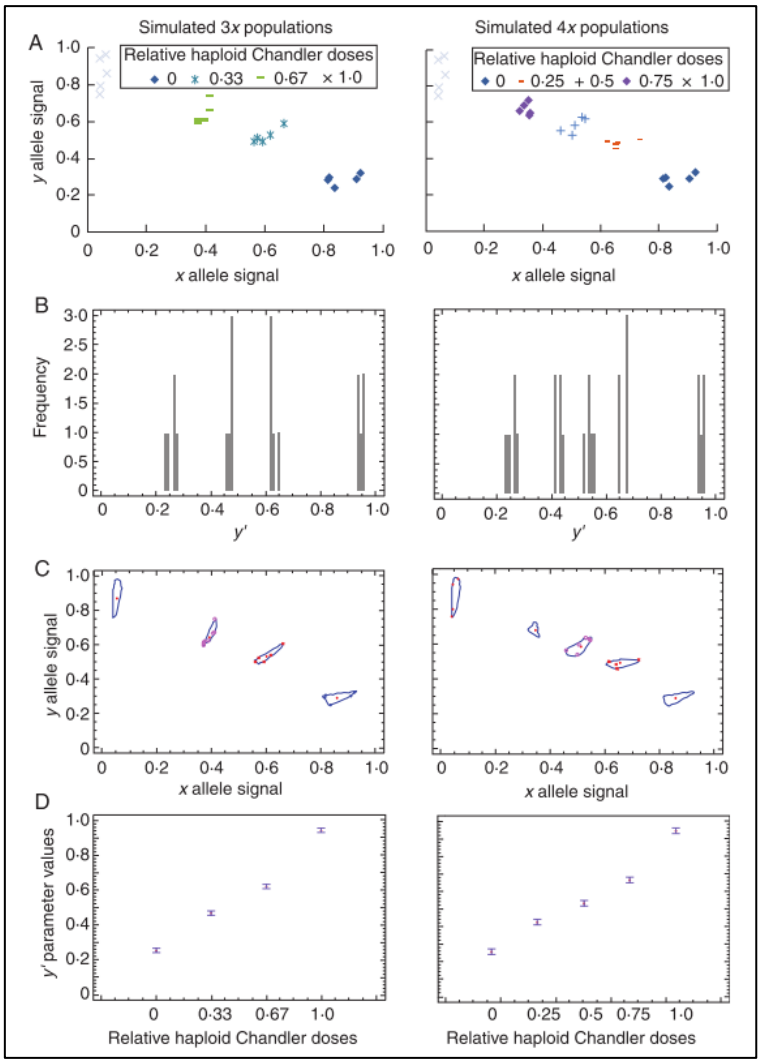

Figure 1.3. Study of simulating triploid and tetraploid populations for the CiC2840-01 SNP marker. (A) Plot of normalized x, y allele signals. (B) Frequency histogram for the y′ parameter. (C) Cluster analysis. (D) Least significant difference intervals for the mean obtained from ANOVA.

Moreover, all expected homogeneous groups were formed by cluster analysis using the farthest-neighbour method with standardised squared euclidean distances. All of the triploid sample replications and 99.43% of the tetraploid ones were correctly classified; only one replication for the CiC5089-06 SNP marker was classified into an incorrect cluster (Table 1.2).

Allelic configuration of triploid populations

Accuracy of genotype calling for duplicated triploid samples:

Thirty-nine triploid hybrids arising from 'Fortune' 2*n* gametes in the 'Fortune × Willowleaf' hybridization were analyzed for seven SNPs, with two technical replications. All SNPs had the following allelic configuration: 'Fortune': *ab* and 'Willowleaf' mandarin: *aa*.

Therefore, depending on the origin of the diploid gamete, three genotypic clusters were expected: *aaa*, *aab*, and *abb*. Samples with replications that were classified in the same cluster and thus genotyped with the same allelic conformation reached 97.44%. Errors in classification were observed in five of the seven SNP markers analyzed. Considering a replicate for the same DNA sample (with different allele calling between replicates) to be classified correctly, the average error rate for further routinely genotyping without replicates was estimated to be 1.28%.

To perform the genotype calling of triploid progeny, cluster analyses were performed according to the expected genotypes for each population and the parental-specific allelic configuration of each marker.

2x × 2x triploid progeny:

When crossing a heterozygous female parent (*ab*) with homozygous parents (*aa*), maternal heterozygosity restitution (HR) is reflected in the duplex (*aab*) triploid hybrids. Under the SDR restitution mechanism, HR is directly linked with the distance from the locus under consideration to the centromere and, therefore, the frequency of HR can be estimated from this distance, as proposed by Cuenca *et al.* (2011). To validate genotype calling for triploids resulting from a 2*x* × 2*x* cross, three SNP markers (CiC3440-07, CiC5785-01 and CiC6278-01) mapped in chromosome II were selected. Indeed, Cuenca *et al.* (2011) located the centromere position for the corresponding linkage group at 59.6 cM of the current clementine's reference genetic map (Ollitrault *et al.*, 2012a) using the $Cx(Co)^4$ partial interference model. The expected HR for the three considered markers (also mapped in the clementine map) was estimated using the same partial interference model.

Cluster analyses were performed from *y'* parameter values of each hybrid over 11 analyzed SNPs (including the three markers on chromosome II) within the 'Clementine × Nadorcott' population to carry out genotype calling. Figure 1.4A shows an example of cluster analysis for the CiC2840-01 SNP marker.

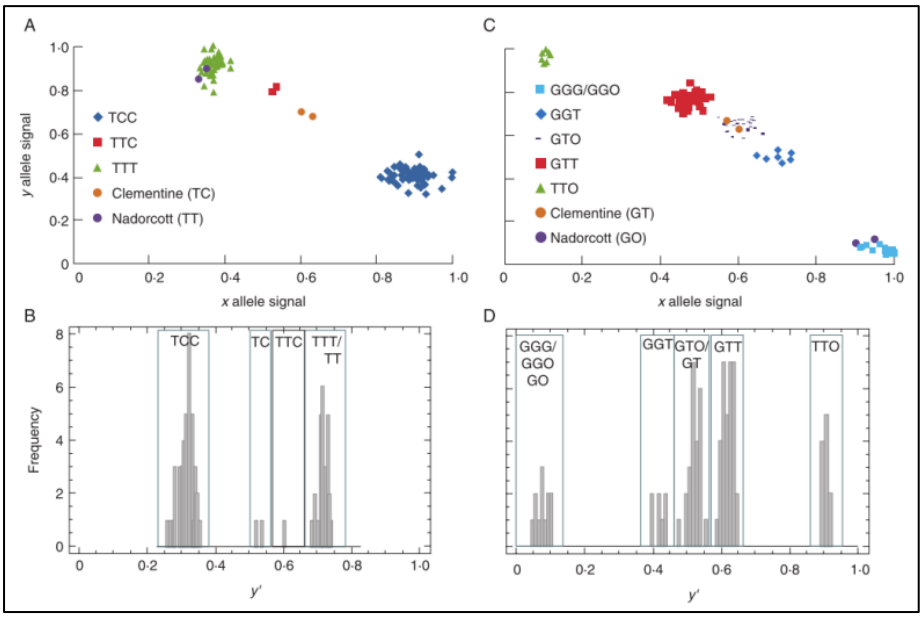

Figure 1.4. Plot of normalized *x*, *y* allele signals and histogram representing genotype calling from cluster analysis over 85 triploids from the 'Clementine × Nadorcott' cross. (A, B) CiC2840-01 (*ab* × *aa*); (C, D) CiC1749-05 (*ab* × *a0*).

The cluster analysis allowed the detection of null alleles in the male parent for two markers. Indeed, if the supposed homozygous parent in fact had heterozygosity for a null allele, five clusters should be obtained (*ab*x*a0*: *aab*, *ab0*, *aaa*/*aa0*, *abb*, and *bb0*), where one cluster contains both *ab0* triploids and *ab* diploid genotypes. Such a cluster configuration was observed for the CiC0610-01 and CiC1749-05 SNP markers (Figure 1.4C,D).

On average, over all of the markers, 99.37% of the samples were assigned to a cluster and, therefore, could be accurately called.

For the three markers of LGII, the observed HR values were not significantly different than those estimated from respective markers and centromere locations in the clementine map (Table 1.3). This provides additional validation of the accuracy of polyploid genotype calling using the method presented in this study.

Table 1.3. Results of 3-SNP genotype calling for the 'Clementine × Nadorcott' triploid population from 2*n* gametes for markers heterozygous for 'Clementine' and homozygous for 'Nadorcott', including the conformity (χ^2 test) of the observed %HR with the theoretical one calculated from the distance from each locus to the centromere on chromosome II.

SNP marker	Map position (cM)	NI	aaa	aab	abb	%HR (aab) observed	%HR estimated from map position (centromere at 59.6 cM)	X^2 (p value)
CiC3440-07	67.22	86	37	10	39	11.63%	13.59%	0.282; NS (p=0.5954)
CiC5785-01	44.73	84	32	17	35	20.24 %	28.23%	2.648; NS (p=0.1037)
CiC6278-01	57.01	86	46	0	40	0.00%	4.21%	3.780; NS (p=0.0519)

NI: number of individuals genotyped; *aaa*, *aab* and *abb*: number of individuals of each genotype; NS: non significantly different at alpha=0.05)

For the eight remaining markers, the observed HR allowed us to estimate the relative distances of these markers from the centromere (Table 1.4), revealing markers with centromeric locations (less than 5 cM distance from centromere: CiC1380-05, CiC2840-01, and CiC4581-01).

Table 1.4. Results of 8-SNP genotype calling for the 'Clementine × Nadorcott' triploid population from 2*n* gametes for markers heterozygous for 'Clementine' and homozygous or heterozygous with null allele for 'Nadorcott', showing the estimated marker-centromere distance.

SNP marker	NI	aaa	aab	abb	ab0	bb0	% HR (aab+ab0) observed	Estimated distance to centromere (cM)
CiC0610-01	86	22	22	14	22	6	51.16%	27.42
CiC0868-01	86	39	14	33	-	-	16.8%	9.23
CiC1380-05	86	43	1	42	-	-	1.16%	0.75
CiC1749-05	85	11	33	7	24	10	67.06%	40.61
CiC1757-02	86	42	12	32	-	-	13.95%	7.80
CiC2840-01	85	46	2	37	-	-	2.35%	1.49
CiC4581-01	86	37	6	43	-	-	6.98%	4.15
CiC5089-06	84	27	40	17	-	-	47.62 %	25.23

NI: number of individuals genotyped; aaa, aab, abb, ab0 and bb0: number of individuals of each genotype.

4x × 2x triploid progeny:

Triploid genotyping was also performed for progeny arising from a cross between doubled diploid 'Clementine' (Aleza *et al.*, 2009a) and 'Pink' pummelo. When crossing a duplex tetraploid parent (*aabb*) with a homozygous diploid parent (*aa*), three clusters can be expected (triplex-*aaa*, duplex-*aab*, and simplex-*abb*), where maternal heterozygosity restitution is reflected into the duplex (*aab*) triploid hybrids. On average, over all of the markers, 97.08% of the samples were assigned to a cluster and, therefore, could be accurately called.

The genetic structure of the triploid progeny arising from 'Clementine 4*x*' × 'Pink' is shown in Table 1.5, which indicates the percentage of heterozygosity restitution. For the other loci, HR values varied, ranging from 52.94% for CiC4581-01 to 69.41% for CiC5089-06. The average HR value over all loci was 59.64%.

Table 1.5. Results of 7-SNP genotype calling for the 'Clementine 4*x* × Pink' triploid population, indicating the heterozygosity restitution in Clementine 4*x* (%HR) at each locus.

SNP marker	NI	aaa	aab	abb	%HR (aab) observed
CiC3440-07	88	17	50	21	56.82%
CiC0868-01	87	11	58	18	66.67%
CiC1380-05	87	19	53	15	60.92%
CiC4581-01	85	19	45	21	52.94%
CiC5089-06	85	15	59	11	69.41%
CiC5785-01	80	22	43	15	53.75%
CiC6278-01	86	22	49	15	56.98%

NI: number of individuals genotyped; *aaa*, *aab* and *abb*: number of individuals of each genotype.

Discussion

The KASPar method is a powerful technique for assigning SNP allelic configurations in polyploid progeny

Several techniques have been used to estimate allele dosage in polyploids, such as the MAC-PR method (Esselink *et al.*, 2004) for SSR markers, and techniques for SNP genotyping, including allele-specific primer extension (Kwok, 2001), temperature-switch PCR (Tabone *et al.*, 2009), array methodologies (Ishikawa *et al.*, 2005), and targeted pyrosequencing™ (Ahmadian *et al.*, 2000). Our study demonstrates that the KASPar technique (Cuppen, 2007) is an alternative method to infer SNP allelic configurations in polyploid plants that offers a wider spectrum of genotyping possibilities. The KASPar method is simple and cost effective compared to other SNP genotyping assays and is well adapted to low- to medium-throughput genotyping projects. In addition to the markers published herein, 51 KASPar markers were successfully developed to analyze triploid and tetraploid citrus populations (Aleza *et al.*, 2012a, 2012b; Cuenca *et al.*, 2012). KASPar markers were also successfully developed (41 over 42 tested) and transferred in the true citrus group (*Citrus, Fortunella, Poncirus, Microcitrus* and E*remocitrus* genera) from SNPs mining by sequencing within a *Citrus* collection (García-Lor *et al.*, 2013a). When SNPs are mined in large discovery panel, this offers the opportunity to selects markers without additional variation in the flanking DNA sequence used as template for the competitive PCR of the KASPar assay and therefore to have a high success in marker development. KASPar markers were successfully developed in a large range of plant (Cortes *et al.*, 2011; Rosso *et al.*, 2011; Byers *et al.*, 2012) and animal (Nijman *et al.*, 2008; Murad *et al.*, 2009; Luciano *et al.*, 2010) species demonstrating its universal applicability.

The SNP genotyping and data analysis method presented in this study is simple and effective for genotyping triploid and tetraploid progeny and can be also used in the quantitative analysis of allele-specific expression. Allele signals (x, y) obtained from KBioscience® can easily be transformed into y' [$y'=y/(x+y)$; $0 \leq y' \leq 1$], which is a very useful parameter to cluster analyzed samples. Theta angles [$\theta=\tan^{-1}(y/x)$; $0º \leq \theta \leq 90º$] can also be used to analyse data, but the y' parameter offers better clustering results. Quantitative analyses for correlation of the allele signals, and the allele doses and sample clustering carried out in this work, were powerful techniques for assigning allelic configurations in simulated triploid and tetraploid citrus genotypes for all SNP markers that were analyzed (100% of the triploids were correctly classified triploids as well as 99.43% of the tetraploids). The analysis of concrete triploid hybrids with technical replications confirmed the high degree of accuracy of the technique (error < 1.5%). This SNP genotyping and data analysis method allowed us to distinguish among very close allele ratios, and it can also be efficiently employed for the analysis of higher ploidy levels. Moreover, the segregations observed with this technique have allowed us to identify heterozygous null alleles in one parent for some of the markers. Diploid progeny genotyping confirmed these conclusions for null alleles in 'Pink' pummelo (Ollitrault *et al.*, 2012a).

PCR drift can affect allelic configuration inference in natural polyploid germplasm

Interpretation of relative allele dosage for markers based on relative polymerase chain reaction (PCR) product intensities has been reported for various plants (Buteler *et al.*, 1999; Julier *et al.*, 2003; Landergott *et al.*, 2006; Martins *et al.*, 2009) and animals (McQuown *et al.*, 2002). The limits of such direct allele doses evaluation are associated with PCR selection caused by differential primer affinity and PCR drift resulting from random events during early cycles of PCR (Wagner *et al.*, 1994).

In this study, such PCR drift has been observed for some markers, displaying incongruence between the lineal regressions established from the mixed DNA pools and the pure sample. However, the correlations between allele signals and allele doses in the DNA pools remained high for these markers. Therefore, as lineal regression appears to offer a good approximation of the doses/relative signal relationship, a control with two dosage points should be sufficient to establish a function that correlates both parameters.

Heterozygous diploid genotypes are suitable for determining the 1:1 ratios that are used as a baseline for calculations of allele quantification in the other heterozygous genotype. In the analysis of citrus triploid progeny, the location of different clusters relative to the heterozygous diploid parent allowed us to assign the alternate theoretical triploid heterozygous allelic configuration.

The situation is much more complicated when analysing polyploid germplasm of unknown origin. Indeed, the variability in the flanking regions of the SNPs that were studied (where the primers were defined) should result in different levels of relative PCR competition and, therefore, should avoid proper allele dose identification from relative *x*/*y* signals. This is inherent to all PCR genotyping methods. Perhaps, as suggested by (Landergott *et al.*, 2006) for the MAC-PR method, the KASPar assay may be very useful for determining the allelic configuration within crossing families, but it would not be generally applicable for estimating allelic dosage in polyploid germplasm without previous verification of the stability of relative allele amplification. An approach to limit the PCR drift associated with variation in the flanking area of the studied SNP should be to select SNPs flanked by conserved sequences. Such information is available in SNP mining studies where large discovery panels are used while there is generally no information of flanking sequence of microsatellite markers. This should be an important advantage of using SNPs rather than microsatellite markers for assignment of allelic configuration in polyploids.

Potential of KASPar for semi-quantitative estimation of allele-specific expression analysis or allelic frequency estimation in DNA extracted from pools

Many genetic variants resulting in phenotypic differences are mediated through changes in gene expression. Variation in gene expression can be due to polymorphisms either at the gene locus (cis) or in other genes that influence gene expression (trans) or cis/trans interactions (Rockman and Kruglyak, 2006). Allele-specific expression (ASE) studies have introduced a creative method to uncover the respective contributions of cis- and trans-regulatory variation (Ronald *et al.*, 2005; Main *et al.*, 2009). Allelic imbalance in non-imprinted genes has been shown to be common in humans, maize, and Arabidopsis (Lo *et al.*, 2003; Guo *et al.*, 2004; Zhang *et al.*, 2009; Zhang and Borevitz, 2009). Moreover, ASE analysis should enable the integration of potentially differential allelic functionality in association models between gene expression and phenotype. Therefore, gene expression analysis is a critical step for better understanding genotype/phenotype relationships.

Analysis of allele-specific expression in relation to genomic structure requires the assessment of DNA and RNA allele dosage. It can be done by different methods: Northern (Guo *et al.*, 1996), RNA-FISH (Herzing *et al.*, 2002), SNP-specific array-based (Bjornsson *et al.*, 2008), Solexa (Main *et al.*, 2009) or RNA-seq (Rozowsky *et al.*, 2011).

Furthermore, the estimation of allelic frequencies on pooled DNA is of great interest both in ecological studies of plants (Ritland, 2002), animals (Shaw *et al.*, 1998; Coop *et al.*, 2010; Grant, 2010) or micro-organisms (Brauer *et al.*, 2006; Wenger *et al.*, 2010), and in bulk segregant analysis to locate genes involved in phenotypic variation (Quarrie *et al.*, 1999; Tabor *et al.*, 2000; Yang *et al.*, 2007).

The high correlation coefficient values between relative allele dosage and SNP allele signals obtained with the KASPar technique, and the ability of this technique to distinguish between close relative allele dosages at the DNA level, has been demonstrated in this study. Moreover were able to detect a 0.1 allele frequency within DNA pools. This technique is a promising method for performing semi-quantitative analysis of relative allele-specific expression by analysing cDNA compared to genomic DNA, to complement global gene expression studies performed by real-time PCR. The KASPar technique may also be useful for allele frequency estimation in populations from DNA pools as mentioned before. For such studies, it should be interesting to extend the range of relative allele dosages to estimate the lowest differences distinguishable with this technique.

Application for citrus genetics and breeding

Triploid citrus breeding is one of the most efficient techniques for the production of seedless mandarins (Ollitrault *et al.*, 2008; Aleza *et al.*, 2010b, 2012c, 2012d), and tetraploid rootstocks are promising tools that enable plants to adapt to various abiotic stresses (Saleh *et al.*, 2008; Dambier *et al.*, 2011). Triploid populations in citrus can arise from 2*x* × 2*x* crosses or from interploid crosses. Discriminating between different types of heterozygotes within triploid progeny is especially useful for population genetic structure studies and marker/trait association analysis.

Knowing the allelic configuration in triploid and tetraploid progeny is also necessary to identify the mechanism of 2*n* gamete formation. The maternal heterozygosity restitution values of under 50% obtained in this study, which were estimated from a 'Clementine × Nadorcott' progeny for nine markers (CiC0868-01, CiC1380-05, CiC1757-02, CiC2840-01, CiC3440-07, CiC4581-01, CiC5089-06, CiC5785-01, and CiC6278-01), confirm the conclusion of (Luro *et al.*, 2004), that the 2*n* gamete in clementine arose from SDR, as in the 'Fortune' mandarin (Cuenca *et al.*, 2011), while (Chen *et al.*, 2008a) proposed FDR for sweet orange. Moreover, this study allowed us to identify several centromeric markers that should be very useful for further analyses of the origin of 2*n* gametes in different cultivars and genotypes, as it was done for potatoes (Douches and Quiros, 1988; Werner *et al.*, 1992).

Most tetraploid citrus germplasm arose from chromosome duplication of nucellar cells (Aleza *et al.*, 2011) or were obtained by bud chemical treatment (Aleza *et al.*, 2009b) of diploid genotypes. These tetraploids are, therefore, doubled diploids with the same *aabb* genomic structure at each heterozygous locus (*ab*) of the parental diploid line. For such tetraploids, the parental restitution (PR) of the heterozygosity to the diploid gamete depends on preferential pairing between chromosomes. In the case of total preferential pairing (disomic segregation), parental heterozygosity is transferred to all gametes (PR=100%). In the case of total random pairing (tetrasomic segregation), the PR ranged from 55% to 66%, depending on the double reduction frequency (Marsden *et al.*, 1987). In this study, the PR results for the tetraploid (doubled diploid) Clementine ranged from 52.94% for the CiC4581-01 marker to 69.41% for the CiC5089-06 marker, which is in agreement with the expected PR values under tetrasomic segregation (Kamiri *et al.*, 2011).

In the case of triploid and tetraploids obtained by somatic hybridization (Dambier *et al.*, 2011; Grosser and Gmitter, 2011), the assignment of allelic configuration will be useful for revealing genome regions acquired from each parent, as well as potential chromosome fragment elimination or duplication.

Conclusions

This work demonstrates that the KASPar SNP genotyping technique, combined with the cluster analysis method we proposed, enables the efficient assignment of heterozygous allelic configuration within polyploid populations. This method is accurate, simple, and cost effective. It has been successfully applied to two citrus triploid populations arising from 2*n* gametes and interploid crosses. Moreover, correlation studies, cluster analysis and ANOVA support the usefulness of this method for performing relative quantitative studies, such as relative allele-specific expression analysis or, eventually, bulk segregant analysis.

Supplementary information

Table 1.S1. SNP information and genotyping of the parents used in this study
Description: Detailed information of the 15 SNPs used in this work, including the GeneBank accessions, genomic sequences surrounding each SNP and the genotyping of the progeny parents.

Table 1.S1. SNP information and genotyping of the parents used in this study.

SNP marker	GENEBANK ACCESION	Genomic sequence surrounding the SNP	Haploid Chandler	Haploid Clementine	Fortune	Willowleaf	Clementine 2x	Clementine 4x	Nadorcott	Pink
CiC0610-01	ET093305	CAACTTAAAATCCCACTTTACTCTTTCCACATGGGCCGTCATGA TGGTGTTTTTGACTGC[A/G]CTTGTTACTTCCACTAAACCCTTTT AAAGATTGCAGATCTCTTTGTTTGAGTTAAAATAT	-	-	-	-	AG	-	G0	-
CiC0868-01	ET095107	ACAGTTTCTGGGATCCAYATTTGATACCAACTTCATTTCAAGGG TAACTTTCAGGAACTA[A/G]GATACAGTATGTGCCTTTAAATTCT ATTAATCACGTGGGATCATTACGATGTCTTGCTTT	-	-	-	-	CT	CCTT	CC	TT
CiC1380-05	ET072553	TTGCTCGAGAAGTCAAAACTTTACAGGAATGCATTACTGGCCT ATCTGGAGAGAAAGCTC[A/G]AGTACAGAAGGACTGCAATGAG TTGAGGGCCATAAACGAGGATTTGTTATCCCGACAAAA	-	-	-	-	CT	CCTT	CC	00
CiC1749-05	ET097636	ACATTACTCTGTCTACCCCGCAGATACTAATAAGGTATTTATGA TCATGTTTTTGCTTCT[A/C]ATATTGCTGACTTCTGCTGTTTTTG TTTGGAATTGATGTTCTGTGATTGATGTGGAGTCA	-	-	-	-	GT	-	G0	-
CiC1757-02	ET097717	GTTTTCAGCTTTGAAAAAATCAGATACTTCCTCTCGAGTCAAGG TAGGTTCTTATCACTT[A/C]CCATCTTCTATATCTGATGCATCTC TGTTAAAGAAGCGATCTGAACTTGAAATTTCAGAT	-	-	-	-	AC	-	CC	-
CiC2840-01	ET103429	TTTATTTATTTACTTATGTATTTGTTTTGTAAATTTTTAATATTTC AGATCCGCACAYAT[A/G]GCGAAAAGAGCGTTGTACGGTTCGT TGAGGCTCACGAGTTCATTGACTGTGGACATATCC	C	T	CT	CT	CT	-	TT	-
CiC3440-07	ET077539	TTTTTATTTTTATTTTTATTCTTATTTTTGCCTCTTGAATCCATCG GTATTCCTGGCTAT[A/G]AAAGAGAAGTTCAACTTCTAGCCGAA TTAACTATATTAACTGCCATTCCAAATATGTTGA	-	-	-	-	AG	AAGG	GG	G0
CiC4581-01	ET109034	CCATTCAAATAAAAATCCTAAACATTCAATCATTAACAATCACA GAGATGAAATTCCCAG[A/C]GGTTCAATAAGGCTGGTGGTTCT CAAAGCACTGAACTGCATAATATAAAATAGGAAAAAG	-	-	-	-	AC	AACC	AA	AA

Table 1.S1-cont. SNP information and genotypes of the parents used in this study.

SNP marker	GENEBANK ACCESION	Genomic sequence surrounding the SNP	Haploid Chandler	Haploid Clementine	Fortune	Willowleaf	Clementine 2x	Clementine 4x	Nadorcott	Pink
CiC5089-06	ET111533	TCAGAACAACATAAAAATACCACTTTTKTTTTTTAGAAGTGGAGGAAATTTCAGTGAAGA[A/G]CAGGAAGGCATCACATCTGATTCAAACAGTGAGATTGATGGGAAAGAAAATTCTAACCAG	C	T	CT	CT	CT	CCTT	CC	CC
CiC5785-01	ET082673	TCAACCACCTTGATAAATGAACAAAAGGAGGCTTAAACTTTGACTTCAATTTTAGCCATT[A/T]GTTATGTTCAATCTAATGATTATAATGATTGCAATTTAAGACGCTTTGATTGAATTACTA	T	A	AT	AT	AT	AATT	TT	TT
CiC6278-01	ET085551	CCTATTATGTTTATTTAAATTGTTGGTGGACGATAAAATTTATTGGCCAATTCAGAATTC[A/C]TAGTCATAATGAAATTCTGTACACATAAAAAGCCAAATGGGTGACTCAGTTCTTTGTGTA	-	-	-	-	AC	AACC	AA	00
DXS M618	DN959423	GGTGCAGGACATAGTTCCACAAGCATCTCTGCTGGTCTTGGTATGTACTTC[G/A]CTCTCTTAATATTTTCCTTTCATCAATCTAGAGAAATTGTAGGATGCAGAATAC	A	G	AG	AG	-	-	-	-
F3H-M309	JX630066	CTAAGCCGTCGAGTTTTTGTGACCAAAGGGACAGAATCTAATGAGTTTAAGGA[T/C]ATGGTGGTAGAGCTCATGACGTCAGCTGGATTTTTCAACATTGGTGATTTTATACCC	C	T	CT	CT	-	-	-	-
FLS-M400	AB011796	GGGTTGATCATCTCTTCCACAGGGTTTGGCCTCCGTCTTCTATCAACTACCGCTT[T/C]TGGCCCAACAACCCTCCTTCTTACCGGTGAATGTT	C	T	CT	CT	-	-	-	-
TRPA-M593	EF028327	GGCACGACAGCATTCGTGAGCATGTGGATGCATAACGTGGCAGCAGCAGTGAT[G/C]ATGATGCCAGTGGCCACTGGGATCTTACAGAACTTGCCAGAGGTTC	C	G	CG	GG	-	-	-	-

CHAPTER 2

Multilocus half-tetrad analysis and centromere mapping in citrus: evidence of SDR mechanism for 2*n* megagametophyte production and partial chiasma interference in mandarin cv 'Fortune'

José Cuenca, Yann Froelicher, Pablo Aleza, José Juárez, Luis Navarro and Patrick Ollitrault

Heredity (2011) 107: 462 – 470. doi:10.1038/hdy.2011.33

Abstract

The genetic structure of 2*n* gametes and, particularly, the parental heterozygosity restitution at each locus depends on the meiotic process by which they originated, with first division restitution and second-division restitution (SDR) being the two major mechanisms. The origin of 2*n* gametes in citrus is still controversial, although sexual polyploidization is widely used for triploid seedless cultivar development. In this study, we report the analysis of 2*n* gametes of mandarin cv 'Fortune' by genotyping 171 triploid hybrids with 35 simple sequence repeat markers. The microsatellite DNA allele counting-peak ratios method for allele dosage evaluation proved highly efficient in segregating triploid progenies and allowed half-tetrad analysis (HTA) by inferring the 2*n* gamete allelic configuration. All 2*n* gametes arose from the female genitor. The observed maternal heterozygosity restitution varied between 10 and 82%, depending on the locus, thus SDR appears to be the mechanism underlying 2*n* gamete production in mandarin cv 'Fortune'. A new method to locate the centromere, based on the best fit between observed heterozygosity restitution within a linkage group and theoretical functions under either partial or no chiasmata interference hypotheses was successfully applied to linkage group II. The maximum value of heterozygosity restitution and the pattern of restitution along this linkage group would suggest there is partial chiasma interference. The implications of such a restitution mechanism for citrus breeding are discussed.

Introduction

Sexual polyploidization has been widely exploited in several plant-breeding programmes (Ramanna and Jacobsen, 2003). Diploidy is the general rule in *Citrus* and related genera, with a basic chromosome number *x*=9. However, sexual polyploidization is currently a central approach used in triploid citrus-breeding programs, aiming to develop new seedless 'mandarin type' cultivars (Ollitrault *et al.*, 2008); very large triploid progenies have been obtained from 2*x* × 2*x* crosses and several cultivars patented (Aleza *et al.*, 2010b). The recovery of triploid citrus hybrids arising from 2*n* megagametophytes produced by diploid plants was described in the 1970s (Esen and Soost, 1971, 1973; Aleza *et al.*, 2010b). Cytogenetic studies (Esen and Soost, 1971) showed that triploid embryos are associated with pentaploid endosperm, indicating that triploid hybrids result from the fertilization of unreduced ovules by normal haploid pollen. According to the genotype, the frequency of duplication in the female gametes can range from below 1% to over 20%. Esen *et al.* (1979) proposed that 2*n* eggs result from the abortion of the second meiotic division in the megaspore (SDR) in citrus. This hypothesis has been corroborated for clementine (*Citrus clementina* Hort. ex Tan.) by molecular marker analysis (Luro *et al.*, 2004). However, Chen *et al.* (2008a) proposed that 2*n* eggs of sweet orange (*C. sinensis* (L.) Osb.) resulted from first meiotic division restitution (FDR). The genetic configuration of 2*n* gametes depends on the mechanism of their formation (Figure 2.1), and the rate of parental heterozygosity restitution is directly linked with the rate of effective chiasma between the centromere and the considered locus (Mendiburu and Peloquin, 1976; Park *et al.*, 2007). It is, total and null until the first chiasma for FDR and SDR, respectively. It is, thus, essential to gain a better understanding of the mechanisms underlying 2*n* gamete formation to optimise sexual polyploidy breeding schemes and to carry out trait-association studies of breeding populations.

The objective of the present work was to shed light on the mechanism underlying 2*n* gamete formation in 'Fortune' mandarin (*C. clementina* × *C. tangerina* Hort. ex Tan.) by Simple Sequence Repeat (SSR) marker analysis. 'Fortune' mandarin is widely used in triploid breeding because of its fruit qualities, late maturing period and relatively high percentage of 2*n* eggs. This diploid cultivar is highly fertile, producing an average of 18.5 seeds per fruit, including 6.5% of seeds arising from 2*n* gametes (our unpublished data). Without the previous knowledge of centromere position, and to avoid the risk of misinterpreting data due to an insufficient or biased set of markers, we selected 35 SSR markers according to their position on the clementine map (Ollitrault *et al.*, 2011). A general approach is proposed to estimate the centromere position by best-fit value between observed data and theoretical functions of heterozygosity restitution for no interference and partial interference models. For this purpose, we used the functions developed by Zhao and Speed (1998b) for ordered tetrads, based on the random spindle–centromere attachment hypothesis (Griffiths *et al.*, 1996), and extended by the same authors to half-tetrad analysis (HTA; (Zhao and Speed, 1998a). The results obtained were compared with the method proposed by Tavoletti *et al.* (1996) using the multilocus structure of half-tetrads.

Finally, we discuss the implications of the 2*n* gamete restitution mechanisms for citrus triploid breeding.

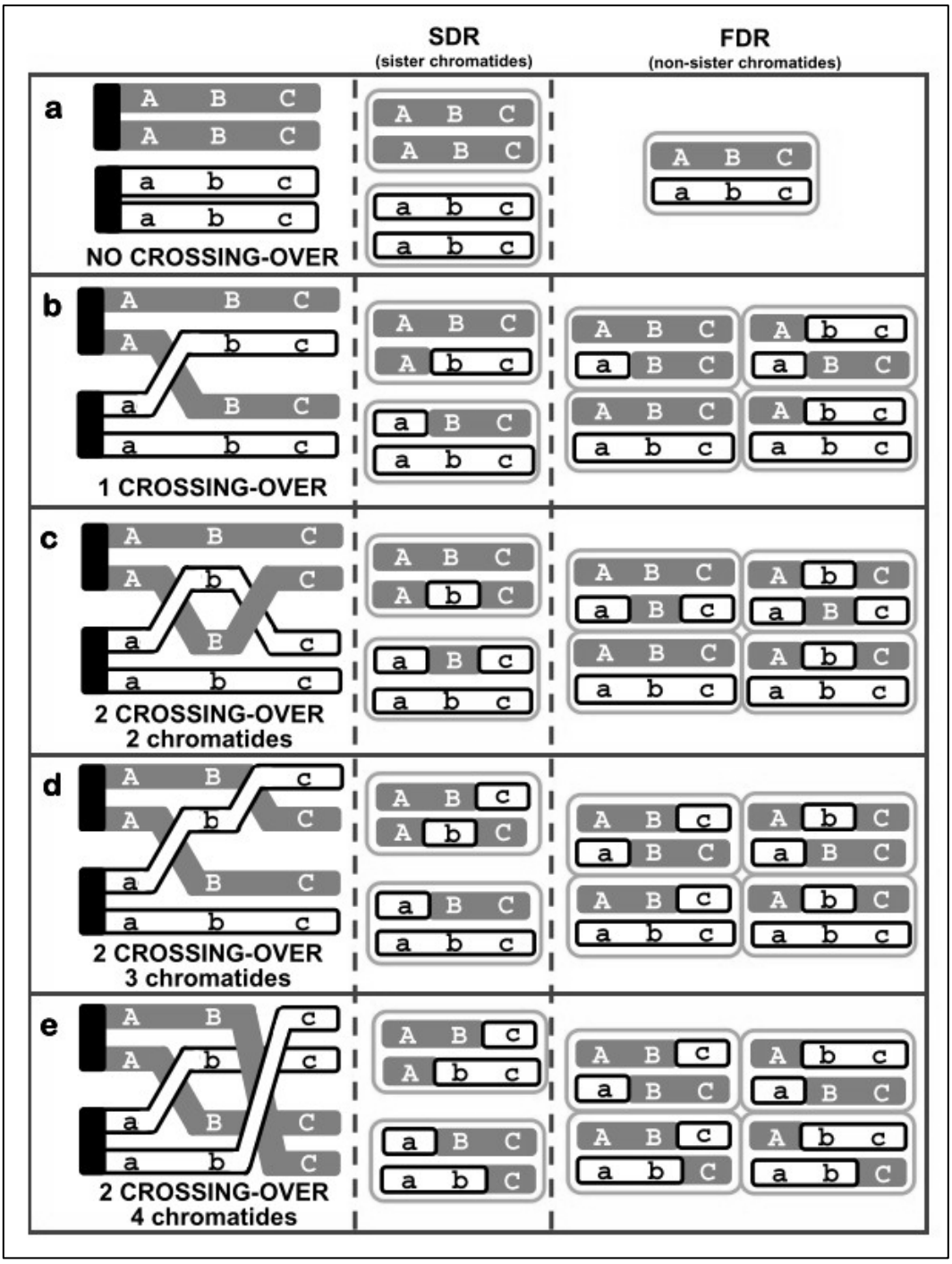

Figure 2.1. SDR and FDR 2*n* gametes resulting from: (a) no crossover;(b) one crossover; (c) two crossovers involving two chromatids; (d) two crossovers involving three chromatids and (e) two crossovers involving four chromatids.

Materials and methods

Plant material

The mechanism underlying 2*n* gamete formation in 'Fortune' mandarin was investigated in 171 triploid hybrids, derived from four crosses between 'Fortune' mandarin (*C. clementina* Hort. ex Tan. × *C. tangerina* Hort. ex Tan.) as female diploid genitor and 'Ellendale' (*C. reticulata* Blanco × *C. sinensis* (L.) Osb.), 'Common Mandarin' (*C. deliciosa* Ten.), 'Minneola' (*C. paradisi* Macf × *C. tangerina*) or 'Murcott' (*C. reticulata* × *C. sinensis*) as male diploid genitors. Parental accessions and hybrids were grown at the 'Instituto Valenciano de Investigaciones Agrarias' orchards in Moncada, Valencia, Spain. Practical details on how the triploid populations were established from diploid × diploid crosses by embryo rescue and triploid selection by flow cytometry, can be found in Aleza *et al.* (2010b). Genomic DNA of triploid hybrids and their parents was isolated using the Plant DNAeasy kit from Qiagen Inc. (Valencia, CA, USA), following the manufacturer's protocol.

Triploid progeny genotyping

SSR markers that proved heterozygous for 'Fortune' were selected to genotype each triploid progeny, depending on the polymorphism existing between 'Fortune' and the male genitor. Thirty-five microsatellite loci were used to analyze the triploid progenies: *CAC15*, *TAA41* (Kijas *et al.*, 1997), *CX6F03*, *CX6F23* (Chen *et al.*, 2008b), *mest121* (Luro *et al.*, 2008), *mest56*, *mest192*, *mest123* (Aleza *et al.*, 2011), *mest104*, *mest110*, *mest247*, *mest488* (François Luro, personal communication; mail to luro@corse.inra.fr for further information), *mCrCIR07F11* (Kamiri *et al.*, 2011), *Ci01C07*, *Ci02B07*, *Ci08C05*, *mCrCIR01E02*, *mCrCIR01F04a*, *mCrCIR06A12*, *mCrCIR06B05*, *mCrCIR06B07*, *mCrCIR07E12* (Froelicher *et al.*, 2007) and thirteen new designed primers (*mCrCIR01C06*, *mCrCIR02A09*, *mCrCIR02D09*, *mCrCIR02F12*, *mCrCIR02G01*, *mCrCIR02G02*, *mCrCIR03B07*, *mCrCIR03C08*, *mCrCIR03G05*, *mCrCIR04H06*, *mCrCIR05A05*, *mCrCIR07D05*, *mCrCIR07D06*; Table 2.1). The polymerase chain reactions (PCRs) were performed with wellRED oligonucleotides (Sigma®) with the following protocol: Eppendorf® Mastercycler epgradient S; reaction volume 15 µl: 0.8 U Taq polymerase (N.E.E.D.®), reaction buffer -750 mM Tris HCl (pH 9), 50 mM KCl, 200 mM $(NH_4)_2SO_4$, 0.001% BSA-, 0.1 mM of each dNTP, 5 mM $MgCl_2$, 3 µM of each primer, 30 ng DNA; PCR program: 94ºC 5 min; 40 cycles of 30 sec at 94ºC, 1 min at 50-55ºC and 30 sec at 72ºC; final elongation 10 min at 72ºC). Separation was carried out by Capillary Gel Electrophoresis (CEQ™ 8000 Genetic Analysis System; Beckman Coulter Inc.). Data collection and analysis were carried out with GenomeLab™ GeXP version 10.0 software.

Table 2.1. New primers designed to amplify the markers used in this study

Marker name	EMBL accession	Sequence Forward 5'-3'	Sequence Reverse 5'-3'	Annealing Temperature (ºC)	Observed alleles in Fortune
mCrCIR01C06	AJ567393	GGACCACAACAAAGACAG	TGGAGACACAAAGAAGAA	50	133-165
mCrCIR02A09	FR677568	ACAGAAGGTAGTATTTTAGGG	TTGTTTGGATGGGAAG	50	160-162
mCrCIR02D09	FR677569	AATGATGAGGGTAAAGATG	ACCCATCACAAAACAGA	55	231-239
mCrCIR02F12	FR677570	GGCCATTTCTCTGATG	TAACTGAGGGATTGGTTT	55	121-123
mCrCIR02G01	FR677571	ATACCAAAACCCCAAAG	CTTTGACCCAAGCAAG	55	291-296
mCrCIR02G02	FR677572	CAATAAGAAAACGCAGG	TGGTAGAGAAACAGAGGTG	55	112-122
mCrCIR03B07	FR677573	CACCTTTCCCTTCCA	TGAGGGACTAAACAGCA	55	264-278
mCrCIR03C08	FR677576	CAGAGACAGCCAAGAGA	GCTTCTTACATTCCTCAAA	55	210-223
mCrCIR03G05	FR677578	CCACACAGGCAGACA	CCTTGGAGGAGCTTTAC	50	199-228
mCrCIR04H06	FR677579	GGACATAGTGAGAAGTTGG	CAAAGTGGTGAAACCTG	55	190-196
mCrCIR05A05	FR677580	ATACCTGTGAGCGTGAG	CCTCTTCCCTTCCATT	50	144-162
mCrCIR07D05	FR677574	TCGTTCTTGCTTTTCCAC	GAATCAAACTACCCTCCAAT	55	206-208
mCrCIR07D06	FR677581	CCTTTTCACAGTTTGCTAT	TCAATTCCTCTAGTGTGTGT	55	166-188

2n gamete allelic structure inference from triploid hybrid genotypes

For a locus bearing totally different parental allelic configurations ($A_1A_2XA_3A_4$), the genotype of the 2*n* gamete was directly read from the triploid hybrid structure. When the male and female genitor shared one allele ($A_1A_2XA_2A_2$ and $A_1A_2XA_2A_3$), the inference of the 2*n* female gamete structure was carried out from the measured allele dosage by the microsatellite DNA allele counting-peak ratio method [MAC-PR; (Esselink *et al.*, 2004)], for the triploid hybrids that have inherited the common allele from the male genitor. The validation of the MAC-PR method for the analyzed loci and populations is given as Supplementary Information.

Identification of 2n gamete origin by single-locus analysis

Once female gamete structures were inferred, the percentages of heterozygosity were calculated for each locus in the whole population and for each genotype over all analyzed loci. Without previous knowledge about relative markers/centromere position, the observation of heterozygosity restitution over 50% for a single locus is not informative because it could have come from either FDR or SDR; however, theoretical heterozygosity restitutions lower than 50% are only found for SDR (Park *et al.*, 2007). When such low values of heterozygosity restitution were observed for a marker, we compared the highest probability of such a population structure under the SDR and FDR hypothesis. Under SDR, the highest probability for such an observation is obtained for a centromere position, leading to a theoretical proportion of *h* heterozygous gametes, whereas the best theoretical proportion of heterozygous gametes to fit with such observed data is 0.5 in the case of FDR. Thus, logarithm of the odds ratios (LODs) were estimated as:

$$LOD = \log_{10}\left[\frac{p(SDR)}{p(FDR)}\right] = \log_{10}\left[\frac{h^{nh}\cdot(1-h)^{(1-h)\cdot n}}{0.5^{nh}\cdot(0.5)^{(1-h)\cdot n}}\right]$$

with *h* being the heterozygosity transmission observed for the marker and *n*, the number of genotypes analyzed with this marker.

Comparison of observed heterozygosity restitution within a linkage group with theoretical functions to infer the 2n gamete formation mechanism and centromere position

To integrate the information on all loci of the same linkage group coming from different populations, we propose a method based on the comparison between observed heterozygosity restitution among a linkage group with theoretical restitution functions under different models. We have tested FDR and SDR mechanisms, for no-interference and partial chiasma interference on a chromosome arm (and no between-arm interference) assuming several centromere positions (interval of 0.5 cM between two theoretical adjacent positions of the centromere). Discrepancies between the different models and the observed data were estimated by the sum of the squared differences between observed and theoretical values at the marker map positions. Under one model of restitution (FDR or SDR), let Fit(*c*) be the value of the sum of the squared distance for each position of the centromere; the best theoretical centromere position under this model is deduced by searching *c*, which minimises Fit(*c*). The confidence interval (95%) for the centromere position was estimated by bootstrap on the loci (500 bootstraps). For this analysis, we have used the marker position of the clementine genetic map (Ollitrault *et al.*, 2011), which should be very similar to the Fortune map, because 'Fortune' is a hybrid between Clementine and Dancy mandarin.

<u>*No interference model*</u>. According to Zhao and Speed (Zhao and Speed, 1998b), assuming that the number of chiasmata in an interval follows a Poisson process (no interference model corresponding to Haldane's map function), the probabilities of a tetrad displaying a first division segregation (FDS; (Griffiths *et al.*, 1996) pattern and a second division segregation (SDS; (Griffiths *et al.*, 1996) pattern are as follows: $p(FDS) = \frac{1}{3}(1+2e^{-3d})$ and $p(SDS) = \frac{2}{3}(1-e^{-3d})$, where *d* is the genetic distance in Morgan units (Haldane's map function) between a given locus and the centromere. Under the FDR mechanism for 2*n* gamete formation, the FDS tetrads and half of the SDS tetrads transfer the parental heterozygosity, while under the SDR mechanism, the restitution of parental heterozygosity will result from SDS tetrads (Zhao and Speed, 1998a).

We can thus derive the rates of heterozygosity transmission (H) as function of the distance to the centromere (*d*): $H(d)=\frac{1}{3}(1+2e^{-3d})+\frac{1}{3}(1-e^{-3d})=\frac{1}{3}(2+e^{-3d})$ for FDR and $H(d)=\frac{2}{3}(1-e^{-3d})$ for SDR. According to this model, from the centromere to the telomere, H varies between 1 and 2/3 for FDR and from 0 to 2/3 for SDR (Figure 2.2).

