DEPARTAMENTO GENETICA

DESARROLLO DE MARCADORES MOLECULARES DE APLICACIÓN EN GENOMICA Y PROGRAMAS DE MEJORA DE CITRICOS

CARLOS RUIZ LAFORA

UNIVERSITAT DE VALENCIA
Servei de Publicacions
2002

Aquesta Tesi Doctoral va ser presentada a València el dia 20 de Febrer de 2002 , davant un Tribunal format per:

- Luis Navarro Lucas
- Manuel Pérez Alonso
- Juan Former Valero
- José Luis García Martínez
- Amparo Latorre Castillo

Va ser dirigida per:
Prof. Dr. Mº José Asíns Cebrián

Depòsit legal:
I.S.B.N.: 84-370-5552-0
Edita: Universitat de València
Servei de Publicacions
C/ Artes Gráficas, 13 bajo
46010 València
Spain
Telèfon: 963864115

UNIVERSIDAD DE VALENCIA
FACULTAD DE CIENCIAS BIOLÓGICAS

DESARROLLO DE MARCADORES MOLECULARES DE APLICACIÓN EN GENÓMICA Y PROGRAMAS DE MEJORA DE CÍTRICOS

TESIS DOCTORAL

Carlos Ruiz Lafora

Valencia, 2002

UNIVERSITAT DE VALÈNCIA

FACULTAD DE CIENCIAS BIOLÓGICAS

Desarrollo de Marcadores Moleculares de Aplicación en Genómica y Programas de Mejora de Cítricos

Tesis Doctoral presentada por **Carlos Ruiz Lafora** para optar al grado de Doctor en Ciencias Biológicas por la Universitat de València.

Dirigida por **Mª José Asíns Cebrián**

Dª María José Asíns Cebrián. Investigadora del Instituto Valenciano de Investigaciones Agrarias.

CERTIFICA:

que D. Carlos Ruiz Lafora, Licenciado en Ciencias Biológicas, ha realizado bajo su dirección el proyecto de investigación con el título **"Desarrollo de Marcadores Moleculares de Aplicación en Genómica y Programas de Mejora de Cítricos".** Dicho proyecto está en condiciones de ser presentado para optar al grado de Doctor en Ciencias Biológicas.

Y para que así conste a todos los efectos del interesado se expide el presente certificado en Moncada a 5 de Noviembre de 2001.

Fdo. María José Asíns Cebrián

A mi familia
y a Susana

AGRADECIMIENTOS

En primer lugar quiero agradecer a mi directora, la Dra. Mª José Asíns, la oportunidad que me ha dado y todo lo que me ha enseñado en estos años. También quiero darle gracias por la paciencia que ha tenido en algunos momentos.

Me gustaría mostrar mi agradecimiento a algunos profesores que en el pasado despertaron en mí interés por la ciencia y en particular por la genética.

En los agradecimientos de una tesis realizada en el IVIA no puede faltar nunca el Dr. Emilio Carbonell por su ayuda para solucionar cualquier problema estadístico.

A todos mis amigos del laboratorio, especialmente a Pepe, Guille e Inma por haberme aguantado durante estos años y por todos los buenos ratos que hemos pasado juntos. Pero no me puedo olvidar de otros que han pasado por el laboratorio, como Sari o Pere, de los que he aprendido mucho. Gracias a Pau por su amistad y por que ha sido un gran apoyo en las largas tardes de invierno.

A toda la gente del departamento, Mónica, Mª Carmen, Susana, Pablo, Olga y todos los demás con los que he compartido muchos cafés y buenas charlas que servían para desconectar un poco del trabajo.

A mis amigos "no científicos" porque en la vida hay algo más que ciencia y junto a ellos he descubierto muchas cosas buenas y otras no tanto pero que te forman como persona. Espero que sigamos estando Siempre Unidos.

También quiero dar las gracias a mi familia que sin saber nunca a que me dedicaba, me han apoyado y siempre han estado a mi lado.

Hay una persona sin la cual no hubiera sido posible esta Tesis y con la cual he compartido muchos años y comparto todas mis ilusiones del futuro, gracias Susana.

No puedo dejar de mencionar al IVIA como institución, no sólo por haberme concedido las becas que he disfrutado durante estos años, sino también por permitir que me haya formado como investigador poniendo a mi disposición todos los medios necesarios.

Y por último quiero agradecerte a ti que estés leyendo esta tesis, por que no ha sido escrita para que se llene de polvo en una estantería olvidada.

Gracias a todos.

ÍNDICE

1. INTRODUCCIÓN

1.1 LOS CÍTRICOS

Los cítricos pertenecen a la familia de las Rutáceas, subfamilia Aurantoideas. Los seis géneros de cítricos verdaderos se encuentran dentro de la subtribu Citrinae y sólo tres de ellos tienen importancia comercial: *Poncirus* (naranjo trifoliado) da frutos no comestibles y es empleado en algunos países como patrón, *Fortunella* (Kumquat) y *Citrus*, estos dos últimos con especies de frutos comestibles. Todos ellos son árboles diploides (2n=18), excepto *F. hindsii* que es tetraploide (2n=36). Los frutos son hesperidios y las semillas, sin endospermo, presentan dos o más embriones nucelares. Sólo en el pummelo (*C. grandis*), el cidro (*C. medica*) y algunos mandarinos, como los clementinos, las semillas presentan un único embrión sexual llamado zigótico.

Desde el punto de vista económico, los cítricos poseen una gran importancia dado que su producción es mayor que la de cualquier otro frutal. La producción mundial de cítricos supera los 105 millones de toneladas, siendo superior a la del plátano, uva y manzana (FAO 2001).

España presenta una superficie de cultivo de unas 285.000 Ha, siendo el cuarto país productor de naranja para consumo en fresco del mundo, por detrás de Brasil, USA y China, con una producción superior a los 5 millones de toneladas en el año 2000 (FAO, 2001). Además, España es el primer país exportador de naranja para consumo directo, exportándose aproximadamente el 50% de la producción. Los grupos de cítricos más importantes desde el punto de vista económico son las naranjas, las mandarinas, los limones y los pomelos. La producción española estimada para el año 2001 es de 2.723.600 toneladas de naranjas, 1.641.500 toneladas de mandarinas y 951.800 toneladas de limones (Estadísticas del MAPA, 2001). Esta producción supone para España, además de un gran beneficio social, una parte importante de su PIB. De la adecuación al

mercado y de la calidad de esta producción depende, en cierta forma, este beneficio económico y social.

1.2 MEJORA GENÉTICA

La mejora genética vegetal se puede definir como "la ciencia cuyo objetivo es cambiar el genotipo, mejorándolo para un determinado medio y según el aprovechamiento para el que se vaya a destinar de acuerdo con las necesidades del hombre" (Frankel, 1958).

Smith (1986), la define como "El arte y la ciencia de mejorar el genotipo de las plantas en relación con su utilización económica".

Johnson (1981) la define de forma similar pero más concreta como "la utilización de un sistema organizado de manipulación genética para modificar una especie vegetal, con el fin de hacerla más útil o aceptable para un uso específico".

La mejora genética vegetal es esencialmente una elección hecha por el hombre de las mejores plantas escogidas dentro de una población en la cual exista variabilidad. En otras palabras, la selección es posible gracias a la existencia de variabilidad (Sánchez Monge, 1993).

A partir de todas estas definiciones, se pueden establecer las tres premisas más importantes para el planteamiento de cualquier programa de mejora genética vegetal:

1. La existencia de variabilidad o bien la capacidad para crearla se convierte de esta forma en el primer requisito de todo programa de mejora.

2. La capacidad de detectar dicha variabilidad, o lo que es lo mismo, la habilidad del mejorador para observar las diferencias que puedan tener

valor económico entre plantas de la misma especie y/o la existencia de técnicas capaces de medirlas.

3. La capacidad para manipular dicha variación para producir un nuevo cultivar estable.

Una de las dificultades para la mejora de cítricos es su compleja biología reproductiva. Muchas especies de cítricos son apomícticas y sus semillas producen embriones nucelares que limitan el desarrollo de embriones zigóticos. Esto dificulta la obtención de progenies de origen híbrido donde seleccionar para los caracteres deseados. Por otro lado, muchos genotipos de interés presentan esterilidad total o parcial en óvulos y/o polen, lo cual impide que puedan ser empleados como parentales en programas de mejora. Además, también se presentan casos de autoincopatibilidad e incompatibilidad entre genotipos, y la elevada heterozigosis (Herrero et al 1996) presente en la mayoría de las especies cultivadas, como el limón, naranjo dulce y clementino, provoca una segregación muy amplia de fenotipos, en su gran mayoría indeseables.

El periodo de juvenilidad es igualmente un factor limitante en la mejora de cítricos (Frost y Soots 1968). Muchos genotipos presentan un periodo de juvenilidad muy largo, de forma que requieren entre 5 y 8 años para empezar a producir. Dado que la mayoría de los caracteres que se desean mejorar se refieren al fruto, en la mejora de variedades, o a la propagación por semilla, en el caso de la mejora de patrones, la larga juvenilidad limita el número de generaciones y el número de plantas en los programas de mejora.

Como consecuencia de todas estas dificultades, el estudio genético de cítricos se ha visto muy limitado, y en la actualidad se conoce poco sobre el modo en que se heredan los caracteres más importantes de los cítricos, lo que dificulta la elección de los parentales y los métodos de mejora y selección más adecuados.

Dentro del área de la biotecnología, el inmenso desarrollo de las técnicas de marcadores moleculares ha permitido su aplicación en fases concretas del programa de mejora, que son limitantes para su eficacia y rentabilidad.

1.3 MARCADORES MOLECULARES

Un marcador es un carácter de determinación sencilla empleado en los programas de mejora cuando el carácter de interés que se desea seleccionar posee una determinación complicada. Los atributos ideales de un marcador son: (a) polimorfismo; (b) herencia mendeliana y no epistasia; (c) insesibilidad a la influencia y efectos ambientales; (d) ausencia de efectos en el desarrollo de la planta, es decir, comportamiento como un gen neutro; (e) simplicidad en la identificación y análisis; (f) codominancia; y (g) posibilidad de detección en las primeras fases del desarrollo de la planta.

Los primeros marcadores que se emplearon fueron aquellos que afectaban a la morfología de las plantas. Estos marcadores generalmente son debidos a la mutación de un gen que provoca un fenotipo marcadamente diferente al fenotipo normal. Su empleo es limitado debido a su escasez y a que suelen presentar dominancia, pleiotropía, epistasia e incluso algunos son deletéreos.

El desarrollo de los marcadores moleculares ha supuesto un importante avance en la mejora de plantas. Los marcadores moleculares presentan una serie de ventajas frente a los marcadores morfológicos: se encuentran en cualquier especie; se detectan en estadíos tempranos del desarrollo y en cualquier tipo de tejido; su análisis no depende de las condiciones experimentales; son fenotípicamente neutros; y además, la interacción entre ellos es pequeña o nula (ausencia de epistasia).

Los primeros marcadores moleculares que se desarrollaron fueron los isoenzimas, que son proteínas, y los RFLPs, que emplean ADN. Después, con la aparición de la técnica de reacción en cadena de la polimerasa (PCR), el tipo de marcadores aumentó considerablemente, y en la actualidad, hay un gran número de tipos de marcadores basados en la PCR. Cada tipo de marcador tiene unas características específicas que lo hacen adecuado para unas determinadas aplicaciones.

Isoenzimas

Los isoenzimas son enzimas que catalizan la misma reacción bioquímica, sin embargo, presentan algunas diferencias en la cadena de aminoácidos, las cuales se pueden detectar mediante electroforesis no desnaturalizante de un extracto crudo de proteínas y posterior tinción con un sustrato específico para cada enzima en cuestión (Soltis y Soltis 1989). Los isoenzimas son marcadores generalmente codominantes, fenotípicamente neutros, y la metodología empleada para su análisis es rápida, sencilla y económica; sin embargo, el número es muy limitado y se restringe a regiones codificantes del genoma. Otro problema es que su detección depende, en gran medida, del tejido que se analice.

Los isoenzimas fueron los primeros marcadores moleculares que se emplearon en mejora genética para múltiples aplicaciones, como por ejemplo la diferenciación entre individuos zigóticos y nucleares, sin embargo, en la actualidad están más en desuso.

RFLPs

Los RFLPs (Restriction Fragment Length Polymorphism) han sido el primer tipo de marcadores basados en ADN. La técnica consiste en la digestión del ADN con enzimas de restricción y una posterior separación de los fragmentos

mediante electroforesis en geles de agarosa. Tras esto, se realiza la transferencia del ADN a una membrana de nylon y se hibrida con un fragmento de ADN marcado o sonda, cuya detección da como resultado una serie de bandas que pueden dar lugar a polimorfismos.