Let *p* and *c*, respectively, be the positions of a locus and of the centromere in a linkage group. As Haldane's map function is additive, the distance between the considered locus and the centromere is $d=|p-c|$. The heterozygosity restitution H(d) as a function of the distance to the centromere (*d*) can thus be applied to any locus position (*p*) on the clementine's genetic map after transposition according to each theoretical position of the centromere (*c*) on the linkage group: $H(p)=\frac{1}{3}(2+e^{-3|(p-c)|})$ for FDR and $H(p)=\frac{2}{3}(1-e^{-3|(p-c)|})$ for SDR. Theoretical curves of *H(p)* are presented in Figure 2 for FDR and SDR models.

Partial interference models. There are several proposals in the literature to incorporate chiasma interference in relating the map distance and SDS ordered tetrad proportion. (Zhao and Speed, 1998b) developed functions for ordered tetrad frequencies derived from the chi-square chiasma interference models, which provide good fits to data from different organisms (Zhao *et al.*, 1995). Moreover, most map functions could be approximated by one of the chi-square models (Zhao and Speed, 1996). The model is represented in the form $Cx(Co)^m$, where *m* is a parameter positively related to the interference level. For m>0, the SDS proportion rises above 2/3 and as *m* increases, the maximal frequency of SDS increases (Zhao and Speed, 1998b). Likewise for the no interference model, the restitution of heterozygosity by half-tetrad under FDR and SDR can be directly derived from the SDS tetrad frequency for the different chi-square models; as examples, Figure 2 shows the pattern of heterozygosity restitution along a linkage group for the Cx(Co) and $Cx(Co)^4$ models. The heterozygosity restitution at any position (*p*) on the Clementine's genetic map can therefore be inferred for each theoretical position of the centromere (*c*) on the linkage group and thus, Fit(c) evaluated. The genetic map is established according to the map function corresponding to each considered chi-square model (Foss *et al.*, 1993).

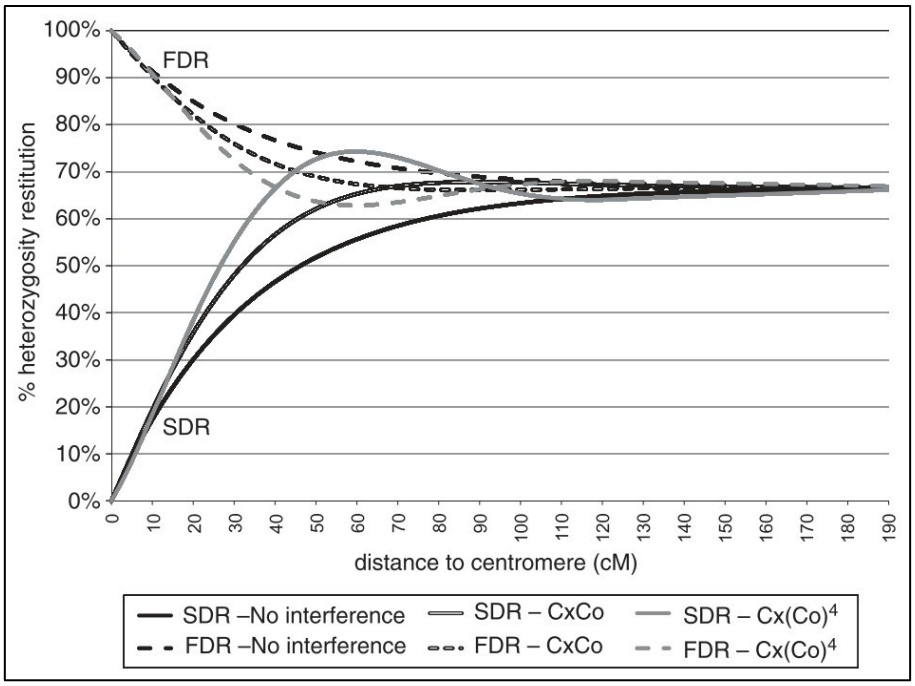

Figure 2.2. Theoretical curves of heterozygosity restitution as function of genetic distance to the centromere for FDR and SDR without chiasma interference and two χ^2-(Cx(Co)^m) models of partial interference.

Identification of the mode of 2n gamete formation and localization of the centromere by half-tetrad multilocus structure analysis

Multilocus analyses were performed on 87 genotypes and five loci mapped in the linkage group whose order was obtained from the current map of clementine (Ollitrault *et al.*, 2011). The analysis was made according to Tavoletti *et al.* (1996) assuming that multiple crossovers do not occur between contiguous markers. Under this hypothesis, each crossover between two markers, *i* and *i+1*, leads to a change from homozygosity to heterozygosity in the case of SDR and a half change from heterozygosity to homozygosity in the case of FDR. Thus, the distance between two adjacent markers ($d_{M_iM_{i+1}}$) can be estimated by the proportion of 2*n* gametes with changes (homozygosity versus heterozygosity; $C_{M_iM_{i+1}}$) between the two markers. For FDR, $d_{M_iM_{i+1}} = C_{M_iM_{i+1}}$ and for SDR, $d_{M_iM_{i+1}} = ½\ C_{M_iM_{i+1}}$.

The probability of the observed multilocus progeny under the different models (FDR or SDR with different centromere positions) was calculated according to Tavoletti *et al.* (1996). Detailed formulas can be found in supplementary information. Both under FDR and SDR models, the centromere position was virtually moved from before the first considered locus along the linkage group to after the last considered locus (intervals of 0.5 cM) and the relative probability was estimated for each position. The LOD between best position under SDR and FDR was calculated in order to determine the mode of 2*n* gamete restitution and the position of the centromere was considered as the one producing the highest relative probability under the identified mode of restitution. The confidence interval was estimated following the LOD drop-off

method (Lander and Botstein, 1989). After determining the centromere position, chiasma interference can be estimated for each chromosome arm with three point linkage mapping as follows:

Let $r_{M_1M_2}$ denote the observed recombination rate (heterozygous to homozygous and vice versa) between the locus *1* and *2*; $r_{M_2M_3}$ the observed one between locus *2* and *3*; and *rd* the observed rate of double recombination between the three loci. Thus, chromosome interference (I) is: $I = 1 - \left[\frac{rd}{r_{M_1M_2} \cdot r_{M_2M_3}} \right]$

Results

Genotyping of triploid progenies

Thirty-five heterozygous SSR markers in 'Fortune' were selected and used for genotyping the different triploid families according to their polymorphism between Fortune and the male genitors. Overall, 22, 18, 21 and 26 of these markers were polymorphic between the two parents for the families with 'Minneola', 'Common Mandarin', 'Ellendale' and 'Murcott' (Table 2.2) as male parent.

Table 2.2. Polymorphism observed among 'Fortune' mandarin and the male genitors for the 35 heterozygous loci in 'Fortune'.

	Heterozygous				Homozygous			No Data
	Num	TP	1 CA	2 CA	Num	TP	1 CA	
Minneola (P1)	25	0	16	9	6	1	5	4
Common mandarin (P2)	19	1	10	8	7	1	6	9
Ellendale (P3)	19	4	9	6	8	4	4	8
Murcott (P4)	13	6	6	1	14	7	7	8

Num: number of SSR markers for each class; TP: SSRs totally polymorphic between Fortune and male parent; CA: SSRs with common alleles between Fortune and male parent. Numbers in grey cells indicate primers used for genotyping each population. Codes in brackets indicate population (e.g. P1: Fortune X

The unambiguous differentiation of allele dosage in heterozygous triploids has been confirmed by the very clear bimodal distribution of the peaks area ratio of the different triploid hybrids for all markers (Supplementary information). The loci with total allelic differentiation between female and male genitors enabled (based on heterozygosity transmission or allele dosage estimation) the genitor producing the 2*n* gametes for each triploid hybrid to be identified unequivocally. Maternal origin of the 2*n* gamete has been observed for all the analyzed triploids; therefore, based on the maternal origin of the 2*n* gamete and allele dosage, it was possible to infer the 2*n* gamete structure from the triploid hybrid genotypes. Potential allelic segregation

distortion in the 2*n* gamete population was tested for each marker by χ^2 analysis (Table 2.3); only one of them (*Ci02B07*) showed significant distortion.

Table 2.3. Heterozygosity restitution (% HR) for each analyzed locus, X^2 test for allelic segregation, and LOD SDR/FDR for markers with less than 50% of heterozygosity restitution, linkage group (LG), number of individuals (NI) and populations (NP) analyzed.

Marker name	% HR	X^2	LOD SDR/FDR	LG	NI (NP)	Marker name	% HR	X^2	LOD SDR/FDR	LG	NI (NP)
mCrCIR06B05	16.1	1.76	16.29	I	149 (P1-P2-P3-P4)	**CX6F03**	65.4	0.24		V	153 (P1-P2-P3-P4)
mest121	36.9	0.60	0.97	I	65 (P2-P3)	mCrCIR06A12	70.1	0.00		V	67 (P4)
CAC15	17.58	0.17	9.02	II	91 (P1-P4)	**mCrCIR07E12**	67.3	0.38		V	101 (P2-P4)
Ci01C07	75.26	0.75		II	97 (P1-P2-P3)	**mest56**	69.7	0.45		V	132 (P1-P3-P4)
CX6F23	14.29	0.14	16.35	II	133 (P1-P3-P4)	**mest104**	59.4	0.96		V	128 (P1-P3-P4)
mCrCIR02D09	82.46	0.27		II	171 (P1-P2-P3-P4)	**mCrCIR01C06**	64.3	0.39		VI	129 (P1-P3-P4)
mCrCIR02G01	79.41	0.07		II	34 (P2)	**mCrCIR01E02**	68.5	0.32		VI	124 (P1-P3-P4)
mCrCIR03C08	34.33	0.73	1.45	II	67 (P4)	**mCrCIR02F12**	74.4	0.43		VI	164 (P1-P2-P3-P4)
mCrCIR04H06	61.54	0.00		II	130 (P1-P3-P4)	**mest123**	65.2	0.20		VI	66 (P4)
mCrCIR05A05	78.03	0.02		II	132 (P1-P3-P4)	**mest192**	67.0	0.24		VI	103 (P2-P4)
mCrCIR06B07	73.13	0.44		II	67 (P2-P3)	**mest488**	74.1	0.24		VI	139 (P1-P3-P4)
mCrCIR07D05	26.67	0.00	1.47	II	30 (P1)	**mCrCIR03B07**	34.1	2.97	3.03	VII	135 (P2-P3-P4)
mest110	73.53	0.17		II	102 (P1-P2-P3)	**mCrCIR01F04a**	39.8	0.57	1.07	VIII	118 (P1-P3-P4)
mest247	80.00	0.07		II	35 (P2)	**mCrCIR02A09**	65.1	0.60		VIII	86 (P1-P4)
TAA41	73.74	0.31		II	99 (P3-P4)	**mCrCIR02G02**	10.3	0.03	13.62	VIII	87 (P2-P4)
mCrCIR03G05	72.2	0.02		IV	97 (P3-P4)	**Ci02B07**	72.3	4.57		IX	101 (P1-P2-P4)
mCrCIR07D06	11.94	0.02	19.05	IV	134 (P1-P2-P4)	**Ci08C05**	17.4	0.16	13.85	IX	138 (P1-P2-P4)
						mCrCIR07F11	41.6	0.01	0.77	IX	125 (P1-P3-P4)

Maternal heterozygosity restitution in the 2n eggs

Restitution of maternal heterozygosity in each 2*n* gamete (based on all analyzed loci) varied between 15.38% and 100%, with 54.98% as mean value. The unimodal distribution of heterozygosity restitution in the 2*n* megaspores among the analyzed genotypes suggests that all these 2*n* gametes arise from the same mechanism (Figure 2.3).

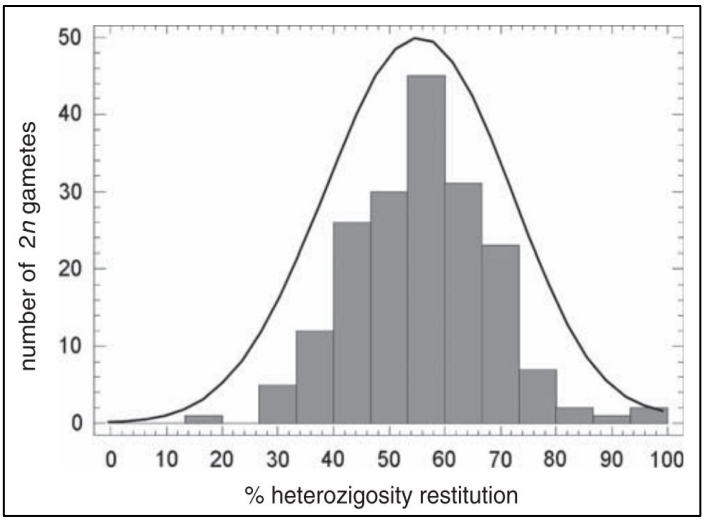

Figure 2.3. Distribution and density trace of the heterozygosity restitution rates estimated for each 2*n* megaspore (estimation of restitution rates based on all markers).

Global heterozygosity restitution for each marker

Rate of maternal heterozygosity restitution was calculated for 35 loci, covering eight out of nine linkage groups from the current map for Clementine (Table 2.3); the average rate was 55.23% and values varied from 10.34% (*mCrCIR02G02* marker) up to 82.46% (*mCrCIR02D09* marker). Twelve of the analyzed markers displayed < 50% maternal heterozygosity restitution. For these markers, LODs of SDR/FDR probabilities were calculated and varied between 0.77 and 19.05. These observations could only fulfil the SDR hypothesis, with markers that are close to the centromere, and rule out the FDR scenario.

Pattern of heterozygosity restitution within linkage group II and centromere mapping

Twelve of the analyzed markers in linkage group II of the Clementine genetic map (Ollitrault *et al.*, 2011) were used to perform this analysis. Taking into account marker order and the distances between them (Haldane's map function), the pattern of heterozygosity is represented in Figure 2.4 (dots). From one side of the linkage group to the other, the heterozygosity decreases from 82.46% for *mCrCIR02D09* marker to 14.29% for *CX6F23* marker and increases again up to 80% as the position rises on the map. Such a pattern within a linkage group is totally incompatible with the FDR model, in which the opposite variation is expected and the lowest theoretical restitution value would be 50%. Therefore, the search for the centromere position producing the best fit between the theoretical function of heterozygosity restitution and the observed data was only conducted under the SDR hypothesis. Under the SDR hypothesis with no interference, the best estimation of centromere position is 10.1 ± 6.4 cM from *CX6F23* marker (Figure 2.4). Without chiasma interference, the maximum theoretical heterozygosity restitution is 2/3, while we have observed values over 75% for five telomeric

markers (two of them with very low *p*-value: mCrCIR02D09, $p=10^{-5}$ and mCrCIR05A05, p=0.006), suggesting the existence of chiasma interference. On the other hand, the maximum restitution found over all analyzed markers was 82.5% while 100% should be expected for telomeric markers in case of total interference. Therefore, among the χ^2 models proposed by Zhao and Speed (1998a) for partial interference, we have tested the $Cx(Co)^4$ model, allowing a maximum restitution of heterozygosity close to 75%. Under this model of chiasma interference and SDR hypothesis, the best estimation of centromere position is 11.8 ± 7.1 cM from *CX6F23* marker (Figure 2.5). This model fits better with the observed data than the no interference model (Fit(*c*) = 0.08 and Fit(*c*) = 0.24 respectively).

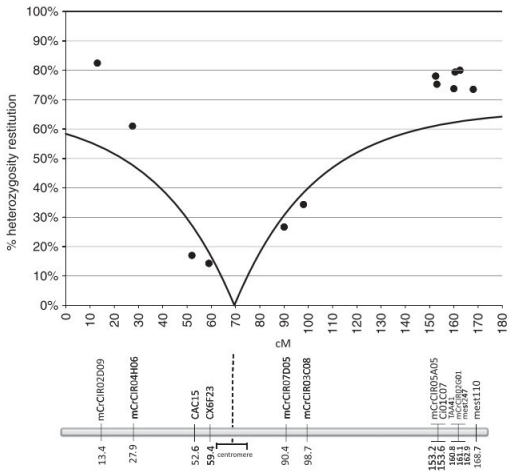

Figure 2.4. Observed heterozygosity restitution values for markers in linkage group II (dots) and theoretical heterozygosity restitutions (line) for the best-fitting centromere position under SDR model without chiasma interference (markers in x axis are positioned according to Clementine genetic map using Haldane's map function).

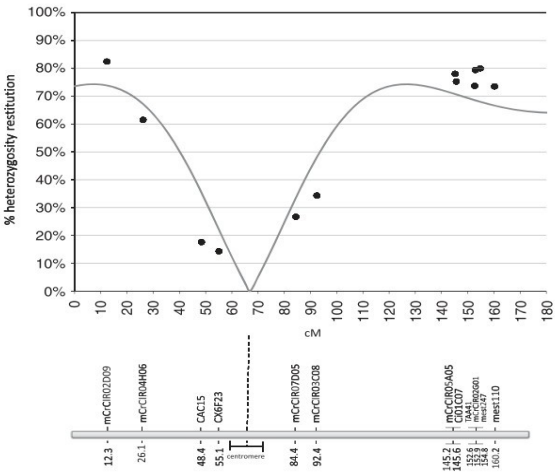

Figure 2.5. Observed heterozygosity restitution values for markers in linkage group II (dots) and theoretical heterozygosity restitutions (line) for the best-fitting centromere position under SDR and the $Cx(Co)^4$ model of partial interference (markers in x axis are positioned according to Clementine genetic map using the corresponding $Cx(Co)^4$ map function).

Multilocus half-tetrad structure analysis in linkage group II

Considering homozygosity and heterozygosity at each locus, 15 different multilocus profiles have been observed. These profiles and the number of corresponding 2*n* gametes are given in Table 2.4.

Table 2.4. Heterozygous (HE) and homozygous (HO) profiles for 87 genotypes and five simple-sequence repeat markers within linkage group II.

N	M1	M2	M3	M4	M5
8	HE	HE	HE	HE	HE
1	HE	HE	HE	HE	HO
3	HE	HE	HE	HO	HE
5	HE	HE	HO	HO	HO
30	HE	HE	HO	HO	HE
1	HE	HE	HO	HE	HE
7	HE	HO	HO	HO	HO
18	HE	HO	HO	HO	HE
1	HE	HO	HE	HE	HE
1	HO	HO	HO	HO	HO
7	HO	HO	HO	HO	HE
1	HO	HO	HE	HE	HE
1	HO	HE	HO	HE	HO
2	HO	HE	HE	HO	HE
1	HO	HE	HO	HO	HE
H	85.06	59.77	18.39	14.94	82.76

N, Number of genotypes for each profile; M1, mCrCIR02D09; M2, mCrCIR04H06; M3, CAC15; M4, CX6F23; M5, mCrCIR05A05; H, heterozygosity restitution percentage for 87 genotypes. Gray shading indicates heterozygous regions.

Probabilities of the observed 2*n* gamete population under the FDR and SDR hypotheses for moving centromere positions have been calculated from this table. The LOD value of the SDR highest probability / FDR highest probability was 6.8, confirming that SDR is the most probable model. Under the SDR hypothesis, the probability variation as a function of the centromere position suggests that the most probable position is between 2.25 cM and 7.00 cM (Morgan's map function; which is between 2.30 cM and 7.54 cM with Haldane's map function) close to the *CX6F23* marker, between the former and the *mCrCIR05A05* marker. This confidence interval overlaps the ones of the estimations done by the best fit method. If up to two crossovers per chromosome occur, it is possible to observe phase-changing between two markers when complementary crossovers take place (*i.e.* two crossovers involving all four chromatids) via the SDR restitution mechanism (Figure 2.1). The HTA analysis detected 16 complementary crossovers, by revealing allelic phase changes between markers in homozygosity, as shown in Table 2.5; only two double crossovers affecting a chromatid pair have been identified.

Table 2.5. Number of observed crossing over events on each arm in chromosome II for 87 genotypes and five markers

		Arm 1					
	Number c.o.	**0**	**1**	**2**	**3**	**4**	
	0	1	9	0	1(1)	0	**12.64%**
Arm 2	**1**	3	56	7(7)	5(3)	1(1)	**82.76%**
	2	0	4(4)	0	0	0	**4.60%**
		4.60%	**79.31%**	**8.05%**	**6.90%**	**1.15%**	

Numbers in brackets indicate detected complementary crossovers; percentages of crossover events (0, 1, 2, 3, 4) in each chromosome arm are given in bold.

Interference analysis

Considering centromere position between *CX6F23* and *mCrCIR07D05* markers, the interference coefficient was estimated for each arm of chromosome II with three markers per side. For arm 1, analyzing 87 genotypes for *mCrCIR02D09*, *mCrCIR04H06* and *CX6F23* markers, the interference coefficient was found to be 0.73. For the other arm, the interference estimation was 0.53, analysing 66 2*n* gametes for *mCrCIR03C08*, *mCrCIR05A05* and *TAA41* markers.

Discussion

MAC-PR is an efficient method to determine allelic configurations in triploid citrus segregating progenies

In this work, the HTA is based on gamete allelic configuration inferred from triploid progeny genotyping. MAC-PR has been proposed to deal with differential amplification intensities among alleles in polyploid plant species (Esselink *et al.*, 2004) for allele dosage estimation. Under the PCR conditions used for citrus SSR analysis, we have successfully verified the correlation between allele dosage and PCR product ratio. Finally, the very clear bimodal distribution of estimated doses for the 35 analyzed SSR markers among triploid hybrids ruled out the occurrence of random PCR drift in our amplifications and validated the MAC-PR approaches for triploid citrus progenies genotyping. A basic assumption of the MACPR method is the repeatability of relative allelic amplification intensities among individuals and, thus, the total homology of primer sites for DNA fragments producing the same allele (same PCR product size). Homoplasy of SSR markers was found at interspecific levels in citrus (Barkley *et al.*, 2009), and could limit general application of MAC-PR. 'Fortune' mandarin and most of the male genitors used in our study are closely related, which should explain why we have not encountered this difficulty because of the by-descent homology of alleles.

Origin of 2n gamete producing triploids in citrus

We observed that all the 171 analyzed triploid hybrids show maternal heterozygosity restitution for at least one marker. This confirms that all triploid hybrids found in the progenies of 2*x* × 2*x* crosses with 'Fortune' mandarin as female genitor arose from 2*n* megaspores. This result is in agreement with the cytogenetic observations done by Esen and Soost (1971), and with previous molecular observations (Luro *et al.*, 2004; Chen *et al.*, 2008a; Ferrante *et al.*, 2010). In the present work, the restitution heterozygosity rates significantly lower than 50% for several markers (distributed in 6 of the 8 represented linkage groups) and the pattern of heterozygosity inside linkage group II, definitively ruled out the FDR hypothesis. Multilocus 2*n* gamete allelic configuration in the same linkage group also revealed that SDR was much more likely to be the mechanism underlying unreduced gamete formation than FDR (LOD=6.8). Moreover, this analysis enabled us to detect four strand (complementary) double crossovers by phase changes of several homozygous markers. Such phase changes between homozygous positions are only possible under the SDR hypothesis if up to two crossovers per arm are considered. This conclusion for SDR is in agreement with that proposed by (Luro *et al.*, 2004), who observed low heterozygosity restitution in *C. clementina* 2*n* megagametophyte. The conclusion of FDR given for sweet orange (Chen *et al.*, 2008a) is questionable because of the low number of analyzed markers. Indeed, the unambiguous identification of FDR without previous location of the centromere must be based on a large set of markers with good genome coverage. In the same way, the results of Ferrante *et al.* (2010), based on a very low number of individuals and markers for each parental genotype are not sufficient to prove the authors's conclusions of SDR for 'Fortune' and 'Wilking' mandarin and FDR for lemon. Systematic analysis of 2*n* gamete allelic configuration with the same set of loci, close and far from the centromere, will shed light on whether SDR is the only mechanism underlying 2*n* egg formation in citrus, or whether there is a different mechanism depending on genotype or environmental conditions.

Analysis of the pattern of heterozygosity restitution within a linkage group is an alternative way to map the centromere compared with half tetrad multilocus allelic configuration analysis

Centromere mapping has been carried out in several crops (Douches and Quiros, 1988; Okagaki *et al.*, 2008) and animals (Kauffman *et al.*, 1995; Lindner *et al.*, 2000), using half-tetrad analysis. In the present work, half-tetrad analysis has been carried out in two ways: by multilocus allelic conformation analysis, as described in Tavoletti *et al.* (1996) and by comparison of observed and theoretical pattern of heterozygosity restitution rate within the linkage group under several models. Both methods estimated the centromere position to be between *CX6F23* and *mCrCIR07D05* markers. Confidence intervals of the positions obtained with the two methods overlapped, validating the best fit method. The fitting curve adjustment has the advantage of potential application to a set of loci analyzed in different progenies

(between a same parent producing the 2*n* eggs but different male parents), potentially enlarging the usable set of markers to all those heterozygous for the female 2*n* gamete producer. It could potentially be used to compare a large range of interference model functions. Furthermore, this method should be easily applied to dominant markers by estimating the heterozygous restitution as 1-2*f* (with *f* being the frequency of homozygous progeny for the recessive allele). However, it requires the use of an existing genetic map and assumes that crossover distribution is similar during normal meiosis and 2*n* gamete formation. Multilocus allelic HTA is advantageous in that it can be used without a previous genetic map of the markers and can be applied with a predefined order of markers (Tavoletti *et al.*, 1996; Da *et al.*, 1995) as we have done, or without any previous information about marker position (Da *et al.*, 1995). An excel template has been developed for easy identification of restitution model (FDR or SDR) and positioning of centromere within a linkage group, from heterozygosity restitution data for a set of mapped loci. It includes an estimation of confidence interval for the centromere position by bootstrap on the loci. It is available under request to the authors.

Evidence for positive chiasma interference in citrus cv 'Fortune'

Many models of half-tetrad analysis are based in the hypothesis of complete interference. In the present work, the analysis of multilocus configuration in linkage group II revealed the occurrence of up to four crossovers in the same chromosome arm and thus, incomplete interference. It is confirmed by the maximum restitution values between 75% and 82% observed for five of the markers, while for SDR, the maximum restitution should be 66.6% under the no interference hypothesis and should reach 100% for total interference. A better adjustment was found between observed data and theoretical curve with the $Cx(Co)^4$ chi-square model for partial interference than the no interference one. Interference values were estimated by three point analysis for each arm of linkage group II, with results suggesting that it should be higher for one arm (0.73) than for the other (0.53). Such variation of interference level between different parts of the genome has also been observed in Arabidopsis (Drouaud *et al.*, 2007), in human (Lian *et al.*, 2008) and in mouse (Broman *et al.*, 2002).

Implications for citrus breeding

Seedlessness is one of the most important characteristics for the citrus fresh fruit market. An efficient way to achieve this aim is to obtain triploid mandarin varieties. Sexual triploidization is a classical method to obtain triploid citrus hybrids (Ollitrault *et al.*, 2008). Indeed, some genotypes such as 'Temple', 'Wilking', and 'Fortune' mandarin (Esen and Soost, 1973; Aleza *et al.*, 2010b) are well known for their production of diploid megagametophytes. The other classical method to create triploid citrus progenies is inter-ploidy hybridization with doubled-diploid. Assuming an SDR origin of 2*n* gametes in 'Fortune' mandarin, sexual polyploidization may lead to lower average of heterozygosity restitution than interploid

hybridization, whatever the segregation model considered for the doubled diploid (Marsden *et al.*, 1987). As heterozygosity and epistatic interactions are maintained for a great number of individuals in the progeny from interploid crosses with doubled-diploid, this triploid breeding strategy should be more efficient than 2*x* × 2*x* hybridization for developing new cultivars that are phenotypically close to 'Fortune mandarin' genitor. Conversely, 2*x* × 2*x* hybridization should produce more polymorphic progenies, by creating larger number of new multilocus allelic combinations (David *et al*., 1995), providing the opportunity to select innovative products within the perspective of market segmentation as a commercial strategy.

Supplementary information

Validation of MAC-PR estimation of allele dosage

The inference of a diploid gamete structure when the two parents share one allele requires the ability to estimate allelic dosage in triploid progenies. In this paper we have used the MAC-PR method (Esselink *et al.*, 2004) to estimate allelic dosage after verification of its applicability in our populations and markers as described below.

In order to verify that the MAC-PR method can be applied to estimate the allelic dosage of our markers in triploid genotypes, with our PCR conditions, a preliminary trial was performed with five markers. The DNA of two haploid lines (from clementine cv 'Nules' and pumello cv 'Chandler') displaying allelic variability were mixed at different ratios (1:2, 1:1, 2:1 and 3:1). Ten replications (PCR and run in the genetic fragment analyzer) were done for each ratio and marker. Area ratios among the two observed peaks (alleles) were calculated for all runs and the correlation between DNA ratio and peak area ratio was analyzed. An illustration is given for the *mest56* marker (Figure 2.S1), showing a very high correlation among DNA proportion and peak area ratio (r^2=0.9902). Similar results were observed for the four other markers.

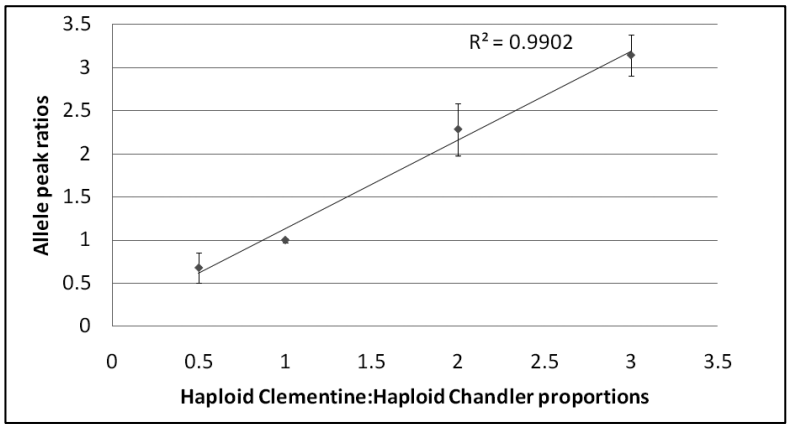

Figure 2.S1. Relationship between several DNA proportions of 'Clementine' and 'Chandler' haploids and peak area ratio for three repetitions of the *mest56* marker.

To avoid misinterpretation of allele dosage associated with PCR allele competition, the triploid hybrid peak area ratios were systematically corrected by 'Fortune' genitor peak ratio as:

$$SD = \frac{S_1 \cdot F_2}{S_2 \cdot F_1}$$

where *SD* is the estimated allele dosage ratio in the sample; S_i and F_i are the peak area for allele *i* for the analyzed sample and 'Fortune' diploid cultivar, respectively (Figure 2.S2).

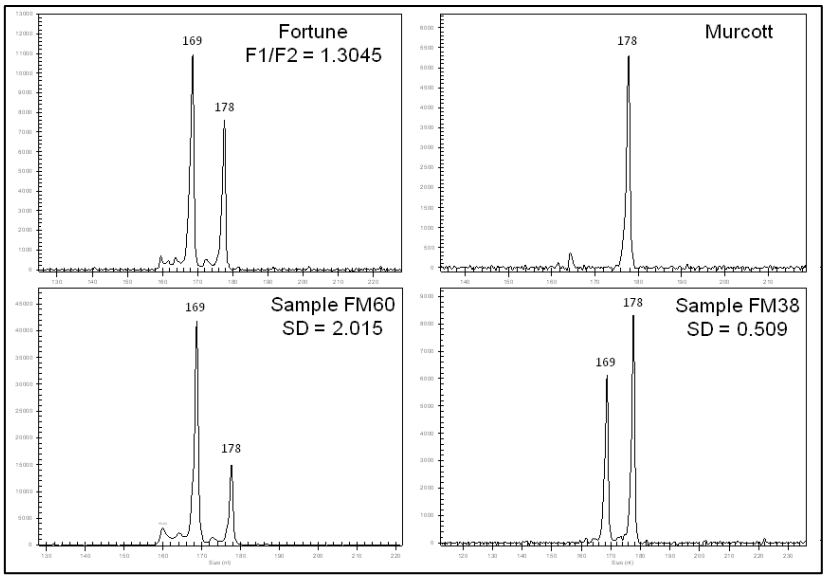

Figure 2.S2. Observed peaks for the CAC15 marker in 'Fortune', 'Murcott', the FM60 hybrid and the FM38 hybrid showing respectively ratio 2:1 and 1:2 between alleles 169 and 178. SD=(S1·F2)/(S2·F1)

Ten technical replications (independent PCR) were done for *mCrCIR06B05* with two triploid genotypes with ratios 1:2 and 2:1 to determine the discrimination power of the assignment of observed peak area ratios. The results (Figure 2.S3a) testified the high stability of allele dosage evaluation.

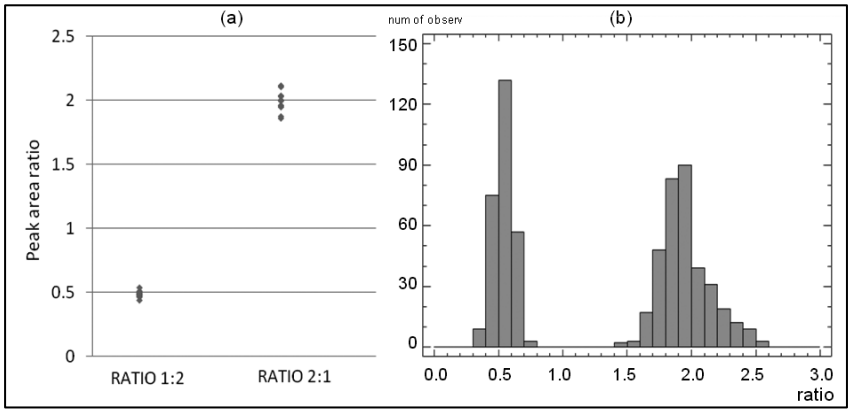

Figure 2.S3. Scatterplot of peak area ratios for ten PCR repetitions of two genotypes with ratios 1:2 and 2:1 for two alleles of *mCrCIR06B05* marker (a) and distribution of peak area ratios for all the genotypes and markers showing 1:2 and 2:1 ratios (b).

The unambiguous differentiation of allele dosage in heterozygous triploids has been confirmed during progeny analysis by the very clear bimodal distribution of the peaks area ratio of the different triploid hybrids for all markers (Figure 2.S3b). These results validate the method of area peak ratios used for genotyping the triploid progenies in this study.

Method for Half Tetrad Analysis

The analysis was made according to Tavoletti *et al*. (1996) assuming that multiple crossovers do not occur between contiguous markers. Under this hypothesis, each crossover between two markers leads to a change from homozygosity to heterozygosity between marker *i* and *i+1* in the case of SDR and a half change from heterozygosity to homozygosity in the case of FDR. Thus, the distance between two adjacent markers (d_{MiMi+1}) can be estimated by the proportion of 2*n* gametes with changes (homozygosity versus heterozygosity; C_{MiMi+1}) between the two markers. For FDR, $d_{MiMi+1} = C_{MiMi+1}$ and for SDR, $d_{MiMi+1} = ½\ C_{MiMi+1}$.

The probability of each observed genotype profile (G*j*), considering that the locus is homozygous (Ho) or heterozygous (He), was calculated as:

$$P(G_j) = (\Pi P_{MiMi+1})(P_{CMa} \cdot P_{CMa+1})$$

where:

- P_{MiMi+1} is the probability of configuration for two adjacent loci whose interval does not include the centromere; being $P_{MiMi+1} = C_{MiMi+1}$ in case of configuration change and $P_{MiMi+1} = 1 - C_{MiMi+1}$ when no change occurs between the adjacent loci.
- P_{CMa} and P_{CMa+1} are the probability of the observed configuration for the two loci flanking the centromere, P_{CMa} (and identically P_{CMa+1}) values being ($1-d_{CMa}$), (d_{CMa}), ($2d_{CMa}$) and ($1-2d_{CMa}$) respectively for the following situation (He, FDR), (Ho, FDR), (He, SDR), (Ho, SDR); with d_{CMa} being the mapped distance between the centromere and the given flanking locus.

Under FDR or SDR models and different centromere positions, the probability of the observed 2*n* gamete population (P) is:

$$P = C\prod_{i=1}^{j}(G_j)^n$$

where *n* is the number of 2*n* gametes with the genotype *j* and C a combinatory coefficient constant for the set of observed data.

Both under FDR and SDR models, the centromere position was virtually moved from before the first considered locus along the linkage group to after the last considered locus (intervals of 0.5 cM) and relative probability (P/C) was estimated for each position. The LOD between best position under SDR and FDR was calculated in order to determine the mode of 2*n* gamete restitution and the position of the centromere was considered as the one producing the highest relative probability under the identified mode of restitution.

CHAPTER 3

Maximum-likelihood method based on parental heterozygosity restitution of centromeric loci identifies SDR as the predominant mechanism leading to 2*n* megagametophytes in *C. reticulata*

José Cuenca, Pablo Aleza, José Juárez, Andrés García-Lor, Yann Froelicher, Luis Navarro and Patrick Ollitrault

Annals of Botany, Submitted

Abstract

Background: The use of unreduced gametes, resulting in the establishment of sexual polyploids is currently a major strategy for triploid citrus breeding. The origin mechanism of 2*n* gamete formation, i.e., FDR (First-Division Restitution) or SDR (Second-Division Restitution), greatly impacts the gametic structures and, therefore, the polyploid populations and the efficiency of breeding strategies. Methods previously developed to identify the underlying mechanism require the analysis of a large set of markers over large progeny.

Scope: Simulating populations were used to check the power of the method in terms of individual and marker numbers needed to obtain significant conclusions. This work develops a new maximum-likelihood method to identify the unreduced gamete formation mechanism both at the population and individual levels using independent centromeric markers. This new method was applied to investigate the mechanism in sixteen mandarin genotypes used as female parents in triploid citrus breeding.

Conclusions: Knowledge of marker-centromere distances greatly improves the statistical power of the comparison between the SDR and FDR hypotheses**.** The importance of selecting markers very close to the centromere to obtain significant conclusions at the individual level has been demonstrated by simulating data and illustrated by our results. All triploid hybrids analyzed originated from 2*n* megagametophytes. SDR was identified as the unreduced gamete formation mechanism in all mandarin genotypes analyzed. Moreover, SDR was identified as the restitution mechanism for 85.3% of the analyzed triploid hybrids, whereas 0.6% of the analyzed triploids were derived from FDR; for 14.1% of the hybrids, no significant conclusions were obtained.

Introduction

Polyploidization is a key source of species diversification and speciation in plants (Harlan and DeWet, 1975; Otto and Whitton, 2000; Soltis and Soltis, 2009) and may occur by somatic chromosome doubling (somatic polyploidization) or sexually through gametic nonreduction (sexual polyploidization) (De Storme and Geelen, 2013). Currently, most researchers consider sexual polyploidization to be the main mechanism of polyploidization in plants (Bretagnolle and Thompson, 1995; Otto and Whitton, 2000; Ramsey and Shemske, 2002).

In sexual polyploidization, polyploids are generated by the formation of diploid gametes, i.e., pollen or eggs that have the somatic rather than the gametophytic chromosome number (Harlan and DeWet, 1975; Bretagnolle and Thompson, 1995; Ramsey and Schemske, 1998; Otto and Whitton, 2000). In most cases, diploid gametes result from a restitution of the meiotic cell cycle (Bretagnolle and Thompson, 1995). In this process, meiotic cell division is converted into a mitosis-like nonreductional process, generating dyads (unreduced gametes) instead of the normal tetrads at the end of meiosis II. This phenomenon is referred to as 'meiotic restitution' and is the predominant mechanism of unreduced gamete formation in plants (Otto and Whitton, 2000; Soltis and Soltis, 2009).

Meiotic aberrations related to spindle formation, spindle function and cytokinesis can lead to unreduced gamete formation in plants. Up to seven major mechanisms of 2*n* gamete formation have been cytogenetically characterised: premeiotic doubling, first-division restitution (FDR), chromosome replication during the meiotic interphase, second-division restitution (SDR), postmeiotic doubling, indeterminate meiotic restitution, and apospory (Peloquin *et al*., 1989; Lim *et al*., 2004; Dewitte *et al*., 2012). FDR and SDR are the predominant mechanisms of 2*n* gamete formation. Failure of the first (FDR) or second (SDR) divisions leads to the formation of restitution nuclei with an unreduced chromosome number. An FDR 2*n* gamete contains non-sister chromatids, while an SDR 2*n* gamete contains two sister chromatids (Bretagnolle and Thompson, 1995; Tavoletti *et al*., 1996; Cai and Xu, 2007).

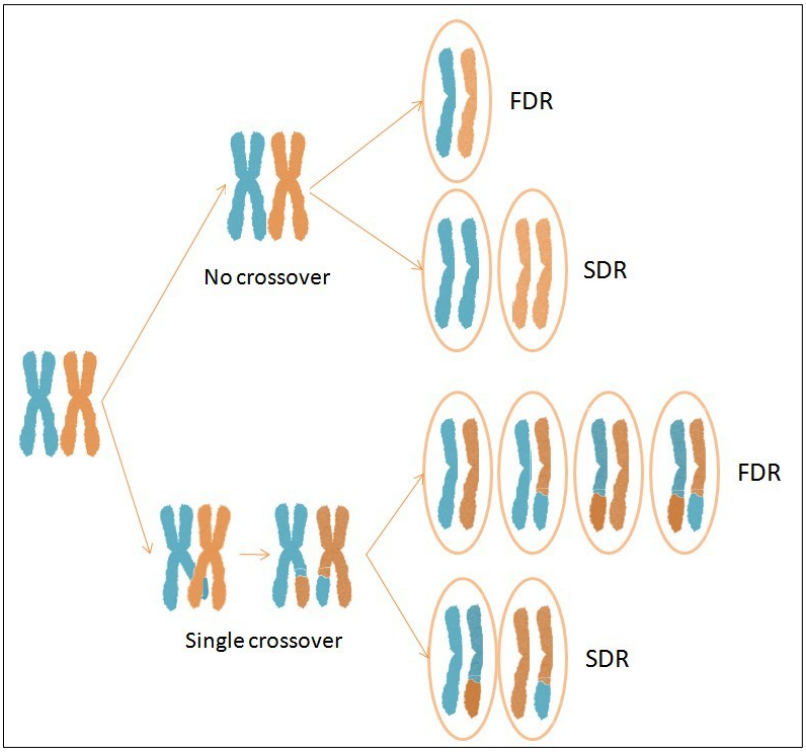

Figure 3.1. Half tetrads resulting from no crossover and single crossover events under FDR and SDR mechanisms of unreduced gamete formation.

The use of unreduced gametes in plant breeding (Ramanna and Jacobsen, 2003; Dewitte *et al.*, 2012), resulting in the establishment of sexual polyploids, is useful for improvement of crops such as lily (Lim *et al*., 2004; Barba-González *et al.*, 2004; 2005), maize (Rhoades and Dempsey, 1966), potato (Mok and Peloquin, 1975; Jongedijk *et al*., 1991), rose (Crespel and Gudin, 2003), rye (Lelley *et al*., 1987), alfalfa (Ortiz and Peloquin, 1991; Barcaccia *et al.*, 2003) and banana (Ortiz, 1997; Ssebuliba *et al*., 2008).

Diploidy is the general rule in *Citrus* and its related genera, with a basic chromosome number x=9 (Krug, 1943). However, triploid breeding has become an important strategic tool in the development of new seedless citrus commercial varieties (Starrantino, 1992; Ollitrault *et al*., 2008; Recupero *et al*., 2005; Aleza *et al*., 2010a; Cuenca *et al.*, 2010). Indeed, seedlessness is one of the most important economic traits related to fruit quality for fresh-fruit marketing of mandarins (Navarro *et al*., 2005; Ollitrault *et al*., 2008).

Spontaneous occurrences of citrus triploid hybrids arising from the union of 2*n* megagametophytes with haploid pollen have been noted since the seventies (Esen and Soost, 1971, 1973; Luro *et al*., 2004; Chen *et al*., 2008a; Aleza *et al*., 2010b). However, the frequency of such events is generally low (Cameron and Frost, 1968; Esen and Soost, 1971 Geraci *et al*., 1975) and extensive breeding programs based on this type of hybridization require very effective methodologies for embryo rescue and ploidy evaluation of large progenies (Ollitrault *et al*., 2008; Aleza *et al*., 2010b). To date, very few cases of citrus triploid hybrid occurrence in 2*x* × 2*x* crosses from unreduced pollen have been reported (Luro *et al*., 2004; Chen *et al*., 2008a).

Esen *et al*. (1979) proposed that, in citrus, 2*n* eggs result from the abortion of the second meiotic division (SDR) in the megaspore. This hypothesis was corroborated by molecular marker analysis for clementine (*Citrus clementina* Hort. ex Tan.) (Luro *et al*., 2004; Aleza *et al*., 2012b) and for 'Fortune' mandarin (*C. clementina* × *C. tangerina*) (Cuenca *et al*., 2011). By contrast, Chen *et al*. (2008a) proposed that 2*n* eggs of sweet orange (*C. sinensis* (L.) Osb.) resulted from first meiotic division restitution (FDR).

The origin of 2*n* gamete formation greatly impacts the gametic structures and, therefore, the polyploid populations and the efficiency of breeding strategies. Under FDR, non-sister chromatids retain parental heterozygosity from the centromere to the first crossover point, and the gametes thus transfer a large proportion of this parental heterozygosity to the progeny. Under SDR, the two sister chromatids are homozygous between the centromere and the first crossover point, and the resultant gametes have lower levels of heterozygosity than FDR gametes (Figure 3.1, Bretagnolle and Thompson, 1995). Several studies based on genetic markers indicate that FDR gametes transmit 70–80% of the parental heterozygosity, but SDR gametes transmit only 30–40% (Barone *et al*., 1995; Douches and Quiros, 1988; Vorsa and Rowland, 1997; Crespel and Gudin, 2003; Buso *et al*., 1999; Dewitte *et al*., 2012). Thus, a tighter distribution is expected in FDR-derived populations than in SDR ones because a higher percentage of the parental genome is transferred intact, resulting in a more uniform gamete production (Douches and Maas; 1998). Therefore, insights into the meiotic nuclear restitution mechanisms that produce unreduced gametes are crucial for the optimization of breeding strategies based on sexual hybridization (Errico *et al*., 2005).

Molecular marker analysis is a valuable tool for the estimation of heterozygosity transmission through diploid gametes to polyploid progenies and, therefore, to identify the mechanisms underlying unreduced gamete formation (Barone *et al*., 1995; Vorsa and Rowland, 1997; Bastiaanssen *et al*., 1998; Barcaccia *et al*., 2003; Luro *et al*., 2004; Chen *et al*., 2008a; Hayashi *et al*., 2009). The models described therein are all based on population analysis and suppose complete chiasma interference.

Cuenca *et al*. (2011) proposed an approach that takes into account different models of chromosome interference (i.e., no interference, partial interference or complete chiasma interference) when testing for FDR and SDR, and for mapping centromeres to linkage groups. This approach is based on functions of heterozygosity restitution (HR) at the population level along a chromosome in relation to locus-centromere distance (d) (Zhao and Speed, 1998a). Indeed, under FDR or SDR, HR is a direct function of the crossing over frequency between the considered locus and the centromere. It is, therefore, possible to implement the function (HR=f(d)) according to the FDR and SDR hypotheses while also taking into account different models of chromosome interference (Figure 3.2). This approach was successfully applied in populations of 2*n* ovules of 'Fortune' mandarin and 'Fina' clementine, and it was concluded that SDR was the main restitution mechanism and that partial chromosome interference occurs

(Cuenca *et al.,* 2011; Aleza *et al.,* 2012b). That study also contributed to the discovery of centromere locations in the nine citrus chromosomes.