Los RFLPs suelen segregar como marcadores codominantes, pueden cubrir cualquier región del genoma, son fenotípicamente neutros, no presentan interacciones epistásicas, son altamente reproducibles y su número es ilimitado. Su principal inconveniente es su elevado coste económico y de tiempo, además de su escaso rendimiento en polimorfismos, por lo que actualmente también están siendo desplazados por los marcadores basados en la PCR.

RAPDs

Los RAPDs (Random Amplified Polymorphic DNA) se basan en la amplificación del ADN mediante PCR con un sólo cebador de 10 nucleótidos de secuencia aleatoria. La reacción de amplificación da como resultado una serie de fragmentos de ADN de diferentes tamaños que pueden dar lugar a polimorfismos del tipo presencia/ausencia al ser separados por electroforesis. Se considera que la presencia de bandas con igual movilidad se corresponde con la misma posición del genoma, aunque esto no siempre es cierto cuando trabajamos con distintas especies.

Los RAPDs son virtualmente ilimitados en número y su coste es inferior al de los RFLPs. Presentan la ventaja de que la metodología es rápida y sencilla, lo que permite obtener un gran número de marcadores en poco tiempo. Otra ventaja es que no es necesaria ninguna información sobre el genoma de la especie a estudiar, ya que se emplean cebadores con secuencia aleatoria. Además de mostrar en algunos casos problemas de reproducibilidad, su gran inconveniente es que presentan dominancia; no se pueden diferenciar los

heterozigotos de los homozigotos para la presencia de banda, lo que les convierte en marcadores de calidad inferior en cuanto a contenido informativo.

SCARs

Los SCARs (Sequence Characterized Amplified Regions) se desarrollaron para solucionar algunos de los problemas que presentan los RAPDs. Consiste en clonar y secuenciar los fragmentos polimórficos de los RAPDs y diseñar cebadores más largos y específicos para esta zona del genoma. Con esto se consigue, en ocasiones, transformar el polimorfismo en codominante y se eliminan los problemas de reproducibilidad.

MICROSATÉLITES

Los microsatélites o SSRs (Simple Sequence Repeats) son secuencias de repeticiones en támdem cuyas unidades de repetición, constituidas entre uno y seis pares de bases, se presentan de forma consecutiva en un número que varía de dos a cientos de veces. Las repeticiones del tipo (TA/AT)n ó (GA/CT)n son las más abundantes en las plantas, siendo inferior el número medio de repeticiones que en animales. El nivel de polimorfismo de estos marcadores es el más alto cuando se compara con el resto de marcadores (Russell et al, 1997). Los polimorfismos se producen por un cambio en el número de repeticiones. La mayoría de los SSRs se comportan como codominantes y están distribuidos aleatoriamente en el genoma de los eucariotas. Esto los hace muy útiles en la construcción de mapas genéticos saturados y altamente informativos.

La metodología empleada consiste en la ampliación de los fragmentos de ADN utilizando cebadores diseñados a ambos lados del microsatélite. Los fragmentos generados se separan en un gel de acrilamida donde se ponen de manifiesto los posibles polimorfismos. El principal inconveniente de estos marcadores es la dificultad para obtenerlos en comparación con otros

marcadores de ADN (Rafalski y Tingey 1993), ya que se debe poseer información sobre la secuencia adyacente al microsatélite para diseñar los cebadores.

ISSRs

Los ISSRs (Inter Simple Sequence Repeat) también están basados en las secuencias repetidas en támdem, pero en este caso no se amplifica el microsatélite, sino que éste se utiliza para diseñar cebadores cuya secuencia posee una zona homóloga a un microsatélite y además se le añaden dos nucleótidos aleatoriamente. Con esto se consigue amplificar la zona genómica que se encuentra entre dos microsatélites que estén lo suficientemente cerca y que posean los dos nucleótidos a continuación de la repetición.

Este tipo de marcador no es tan polimórfico como los SSRs, ya que su polimorfismo no se basa en el cambio de número de repeticiones del microsatélite, sin embargo, tiene la ventaja de que no es necesario tener información previa de la secuencia.

IRAPs

Los IRAPs (Inter-Retrotransposon Amplified Polymorphism) son marcadores moleculares que se basan en la amplificación de la zona del genoma que se encuentra entre dos retrotransposones. Para ello se emplean cebadores específicos con una secuencia homóloga a la de un retrotransposón, de tal manera que, el extremo 3' del cebador se sitúe próximo al extremo de la secuencia del retrotransposón.

Los retrotransposones son elementos móviles que se encuentran de forma abundante en el genoma de las plantas. Se ha demostrado que se distribuyen de forma aleatoria por todo el genoma, aunque existen ciertas zonas donde se producen más inserciones, los llamados puntos calientes de inserción.

La ventaja de este tipo de marcador es que con pocas combinaciones de cebadores se puede conseguir una gran cantidad de bandas polimórficas, aunque la mayoría de ellas dominantes.

AFLPs

Los AFLPs (Amplification Fragment Length Polymorphisms) detectan fragmentos de restricción de ADN por medio de la amplificación por PCR. La tecnología de AFLPs consiste en una digestión del ADN con dos enzimas de restricción, uno con una diana de 6 nucleótidos y el otro con una diana de 4; después se realiza una ligación con unos adaptadores que se unen a los fragmentos de restricción; la amplificación de estos fragmentos se realiza con cebadores homólogos a la secuencia del adaptador y de la diana de restricción, además, se añaden 2 ó 3 nucleótidos selectivos para que se amplifique sólo un subconjunto de los fragmentos. Los amplificados se separan en un gel de poliacrilamida desnaturalizante y se visualiza por medio de autorradiografía o por fluorescencia en un secuenciador automático.

Las ventajas de los AFLPs es que generan muchas bandas por reacción, entre 50 y 100 fragmentos; no se requiere información sobre la secuencia de la especie; y son muy abundantes. Entre las desventajas, hay que destacar que los marcadores son dominantes y que fragmentos con el mismo tamaño pueden no ser homólogos.

S-SAPs

Los S-SAPs (Sequence-Specific Amplified Polymorphisms) son una combinación de los dos últimos tipos de marcadores expuestos. Son muy similares a los AFLPs, pero como cebadores se utiliza uno del AFLP y el otro es un cebador con secuencia homóloga a la secuencia de un retrotransposón.

Normalmente, se producen menos fragmentos que con los AFLPs aunque más polimorfismos.