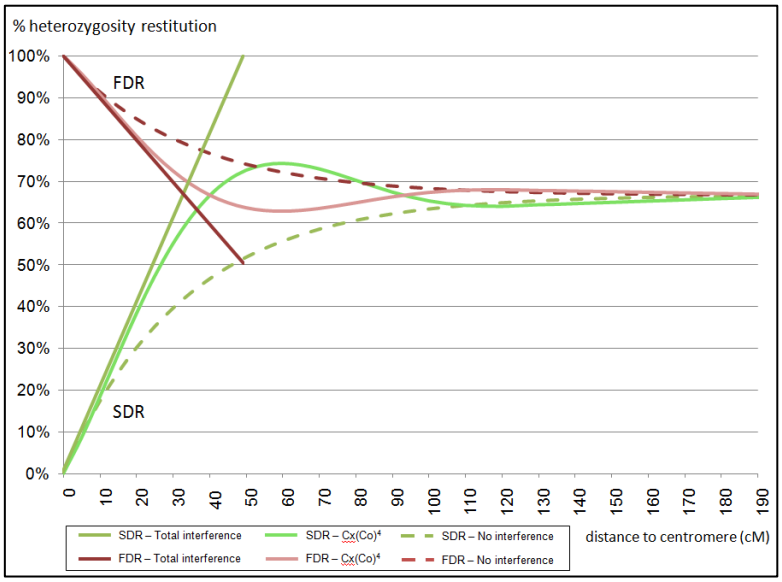

Figure 3.2. Rate (percentage) of heterozygosity restitution in the unreduced gametes under FDR and SDR mechanisms in function of the locus-centromere distance considering the total interference model, the no interference model and the $Cx(Co)^4$ partial interference model. Adapted from Cuenca *et al*. (2011).

In the present work, we propose a maximum-likelihood approach to test the SDR/FDR mechanism based on the HR of unlinked markers located close to the centromere of different chromosomes. This approach can be applied at the individual or population level. We simulated 2*n* gamete populations arising from FDR or SDR. This enabled us to identify the number of independent markers necessary to test in order to draw significant conclusions at the individual level in relation to marker/centromere distances, as well as the minimum population size necessary to be able to draw significant conclusions when analysing a defined number of markers.

Taking advantage of the known centromere locations (Aleza *et al.,* 2012b) within the nine linkage groups of the clementine reference genetic map (Ollitrault *et al.,* 2012a), we selected centromeric markers and used this maximum-likelihood method to (i) check the potential variability of origin between individuals for two genotypes in which SDR was proposed to be the predominant polyploidization mechanism as determined by population analysis (‘Fortune‘ mandarin, (Cuenca *et al*., 2011)) and clementines, (Luro *et al., 2004,* Aleza *et al.,* 2012b) and (ii) shed light on the mechanism leading to unreduced gamete formation in a range of mandarin genotypes used as female parents in 2*x* × 2*x* triploid breeding programs.

Materials and Methods

Plant materials

Analyses were performed using 497 triploid hybrids derived from 16 different mandarin genotypes as female parents in 2*x* × 2*x* cross populations (Table 3.1). The genotypes include six clementine and ten hybrid mandarin genotypes. Triploid hybrids were grown at the 'Instituto Valenciano de Investigaciones Agrarias' orchards in Moncada, Valencia, Spain. Practical details for the establishment of triploid populations from 2*x* × 2*x* crosses by embryo rescue and triploid selection by flow cytometry can be found in Aleza *et al*. (2010b). All triploid genotypes in the present study were selected after probation of their hybrid status by molecular marker analysis (data not shown). Taxonomic information about both female and male parental accessions is given in Table S3.1 according to the standard classification system for the *Citrus* genus (Swingle and Reece, 1967; Tanaka, 1977).

Table 3.1. Number of hybrids within each population analyzed in this study

# population	Population	Number of hybrids	# population	Population	Number of hybrids
1	Bruno × Chandler	17	*7d*	*Fortune × Willowleaf*	*37*
2	Clemenules × Nadorcott	23	8	Guillermina × Chandler	14
3	Ellendale × Fortune	69	9	Hernandina × Nadorcott	22
4	Encore × Ellendale	3	10	Honey × N'15	1
5	Fallglo × N'15	3	11	Imperial × Moncada	24
6	Fina × Nadorcott	87	12	Kiyomi × Nadorcott	21
7	Fortune × 4 male parents	197	13	Loretina × Chandler	2
7a	*Fortune × Ellendale*	*58*	14	Moncada × Ellendale	8
7b	*Fortune × Minneola*	*35*	15	Umatilla × Simeto	5
7c	*Fortune × Murcott*	*67*	16	Wilking × Fina	1

Selection of centromeric markers for the analysis of 2n gamete origin and formation mechanisms

Triploid citrus hybrids obtained in 2*x* × 2*x* hybridizations arise from unreduced megagametophytes (Esen and Soost, 1971, 1973; Geraci *et al*., 1975; Luro *et al*., 2004; Chen *et al*., 2008a; Aleza *et al*., 2010b; Cuenca *et al*., 2011). Therefore, markers heterozygous for the female parent and displaying polymorphism between the two parents were primarily selected for the molecular characterization of triploid hybrids and analysis of 2*n* gamete origin.

Centromere positions in all nine clementine chromosomes are known (Aleza *et al*., 2012b). Molecular markers within 20 cM of the centromere were used in this study because centromere-proximal markers are more informative with regard to the mechanisms of 2*n* gamete formation than centromere-distal markers (Ollitrault *et al*., 2012a). Within this range, the lowest expected HR rate is greater than 80% for FDR, while the highest HR for SDR is 40% (Figure

3.2). Twenty-four markers were selected for genotyping the triploid progeny. Between four and seven of these centromeric markers were used for genotyping each population (Table 3.2).

Table 3.2. Centromeric markers used for genotyping each triploid population.

LG	Centromere Position (cM)	Marker id	Marker type	Reference	Marker Position (cM)	Centomere-distance (cM)	Populations analyzed (#)
1	60.66	mCrCIR06B05	SSR	Froelicher *et al.*, 2008	50.27	10.39	7
		CID0806	InDel	Ollitrault *et al.*, 2012	55.17	5.49	8
		CIBE5720	SSR	Ollitrault *et al.*, 2010	58.45	2.21	4,5,11,12,14
		MEST539	SSR	In preparation	61.82	1.16	6
		MEST001	SSR	Luro *et al.*, 2008	70.60	9.94	10,15,16
		mCrCIR07D05	SSR	Cuenca *et al.*, 2011	75.60	14.94	1,13,14,15
2	56.87	CX2004	SSR	Chen *et al.*, 2008	46.67	10.20	15
		CX6F23	SSR	Chen *et al.*, 2006	49.53	7.34	1,2,4,5,6,7,9,10,11,12,13,15,16
3	90.59	CiC4225-01	SNP	Ollitrault *et al.*, 2012	86.33	4.26	4
		MEST470	SSR	In preparation	88.76	1.83	6
		CiC2167-02	SNP	Ollitrault *et al.*, 2012	90.60	0.01	12
		CX0124	SSR	Chen *et al.*, in prep	110.28	19.69	13,14,15
4	16.14	mCrCIR07D06	SSR	Cuenca *et al.*, 2011	16.33	0.19	1,7,8,13
		CF-ACA01	SSR	In preparation	24.41	8.27	2,4,6,9,11,15
5	23.12	CiC0004-01	SNP	Ollitrault *et al.*, 2012	20.90	2.22	6
		CiC0245-11	SNP	Ollitrault *et al.*, 2012	20.94	2.18	2,5,9
		MEST104	SSR	Garcia-Lor *et al.*, 2012	40.46	17.34	1,3,8,12,13,14,15,16
6	6.4	MEST191	SSR	In preparation	10.86	4.46	1,5,8,10,11,12,13,16
7	96.43	mCrCIR03B07	SSR	Cuenca *et al.*, 2011	83.39	13.04	7
		CX0114	SSR	Chen *et al.*, in prep	94.97	1.46	3
		CI07C07	SSR	Froelicher *et al.*, 2008	98.02	1.59	2,3,6,9,10,14
8	54.21	mCrCIR07B05	SSR	Froelicher *et al.*, 2008	31.70	22.51	3
9	52.16	mCrCIR07F11	SSR	Kamiri *et al.*, 2011	49.57	2.59	1,2,3,4,5,8,9,10,11,13,14,15,16
		CI08C05	SSR	Froelicher *et al.*, 2008	55.14	2.98	7

Genotyping of triploid hybrids

DNA extraction: Leaf DNA of triploid hybrids and their parents was isolated using the Plant DNAeasy kit from Qiagen Inc. (Valencia, CA, USA), following the manufacturer's protocol.

SSR and InDel analyses: Polymerase chain reactions (PCRs) were performed with wellRED oligonucleotides (Sigma-Aldrich®, St Louis, MO, USA) in a Mastercycler epgradient S (Eppendorf Scientific Inc., Westbury, NY, USA). The reaction (volume, 15 µl) contained 0.8 U Taq polymerase (Fermentas®, Burlington, VT, USA), 0.1 mM of each dNTP, 5 mM $MgCl_2$, 3 mM of each primer, and 30 ng of DNA in buffer containing 750 mM Tris-HCl (pH 9), 50 mM KCl, 200 mM $(NH_4)_2SO_4$, and 0.001% bovine serum albumin. The PCR program was 94ºC for 5 min; 40 cycles of 30 s at 94ºC, 1 min at 55ºC and 30 s at 72ºC, and a final elongation of 10 min at 72ºC. Separation was carried out by capillary gel electrophoresis (CEQ 8000 Genetic Analysis System; Beckman Coulter Inc., Fullerton, CA, USA). Data collection and analysis were carried out with GenomeLab GeXP (Beckman Coulter Inc.) version 10.0 software. Identification of allele

doses in heterozygous triploid hybrids was carried out using the MAC-PR method (Esselink *et al*., 2004) adapted in *Citrus* by Cuenca *et al*. (2011).

SNP analyses: SNP genotyping was performed by Kbioscience® services, using the KASPar technique. Experimental details were as described in Cuppen (2007). Identification of allele doses in heterozygous triploid hybrids was carried out using the relative allele signals as previously described for SNP markers (Cuenca *et al*., 2013a), based on competitive allele-specific PCR.

Analysis of the parental origin of the 2n gametes producing the triploid hybrids

For each hybrid, determination of the 2*n* gamete origin was carried out by identifying the parent that passed double genetic information to the hybrid. For markers displaying $A_1A_2 \times A_1A_1$ or $A_1A_2 \times A_1A_3$ configurations, the identification of $A_1A_2A_2$ or $A_2A_2A_3$ (i.e., double dosage of A_2, the allele specific to the female parent) configurations in the hybrid would imply a female origin of the 2*n* gamete. For the second combination, the observation of $A_1A_3A_3$ or $A_2A_3A_3$ (i.e., double dosage of A_3, the allele specific to the male parent) would indicate a male origin.

For markers displaying $A_1A_2 \times A_3A_3$ configurations in the parents, the identification of $A_1A_2A_3$, $A_1A_1A_3$, or $A_2A_2A_3$ configurations in the hybrid resulted from a maternal origin of the unreduced gamete, while $A_1A_3A_3$ or $A_2A_3A_3$ resulted from a paternal origin.

For markers with $A_1A_2 \times A_3A_4$ parental configuration, the identification of the following genotypes ($A_1A_1A_3$, $A_1A_1A_4$, $A_1A_2A_3$, $A_1A_2A_4$, $A_2A_2A_3$, $A_2A_2A_4$) and ($A_1A_3A_3$, $A_2A_3A_3$, $A_1A_3A_4$, $A_2A_3A_4$, $A_1A_4A_4$, $A_2A_4A_4$) implied, respectively, female and male origin of the 2*n* gamete.

Once the parental origin of the 2*n* gamete was identified, the inference of the allelic configurations of the unreduced gametes from triploid hybrid genotyping was carried out as previously described by Cuenca *et al*. (2011).

Identification of the restitution mechanism at an individual level

For loci heterozygous for the parent producing the 2*n* gamete, the probabilities of a 2*n* gamete being heterozygous or homozygous as a consequence of FDR or SDR mechanisms are direct functions of the marker-centromere distance.

To estimate such probabilities, the function relating HR rate and locus-centromere distance (Cuenca *et al*., 2011), derived from the $Cx(Co)^4$ partial chiasma interference model developed by Foss *et al*. (1993) and Zhao and Speed (1998a), could be used. Indeed, Cuenca *et al*. (2011) showed that this model fit better to 'Fortune' mandarin data (SDR mechanism) than total or no interference models. However, since selected markers are located close to centromeres (as explained above), for our data, the $Cx(Co)^4$ model and the total interference

model are equivalent (Figure 3.2). To simplify mathematical calculations of probabilities, the total interference model was used. Marker-centromere distances (d) were estimated from the clementine reference genetic map (Ollitrault *et al*., 2012a) and centromere locations in the clementine map were determined by Aleza *et al*. (2012b).

The probabilities of a marker being inherited as heterozygous under the SDR [$P_{SDR}(M_{He})$] or FDR [$P_{FDR}(M_{He})$] mechanisms were directly estimated from the total interference model functions as $P_{SDR}(M_{He})=2d$ and $P_{FDR}(M_{He})=(1-d)$. The probabilities of a marker being inherited as homozygous under SDR and FDR were estimated as $P_{SDR}(M_{Ho})=(1-2d)$ and $P_{FDR}(M_{Ho})=d$, respectively.

Therefore, the LOD values used to compare the probabilities of a heterozygous or a homozygous diploid gamete occurring at a locus, under the two models (SDR/FDR), were calculated respectively as:

$$LOD(M_{He}) = \log [P_{SDR}(M_{He})/P_{FDR}(M_{He})] = \log (2d/(1-d)) \quad \text{and}$$

$$LOD(M_{Ho}) = \log [P_{SDR}(M_{Ho})/P_{FDR}(M_{Ho})] = \log ((1-2d)/d)$$

For each restitution model, the probability of a single unreduced gamete [P(G)] presenting the observed allelic configuration for *i* unlinked markers (M_i) is the product of the probabilities of the observed genotype at each locus, $P(G)=\prod P_{Mi}$, and therefore the LOD value to compare the SDR/FDR models is the sum of the LOD at each locus,

$$LOD(G)=\sum LOD_{Mi},$$

where P_{Mi} and LOD_{Mi} are the probability and the LOD value of the observed genotype at the locus *I*, respectively.

As an example, if three unlinked loci (M_1, M_2 and M_3) were heterozygous, homozygous and homozygous, respectively, the probabilities of observing such gametes [P(G); (M_{1He}-M_{2Ho}-M_{3Ho})] are, respectively,

$$P_{SDR}(G) = 2d_1 \times (1-2d_2) \times (1-2d_3) \quad \text{under SDR and}$$

$$P_{FDR}(G) = (1-d_1) \times d_2 \times d_3 \quad \text{under FDR}$$

The LOD value used to compare the probabilities of SDR/FDR models is

$$LOD(G) = \log (2d_1/(1-d_1)) + \log ((1-2d_2)/d_2) + \log ((1-2d_3)/d_3)$$

where d_i is the distance from the locus *i* to its centromere.

LOD scores greater than 3 (the probability of the observed gamete is more than 1000-fold higher under the SDR model than the FDR one; LOD3) or greater than 2 (the probability of the observed gamete is more than 100-fold higher under the SDR model than the FDR one; LOD2) were considered as thresholds indicating that SDR was the mechanism involved in the single unreduced gamete formation, whereas LODs below -3 (or -2) indicate that FDR was the underlying mechanism; for LOD scores between -3 and 3 (or between 2 and -2), we considered that the mechanism could not be determined significantly.

Identification of the restitution mechanism at population level

Considering an infinite population of 2*n* gametes and a single locus, the probability of observing a sample of gametes [P(Pop)] with *j* heterozygous and *k* homozygous individuals under the SDR and FDR model are, respectively:

$$P_{SDR}(Pop) = C \times P_{SDR}(M_{He})^{j} \times P_{SDR}(M_{Ho})^{k} = C \times (2d)^{j} \times (1-2d)^{k}$$

$$P_{FDR}(Pop) = C \times P_{FDR}(M_{He})^{j} \times P_{FDR}(M_{Ho})^{k} = C \times (1-d)^{j} \times (d)^{k}$$

where C is a combinatory coefficient constant for the observed sample. Therefore,

$$LOD(Pop) = \frac{(2d)^{j} \times (1-2d)^{k}}{(1-d)^{j} \times (d)^{k}}$$

If *i* independent loci are analyzed, the probabilities of the observed sample of gametes occurring under the SDR [$P_{SDR}(Pop)$] or FDR [$P_{FDR}(Pop)$] models are the products of the probabilities of the observed sample at each locus

$$P_{SDR}(Pop) = \prod C_i \times P_{SDR}(M_{iHe})^{ji} \times P_{SDR}(M_{iHo})^{ki} = \prod C_i \times (2d_i)^{ji} \times (1-2d_i)^{ki}$$

$$P_{FDR}(Pop) = \prod C_i \times P_{FDR}(M_{iHe})^{ji} \times P_{FDR}(M_{iHo})^{ki} = \prod C_i \times (1-d_i)^{ji} \times (d_i)^{ki}$$

and therefore,

$$LOD(Pop) = \sum \frac{(2d_i)^{ji} \times (1-2d_i)^{ki}}{(1-d_i)^{ji} \times (d_i)^{ki}}$$

where $P(M_{iHe})$,$P(M_{iHo})$, *ji*, *ki* and d_i are, respectively, the probability of heterozygous individuals, probability of homozygous individuals, number of heterozygous individuals, number of homozygous individuals and distance to centromere for the locus *i*.

At the population level, LOD scores greater than 3 were considered to indicate that SDR was the mechanism involved in unreduced gamete formation, whereas LODs below -3 indicated

that FDR was the underlying mechanism. When LOD scores between -3 and 3 were obtained, we considered that the mechanism could not be significantly determined.

Studies to check the power of the method

We assessed the power of our method using simulated samples of diploid gametes arising from either the FDR or SDR mechanisms. From a theoretical infinite population with heterozygous and homozygous genotype frequencies directly defined by the considered locus-centromere distances ([$P_{FDR}(M_{He})=(1-d)$; $P_{FDR}(M_{Ho})=d$; $P_{SDR}(M_{He})=2d$; $P_{SDR}(M_{Ho})=(1-2d)$] as explained above), individual gametes with information for nine markers (the haploid number of chromosome in *Citrus*) were randomly generated. Then, the LOD values of these gametes were calculated as described above. We estimated the proportion of gametes with significant solutions at LOD3 (LOD value > 3 or < -3) and LOD2 (LOD value > 2 or < -2) when analysing 1–9 markers mapped at the same centromere distance, but in different chromosomes, and for distances ranging from 0 to 20 cM.

Gamete populations were also generated in order to estimate the theoretical number of hybrids that would need to be analyzed in order to obtain significant conclusions for a mechanism, depending on the number of markers used and the marker-centromere distances. From each theoretical population (FDR and SDR populations), 200 replicates of populations (with 1–100 gametes/population) were randomly generated. The generated population LODs were calculated as described above and, for each number of considered markers at a given centromere distance, we identified the minimum number of gametes needed in order to be able to reach a true significant conclusion for at least 99% of the generated populations (99% of replicates with LOD > 3 for SDR or LOD < -3 for FDR).

Results

Simulation results

From 1000 randomly selected gametes with nine independent markers (at the same distance from their respective centromere) from a theoretical SDR and FDR infinite population, we analyzed the percentage of replicates with significant LOD value (i.e., LOD3 and LOD2) at a given distance considering the data from 1–9 markers.

Curves corresponding to a significant true answer are shown in Figure 3.3. All curves display a vertical drop to 0, corresponding to the distance when the maximum theoretical LOD score (when all considered markers are in the most favourable combination for the model) is below the considered threshold. Compared with LOD3, the LOD2 threshold allows maintenance of the progressive decrease of the significant answer with increasing distance. As distance increases, more markers are needed to maintain a high level of significance.

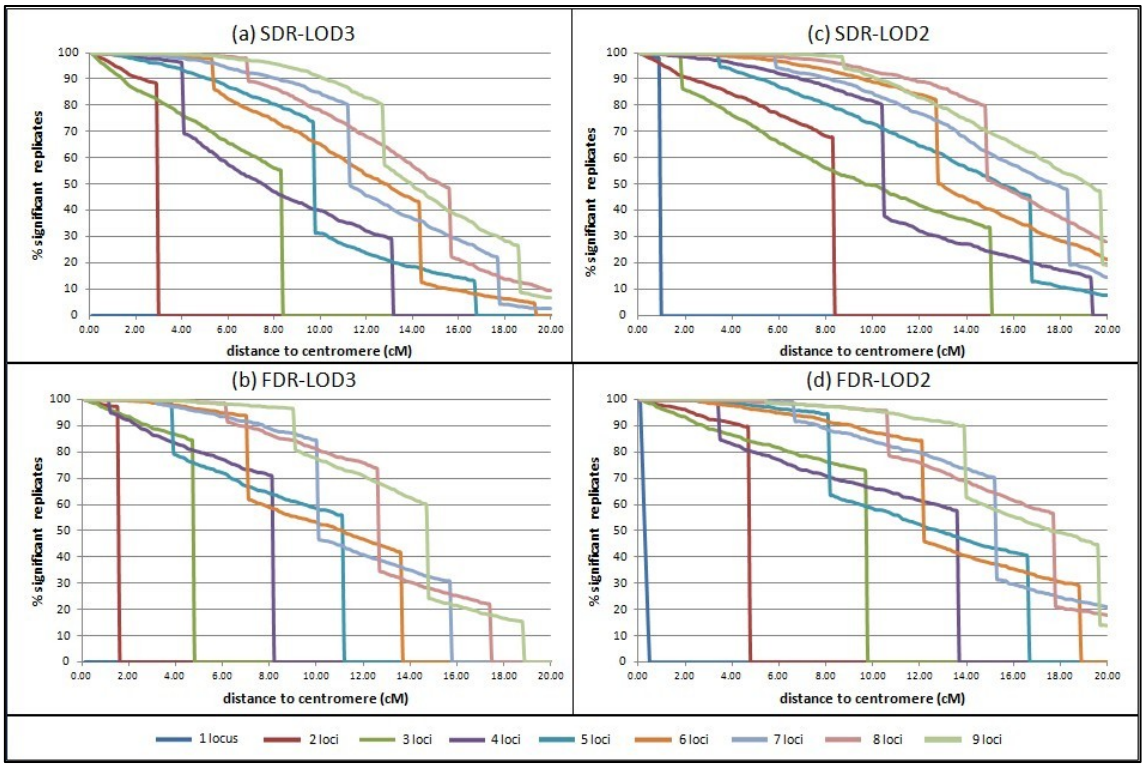

Figure 3.3. Percentage of replicates with significant LOD value considering a LOD3 for (a) theoretical SDR and (b) FDR populations, and considering a LOD2 for (c) SDR and (d) FDR populations.

At LOD3, the usefulness of only one marker is null for both the SDR (Figure 3.3a) and the FDR (Figure 3.3b) models at a very low marker distance from the centromere (0.1 cM). At 5 cM, at least five (for SDR) and six (for FDR) markers are necessary to maintain a 90% true significant identification of the mechanism. When all markers were at least 10 cM from centromeres, nine markers were necessary to provide a 90% true significant answer for the SDR population, but only 78% significant true answers were obtained with nine markers for a FDR population. At 15 cM and nine markers, the true identification rates fall to 44% and 24% for SDR and FDR, and, at 20 cM, to 6.6% and 0%, respectively.

If the LOD2 threshold is considered, a single marker was informative in the first cM interval for the SDR model (Figure 3.3c) but significant replicate number decreases very quickly for FDR (Figure 3.3d). At 5 cM, at least four and five markers were necessary to provide 90% of true significant identification for SDR and FDR populations, respectively. With all markers 10 cM from centromeres, at least eight markers were necessary to provide 90% true significant answers with an SDR or FDR population. For nine markers, the rate of true significant identification is improved for the SDR population at 15 cM and 20 cM (70% and 19%, respectively) as well as for the FDR population (59% and 14%, respectively) when compared with LOD3.

The rate of false identification (FDR significant conclusion [i.e., LOD<-3 or LOD<-2] for a SDR population, or reciprocally) is very low for both models (SDR or FDR), whatever the

centromere distance and the number of considered loci. At LOD3, it is under 0.1% for all conditions and it remains below 1% for the LOD2 threshold (Figure 3.S1).

At the population level (Figure 3.4), due to the probabilities of the 2*n* gamete genotypic structure under FDR and SDR models becoming similar as the distance to centromere rises, the number of hybrids needed to obtain significant conclusions for a mechanism increases as an exponential function and is more pronounced when analysing a single marker only.

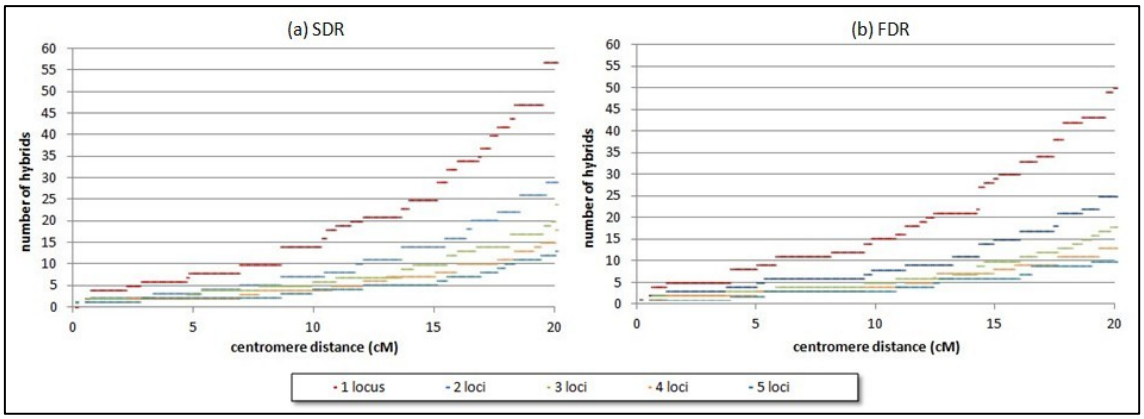

Figure 3.4. Number of hybrids needed to obtain significant conclusions for (a) SDR and (b) FDR mechanisms.

For a concrete locus-centromere distance, the number of hybrids (h_m) needed is related to the number of markers analyzed as: $h_m=h_1/m$, being h_1 the number of hybrids needed for one marker and *m*, the number of markers analyzed. For example, for a SDR population model, at 20 cM, 58 hybrids are necessary if analysing only one marker, 29 are necessary for two markers, and 20 are necessary for three markers. The number of hybrids needed to provide the same level of conclusive answer is slightly lower for FDR (50 hybrids for one marker at 20 cM). With these population sizes, no false mechanism identification occurred for the generated populations.

Inference of allelic configuration of triploid hybrids and corresponding 2n gametes

Assignment of allelic configuration in heterozygous triploid hybrids was performed using the MAC-PR method for SSR markers [(Esselink *et al*., 2004); Figure 3.S2] adapted for *Citrus* by Cuenca *et al*. (2011) and from relative allele signal for SNP markers as proposed by Cuenca *et al*. (2013a), based on competitive allele-specific PCR (Figure 3.S3). However, both these methods use a 1:1 dosage correction from the relative allele signals for heterozygous diploid parents (A_1:A_2, A_1:A_3 or A_3:A_4). Therefore, for markers displaying $A_1A_2 \times A_1A_3$ configuration in the parents, among the heterozygous triploid hybrids only the $A_1A_2A_2/A_1A_1A_2$ or $A_1A_3A_3/A_1A_1A_3$ configurations can be determined using these methods, while no direct allele dosage estimation

can be obtained for a triploid with A_2/A_3 heterozygosity without a reference for the relative A_2/A_3 allele signal. Similarly, for markers displaying the $A_1A_2 \times A_3A_4$ configuration, it is not possible to directly estimate allele dosage for the heterozygous triploid hybrids.

When the parental origin of a 2*n* gamete can be demonstrated using one marker, this allows the estimation of the relative signals of the alleles for another locus in the hybrid under consideration. As an example, consider two loci, A and B, with the parental configurations $A_1A_2 \times A_3A_3$ and $B_1B_2 \times B_1B_3$, which produce a triploid hybrid with $A_1A_2A_3$ and B_2B_3 allelic configurations. Configuration of the A locus unequivocally reveals a maternal origin of the 2*n* gamete. Based on the maternal origin of the unreduced gamete, the only possibility for the B locus is $B_2B_2B_3$. Thus, the allelic pattern for this triploid genotype can be used as a reference of the relative allele signal to infer allele dosage for the other triploid hybrids in the same progeny that display B_2/B_3 heterozygosity. A summary of triploid genotypes allowing inference of the 2*n* gamete genotype and origin, either directly or by inferring allele doses from diploid parents or reference triploid hybrids, is given in additional Table 3.S2. Loci with complete differentiation between the parents ($A_1A_2xA_3A_4$ or $A_1A_2xA_3A_3$) are by far the best configurations as they allow unequivocal identification of the 2*n* gamete parent and unambiguous determination of 2*n* gamete structure. When the parental origin of a 2*n* gamete has been determined by triploid patterns at other loci, the 2*n* gamete structure can be inferred for all triploid hybrids for the loci sharing a single allele between the two parents.

Identification of the unreduced gamete parental origin

Allelic patterns of the markers allowed unequivocal identification of the origin of the double dosage for each analyzed triploid hybrid. Female parents were the unreduced gamete producers leading to triploid hybrids for all studied parental combinations. No triploid hybrid arising from unreduced pollen was found. It was therefore possible to infer the maternal 2*n* gamete genotypes for all hybrids and loci.

Identification of the restitution mechanism at the individual level

LOD score testing the SFR/FDR hypothesis was estimated for each individual 2*n* gamete from its inferred genotype, as described in the Materials and Methods. Positive LODs were found for 482 hybrids of the 497 analyzed (Figure 3.5), suggesting a large global predominance of the SDR mechanism. The LOD distribution for clementine 2*n* gametes is displaced to higher values when compared with the distribution for ‘Fortune’ and other mandarin 2*n* gametes. Fifty-seven diploid gametes occur with LOD between 9 and 10, and these correspond mostly to the ‘Fina’ clementine progeny.

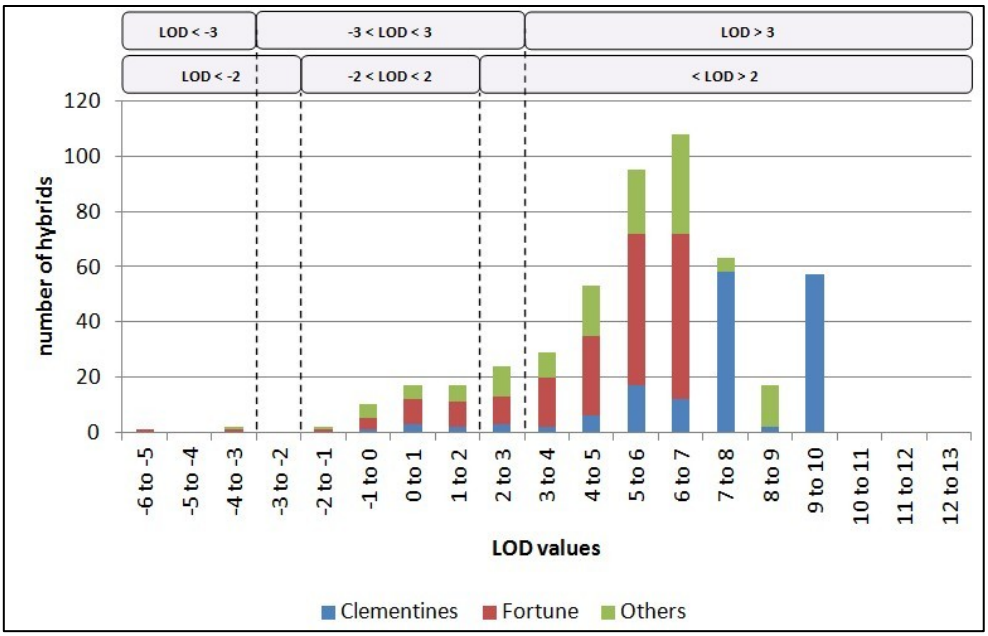

Figure 3.5. Frequency histogram of LOD values obtained for each individual 2*n* gamete, indicating those arising from clementines, 'Fortune' mandarin and other mandarins analyzed in this study.

When using LOD3 as the threshold, SDR was found to be the restitution mechanism underlying unreduced megagametophyte production for 424 (85.3%) of the analyzed triploid hybrids (Table 3.3). For one triploid hybrid arising from 'Ellendale' and two arising from 'Fortune' (0.6%), the FDR mechanism was implicated. The other 70 (14.1%) triploid hybrids did not give significant conclusions for either the SDR or FDR mechanisms. All unreduced gametes arising from 'Encore', 'Fallgo', 'Guillermina', 'Honey', 'Loretina' and 'Wilking' were identified as having an SDR origin, whereas for 33 unreduced gametes arising from 'Fortune' (16.7%) no significant conclusions were obtained (Table 3.3).

When using LOD2 as the threshold, the percentage of gametes with unidentified origins decreased to 9%. Gametes attributed to SDR increased to 90.1%, with significance achieved for an additional three clementine gametes, another ten from 'Fortune' and an extra 11 from other mandarins. No additional 2*n* gametes arising from FDR were identified.

Identification of the restitution mechanism at population level

At the population level, all LOD scores were greater than 3, even for small populations with fewer than five hybrids. Therefore, SDR was identified as the preeminent restitution mechanism producing 2*n* megagametophyte for all female parents analyzed (Table 3.3).

Table 3.3. LOD scores for progeny of 16 female parents analyzed at population level and individuals within each population originated by SDR, FDR, or with unidentified origin.

Group	Female parent	Nh	Nm	Cd average (cM)	LODs > +3		LODs -3 – +3		LODs < -3		Population LOD P_{SDR} / P_{FDR}
					number	(%)	number	(%)	number	(%)	
Clementine	'Bruno'	17	6	7.8	15	(88.2)	2	(11.8)	0	(0.0)	98.9
	'Clemenules'	23	5	4.4	22	(95.6)	1	(4.4)	0	(0.0)	143.3
	'Fina'	87	6	3.7	83	(95.4)	4	(4.6)	0	(0.0)	699.3
	'Guillermina'	14	6	8.1	14	(100.0)	0	(0.0)	0	(0.0)	91.1
	'Hernandina'	22	5	4.4	20	(90.9)	2	(9.1)	0	(0.0)	139.0
	'Loretina'	2	7	9.5	2	(100.0)	0	(0.0)	0	(0.0)	10.3
Mandarin	'Imperial'	24	5	5.0	23	(95.8)	1	(4.2)	0	(0.0)	138.5
Hybrid mandarin	'Ellendale'	69	5	9.1	50	(72.5)	18	(26.1)	1	(1.4)	282.7
	'Encore'	3	5	4.9	3	(100.0)	0	(0.0)	0	(0.0)	17.9
	'Fallglo'	3	5	3.7	3	(100.0)	0	(0.0)	0	(0.0)	21.6
	'Fortune'	197	5	6.8	162	(82.2)	33	(16.7)	2	(1.1)	933.0
	'Fortune' × 'Ellendale'	*58*	*5*	*6.8*	*54*	*(93.1)*	*4*	*(6.9)*	*0*	*(0.0)*	*310.5*
	'Fortune' × 'Minneola'	*35*	*4*	*5.2*	*28*	*(80.0)*	*6*	*(17.1)*	*1*	*(2.9)*	*145.7*
	'Fortune' × 'Murcott'	*67*	*5*	*6.8*	*53*	*(79.1)*	*14*	*(20.9)*	*0*	*(0.0)*	*326.2*
	'Fortune' × 'Willowleaf'	*37*	*4*	*6.6*	*28*	*(75.7)*	*8*	*(21.6)*	*1*	*(2.7)*	*150.6*
	'Honey'	1	4	6.1	1	(100.0)	0	(0.0)	0	(0.0)	5.1
	'Kiyomi'	21	5	6.3	20	(95.2)	1	(4.8)	0	(0.0)	162.9
	'Moncada'	8	4	10.3	4	(50.0)	4	(50.0)	0	(0.0)	22.1
	'Umatilla'	5	4	11.3	1	(20.0)	4	(80.0)	0	(0.0)	9.6
	'Wilking'	1	5	8.3	1	(100.0)	0	(0.0)	0	(0.0)	15.6

Nh: number of hybrids within each population (pop). Nm: number of markers analyzed over each population. Cd: Centromere distance

Discussion

A powerful maximum-likelihood method to compare FDR and SDR hypothesis at the individual and population level has been developed

In sexual polyploidization, polyploids are generated by the formation of unreduced diploid gametes. From the cytogenetic point of view, two types of meiotic nuclear restitution leading to 2*n* gamete formation are considered, FDR and SDR (Bretagnolle and Thompson, 1995; Ramsey and Schemske, 1998; Bastiaanssen *et* al., 1998; Dewitte *et al*., 2012; Brownfield and Kholer, 2011). The subsequent union of unreduced and reduced gametes leads to the formation of polyploids.

The identification of the mechanisms driving the formation of 2*n* gametes is complex. However, the use of cytological or marker analysis on polyploid progeny provide accurate or additional information on these mechanisms (Lim *et al.*, 2001; Crespel and Gudin, 2003; Dewitte *et al.*, 2012). Molecular cytological approaches have been used successfully, including the unequivocal identification of genomes and recombinant segments in the sexual polyploid progenies (Takahashi *et al.*, 1997; Karlov *et al.*, 1999; Lim *et al.*, 2001; Ramanna and Jacobsen, 2003; Barba-Gonzalez *et al.*, 2005). Molecular marker analysis is also useful for the identification of mechanisms underlying unreduced gamete formation, and different approaches based on population analysis have been developed previously. Several methods are based on the analysis of HR rates for randomly chosen unmapped markers (Vorsa and Rowland, 1997; Chen *et al.*, 2008a; Ferrante *et al.,* 2010). These methods require the analysis of a large set of molecular markers to encounter, by chance, the loci with HR lower than 50% that are only found under SDR (Park *et al*., 2007). However, when HR over 50% is observed for all loci, no

definitive conclusion can be reached without a prior knowledge of their location relative to a centromere. Significant FDR conclusions are therefore difficult to obtain with such non-mapped markers. Half-tetrad analysis (HTA; (Mendiburu and Peloquin, 1979)), based on multiple linked loci, is a powerful method for mapping centromeres or for determining the mode(s) of 2*n* gamete formation. Tavoletti *et al*. (1996) developed a multilocus maximum-likelihood method of HTA that permits the estimation of both the relative frequencies of FDR and SDR 2*n* gametes and the centromere location within a linkage group without relying on previously identified centromeric markers. These methods generally assume complete chiasma interference. The method proposed by Cuenca *et al.* (2011), based on the HR restitution curve along a linkage group, allows simultaneous identification of the restitution mechanism, raw centromere location, and comparison of several chromosome interference models. This approach is based on the analysis of genotype frequency in relatively large populations and provides global results of the preeminent mechanism; however, determination of the potential coexistence of the two mechanisms in the same progeny was not possible.

In this study, a maximum-likelihood approach based on marker HR with centromeric loci was developed and successfully applied both at the individual and population levels. Knowledge of marker-centromere distances greatly improves the statistical power of the comparison between the SDR and FDR hypotheses. For example, in this study, the restitution mechanism was identified in 'Fortune' as SDR at the population level with a LOD(SDR/FDR) of 933, whereas for the same population using 12 markers without information regarding marker-centromere distance, but with HR values under 50% (Cuenca *et al.,* 2011), the mechanism was identified as SDR with a LOD value of only 6.8. With the method proposed in the present paper, conclusions at the population level could therefore be obtained from smaller numbers of progeny and fewer markers than with non-located markers.

The theoretical limits of our method were assessed by the simulation of populations arising from FDR or SDR mechanisms. At the population level, considering that the independent markers used are at the same distance from their respective centromeres, the power of the statistical test was directly linked to the product of the number of markers and the number of individuals. That means that the efficiency would be the same for *n* individuals with *m* markers as for 2·*n* individuals with *m*/2 markers. Moreover, the necessary *n*·*m* genotyping points increase exponentially with increasing distance of the marker to the centromere. For example, to obtain a greater than 99% significant answer, it would be necessary for *n*·*m* to be greater than fifty-seven for markers at 20 cM, while an *n*·*m* value greater than eight and four would be sufficient for markers at 5 cM and 1 cM, respectively. The selection of markers as close as possible to their centromere is therefore a key element for successful analysis when low numbers of individuals and markers are used.

In the study of citrus 2*n* gamete progenies, significant results were obtained for all analyzed populations, even populations of fewer than five individuals.

One major improvement of our approach over existing methods is that it allows the identification of the restitution mechanism for each individual unreduced gamete. Simulation studies indicated that the proximity of markers to the centromeres is a key factor. With markers closer than 5 cM, five markers are sufficient to result in 95% significant answers, but significance diminishes to less than 78% and 0% for nine markers at 10 cM and 20 cM from their centromeres, respectively.

The importance of selecting markers very close to the centromere to obtain significant conclusions at the individual level is illustrated by the results of our citrus analysis. Indeed, a very high percentage of significant results at the individual level (95.4%) and with high LODs were obtained for the 'Fina' clementine progeny analyzed with markers closer to centromeres than the other progenies.

2*n* megagametophytes arising from SDR are the preeminent source of triploid occurrence in 2*x* × 2*x* hybrid populations using mandarin-like parents

In this study, the mechanism leading to triploid formation in 2*x* × 2*x* crosses was elucidated, both at individual and population level, for nineteen parental combinations involving, respectively, sixteen and eleven varieties as female and male parents.

All the 497 triploid hybrids analyzed originated from 2*n* megagametophytes and, therefore, no 2*n* pollen contributed to the production of triploids in our parental combinations. These results expand to a large range of genotypes the prior conclusion obtained from cytological studies (Esen and Soost, 1971, 1973) for 'Sukega' (*C. paradisi* × *C. sinensis*), 'Temple' (*C. reticulata* × *C. sinensis* and clementine (*C. clementina*), indicating that in such 2*x* × 2*x* crosses, triploid embryos were associated with pentaploid endosperm. However, the occurrence of triploids arising from 2*n* pollen at very low rates has been previously reported in studies using molecular markers for three selections of clementine ('Caffin', 'Commun' or 'SRA85' and 'Muskat'), and 'King' mandarin pollinated with *C. deliciosa* ('Tardivo di Ciaculi', 'Willow Leaf'), *C. reticulata* ('Hansen', 'Ananas'), *C. paradisi* ('Star Ruby') and *C. sinensis* ('Tarroco Rosso', 'Sanguinelli') (Luro *et al*., 2004) and for *C. sinensis* × *Poncirus trifoliata* hybridizations (Chen *et al*., 2008a).

When using the LOD3 threshold, SDR was identified as the restitution mechanism for 85.3% of the analyzed triploid hybrids, no significant conclusions were obtained for 14.1% of the hybrids, and 0.6% of the analyzed triploids were derived from FDR (one triploid hybrid from arising from 'Ellendale' and two arising from 'Fortune'). When the LOD2 threshold was considered, the percentage of individuals with unidentified origin decreased to 9% and SDR levels increased to 90.1%. Moreover, we conducted individual level analysis of previously studied 'Fortune' mandarin progeny (Cuenca *et al*., 2011) and the progeny arising from 'Fina' Aleza *et al*., 2012b), and we confirmed SDR at the individual level for most hybrids, which concurs with the global-level conclusions proposed in these two studies. In the current study, six

clementine genotypes were also analyzed to discover their unreduced gamete formation mechanism. Results indicate that SDR is the most probable mechanism in the clementine group, in agreement with previous conclusions of Luro *et al*. (2004) and Aleza *et al*. (2012b). For the other mandarin varieties, SDR was also the most probable mechanism at the individual level and, therefore, also at the population level. Taken together, our data and those of others suggest that SDR is the major mechanism underlying unreduced megagametophyte formation in most mandarin genotypes.

The mechanism leading to unreduced eggs or pollen was previously elucidated for several plant species (Ramanna and Jacobsen, 2003; De Storme and Geelen, 2013). Bretagnolle and Thompson (1995) identified that both FDR and SDR are responsible for 2*n* pollen formation, while SDR is more frequent in the formation of 2*n* eggs. In potatoes, 2*n* pollen arises predominantly by FDR (Mok and Peloquin, 1975), while 2*n* megagametophytes arise most frequently by SDR (Stelly and Peloquin, 1986), although SDR-FDR mixture in the formation of 2*n* eggs has been also found (Conicella *et al*., 1991). Bilateral sexual polyploidization can arise either from FDR and SDR in *Lilium* (Lim *et al*., 2004; Errico *et al*., 2005; Nadeem, *et al*., 2010) and alfalfa (Barcaccia *et al*., 2003). Moreover, Bretagnolle and Thompson (1995) described other several examples of single plant species where FDR and SDR may occur simultaneously, underlining the influence of genotype and environment on the expression of meiotic abnormality factors (Pécrix *et al*., 2011; Mason *et al*., 2011).

Implications for citrus triploid breeding

The genetic and phenotypic consequences of FDR and SDR gametes are highly divergent, and are of potential importance for breeding applications, due to the different parental heterozygosity rate that each mechanism transmit to the polyploid progeny (Errico *et al*., 2005; De Storme and Geelen, 2013).

Under FDR, the resulting 2*n* gametes are heterozygous from the centromere to the first crossover point, and hence the gametes retain most parental heterozygosity and epistatic interactions. With the SDR mechanism, the resulting 2n gametes are homozygous from the centromere to the first crossover point, but retain parental heterozygosity on the telomeric regions (Ramanna and Jacobsen, 2003). As a result, SDR-2*n* gametes confer a reduced level of heterozygosity than FDR-2*n* and show a corresponding greater loss of parental epistasis (Bretagnolle and Thompson, 1995; Dewitte *et al*., 2012).

If an SDR origin of 2*n* gametes is assumed for most mandarins, sexual polyploidization may lead to a reduced average of HR and, therefore, loss of epistatic interactions. Therefore, when compared with interploid crosses using doubled diploids (Aleza *et al*., 2012a, 2012c), the sexual polyploidization strategy should produce more polymorphic progeny by creating a larger number of new multilocus allelic combinations (David *et al*., 1995). This provides the opportunity

to select innovative products within the perspective of market segmentation as a commercial strategy.

Consequences of the SDR restitution mechanism would be clearly apparent for a characteristic controlled by a single gene. If the gene is heterozygous in the female parent, most unreduced gametes will be homozygous for that gene if it is located near the centromere, but gametes will be mostly be heterozygous for the gene if it is telomere-proximal (partial interference model; Cuenca *et al*., 2011). Recently, Cuenca *et al*. (2013b) analyzed the inheritance of resistance to *Alternaria* brown-spot fungal disease in triploid progenies arising from crosses between diploid parents. They demonstrated that the resistance was controlled as a recessive trait by a single locus located near a centromere (10.5 cM from the centromere of chromosome 3). If a susceptible female parent is heterozygous, the SDR mechanism leads to approximately 80% homozygous unreduced gametes, half of which present with two resistant alleles. As *Alternaria* resistance is a major selective trait when maternal heterozygous parents are used, sexual polyploidization is a more effective strategy than use of interploid crosses in this case, which results in only 16.7–22.5% of progeny being resistant. For dominant traits controlled by a single centromeric locus, interploid crosses should be more interesting than 2*x* × 2*x* crosses. For characters controlled by major loci more distant than 30 cM from the centromere, the efficiency of the two triploid breeding strategies would be relatively similar.

Supplementary information

Table 3.S1. Taxonomic information on parental genotypes used in this study

Group	Genotype	Swingle and Reece, 1967	Tanaka, 1977
clementine	'Bruno' 'Clemenules' 'Fina' 'Guillermina' 'Hernandina' 'Loretina'	*Citrus reticulata*	*Citrus clementina*
mandarin	'Imperial'		*C. reticulata*
	'Willowleaf'		*C. deliciosa*
hybrid mandarin	'Ellendale'	*C. reticulata* × *C. sinensis*	*C. reticulata* × *C. sinensis*
	'Encore'	*C. reticulata*	*C. nobilis* × *C. deliciosa*
	'Fallglo'		[*C. clementina* × (*C. paradisi* × *C. tangerina*)] × *C. temple*
	'Fortune'		*C. clementina* × *C. tangerina*
	'Honey'		*C. nobilis* × *C. deliciosa*
	'Kiyomi'	*C. reticulata* × *C. sinensis*	*C. unshiu* × *C. sinensis*
	'Minneola'	*C. reticulata*	*C. paradisi* × *C. tangerina*
	'Moncada'		*C. clementina* × (*C. unshiu* × *C. nobilis*)
	'Murcott'	*C. reticulata* × *C. sinensis*	(*C. reticulata* × *C. sinensis*)
	'N'15'	*C. reticulata*	*C. clementina* × (*C. unshiu* × *C. nobilis*)
	'Nadorcott'	-	(*C. reticulata* × *C. sinensis*) × (unknown)
	'Umatilla'	*C. reticulata* × *C. sinensis*	*C. unshiu* × *C. sinensis*
	'Simeto'	*C. reticulata*	*C. unshiu* × *C. deliciosa*
	'Wilking'		*C. deliciosa* × *C. nobilis*
pummelo	'Chandler'	*C. grandis*	*C. maxima*

Table 3.S2. Inferring of the 2*n* gamete parental producer depending on parental configuration and the raw genotype pattern observed.

Parental configuration	Raw genotype pattern observed	Allele doses estimation method	Inferred triploid genotype	Inferred 2*n* gamete configuration	Inferred parental origin	Inferred 2*n* gamete if maternal origin proved	Inferred 2*n* gamete if paternal origin proved
$A_1A_2 \times A_3A_4$	$A_1A_2A_3$	D	$A_1A_2A_3$	A_1A_2	M	A_1A_2	
	$A_1A_2A_4$	D	$A_1A_2A_4$	A_1A_2	M	A_1A_2	
	$A_1A_3A_4$	D	$A_1A_3A_4$	A_3A_4	P		A_3A_4
	$A_2A_3A_4$	D	$A_2A_3A_4$	A_3A_4	P		A_3A_4
	A_1A_3	TH	$A_1A_1A_3$	A_1A_1	M	A_1A_1	
		TH	$A_3A_3A_1$	A_3A_3	P		A_3A_3
	A_1A_4	TH	$A_1A_1A_4$	A_1A_1	M	A_1A_1	
		TH	$A_4A_4A_1$	A_4A_4	P		A_4A_4
	A_2A_3	TH	$A_2A_2A_3$	A_2A_2	M	A_2A_2	
		TH	$A_3A_3A_2$	A_3A_3	P		A_3A_3
	A_2A_4	TH	$A_2A_2A_4$	A_2A_2	M	A_2A_2	
		TH	$A_4A_4A_2$	A_4A_4	P		A_4A_4
$A_1A_2 \times A_1A_1$	A_1	D	$A_1A_1A_1$	A_1A_1	NI	A_1A_1	A_1A_1
	A_1A_2	DP	$A_1A_2A_2$	A_2A_2	M	A_2A_2	
		DP	$A_1A_1A_2$	NI	NI	**A_1A_2**	**A_1A_1**
$A_1A_2 \times A_3A_3$	$A_1A_2A_3$	D	$A_1A_2A_3$	A_1A_2	M	A_1A_2	
	A_1A_3	TH	$A_1A_1A_3$	A_1A_1	M	A_1A_1	
		TH	$A_3A_3A_1$	A_3A_3	P		A_3A_3
	A_2A_3	TH	$A_2A_2A_3$	A_2A_2	M	A_2A_2	
		TH	$A_3A_3A_2$	A_3A_3	P		A_3A_3
$A_1A_2 \times A_1A_3$	$A_1A_2A_3$	D	$A_1A_2A_3$	A_1A_2	M	A_1A_2	
	A_1	D	$A_1A_1A_1$	A_1A_1	NI	A_1A_1	A_1A_1
	A_1A_2	DP	$A_1A_1A_2$	NI	NI	**A_1A_2**	**A_1A_1**
		DP	$A_1A_2A_2$	A_2A_2	M	A_2A_2	
	A_1A_3	DP	$A_1A_1A_3$	NI	NI	**A_1A_1**	**A_1A_3**
		DP	$A_1A_3A_3$	A_3A_3	P		A_3A_3
	A_2A_3	TH	$A_2A_2A_3$	A_2A_2	M	A_2A_2	
		TH	$A_2A_3A_3$	A_3A_3	P		A_3A_3

D: direct visualization of allelic constitution; DP: inference of allelic doses based on diploid parent reference; TH: inference of allelic doses based on at least one reference triploid hybrid; NI: not identified. Bold letters indicate gamete structures inferred when parental origin has been proved from other markers.

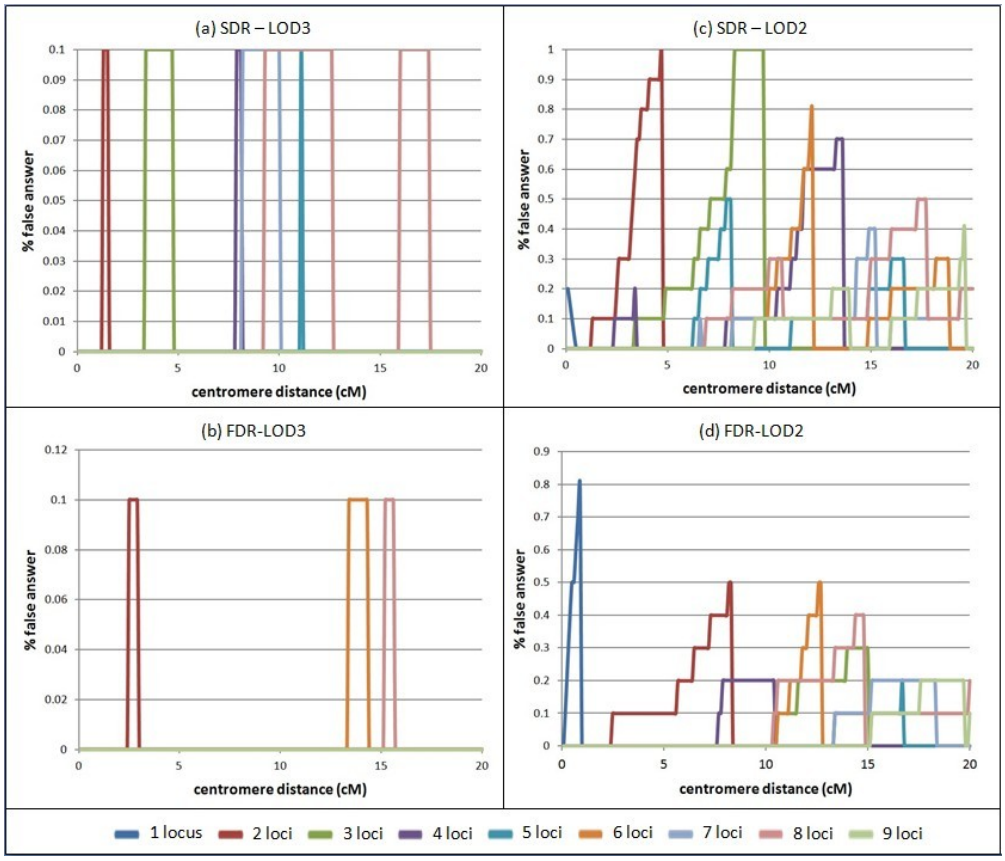

Figure 3.S1. Percentage of significant replicates giving false answer considering a LOD3 for populations arising from (a) SDR and (b) FDR, and considering LOD2 for populations arising from (c) SDR and (d) FDR.

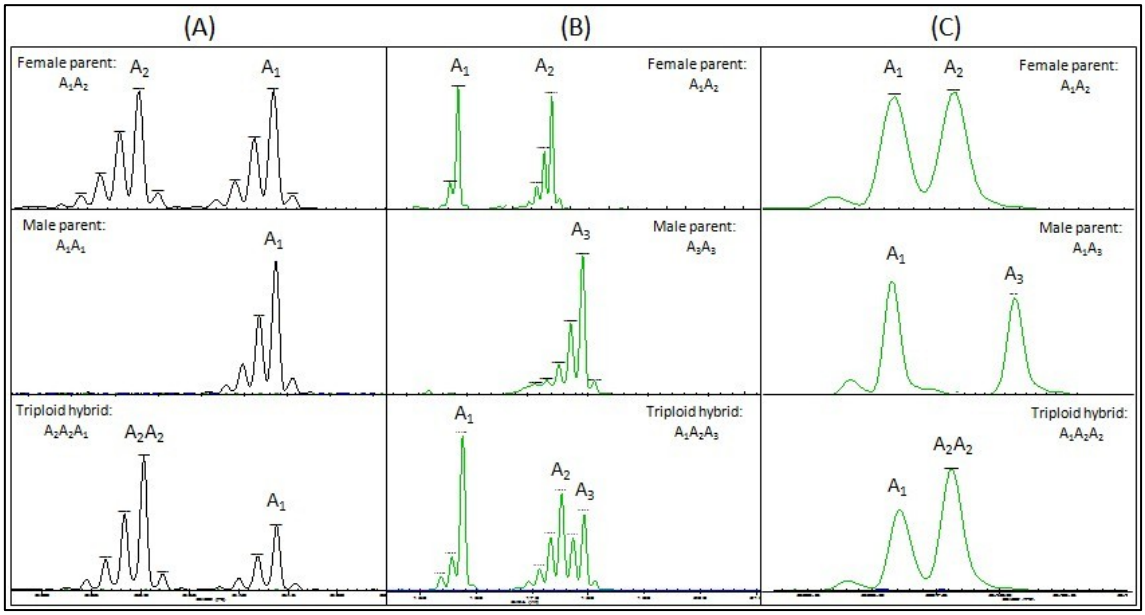

Figura 3.S2. Assignment of allelic configuration in heterozygous triploid hybrids for SSR markers showing (A) $A_1A_2 \times A_1A_1$ (B) $A_1A_2 \times A_3A_3$ and (C) $A_1A_2 \times A_1A_3$ parental configurations.

.

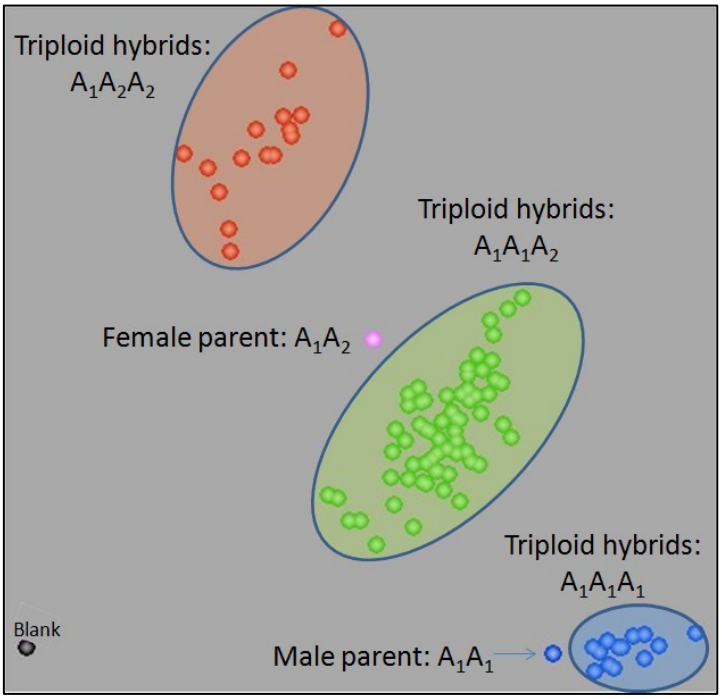

Figura 3.S3. Assignment of allelic configuration in heterozygous triploid hybrids for SNP markers with A_1A_2 × A_1A_1 parental configurations.

CHAPTER 4

Genetically based location from triploid populations and gene ontology of a 3.3-Mb genome region linked to Alternaria brown spot resistance in citrus reveal clusters of resistance genes

José Cuenca, Pablo Aleza, Antonio Vicent, Dominique Brunel, Patrick Ollitrault and Luis Navarro

PLoS ONE (2013), 8(10): e76755. doi: 10.1371/journal.pone.0076755

Abstract

Genetic analysis of phenotypical traits and marker-trait association in polyploid species is generally considered as a challenge. In the present work, different approaches were combined taking advantage of the particular genetic structures of 2n gametes resulting from second division restitution (SDR) to map a genome region linked to Alternaria brown spot (ABS) resistance in triploid citrus progeny. ABS in citrus is a serious disease caused by the tangerine pathotype of the fungus *Alternaria alternata*. This pathogen produces ACT-toxin, which induces necrotic lesions on fruit and young leaves, defoliation and fruit drop in susceptible genotypes. It is a strong concern for triploid breeding programs aiming to produce seedless mandarin cultivars. The monolocus dominant inheritance of susceptibility, proposed on the basis of diploid population studies, was corroborated in triploid progeny. Bulk segregant analysis coupled with genome scan using a large set of genetically mapped SNP markers and targeted genetic mapping by half tetrad analysis, using SSR and SNP markers, allowed locating a 3.3 Mb genomic region linked to ABS resistance near the centromere of chromosome III. Clusters of resistance genes were identified by gene ontology analysis of this genomic region. Some of these genes are good candidates to control the dominant susceptibility to the ACT-toxin. SSR and SNP markers were developed for efficient early marker-assisted selection of ABS resistant hybrids.

Introduction

Alternaria brown spot (ABS) is a serious disease that induces necrotic lesions on fruit and young leaves, defoliation and fruit drop in susceptible citrus genotypes (Akimitsu *et al.*, 2003). The disease was first observed in Australia in 1903 on the 'Emperor' mandarin (Pegg, 1966), and was subsequently detected in citrus-growing regions in America, the Mediterranean Basin, South Africa, Iran and China (Timmer *et al.*, 2003; Golmohammadi *et al.*, 2006; Wang *et al.*, 2010). In Spain, the disease was first detected in 1998 (Vicent *et al.*, 2000), and it is currently widespread in all citrus-growing areas, affecting mainly 'Fortune' and 'Nova' mandarin hybrids. The disease is caused by the tangerine pathotype of the fungus *Alternaria alternata* (Fr.) Keissl., which carries a gene cluster (*ACTT*) located in a small (<2.0 Mb) conditionally dispensable chromosome responsible for ACT-toxin biosynthesis (Ajiro *et al.*, 2010). This host-specific toxin is released during the germination of conidia, rapidly affecting the plasma membrane integrity of susceptible host cells (Kohmoto *et al.*, 1993). There is also indirect evidence suggesting the presence of toxin receptors in susceptible citrus genotypes (Tsuge *et al.*, 2012). In addition, recent studies indicate that the mitigation of reactive oxygen species (ROS) produced by the host plants is essential for pathogenicity (Yang and Chung, 2012). The pathogen sporulates on affected tissues, and conidia are disseminated by air currents and rain splash. Warm temperatures and prolonged wetness on the tree are required for infection. The incubation period is very short, and lesions are visible just 1 or 2 days after infection is initiated, due to the rapid effects of the ACT-toxin (Canihos *et al.*, 1999). The disease causes severe epidemics in humid areas, as well as in semi-arid regions, due to its environmental flexibility (Timmer *et al.*, 2003). Currently, ABS control is primarily based on the application of fungicides. Sprays must be scheduled to protect susceptible organs during the critical periods for infection. Depending on the climate of the region and the susceptibility of the cultivar, between four and ten fungicide sprays per year are needed to produce quality fruit for the fresh market (Bhatia *et al.*, 2003; Peres and Timmer, 2006; Vicent *et al.*, 2007).

Despite this large number of sprays, disease control is not always satisfactory, and cultivation of susceptible cultivars such as the 'Fortune' mandarin has declined significantly in Spain during recent years. In addition, systematic application of fungicides for ABS control over many years may create environmental problems and public health concerns (Vicent *et al.*, 2009). Moreover, in the context of the changing global climate, plant breeding is especially focused on improving resistance to biotic and abiotic stresses.

Several studies have been carried out to determine the resistance or susceptibility to ABS in citrus genotypes (Hutton and Mayers, 1988; Goes *et al.*, 2001; Vicent *et al.*, 2004; Dalkilic *et al.*, 2005; Reis *et al.*, 2007; de Souza *et al.*, 2009). Although there are some discrepancies among the results of these studies, resistance is clearly present in clementine (*Citrus clementina* Hort. ex Tan), 'Willowleaf' (*C. deliciosa* Ten) and satsuma (*C. unshiu* Mark) mandarins. Susceptibility has also been well established for 'Dancy' (*C. tangerina* Hort. ex Tan) and 'Fortune' (supposed *C. clementina* Hort. ex Tan × *C. tangerina* Hort. ex Tan) mandarins;

'Orlando', 'Minneola' and 'Nova' tangelos (mandarin × grapefruit hybrids); and the 'Murcott' tangor (supposed mandarin × sweet-orange hybrid). Other cultivars such as the 'Ellendale' tangor and some sweet oranges and grapefruits have been characterised as sensitive or resistant by different authors. From diploid progeny analysis, it has been proposed that inheritance of ABS resistance in citrus is controlled by a single recessive allele (Dalkilic *et al.*, 2005; Gulsen *et al.*, 2010). Resistance to the strawberry and pear *Alternaria* pathotypes, which produce toxins structurally analogous to those of the tangerine pathotype, as well as resistance to the apple pathotype, is controlled in the same way, by a single recessive allele (Tsuge *et al.*, 2012). Therefore, resistant cultivars are considered to be recessive homozygous for this locus, whereas susceptible cultivars could be heterozygous or homozygous dominant.

Diploidy is the general rule in *Citrus* and related genera; however, polyploidy manipulation is currently widely used in triploid citrus breeding programs aimed at developing new seedless mandarin cultivars (Ollitrault *et al.*, 2008). Many of these breeding programs (Mourao Fo *et al.*, 1996; McCollum, 2007; Aleza *et al.*, 2010a, 2010b; Cuenca *et al.*, 2010; Grosser *et al.*, 2010; Aleza *et al.*, 2012c, 2012d; Froelicher *et al.*, 2012; Navarro *et al.*, 2012) use ABS-susceptible cultivars as parents, due to their utility with regard to other important traits (fruit quality, maturing period, production) and particular reproductive biology (monoembryony, high rate of triploid production). The inheritance and efficient selection of resistance to ABS is therefore of central importance to triploid mandarin breeding projects.

Genetic analysis of phenotypical traits and marker-trait association in polyploid species is generally considered as a challenge due to complex segregation, dosage effects and potential non Mendelian inheritance associated with epigenetic variations.

The main factor affecting trait inheritance in triploid families is the strategy used for triploid breeding (Ollitrault *et al.*, 2008), with significant differences between the sexual polyploidization approach (2*x* × 2*x* crosses with unreduced -2*n*- gamete formation) and interploid crosses (2*x* × 4*x* or 4*x* × 2*x*). Indeed, the choice of strategy affects the transmission of parental heterozygosity to the diploid gamete.

In sexual polyploidization, two factors affect the transmission of parental heterozygosity to the offspring: the mechanism of 2*n* gamete formation (i.e., first-division restitution [FDR] or second division restitution [SDR]) and the genetic distance from the locus of interest to the centromere (Douches and Quiros, 1988). Therefore, Half-Tetrad Analysis (HTA) based on 2*n* gametes is an efficient means of genetic mapping (Mendiburu and Peloquin, 1979; Douches and Quiros, 1987; Tavoletti *et al.*, 1996). In 2*x* × 2*x* citrus crosses, the diploid (unreduced) gamete is transmitted by the female parent (Esen and Soost, 1971; 1973). SDR has been proposed for diploid megagametophyte development in clementines (Luro *et al.*, 2004) and 'Fortune' mandarins (Cuenca *et al.*, 2011), whereas FDR has been reported in sweet oranges (Chen *et al.*, 2008a). Recent studies have revealed that SDR is the main mechanism involved in unreduced gamete formation in the majority of citrus cultivars (Aleza *et al.*, 2012b). For

interploid crosses, most of the tetraploid parents used in citrus breeding arise from chromosome doubling in nucellar cells of apomictic diploid parents (Aleza *et al.*, 2011). Because mandarins are one of the ancestral species of cultivated citrus (García-Lor *et al.*, 2012), doubled-diploid mandarins should be considered as autotetraploid, and tetrasomic inheritance should be expected (Kamiri *et al.*, 2011). In such a situation, the frequency of diploid gametes that receive a locus in heterozygosis from the tetraploid parent varies between 0.55 and 0.66, depending on the double-reduction frequency (Marsden *et al.*, 1987).

In addition to the particular transmission of chromosome fragments and parental heterozygosity, the phenotypic trait inheritance in polyploids can be affected by dosage effects (Guo *et al.*, 1996; 2004) and even by neoregulation of gene expression due to epigenomic reformatting (Liu and Wendel, 2003; Osborn *et al.*, 2003), eventually leading to non- Mendelian segregation. Moreover, polyploidy induces morphological variations in leaves and fruits (Starrantino, 1992; Otto and Whitton, 2000; Allario *et al.*, 2011) that should affect fungus colonization. In this context, no data have yet been published regarding the inheritance of ABS resistance in triploid progenies.

Due to its direct applicability in marker-assisted selection, the identification of molecular markers linked to phenotypic variation, e.g., related to disease resistance, is a key step in most breeding programs. Bulked segregant analysis [BSA; (Michelmore *et al.*, 1991)] can be used to identify molecular markers in a genomic region associated with a specific phenotype rapidly. This method is based on linkage disequilibrium between the gene and linked markers in segregating progeny, and the genetic linkage between markers and the causal gene is determined by differences in marker-allele frequencies between resistant and susceptible bulks. For characters controlled by one or a few genes, BSA is an effective technique for detecting alleles linked to phenotypes in a large sample of progeny at a relatively low cost, where the only requirement is that the genotyping technique and molecular markers utilised provide quantitative measurements of allelic frequencies (Liu *et al.*, 2012). This approach should be optimised by coupling BSA with a high-throughput genotyping method using markers covering the whole genome. Genome-wide association studies of pooled DNA samples have been valuable tools in the fast, scalable and economical identification of candidate single nucleotide polymorphisms (SNPs) associated with a phenotype (Brauer *et al.*, 2006; Akhunov *et al.*, 2009; Hyten *et al.*, 2009; Becker *et al.*, 2011; Ricci *et al.*, 2011; Szelinger *et al.*, 2011; Swinnen *et al.*, 2012; Trick *et al.*, 2012). In citrus, very large SNP resources are becoming available from extensive citrus sequencing projects (Terol *et al.*, 2007, 2008; Gmitter *et al.*, 2012; Ollitrault *et al.*, 2012b); meanwhile, new technologies have been developed for very rapidly genotyping large numbers of SNPs in DNA samples. One such technology is the GoldenGate assay from Illumina™ (Fan *et al.*, 2003; Hyten *et al.*, 2009; Yan *et al.*, 2010), which proved useful in citrus by allowing mapping of 677 SNP markers onto the clementine's consensus map (Ollitrault *et al.*, 2012a).

The objectives of this study were (i) to confirm Mendelian monolocus inheritance of ABS resistance in triploid progenies, by analysing the segregation of resistance in different interploid crosses, and to confirm the dominance of susceptibility by analysing segregation of resistance in progeny produced by sexual polyploidization and interploid hybridization; (ii) to locate the chromosome region associated with the ABS resistance using a genome scan assay coupled with BSA, followed by targeted genetic mapping by HTA in triploid progenies arising from 2n gametes; and (iii) to identify candidate resistance genes in the located region, taking advantage of the recently released reference whole genome sequence of *C. clementina* (Wu *et al.*, 2013). A more applied objective was identification of molecular markers for marker-assisted selection (MAS) in citrus breeding programs.

Material and methods

Plant material

Seven hundred and fourteen triploid hybrids arising from four 2*x* × 2*x* crosses, three 2*x* × 4*x* crosses and their parents were evaluated for field and *in vitro* infection by *A. alternata*. Parental genotypes included in the citrus germplasm bank and hybrids were grown at the 'Instituto Valenciano de Investigaciones Agrarias' (I.V.I.A.) orchards in Moncada, Valencia, Spain. The plantings were very dense, with conditions very favourable for the development of ABS infection.

Information about parental accessions, their origin, ABS phenotype and references are shown in Table 4.1. The genetic configuration of the *ABSr* locus ('A', dominant susceptible allele; 'a', recessive resistant allele) for each parental accession (also given in the table) has been deduced from information about ABS resistance/susceptibility of diploid genotypes, their pedigree and segregation data at the diploid level, under the hypothesis of single locus inheritance. The tetraploids 'Nova' and 'Orlando' resulted from chromosome stock doubling of the Nova and Orlando diploids, respectively (Aleza *et al.*, 2011). The diploid lines are considered to be 'Aa' at the *ABSr* locus; therefore, the genotypes of the two tetraploid parents should be 'AAaa'.

Table 4.1. Parental genotypes used in this study, phenotypic information on ABS resistance and deduced *ABSr* locus genotyping.

Genotype	Origin	Phenotype	Reference	*ABSr* locus genotype
'Fortune'	*C. clementina* × *C. tangerina*	S	(Vicent *et al.*, 2000) (Reis *et al.*, 2007)	Aa
'Minneola'	*C. paradisi* × *C. tangerina*	S	(Solel, 1991) (Solel and Kimchi, 1997) (Peever *et al.*, 1999) (Elena, 2006)	AA
'Orlando'	*C. paradisi* × *C. tangerina*	S	(Solel and Kimchi, 1997) (Gulsen *et al.*, 2010)	2x: Aa 4x: AAaa
'Nova'	*C. clementina* × (*C. paradisi* × *C. tangerina*)	S	(Dalkilic *et al.*, 2005)	2x: Aa 4x: AAaa
'Murcott'	(unknown)	S	(Solel and Kimchi, 1997) (Dalkilic *et al.*, 2005)	Aa
'Willowleaf'	*C. deliciosa*	R	(Solel and Kimchi, 1997)	aa
'Clemenules' 'Fina'	*C. clementina*	R	(Kohmoto *et al.*, 1991) (Solel and Kimchi, 1997) (Elena, 2006) (Reis *et al.*, 2007) (Gulsen *et al.*, 2010)	aa
'Nadorcott'	'Murcott' × unknown	R	Our unpublished data	aa

(S) Susceptible phenotype; (R) Resistant phenotype; (A) Susceptible allele; (a) Resistant alelle

Three of the 2*x* × 2*x* crosses share 'Fortune' as the female parent, with 'Willowleaf' mandarin (93 hybrids), 'Minneola' tangelo (127 hybrids) and 'Murcott' (148 hybrids) as male parents. The other 2*x* × 2*x* cross was 'Fina' × 'Nadorcott' (50 hybrids). Details on procedures for establishing the triploid populations from 2*x* × 2*x* crosses by embryo rescue and triploid selection by flow cytometry can be found in Aleza *et al.* (2010b).

Two of the 2*x* × 4*x* crosses share 'Orlando 4x' as the male parent, with 'Clemenules' (180 hybrids) and 'Fortune' (116 hybrids) as female parents. The other 2*x* × 4*x* population was 'Clemenules' × 'Nova 4x' (100 hybrids). Information about procedures for establishing the 2*x* × 4*x* populations can be found in Aleza *et al.* (2012c).

Moreover, five additional triploid populations arising from 2*x* × 2*x* and 2*x* × 4*x* crosses (114 hybrids) were also evaluated for ABS resistance to extend the experiments to other genetic backgrounds. Due to the relatively low number of triploid hybrids within each population, the resultant data have been included as supplementary material (Table 4.S1).

Evaluation of ABS resistance

Previous studies in diploid genotypes have shown a range of susceptibility level among citrus germplasm, but suggest that immune response could be controlled by a single recessive allele (Dalkilic *et al*., 2005; Gulsen *et al*., 2010). In the present study, genotypes have been considered as resistant if no symptoms have been observed neither under field evaluations nor leaf inoculations. Therefore, in this study as in the previous ones at diploid level (Dalkilic *et al*., 2005; Reis *et* al., 2007; De Souza *et* al., 2009; Gulsen *et al*., 2010), the resistant phenotype corresponds to immune symptom.

Field evaluation: Symptoms of *A. alternata* were evaluated for all genotypes on trees grown at the I.V.I.A. orchards in spring, when young leaves are more susceptible to ABS and environmental conditions are highly favourable for infection (Vicent *et al.*, 2009). Presence or absence of ABS symptoms on the leaves was recorded in a qualitative manner. For each tree, observations were carried out over three consecutive years (2010, 2011 and 2012).

In vitro inoculation of detached leaves:

Inoculum production: A virulent single-spore isolate of *A. alternata* (IVIA-A005) isolated from an infected 'Fortune' fruit from Valencia (Spain) was used for inoculations. Abundant conidia were obtained by a method adapted from (Everts and Lacy, 1996). The isolate was grown on potato dextrose agar (PDA) plates at 25°C in darkness for 8–10 days, illuminated with fluorescent lamps (Philips TLD 18W/33) at 25°C for 8 h to initiate conidiophore formation, and then placed in the dark at 18°C for 12 h. Conidial suspensions were prepared by pouring sterile water over the colonies and gently rubbing the surface with a sterile glass rod. The suspension was filtered through two layers of cheesecloth, and the spore concentration was adjusted to 10^5 conidia·ml^{-1} with a haemocytometer. Suspensions with conidial germination lower than 90% were discarded.

Leaf inoculations: Bioassays were performed immediately after leaf harvest. Young leaves (about 50% developed) were inoculated with 10^5 conidia·ml^{-1} (Kohmoto *et al.*, 1991). This suspension was sprayed over both upper and lower surfaces of each leaflet, using five leaves per genotype. Controls were inoculated by spraying sterile distilled water. Leaves were incubated in a moist chamber in the dark at 27°C, and the results were evaluated 48 h after inoculation. In susceptible genotypes, leaf symptoms appear during the second day after inoculations and very clear necrosis induced by the ACT-toxin can be observed after 48h (Figure 4.1).

A genotype was considered resistant when no symptoms of ABS were observed in any leaf, whereas presence of infection was recorded when a clear symptom of ABS was observed in any leaf. The inoculations were repeated when there was doubt regarding interpretation. The complete experiments were carried out twice during spring of 2010 and twice during spring of 2011.

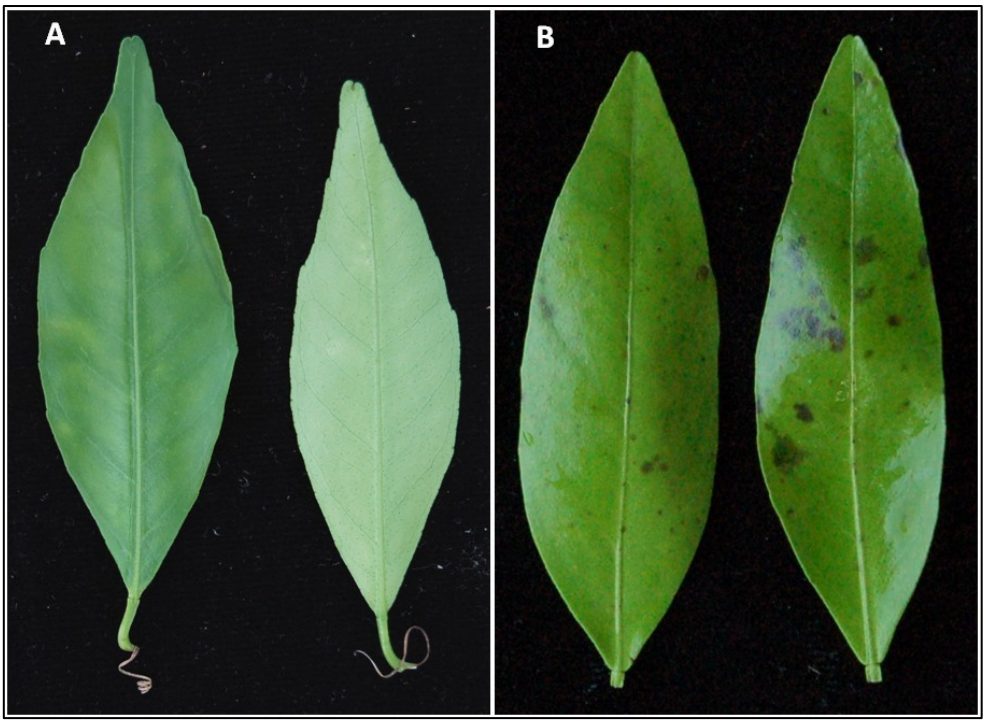

Figure 4.1. Leaves of resistant genotype 'Willowleaf' mandarin (A) and susceptible genotype 'Fortune' mandarin (B) showing ABS symptoms 48h after inoculation with a suspension of 10^5 conidia·ml^{-1}.

Two triploid populations derived from 2*x* × 2*x* crosses ['Fortune' ('Aa') × 'Minneola' ('AA') and 'Fina' ('aa') × 'Nadorcott' ('aa')] and three triploid populations derived from 2*x* × 4*x* crosses ['Clemenules' ('aa') × 'Orlando 4x' ('AAaa'), 'Fortune' ('Aa') × 'Orlando 4x' ('AAaa') and 'Clemenules' ('aa') × 'Nova 4x' ('AAaa')] were phenotyped to compare the expected and observed proportions of resistant and susceptible genotypes, to confirm the monolocus inheritance and dominance of the ABS susceptibility. For the 'Aa' × 'AA' cross, all segregation progeny are expected to be susceptible to ABS ('AAA', 'AAa' or 'Aaa'), whereas for the 'aa' × 'aa' cross, all segregation progeny are expected to be ABS resistant ('aaa'). In case of the 'aa' × 'AAaa' and 'Aa' × 'AAaa' crosses, the resistant and susceptible proportions depend on heterozygosity restitution (HR) from the tetraploid parent to the progeny, which varies between 0.55 and 0.66 depending on the double-reduction frequency (Marsden *et al.*, 1987). Therefore, in these cases, the resistant proportions are expected to be between 0.1667 and 0.225 for the 'aa' × 'AAaa' cross and between 0.0833 and 0.1125 for the 'Aa' × 'AAaa' cross (Table 4.2). χ^2 tests were conducted on the observed and expected frequencies. In cases of diploid × tetraploid crosses, where expected frequencies are included in an interval (according to the double-reduction frequency), if the observed value was found to be out of the interval, the observed value was compared with the closest value flanking this interval.

Table 4.2. Expected proportions of ABS locus allelic configuration (AAA, AAa, Aaa or aaa) for each population evaluated.

POPULATION	SUSCEPTIBLE				RESISTANT
	AAA	AAa	Aaa	TOTAL SUSCEPTIBLE	aaa
'Fortune' × 'Minneola' (Aa × AA)	$\frac{(1-HR)}{2}$	HR	$\frac{(1-HR)}{2}$	*1*	*0*
'Fina' × 'Nadorcott' (aa × aa)	-	-	-	0	1
'Clemenules' × 'Orlando 4x' (aa × AAaa)	-	0.1667 – 0.225	0.55 – 0.66	0.775 – 0.833	0.1667 – 0.225
'Fortune' × 'Orlando 4x' (Aa × AAaa)	0.0833 – 0.1125	0.3875 – 0.4167	0.3875 – 0.4167	0.8875 – 0.9167	0.0833 – 0.1125
'Clemenules' × 'Nova 4x' (aa × AAaa)	-	0.1667 – 0.225	0.55 – 0.66	0.775 – 0.833	0.1667 – 0.225

HR: heterozygosity restitution

Estimation of the locus-centromere genetic distance under the hypothesis of monolocus inheritance

Two segregating triploid progeny derived from crosses between 'Fortune' ('Aa') as the female parent and 'Willowleaf' ('aa') and 'Murcott' ('Aa') as male parents have been used to estimate the locus-centromere distance. Because SDR is the mechanism leading to unreduced gamete formation in 'Fortune' (Cuenca *et al.*, 2011), the maternal HR frequency varies between 0 at the centromere to 0.66 if a model of no chromosome interference is assumed. However, Cuenca *et al.*, (2011) demonstrated that the Cx(Co)^4 model assuming partial chromosome interference (Zhao and Speed, 1998b) was better adapted to the observed HR in Fortune 2*n* gametes.

The functions for estimating the frequency of diploid gametes that would be heterozygous for a given locus according to its distance from the centromere can be easily modified to estimate the expected genotypic frequency within resultant triploid progeny and even the expected segregation of phenotypic traits with monolocus inheritance. Considering that ABS resistance is a recessive trait controlled by a single locus, susceptible triploid genotypes may have 'AAA', 'AAa' or 'Aaa' allele configurations, whereas resistant triploid genotypes should present only the 'aaa' configuration for this locus (Table 4.3). Therefore, the frequency of resistant genotypes within each population is informative for HR estimation, and therefore for determination of the locus-centromere distance. The relation between centromere distance and percentage of resistant hybrids in controlled progeny have been represented (Figure 4.2) for the two models of crosses corresponding to the 'Fortune' × 'Willowleaf' and 'Fortune' × 'Murcott' crosses ('Aa' × 'aa' and 'Aa' × 'Aa', respectively) under two models of chromosome interference (no interference and partial interference). It should be noted that under the Cx(Co)^4 model of partial chromosome interference, the frequencies of resistant hybrids under 20% and 10% for the 'Aa' × 'aa' and 'Aa' × 'Aa' crosses, respectively, can correspond to two different distances from the centromere.

Table 4.3. Expected susceptible and resistant proportions for 'Fortune' ('Aa') × 'Willowleaf' ('aa') and 'Fortune' ('Aa') × 'Murcott' ('Aa') populations.

POPULATION	SUSCEPTIBLE				RESISTANT
	AAA	AAa	Aaa	TOTAL SUSCEPTIBLE	aaa
'Fortune' × 'Willowleaf' (Aa × aa)	-	$\frac{(1-HR)}{2}$	HR	$\frac{(1+HR)}{2}$	$\frac{(1-HR)}{2}$
'Fortune' × 'Murcott' (Aa × Aa)	$\frac{(1-HR)}{4}$	$\frac{(1-HR)}{4}+\frac{HR}{2}=\frac{(1+HR)}{4}$	$\frac{(1-HR)}{4}+\frac{HR}{2}=\frac{(1+HR)}{4}$	$\frac{(3+HR)}{4}$	$\frac{(1-HR)}{4}$

HR: heterozygosity restitution

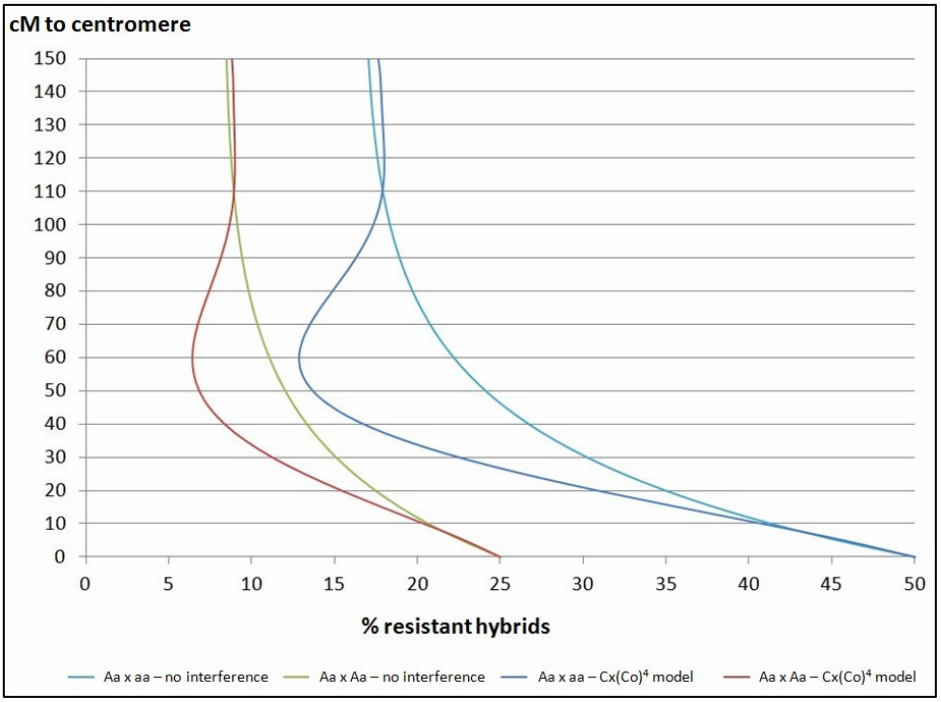

Figure 4.2. Locus-centromere distance estimated from the proportion of resistant hybrids observed in 'Aa' × 'aa' and 'Aa' × 'Aa' crosses under a model of no chromosome interference and the $Cx(Co)^4$ model of partial chromosome interference.

Bulk segregant analysis coupled with genome scan

BSA (Michelmore *et al.*, 1991) has been used to identify genomic regions linked to ABS resistance. To simplify the analysis with only one parental segregation of ABS resistance, the 'Aa' × 'aa' population (rather than 'Aa' × 'Aa') was selected. Triploid hybrids from the 'Fortune' ('Aa') × 'Willowleaf' ('aa') population yielding conclusive phenotypes (resistant or susceptible) in both field and *in vitro* evaluations were selected for this purpose. Genomic DNA of triploid hybrids and their parents was isolated using the Plant DNeasy kit from Qiagen, Inc. (Valencia, CA, USA), following the manufacturer's protocol. DNA concentrations were estimated with PicoGreen® and adjusted to 30 ng/µl. Four resistant and four susceptible DNA bulks were established by mixing DNA from five resistant or susceptible hybrids. Each bulk and the parents were genotyped using an Illumina GoldenGate™ array platform, which contains 1536 SNP

markers (Ollitrault *et al.*, 2012b). Six hundred and seventy-seven of these SNPs are mapped in the clementine's reference genetic map (Ollitrault *et al.*, 2012a).

For the mapped markers that were heterozygous in the 'Fortune' genotype, we estimated the relative allele signal in each bulk by allelic composition measurement, called the "B allele frequency" (BAF) by Illumina™ (Fan *et al.*, 2003), using the Illumina® GenomeStudio 2009. The BAF parameter varies between 0 and 1 and is related to the proportion of the B allele versus A+B (SNP genotyping in GoldenGate™ array is diallelic). For pooled samples, this parameter provides useful information on the BAF in the bulk.

ANOVA were performed using BAF information, and the significance of the differentiation between the resistant and susceptible bulks was tested by the F statistic. The pattern of this F parameter along the genome allowed identification of genomic regions with high probability of association with phenotype variation.

Individual genotyping and mapping of the ABS resistance gene

Ninety-three triploid hybrids for the 'Fortune' × 'Willowleaf' population and their diploid parents were genotyped using available SSR and SNP markers already mapped (Ollitrault *et al.*, 2012a) in the interval identified by the BSA analysis or developed from the clementine genomic sequence as described below.

New SSR and SNP marker development: We have taken advantage of the recent release of the reference citrus genome sequence (haploid Clementine genome publicly available at http:://www.phytozome.net/clementine) by the International Citrus Genomics Consortium (ICGC) to develop new markers in the genomic region surrounding the SNPs identified by BSA genome scan as linked to ABS resistance. Microsatellites motifs were searched using Sputnik software (http://espressosoftware.com/sputnik/) and new SSR markers were developed and tested for useful polymorphisms. Moreover, 4.47 kb corresponding to four DNA fragments within this region were sequenced in 'Fortune' and 'Willowleaf' to find SNPs that could be heterozygous in 'Fortune' and homozygous in 'Willowleaf' mandarin (information on location of the corresponding sequences on the haploid Clementine reference genome and primers used to amplify these DNA fragments is given in Table 4.S2).

SSR analyses: Polymerase chain reactions (PCRs) were performed with wellRED oligonucleotides (Sigma-Aldrich®, St Louis, MO, USA) using the following protocol: Mastercycler ep Gradient S (Eppendorf Scientific Inc., Westbury, NY, USA); reaction volume, 15 µl; 0.8 U Taq polymerase (Fermentas®, Burlington, VT, USA); reaction buffer: 750 mM Tris-HCl (pH 9), 50 mM KCl, 200 mM $(NH_4)_2SO_4$, 0.001% bovine serum albumin, 0.1 mM of each dNTP, 5 mM $MgCl_2$, 3 mM of each primer, 30 ng DNA. The PCR program was as follows: 94°C for 5 min; 40 cycles of 30 s at 94°C, 1 min at 55°C and 30 s at 72ºC; final elongation 10 min at 72ºC. Separation was carried out by capillary gel electrophoresis (CEQ 8000 Genetic Analysis

System; Beckman Coulter Inc., Fullerton, CA, USA). Data collection and analysis were carried out using the GenomeLab GeXP (Beckman Coulter Inc.) version 10.0 software.

SNP analyses: SNP genotyping was performed by Kbioscience® services, using the KASPar technique. Detailed explanation of specific conditions and reactives can be found in Cuppen (2007).

Assignment of allelic configuration in heterozygous triploid hybrids was carried out using the MAC-PR method for SSR markers (Esselink *et al.*, 2004), or using relative allele signal as proposed by Cuenca *et al.* (2013a) for SNPs genotyped by the KASPar (KBioscience®, UK) technique. Maternal HRs within the triploid progeny were used for *de novo* mapping of the markers in relation to the centromere position, using the $Cx(Co)^4$ model for SDR with partial interference (Zhao and Speed, 1998b; Cuenca *et al.*, 2011).

Allelic phase of linked marker loci was inferred from the preferential association at the population level between the phenotype (resistant/susceptible) and the maternal alleles. Marker alleles linked with susceptibility were codified as 'a' alleles, and those linked with resistance as 'b' alleles. The global coherence of this phase attribution was checked by performing a correlation (Pearson's coefficient) from an individual/loci matrix with values of 1, 0.5, and 0 for the 'aa', 'ab', and 'bb' genotypes, respectively. These correlation values were also used to determine the locations of the various markers in the relative chromosome arms (i.e., on either side of the centromere) in the *de novo* mapping process.