1.4 UTILIDADES DE LOS MARCADORES

Los marcadores moleculares presentan múltiples aplicaciones en la mejora genética de plantas. Previamente hay que plantearse que tipo de marcador es el más apropiado a nuestras necesidades y para ello nos debemos plantear una serie de cuestiones como: ¿qué objetivos tenemos?, ¿qué nivel de resolución requerimos? y también debemos preguntarnos si nuestros recursos a nivel de personal, equipamiento y tiempo son los adecuados.

Identificación

El hecho de que los marcadores moleculares puedan ser analizados en estadíos tempranos del desarrollo les hace especialmente indicados para determinadas situaciones en las que se requiere la identificación temprana del origen y/o naturaleza de los diferentes individuos de una población. Algunas de estas situaciones son: distinción entre individuos zigóticos y nucelares en plantas que presentan apomixis, como sucede con los cítricos (Torres et al, 1982); separación de individuos provenientes de autofecundación o contaminantes en una población de origen híbrido (Arús et al, 1982); detección temprana de híbridos interespecíficos (Parfitt et al, 1985); o distinción entre diploides y dobles haploides en poblaciones de individuos obtenidos mediante cultivo de anteras o microsporas (Zamir et al, 1981).

También hay que mencionar el uso de marcadores en la distinción de cultivares comerciales. Los marcadores pueden ayudar a solucionar problemas tan importantes como la distinción de cultivares en estadíos tempranos no productivos, lo que en algunas especies resulta seriamente difícil, como en los clementinos, y es de gran importancia para la protección de los derechos del

obtentor o del viverista, de cara a posibles problemas de tipo legal que puedan surgir.

Es necesario destacar aquí un uso rutinario, pero fundamental, de los marcadores como es la comprobación de los genotipos de plantas sometidas a diversos procesos que pueden dar lugar a errores en su denominación posterior, tales como: importación, exportación, multiplicación, conservación, etc.

Localización de genes

Algunos caracteres de interés agronómico se encuentran bajo el control de uno o pocos genes (Gottlieb, 1986). Este suele ser el caso de resistencia a enfermedades, autoincompatibilidad, y otros caracteres relacionados con distintos aspectos de la morfología de los diferentes órganos de la planta. Para obtener marcadores ligados a caracteres monogénicos se suelen realizar análisis de cosegregación, utilizando la estrategia de segregación masal. Una vez que ha sido establecida la asociación entre el gen y los marcadores ligados, estos pueden ser usados en mejora genética (selección asistida por marcadores) y en la clonación basada en el mapeado fino del gen.

También se pueden identificar los genes que controlan los caracteres cuantitativos o QTLs (Quantitative Trait Locus) si previamente disponemos de un mapa de ligamiento que cubra ampliamente el genoma de los genotipos parentales. El estudio de la cosegregación entre marcadores y caracteres cuantitativos en poblaciones de mejora, permite la detección de zonas del cromosoma en las que se encuentran uno o más genes que afectan al carácter de interés,así como estimar los efectos de los alelos en dichos QTLs e incluso estudiar el efecto de las distintas condiciones de cultivo y presencia de patógenos sobre los efectos génicos. Dado que la mayoría de los caracteres de interés agronómico son cuantitativos, el análisis de QTL tiene una gran relevancia en la mejora de plantas.

Selección asistida por marcadores

La transferencia de caracteres de interés agronómico desde una especie donadora (silvestre) a una receptora (cultivada) se ha venido realizando tradicionalmente por el método del retrocruzamiento: tras la obtención de una generación F1 proveniente del cruzamiento entre los dos parentales (donador y recurrente), esta F1 y las subsiguientes generaciones se cruzan con el parental recurrente tantas veces como sea necesario hasta conseguir un individuo con el mismo genotipo que el parental recurrente, con la excepción del carácter que ha sido transferido desde el donador. Según Allard (1960), para completar el proceso son necesarias al menos cinco o seis generaciones de retrocruzamiento y, según en que casos, el proceso puede ser muy largo y costoso.

Este método se puede simplificar considerablemente con el uso de marcadores moleculares: por un lado, la selección para un marcador asociado al carácter, que además de poderse realizar en etapas tempranas del desarrollo de la planta, suele ser un proceso menos laborioso que la evaluación del carácter; y por otro lado, la selección para marcadores moleculares no ligados al gen de interés, puede contribuir a la reducción de hasta la mitad del número de generaciones de retrocruzamiento necesarias para la recuperación del genotipo parental recurrente (Tanksley et al, 1989). Además, los marcadores moleculares también son de gran utilidad para la eliminación de genes con efectos deletéreos que puedan encontrarse ligados al carácter de interés en la especie donadora (Young y Tanksley, 1995), eliminación que es muy difícil de conseguir a base de retrocruzamientos (Michelmore, 1995).

La aplicación de los marcadores moleculares a la mejora genética puede resultar especialmente ventajosa en especies con largos periodos de generación, como es el caso de las leñosas. En estas especies, con periodos de juvenilidad que rondan los 5-7 años (Torres, 1983), la reducción del número de

generaciones, y sobre todo de gastos de mantenimiento del material vegetal que se consigue con el uso de los marcadores, puede transformar un proceso de duración excesiva en otro de duración más razonable.

Mapa genético

El mapeo genético tiene como objetivo localizar marcadores dentro de un cromosoma y colocarlos en sus posiciones relativas de unos con respecto a otros. La localización cromosómica de cada marcador constituye su locus. Distintos alelos de un mismo marcador mapean en el mismo locus. El mapeo genético utiliza la frecuencia de recombinación entre dos loci para determinar la distancia cromosómica que hay entre ellos. Al aumentar la distancia, la probabilidad de recombinación es mayor.

La unidad de distancia genética es el centiMorgan (cM), que corresponde al 1 por 100 de recombinación. La función más simple para calcular la distancia es la de Haldane (1919), que asume no interferencia, es decir, que la recombinación en un intervalo es independiente de la recombinación en otro intervalo. La función Kosambi (1944) permite interferencia parcial, por lo que es más precisa.

Dado que la construcción de un mapa genético con más de tres loci es computacionalmente complicado, se han desarrollado programas informáticos para ello, tales como MAPMAKER/EXP (Lander et al. 1987), JOINMAP (Stam yvan Ooijen 1995), etc. Estos programas permiten el análisis multipunto de los loci ligados, con lo que se consigue una mayor exactitud en la ordenación de los marcadores.

Los cítricos presentan algunas características que facilitan la construcción de los mapas genéticos. Son diploides y presentan un número haploide de cromosomas relativamente bajo, n=9, y un genoma pequeño, 1C=0.62 pg

(Guerra 1984). Por otro lado, las hibridaciones intergenéticas e interespecíficas son empleadas comúnmente por los mejoradores. Estos híbridos son heterozigotos para un gran porcentaje de loci.

Como ya se ha dicho anteriormente, los mapas genéticos son esenciales en la localización de genes y QTLs. Otro de sus posibles usos es la comparación de los genomas de distintas especies para estudiar la colinearidad de los genes y aprovechar los conocimientos generados por proyectos "genoma" de una especie en especies relacionadas.

En esta introducción se ha realizado una rápida revisión sobre los principales tipos de marcadores moleculares y algunas de sus múltiples utilidades, centrándonos en la problemática de los cítricos y cómo la aplicación de marcadores moleculares puede ayudar a superar gran parte de las limitaciones existentes en su mejora. A continuación se detallan los objetivos concretos de esta Tesis doctoral.

2. OBJETIVOS

Como ya se ha mencionado anteriormente, la mejora genética de cítricos no se ha desarrollado mucho debido al problema que supone trabajar con especies leñosas, que además presentan una biología reproductiva compleja. Por ello, y para avanzar en este tema nos planteamos el desarrollo de nuevos marcadores moleculares y su utilización en diversas etapas de la mejora genética de cítricos.

El punto de partida de todo programa de mejora de cítricos originado por cruzamientos controlados con algún genotipo apomíctico es la diferenciación entre individuos zigóticos y nucelares.

Los nuevos marcadores se han empleado en la construcción de mapas de ligamiento en 3 especies. Estos mapas podrán servir como punto de partida para nuevos estudios genéticos sobre localización de genes, tanto cualitativos como cuantitativos, para selección asistida por marcadores y para iniciar estudios de genómica comparada.

Otra aplicación de los marcadores generados fue el estudio de la naturaleza molecular de la variación natural dentro del grupo de los clementinos, lo que permitirá diseñar métodos de mejora que emulen su variabilidad en la naturaleza.

Por todo ello, los objetivos de la presente tesis son los siguientes:

- Buscar un método sencillo de evaluación genética para la selección eficaz de individuos nucelares y zigóticos en familias segregantes.
- Desarrollar marcadores moleculares con alto contenido informativo y nivel de polimorfismo, de análisis rápido, sencillo y repetible para saturar los mapas genéticos de cítricos.
- Estudiar la distribución genómica de retrotransposones en distintas especies de cítricos.

- Realizar un análisis comparativo de los mapas genéticos de *Citrus aurantium*, *Citrus volkameriana* y *Poncirus trifoliata*.
- Estudiar comparativamente los mecanismos genéticos implicados en la variación somática de los clementinos.

A QUICK METHODOLOGY TO IDENTIFY SEXUAL SEEDLINGS IN CITRUS BREEDING PROGRAMS USING SSR MARKERS

CARLOS RUIZ, Mª PAZ BRETÓ y Mª JOSÉ ASÍNS

Instituto Valenciano de Investigaciones Agrarias (I.V.I.A.)

Apartado Oficial, E-46113 Moncada (Valencia).

Euphytica **112: 89-94 (2000)**

ABSTRACT:

In citrus breeding and genetics, it is very important to distinguish between zygotic and nucellar seedlings in order to eliminate unwanted genotypes. Usually, isozyme markers have been employed to determine the genetic origin of young plants. In this work we propose the use of SSR markers as an alternative methodology and compare them with isozymes in this kind of screenings. Two different populations were analysed: one derives from an interspecific cross and the other from selfing. We conclude that, in most cases, microsatellites are more efficient than isozymic markers to identify the sexual origin of citrus seedlings, given their higher level of polymorphism and the scarce number of polymorphic isozymes in some populations. We describe a quick and efficient methodology for SSR analysis, including a fast DNA extraction in microcentrifuge tubes, and visualization through silver staining, which eliminates the need for a labelling step.

KEY WORDS: Citrus, Microsatellites, Zygotic, Molecular markers, Fruit breading, Isozymes

INTRODUCTION

Citrus are one of the major fruit crops in the world. Citrus species are trees with persistent leaves, hesperidium fruits, and seeds without endosperm, and often with two or more nucellar embryos. Many members of the genus *Citrus* and some closely related genera belonging to Rutaceae reproduce apomictically by seed, through nucellar embryony (Frost, 1943). Since nucellar embryos develop asexually by ordinary mitotic division of cells of the nucellus and no male gamete contribute to their formation, nucellar seedlings are identical to the seed parent. Nucellar embryony has very important consequences for evolution, breeding, and culture of citrus fruits. Citrus are almost universally propagated by budding onto seedling rootstocks. Uniformity of rootstock genotypes is essential for reliable performance following budding and orchard establishment. Therefore, propagation of citrus rootstocks depends upon the production of clonal plants from nucellar seedlings. Most rootstock cultivars are polyembryonic, producing seeds that contain both nucellar and zygotic embryos. The proportion of zygotic embryos that develop and give rise to seedlings varies with both genotype and environment (Cameron and Frost, 1968; Cameron, 1979) and uniformity is seldom complete. Off-type seedlings are generally rogued in the nursery, but the efficiency of such roguing is not known for many rootstocks. Anderson et al. (1991) found, by isozyme analysis, that roguing based primarily on the size and growth habit of the seedlings is effective in removing some, but not all, zygotic seedlings. They found zygotic rootstocks escaped from the roguing in two of the three groves studied, and in some instances an apparent incompatibility was developing in young trees. In any case, zygotic seedlings may affect the performance of trees budded on them (Weber, 1932; Roose and Traugh, 1988). Therefore, one of the most important traits to be considered within the breeding programs for citrus rootstocks is apomictic reproduction by seed, i.e. that the new genotype yield the least number of zygotic seedlings. But apomixis is also an

important problem in citrus breeding because it greatly limits the genetic variability obtained by controlled pollination. In most crossing combinations it would not be possible to determine the genetic origin of young seedlings from morphology. Yet, such determination is essential in order for the breeder to avoid the 5- to 10-year costs of growing and maintaining until fruiting unwanted nucellar seedlings that are genetically identical to already available seed parents. Thus, zygotic seedlings, sexually originated, have to be identified to quickly discard the nucellar ones; otherwise costs would make citrus breeding programs unaffordable. Molecular markers are very useful tools to help plant breeders in facing this and other issues: they allow the study of genetic variation, the search of new sources of agronomically important genes, the identification of cultivars and the efficient management of segregant generations during plant breeding programs through marker-assisted selection (Asíns et al., 1998).