The relative position of the *ABSr* locus and markers were analyzed by performing a multiple correspondence analysis (MCA), considering markers as individuals and the various 2*n* gametes as variables. From the previous matrix, we established the qualitative matrix for the factorial analysis by grouping 1 and 0.5 as the same modality (presence of the 'a' allele linked with the dominant susceptibility allele in 'Fortune') and considering the absence of the 'a' allele as the other modality. XLSAT was used to calculate the Pearson's correlation coefficient and to perform the MCA.

Gene ontology

All genes encountered within the genomic region between the two markers flanking the estimated location of the *ABSr* locus were searched in the clementine whole genome assembly delivered by the ICGC and publicly available at http:://www.phytozome.net/clementine. The corresponding annotation data were then processed with Blast2GO (Conesa *et al.*, 2005) to provide a global description of the cellular components and biological processes of the genes identified in this genome region.

Results

Segregation of ABS resistance in various triploid progeny arising from sexual polyploidization and interploid crosses

Field and in vitro evaluation of ABS resistance

Symptoms of ABS were evaluated for all parental accessions and hybrids both from visual inspection of the trees grown at orchards and by *in vitro* inoculations with a conidial suspension of the pathogen. Results obtained for parental genotypes were according to those cited in the literature: 'Fortune', 'Minneola', 'Murcott' and 'Orlando' exhibited symptoms of ABS both in the field and *in vitro* in all evaluations; 'Clemenules', 'Fina', 'Willowleaf' and 'Nadorcott' did not exhibit any ABS symptoms on their leaves at any time. Triploid hybrids derived from various evaluated crosses were susceptible (exhibiting typical ABS symptoms) or resistant at proportions depending on the progeny evaluated. No resistant genotypes were found within the 'Fortune' × 'Minneola' population, whereas all triploid hybrids from the cross between two resistant genotypes ('Fina' × 'Nadorcott') were resistant to ABS. Total concordance between field and *in vitro* evaluations was observed for all evaluated populations with the exception of 'Clemenules' × 'Orlando 4x', where it was over 97% (Table 4.4).

Table 4.4. Results of field and *in vitro* phenotyping for Alternaria brown spot, showing the number of resistant hybrids within each population and the concordance between both types of evaluation.

	'Fortune' × 'Willoleaf'	'Fortune' × 'Murcott'	'Fortune' × 'Minneola'	'Clemenules' × 'Orlando 4x'	'Clemenules' × 'Nova 4x'	'Fortune' × 'Orlando 4x'	'Fina' × 'Nadorcott'
Nmber of hybrids evaluated	93	148	127	180	100	116	50
Resistant hybrids by field evaluation	37 (39.78%)	26 (17.57%)	0 (0%)	46 (25.55%)	16 (16%)	12 (10.34%)	50 (100%)
Resistant hybrids by in vitro evaluation	37 (39.78%)	26 (17.57%)	0 (0%)	41 (22.78%)	16 (16%)	12 (10.34%)	50 (100%)
Field-in vitro concordance (%)	93/93 (100%)	148/148 (100%)	127/127 (100%)	175/180 (97.22%)	100/100 (100%)	116/116 (100%)	50/50 (100%)
Consensus hybrids evaluated	93	148	127	175	100	116	50
Consensus resistant hybrids	37/93 (39.78%)	26/148 (17.57%)	0/127 (0%)	41/175 (23.43%)	16/100 (16%)	12/116 (10.34%)	50/50 (100%)

Inheritance of ABS resistance

Observed resistant proportions within the 'Fortune' ('Aa') × 'Minneola' ('AA'), 'Clemenules' ('aa') × 'Orlando 4x' ('AAaa'), 'Clemenules' ('aa') × 'Nova 4x' ('AAaa'), 'Fortune' ('aa') × 'Orlando 4x' ('AAaa') and 'Fina' ('aa') × 'Nadorcott' ('aa') populations are shown in Table 4.5. As expected, no resistant genotypes were observed within the 'Fortune' × 'Minneola' triploid population, whereas no susceptible ones were observed within the 'Fina' × 'Nadorcott' population. Regarding the interploid crosses, 41/175 (23.43%) and 16/100 (16%) triploid hybrids were phenotyped as resistant within the 'Clemenules' × 'Orlando 4x' and 'Clemenules' × 'Nova 4x' populations, respectively. These values are not significantly different (χ^2=0.087, p-value=0.769 and χ^2=0.032, p-value=0.857, respectively) to the closest value of the theoretical interval (16.67–22.5%) under the hypothesis of single locus recessive inheritance of resistance in an 'aa' × 'AAaa' cross. In the same way, the observed proportion of resistant hybrids in 'Fortune' × 'Orlando 4x' is within the theoretical interval under the same hypothesis for an 'Aa' × 'AAaa' cross.

Table 4.5. Expected and observed frequencies of Alternaria brown spot resistant hybrids under the hypothesis of single dominant inheritance within each population and significances of χ^2 conformity tests.

	'Fortune' × 'Minneola' ('Aa' × 'AA')	**'Clemenules' × 'Orlando 4x' ('aa' × 'AAaa')**	**'Clemenules' × 'Nova 4x' ('aa' × 'AAaa')**	**'Fortune' × 'Orlando 4x' ('Aa' × 'AAaa')**	**'Fina' × 'Nadorcott' ('aa' × 'aa')**
Number of evaluated hybrids	127	175	100	116	50
Expected resistant proportion (%)	0%	16.67% - 22.5%	16.67% - 22.5%	8.33% - 11.25%	100%
Observed resistant proportion (%)	0%	23.43%	16.00%	10.34%	100%
χ2 test; p-value	**NS**	**0.087; 0.769 (NS)**	**0.032; 0.857 (NS)**	**WTI**	**NS**

WTI: within theoretical interval; NS: no significantly different than theoretical segregation or identical.

These results confirmed the single dominant inheritance of the ABS susceptibility in triploid populations. Moreover, results of the five additional triploid populations evaluated for ABS resistance (Table 4.S1) also confirm the single recessive inheritance of ABS resistance.

Estimation of the genetic distance of the ABS resistance locus (*ABSr*) to the centromere

For the two triploid populations arising from 2*x* × 2*x* crosses ('Fortune' × 'Willowleaf' and 'Fortune' × 'Murcott'), the proportions of resistant and susceptible hybrids are related to the *ABSr* locus-centromere distance. To estimate the locus-centromere distance, we used a simple cross model in which only the 2*n* gametes segregate for the *ABSr* locus ('Fortune' × 'Willowleaf'). This allows avoiding any eventual bias associated with distorted segregation from the male parent.

The proportion of resistant hybrids in the 'Fortune' × 'Willowleaf' population was 39.78%, corresponding to an HR estimation value of 0.2043, assuming that 'aa' and 'AA' 2*n* gametes were equally represented. Only one value for centromere distance is associated with the observed proportion of resistant hybrids when the functions presented in Material and Methods are applied (Figure 4.2). Moreover, no interference and partial chromosome interference models gave very similar estimates for the centromere distance, which has been estimated to be 10.5 cM.

With such an *ABSr* locus-centromere distance, the expected proportion of the resistant genotype in 'Fortune' × 'Murcott' progeny ('Aa' × 'Aa') should be 19.9%. The observed value (17.57%) is not significantly different (χ^2=0.501), confirming the proximity of the *ABSr* locus to a centromere.

Bulk segregant analysis coupled with genome scan

BSA over the 'Fortune' ('Aa') × 'Willowleaf' ('aa') population has been used to identify a genomic region linked to the ABS resistance gene. Four resistant and four susceptible bulks were genotyped for 1536 SNP markers using a GoldenGate™ array platform. Of these, 429 SNP markers were heterozygous for the 'Fortune' mandarin and were used to perform ANOVA analyses over relative allele signal for each bulk; significance of the differentiation between the resistant and susceptible bulks was tested by the F statistic. A graphical example for the CiC3248-06 and CiC6243-03 markers, which differentiate resistant and susceptible bulks, is shown in Figure 4.3.

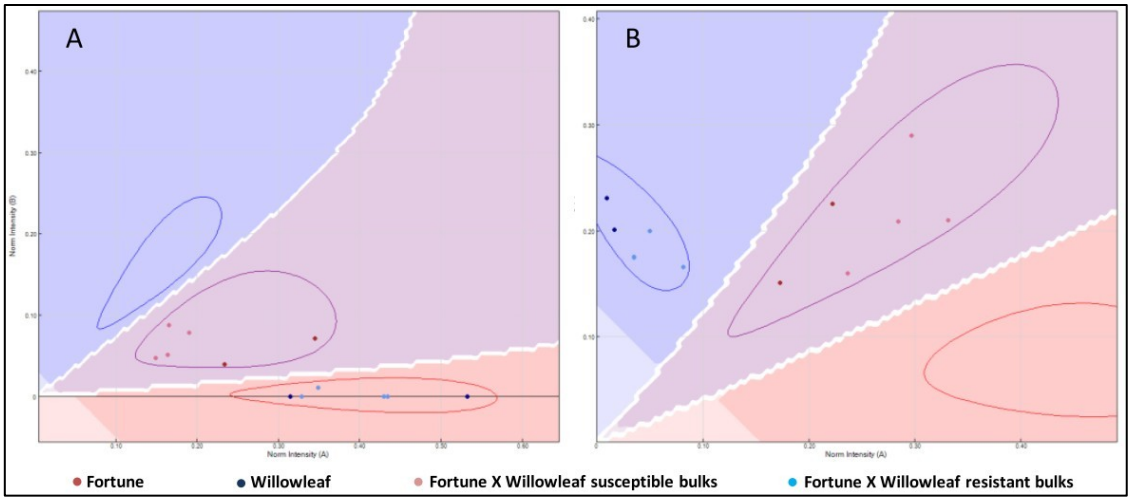

Figure 4.3. Plot showing Bulked Segregant Analysis results for the CiC3248-06 (A) and CiC6243-03 (B) markers, distinguishing between susceptible and resistant genotypes and bulks.

The pattern of this F parameter along the nine linkage groups of the clementine's genetic map (Ollitrault *et al.*, 2012a) led us to discard most genomic regions (Figure 4.S1) and allowed identification of a region containing numerous markers with a high probability (>99%) of association with phenotype variation, located on chromosome III (Figure 4.4). This region includes 25 significant SNP markers within an interval of 13.1 cM between markers CiC4831-03 (at 84.66 cM) and CiC1875-01 (at 97.76 cM) on the clementine's map. The maximum F value within this region is attained by marker CiC4681-02, located at 92.78 cM (F=2055). The genomic region between these two markers contains around 15 Mb. No significant marker clusters were found in any other area of the genome.

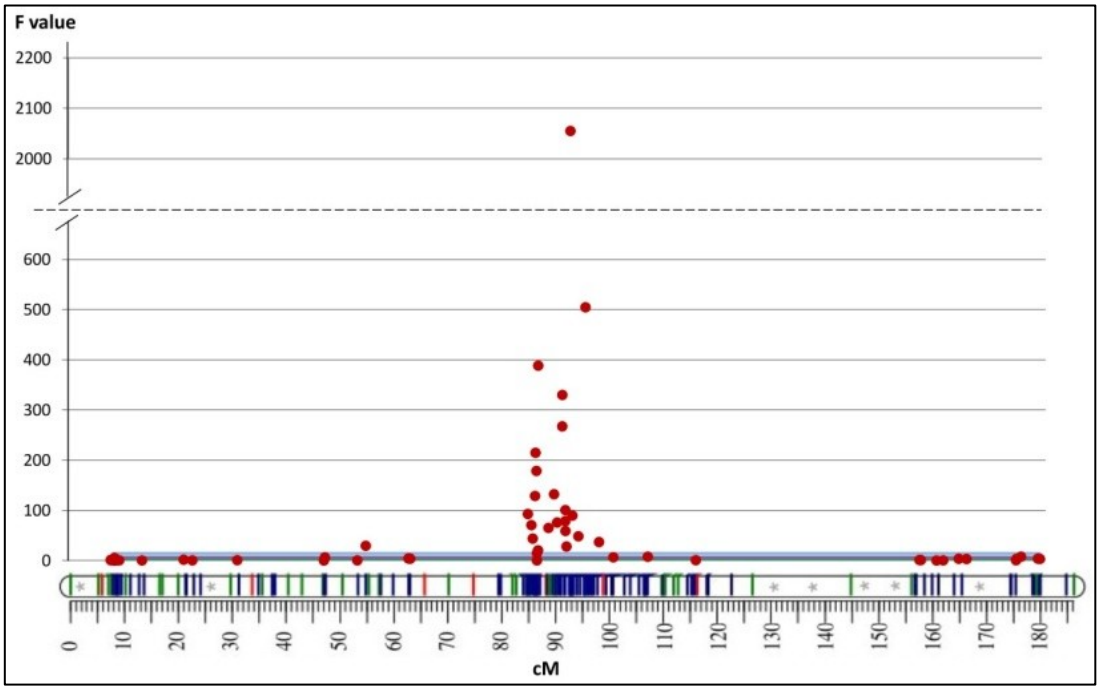

Figure 4.4. Pattern of F statistic from ANOVA along chromosome III (the linkage group map under the F value graph is taken from the genetic map of clementine (Ollitrault *et al.*, 2012a). The blue line indicates the least significant value for F at p<0.01.

Genetic mapping of the genomic region surrounding the ABS locus

Among the SNP markers with significant linkage to ABS resistance, five displayed the most convenient allelic conformation in the parents (heterozygous in 'Fortune' and homozygous in 'Willowleaf') for genetic mapping by individual genotyping of the 'Fortune' × 'Willowleaf' progeny. One SSR marker included in this segment in the clementine's genetic map (Ollitrault *et al.*, 2012a) also displayed useful allelic polymorphisms between parents.

To develop additional markers with useful allelic conformation, the genomic region (from www.phytozome.net) surrounding the 25 significant SNPs was scanned to find new microsatellites and develop new SSR markers. Among 42 SSRs tested, four new SSR markers

provided useful polymorphisms. Moreover, 4.47 kb (Table S2) within this region in 'Fortune' and 'Willowleaf' and two SNPs heterozygous in 'Fortune' and homozygous in 'Willowleaf' were sequenced. More detailed information on all markers used in this study is available as supplementary material in Table 4.S3 and Table 4.S4.

Next, five mapped SNP markers (Ollitrault *et al.*, 2012a), one mapped SSR marker [CX0038: (Chen *et al.*, 2008a)], and six newly developed markers (four SSRs and two SNP markers) were used to genotype all 93 triploid hybrids of the 'Fortune' × 'Willowleaf' population. Because the male parent was homozygous or different from the female parent at each selected locus, the genetic structure of the diploid female gamete (Table 4.S5) was deduced from the triploid hybrid genotyping [see (Cuenca *et al.*, 2011) for details], and the marker HRs were estimated.

We took advantage of the direct link between HR in 2*n* gametes and the locus-centromere distance for *de novo* mapping of genetic markers in relation to the centromere position, using the $Cx(Co)^4$ model for SDR with partial interference (Cuenca *et al.*, 2011). No recombination was observed between the centromere and the CiC1229-05 and CiC6116-04 markers. The markers next closest to centromere were SNP-ALT1 and SNP-ALT2 (with the same HTA data), 0.54 cM away; the next closest marker was CX0038 (2.7 cM). To determine whether this marker was located at one side or the other of the centromere, we checked its correlation with the markers in distal positions on the draft map (at this step, CiC1229-05 and CiC6116-04 on one side, and SNP-ALT1 and SNP-ALT2 on the other side). Because the lower correlation was for SNP-ALT1 and SNP-ALT2, CX0038 was positioned on the opposite chromosome arm. The same process was applied at each subsequent step of marker addition to the map, proceeding according to increasing distance from the centromere. The order of the mapped markers in the *de novo* map was the same as in the map of clementine (Ollitrault *et al.*, 2012a), and new (non-mapped) markers maintained the expected order of the assembled sequence available at www.phytozome.net (Figure 4.5a,b). However, the estimated genetic distances were higher than those on the clementine's map, suggesting that the recombination rate in this genomic region during the production of the 2*n* gamete was higher in 'Fortune' than in clementine. A logically important modification of the slope of the physical distances according to genetic map is observed in the centromeric region (with lower recombination by physical distance unit).

No recombination was observed between the AT21 marker and the *ABSr* locus. The two flanking markers (TTC8 and CiC3248-06) were found at 3.77 and 1.71 cM, respectively, from the *ABSr* locus, delimiting a 3.3 Mb genome region. This position of the *ABSr* locus was checked by an MCA based on a qualitative matrix (see Material and Methods). Most of the matrix diversity was represented in the first axis (72.9%; Figure 4.5c), where the order of markers and the relative position of the *ABSr* locus was identical to the *de novo* mapping, based on HR.

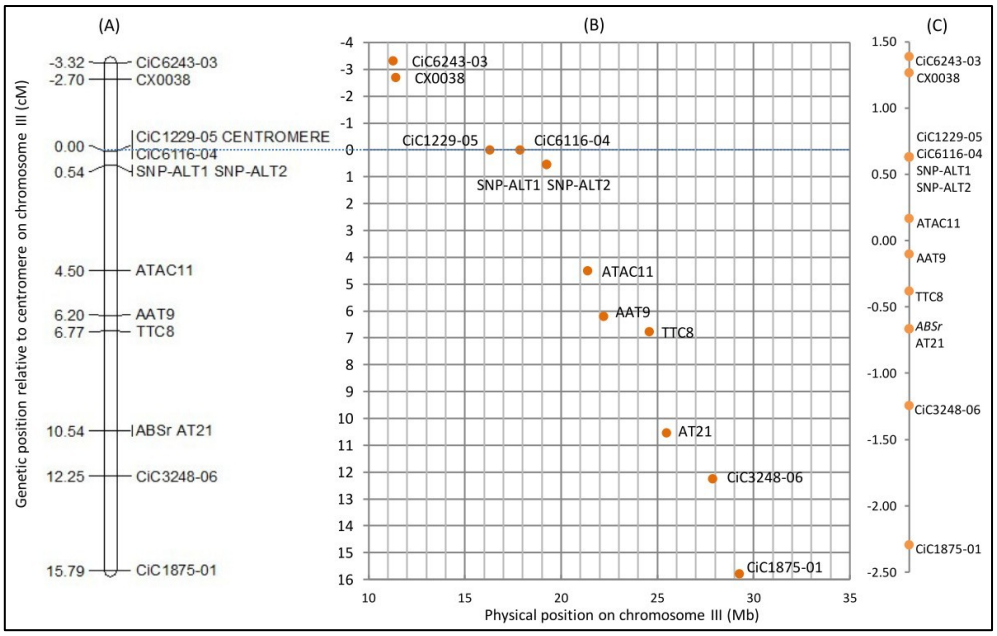

Figure 4.5. Order and location of markers and *ABSr* locus. (A) De novo genetic mapping (cM) of markers and the *ABSr* locus on chromosome III relative to the centromere by half-tetrad analysis, (B) relation between genetic and physical location in the clementine's reference genome (www.phytozome.net/clementine), and (C) representation of the markers on the first axis of the multiple correspondence analysis.

Gene annotations around the *ABSr* locus

The assembled sequence (www.phytozome.org) of the region of chromosome III between the two markers flanking the *ABSr* locus (TTC8 and CiC3248-06) was examined for gene annotations. The results revealed several disease resistance genes along and at the extremes of the analyzed region, so the analysis was extended 1.5 Mb down from the TTC8 marker and 1.7 Mb up from the CiC3248-06 marker. Ninety-five genes annotated as homologous to disease resistance genes were found within the corresponding 6.5 Mb region. A genome-wide analysis of disease resistance gene homologs revealed that 17% of them are located within this region on chromosome III.

Within the 3.3-Mb region defined by the two flanking markers, 177 annotated genes were found (Table 4.S6). Gene ontology (GO) analysis of biological processes revealed that 69.1% of these genes are involved in metabolic processes and 21.9% are related to response or cell death (Figure 4.6a). GO also indicated that 25% of the genes are intrinsic to the membrane (Figure 4.6b), which is the target of the ACT-toxin produced by the tangerine pathotype of *A. alternata*.

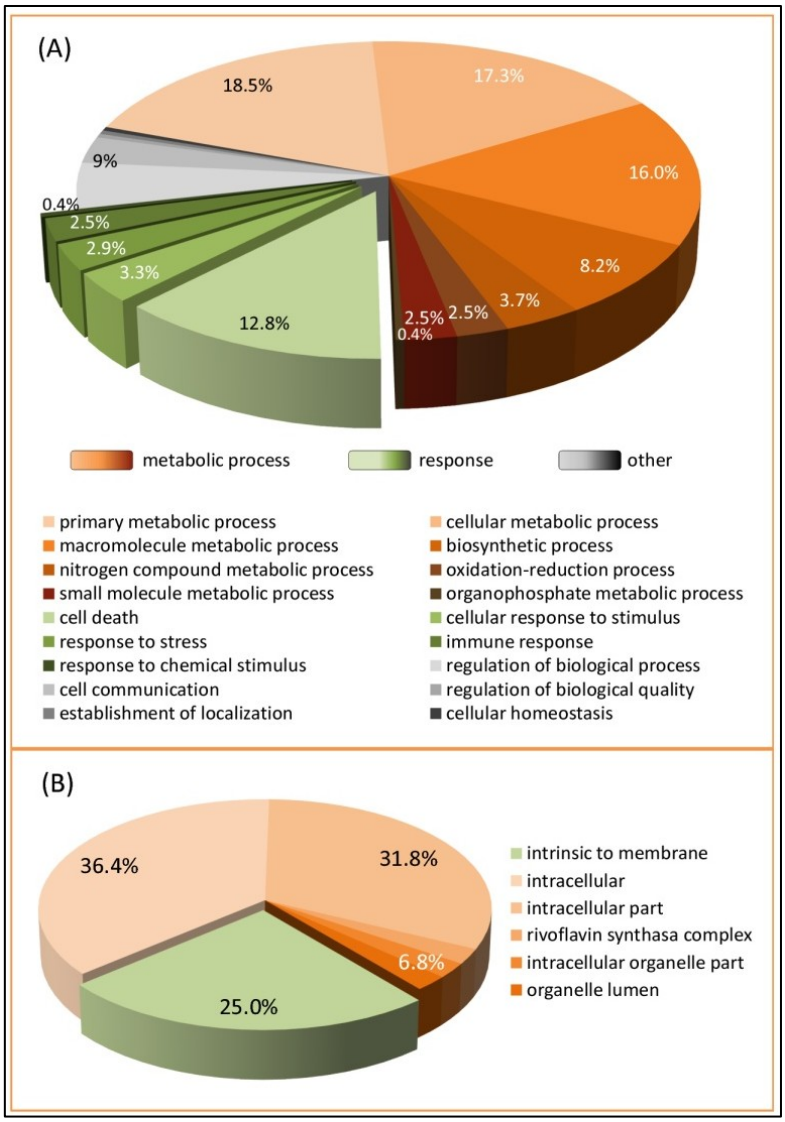

Figure 4.6. Classification of genes annotated between the TTC8 and CiC3248-06 markers according to gene ontology (GO) functional categories. (A) GO biological process categories. (B) GO cellular component categories.

In the region defined by the two flanking markers, which includes the *ABSr* locus, 33 disease resistance homologous genes were found (Figure 4.7). Thirty of these genes encode proteins predicted to have a central nucleotide-binding site (NBS) domain, 28 are involved in apoptosis, and 29 have a C-terminal leucine-rich repeat (LRR) domain. Six of the 30 NBS-containing genes have transmembrane activity. Among the resistance genes identified, 15 are homologous to the *LOV1* gene, which has been implicated in dominant susceptibility of *Arabidopsis* to the victorin toxin produced by *Cochliobolus victoriae* Nelson (Lorang *et al.*, 2007). Other three of these resistance genes belong to the *mlo* family, which confer durable broad-spectrum resistance against the powdery mildew pathogen in barley (Büschges *et al.*, 1997).

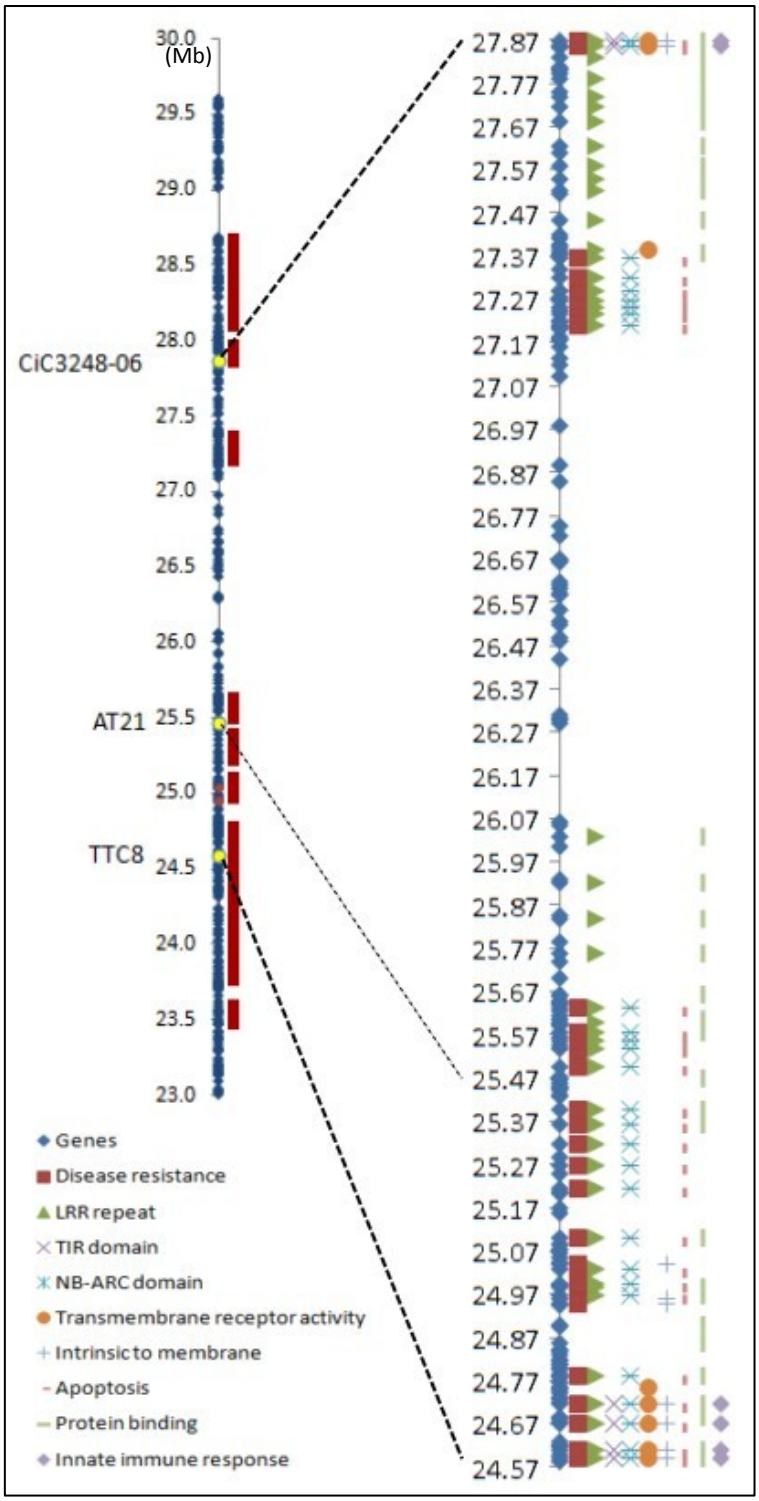

Figure 4.7. Genes found in the region around the *ABSr* locus, indicating the disease resistance gene homologous and their major domains and annotations.

Discussion

The monolocus inheritance and recessivity of ABS resistance was confirmed in citrus triploid progeny, and the *ABSr* locus was mapped in the chromosome III genetically close to the centromere

Several studies have reported the single dominance inheritance of ABS susceptibility in diploid citrus genotypes (Dalkilic *et al.*, 2005; Gulsen *et al.*, 2010); however, no data have been previously published regarding triploid progeny. In this present study, inheritance of resistance to the ABS pathogen has been analyzed in triploid progeny produced by different strategies (sexual polyploidization and interploid crosses) over a wide range of genetic backgrounds. The segregations (resistance/susceptibility) observed for all the triploid populations we evaluated confirm the monolocus inheritance and recessivity of the disease in a triploid context. All progeny arising from a homozygous susceptible cultivar, such as 'Minneola', were susceptible, whereas in cases of heterozygous parents, segregations were as expected, depending on the hybridization strategy.

Genetic mapping and marker-trait association in polyploids are complicated by the diversity of the meiotic process (and therefore, recombination mechanisms) involved, as well as the distribution of markers into multiple dosage classes (Da Silva and Sorrells, 1996). Despite these limitations, genetic maps based on segregating molecular markers have been generated for a wide range of polyploids including tetraploid cotton (Reinisch *et al.*, 1994), hexaploid tall fescue (Xu *et al.*, 1995), hexaploid wheat (Gill *et al.*, 1991) and octoploid sugar-cane mandarin (Da Silva and Sorrells, 1996; Ming *et al.*, 2001).

In this study, the first information regarding the location of the *ABSr* locus was given directly by the analysis of the segregation between susceptible and resistant triploid hybrids in progeny obtained by sexual polyploidization (2*n* gametes), assuming monolocus recessive determination of the resistance. Indeed, for 2*n* gametes resulting from either FDR or SDR, as in the 'Fortune' mandarin (Cuenca *et al.*, 2011), there is a direct linkage between parental HR (and therefore, the proportion of resistant and susceptible hybrids) and the genetic distance to the centromere. HTA is therefore an efficient way to map loci relative to a centromere (Johnson *et al.*, 1995; Zhao and Speed, 1998a; Lindner *et al.*, 2000). At this step, we identified that the *ABSr* locus was relatively close (10.5 cM) to one centromere.

The *ABSr* locus was then located by combining two approaches. The first one was to perform BSA coupled with genome-scan using SNP markers mapped in the clementine's reference genetic map (Ollitrault *et al.*, 2012a). This approach allowed localization of the ABS resistance locus within a 13.1 cM area on chromosome III of the clementine's genetic map, corresponding to 15 Mb of scaffold 3 of the current clementine whole genome assembly in pseudomolecules (www.phytozome.net). Our results confirm the potential for raw location of major genes involved in phenotypic trait variability by coupling BSA strategies with genomic

scanning, as previously proposed by Brauer *et al.* (2006), and demonstrate that it can be successfully applied in a polyploid segregating population.

This genomic region was further examined for additional SSR and SNP markers, which were used for d*e novo* mapping of the area by HTA of 2*n* gamete inferred from triploid hybrid genotypes. Mapping functions relating HR and centromere distance (Zhao and Speed, 1998a; Cuenca *et al.*, 2011) and an approach based on a correlation matrix have given convergent results in identification of the closest markers flanking the candidate ABS resistance gene. The AT21 marker appeared to be tightly linked to the *ABSr* locus. However, it would be necessary to analyse many more progeny to estimate its linkage with the *ABSr* locus precisely. The two flanking markers of AT21 and the *ABSr* locus are TTC8 and CiC3248-06, respectively, at 3.77 and 1.71 cM. This marker frame is much more closely linked than previous markers identified from diploid segregating progenies. Two RAPD markers are in loose linkage with the *ABSr* locus (15.3 cM and 36.7 cM far from the ABS locus in the same side) (Dalkilic *et al.*, 2005). A more recent study identified two flanking SRAP markers at 3 cM and 13 cM, and the authors of that study proposed that the genomic region of the *ABSr* locus should display low polymorphism, explaining the difficulty of obtaining markers very close to the gene (Gulsen *et al.*, 2010). The availability of the reference genetic map (Ollitrault *et al.*, 2012a) and whole genome sequence (Xu *et al.*, 2013) of clementine clearly increases the potential for marker-trait association studies in citrus, with co-dominant markers located both in the physical and genetic maps.

Candidate genes for resistance to ABS were identified

The *ABSr* locus appears to be included in a genomic region very rich in disease resistance gene homologs. Indeed, 17% of all resistance genes annotated in the citrus reference genome (www.phytozome.net) are found in a 6.5 Mb region (2.2% of the whole genome) of chromosome III, surrounding the *ABSr* locus. In the 3.3 Mb region defined by the two flanking markers, 33 disease resistance gene homologs were identified. Six of them are considered to be intrinsic to the membrane, including three belonging to the *mlo* family and 28 related to apoptosis. These resistance genes are organised in clusters, as generally described in many crop species (Michelmore and Meyers, 1998; Yi and Richards, 2007), and as already demonstrated in citrus for resistance to Tristeza virus found in the *Poncirus* genome (Fang *et al.*, 1998).

A. alternata is a necrotroph pathogen, which first kills host cells before parasitizing them and metabolising their contents. If the toxins used to kill host cells are not released at the right time, place, or concentration, or if a particular host genotype is insensitive to the toxin, the host cells will not die, the necrotroph will be unable to infect or reproduce, and the plant will be resistant (Guest and Brown, 1997). *A. alternata,* like other necrotrophs, produces host-selective toxins (ACT-toxins), defined as pathogen effectors, which induce toxicity and promote disease

only in the host species expressing a specific and often dominant susceptibility gene (Friesen *et al.*, 2008). Their pathogenic ability is conditioned by a gene in the pathogen that encodes production of the toxin and by a gene in the host that promotes sensitivity to the toxin. For this type of pathogen, plant resistance can be achieved via the loss or modification of the toxin's target or through detoxification (Hammond-Kosack and Jones, 1997).

Inheritance of ABS resistance in citrus has been described as monogenic (Dalkilic *et al.*, 2005; Gulsen *et al.*, 2010), controlled by a single recessive allele. The results of this study corroborate this hypothesis by demonstrating the predicted proportions of resistant and susceptible genotypes obtained from various crosses and the identification, by BSA, of a single genomic region highly associated with resistance. ACT-toxins from the tangerine pathotype of *A. alternata*, as well as AF-toxin from the strawberry pathotype and AK-toxin from the Japanese pear pathotype, have an epoxy-decatrienoic acid structure and exert their primary effect on the plasma membrane of susceptible cells, causing a rapid increase in electrolyte loss from tissues and invaginations in plasma membranes (Otani *et al.*, 1995). Varietal resistance to ACT-toxin in citrus is very highly correlated with ABS resistance. Therefore, a probable function for the gene of interest is to encode a protein involved in ACT-toxin recognition, which would allow the toxin to cause cell death. Such a dominant gene should be present (homozygously or heterozygously) in susceptible cultivars, and absent or defective in resistant cultivars.

The most obvious candidate for providing recognition specificity to the pathogen effector is the LRR domain, which binds a corresponding ligand (Hammond-Kosack and Jones, 1997) with a putative nucleotide-binding (NB) site; these genes are classified as 'NB-LRR' genes (Toyoda *et al.*, 2002). This class includes members that carry either N-terminal homology to the Toll protein and interleukin-1 receptor (TIR-NB-LRR) or a putative coiled-coil (CC) at the N-terminus (CC-NB-LRR). Resistance (R) genes from both of these subclasses confer resistance against fungi, and several fungal resistance genes have been reported and used in crop improvement programs. NB-LRR genes have been identified that confer resistance against flax rust, maize rust, barley powdery mildew, rice blast and Fusarium wilt and downy mildew of tomato (Dangl and Jones, 2001). However, sequence variation within the central LRR domain, as well as variation in LRR copy number, plays an important role in determining recognition specificity (Gururani *et al.*, 2012). Likewise, *R* genes, first identified as dominant resistance genes, could be targets of pathogen effectors and therefore play roles in susceptibility (Lorang *et al.*, 2012). Thus, avirulence (Avr) elicitors and HST may be recognising the same resistance genes in plants, leading to evolutionary outcomes that differ between necrotrophs and biotrophs while affecting the evolution of the corresponding R genes (Stukenbrock and McDonald, 2009). In *Arabidopsis*, victorin (an HST produced by *C. victoriae*) sensitivity and disease susceptibility is conferred by the *LOV1* gene, which encodes a NB-LRR protein. *LOV1* is targeted by victorin, the pathogen effector, and this interaction results in disease susceptibility (Lorang *et al.*, 2007). These NB-LRR proteins recognise specific pathogen-derived products and initiate a resistance response that often includes a type of cell death known as the hypersensitive response (Moffett

et al., 2002). In the same way, the *Pc* locus of sorghum, which contains genes encoding NB-LRR proteins, determines dominant susceptibility to HSTs produced by the necrotroph fungus *Periconia circinata* (L). Mangin Sacc. (Nagy *et al.*, 2007; Nagy and Bennetzen, 2008). Together, these results suggest that for necrotroph fungi, the disease is favoured by inducing the resistance response (Otani *et al.*, 1995; Lorang *et al.*, 2012); this mode of susceptibility could also apply to *A. alternata*. In this study, thirty disease resistance gene homologous encoding proteins with NBSs were found in the ABS locus region, and 15 of them are homologous to the *LOV1* gene. Therefore, disease resistance gene homologous should be considered as candidate genes for inducing susceptibility, especially in the case of *LOV1* homologs found in this region.

Another class of resistance genes, belonging to the *mlo* family, has also been implicated in susceptibility to barley powdery mildew produced by *Blumeria graminis* f. sp. *hordei* (Büschges *et al.*, 1997). Mlo proteins are localised in the plasma membrane and possess seven transmembrane domains; it has been suggested that they function as receptors in plants (Jørgensen and Wolfe, 1994; Büschges *et al.*, 1997; Devoto *et al.*, 1999). The resistance trait conferred by *mlo* is recessively inherited and non-race-specific, because it is effective against all isolates of the fungus *B. graminis* (Jørgensen, 1977; Qu *et al.*, 2006). Three resistance genes found in the *ABSr* locus region belong to this class. However, in citrus, two pathotypes of *A. alternata* have been described that produce HSTs that affect a narrow range of genotypes (ACT-toxin to tangerines, ACR-toxin to Rough lemon [*C. jambhiri* Lush.] and Rangpur lime [*C. limonia* Osbeck]), and resistance found in the germplasm was pathotype-specific (Timmer *et al.*, 2003).

The identification of the gene for ABS resistance will involve fine mapping with large diploid populations. SNP markers are currently being developed from each candidate gene for this purpose. From the reduced set of candidate genes that would result from this fine genetic mapping, functional validation could be performed by genetic transformation (Cervera *et al.*, 2005) or viral vector-induced gene silencing (Folimonov *et al.*, 2007; Agüero *et al.*, 2013).

For susceptible genotypes it is probable that additional genes, but also environmental factors, affect the susceptibility level. QTLs analyses conducted in susceptible progeny should be necessary to decipher this quantitative component of susceptibility.

Toward efficient breeding for ABS resistance

ABS is a major fungal disease in certain mandarin cultivars around the world; the disease causes a substantial loss of production and fruit quality (Akimitsu *et al.*, 2003; Timmer *et al.*, 2003). Currently, ABS management relies mainly on the application of fungicides (Bhatia *et al.*, 2003; Peres and Timmer, 2006; Vicent *et al.*, 2007), but this control is expensive, not environmentally friendly, and not always efficient. As a consequence, the production of

susceptible cultivars, such as 'Fortune' and ‘Nova’ among others, has declined significantly during recent years, and many trees of the most susceptible varieties have been removed and replaced by resistant cultivars that may lack some of the interesting agronomic traits of the susceptible cultivars (Navarro *et al.*, 2005). Therefore, ABS resistance must be considered as a major selection criterion in mandarin breeding programs.

Our results demonstrate that it is possible to use susceptible parents heterozygous for the resistance gene to breed resistant triploid varieties. For instance, the susceptible cultivar ‘Fortune’, which is a very efficient female parent in producing high-quality triploid hybrids in 2*x* × 2*x* hybridization (Aleza *et al.*, 2010b), should not be discarded. Indeed, the 39% and 19% of resistant triploid hybrids produced when crossed with resistant or heterozygous susceptible genotypes, respectively, are acceptable if combined with early selection by controlled inoculation phenotyping or MAS. On the other hand, parents homozygous for the susceptible allele, such as ‘Minneola’, should be definitively ruled out. Our results also demonstrate that when heterozygous susceptible parents are used as producers of diploid gametes, it is much more efficient to integrate them in a 2*x* × 2*x* strategy rather than to use them as doubled-diploid parents in interploid crosses. Indeed, the heterozygosity transmission of the *ABSr* locus (associated with susceptibility transmission to the triploid progeny) is lower in the 2*n* gametes than in the diploid gametes produced by doubled-diploids, due to its location close to the centromere of chromosome III and the SDR origin of unreduced gamete formation in most citrus genotypes (Aleza *et al.*, 2012b).

HTA has permitted identification two flanking markers at 3.77 and 1.71 cM of the *ABSr* locus, as well as a third marker that did not exhibit any recombination with the *ABSr* locus within the analyzed population. These markers should be used together for efficient early MAS for different parental combinations when the markers are heterozygous in the susceptible parent and polymorphic between the two parents. We are currently sequencing DNA fragments between these two markers to identify SNP loci that provide a useful allelic combination for the various crosses of our mandarin breeding program. These are examples of the very few identified markers for MAS in citrus, which include the SSR markers flanking the Citrus Tristeza virus resistance gene(s) of *Poncirus* [(Yang *et al.*, 2003); Mikeal Roose, personal communication] and the dominant PCR assay for the anthocyanin content of pulp of blood orange due to a transposable element in the 5’extremity of the Ruby gene (Butelli *et al.*, 2012).

Supplementary information

Table 4.S1. Number of individuals (N) evaluated within each population, and percentage of hybrids evaluated as resistant (%R)

	Population				
	'Orri' × 'Oronules'	**'Clemenules' × 'Nova'**	**'Fortune' × 'Scarlett'**	**'Fina' × 'Fairchild 4x'**	**'Fortune' × 'Kara 4x'**
Segregation	<aa × aa>	<aa × Aa>	<Aa × aa>	<aa × AAaa>	<Aa × aaaa>
N	16	11	47	15	25
% R	100%	54.55%	38.3%	20%	52%

Table 4.S2. Primers used to sequence the 4.47 kb genomic region on scaffold 3 (http:://www.phytozome.net/clementine) surrounding the SNPs identified by BSA-genome scan as linked to ABS resistance

Primer Forward	Primer Reverse	Product size	Position in scaffold 3 (bp)
CAATTTGAGCTCGCTTATTT	GGTTCATCTAGGTCACCTTCT	1154	19240437 to 19241590
TAAAACTTGGCATGGATCTT	CATATGGAATCTTCCCAGTC	1176	19241527 to 19242702
TGCCAGCTATGATAAGAACA	AGACAAAATTATCCCACTGTGT	1170	19242617 to 19243786
ATTTAAATGATGAATTTGATGC	TTATCTTTGCTGCATTTGAA	1175	19243732 to 19244906

Table 4.S3. Information about SSR and SNP markers used in this study, indicating the alleles in the parental lines and the expected genotypes within the 'Fortune' (F) × 'Willowleaf' (WL) triploid progeny

Marker id	Marker type	F alleles	WL alleles	Expected 3x genotypes	Reference
CiC1229-05	SNP (T/G)	TG	TT	TTT, TTG, TGG	Ollitrault *et al*., 2012a
CiC1875-01	SNP (T/C)	TC	CC	CCC, TCC, TTC	Ollitrault *et al*., 2012a
CiC3248-06	SNP (A/G)	AG	AA	AAA, AAG, AGG	Ollitrault *et al*., 2012a
CiC6116-04	SNP (T/C)	TC	CC	CCC, TCC, TTC	Ollitrault *et al*., 2012a
CiC6243-03	SNP (A/G)	AG	GG	GGG, AGG, AAG	Ollitrault *et al*., 2012a
CX0038	SSR	337/339	339/339	319/319/321, 319/321/321, 321/321/321	Chen *et al*., 2008
ATAC11	SSR	248/264	256/264	248/248/256, 248/248/264, 248/256/264, 248/264/264, 256/264/264, 264/264/264	new
AAT9	SSR	260/266	266/269	260/260/266, 260/260/269, 260/266/266, 260/266/269, 266/266/266, 266/266/269	new
TTC8	SSR	197/203	203/null	197/197/203, 197/197/null, 197/203/203, 197/203/null, 203/203/203, 203/203/null	new
AT21	SSR	193/205	177/205	177/193/193, 193/193/205, 177/193/205, 193/205/205, 177/205/205, 205/205/205	new
SNPALT1-Y	SNP (C/T)	CT	CC	CCC, CCT, CTT	new
SNPALT2-K	SNP (G/T)	GT	TT	GGT, GTT, GGT	new

Table 4.S4. Information about new SSR and SNP markers developed

Marker id	Position in scaffold 3 (bp)	Allele polymorphism or repeat motif	Sequence flaking the SNP	Forward primer	Reverse primer	Melting Temp (ºC)
SNPALT1-Y	19241709	C/T	CCAGACTCGTCACCACCACGCCTCCTTCCATCCAAATCGGCTGCACCTAATGTTGATGATACCACGCTGGCT**[C/T]**TAACTGTTGCCCAAGCCCGCCAAACCCAATCTAGGCCCATTGACCCCAGCCAACAC			
SNPALT2-K	19243610	G/T	GTATATGAATTTTTTTAGTTTTACGATGGATCTATT**[G/T]**TCACGTGAAAATTGGTTTCCAATTGCCTTCGCCTCACATACCATATG			
ATAC11	21375818 to 21376080	ATAC		GTCGGATTCCTCTATCAACA	TCAAGCAAGCATTTCAATAA	55
AAT9	22207742 to 22208005	AAT		TTACTTCACCTCCCTGAAAA	CAAGAATTGGGACAACTGAT	55
TTC8	24579024 to 24579224	TTC		TACATATCAAGCGCACAGAC	GACAGAGCCGAATAGAGATG	55
AT21	25473601 to 25473785	AT		TAAAATTCTGCACCGATGA	GGCTTCATTTTATTGCTTGT	55

Positions from http:://www.phytozome.net/clementine

Table 4.S5. Allelic configuration for the analyzed markers of 93 diploid female gametes within the 'Fortune' × 'Willowleaf' population

CiC6243-03	CX0038	CiC1229-05	CiC6116-04	SNP-ALT1	SNP-ALT2	ATAC11	AAT9	TTC8	AT21	CiC3248-06	CiC1875-01	*ABSr*
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	ab	bb
ab	ab	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	aa	ab	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	ab	ab	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	a-
bb	bb	bb	bb	bb	bb	aa	aa	aa	aa	bb	ab	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	ab	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	ab	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	ab	ab	ab	ab	ab	ab	a-
bb	bb	bb	bb	bb	bb	ab	ab	ab	ab	ab	ab	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	ab	ab	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	ab	ab	ab	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	ab	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
ab	ab	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	ab	aa	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	ab	ab	ab	ab	ab	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	ab	a-
aa	aa	aa	aa	ab	ab	ab	ab	ab	ab	ab	ab	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
ab	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	ab	ab	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	ab	bb
bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb	bb
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
bb	bb	bb	bb	bb	bb	bb	bb	aa	ab	ab	ab	a-
bb	bb	bb	bb	bb	bb	bb	ab	ab	ab	ab	ab	a-
bb	bb	bb	bb	bb	bb	bb	ab	ab	ab	ab	ab	a-
aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	a-
6.45%	5.38%	0.00%	0.00%	1.08%	1.08%	8.60%	11.83%	12.90%	20.43%	23.91%	31.11%	20.43%
11.26	11.40	16.29	17.86	19.24	19.24	21.38	22.21	24.57	25.47	27.87	29.26	-

%HR (ab)

Physical position (Mb)

Table 4.S6. Annotations between 24.57 Mb (TTC8 marker) and 27.87 Mb (CiC3248-06 marker) in scaffold 3 (www.phytozome.net). (Annotations related to resistance response are indicated in bold letters).