A molecular marker is a genetic difference that may exist between two plants, visualized by means of biochemical methods and located at a certain chromosomal position. When the genotype for one or more marker loci is known for two plants, inferences can be made about the genotypes that may appear in the progeny. Isozyme analysis is the usual technique in the separation of zygotic from nucellar plants among progenies (Iglesias et al., 1974; Spiegel-Roy et al., 1977; Torres et al., 1978, 1982; Ashari et al., 1988; Xiang and Roose, 1988; Anderson et al., 1991; Mestre et al., 1997), since it is a fast and cheap methodology. Isozyme variation is limited and there are few isozymic loci available for marker analysis. A fingerprinting probe has also been reported for application in the separation of zygotic from nucellar progeny (Orford et al., 1995). That methodology is expensive and time consuming, because it requires DNA extractions, digestions, transfer and southern blot hybridization. The advent of the polymerase chain reaction (PCR) technique (Mullis et al., 1986) has accelerated the development of new DNA marker systems, such as

microsatellites or SSRs (Simple Sequence Repeat Pomorphism). Whenever SSRs have been compared to other systems they have always revealed the highest levels of polymorphism (Russell et al., 1997). Their nature (stretches of tandemly repeated short nucleotide motifs) produces high levels of length polymorphism, which results from variation in the number of repeated units at the locus (Litt and Lutty, 1989; Tautz, 1989; Weber and May 1989; Morgante and Olivieri, 1993).

This article describes a new, quick and efficient method to identify the sexual origin of citrus seedlings based on the analysis of SSRs without radioactive-labelling.

MATERIAL AND METHODS

Plant material

Two populations of citrus were screened in order to distinguish zygotic from nucellar descents and to confirm the self-pollination or hybrid origin of the plants: one was a cross between a single descent of the selfing of *Poncirus trifoliata* (L.) Raf. var. "Flying Dragon" and the tangor "Ortanique" (*Citrus reticulata* (Blanco) x *C. sinensis* (L.) Osb.) (PxO population), consisting of 46 seedlings. The other one was derived by selfing *Fortunella crassifolia* Swing. (Fc population) and consisted of 106 seedlings. Plant material proceeds from the citrus germplasm bank at IVIA.

Isozyme analysis

Five enzymatic systems were analysed according to Asíns et al. (1995): Phosphoglucoisomerase (PGI), phosphoglucomutase (PGM), glutamate oxaloacetate transaminase (GOT), 6-phosphogluconate dehydrogenase (6PG) and peroxidase (PRX). For the Fc population we referred to a previous study

showing that *Fortunella crassifolia* is heterozygous for *Pgm, Pgi* and *Got-1*, and that it has a complex pattern with potentially polymorphic bands for PRX and 6-PG (Herrero et al.,1996). Only PGM, PGI and 6PG were analysed in the PxO family.

DNA minipreps in microcentrifuge tubes

The tissue for the DNA extraction is collected by using the microcentrifuge tube cap to obtain 8 leaf punches from each sample. Tubes are placed on ice and the tissue is homogenised with a screw-driver after adding a little liquid N_2 in the tube; then, resuspended in 0.5 ml extraction buffer (100 mM Tris, 50 mM EDTA, 500 mM NaCl, 10 mM β-mercaptoethanol, pH 8.0), 33 µl 20 % SDS are added and tubes are vigorously shaked. Then, tubes are incubated for 10 min at 65 ºC, and 166 µl 5M potassium acetate added and mixed gently. The tubes are left on ice for 20 min, centrifuged in a microcentrifuge at top speed and the supernatant filtered through Miracloth and transferred to a new tube. Isopropanol is added to fill up the tube. Tubes are then mixed by inversion, left at room temperature for 10 min and centrifuged 5 min. Supernatant is discarded and the pellet is washed with 70% ethanol and resuspended in 100 µl TE. Tubes are centrifuged 10 min and the supernatant transferred to a clean tube. 11 µl 3M sodium acetate and 72 µl isopropanol are added, tubes mixed by inversion and centrifuged 3 min. Supernatant is decanted and pellet washed with ethanol as before, dried in Speed-Vac and resuspended in 30 µl of TE.

Microsatellite analysis

Nine microsatellites described in Kijas et al. (1997) were analysed in the genitors of the families: five of them showed more than one band in *F. crassifolia* (TAA41, TAA45, TAA33, CAC23 and TAA27) and were screened in the whole Fc population. Only marker TAA41 was analysed in the whole PxO population.

PCR amplifications were performed as described in Kijas et al. (1997) with minor modifications, using 300 ng of genomic DNA per 25 μl reaction, in a MJ Research thermocycler. PCR products were mixed with 6 μl of 5X loading buffer (50% v/v Glycerol, 1x TAE, 10% v/v saturated Bromophenol Blue, 0.2% w/v Xylene Cyanole) and analysed by electrophoresis in sequencing-type 10% polyacrylamide gels (acrylamide: N,N'-methylenebisacrylamide, 29:1) under nondenaturing conditions, in 1X TBE buffer (90 mM Tris borate, 2 mM EDTA, pH 8.3). The gel was run at a constant 40 W, 45ºC for 2-3 h in a Genomyx LR (Genomyx Corporation) or at free temperature in a S2 Sequencing System (Gibco BRL/Life Technologies). Gels were fixed to one of the glass plates and silver stained according to Beidler et al. (1982), with minor modifications: 20 min 10% Ethanol, 5 min 1% Nitric Acid, 2 water washes, 30 min 0.2% $AgNO_3$, 1 water wash, 1 and 5 min in developer solution (3% Na_2CO_3, 0.02% Formaldehyde), 30 s 5% Acetic Acid and 3 water washes.

RESULTS AND DISCUSSION

The seedlings from the PxO population were screened for the SSR TAA41, and only 6 of them were nucellar. This microsatellite was able to distinguish the four alleles present in the population, therefore the genotype of every plant could be unmistakably assessed (Fig. 1). The three remaining plants showed to be zygotic when analysed with 4 more SSRs. In order to compare the efficiency of SSRs and isozymes to discriminate between zygotic and nucellar plants, 14 of the zygotic individuals were studied using 3 isozymic systems. All of them were able to determine the non-nucellar origin of the plants.