Initial position (bp)	final position (bp)	Locus name	Annotations
24584262	24588809	Ciclev10023819m.g	Aluminium activated malate transporter family protein
24593002	24597811	Ciclev10023283m.g	**disease resistance protein (TIR-NBS-LRR class), putative**
24598362	24600642	Ciclev10020543m.g	alpha/beta-Hydrolases superfamily protein
24611867	24615456	Ciclev10024092m.g	**disease resistance protein (TIR-NBS-LRR class), putative**
24617312	24619000	Ciclev10021706m.g	alpha/beta-Hydrolases superfamily protein
24620623	24627357	Ciclev10020009m.g	Enhancer of polycomb-like transcription factor protein
24629370	24637605	Ciclev10018715m.g	Alkaline-phosphatase-like family protein
24669795	24674192	Ciclev10018914m.g	**disease resistance protein (TIR-NBS-LRR class), putative**
24674735	24676674	Ciclev10021413m.g	alpha/beta-Hydrolases superfamily protein
24683380	24686055	Ciclev10023561m.g	FAR1-related sequence 5
24687505	24689173	Ciclev10023487m.g	F-box/RNI-like superfamily protein
24717374	24720944	Ciclev10023733m.g	**disease resistance protein (TIR-NBS-LRR class), putative**
24722080	24724113	Ciclev10021643m.g	alpha/beta-Hydrolases superfamily protein
24726485	24731694	Ciclev10021199m.g	Enhancer of polycomb-like transcription factor protein
24732324	24734809	Ciclev10023873m.g	Alkaline-phosphatase-like family protein
24741494	24743967	Ciclev10021112m.g	RHOMBOID-like protein 14
24744597	24750079	Ciclev10019165m.g	DEA(D/H)-box RNA helicase family protein
24750794	24752907	Ciclev10022561m.g	Ribosomal L18p/L5e family protein
24753302	24753667	Ciclev10023727m.g	Plant self-incompatibility protein S1 family
24756088	24758749	Ciclev10018885m.g	cation/H+ exchanger 19
24768060	24768577	Ciclev10023601m.g	NAC domain containing protein 46
24777286	24779149	Ciclev10024369m.g	Ankyrin repeat family protein
24779630	24784345	Ciclev10018504m.g	**NB-ARC domain-containing disease resistance protein**
24785318	24795064	Ciclev10019018m.g	MUTL-homologue 1
24796351	24800588	Ciclev10019689m.g	serine hydroxymethyltransferase 3
24802124	24804993	Ciclev10022715m.g	Surfeit locus protein 5 subunit 22 of Mediator complex
24806596	24809323	Ciclev10023969m.g	ribosomal protein L24
24809538	24814289	Ciclev10020286m.g	Transducin/WD40 repeat-like superfamily protein
24815262	24816480	Ciclev10023234m.g	bonsai
24819124	24831081	Ciclev10018511m.g	Transducin/WD40 repeat-like superfamily protein
24831949	24837740	Ciclev10020270m.g	methionine aminopeptidase 2B
24837938	24840296	Ciclev10021676m.g	Nucleic acid-binding, OB-fold-like protein
24840386	24845254	Ciclev10019406m.g	methylenetetrahydrofolate reductase 2
24856909	24862275	Ciclev10024071m.g	histone-lysine N-methyltransferase ASHH3
24893551	24898675	Ciclev10019906m.g	Peptidase family C54 protein
24899223	24901003	Ciclev10020147m.g	Galactose oxidase/kelch repeat superfamily protein
24943610	24945394	Ciclev10022372m.g	Ribosomal protein L6 family
24949255	24954319	Ciclev10019784m.g	**Seven transmembrane MLO family protein**
24958927	24963659	Ciclev10023336m.g	**Seven transmembrane MLO family protein**
24966135	24969178	Ciclev10024611m.g	**NB-ARC domain-containing disease resistance protein**
24985138	24987298	Ciclev10023902m.g	**LRR and NB-ARC domains-containing disease resistance protein**
24991889	24992787	Ciclev10024232m.g	**NB-ARC domain-containing disease resistance protein**
24993845	24995370	Ciclev10024038m.g	**NB-ARC domain-containing disease resistance protein**
25026668	25030809	Ciclev10024551m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25039430	25044580	Ciclev10020313m.g	**Seven transmembrane MLO family protein**
25044600	25051607	Ciclev10020879m.g	protein serine/threonine kinases;ATP binding;catalytics
25053425	25057795	Ciclev10020158m.g	guanosine nucleotide diphosphate dissociation inhibitor 1
25067469	25072314	Ciclev10019310m.g	Ankyrin repeat family protein
25074163	25076367	Ciclev10022861m.g	auxin-regulated gene involved in organ size
25085745	25086451	Ciclev10023183m.g	
25087317	25089114	Ciclev10020871m.g	Late embryogenesis abundant protein, group 2
25100778	25104843	Ciclev10018688m.g	**NB-ARC domain-containing disease resistance protein**
25158172	25162574	Ciclev10019028m.g	Prolyl oligopeptidase family protein
25169542	25170675	Ciclev10023618m.g	
25204996	25205221	Ciclev10024113m.g	Chaperone DnaJ-domain superfamily protein
25205365	25207063	Ciclev10023518m.g	
25216526	25220629	Ciclev10024119m.g	**LRR and NB-ARC domains-containing disease resistance protein**

25248367	25251368	Ciclev10023445m.g	indole-3-acetate beta-D-glucosyltransferase
25269003	25272836	Ciclev10018531m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25286483	25288105	Ciclev10019912m.g	UDP-glucosyl transferase 75B2
25318178	25322601	Ciclev10018594m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25361089	25365339	Ciclev10018499m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25398286	25402615	Ciclev10023525m.g	**NB-ARC domain-containing disease resistance protein**
25426124	25427403	Ciclev10024201m.g	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
25432861	25433406	Ciclev10024286m.g	BED zinc finger ;hAT family dimerization domain
25441988	25444731	Ciclev10019183m.g	RAP
25445568	25448031	Ciclev10023789m.g	ADP-glucose pyrophosphorylase small subunit 2
25448858	25451858	Ciclev10022149m.g	6,7-dimethyl-8-ribityllumazine synthase / DMRL synthase / lumazine synthase / riboflavin synthase
25452062	25455844	Ciclev10020914m.g	Dihydrodipicolinate reductase, bacterial/plant
25458171	25463466	Ciclev10019560m.g	Poly (ADP-ribose) glycohydrolase (PARG)
25463711	25465970	Ciclev10020069m.g	FBD, F-box, Skp2-like and Leucine Rich Repeat domains containing protein
25470795	25473216	Ciclev10020065m.g	F-box/RNI-like/FBD-like domains-containing protein
25474880	25476924	Ciclev10024256m.g	Thioredoxin superfamily protein
25495290	25499648	Ciclev10023260m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25539228	25543442	Ciclev10018540m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25546089	25546373	Ciclev10023953m.g	mitochondrial ribosomal protein L11
25558189	25563873	Ciclev10018510m.g	**LRR and NB-ARC domains-containing disease resistance protein**
25563004	25563956	Ciclev10024474m.g	
25577398	25580479	Ciclev10023481m.g	**NB-ARC domain-containing disease resistance protein**
25591667	25592171	Ciclev10022922m.g	
25596184	25598211	Ciclev10019166m.g	
25598722	25601734	Ciclev10020079m.g	F-box family protein
25605171	25606024	Ciclev10023014m.g	F-box/RNI-like superfamily protein
25612741	25615635	Ciclev10023374m.g	uridine-ribohydrolase 2
25625682	25630835	Ciclev10019447m.g	inositol 1,3,4-trisphosphate 5/6-kinase 4
25633452	25636420	Ciclev10018897m.g	**Disease resistance protein (CC-NBS-LRR class) family**
25639011	25644826	Ciclev10019649m.g	RNA-binding protein
25645230	25649184	Ciclev10021021m.g	chloroplast outer envelope protein 37
25649848	25653317	Ciclev10024361m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
25657664	25663490	Ciclev10024293m.g	Endonuclease/exonuclease/phosphatase family protein
25663967	25668616	Ciclev10019293m.g	Ankyrin repeat family protein
25702085	25702716	Ciclev10023674m.g	
25740008	25740471	Ciclev10023198m.g	Pectin lyase-like superfamily protein
25756520	25759773	Ciclev10018637m.g	Leucine-rich repeat receptor-like protein kinase family protein
25784246	25784449	Ciclev10023998m.g	
25838882	25842061	Ciclev10023511m.g	Leucine-rich repeat receptor-like protein kinase family protein
25844828	25845232	Ciclev10024127m.g	Plant self-incompatibility protein S1 family
25920662	25925208	Ciclev10023567m.g	Leucine-rich repeat receptor-like protein kinase family protein
25929095	25929878	Ciclev10024445m.g	Leucine-rich repeat receptor-like protein kinase family protein
26006287	26009481	Ciclev10024013m.g	Leucine-rich repeat receptor-like protein kinase family protein
26027989	26031836	Ciclev10024332m.g	Leucine-rich repeat receptor-like protein kinase family protein
26054584	26055280	Ciclev10023832m.g	Leucine-rich repeat receptor-like protein kinase family protein
26061449	26062187	Ciclev10023220m.g	gamma-glutamyl hydrolase 1
26288659	26290508	Ciclev10021937m.g	response regulator 9
26297949	26304739	Ciclev10021153m.g	Plant protein 1589 of unknown function
26303573	26304130	Ciclev10023069m.g	
26308767	26309588	Ciclev10023742m.g	FAR1-related sequence 5
26439998	26440627	Ciclev10024302m.g	FAR1-related sequence 5
26480614	26486649	Ciclev10020055m.g	cystathionine beta-lyase
26487045	26487269	Ciclev10023935m.g	Ribosomal protein L39 family protein
26517176	26525778	Ciclev10019027m.g	Trimeric LpxA-like enzyme
26526425	26531920	Ciclev10021922m.g	tubulin folding cofactor B
26552200	26563660	Ciclev10018515m.g	binding
26585748	26586352	Ciclev10023130m.g	
26591205	26592308	Ciclev10024114m.g	Glutaredoxin family protein
26603126	26608232	Ciclev10021999m.g	Eukaryotic rpb5 RNA polymerase subunit family protein
26610747	26615954	Ciclev10021431m.g	
26618336	26621274	Ciclev10022600m.g	Tetratricopeptide repeat (TPR)-like superfamily protein

26660974	26662413	Ciclev10021720m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
26666978	26668075	Ciclev10024496m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
26671216	26672059	Ciclev10024436m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
26724892	26726825	Ciclev10023900m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
26747398	26748527	Ciclev10023903m.g	TTF-type zinc finger protein with HAT dimerization domain
26847877	26850055	Ciclev10021728m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
26886018	26886467	Ciclev10023695m.g	
26976583	26978658	Ciclev10021734m.g	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
27092414	27093612	Ciclev10024663m.g	Zinc-binding dehydrogenase family protein
27120138	27122618	Ciclev10024396m.g	S-locus lectin protein kinase family protein
27132311	27137357	Ciclev10023678m.g	cysteine-rich RLK (RECEPTOR-like protein kinase) 8
27160750	27163015	Ciclev10024040m.g	Na+/H+ antiporter 6
27174902	27177223	Ciclev10023619m.g	MuDR family transposase
27179651	27180616	Ciclev10024633m.g	
27182222	27186798	Ciclev10018865m.g	Prolyl oligopeptidase family protein
27202252	27203036	Ciclev10023045m.g	hemoglobin 1
27203532	27205968	Ciclev10024617m.g	Zinc-binding dehydrogenase family protein
27206738	27207496	Ciclev10023412m.g	Zinc-binding dehydrogenase family protein
27210508	27212925	Ciclev10023669m.g	**NB-ARC domain-containing disease resistance protein**
27218280	27219176	Ciclev10023093m.g	
27222437	27223126	Ciclev10022330m.g	Ribosomal protein L2 family
27236430	27242516	Ciclev10018492m.g	**LRR and NB-ARC domains-containing disease resistance protein**
27243916	27244600	Ciclev10023862m.g	
27252187	27258492	Ciclev10023613m.g	**LRR and NB-ARC domains-containing disease resistance protein**
27267893	27272240	Ciclev10018507m.g	**LRR and NB-ARC domains-containing disease resistance protein**
27274447	27274841	Ciclev10023703m.g	Ribosomal protein L2 family
27288919	27290352	Ciclev10023361m.g	Zinc-binding dehydrogenase family protein
27291189	27295420	Ciclev10024454m.g	**NB-ARC domain-containing disease resistance protein**
27320758	27322899	Ciclev10024530m.g	**NB-ARC domain-containing disease resistance protein**
27338257	27339658	Ciclev10023645m.g	Ribosomal protein L2 family
27365802	27371376	Ciclev10018509m.g	**LRR and NB-ARC domains-containing disease resistance protein**
27373577	27374040	Ciclev10024425m.g	Ribosomal protein L2 family
27378259	27383779	Ciclev10019887m.g	Integrin-linked protein kinase family
27385466	27391796	Ciclev10018983m.g	LETM1-like protein
27392621	27394211	Ciclev10020021m.g	HXXXD-type acyl-transferase family protein
27395608	27398058	Ciclev10020092m.g	UDP-glucosyltransferase 74F2
27411566	27413326	Ciclev10023826m.g	UDP-glucosyltransferase 74F2
27420861	27421928	Ciclev10024366m.g	Protein kinase family protein with leucine-rich repeat domain
27453996	27456865	Ciclev10023662m.g	Leucine-rich repeat receptor-like protein kinase family protein
27512875	27514002	Ciclev10024240m.g	Leucine-rich repeat receptor-like protein kinase family protein
27522643	27525763	Ciclev10023288m.g	Leucine-rich repeat receptor-like protein kinase family protein
27550207	27553496	Ciclev10024387m.g	Leucine-rich repeat receptor-like protein kinase family protein
27579560	27591874	Ciclev10024208m.g	Leucine-rich repeat receptor-like protein kinase family protein
27610925	27611502	Ciclev10023474m.g	Leucine-rich repeat receptor-like protein kinase family protein
27624874	27626398	Ciclev10021536m.g	Protein kinase family protein with leucine-rich repeat domain
27682550	27684073	Ciclev10024553m.g	Protein kinase family protein with leucine-rich repeat domain
27717305	27722483	Ciclev10023899m.g	Protein kinase family protein with leucine-rich repeat domain
27737586	27740735	Ciclev10018816m.g	Leucine-rich repeat receptor-like protein kinase family protein
27750059	27750755	Ciclev10022950m.g	Ribosomal protein S5/Elongation factor G/III/V family protein
27780847	27783988	Ciclev10018798m.g	Leucine-rich repeat receptor-like protein kinase family protein
27792547	27797748	Ciclev10020409m.g	BTB and TAZ domain protein 3
27798973	27801782	Ciclev10023071m.g	Small nuclear ribonucleoprotein family protein
27802944	27807050	Ciclev10019655m.g	Thioesterase/thiol ester dehydrase-isomerase superfamily protein
27807172	27813248	Ciclev10018774m.g	formin homology 1
27827995	27830405	Ciclev10019217m.g	**Leucine-rich repeat receptor-like protein kinase family protein**
27852621	27853757	Ciclev10023555m.g	Eukaryotic release factor 1 (eRF1) family protein
27856160	27856424	Ciclev10024351m.g	
27857215	27861301	Ciclev10018573m.g	**Disease resistance protein (TIR-NBS-LRR class) family**
27867691	27874119	Ciclev10018528m.g	**Disease resistance protein (TIR-NBS-LRR class), putative**

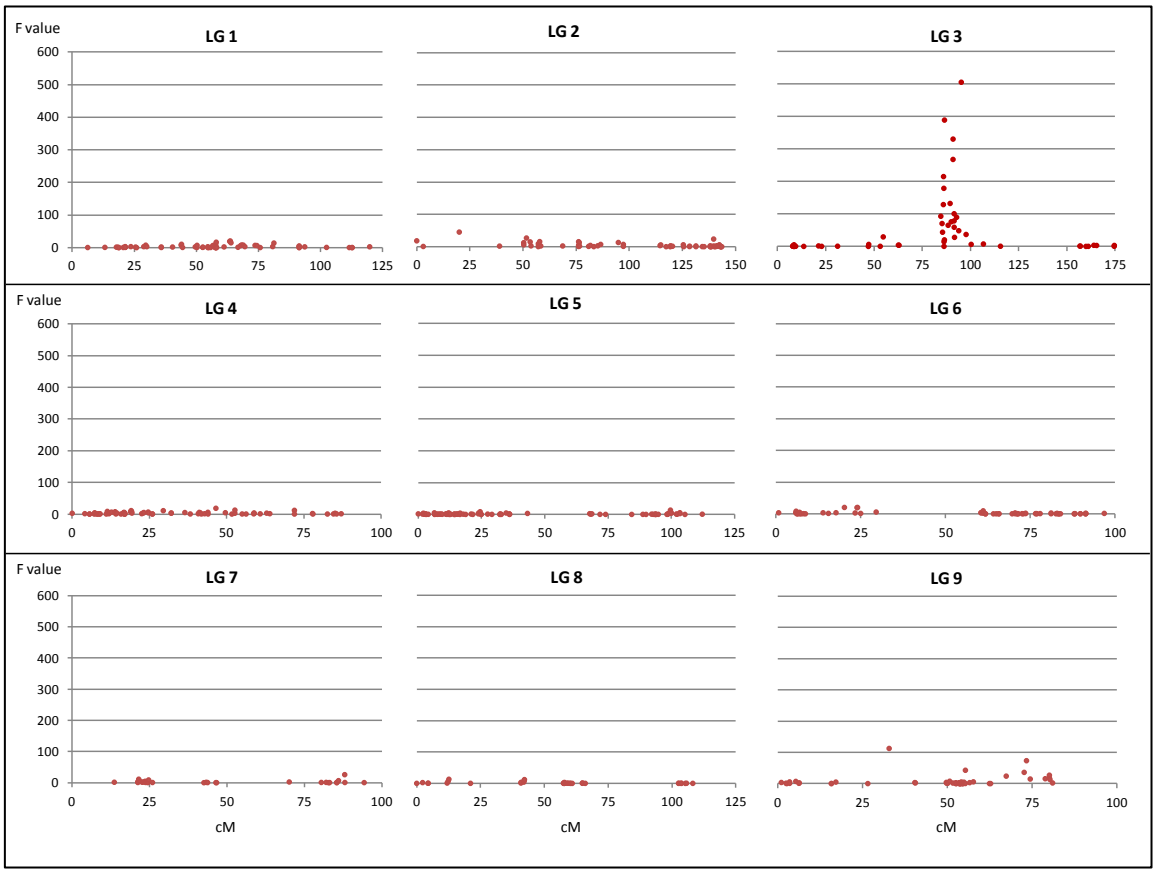

Figure 4.S1. Values of F parameter from ANOVA along the nine linkage groups of the clementine's genetic map (Ollitrault *et al.*, 2012a)

General discussion

Background

Polyploidy is a major component of angiosperms evolution (Grant, 1981; Soltis and Soltis, 1993; Wendel and Doyle, 2005) and sexual polyploidization is considered as the major mechanism leading to polyploidy (Harlan and DeWet, 1975; Bretagnolle and Thompson, 1995; Ramsey and Schemske, 1998). In sexual polyploidization, polyploids are generated by the formation of diploid (2*n*) gametes, i.e., pollen or eggs having the somatic chromosome number rather than the gametophytic number (Harlan and DeWet, 1975; Bretagnolle and Thompson, 1995; Ramsey and Schemske, 1998; Otto and Whitton, 2000), mainly resulting from a restitution of the meiotic cell cycle. This phenomenon is referred to as meiotic restitution and it is the predominant mechanism of unreduced gamete formation in plants (Harlan and DeWet, 1975; Bretagnolle and Thompson, 1995; Otto and Whitton, 2000; Soltis and Soltis, 2009). The types of unreduced gametes produced result essentially from one of two basic processes depending on the mode of meiotic restitution (Bretagnolle and Thompson, 1995; Tavoletti *et al.*, 1996), i.e., first-division restitution (FDR) and second-division restitution (SDR), which occur, respectively, during abnormal development of the first and second meiotic division.

A FDR-2*n* gamete contains non-sister chromatids, while a SDR-2*n* gamete contains two sister chromatids (Bretagnolle and Thompson, 1995; Tavoletti *et al.*, 1996; Cai and Xu, 2007). Under FDR, non-sister chromatids retain the parental heterozygosity from the centromere to the first crossover point, and hence, the gametes transfer a large part of this parental heterozygosity and epistatic interactions to the progenies. Under SDR, the two sister chromatids are homozygous between the centromere and the first crossover point, and the resultant gametes have reduced levels of heterozygosity compared with FDR ones (Bretagnolle and Thompson, 1995). Therefore, the genetic constitution of unreduced gametes depends on the mechanism of their formation: a tighter distribution is expected in FDR-derived populations than in SDR ones because a higher percentage of the parental genome is transferred intact, resulting in a more uniform gamete production (Douches and Maas, 1998). In this context, insights in the mechanism underlying meiotic nuclear restitution producing unreduced gametes are crucial to optimize breeding strategies based on sexual hybridization (Errico *et al.*, 2005).

Sexual polyploidization is currently a central approach in citrus triploid breeding programs, aiming to develop new seedless mandarin cultivars (Ollitrault *et al.*, 2008). Citrus triploid hybrids can be recovered mainly from sexual hybridizations of 2*x* × 2*x* [through unreduced gametes; (Aleza *et al.*, 2010b)], 2*x* × 4*x* (Aleza *et al.*, 2012c) and 4*x* × 2*x* (Cameron and Bernett, 1978; Esen *et al.*, 1978; Aleza *et al.*, 2012d).

Esen *et al.* (1979) proposed that unreduced gametes in citrus result from the abortion of the second meiotic division (SDR) in the megaspore. This hypothesis has been corroborated by molecular marker analysis for clementines (*Citrus clementina*) (Luro *et al.*, 2004; Aleza *et al.*, 2012b). However, Chen *et al.* (2008a) proposed that 2*n* eggs of sweet orange (*C. sinensis*) resulted from first meiotic division restitution (FDR).

Molecular marker analysis has proved as a very useful tool to estimate the heterozygosity transmission through the diploid gametes to polyploid progenies and, therefore, to identify the mechanism underlying unreduced gamete formation. This strategy was previously successfully applied for several crops (Barone *et al.*, 1995; Vorsa and Rowland, 1997; Bastiaanssen *et al.*, 1998; Barcaccia *et al.*, 2003; Luro *et al.*, 2004; Chen *et al.*, 2008a; Hayashi *et al.*, 2009). However, the estimation of molecular marker allele copy number has long been considered as a challenge for polyploid species with polysomic inheritance. Assignment of allelic configurations for different types of heterozygous polyploids is essential for accurate genetic studies, such as segregation analyses and marker-trait association. Moreover, single nucleotide polymorphisms (SNPs) have emerged as the most widely used genotyping markers due to their abundance in the genome and increasing available data from sequencing projects. In addition, newly developed genetic maps have utilized information on SNP marker segregations. In this context, new methods to assign SNP allele configuration in polyploids should be considered.

On the other hand, trait segregations in triploid populations are still not well understood, due to the complexity of polyploid genomes and eventually non-Mendelian segregations. In the case of citrus, very few characters have been studied on the basis of triploid genomes, and genetics underlying their segregations have not been characterized in any case.

In this thesis, we have (i) developed a new method for genotyping heterozygous polyploid plants with SNP markers, (ii) proposed two methods for the identification of the origin of 2*n* gametes from genetic markers segregation data and (iii) applied them to study the origin of 2*n* ovules in mandarin germplasm. Finally, (iv) we analyzed the inheritance of resistance to the fungus causing Alternaria brown spot as a case of single trait segregation on triploid populations. We have taken advantage of the acquired knowledge on 2*n* gamete origin to locate a genome region controlling the resistance trait by half tetrad analysis, and to identify candidate genes. From these data, we have developed SSR and SNP markers for MAS.

Competitive Allele-Specific PCR is an efficient method to assign SNP allelic configuration in heterozygous polyploid genotypes

When analysing microsatellite markers (SSRs), the microsatellite allele counting – peak ratios method [MAC-PR; (Esselink *et al.*, 2004)] is especially useful in assignment of allelic configurations in polyploids. However, SSR analysis remains relatively costly and time consuming compared with SNP genotyping methods. In this sense, several SNP genotyping methods have been developed (Ronaghi *et al.*, 1998; Sapolsky *et al.*, 1999; Ahmadian *et al.*, 2000; Fan *et al.*, 2003; Kwok and Xiao, 2003; Lavebratt *et al.*, 2004; Ishikawa *et al.*, 2005; Tabone *et al.*, 2009). The KBiosciences Competitive Allele-Specific PCR SNP genotyping system [KASPar; (Cuppen, 2007)] is a simple and cost-effective system compared with other

SNP genotyping assays and is well adapted to low- to medium- throughput genotyping projects (Chen *et al.*, 2010). In this thesis, we have demonstrated that the relative allele signal of KASPar technique was highly correlated with relative allele doses and this method, combined with analysis of the allele signal data, has been useful for genotyping citrus triploid and tetraploid progenies (Cuenca *et al.*, 2013a). Moreover, it could also be used in the semi-quantitative analysis of relative allele-specific expression. KASPar technique is much more adapted for small- to medium- scale studies than the SNP genotyping methods previously used in polyploidy species (Bérard *et al.*, 2009; Trick *et al.*, 2009; Oliver *et al.*, 2011; Trebbi *et al.*, 2011; Voorrips *et al.*, 2011; Byers *et al.*, 2012; Han *et al.*, 2012). Indeed, these methods are more suitable for high-throughput genotyping than for performing small-scale analysis, and lack flexibility in terms of the numbers and panels of SNP loci that can be analyzed, while it is possible to design targeted KASPar assay for each specific study. Moreover, several of the previously used methods (Bérard *et al.*, 2009; Voorrips *et al.*, 2011) base the assignment of allele doses in polyploidy plants on statistical approach, assuming Hardy-Weinberg equilibrium, which is often not respected, particularly when working with progenies arising from 2*n* gametes. In our work, we have demonstrated that Competitive Allele-Specific PCR is an effective and versatile method to infer SNP allelic configurations in polyploid plants and developed KASPar markers that are currently routinely used for triploid and tetraploid genotyping in the IVIA citrus breeding program.

Two new alternative approaches to identify the mechanism leading to unreduced gamete formation based on molecular marker (HR restitution) analysis have been developed

Identifying the restitution mechanism using molecular markers can be achieved by several approaches. Half-tetrad analysis (HTA) developed by Mendiburu and Peloquin (1979) is a powerful method for mapping centromeres, or for determining the mode(s) of 2*n* gamete formation. Tavoletti *et al.* (1996) developed a multilocus maximum likelihood method of HTA that allows the estimation of both the relative frequencies of FDR and SDR 2*n* gametes and the centromere location within a linkage group without relying on previously identified centromeric markers. However, all these models assume complete chiasma interference.

In this study, two alternative approaches have been proposed to check FDR/SDR hypothesis and, eventually, map the centromeres within linkage groups. For these two methods, molecular markers are selected to be heterozygous in the 2*n* gamete producer parent and to be polymorphic between the two parents. With this marker configuration, it is possible to infer the 2*n* gamete genotypes from the triploid progeny ones. The first developed approach is based on functions of parental heterozygosity restitution in the 2*n* gamete (HR), along a chromosome in relation with locus-centromere distance (d) (Zhao and Speed, 1998a). Indeed, under FDR or SDR, HR is a direct function of the crossing over frequency between the considered locus and the centromere. It is therefore possible to implement functions (HR=f(d)) according to the FDR

and SDR hypothesis taking into account different models of chromosome interference (no interference, partial interference or complete chiasma interference). This strategy has been successfully applied during this thesis in 'Fortune' mandarin. The second alternative proposed to check FDR/SDR hypothesis is a maximum likelihood approach, based on the HR of independent markers closely linked to centromere of different chromosomes. Taking advantage of the centromere location (Aleza *et al.*, 2012b) within all nine linkage groups of the current clementine's reference genetic map (Ollitrault *et al.*, 2012a), centromeric markers were selected to analyze the predominant mechanisms for 16 mandarin genotypes used as female parents and to check potential variability of origin between 2*n* gametes of a same parent.

The first newly developed method allows the identification of the unreduced gamete formation mechanism without any previous knowledge about centromere location, but it is necessary to analyze numerous markers with good dispersion in the chromosomes in large populations. In this method, the information on the mechanism is obtained at population level. Compared with previously published approaches of HTA (Mendiburu and Peloquin, 1979; Werner *et al.*, 1992; Johnson *et al.*, 1995; Tavoletti *et al.*, 1996; Park *et al.*, 2007) this new approach has the advantage to allow testing for no interference or partial interference models, but needs a set of previously mapped markers; on the other hand, HTA models assume total interference, but they could be applied without any previous information on marker order. In addition, this new method allows the identification of the restitution model and the estimation of the centromere location at the same time. During this thesis, it has been used to locate the centromere on chromosome II of 'Fortune' mandarin (Cuenca *et al.*, 2011) and recently to map the centromeres on all 9 clementine chromosomes (Aleza *et al.*, 2012b).

The second developed approach requires a previous location of centromeres, but is much more efficient than all previous methods. Indeed, it need a limited number of centromeric markers and has the advantage that allows deducing the restitution mechanism both at individual and population level, while the first developed one is only applicable at population level, requiring larger number of genotyping points (numerous markers and individuals) to give robust conclusions.

SDR is the preeminent mechanism leading to unreduced ovule formation in mandarins

In this thesis, the restitution mechanism leading to unreduced gamete formation has been uncovered for 16 mandarin genotypes by molecular marker analysis. Firstly, all the analyzed triploid hybrids in the progenies of 2*x* × 2*x* crosses were found to arise from unreduced megaspores. This result is in agreement with the cytogenetic observations of Esen and Soost (1971), and with previous molecular observations (Luro *et al.*, 2004; Chen *et al.*, 2008a; Ferrante *et al.*, 2010). To date, very few cases of citrus triploid hybrids occurrence in 2*x* × 2*x* crosses from unreduced pollen has been reported (Luro *et al.*, 2004; Chen *et al.*, 2008a).

The first approach described above allowed concluding that SDR, rather than FDR, was much more likely to be the mechanism underlying unreduced gamete formation in 'Fortune' mandarin and evidenced positive chiasma interference. Using the second approach, SDR has been identified as the preeminent restitution mechanism leading to unreduced gamete formation in all mandarin genotypes analyzed. Indeed among 497 triploid hybrids arising from 16 female genotypes studied, significant conclusions for SDR were obtained for 85.3%, but only 0.6% for FDR. The un-conclusive results for 14.1% of the analyzed triploids can be explained by insufficient number of markers analyzed in some families and/or the higher distance of some of these markers to the centromere.

This global conclusion for SDR restitution mechanism is in agreement with that proposed by Luro *et al.* (2004), who observed low heterozygosity restitution in clementine 2*n* megagametophyte. The conclusion of FDR given for sweet orange (Chen *et al.*, 2008a) is questionable because of the low number of analyzed markers without knowledge of their distance to centromere. Indeed, the unambiguous identification of FDR without previous location of the centromere must be based on a large set of markers with good genome coverage. In the same way, the results of Ferrante *et al.* (2010), based on a very low number of individuals and markers for each parental genotype are not sufficient to prove the authors's conclusions of SDR for 'Fortune' and 'Wilking' mandarin and FDR for lemon.

Triploid recovery strategy and restitution mechanism leading to unreduced gamete formation directly affect single character segregations in triploid progenies. The case of Alternaria brown spot resistance

The main factor affecting trait inheritance in triploid progenies is the strategy of triploid breeding (Ollitrault *et al.*, 2008) with strong differences between the sexual polyploidization approach (2*x* × 2*x* crosses with unreduced gamete formation) and interploid crosses (2*x* × 4*x* or 4*x* × 2*x*). Indeed, it affects the transmission of the parental heterozygosity to the diploid gamete.

In sexual polyploidization, two factors affect the transmission of the parental heterozygosity to the offspring: the mechanism of the 2*n* gamete formation (FDR or SDR), and the genetic distance from the considered locus to the chromosome centromere (Douches and Quiros, 1988). In 2*x* × 2*x* citrus crosses, most studies have demonstrated that triploid progeny arose from diploid megagametophyte (Esen and Soost, 1971, 1973). In this thesis, we have demonstrated that 2*n* megagametophyte resulted from SDR in the mandarin group, as previously proposed by Luro *et al.* (2004) for clementines. For interploid crosses, most of the tetraploid parents used in citrus breeding arise from chromosome doubling in nucellar cells of apomictic diploid parents (Aleza *et al.*, 2011). Therefore, doubled diploid mandarins should be considered as autotetraploids and tetrasomic inheritance should be expected (Kamiri *et al.*, 2011). In such situation, the frequency of diploid gametes that receive a locus in heterozygosis

from the tetraploid parent varies between 0.55 and 0.66 depending on the double reduction frequency (Marsden *et al.*, 1987).

Segregations for the single inherited character Alternaria brown spot (ABS) resistance have been analyzed in this study for several triploid recovery strategies and parental combinations. ABS is a major fungal disease in some mandarin cultivars that causes a substantial loss of production and fruit quality worldwide (Akimitsu *et al.*, 2003; Timmer *et al.*, 2003). As a consequence, the production of susceptible cultivars, such as 'Fortune' or 'Nova', among others, has declined significantly during recent years and many trees of the most susceptible varieties have been removed and replaced by resistant cultivars, which may lack some of the interesting agronomic traits of these susceptible cultivars (Navarro *et al.*, 2005). In addition, many of actual triploid breeding programs (Mourao Fo *et al.*, 1996; McCollum, 2007; Aleza *et al.*, 2010a, 2010b; Cuenca *et al.*, 2010; Grosser *et al.*, 2010; Aleza *et al.*, 2012c, 2012d; Froelicher *et al.*, 2012; Navarro *et al.*, 2012) use ABS susceptible cultivars as parents, due to their interest for other important traits (fruit quality, maturing period, production) and particular reproductive biology (monoembryony, high rate of unreduced 2*n* gamete formation). Therefore, ABS resistance must be considered as a major selection criterion in the mandarin breeding programs and early efficient selection of ABS resistant hybrids is of central importance.

In this thesis, the monolocus dominant inheritance of the susceptibility, proposed from diploid population studies (Dalkilic *et al.*, 2005; Gulsen *et al.*, 2010), was corroborated in triploid progenies by studying the segregations in several parent combinations and triploid recovery strategies. Moreover, a genome region including the ABS resistance (*ABSr* locus) was identified by genetic studies based in linkage disequilibrium.

The first information regarding the location of the *ABSr* locus was given directly by the analysis of the segregation between susceptible and resistant triploid hybrids in progeny obtained by sexual polyploidization (2*n* gametes), assuming monolocus recessive determination of the resistance. Indeed, for 2*n* gametes, there is a direct linkage between parental HR (and therefore, the proportion of resistant and susceptible hybrids) and the genetic distance to the centromere. At this step, *ABSr* locus was identified to be relatively close (10.5 cM) to one centromere.

The *ABSr* locus was then located performing Bulk Segregant Analysis (BSA) coupled with genome-scan using SNP markers mapped in the clementine's reference genetic map (Ollitrault *et al.*, 2012a). This approach allowed localization of the ABS resistance locus within a 13.1 cM area on chromosome III of the genetic map of clementine, corresponding to 15 Mb of scaffold 3 of the current clementine whole genome assembly (www.phytozome.net). This genomic region was further examined for additional SSR and SNP markers, which were used for d*e novo* mapping of the area by HTA of 2*n* gamete inferred from triploid hybrid genotypes. Results allowed identifying two flanking markers, located at 3.77 and 1.71 cM of the candidate

ABS resistance gene, and another one that appeared to be tightly linked to the *ABSr* locus. This marker frame is much more closely linked than previous markers identified from diploid segregating progenies by Dalkilic *et al.* (2005) (15.3 and 36.7 cM) and Gulsen *et al.* (2010) (3 and 13 cM). A 3.3 Mb is now delimited and it has been found that a large number of resistance genes, organized in several clusters, which have been annotated.

The closest markers are being currently used for molecular marker-assisted selection in our triploid breeding program. However, it would be necessary to analyze many more hybrids to accurately estimate their linkage with the *ABSr* locus and to perform fine mapping with additional markers mined in the corresponding genomic sequence.

Due to the centromeric position of this gene and the recessivity of the resistant trait, the 2x × 2x breeding strategy appears much more efficient to obtain resistant hybrids than the interploid strategy when susceptible heterozygous parents need to be used as diploid gamete producer. Indeed, with the 2x × 2x, about 40% of resistant genotypes are obtained, whereas this percentage varies between 16.7% and 22.5% with the interploid strategy.

Global implication of the diploid gamete origin on triploid breeding strategy

Cultivar breeding implies selection on multiple traits, many of them having a complex genetic determinism. If restrictive traits such as ABS susceptibility can orient a strategy when heterozygous susceptible parent is used, the choice of the breeding scheme should be globally more conducted by the targeted kind of innovation. Assuming an SDR origin of 2*n* gametes in mandarins, sexual polyploidization may lead to lower average of heterozygosity restitution than interploid hybridization, whatever the segregation model considered for the doubled diploid (Marsden *et al.*, 1987). As heterozygosity and epistatic interactions are maintained for a great number of individuals in the progeny from interploid crosses with doubled-diploid, this triploid breeding strategy should be more efficient than 2*x* × 2*x* hybridization for developing new cultivars that are phenotypically close to the genitor giving the diploid gamete. Conversely, 2*x* × 2*x* hybridization should produce more polymorphic progenies, by creating larger number of new multilocus allelic combinations (David *et al.*, 1995), providing the opportunity to select innovative products within the perspective of market segmentation as a commercial strategy.

Perspectives

An alternative way for SNP genotyping in polyploids, based in Competitive Allele-Specific PCR, has been proposed and successfully applied in triploid populations. This approach is very powerful and particularly adapted for studies needing a limited number of markers (a few hundred). However, for pangenomic studies (particularly, in progeny of two highly heterozygous parents as in *Citrus*), methods based on next-generation sequencing such as Genotyping-by-sequencing (GBS) should be much more adapted. With enough sequencing coverage, these method should be applied efficiently to estimate allele doses in polyploid plants in a pangenomic way and should be the preeminent approach for marker-trait association studies both in diploid and polyploid citrus.

The mechanism of unreduced megagametophyte formation has been elucidated for several monoembryonic mandarin genotypes. Further studies on polyembryonic mandarin genotypes and other ancestral and secondary species, such as citron, pummelo, grapefruit, sweet orange, lemon or lime could determine whether the same mechanism or any other is involved in the unreduced female gamete formation in citrus germplasm. The use of 2*n* pollen, revealed at low frequency by a few works should also be explored to produce triploid progeny in 2*x* × 2*x* crosses. Moreover, further knowledge on environment influence as well as molecular insights on unreduced gamete formation in citrus would aid in altering frequencies of unreduced gamete formation.

Segregations for the single inherited character Alternaria brown spot (ABS) resistance have been analyzed in this study for several triploid recovery strategies and parental combinations. In addition, molecular markers have been developed for marker-assisted selection. With the genomic region linked to ABS resistance indentified, further sequence analyses would allow selecting markers closer to the ABS resistance gene for selection within many parental combinations. Several candidate genes have been identified within this region; therefore, fine mapping based on large progeny genotyped with additional markers and functional analyses of the genes remaining between newly identified flanking markers would permit identifying the gene(s) controlling ABS resistance in citrus. In the same way, inheritance in other agronomic traits should be studied, even those with complex determinism.

Conclusions

A new method for genotyping heterozygous polyploids using SNP markers by competitive-allele specific PCR (KASPar) has been developed

KASPar technique combined with analysis of the allele signal data is an alternative method to infer SNP allelic configurations in polyploid plants that offers a wider spectrum of genotyping possibilities. Moreover, quantitative analyses for correlation of the allele signals and the allele doses and sample clustering carried out in this work were powerful techniques for assigning allelic configurations. Therefore, it could also be used in the quantitative analysis of allele-specific expression.

Two alternative approaches identify SDR as the mechanism leading to unreduced gamete formation in mandarins, based on molecular marker heterozygosity restitution analysis

The first approach is based on functions of heterozygosity restitution along a chromosome in relation with locus-centromere distance. This strategy has been successfully applied in this study for 'Fortune' mandarin and allowed mapping the centromere on chromosome II of the current clementine's genetic map. This strategy was used to conclude that SDR, rather than FDR, is much more likely to be the mechanism underlying unreduced gamete formation in 'Fortune' mandarin and evidenced positive chiasma interference. This same strategy has been later applied by Aleza *et al.* (2012b) to map centromeres on all nine citrus chromosomes.

The second approach proposed to check FDR/SDR hypothesis is a maximum likelihood approach, based in the heterozygosity restitution of independent markers closely linked to centromere of different chromosomes. This method has the advantage that allows deducing the restitution mechanism at population level or even at individual level. Using this second approach, SDR has been identified as the restitution mechanism leading to unreduced gamete formation in a wide range of mandarin genotypes used as female parents in triploid breeding programs.

The monolocus inheritance and recessivity of the Alternaria brown spot (ABS) resistance was confirmed in triploid progeny and the *ABSr* locus was mapped in the chromosome III genetically close to the centromere

The monolocus inheritance of the resistance to ABS, proposed on the basis of diploid segregation studies was corroborated in triploid progeny by studying the segregations in several parent combinations and triploid recovery strategy. Taking advantage of the particular genetic structures of 2*n* gametes resulting from second division restitution (SDR), segregation analysis allowed mapping the *ABSr* locus genetically at 10.5 cM close to a centromere at a first step.

Bulk segregant analysis coupled with genome-scan using SNP markers mapped in the reference genetic map allowed locating the ABS resistance locus within a 13.1 cM area on chromosome III, corresponding to 15 Mb of the scaffold 3 in the current whole genome assembly.

Molecular markers flanking the *ABSr* locus were selected for use in marker-assisted selection (MAS)

Two flanking markers at 3.77 and 1.71 cM of the *ABSr* locus, as well as a third marker that did not exhibit any recombination are currently being used in marker-assisted selection in the IVIA mandarin breeding program. These markers should be used together for efficient early MAS for different parental combinations when the markers are heterozygous in the susceptible parent and polymorphic between the two parents. We are currently sequencing DNA fragments between these two markers to identify SNP loci that provide a useful allelic combination for a wider range of crosses of our mandarin breeding program.

Gene ontology reveals clusters of resistance genes in the *ABSr* locus: a first step towards identification of ABS resistance gene

The *ABSr* locus is included in a genomic region very rich in disease resistance homologous genes, organized in clusters. A probable function for the gene of interest is to encode a protein involved in ACT-toxin recognition, which would allow the toxin to cause cell death. With the genomic region containing *ABSr* locus mapped, the identification of the gene for ABS resistance will involve fine mapping with large diploid populations. For this purpose, SNP markers are currently being developed from each candidate gene. From the reduced set of candidate genes that would result from this fine genetic mapping, functional validation could be performed by genetic transformation or viral vector-induced gene silencing.

Annex

Other scientific publications and congress communications developed during the thesis

Publications

Aleza P, Cuenca J, Juárez J, Pina JA, Navarro L. 2010. 'Garbi' mandarin: a new late-maturing triploid hybrid. *HortScience* **45(1)**: 139-141.

Aleza P, Juárez J, Cuenca J, Ollitrault P, Navarro L. 2010. Recovery of citrus triploid hybrids by embryo rescue and flow cytometry from 2x x 2x sexual hybridisation and its application to extensive breeding programs. *Plant Cell Reports* **29(9)**: 1023-1034.

Aleza P, Juárez J, Cuenca J, Ollitrault P, Navarro L. 2012. Extensive citrus triploid hybrid production by 2x x 4x sexual hybridizations and parent-effect on the length of the juvenile phase. *Plant Cell Reports* **31(9)**: 1723-1735.

Cuenca J, Aleza P, Juárez J, Pina JA, Navarro L. 2010. 'Safor' mandarin: a new citrus mid-late triploid hybrid. *HortScience* **45(6)**: 977-980.

Ollitrault P, Terol J, Chen C, Federici C, Lotfy S, Hippolyte I, Ollitrault F, Bérard A, Chauveau A, Cuenca J, Costantino G, Kacar Y, Mu L, García-Lor A, Froelicher Y, Aleza P, Boland A, Billot C, Navarro L, Luro F, Roose M, Gmitter FG, Talón M, Brunel D. 2012. A reference genetic map of *C. clementina* Hort. ex Tan.; citrus evolution inferences from comparative mapping. *BMC Genomics* **13**: 593.

Congress comunications

Aleza P, Cuenca J, García A, Arregui JM, Juárez J, Froelicher Y, Luro F, Ollitrault P, Navarro L. 2010. Mejora genética de cítricos. Obtención de híbridos triploides. *V Congreso de Mejora Genética de Plantas. Madrid (España): 7-9 Julio 2010.*

Aleza P, Cuenca J, Juárez J, Ollitrault P, Navarro L. 2012. Differences in the genetic structure of citrus triploid hybrids recovered from 2*x* X 2*x* and 4*x* X 2*x* sexual hybridisations. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain. International Society of Citriculture.*

Aleza P, Cuenca J, Juárez J, Navarro L, Ollitrault P. 2012. Mechanism of 2*n* gametes formation and centromere mapping in citrus. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain. International Society of Citriculture.*

Aleza P., Juárez J., Cuenca J., Navarro L., Ollitrault P. 2012. Differences in the genetic structure of citrus triploid hybrids recovered from 2x x 2x and 4x x 2x sexual hybridizations. *19th EUCARPIA General Congress "Plant Breeding for future generations". Budapest (Hungary): 21-24 Mayo 2012. Plant breeding for future generations, 341.*

Bassimba DDM., Vicent A., Cuenca J., Aleza P., Navarro L. 2012. Screening for Alternaria Brown Spot resistance in the triploid mandarin breeding program in Spain. *XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.*

Cuenca J, Navarro L, Ollitrault P. 2008. Parental heterozygosity and origin of 2*n* gametes in mandarin: two key parameters for citrus triploid breeding projects. *18th EUCARPIA General*

Congress "Modern Variety Breeding for Present and Future Needs". Valencia (Spain). 9-12 Septiembre 2008. Modern Variety Breeding for Present and Future Needs, 697-698.