The isozymic analysis of the Fc population, however, revealed only three zygotic progeny when using PGM and PGI (Fig. 2a), and none with the other three isozymes analysed (although 3 of these plants could not be genotyped for GOT because they died). The same population was analysed with microsatellites,

and 6 zygotic plants were found, with 4 out of 5 SSRs detecting such plants (Fig.2b). Table 1 shows how screening of *Got-1* in the dead plants would have revealed only one more zygotic individual through isozyme analysis. Zygotic seedlings are very weak in *Citrus* and *Citrus* related species, and die very early, specially when they originate by self pollination. On the contrary, DNA extractions from them are still available, and could be analysed for additional DNA markers. In this case, SSRs proved to be more efficient than isozymes in the identification of zygotic individuals, since more zygotic plants were found and more microsatellite loci were able to detect them. Only 5.6% of the plants have a sexual origin in the FC population, a very low percentage compared to that of the PxO family (87%). Low percentages of zygotic progeny have been found previously in citrus populations (Ashari et al., 1988), and they are known to depend on the seed parent (Spiegel-Roy et al., 1977), the pollen origin (Cameron and Soost, 1969) and also subjected to environmental influences (Khan and Roose, 1988; Moore and Castle, 1988).

Loci	**Zygotic individuals code**					
	26	38	40	70	78	96
TAA41	**Z**	**Z**	n	n	n	n
TAA45	n	**Z**	**Z**	n	n	n
TAA33	n	**Z**	n	**Z**	**Z**	n
CAC23	n	n	n	n	n	n
TAA27	n	**Z**	n	-	**Z**	**Z**
*P*gm	n	n	n	n	**Z**	**Z**
Pgi	n	n	**Z**	n	**Z**	n
PRX	n	n	n	n	n	n
Got-1	n	n	n	-	-	-
6-PG	n	n	n	n	n	n

Table 1: Zygotic individuals from the selfing of *Fortunella crassifolia*. (**Z**) indicates a genotype different from that of the mother at a given locus. (n) refers to the same genotype as the mother. (-) refers to lack of data.

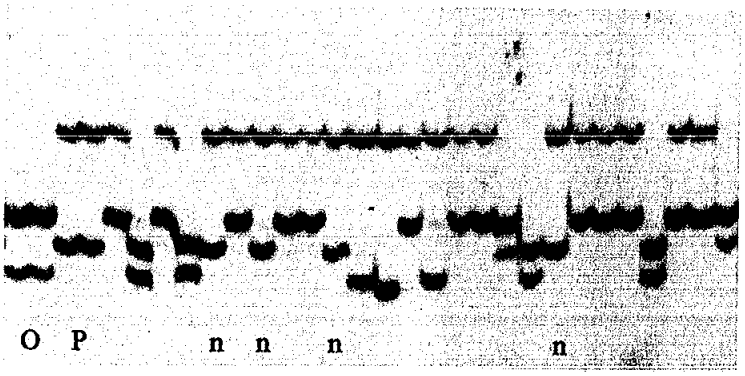


Figure 1.- Silver stained acrylamide gel of the SSR locus TAA 41 in the PxO progeny. From left to right there are two lanes of each parent, **O** for "Ortanique" and **P** for *P. trifoliata*, followed by individual plants of the progeny. Lanes corresponding to nucellar plants are marked **n**.

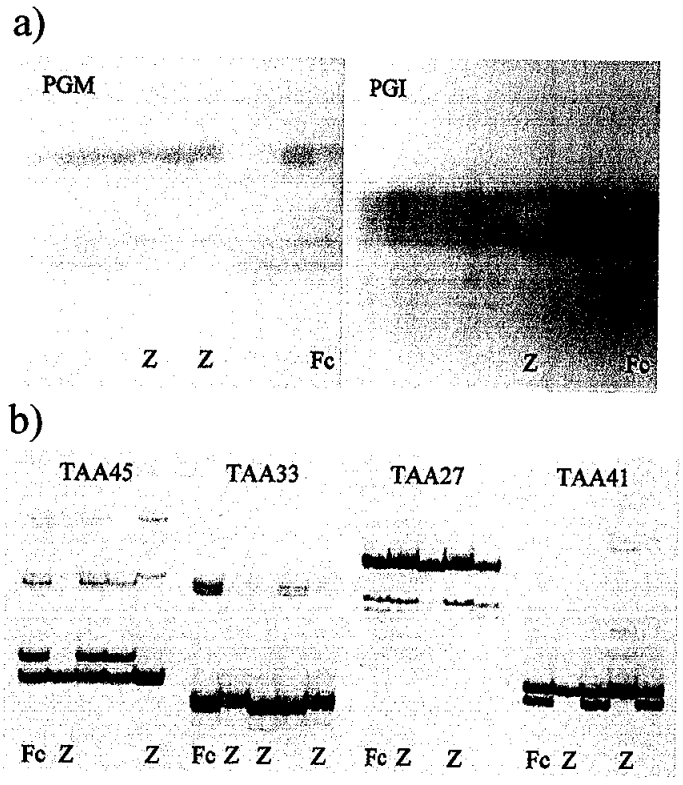


Figure 2.- Marker analysis of Fc progeny. a) PGM and PGI zymograms. b) Silver stained acrylamide gel of the SSR loci TAA45, TAA33, TAA27 and TAA41. Lanes marked as **Fc** corespond to *F. crassifolia* and **Z**, to zigotic or sexual seedlings from the Fc progeny.

One of the main limitations of isozyme analysis in this kind of studies when compared to other molecular markers is the small number of available loci in the genome, and the scarce variability at them. The limitation in the number of available loci can be overcome by using DNA markers such as RAPDs, AFLPs or SSRs. The two first are less informative because of their dominant nature (heterozygous and homozygous individuals for the presence of a band can not be distinguished), which reduces in half the ability to detect zygotic plants in some progenies, and then require to score a higher number of loci. SSRs, on the other hand, are co-dominant markers as are isozymes, but much more variable both in the number of polymorphic loci and in the number of alleles at each locus (Karhu et al., 1996; White and Powell, 1997; Raybould et al., 1998). Besides, their genetic interpretation is much simpler than that of isozymes. *Fortunella crassifolia* has low heterozygosity, and the number of known isozymic loci for marker analysis is very limiting. Mean observed heterozygosity for SSR loci has been found to exceed levels of diversity detected using isozymes or RFLPs as marker systems (White and Powell, 1997). Therefore, for some rootstocks, as "Cleopatra" mandarin, where isozyme analysis has failed in detecting zygotic seedlings because its high proportion of homozygous loci (Ashari et al., 1988), SSRs are the markers of choice.