Cuenca J, Navarro L, Ollitrault P. 2009. Origin of 2*n* gametes in C. reticulata cv Fortune mandarin. *International Conference on Polyploidy, Hybridization and Biodiversity. Saint Malo (France). 17-20 Mayo 2009*

Cuenca J, Navarro L, Ollitrault P. 2010. Mecanismo de origen de los gametos 2*n* en el mandarino 'Fortune' y su implicación en las estructuras genéticas y fenotípicas de sus poblaciones de híbridos triploides. *V Congreso de Mejora Genética de Plantas. Madrid (España): 7-9 Julio 2010.*

Cuenca J, Aleza P, Navarro L, Ollitrault P. 2012. Assignment of SNP allelic configuration in polyploids using Competitive Allele-Specific PCR: application to triploid citrus progenies. *XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.*

Cuenca J, Aleza P, Iborra E, Vicent A, Ollitrault P, Navarro L. 2012. Location of a chromosome region linked to Alternaria Brown Spot resistance from the evaluation of triploid mandarin populations. *XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.*

Cuenca, J, Aleza, P, Juarez, J, Pina, JA, Navarro, L. 2012. Two new IVIA triploid hybrids of mandarin; IVIA-592 and IVIA-599. *XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.*

Cuenca J, Iborra E, Aleza P, Vicent A, Ollitrault P, Navarro L. 2012. Location of a chromosome region linked to Alternaria alternata pv. citri resistance from the evaluation of triploid mandarin populations. *19th EUCARPIA General Congress "Plant breeding for future generations". Budapest (Hungary): 21-24 Mayo 2012. Plant breeding for future generations 135-140*

Navarro L, Aleza P, Cuenca J, Juárez J, Pina JA, Pensabene-Bellavia G, Hernández M, Ortega C, Navarro A, Ortega V. 2009. The mandarin triploid breeding program in Spain. *Second International Citrus Biotechnology Symposium. Catania (Italy)*

Navarro L, Aleza P, Cuenca J, Juárez J, Pina JA. 2010. 'Garbí´ and `Safor´ mandarins, two new mandarin triploid hybrids obtained by embryo rescue and flow cytometry. *28th International Horticultural Congress. Lisboa (Portugal): 22-27 Agosto 2010.*

Navarro L, Aleza P, Cuenca J, Juárez J, Pina JA, Ortega C, Navarro A, Ortega V. 2012. The triploid mandarin breeding program in Spain. *XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.*

Ollitrault P, Terol J, Chen C, Federici CT, Lotfy S, Hippolyte I, Ollitrault F, Bérard A, Chauveau A, Costantino G, Kacar Y, Mu L, Cuenca J, Garcia A, Froelicher Y, Aleza P, Boland A, Billot C, Navarro L, Luro F, Roose ML, Gmitter FG Jr, Talon M, Brunel D. 2011. A reference linkage map of *C. clementina* based on SNPs, SSRs and Indels. *Plant and Animal Genome XIX Conference* San Diego, CA, January 15-19, 2011: P477

Ollitrault P, Terol J, Chen C, Federici CT, Lotfy S, Hippolyte I, Ollitrault F, Bérard A, Chauveau A, Cuenca J, Costantino G, Kacar Y, Mu L, García-Lor A, Froelicher Y, Aleza P, Boland A, Billot C, Navarro L, Luro F, Roose ML, Gmitter FG, Talón M, Brunel D. 2012. A reference genetic map of *Citrus clementina*; citrus evolution inferences from comparative

mapping. XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.

Sdiri S, Navarro P, Cuenca J, Pardo J, Salvador A. 2012. Postharvest behavior of new mandarins cultivars obtained in the IVIA. XII International Citrus Congress. Valence (Spain): 18-23 November 2012. International Society of Citriculture.

Sdiri S, Navarro P, Monterde A, Salvador A, Cuenca J, Aleza P, Ben Abda J. 2012. Postharvest behavior of 'Garbi' and 'Safor' - new triploid mandarins. Acta Hort. (ISHS) 945:255-262.

Literature cited

Agarwal M, Shrivastava N, Padh H. 2008. Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* **27**: 617-631.

Agüero J, Vives M, Velázquez K, Ruiz-Ruiz S, Juárez J, Navarro L, Moreno P, Guerri J. 2013. Citrus leaf blotch virus invades meristematic regions in *Nicotiana benthamiana* and citrus. *Molecular Plant Pathology*: 10.1111/mpp.12031.

Ahmadian A, Gharizadeh B, Gustafsson AC, Sterky F, Nyren P, Uhlen M, Lundeberg J. 2000. Single-nucleotide polymorphism analysis by pyrosequencing. *Analytical Biochemistry* **280**: 103-110.

Ajiro N, Miyamoto Y, Masunaka A, Tsuge T, Yamamoto M, Ohtani K, Fukumoto T, Gomi K, Peever TL, Izumi Y, Tada Y, Akimitsu K. 2010. Role of the host-selective ACT-toxin synthesis gene ACTTS2 encoding an enoyl-reductase in pathogenicity of the tangerine pathotype of *Alternaria alternata*. *Phytopathology* **100**: 120-126.

Akhunov E, Nicolet C, Dvorak J. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theoretical and Applied Genetics* **119**: 507-517.

Akimitsu K, Peever TL, Timmer LW. 2003. Molecular, ecological and evolutionary approaches to understanding Alternaria diseases of citrus. *Molecular Plant Pathology* **4**: 435-446.

Aleza P, Juárez J, Hernández M, Pina JA, Ollitrault P, Navarro L. 2009a. Recovery and characterization of a *Citrus clementina* Hort. ex Tan. 'Clemenules' haploid plant selected to establish the reference whole Citrus genome sequence. *BMC Plant Biology* **9**: 110.

Aleza P, Juárez J, Ollitrault P, Navarro L. 2009b. Production of tetraploid plants of non apomictic citrus genotypes. *Plant Cell Reports* **28**.

Aleza P, Cuenca J, Juárez J, Pina JA, Navarro L. 2010a. 'Garbi' mandarin: a new late-maturing triploid hybrid. *HortScience* **45(1)**: 139-141.

Aleza P, Juárez J, Cuenca J, Ollitrault P, Navarro L. 2010b. Recovery of citrus triploid hybrids by embryo rescue and flow cytometry from 2x x 2x sexual hybridisation and its application to extensive breeding programs. *Plant Cell Reports* **29(9)**: 1023-1034.

Aleza P, Froelicher Y, Schwarz S, Agustí M, Hernández M, Juárez J, Luro F, Morillon R, Navarro L, Ollitrault P. 2011. Tetraploidization events by chromosome doubling of nucellar cells are frequent in apomictic citrus and are dependent on genotype and environment. *Annals of Botany* **108(1)**: 37-50.

Aleza P, Cuenca J, Juárez J, Ollitrault P, Navarro L. 2012a. Differences in the genetic structure of citrus triploid hybrids recovered from 2*x* X 2*x* and 4*x* X 2*x* sexual hybridisations. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Aleza P, Cuenca J, Juárez J, Navarro L, Ollitrault P. 2012b. Mechanism of 2n gametes formation and centromere mapping in citrus. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Aleza P, Juárez J, Cuenca J, Ollitrault P, Navarro L. 2012c. Extensive citrus triploid hybrid production by 2x x 4x sexual hybridizations and parent-effect on the length of the juvenile phase. *Plant Cell Reports* **31(9)**: 1723-1735.

Aleza P, Juárez J, Hernández M, Ollitrault P, Navarro L. 2012d. Implementation of extensive citrus triploid breeding programs based on 4x x 2x sexual hybridizations. *Tree Genetics and Genomes* **8(6)**: 1293-1306.

Allario T, Brumos J, Colmenero-Flores J, Tadeo F, Froelicher Y, Talón M, Navarro L, Ollitrault P, Morillon R. 2011. Large changes in anatomy and physiology between diploid Rangpur lime (*Citrus limonia*) and its autotetraploid are not associated with large changes in leaf gene expression. *Journal of Experimental Botany* **62(8)**: 2507-2519.

Alquézar B, Zacarías L, Rodrigo MJ. 2009. Molecular and functional characterization of a novel chromoplast-specific lycopene ß-cyclase from Citrus and its relation to lycopene accumulation. *Journal of experimental botany* **60**: 1783-1797.

Alquézar B, Ayres AJ, Bento JMS, De Noronha NC Jr, Enrique A, Pedreira M, Peña L, Peris J, Postali JR, Rodríguez A, Spósito MB, Teixeira D, Wulff NA, Yamamoto PT. 2012. Repellent compositions and genetic approaches for controlling Huanglongbing. WIPO Patent Application WO/2011/050432.

Arumuganathan K, Earle DE. 1991. Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9(3)**: 208-218.

Asins MJ, Bernet GP, Ruiz C, Cambra M, Guerri J, Carbonell EA. 2004. QTL analysis of Citrus tristeza virus–citradia interaction. *Theoretical and Applied Genetics* **108**:603–611.

Barba-González R, Lokker AC, Lim KB, Ramanna MS, Van Tuyl J. 2004. Use of 2n gametes for the production of sexual polyploids from sterile Oriental— Asiatic hybrids of lilies (*Lilium*). *Theoretical and Applied Genetics* **109**:1125–1132.

Barba-González R, Lim KB, Ramanna MS, Visser R, Van Tuyl J. 2005. Occurrence of 2n gametes in the F1 hybrids of Oriental— Asiatic lilies (*Lilium*): Relevance to intergenomic recombination and backcrossing. *Euphytica* **143**: 67-73.

Barba-González R, Lim KB, Zhou S, Ramanna MS, Van Tuyl JM. 2008. Interspecfic hybridization in Lily: the use of 2n gametes in interspecific lily hybrids. In: JAT Silva, ed. *Floriculture, Ornamental and Plant Biotechnology: Advances and Topical Issues*, 1 edn. Global Science Books: UK.

Barcaccia G, Tavoletti S, Mariani A, Veronesi F. 2003. Occurrence, inheritance and use of reproductive mutants in alfalfa improvement. *Euphytica* **133**.

Barkley NA, Roose ML, Krueger RR, Federici CT. 2006. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theoetical and Applied Genetics* **112**: 1519-1531.

Barkley NA, Krueger RR, Federici CT, Roose ML. 2009. What phylogeny and gene genealogy analyses reveal about homoplasy in citrus microsatellite alleles. *Plant Systematics and Evolution* **282**.

Barone A, Gebhardt C, Frusciante L. 1995. Heterozygosity in 2n gametes of potato evaluated by RFLP markers. *Theoretical and Applied Genetics* **91**.

Barret HC. 1974. Colchicine-induced polyploidy in *Citrus*. *Botanical Gazette* **135(1)**: 29-41.

Barrett HC, Rhodes AM. 1976. A numerical taxonomic study of affinity relationships in cultivated Citrus and its close relatives. *Systematic Botany* **1**: 105-136.

Bastiaanssen HJM, van dB, Lindhout P, Jacobsen E, Ramanna MS. 1998. Postmeiotic restitution in 2*n*-egg formation of diploid potato. *Heredity* **81**: 20-27.

Bauer F, Elbers CC, Adan RA, Loos RJ, Onland-Moret NC, Grobbee DE, van Vliet-Ostaptchouk JV, Wijmenga C, van der Schow YT. 2009. Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference. *American Journal of Clinical Nutrition* **90**.

Bausher MG, Singh ND, Lee SB, Jansen RK, Daniell H. 2006. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biology* **6**: 21.

Becker A, Chao D, Zhang X, Salt DE, Baxter I. 2011. Bulk Segregant Analysis using Single Nucleotide Polymorphism microarrays. *Plos One* **6(1)**: e15993.

Bérard A, Le Paslier MC, Dardevet M, Exbrayat-Vinson F, Bonnin I, Cenci A, Haudry A, Brunel D, Ravel C. 2009. High-throughput single nucleotide polymorphism genotyping in wheat (*Triticum* spp.). *Plant Biotechnology Journal* **7(4)**: 364-374.

Bernet GP, Bretó MP, Asins MJ. 2004. Expressed sequence enrichment for candidate gene analysis of Citrus tristeza virus resistance. *Theoretical and Applied Genetics* **108**:592–602.

Bertrand C, Collar Y, Mackill DJ. 2008. Marker-assisted selection: an approach for precision plant breeding in thetwenty-first century. *Philosophical Transactions of the Royal Society B* **363**: 557-572.

Bhatia A, Roberts PD, Timmer LW. 2003. Evaluation of the alter-rater model for timing of fungicide applications for control of Alternaria Brown Spot of citrus. *Plant Disease* **87**: 1089-1093.

Bjornsson HT, Albert TJ, Ladd-Acosta C, Green RD, Rongione MA, Middle CM, Irizarry RA, Broman KW, Feinberg AP. 2008. SNP-specific array-based allele-specific expression analysis. *Genome Research* **18(5)**: 771-779.

Brauer MJ, Christianson CM, Pai DA, Dunham MJ. 2006. Mapping novel traits by array-assisted bulk segregant analysis in *Saccharomyces cerevisiae*. *Genetics* **173**: 1813-1816.

Bretagnolle F, Thompson JD. 1995. Tansley Review no-78 - Gametes with the somatic chromosome-number - mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist* **129**: 1-22.

Bretó MP, Ruiz C, Pina JA, Asíns MJ. 2001. The diversification of *Citrus clementina* Hort. ex Tan., a vegetatively propagated crop species. *Molecular Phylogenetics and Evolution* **21**: 285-293.

Broman KW, Rowe LB, Churchill GA, Paigen K. 2002. Crossover interference in the mouse. *Genetics* **160**.

Brownfield L, Kohler C. 2011. Unreduced gamete formation in plants: mechanisms and prospects. *Journal of Experimental Botany* **62**: 1659-1668.

Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Topsch S, Vos P, Salamini F, Schulze-Lefert P. 1997. The barley Mlo gene: a novel control element of plant pathogen resistance. *Cell* **88(5)**: 695-705.

Buso JA, Boiteux LS, Tai GCC, Peloquin SJ. 1999. Chromosome regions between centromeres and proximal crossovers are the physical sites of major effect loci for yield in potato: Genetic analysis employing meiotic mutants. *Proceedings of the National Academy of Sciences of the United States of America* **96(4)**: 1773-1778.

Buteler MI, Jarret RL, LaBonte DR. 1999. Sequence characterization of microsatellites in diploid and polyploid Ipomoea. *Theoretical and Applied Genetics* **99**.

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Recupero GR, Martin C. 2012. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* **24(3)**: 1242-1255.

Byers R, Harker D, Yourstone S, Maughan P, Udall J. 2012. Development and mapping of SNP assays in allotetraploid cotton. *Theoretical and Applied Genetics* **124**: 1201-1214.

Cai X, Xu SS. 2007. Meiosis-driven genome variation in plants. *Current Genomics* **8(3)**: 151-161.

Cameron JW, Frost HB. 1968. Genetics, breeding and nucellar embryony. In: W Reuther, LD Batcherlor, HJ Webber, eds. *The Citrus Industry, Vol II*. Univ. Calif. Press: Berkeley, 325-370.

Cameron JW, Bernett RH. 1978. Use of sexual tetraploid seed parents for production of triploid citrus hybrids. *HortScience* **13**: 167-169.

Canihos Y, Peever TL, Timmer LW. 1999. Temperature, leaf wetness, and isolate effects on infection of Minneola tangelo leaves by *Alternaria* spp. *Plant Disease* **83**: 429-433.

Cervera M, Juárez J, Navarro L, Peña L. 2005. Genetic transformation of mature citrus plants. *Methods in Molecular Biology* **286**: 177-188.

Chao CT. 2005. Pollination study of mandarins and the effect on seediness and fruit Size: Implications for seedless mandarin production. *HortScience* **40**: 362-365.

Chapot H. 1975. Citrus, Technical Monograph No.4. *Ciba-Geigy Agrochem*: Basle, Switzerland.

Chen ZJ. 2007. Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids. *Annual Review of Plant Biology* **58**: 377-406.

Chen C, Zhou P, Choi YA, Huang S, Gmitter FG. 2006. Mining and characterizing microsatellites from citrus ESTs. *Theoretical and Applied Genetics* **112**: 1248-1257.

Chen C, Lyon MT, O'Malley D, Federici CT, Gmitter J, Grosser JW, Chaparro JX, Roose ML, Gmitter FG, Jr 2008a. Origin and frequency of 2*n* gametes in *Citrus sinensis* x *Poncirus trifoliata* and their reciprocal crosses. *Plant Science* **174**: 1-8.

Chen C, Bowman KD, Choi YA, Dang PM, Rao MN, Huang S, Soneji JR, McCollum TG, Gmitter FG, Jr. 2008b. EST-SSR genetic maps for Citrus sinensis and Poncirus trifoliata. *Tree Genetics & Genomes* **4**: 1-10.

Chen W, Mingus J, Mammadov J, Backlund JE, Greene T, Thompson S, Kumpatla S. 2010. KASPar: a simple and cost-effective system for SNP genotyping. *Plant and Animal Genomes XVII Conference*: 194.

Cheng FS, Roose ML. 1995. Origin and inheritance of dwarfing by the Citrus rootstock *Poncirus trifoliata* 'Flying Dragon'. *Journal of the American Society for Horticultural Science* **120**: 286-291.

Comai L. 2003. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* **6**: 836-846.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21(18)**: 3674-3676.

Conicella C, Barone A, Del Giudice A, Frusciante L, Monti LM. 1991. Cytological evidences of SDR-FDR mixture in the formation of 2*n* eggs in a potato diploid clone. *Theoretical and Applied Genetics* **81(1)**:.59-63.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK. 2010. Using Environmental Correlations to Identify Loci Underlying Local Adaptation. *Genetics* **doi: 10.1534/genetics.110.114819**.

Copenhaver GP, Housworth EA, Stahl FW. 2002. Crossover Interference in Arabidopsis. *Genetics* **160**: 1631-1639.

Cortés AJ, Chavarro MC, Blair MW. 2011. SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics* **123(5)**: 827-845.

Crespel L, Gudin S. 2003. Evidence for the production of unreduced gametes by tetraploid *Rosa hybrida* L. *Euphytica* **133**: 65-69.

Cuenca J, Aleza P, Juárez J, Pina JA, Navarro L. 2010. 'Safor' mandarin: a new citrus mid-late triploid hybrid. *HortScience* **45**: 977-980.

Cuenca J, Froelicher Y, Aleza P, Juárez J, Navarro L, Ollitrault P. 2011. Multilocus half-tetrad analysis and centromere mapping in citrus: evidence of SDR mechanism for 2n megagametophyte production and partial chiasma interference in mandarin cv 'Fortune'. *Heredity* **107**: 462-470.

Cuenca J, Aleza P, Iborra E, Vicent A, Ollitrault P, Navarro L. 2012. Location of a chromosome region linked to Alternaria Brown Spot resistance from the evaluation of triploid mandarin populations. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Cuenca J, Aleza P, Navarro L, Ollitrault P. 2013a. Assignment of SNP allelic configuration in polyploids using Competitive Allele-Specific PCR: application to triploid citrus progenies. *Annals of Botany* **111**: 731-742.

Cuenca J, Aleza P, Vicent A, Brunel D, Ollitrault P, Navarro L. 2013b. Genetically based location from triploid populations and gene ontology of a 3.3-Mb genome region linked to Alternaria brown spot resistance in citrus reveal clusters of resistance genes. *PLoS ONE* **8(10)**: e76755.

Cuppen E. 2007. Genotyping by Allele-Specific Amplification (KASPar). *Cold Spring Harbor Protocols* **2007**: 172-173.

Da Silva J, Sorrells M. 1996. Linkage Analysis in Polyploids Using Molecular Markers. In: PP Jauhar, ed. *Methods of genome analysis in plants*. CRC Press: Boca Raton, 211-228.

Da Y, Jarrell VL, Wang TL, Fernando RL, Wheeler MB, Lewin HA. 1995. Multilocus analysis for gene centromere mapping using first polar bodies and secondary oocytes. *Genetics* **139(2)**: 1091-1097.

Dalkilic Z, Timmer LW, Gmitter FG. 2005. Linkage of an Alternaria disease resistance gene in mandarin hybrids with RAPD fragments. *Journal of the American Society for Horticultural Science* **130(2)**: 191-195.

Dambier D, Benyahia H, Pensabene-Bellavia G, Kacar YA, Froelicher Y, Belfalah Z, Lhou B, Handaji N, Printz B, Morillon R, Yesiloglu T, Navarro L, Ollitrault P. 2011. Somatic hybridization for citrus rootstock breeding: an effective tool to solve some important issues of the Mediterranean citrus industry. *Plant Cell Reports* **30**.

Dangl J, Jones JD. 2001. Plant pathogens and integrated defence responses to infection. *Nature* **411(6839)**: 826-833.

Danzmann RG, Bogart JP. 1983. Further evidence for a polymorphism in gametic segregation in the tetraploid treefrog *Hyla versicolor* using a glutamate oxaloacetic transaminase locus. *Genetics* **103(4)**: 753-769.

David JL, Boudec P, Gallais A. 1995. Quantitative genetics of 4x-2x hybrid populations with First-Division Restitution and 2nd-Division Restitution 2n gametes produced by diploid parents. *Genetics* **139**.

Davies FS, Albrigo LG. 1994. *Citrus*. Crops Production Science in Horticulture. CAB International, UK.

De Jong WS, De Jong DM, Bodis M. 2003. A fluorogenic 5 ' nuclease (TaqMan) assay to assess dosage of a marker tightly linked to red skin color in autotetraploid potato. *Theoretical and Applied Genetics* **107**.

De Souza MC, Stuchi ES, de Goes A. 2009. Evaluation of tangerine hybrid resistance to *Alternaria alternata*. *Scientia Horticulturae* **123**: 1-4.

De Storme N, Geelen D. 2011. The Arabidopsis mutant jason produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. *Plant Physiology* **155**: 1403-1415.

De Storme N, Geelen D. 2013. Sexual polyploidization in plants – cytological mechanisms and molecular regulation. *New Phytologist* **198**: 670-684.

D'Erfurth I, Jolivet S, Froger N, Catrice O, Novatchkova M, Simon M, Jenczewski E, Mercier R. 2008. Mutations in *AtPS1* (Arabidopsis thaliana Parallel Spindle 1) lead to the production of diploid pollen grains. *PLoS Genetics* **4**: e1000274.

D'Erfurth I, Jolivet S, Froger N, Catrice O, Novatchkova M, Mercier R. 2009. Turning meiosis into mitosis. *PLoS Biology* **7**: e1000124.

Devoto A, Piffanelli P, Nilsson I, Wallin E, Panstruga R, von Heijne G, Schulze-Lefert P. 1999. Topology, subcellular localization, and sequence diversity of the Mlo family in plants. *The Journal of Biological Chemistry* **274(49)**: 34993-35004.

Dewitte A, Van Laere K, Van Huylenbroeck J. 2012. Use of 2n fametes in plant breeding. In: DI Abdurakhmonov, ed. *Plant Breeding*. InTech.

Douches DS, Quiros CF. 1987. Use of 4x-2x crosses to determine gene centromere map distances of isozyme loci in *Solanum* species. *Genome* **29**: 519-527.

Douches DS, Quiros CF. 1988. Genetic strategies to determine the mode of 2n egg formation in diploid potatoes. *Euphytica* **38(3)**: 247-260.

Douches DS, Maas DL. 1998. Comparison of FDR and SDR derived tetraploid progeny from 2x x 4x crosses using haploids of *Solanum tuberosum* L. that produce mixed modes of 2n eggs. *Theoretical and Applied Genetics* **97**.

Drouaud J, Mercier R, Chelysheva L, Berard A, Falque M, Martin O, Zanni V, Brunel D, Mezard C. 2007. Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. *Plos Genetics* **3(6)**: e106. doi:10.1371/journal.pgen.0030106.

Edwards D, Batley J. 2010. Plant genome sequencing: applications for crop improvement. *Plant Biotechnology Journal* **8(1)**: 2-9.

Elena K. 2006. Alternaria brown spot of Minneola in Greece; evaluation of citrus species susceptibility. *European Journal of Plant Pathology* **115(2)**: 259-262.

Errico A, Cammareri M, Conicella C. 2005. Meiotic nuclear restitution mechanisms in a triploid lily. *Caryologia* **58(4)**: 367-373.

Esch E, Weber E. 2002. Investigation of crossover interference in barley (*Hordeum vulgare* L.) using the coefficient of coincidence. *Theoretical and Applied Genetics* **104(5)**: 786-796.

Esen A, Soost RK. 1971. Unexpected triploids in *Citrus*: their origin, identification, and possible use. *Journal of Heredity* **62(6)**: 329-333.

Esen A, Soost RK. 1973. Precocious development and germination of spontaneous triploid seeds in C*itrus*. *Journal of Heredity* **64(3)**: 147-154.

Esen A, Soost RK, Geraci G. 1978. Seed set, size, and development after 4x x 2x and 4x x 4x crosses in citrus. *Euphytica* **27**: 283-294.

Esen A, Soost RK, Geraci G. 1979. Genetic evidence for the origin of diploid megagametophytes in *Citrus*. *Journal of Heredity* **70**: 5-8.

Esselink GD, Nybom H, Vosman B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theoretical and Applied Genetics* **109**: 402-408.

Everts KL, Lacy ML. 1996. Factors influencing infection of onion leaves by *Alternaria porri* and subsequent lesion expansion. *Plant Disease* **80(3)**: 276-280.

Fan J, Oliphant A, Shen R, Kermani B, Garcia F, Gunderson K, Hansen M, Steemers F, Butler S, Deloukas P, Galver L, Hunt S, McBride C, Bibikova M, Rubano T, Chen J, Wickham E, Doucet D, Chang W, Campbell D, Zhang B, Kruglyak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J, Chee M. 2003. Highly parallel SNP genotyping. *Cold Spring Harbor Symposia on Quantitative Biology* **68**: 69-78.

Fang DQ, Federici CT, Roose ML. 1997. Development of molecular markers linked to a gene controlling fruit acidity in citrus. *Genome* **40**.

Fang DQ, Federici CT, Roose ML. 1998. A high-resolution linkage map of the citrus tristeza virus resistance gene region in *Poncirus trifoliata* (L.) Raf. *Genetics* **150(2)**: 883-890.

FAO. 2013. *www.faostat.fao.org.*

Fatta Del Bosco S, Matranga G, Geraci G. 1992. Micro and macro-sporogenesis of two triploid hybrids of citrus. *Proceedings of 7th International Citrus Congress. International Society of Citriculture, Acireale, Italy* **1**: 122-124.

Federici CT, Fang DQ, Scora RW, Roose ML. 1998. Phylogenetic relationships within the genus *Citrus* (*Rutaceae*) and related genera as revealed by RFLP and RAPD analysis. *Theoretical and Applied Genetics* **94**: 812-822.

Ferrante SP, Lucretti S, Reale S, De Patrizio A, Abbate L, Tusa N, Scarano M. 2010. Assessment of the origin of new citrus tetraploid hybrids (2n=4x) by means of SSR markers and PCR based dosage effects. *Euphytica* **173(2)**: 223-233.

Folimonov AS, Folimonova SY, Bar-Joseph M, Dawson WO. 2007. A stable RNA virus-based vector for citrus trees. *Virology* **368(1)**: 205-216.

Foss E, Lande R, Stahl F, Steinberg C. 1993. Chiasma interference as a function of genetic distance. *Genetics* **133(3)**: 681-691.

Friesen T, Faris J, Solomon P, Oliver RP. 2008. Host-specific toxins: effectors of necrotrophic pathogenicity. *Cellular Microbiology* **10(7)**: 1421-1428.

Froelicher Y, Dambier D, Costantino G, Lotfy S, Didout C, Beaumont V, Brottier P, Risterucci A, Luro F, Ollitrault P. 2007. Characterization of microsatellite markers in *Citrus reticulata* Blanco. *Molecular Ecology Note*.

Froelicher Y, Dambier D, Bassene JB, Costantino G, Lotfy S, Didout C, Beaumont V, Brottier P, Risterucci AM, Luro F, Ollitrault P. 2008. Characterization of microsatellite markers in mandarin orange (*Citrus reticulata* Blanco). *Molecular Ecology Resources* **8(1)**: 119-122.

Froelicher Y, Bouffin J, Dambier D, Ollitrault P. 2012. Triploid seedless mandarin breeding in France. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Frost HB. 1948. Genetics and breeding. In: W Reuther, HJ Webber, LD Batchelor, eds. *The Citrus Industry, Vol. 1*. University of California Press, Berkeley, USA, 817-914.

Frost HB, Soost RK. 1968. Seed reproduction, development of gametes and embryos. In: W Reuther, LD Batchelor, HJ Webber, eds. *The Citrus Industry, Vol. 2.* University of California Press, Berkeley, USA, 290-324.

Gallais A. 2003. *Quantitative Genetics and Breeding Methods in Autopolyploids Plants*, 2 edn. INRA: France.

García R, Asins MJ, Forner J, Carbonell EA. 1999. Genetic analysis of apomixis in *Citrus* and *Poncirus* by molecular markers. *Theoretical and Applied Genetics* **99(3-4)**: 511-518.

García-Lor A, Curk F, Snoussi-Trifa H, Morillon R, Ancillo G, Luro F, Navarro L, Ollitrault P. 2012. A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the 'true citrus fruit trees' group (*Citrinae, Rutaceae*) and the origin of cultivated species. *Annals of Botany*: doi: 10.1093/aob/mcs227.

García-Lor A, Ancillo G, Navarro L, Ollitrault P. 2013a. *Citrus* (*Rutaceae*) SNP markers based on Competitive Allele-Specific PCR; transferability across the *Aurantioideae* subfamily . *Applications in Plant Sciences* **1(4)**: doi:10.3732/apps.1200406.

García-Lor A, Luro F, Ancillo G, Ollitrault P, Navarro L. 2013b. Genetic diversity and population-structure analysis of mandarin germplasm by nuclear (SSR, indel) and mitochondrial markers. *Submitted* .

Geraci G, Esen A, Soost RK. 1975. Triploid progenies from 2x x 2x crosses of *Citrus* cultivars. *Journal of Heredity* **66**: 177-178.

Geraci G, De Pasquale F, Tusa N. 1977. Percentages of spontaneous triploids in progenies of diploid lemons and mandarins. *International Citrus Congress (2nd, Orlando, Florida), International Society of Citriculture* **2**: 596-597.

Gill KS, Lubbers EL, Gill BS, Raupp WJ, Cox TS. 1991. A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). *Genome* **34**: 361-374.

Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, Mézard C. 2011. Genome-wide crossover distribution in *Arabidopsis thaliana* meiosis reveals sex-specific patterns along chromosomes. *PLoS Genetics* **7**: e1002354.

Gmitter FG Jr, Ling XB, Deng XX. 1990. Induction of triploid citrus plants from endosperm calli in vitro. *Theoretical and Applied Genetics* **80**: 785-790.

Gmitter FG Jr, Ling XB. 1991. Embryogenesis in vitro and nonchimeric tetraploid plant recovery from undeveloped *Citrus* ovules treated with colchicine. *Journal of the American Society for Horticultural Science* **116(2)**: 317-321.

Gmitter FG Jr, Ling XB, Deng XX. 1991. Colchicine-induced polyploidy in *Citrus* embryogenic cultures, somatic embryos, and regenerated plantlets. *Plant Science* **74(1)**: 135-141.

Gmitter FG Jr, Grosser JW, Moore GA. 1992. Citrus. In: F Hammerschlag, RL Litz, eds. *Biotechnology of perennial crops*. CAB Intl.: Wallingford, Oxon, U.K., 335-369.

Gmitter FG Jr, Xiao SY, Huang S, Hu XL, Garnsey SM, Deng Z. 1996. A localized linkage map of the citrus tristeza virus resistance gene region. *Theoretical and Applied Genetics* **92**.

Gmitter FG, Jr., Chen C, Machado MA, de Souza AA, Ollitrault P, Froehlicher Y, Shimizu T. 2012. Citrus genomics. *Tree Genetics & Genomes* **8(3)**: 611-626.

Goes A, Montes de Oca AG, Reis RF. 2001. Ocurrencia de las mancha de *Alternaria* en mandarina 'Dancy' en el estado de Rio de Janeiro. *Fitopatologia Brasileira* **26**: 386.

Golmohammadi M, Andrew M, Peever TL, Peres NA, Timmer LW. 2006. Brown spot of tangerine hybrid cultivars Minneola, Page and Fortune caused by *Alternaria alternata* in Iran. *Plant Pathology* **55(4)**: 578.

Grant V. 1981. *Plant Speciation*, 2nd ed. Columbia University Press, New York, USA.

Grant W. 2010. A multicountry ecological study of risk-modifying factors for prostate cancer: apolipoprotein E epsilon4 as a risk factor and cereals as a risk reduction factor. *Anticancer Research* **30(1)**: 189-199.

Griffiths A, Miller J, Suzuki D, Lewontin R, Gelbart W. 1996. *An introduction to genetic analysis*, 6th ed. W. H. Freeman and Company, New York, USA.

Grosser JW, Gmitter FG Jr. 1990. Protoplast fusion and citrus improvement. *Plant Breeding Reviews* **8**: 339-374.

Grosser JW, Ollitrault P, Olivares-Fuster O. 2000. Somatic hybridization in citrus: An effective tool to facilitate variety improvement. *In Vitro Cellular & Developmental Biology - Plant* **36**: 434-449.

Grosser JW, An H, Calovic M, Lee D, Chen C, Vasconcellos M, Gmitter FG. 2010. Production of new allotetraploid and autotetraploid citrus breeding parents: focus on zipperskin mandarins. *HortScience* **45(8)**: 1160-1163.

Grosser JW, Gmitter FG Jr. 2011. Protoplast fusion for production of tetraploids and triploids: applications for scion and rootstock breeding in citrus. *Plant Cell Tissue and Organ Culture* **104**: 343-357.

Gruber JD, Colligan PB, Wolford JK. 2002. Estimation of single nucleotide polymorphism allele frequency in DNA pools by using Pyrosequencing. *Human Genetics* **110(5)**: 395-401.

Guest D, Brown J. 1997. Plant defences against pathogens. In: J Brown, H Ogle, eds. *Plant pathogens and plant diseases*. Rockvale Publications, 242-262.

Gulsen O, Roose ML. 2001. Chloroplast and nuclear genome analysis of the parentage of lemons. *J Amer Soc Hort Sci* **126**: 210-215.

Gulsen O, Uzun A, Canan I, Seday U, Canihos E. 2010. A new citrus linkage map based on SRAP, SSR, ISSR, POGP, RGA and RAPD markers. *Euphytica* **173(2)**: 265-277.

Guo M, Davis D, Birchler JA. 1996. Dosage effects on gene expression in a maize ploidy series. *Genetics* **142(4)**: 1349-1355.

Guo M, Rupe MA, Zinselmeier C, Habben J, Bowen BA, Smith OS. 2004. Allelic variation of gene expression in maize hybrids. *Plant Cell* **16(7)**: 1707-1716.

Guo WW, Deng XX, Yi HL. 2000. Somatic hybrids between navel orange (*Citrus sinensis*) and grapefruit (*C. paradisi*) for seedless triploid breeding. *Euphytica* **116**: 281-285.

Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW. 2012. Plant disease resistance genes: Current status and future directions. *Physiological and Molecular Plant Pathology* **78**: 51-65.

Gustafsson Å. 1935. Studies on the mechanisms of parthenogenesis. *Hereditas* **21**: 1-112.

GVA. 2013. *http://www.agricultura.gva.es/la-conselleria/estadisticas. Superficies y producciones de la Comunitat Valenciana (Principales cultivos).*

Hammond-Kosack K, Jones JD. 1997. Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology* **48**: 575-607.

Han Y, Khu D, Monteros MJ. 2012. High-resolution melting analysis for SNP genotyping and mapping in tetraploid alfalfa (*Medicago sativa* L.). *Molecular Breeding* **29(2)**: 489-501.

Harlan JR, DeWet J. 1975. On O. Winge and a prayer: the origins of polyploidy. *The Botanical Review* **41**: 361-390.

Hauber DP, Reeves A, Stack SM. 1999. Synapsis in a natural autotetraploid. *Genome* **42**: 936-949.

Hayashi M, Kato J, Ohashi H, Mii M. 2009. Unreduced 3x gamete formation of allotriploid hybrid derived from the cross of *Primula denticulata* (4x) x *P. rosea* (2x) as a causal factor for

producing pentaploid hybrids in the backcross with pollen of tetraploid *P. denticulata*. *Euphytica* **169**: 123-131.

Hearn CJ. 1984. Development of seedless orange and grapefruit cultivars through seed irradiation. *Journal of the American Society for Horticultural Science* **109**: 270-273.

Henry IM, Dilkes BP, Young K, Watson B, Wu H, Comai L. 2005. Aneuploidy and genetic variation in the Arabidopsis thaliana triploid response. *Genetics* **170**: 1979-1988.

Hensz RA. 1971. 'Star Ruby', a new deep-red-fleshed grapefruit variety with distinct tree characteristics. *Journal of the Rio Grande Valley Horticultural Society* **25**: 54-58.

Herrero M. 1929. Variedades de grapefruit. *Agricultura* **1**: 322-324.

Herrero R, Asins MJ, Carbonell EA, Navarro L. 1996. Genetic diversity in the orange subfamily Aurantioideae. I. Intraspecifies and intragenus genetic variability. *Theoretical and Applied Genetics* **92**: 599-609.

Herzing LBK, Cook EH Jr, Ledbetter DH. 2002. Allele-specific expression analysis by RNA-FISH demonstrates preferential maternal expression of UBE3A and imprint maintenance within 15q11-q13 duplications. *Human Molecular Genetics* **11(15)**: 1707-1718.

Hutten R, Scholberg E, Huigen DJ, Hermsen JGT, Jacobsen E. 1993. Analysis of dihaploid induction and production ability and seed parent x pollinator interaction in potato. *Euphytica* **72**: 61-64.

Hutton DG, Mayers PE. 1988. Brown spot of Murcott tangor caused by *Alternaria alternata* in Queensland. *Australasian Plant Pathololgy* **17**: 69-73.

Hyten DL, Smith JR, Frederick RD, Tucker ML, Song Q, Cregan PB. 2009. Bulked Segregant Analysis using the GoldenGate assay to locate the Rpp3 locus that confers resistance to Soybean Rust in soybean. *Crop Science* **49**: 265-271.

Ishikawa S, Komura D, Tsuji S, Nishimura K, Yamamoto S, Panda B, Huang J, Fukayama M, Jones KW, Aburatani H. 2005. Allelic dosage analysis with genotyping microarrays. *Biochemical and Biophysical Research Communications* **333(4)**: 1309-1314.

Iwamasa M. 1966. Study on the steriliy in genus *Citrus* with special reference to the seedlessness. *Bulletin of the Horticultural Research Station of Japan, Series B* **6**: 2-77.

Iwamasa M, Iwasaki T. 1963. On the sterility phenomenon caused by low temperatures in the Mexican lime *(Citrus aurantifolia* Swing.). *Bulletin of the Horticultural Research Station of Japan, Series B* **2**: 25-45.

Iwamasa M, Nito N, Ling JT. 1988. Intra and intergeneric hybridization in the orange subfamily, *Auranthioideae Proceedings of 6th International Citrus Congress. Balaban, PA: International Society of Citriculture*: 123-130.

Iwasaki T. 1943. On the big leaf and small leaf strain of trifoliate orange (*Poncirus trifoliata*). *Journal of the Horticultural Association of Japan* **14**: 302-305.

Jackson RC, Jackson JW. 1996. Gene segregation in autotetraploids: prediction from meiotic configurations. *American Journal of Botany* **83(6)**: 673-678.

Jaskani M, Husnain S, Bashir MA, Khan LA. 1996. Morphological description of citrus colchiploids. *Proceedings of the International Society of Citriculture* **1**: 130-132.

Jaskani MJ, Omura M, Khan IA. 2007. Cytogenetics. In: IA Khan, ed. *Citrus Genetics, Breeding and Biotechnology*. CABI: Wallingford, UK.

JinPing X, LiGeng C, Ming X, HaiLin L, WeiQi Y. 2009. Identification of AFLP fragments linked to seedlessness in Ponkan mandarin (*Citrus reticulata* Blanco) and conversion to SCAR markers. *Scientia Horticulturae* **121**: 505-510.

Johnson SL, Africa D, Horne S, Postlethwait JH. 1995. Half-tetrad analysis in zebrafish: mapping the ros mutation and the centromere of linkage group I. *Genetics* **139(4)**: 1727-1735.

Jongedijk E, Ramanna MS, Sawor Z, Hermsen JGT. 1991. Formation of first division restitution (FDR) 2n-megaspores through pseudohomotypic division in ds-1 (desynapsis) mutants of diploid potato: routine production of tetraploid progeny from 2xFDR x 2xFDR crosses. T*heoretical and Applied Genetics* **82(5)**: 645-656.

Jørgensen JH, Wolfe M. 1994. Genetics of powdery mildew resistance in barley. *Critical Reviews in Plant Sciences* **13**: 97-119.

Jørgensen JH. 1977. Spectrum of resistance conferred by*ml-o* powdery mildew resistance genes in barley. *Euphytica* **26**: 55-62.

Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C. 2003. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. *BMC Plant Biology* **3**: 9.

Kamiri M, Stift M, Srairi I, Costantino G, El Moussadik A, Hmyene A, Bakry F, Ollitrault P, Froelicher Y. 2011. Evidence for non-disomic inheritance in a *Citrus* interspecific tetraploid somatic hybrid between *C. reticulata* and *C. limon* using SSR markers and cytogenetic analysis. *Plant Cell Reports* **30(8)**: 1415-1425.

Kamiri M, Stift M, Srairi I, Chahidi B, Poullet T, Costantino G, Dambier D, Ollitrault P, Froelicher Y. 2012. Gametic configuration and inheritance of SSR markers in tetraploid interspecific and intergeneric citrus somatic hybrids. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Karlov GI, Khrustaleva LI, Lim KB, Van Tuyl JM. 1999. Homoeologous recombination in 2 n-gametes producing interspecific hybrids of *Lilium* (*Liliaceae*) studied by genomic in situ hybridization (GISH). *Genome* **42(4)**: 681-686.

Kauffman E, Gestl E, Kim D, Walker C, Hite J, Yan G, Rogan P, Johnson S, Cheng K. 1995. Microsatellite–centromere mapping in the Zebrafish (*Danio rerio)*. *Genomics* **30(2)**: 337-341.

Kepiro JL, Roose ML. 2010. AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* x *Poncirus trifoliata*. *Tree Genetics & Genomes* **6**: 1-11.

Khan IA, Kender WJ. 2007. Citrus breeding: introduction and objectives. In: IA Khan, ed. *Citrus Genetics, Breeding and Biotechnology*. CABI: Wallingford, UK.

Kijas J, Thomas MR, Fowler J, Roose ML. 1997. Integration of trinucleotide microsatellites into a linkage map of citrus. *Theoretical and Applied Genetics* **94**: 701-706.

Kirov G, Nikolov I, Georgieva L, Moskvina V, Owen MJ, O'Donovan MC. 2006. Pooled DNA genotyping on Affymetrix SNP genotyping arrays. *BMC Genomics* **7**: 27.

Kobayashi S, Ohgawara T, Saito W, Nakamura Y, Shimizu J. 1995. Fruit characteristics and pollen fertility of citrus somatic hybrids. *Journal of the Japanese Society for Horticultural Science* **64(2)**: 283-289.

Kohmoto K, Itoh Y, Shimomura N, Kondoh Y, Otani H, Kodama M, Nishimura S, Nakatsuka S. 1993. Isolation and biological-activities of 2 Host-Specific Toxins from the tangerine pathotype of *Alternaria alternata*. *Phytopathology* **83**: 495-502.

Kohmoto K, Akimitsu K, Otani H. 1991. Correlation of resistance and susceptibility of citrus to *Alternaria alternata* with sensitivity to Host-Specific Toxins. *Phytopathology* **81**: 719-722.

Koltunow AM, Brennan P, Protopsaltis S. 2000. Regeneration of West Indian limes (*Citrus aurantifolia*) containing genes for decreased seed set. *ISHS Acta Horticulturae. First International Citrus Biotechnology Symposium* **535**: 81-91.

Kono Y, Gardner JM, Takeuchi S. 1986. Structure of the host-selective toxins produced by a pathotype of *Alternaria citri* causing brown spot disease of mandarins. *Agricultural and Biological Chemistry* **50**: 801-804.

Krezdorn AH, Robinson FA. 1958. Unfruitfulness in the 'Orlando' tangelo. *Proceedings of the Florida State Horticultural Society* **21**: 86-91.

Krueger RR, Navarro L. 2007. Citrus germplasm resources. In: IA Khan. *Citrus Genetics, Breeding and Biotechnology*. CABI: Wallingford, UK.

Krug C. 1943. Chromosome number in the subfamily *Aurantioideae* with special reference to the genus *Citrus*. *Botanical Gazette* **104**: 602-611.

Kwok PY, Xiao M. 2003. SNP genotyping and molecular haplotyping of DNA pools. *Cold Spring Harbor Symposia on Quantitative Biology* **68**: 65-67.

Kwok PY. 2001. Methods for genotyping single nucleotide polymorphisms. *Annual Review of Genomics and Human Genetics* **2**: 235-258.

Ladaniya M. 2008. *Citrus fruit: Biology, Technology and Evaluation*. Elsevier Inc., India.

Lander E, Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185-199.

Landergott U, Naciri Y, Schneller JJ, Holderegger R. 2006. Allelic configuration and polysomic inheritance of highly variable microsatellites in tetraploid gynodioecious *Thymus praecox* agg. *Theoretical and Applied Genetics* **113(3)**: 453-465.

Lapin WK. 1937. Investigation on polyploidy in *Citrus*. *USSR All Union Scientific Research Institute* **1(4)**: 1-68.

Lavebratt C, Sengul S, Jansson M, Schalling M. 2004. Pyrosequencing-based SNP allele frequency estimation in DNA pools. *Human Mutation* **23(1)**: 92-97.

Lelley T, Mahmoud AA, Lein V. 1987. Genetics and cytology of unreduced gametes in cultivated rye (*Secale cereale* L.). Genome **29(4)**: 635-638.

Lewis CM, Knight J. 2012. Introduction to genetic association studies. *Cold Spring Harbor Protocols* **2012**: doi:10.1101/pdb.top068163.

Li DD, Shi W, Deng X. 2002. Agrobacterium-mediated transformation of embryogenic calluses of Ponkan mandarin and the regeneration of plants containing the chimeric ribonuclease gene. *Plant Cell Reports* **21**: 153-156.

Li DD, Shi W, Deng XX. 2003. Factors influencing *Agrobacterium*-mediated embryogenic callus transformation of Valencia sweet orange (*Citrus sinensis*) containing the *pTA29-barnase* gene. *Tree Physiology*. **23**: 1209-1215.

Li X, Xie R, Lu Z, Zhou Z. 2010. The origin of cultivated *Citrus* as inferred from internal transcribed spacer and chloroplast DNA sequence and Amplified Fragment Length Polymorphism fingerprints. *Journal of the American Society for Horticultural Science,* **135**: 341-350.

Lian J, Yin Y, Oliver-Bonet M, Liehr T, Ko E, Turek P, Sun F, Martin RH. 2008. Variation in crossover interference levels on individual chromosomes from human males. *Human Molecular Genetics* **17(17)**: 2583-2594.

Liang G, Xiong G, Guo Q, He Q, Li X. 2007. AFLP analysis and the taxonomy of *Citrus*. *Acta Horticulturae* **760**: 137-142.

Lim KB, Ramanna MS, de Jong JH, Jacobsen E, van Tuyl JM. 2001. Indeterminate meiotic restitution (IMR): a novel type of meiotic nuclear restitution mechanism detected in interspecific lily hybrids by GISH. *Theoretical and Applied Genetics* **103**: 219-230.

Lim KB, Shen TM, Barba-González R, Ramanna MS, van Tuyl JM. 2004. Occurrence of SDR 2*n* gametes in Lilium hybrids. *Breeding Science* **54(1)**: 13-18.

Lindner K, Seeb J, Habicht C, Knudsen KL, Knudsen K, Kretschmer E, Reedy, DJ, Spruell P, Allendorf FW. 2000. Gene-centromere mapping of 312 loci in pink salmon by half-tetrad analysis. *Genome* **43(3)**: 538-549.

Ling P, Duncan LW, Deng Z, Dunn D, Hu X, Huang S, Gmitter FG Jr. 2000. Inheritance of citrus nematode resistance and its linkage with molecular markers. *Theoretical and Applied Genetics* **100**: 1010-1017.

Liu B, Wendel JF. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. *Molecular Phylogenetics and Evolution* **29**: 365-379.

Liu S, Yeh C, Tang HM, Nettleton D, Schnable PS. 2012. Gene Mapping via Bulked Segregant RNA-Seq (BSR-Seq). *PLoS ONE* **7**: e36406.

Lo HS, Wang ZN, Hu Y, Yang HH, Gere S, Buetow KH, Lee MP. 2003. Allelic variation in gene expression is common in the human genome. *Genome Research* **13(8)**: 1855-1862.

Longley AE. 1925. Polycarpy, polyspory and polyploidy in *Citrus* and *Citrus* relatives. *Journal of the Washington Academy of Sciences* **15**: 347-357.

Lorang J, Sweat T, Wolpert T. 2007. Plant disease susceptibility conferred by a "resistance" gene. *Proceedings of the National Academy of Sciences of the United States of America* **104(37)**: 14861-14866.

Lorang J, Kidarsa T, Bradford CS, Gilbert B, Curtis M, Tzeng S, Maier CS, Wolpert TJ. 2012. Tricking the guard: exploiting plant defense for disease susceptibility. *Science* **338**: 659-662.

Lotfy S, Luro F, Carreel F, Froelicher Y, Rist D, Ollitrault P. 2003. Application of Cleaved Amplified Polymorphic Sequence method for analysis of cytoplasmic genome among *Aurantioideae* intergeneric somatic hybrids. *Journal of the American Society for Horticultural Science* **128**: 225-230.

Louzada EA, Grosser JW, Gmitter FG Jr, Nielsen B, Chandler JL, Deng XX, Tusa N. 1992. Eight new somatic hybrid citrus rootstocks with potential for improved disease resistance. *HortScience* **27**: 1033-1036.

Louzada ES, Grosser JW, Gmitter FG Jr. 1993. Intergeneric somatic hybridization of sexually incompatible parents: *Citrus sinensis* and *Atalantia ceylanica*. *Plant Cell Reports* **12**: 687-690.

Luciano M, Houlihan L, Harris S, Gow A, Hayward C, Starr J, Deary I. 2010. Association of existing and new candidate genes for anxiety, depression and personality traits in older people. *Behavior Genetics* **40**: 518-532.

Luro F, Lorieux M, Laigret F, Bove JM, Ollitrault P. 1994. Genetic mapping of an intergeneric Citrus hybrid using molecular markers. *Fruits* **49**: 404-408.

Luro F, Rist D, Ollitrault P. 2001. Evaluation of genetic relationships in *Citrus* genus by means of sequence tagged microsatellites. *Proceedings of the International Symposium on Molecular Markers for Characterizing Genotypes and Identifying Cultivars in Horticulture* **546**: 537-542.

Luro F, Maddy F, Jacquemond C, Froelicher Y, Morillon R, Rist D, Ollitrault P. 2004. Identification and evaluation of diplogyny in clementine (*Citrus clementina*) for use in breeding. *Acta Horticulturae* **663(2)**: 841-848.

Luro F, Costantino G, Terol J, Argout X, Allario T, Wincker P, Talon M, Ollitrault P, Morillon R. 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.

Mabberley DJ. 1997. A classiication for edible *Citrus*. *Telopea* **7**: 167-172.

MacQueen JB. 1966. Some Methods for classification and analysis of multivariate observations. *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability* **1**: 281-297.

Maekawa N, Yamamoto M, Nishimura S, Kohmoto K, Kuwata M. 1984. Studies on host-specific AF-toxins produced by Alternaria alternata strawberry pathotype causing Alternaria black spot of strawberry. I. Production of host-specific toxins and their biological activity. *Annals of the Phytopathology Society of Japan* **50**: 600-609.

MAGRAMA. 2013. *Superficies y producciones anuales de cultivos.*

Main BJ, Bickel RD, McIntyre LM, Graze RM, Calabrese PP, Nuzhdin SV. 2009. Allele-specific expression assays using Solexa. *BMC Genomics* **10**: 422.

Marsden JE, Schwager SJ, May B. 1987. Single-locus inheritance in the tetraploid treefrog hyla-versicolor with an analysis of expected progeny ratios in tetraploid organisms. *Genetics* **116(2)**: 299-311.

Martins FA, Carneiro PCS, Guimaraes CT, Magalhaes JV, Carneiro JES, Cruz CD. 2009. Distinction between plant samples according to allele dosage by semiquantitative polymerase chain reaction. *Genetics and Molecular Research* **8(1)**: 319-327.

Mason A, Nelson M, Yan G, Cowling W. 2011. Production of viable male unreduced gametes in *Brassica* interspecific hybrids is genotype specific and stimulated by cold temperatures. *BMC Plant Biology* **11(1)**: 103.

Mather K. 1936. Segregation and linkage in autotetraploids. *Journal of Genetics* **32**: 287-314.

Matsumoto R, Okudai N, Oiyama I, Takahara T, Yamamoto M, Asada K, Ishiuchi D, Murata H. 1991. New citrus cultivar "Tsunokaori". *Bulletin of the Fruit Tree Research Station* **21**: 59-65.

McCollum TG. 2007. Update of the USDA, ARS citrus scion improvement project. *Proc Fla State Hort Soc* **120**: 285-287.

McQuown E, Gall GAE, May B. 2002. Characterization and inheritance of six microsatellite loci in lake sturgeon. *Transactions of the American Fisheries Society* **131(2)**: 299-307.

Meaburn E, Butcher LM, Schalkwyk LC, Plomin R. 2006. Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans. *Nucleic Acids Research* **34(4)**: e28.

Mendes BJ, Filho FM, Farias P, Benedito V. 2001. Citrus somatic hybridization with potential for improved blight and CTV resistance. *In Vitro Cellular & Developmental Biology - Plant* **37**: 490-495.

Mendiburu AO, Peloquin SJ. 1976. Sexual polyploidization and depolyploidization: some terminology and definitions. *Theoretical and Applied Genetics* **48**: 137-143.

Mendiburu AO, Peloquin SJ. 1979. Gene-centromere mapping by 4*X*-2*X* matings in potatoes. *Theoretical and Applied Genetics* **54**: 177-180.

Michelmore RW, Paran I, Kesseli RV. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences, USA* **88**: 9828-9832.

Michelmore RW, Meyers BC. 1998. Clusters of resistance genes in plants evolve by divergent selection and a Birth-and-Death process. *Genome Research* **8**: 1113-1130.

Ming R, Liu ,S.C., Moore PH, Irvine J, Paterson AH. 2001. QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. *Genome Research* **11(12)**: 2075-2084.

Miwa T. 1951. Pollination, fertilization and fruit drop in *Citrus tamurana* Hort. *Bulletin of the Faculty of Agriculture, Miyazaki University* **2**: 1-67.

Moffett P, Farnham G, Peart J, Baulcombe DC. 2002. Interaction between domains of a plant NBS–LRR protein in disease resistance-related cell death. *The EMBO Journal* **21**: 4511-4519.

Mok DWS, Peloquin SJ. 1975. Three mechanisms of 2n pollen formation in diploid potatoes. *Canadian Journal of Genetics and Cytology* **17**: 217-225.

Moore GA, Tozlu I, Weber CA, Guy CL. 2000. Mapping quantitative trait loci for salt tolerance and cold tolerance in *Citrus grandis* (L.) Osb. x *Poncirus trifoliata* (L.) Raf. hybrid populations. *Acta Horticulturae* **535**: 37-45.

Mourao-Fo F, Gmitter FJ Jr, Grosser J. 1996. New tetraploid breeding parents for triploid seedless citrus cultivar development. *Fruit Varieties Journal.* **50**: 76-80.

Muller HJ. 1914. A new mode of segregation in Gregory's tetraploid primulas. *American Naturalist* **48**: 508-512.

Murad A, Lewis S, Smith G, Collin S, Chen L, Hamdy F, Neal D, Donovan J, Martin RM. 2009. PTGS2-899G>C and prostate cancer risk: a population-based nested case-control study (ProtecT) and a systematic review with meta-analysis. *Prostate Cancer and Prostatic Diseases* **12(3)**: 296-300.

Mustard MJ, Lynch SJ, Nelson RO. 1956. Pollination and floral studies of the Minneola tangelo. *Proceedings of the Florida State Horticultural Society* **69**: 277-281.

Nadeem K, Barba-González R, Ramanna MS, Arens P, Visser R, van Tuyl JM. 2010. Relevance of unilateral and bilateral sexual polyploidization in relation to intergenomic recombination and introgression in *Lilium* species hybrids. *Euphytica* **171(2)**: 157-173.

Nagy E, Lee T, Ramakrishna W, Xu Z, Klein P, SanMiguel P, Cheng C, Li J, Devos K, Schertz K, Dunkle L, Bennetzen J. 2007. Fine mapping of the Pc locus of Sorghum bicolor, a gene controlling the reaction to a fungal pathogen and its host-selective toxin. *Theoretical and Applied Genetics* **114(6)**: 961-970.

Nagy E, Bennetzen J. 2008. Pathogen corruption and site-directed recombination at a plant disease resistance gene cluster. *Genome Research* **18(12)**: 1918-1923.

Nakamura M. 1943. Cytological and ecological studies on the genus *Citrus* with special reference to the occurrence of sterile pollen grains. *Memoirs of the Faculty of Science and Agriculture, Taihoku Imperial University* **27**: 53-159.

Nakano M, Nesumi H, Yoshioka T, Yoshida T. 2001. Segregation of plants with undeveloped anthers among hybrids derived from the seed parent, 'Kiyomi' (*Citrus unshiu* x *C. sinensis*). *Journal of the Japanese Society for Horticultural Science* **70(5)**: 539-545.

Nakashima Y, Ueno T, Fukami H, Taga T, Masuda H, Osaki K, Otani H, Kohmoto K, Nishimura S. 1985. Isolation and structures of AK-toxin I and II, host-specific metabolites produced by *Alternaria alternata* Japanese pear pathotype. *Agricultural and Biological Chemistry* **49**: 807-815.

Nakatsuka S, Ueda K, Goto T, Yamamoto M, Nishimura S, Kohmoto K. 1986. Structure of AF-toxin II, one of the host-specific toxins, produced by *Alternaria alternata* strawberry pathotype. . *Tetrahedron Letters* **27**: 2753-2756.

Navarro L, Roistacher CN, Murashige T. 1975. Improvement of shoot-tip grafting in vitro for virus-free citrus. *Journal of the American Society for Horticultural Science* **100(5)**: 471-479.

Navarro L, Olivares-Fuster O, Juárez J, Aleza P, Pina JA, Ballester-Olmos J, Cervera M, Fagoaga C, Durán-Vila N, Peña L. 2004. Applications of biotechnology to citrus improvement in Spain. In: LG Albrigo, VG Sauco, eds.

Navarro L, Juárez J, Aleza P, Pina J, Olivares-Fuster O, Cuenca J, Julve J. 2005. Programa de obtención de híbridos triploides de mandarino en España. *Phytoma* **170**: 36-41.

Navarro L, Aleza P, Cuenca J, Juárez J, Pina J, Ortega C, Navarro A, Ortega V. 2012. The triploid mandarin breeding program in Spain. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Nesumi H, Nakano M, Yoshida T. 2001. Mode of inheritance on the abnormal development of impregnated ovules derived from Mukaku-kishu. *Journal of the Japanese Society of Horticultural Science* **70(2)**: 869.

Neve B, Froguel P, Corset L, Vaillant E, Vatin V, Boutin P. 2002. Rapid SNP allele frequency determination in genomic DNA pools by Pyrosequencing™. *BioTechniques* **32(5)**: 1138-1142.

Nicolosi E. 2007. Origin and taxonomy. In: IA Khan, ed. *Citrus Genetics, Breeding and Biotechnology*. CABI: Wallingford, UK.

Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E. 2000. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theoretical and Applied Genetics* **100**: 1155-1166.

Nijman IJ, Kuipers S, Verheul M, Guryev V, Cuppen E. 2008. A genome-wide SNP panel for mapping and association studies in the rat. *BMC Genomics* **9**: 95.

Nishiura M. 1964. Citrus breeding and bud selection in Japan. *American Society for Horticultural Science Citrus Breeding Symposium, Boulder, Colo, USA* **1964**: 79-83.

Ohgawara T, Kobayashi S, Ohgawara E, Uchimaya H, Ishi S. 1985. Somatic hybrid plants obtained by protoplast fusion between *Citrus sinensis* and *Poncirus trifoliata*. *Theoretical and Applied Genetics* **71**: 1-4.

Oiyama I, Okudai N, Takahara T. 1981. Ploidy levels of seedlings obtained from 2x X 4x crosses in citrus. *Proceedings of the International Society of Citriculture*: 32-34.

Okagaki RJ, Jacobs MS, Stec AO, Kynast RG, Buescher E, Rines HW, Vales MI, Riera-Lizarazu O, Schneerman M, Doyle G, Friedman KL, Staub RW, Weber DF, Kamps TL, Amarillo IFE, Chase CD, Bass HW, Phillips RL. 2008. Maize centromere mapping: a comparison of physical and genetic strategies. *Journal of Heredity* **99**: 85-93.

Olivares-Fuster O, Hernández-Garrido M, Guerri J, Navarro L. 2007. Plant somatic hybrid cytoplasmic DNA characterization by single-strand conformation polymorphism. *Tree Physiology.* **27**: 785-792.

Oliver R, Lazo G, Lutz J, et al. 2011. Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. *BMC Genomics* **12**: 77.

Ollitrault F, Terol J, Martin AA, Pina JA, Navarro L, Talon M, Ollitrault P. 2012. Development of indel markers from *Citrus clementina* (*Rutaceae*) BAC-end sequences and interspecific transferability in *Citrus*. *American Journal of Botany* **99(7)**: e268-73.

Ollitrault P, Vanel F, Froelicher Y, Dambier D. 1995. Creation of triploid citrus hybrids by electrofusion of haploid and diploid protoplasts. *Acta Horticulturae* **535**: 191-198.

Ollitrault P, Dambier D, Jacquemont C, Allent V, Luro F. 1996a. In vitro rescue and selection of spontaneous triploid by flow cytometry for easy peeler citrus breeding. *Proceedings of the International Society of Citriculture* **1**: 254-258.

Ollitrault P, Dambier D, Sudahono S, Luro F. 1996b. Somatic hybridisation in citrus: some new hybrid and alloplasmic plants *Proc Int Soc Citriculture*: 907-912.

Ollitrault P, Guo W, Grosser JW. 2007a. Somatic Hybridization. In: IA Khan, ed. *Citrus Genetics, Breeding and Biotechnology* . CABI: Wallingford, UK, 238.

Ollitrault P, Luro F, Yamamoto M. 2007b. Seedlessness and ploidy manipulations. In: IA Khan. *Citrus Genetics, Breeding and Biotechnology*. CABI: Wallingford, UK.

Ollitrault P, Dambier D, Luro F, Froelicher Y. 2008. Ploidy manipulation for breeding seedless triploid citrus. *Plant Breeding Reviews* **30**: 323-352.

Ollitrault P, Terol J, Chen C, Federici CT, Lotfy S, Hippolyte I, Ollitrault F, Bérard A, Chauveau A, Costantino G, Kacar Y, Mu L, Cuenca J, Garcia A, Froelicher Y, Aleza P, Boland A, Billot C, Navarro L, Luro F, Roose ML, Gmitter FG Jr, Talon M, Brunel D. 2011. A reference linkage map of *C. clementina* based on SNPs, SSRs and Indels. *Plant and Animal Genome XIX Conference* **San Diego, CA, January 15-19, 2011**: P477.

Ollitrault P, Navarro L. 2012. Citrus. In: ML Badenes, DH Byrne, eds. Springer US, 623-662.

Ollitrault P, Terol J, Chen C, Federici C, Lotfy S, Hippolyte I, Ollitrault F, Bérard A, Chauveau A, Cuenca J, Costantino G, Kacar Y, Mu L, García-Lor A, Froelicher Y, Aleza P, Boland A, Billot C, Navarro L, Luro F, Roose M, Gmitter FG, Talón M, Brunel D. 2012a. A reference genetic map of *C. clementina* Hort. ex Tan.; citrus evolution inferences from comparative mapping. *BMC Genomics* **13**: 593.

Ollitrault P, Terol J, García-Lor A, Bérard A, Chauveau A, Froelicher Y, Belzile C, Morillon R, Navarro L, Brunel D, Talon M. 2012b. SNP mining in *C. clementina* BAC end sequences; transferability in the *Citrus* genus (Rutaceae), phylogenetic inferences and perspectives for genetic mapping. *BMC Genomics* **13**: 13.

Ortiz R. 1997. Occurrence and Inheritance of 2n pollen in *Musa*. *Annals of Botany* **79**: 449-453.

Ortiz R, Peloquin SJ. 1991. Breeding for 2n-egg production in haploid-X-species 2x-potato hybrids. *American Potato Journal* **68(10)**: 691-703.

Osawa I. 1912. Cytological and experimental studies in *Citrus. Journal of the College of Agriculture,Tokyo University* **4**: 83-116.

Osborn TC, Chris Pires J, Birchler JA, Auger DL, Jeffery Chen Z, Lee H, Comai L, Madlung A, Doerge RW, Colot V, Martienssen RA. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics* **19**: 141-147.

Otani H, Kohmoto K, Kodama M. 1995. *Alternaria* toxins and their effects on host plants. *Canadian Journal of Botany* **73**: 453-458.

Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* **34**: 401-437.

Pang X, Hu C, Deng X. 2007. Phylogenetic relationships within *Citrus* and its related genera as inferred from AFLP markers. *Genetic Resources and Crop Evolution* **54**: 429-436.

Park T, Kim J, Hutten RCB, van Eck HJ, Jacobsen E, Visser RGF. 2007. Genetic positioning of centromeres using half-tetrad analysis in a 4x-2x cross population of potato. *Genetics* **176**.

Pécrix Y, Rallo G, Folzer H, Cigna M, Gudin S, Le Bris M. 2011. Polyploidization mechanisms: temperature environment can induce diploid gamete formation in *Rosa* sp. *Journal of Experimental Botany*. doi:10.1093/jxb/err052.

Peever TL, Canihos Y, Olsen L, Ibáñez A, Liu YC, Timmer LW. 1999. Population genetic structure and host specificity of *Alternaria* spp. causing brown spot of Minneola tangelo and rough lemon in Florida. *Phytopathology* **89(10)**: 851-860.

Pegg KG. 1966. Studies of a strain of *Alternaria citri* Pierce, the causal organism of brown spot of Emperor mandarin. *Queensland Journal of Agriculture and Animal Science* **23(1)**: 15-28.

Peloquin SJ, Yerk GL, Werner JE, Darmo E. 1989. Potato breeding with haploids and 2n gametes. *Genome* **31(2)**: 1000-1004.

Peloquin SJ, Boiteux LS, Simon PW, Jansky SH. 2008. A chromosome-specific estimate of transmission of heterozygosity by 2*n* gametes in potato. *Journal of Heredity* **99**: 177-181.

Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, Nagano Y. 2013. Phylogenetic relationships of *Citrus* and its relatives based on *matK* gene sequences. *PLoS ONE* **8**: e62574.

Peres NA, Timmer LW. 2006. Alternaria brown spot of citrus in Brazil: Evaluation of the Alter-Rater for spray timing and effects on yield of Murcott tangor. *Crop Protection* **25**: 454-460.

Predieri S. 2001. Induced mutation and tissue culture in fruits. *Plant Cell Tissue Organ Culture* **64(3)**: 185-210.

Qu DY, Zhu DW, Wang DS, Gao ZW. 1995. Genetic analysis of 2*n* gamete formation in potato. *Acta Horticulturae Sinica* **22(1)**: 61-66.

Qu S, Liu G, Zhou B, Bellizzi M, Zeng L, Dai L, Han B, Wang G. 2006. The broad-spectrum Blast resistance gene *Pi9* encodes a Nucleotide-Binding Site–Leucine-Rich Repeat protein and is a member of a multigene family in rice. *Genetics* **172(3)**: 1901-1914.

Quarrie SA, Lazic-Jancic V, Kovacevic D, Steed A, Pekic S. 1999. Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. *Journal of Experimental Botany* **50**: 1299-1306.

Raghuvanshi SS. 1962. Cytogenetical studies in genus *Citrus*. IV. Evolution in genus *Citrus*. *Cytologia* **27**: 172-188.

Ramanna MS. 1974. The origin of unreduced microspores due to aberrant cytokinesis in the meiocytes of potato and its genetic significance. *Euphytica* **23**: 20-30.

Ramanna MS. 1979. A re-examination of the mechanisms of 2*n* gametes formation in potato and its implications for breeding. *Euphytica* **28**: 537-561.

Ramanna MS, Jacobsen E. 2003. Relevance of sexual polyploidization for crop improvement - A review. *Euphytica* **133**: 3-8.

Ramsey J. 2007. Unreduced gametes and neopolyploids in natural populations of *Achillea borealis*. *Heredity* **98**: 143-150.

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29**: 467-501.

Ramsey J, Schemske DW. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* **33**: 589-639.

Raza H, Mumtaz-Khan M, Ali-Khan A. 2003. Review: seedlessness in *Citrus*. *International Journal of Agriculture and Biology* **5**: 388-391.

Recupero GR, Russo G, Recupero S. 2005. New promising citrus triploid hybrids selected from crosses between monoembryonic diploid female ant tetraploid male parents. *HortScience* **40**: 516-520.

Reinisch AJ, Dong JM, Brubaker CL, Stelly DM, Wendel JF, Paterson AH. 1994. A Detailed RFLP Map of Cotton, *Gossypium Hirsutum* X *Gossypium Barbadense*: Chromosome Organization and Evolution in a Disomic Polyploid Genome. *Genetics* **138(3)**: 829-847.

Reis RF, de Almeida TF, Stuchi ES, de Goes A. 2007. Susceptibility of citrus species to *Alternaria alternata*, the causal agent of the Alternaria Brown Spot. *Scientia Horticulturae* **113(4)**: 336-342.

Rhoades MM, Dempsey E. 1966. Induction of chromosome doubling at meiosis by the elongate gene in maize. *Genetics* **54(2)**: 505-522.

Ricci G, Astolfi A, Remondini D, Cipriani F, Formica S, Dondi A, Pession A. 2011. Pooled genome-wide analysis to identify novel risk loci for pediatric allergic asthma. *PLoS ONE* **6**: e16912.

Rickert AM, Premstaller A, Gebhardt C, Oefner PJ. 2002. Genotyping of SNPs in a polyploid genome by pyrosequencing™. *BioTechniques* **32(3)**: 592-593, 596-598 passim.

Ritland K. 2002. Estimation of gene frequency and heterozygosity from pooled samples. *Molecular Ecology Notes* **2**: 370-372.

Rockman MV, Kruglyak L. 2006. Genetics of global gene expression. *Nature Reviews Genetics* **7**: 862-872.

Ronaghi M, Uhlen M, Nyren P. 1998. A sequencing method based on real-time pyrophosphate. *Science* **281(5375)**: 363-365.

Ronald J, Akey JM, Whittle J, Smith EN, Yvert G, Kruglyak L. 2005. Simultaneous genotyping, gene-expression measurement, and detection of allele-specific expression with oligonucleotide arrays. *Genome Research* **15(2)**: 284-291.

Roose ML. 1988. Isozymes and DNA restriction fragment length polymorphisms in *Citrus* breeding and systematics. *Proceedings of the XVI International Citrus Congress, Tel Aviv, Israel* : 155-165-165.

Roose ML, Williams TA, Soost RK, Cameron JW. 2002. US Patents: USPP 15,461, USPP 15,703, USPP 16,289. University of California, Riverside, USA.

Roose ML, Williams TE. 2007. Mutation Breeding. In: IA Khan. *Citrus Genetics, Breeding and Biotechnology*. CABI: Wallingford, UK.

Rosenberg O. 1927. Die Semiheterotypische Teilung und Ihre Bedeutung für die Entstehung verdoppelter Chromozomenzahlen. *Hereditas* **8**: 305-358.

Rosso ML, Burleson S, Maupin L, Rainey K. 2011. Development of breeder-friendly markers for selection of *MIPS1* mutations in soybean. *Molecular Breeding* **28**: 127-132.

Rozowsky J, Abyzov A, Wang J, Alves P, Raha D, Harmanci A, Leng J, Bjornson R, Kong Y, Kitabayashi N, Bhardwaj N, Rubin M, Snyder M, Gerstein M. 2011. AlleleSeq: analysis of allele-specific expression and binding in a network framework. *Molecular Systems Biology* **7:** 522.

Ruiz C, Bretó MP, Asins MJ. 2000. A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. *Euphytica* **112**: 89-94.

Ruiz C, Asins MJ. 2003. Comparison between Poncirus and Citrus genetic linkage maps. *Theoretical and Applied Genetics* **106**: 826-36.

Saintenac C, Falque M, Martin OC, Paux E, Feuillet C, Sourdille P. 2009. Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). *Genetics* **181**: 393-403.

Saleh B, Allario T, Dambier D, Ollitrault P, Morillon R. 2008. Tetraploid citrus rootstocks are more tolerant to salt stress than diploid. *Comptes Rendus Biologies* **331(9)**: 703-710.

Sapolsky RJ, Hsie L, Berno A, Ghandour G, Mittmann M, Fan JB. 1999. High-throughput polymorphism screening and genotyping with high-density oligonucleotide arrays. *Genetic Analysis-Biomolecular Engineering* **14**: 187-192.

Schinor E, Pacheco C, Bastianel M, Azevedo F, Cristofani-Yali M. 2012. New tangors for Brazilian citriculture. *XII International Citrus Congress 19-23 December 2012 Valence, Spain*.

Scora RW. 1975. On the history and origin of citrus. *Bulletin of the Torrey Botanical Club* **102**: 369-375.

Scora RW. 1988. Biochemistry, taxonomy and evolution of modern cultivated citrus. *Citriculture: Proceedings of the Sixth International Citrus Congress, Tel Aviv, Israel, March 6-11, 1988* .

Sdiri S, Bermejo A, Aleza P, Navarro P, Salvador A. 2012. Phenolic composition, organic acids, sugars, vitamin C and antioxidant activity in the juice of two new triploid late-season mandarins. *Food Research International* **49(1)**: 462-468.

Shaw S, Carrasquillo M, Kashuk C, Puffenberger E, Chakravarti A. 1998. Allele frequency distributions in pooled DNA samples: applications to mapping complex disease genes. *Genome Research* **8(2)**: 111-123.

Shu QY. 2009. *Induced plant mutations in the genomics era.* Food and Agriculture Organization of the United Nations: Rome.

Singh S, Sharma SR, Kalia P, Deshmukh R, Kumar V, Sharma P, Sharma TR. 2012. Molecular mapping of the downy mildew resistance gene Ppa3 in cauliflower (*Brassica oleracea* var. botrytis L.). *Journal of Horticultural Science & Biotechnology* **87(2)**: 137-143.

Sjoling A, Walentinsson A, Nordlander C, Karlsson A, Behboudi A, Samuelson E, Levan G, Rohme D. 2005. Assessment of allele dosage at polymorphic microsatellite loci displaying allelic imbalance in tumors by means of quantitative competitive-polymerase chain reaction. *Cancer Genetics and Cytogenetics* **157(2)**: 97-103.

Solel Z. 1991. Alternaria Brown Spot on Minneola tangelos in Israel. *Plant Pathology* **40(1)**: 145-147.

Solel Z, Kimchi M. 1997. Susceptibility and resistance of citrus genotypes to *Alternaria alternata* pv. citri. *Journal of Phytopathology* **145**: 389-391.

Soltis DE, Soltis PS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Science* **12**: 243-273.

Soltis PS, Soltis DE. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences* **97**: 7051-7057.

Soltis PS, Soltis DE. 2009. The role of hybridization in plant apeciation. *Annual Review of Plant Biology* **60**: 561-588.

Soost R. 1987. Breeding citrus-genetics and nucellar embryony. In: AJ Abbott, R Atkin, eds. *Improving vegetatively propagated crops*. Academic Press, London, 83-110.

Spiegel-Roy P. 1990. Economic and agricultural impact of mutation breeding in fruit trees. *Mutation Breeding Reviews. Vienna* **5**: 215-235.

Spiegel-Roy P, Vardi A. 1992. 'Shani,' 'Orah' and 'Winola': Three new selections from our breeding program, *Proceedings of the 7th International Citrus Congress of the International Society of Citriculture. Acireale, Italy*: 72-73.

Ssebuliba RN, Tenkouano A, Pillay M. 2008. Male fertility and occurrence of 2*n* gametes in East African Highland bananas (*Musa* spp.). Euphytica **164(1)**: 53-63.

Starrantino A. 1992. Use of triploids for production of seedless cultivars in citrus improvement program. *Proceedings of the International Society of Citriculture* **1**: 117-121.

Starrantino A, Recupero GR. 1981. Citrus hybrids obtained *in vitro* from 2x females and 4x male. *Proceedings of the International Society of Citriculture* **1**: 31-32.

Stebbins GL. 1932. Cytology of Antennaria. II. Parthenogenetic species. *Botanical Gazzete* **94**: 322-344.

Stebbins GL. 1950. *Variation and Evolution in Plants*. Columbia University Press, New York, USA.

Stebbins GL. 1971. *Chromosomal Evolution in Higher Plants*. Addison-Wesley, London, UK.

Stelly DM, Peloquin SJ. 1986. Diploid female gametophyte formation in 24-chromosome potatoes - genetic-evidence for the prevalence of the 2nd meiotic division restitution mode. *Canadian Journal of Genetics and Cytology* **28(1)**: 101-108.

Steer S, Abkevich V, Gutin A, Cordell HJ, Gendall KL, Merriman ME, Rodger RA, Rowley KA, Chapman P, Gow P, Harrison AA, Highton J, Jones PBB, O'Donnell J, Stamp L, Fitzgerald L, Iliev D, Kouzmine A, Tran T, Skolnick MH, Timms KM, Lanchbury JS, Merriman TR. 2007. Genomic DNA pooling for whole-genome association scans in complex disease: empirical demonstration of efficacy in rheumatoid arthritis. *Genes and Immunity* **8(1)**: 57-68.

Stift M, Berenos C, Kuperus P, van Tienderen PH. 2008. Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to *Rorippa* (yellow cress) microsatellite data. *Genetics* **179(4)**: 2113-2123.

Stukenbrock E, McDonald BA. 2009. Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions. *Molecular Plant-Microbe Interactions* **22(4)**: 371-380.

Swart SH, Wingfield MJ, Swart WJ, Schute GC. 1998. Chemical control of Alternaria brown spot on Minneola tangelo in South Africa. *Annals of Applied Biology* **133**: 17-30-30.

Swingle WT. 1943. The botany of *Citrus* and its wild relatives in the orange subfamily (family Rutaceae, subfamily Aurantioideae). University of California Press, Berkeley, USA.

Swingle WT, Reece PC. 1967. The botany of *Citrus* and its wild relatives. In: W Reuther, H Webber, L Batchelor, eds. *The Citrus Industry, Vol 1*. University of California, Berkeley, USA, 190-430.

Swinnen S, Schaerlaekens K, Pais T, Claesen J, Hubmann G, Yang Y, Demeke M, Foulquié-Moreno M, Goovaerts A, Souvereyns K, Clement L, Dumortier F, Thevelein J. 2012. Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. *Genome Research* **22**: 975-984.

Sybenga J. 1996. Chromosome pairing affinity and quadrivalent formation in polyploids: do segmental allopolyploids exist? *Genome* **39(6)**: 1176-1184.

Szelinger S, Pearson J, Craig D. 2011. Microarray-based genome-wide association studies using pooled DNA. In: JK DiStefano, ed. *Disease Gene Identification: Methods and Protocols*. Humana Press, 49-60.

Tabone T, Mather DE, Hayden MJ. 2009. Temperature Switch PCR (TSP): Robust assay design for reliable amplification and genotyping of SNPs. *BMC Genomics* **10**: 580.

Tabor G, Kubisiak T, Klopfenstein N, Hall R, McNabb HS. 2000. Bulked Segregant Analysis Identifies Molecular Markers Linked to Melampsora medusae Resistance in Populus deltoides. *Phytopathology* **90(9)**: 1039-1042.

Tachikawa T, Tanaka Y, Hara S. 1961. Investigation on the breeding of *Citrus* trees. Study on the breeding of triploid *Citrus* varieties. *Bulletin of the Shizuoka Prefectural Citrus Experiment Station* **4**: 33-44.

Tai G, DeJong H. 1997. A comparison of performance of tetraploid progenies produced by diploid and their vegetatively doubled (tetraploid) counterpart parents. *Theoretical and Applied Genetics* **94**: 303-308.

Takahashi C, Leitch IJ, Ryan A, Bennett MD, Brandham PE. 1997. The use of genomic in situ hybridization (GISH) to show transmission of recombinant chromosomes by a partially fertile bigeneric hybrid, *Gasteria lutzii* × *Aloe aristata* (Aloaceae), to its progeny. *Chromosoma* **105(6)**: 342-348.

Tanaka T. 1977. Fundamental discussion of *Citrus* classification. *Studia Citrologica* **14**: 1-6.

Tang XY, Luo ZR. 2002. Cytology of 2n pollen formation in nonastringent persimmon. *Agricultural Sciences in China* **1(6)**: 680-683.

Tavoletti S, Bingham ET, Yandell BS, Veronesi F, Osborn TC. 1996. Half tetrad analysis in alfalfa using multiple restriction fragment length polymorphism markers. *Proceedings of the National Academy of Sciences of the United States of America* **93(20)**: 10918-10922.

Terol J, Conesa A, Colmenero JM, Cercos M, Tadeo F, Agusti J, Alos E, Andres F, Soler G, Brumos J, Iglesias DJ, Gotz S, Legaz F, Argout X, Courtois B, Ollitrault P, Dossat C, Wincker P, Morillon R, Talon M. 2007. Analysis of 13000 unique *Citrus* clusters associated with fruit quality, production and salinity tolerance. *BMC Genomics* **8**: 31.

Terol J, Naranjo MA, Ollitrault P, Talon M. 2008. Development of genomic resources for *Citrus clementina*: Characterization of three deep-coverage BAC libraries and analysis of 46,000 BAC end sequences. *BMC Genomics* **9**: 423.

Timmer LW, Peever TL, Solel Z, Akimitsu K. 2003. Alternaria diseases of citrus-novel pathosystems. *Phytopathologia Mediterranea* **42**: 99-112.

Ton LD, Krezdorn AH. 1967. Growth of pollen tubes in three incompatible varieties of *Citrus*. *Proceedings of American Society of Horticultural Science* **89**: 211-215.

Torres AM, Soost RK, Diendenhofen U. 1978. Leaf isozymes as genetic markers in *Citrus. American Journal of Botany* **65**: 869-881.

Torres AM, Soost RK, Mau-Lastovicka T. 1982. Citrus isozymes: Genetics and distinguishing nucellar from zygotic seedlings. *Journal of Heredity* **73**: 335-338.

Toyoda K, Collins N, Takahashi A, Shirasu K. 2002. Resistance and Susceptibility of Plants to Fungal Pathogens. *Transgenic Research* **11**: 567-582.

Trebbi D, Maccaferri M, De Heer P, Sorensen A, Giuliani S, Salvi S, Sanguineti M, Massi A, Van der-Vossen E, Tuberosa R. 2011. High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf). *Theoretical and Applied Genetics* **123**: 555-569.

Trick M, Long Y, Meng J, Bancroft I. 2009. Single nucleotide polymorphism (SNP) discovery in the polyploid Brassica napus using Solexa transcriptome sequencing. *Plant Biotechnology Journal* **7**: 334-346.

Trick M, Adamski NM, Mugford SG, Jiang C, Febrer M, Uauy C. 2012. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biology* **12**: 14.

Tsuge T, Harimoto Y, Akimitsu K, Ohtani K, Kodama M, Akagi Y, Egusa M, Yamamoto M, Otani H. 2012. Host-selective toxins produced by the plant pathogenic fungus *Alternaria alternata*. *FEMS Microbiology Reviews* **37(1)**: 44-66.

USDA. 2013. *http://www.fas.usda.gov/. Citrus: World Markets and Trade*.

Uzun A, Yesiloglu T. 2012. *Genetic diversity in Citrus*. Genetic Diversity in Plants. In M Caliskan ed., InTech, doi: 10.5772/32885.

Vardi A, Elhanati A, Frydman-Shani A, Neumann H. 1996. Strategies and considerations in mandarin improvement programmes. *Proceedings 8th Citrus Congress International. Society of Citriculture, Sun City, Florida,* USA. 109-112.

Vardi A, Levin I, Carmi N. 2008. Induction of seedlessness in *Citrus*: From classical techniques to emerging biotechnological approaches. *Journal of the American Society for Horticultural Science* **133**: 117-126.

Veilleux R. 1985. Diploid and polyploid gametes in crop plants: mechanisms of formation and utilization in plant breeding. *Plant Breeding Reviews* **3**: 253-288.

Vicent A, Armengol J, Sales R, García-Jiménez J, Alfaro-Lassala F. 2000. First Report of Alternaria Brown Spot of Citrus in Spain. *Plant Disease* **84(9)**: 1044.

Vicent A, Badal J, Asensi MJ, Sanz N, Armengol J, García-Jimenez J. 2004. Laboratory evaluation of citrus cultivars susceptibility and influence of fruit size on Fortune mandarin to infection by *Alternaria alternata* pv. citri. *European Journal of Plant Pathology* **110**: 245-251.

Vicent A, Armengol J, García-Jimenez J. 2007. Rain fastness and persistence of fungicides for control of Alternaria Brown Spot of citrus. *Plant Disease* **91**: 393-399.

Vicent A, Armengol J, García-Jiménez J. 2009. Protectant activity of reduced concentration copper sprays against Alternaria brown spot on 'Fortune' mandarin fruit in Spain. *Crop Protection* **28**: 1-6.

Voorrips RE, Gort G, Vosman B. 2011. Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC Bioinformatics* **12**: 172.

Vorsa N, Rowland LJ. 1997. Estimation of 2*n* megagametophyte heterozygosity in a diploid blueberry (*Vaccinium darrowi* Camp) clone using RAPDs. *Journal of Heredity* **88(5)**: 423-426.

Wagner A, Blackstone N, Cartwright P, Dick M, Misof B, Snow P, Wagner GP, Bartels J, Murtha M, Pendleton J. 1994. Surveys of gene families using Polymerase Chain-Reaction - PCR selection and PCR drift. *Systematic Biology* **43(2)**: 250-261.

Wakana A, Iwamasa M, Uemoto S. 1981. Seed development in relation to ploidy of zygotic embryo and endosperm in polyembryonic *Citrus*. *Proceedings of the International Society of Citriculture* **1**: 135-139.

Wang DY, Chang CJ. 1978. Triploid *Citrus* plantlets from endosperm. *Scientia Sinica* **21**: 822-827.

Wang XF, Li ZA, Tang KZ, Zhou CY, Yi L. 2010. First report of Alternaria Brown Spot of Citrus caused by *Alternaria alternata* in Yunnan Province, China. *Plant Disease* **94(3)**: 375.

Wasson J, Skolnick G, Love-Gregory L, Permutt MA. 2002. Assessing allele frequencies of single nucleotide polymorphisms in DNA pools by Pyrosequencing™ technology. *BioTechniques* **32(5)**: 1144-1146, 1148, 1150 passim.

Webber HJ. 1967. History and development of the *Citrus* industry. In: W Reuther, HJ Webber, LD Batchelor, eds. *The Citrus Industry, Vol. 1*. University of California Press: Berkeley, USA, 1-39.

Wendel JF, Doyle JJ. 2005. Polyploidy and evolution in plants. In: R Henry, ed. *Plant Diversity and Evolution*. CABI Publishing: Wallingford, UK, 97-117.

Wenger JW, Schwartz K, Sherlock G. 2010. Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*. *PLoS Genetics* **6**: e1000942.

Werner JE, Peloquin SJ. 1991. Occurence and mechanisms of 2*n* egg formation in 2x potato. *Genome* **34(6)**: 975-982.

Werner JE, Douches DS, Freyre R. 1992. Use of half-tetrad analysis to discriminate between 2 types of 2*n* egg formation in a potato haploid. *Genome* **35**: 741-745.

Werner JE, Peloquin SJ. 1990. Inheritance and 2 mechanisms of 2*n* egg formation in 2x potatoes. *Journal of Heredity* **81(5)**: 371-374.

Williams TE. 2012. Experiences in the development, release and commercialization of new irradiated citrus varieties from the citrus breeding program at the University of California, Riverside. *XII International Citrus Congress. 19-23 December 2012. Valence, Spain.*

Wu SS, Wu R, Ma ,C.X., Zeng ZB, Yang MC, Casella G. 2001a. A multivalent pairing model of linkage analysis in autotetraploids. *Genetics* **159(3)**: 1339-1350.

Wu R, Gallo-Meagher M, Littell ,R.C., Zeng ZB. 2001b. A general polyploid model for analyzing gene segregation in outcrossing tetraploid species. *Genetics* **159(2)**: 869-882.

Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, Perrier X, Ruiz M, Scalabrin S, Takita MA, Labadie K, Poulain J, Couloux A, Jabbari K, Cattonaro F, Del Fabbro C, Pinosio S, Zuccolo A, Chapman J, Grimwood J, Terol J, Tadeo FR, Estornell LH, Muñoz-Sanz JV, Ibánez V, Herrero-Ortega A, Aleza P, Pérez Pérez J, Ramón D, Brunel D, Luro F, Chen C, Farmerie WG, Desany B, Kodira C, Mohiuddin M, Harkins T, Fredrikson K, Burns P, Lomsadze A, Borodovsky M, Reforgiato G, Freitas-Astúa J, Quetier F, Navarro L, Roose M, Wincker P, Schmutz J, Morgante M, Machado MA, Talon M, Jaillon O, Ollitrault P, Gmitter F, Rokhsar D. 2013. Genome sequencing reveals complex history of admixture during citrus domestication. *Nature Genetics,* submitted.

Xu WW, Sleper DA, Chao S. 1995. Genome mapping of polyploid tall fescue (Festuca arundinacea Schreb.) with RFLP markers. *Theoretical and Applied Genetics* **91**: 947-955.

Xu Q, Chen L, Ruan X, Chen D, Zhu A, Chen C, Bertrand D, Jiao W, Hao B, Lyon MP, Chen J, Gao S, Xing F, Lan H, Chang J, Ge X, Lei Y, Hu Q, Miao Y, Wang L, Xiao S, Biswas MK, Zeng W, Guo F, Cao H, Yang X, Xu X, Cheng Y, Xu J, Liu J, Luo OJ, Tang Z, Guo W, Kuang H, Zhang H, Roose ML, Nagarajan N, Deng X, Ruan Y. 2013. The draft genome of sweet orange (*Citrus sinensis*). *Nature Genetics* **45**: 59-66.

Yamamoto M, Kobayashi S, Yoshioka T, Matsumoto R. 2001. Cybridization in *Citrus unshiu* Marc. (satsuma mandarin) and *C. sinensis* (L.) Osbeck (sweet orange). In: T Nagata, YPS Bajaj, eds. Springer Berlin Heidelberg, 124-138.

Yan W, Ye S, Jin Q, Zeng L, Peng Y, Yan D, Yang W, Yang D, He Z, Dong Y, Zhang X. 2010. Characterization and mapping of a novel mutant *sms1* (senescence and male sterility 1) in rice. *Journal of Genetics and Genomics* **37(1)**: 47-55.

Yang Z, Ye X, Molina J, Roose M, Mirkov T. 2003. Sequence analysis of a 282-kilobase region surrounding the citrus Tristeza virus resistance gene (*Ctv*) locus in *Poncirus trifoliata* L. Raf. *Plant Physiology* **131(2)**: 482-492.

Yang J, Zhu J, Williams RW. 2007. Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics* **23(12)**: 1527-1536.

Yang SL, Chung KR. 2012. The NADPH oxidase-mediated production of H_2O_2 and resistance to oxidative stress in the necrotrophic pathogen *Alternaria alternata* of citrus. *Molecular Plant Pathology* **13**: 900-914.

Yi H, Richards EJ. 2007. A cluster of disease resistance genes in *Arabidopsis* is coordinately regulated by transcriptional activation and RNA silencing. *The Plant Cell Online* **19**: 2929-2939.

Zaragoza S. 2007. *Aproximación a la historia de los cítricos. Origen, dispersión y evolución de su uso y cultivo*, Universidad Politécnica de Valencia. Valencia, España.

Zhang WC, Shao ZY, Lo JH, Deng CH, Deng SS, Wang F,. 1988. Investigation and utilization of citrus varietal resources in China. *Citriculture Sixth international citrus congress, Middle East, Tel Aviv, Israel, 6-11 March 1988* **1**: 291-294.

Zhang CH, Shen SX, Wang YH. 2000. Genetic analysis of 2*n* gamete formation in Chinese cabbage. *Acta Horticulturae Sinica* **27(4)**: 295-296.

Zhang CH, Xuan SX, Zhu HY. 2007. Studies on the formation and transmission of n+1 gametes of cabbage primary trisomics. *Agricultural Sciences in China,* **6(7)**: 786-791.

Zhang X, Borevitz JO. 2009. Global Analysis of Allele-Specific Expression in *Arabidopsis thaliana*. *Genetics* **182**: 943-954.

Zhang G, Chen L, Xiao G, Xiao Y, Chen X, Zhang S. 2009. Bulked segregant analysis to detect QTL related to heat tolerance in rice (*Oryza sativa* L.) using SSR markers. *Agricultural Sciences in China* **8(4)**: 482-487.

Zhao H, Speed TP, McPeek MS. 1995. Statistical analysis of crossover interference using the Chi-Square model. *Genetics* **139**: 1045-1056.

Zhao H, Speed T. 1996. On genetic map functions. *Genetics* **142**: 1369-1377.

Zhao H, Speed TP. 1998a. Statistical analysis of half-tetrads. *Genetics* **150(1)**: 473-485.

Zhao H, Speed TP. 1998b. Statistical analysis of ordered tetrads. *Genetics* **150(1)**: 459-472.

Zheng BB, Wu X, Ge X, Deng X, Grosser JW, Guo W. 2012. Comparative transcript profiling of a male sterile cybrid pummelo and its fertile type revealed altered gene expression related to flower development. *PLoS ONE* **7**: e43758.