Another disadvantage of isozyme analysis is that, in a preliminary search for polymorphic loci, one can find proteozymes where several bands are present and their genetic control is unknown. If a progeny is derived by selfing, one could expect some of these bands to be segregating, but in our Fc population, where the genitor shows complex zymograms for 4 out of 10 enzymatic systems (Herrero et al., 1996), PRX and 6-PG were not able to detect any zygotic individual, which could be due to that there is no segregating locus behind the variation for the enzymatic patterns.

It has been proposed that a major drawback for PCR derived markers when compared to isozymes is the need of a DNA extraction step, but the obtained DNA can be later used for future analysis, not depending on the phenological stage of the plant, as isozymes are. We propose a DNA miniprep protocol in microcentrifuge tubes which simplifies and shortens the process. Though this is a very crude DNA preparation, DNA quality has shown to be good enough for the analysis of SSR markers. Another criticism of the use of SSRs versus isozymes is that one can analyse 3-4 enzymatic systems in the same gel. But the number of samples that can be analysed at a time is 30-35 in isozymes, while we visualise 96 samples in each sequencing gel for microsatellites. Silver staining of the gels has proven to be highly sensitive, and eliminates the need for fluorescent or radioactive labelling, thus there is no need for the manipulation of radioactive isotopes and residues, nor special equipment or installations. The efficiency of the proposed technique can be further increased by running several SSRs in the same gel, as long as they yield bands of different size. In the case of citrus, sequences for microsatellite primers are available from Kijas et al. (1997) and a new set of primers is being developed in our laboratory.

The screening for zygotic individuals can require a different strategy depending on the population under study. We have seen that the efficiency of detection varies with the population type and the percentage of zygotic plants in the population. When a population derives from a cross, all of the zygotic individuals show a genotype different to that of the mother at any discriminating locus, provided that the father has alleles different from those of the mother. Even in the event of both parents being homozygous, any co-dominant marker would identify every zygotic plant as long as they have different alleles. Thus, in population PxO the first scored SSR was able to detect all of the screened zygotic plants. In this cross, the parents belong to species distantly related, which favours the existence of a variety of alleles at each locus, so that isozymes

showed a similar efficiency to that of SSRs in detecting zygotic plants. But if the cross is carried out between two closely related parents, it would be more difficult to find an allozyme distinguishing the parents; in such a case the use of SSRs would be the strategy of choice.

On the other hand, when the population is obtained by selfing, any co-dominant marker (Isozymes, SSRs, ...) will detect only 50% of the zygotic individuals, since half of the offspring has the same genotype as the mother and makes them indistinguishable from nucellar plants. That implies that several marker loci have to be scored in order to identify every zygotic individual (5 markers would identify 97% of the zygotic plants and 7 markers would identify over 99%). It is important to point out here that zygotic seedlings from most citrus rootstocks in open-pollination appear to arise from self-pollination (Moore and Castle, 1988). Therefore, the detection of zygotic seedlings derived by self-pollination, or open-pollination of autogamous rootstocks, will always be more efficient using SSR markers, specially when the level of heterozygosity is low.

AKNOWLEDGEMENTS

This work was supported in part by research grants Conselleria de Cultura, Educación y Ciencia (GV-3114/95), I.V.I.A. (C.R.) and Ministerio de Educación y Cultura (M.P.B.)

COMPARISON BETWEEN GENETIC LINKAGE MAPS FROM *PONCIRUS* AND *CITRUS* SPECIES

CARLOS RUIZ y Mª JOSÉ ASÍNS

Instituto Valenciano de Investigaciones Agrarias (IVIA)

Apdo. Oficial 46113 Moncada (Valencia)

Submitted to *Theoretical and Applied Genetics*

ABSTRACT

Five genetic linkage maps were constructed for the parents of three progenies: *Citrus aurantium* (A) x *Poncirus trifoliata* var Flying Dragon (Pa); *C. volkameriana* (V) x *P. trifoliata* var. Rubidoux (Pv) and a self-pollination of *P. trifoliata* var Flying Dragon (Pp). The number of polymorphic markers assayed ranged from 48 for Pa to 120 for A, according to the heterozygosity of each parent. Focused on genome comparison, most of the markers were newly generated SSRs. IRAPs based on 4 retrotransposon sequences isolated from *Citrus* spp were also used to saturate the maps. These polymorphisms were much more frequent in A (53) than in Pa (15) and randomly distributed throughout both genomes. Since comparative genomics and QTL-analysis applicability depends on the reliability of marker ordering, causes of variation in marker order were investigated. Three artificial causes were found: the introduction of new markers in the map, lowering the LOD score and the mapping software (JOINMAP and MAPMAKER). Around 25% of the markers show gametal segregation distortions. The most important reorganisations affect homeologous linkage groups 3, 7 and 11, where highly distorted markers are found. Segregation distortions were also observed at the zygotic level in the family obtained by self-pollination towards a reduction of the observed frequency of homozygotes from the expected one. Genomic reorganisations were observed when comparing the colinearity between *Citrus* and *Poncirus*, and also between *P. trifoliata* varieties, underlying the importance of constructing a linkage map for each parent separately instead of working with consensus maps. These results put into question the robustness of linkage maps using this kind of wide crosses and the available mapping software.

Key Words: *SSR, IRAP, Genomics, Segregation distortion, Fruit breeding*

INTRODUCTION

Citrus are economically the most important fruit trees on an international scale because their production exceeds 95 million tones (FAO 1999). Citrus cultivars belong to several species, in fact, the fruit originates from a scion which is grafted onto an apomictic rootstock propagated by seed to ensure uniform production and freedom from diseases. The scion belongs mainly to the following species (or crosses between them): *C. sinensis* (L.) Osb. (sweet orange), *C. clementina* Hort. ex Tan. (clementine mandarin), *C. unshiu* (Mak.) Marc. (satsuma mandarin), *C. paradisi* Macf. (grapefruit) and *C. limon* L. Burm. f. (lemon). Amongst the most important species used as rootstock, or in rootstock improvement programs, *P. trifoliata* (L.) Raf. (trifoliate orange), *C. aurantium* L. (sour orange), *C. volkameriana* Ten. (Volkamer lemon) and *C. reshni* Hort. ex Tan. (Cleopatra mandarin) are all notable. However, the number of available rootstock cultivars is very limited in several countries. Thus, in the East of Spain more than 85 % of new sweet orange and mandarin varieties are grafted onto just one rootstock genotype, the citrange Carrizo (*C. sinensis* x *P. trifoliata*) (Pina et al 2000). This is the reason for its vulnerability to attack from new pathogenic agents, or the evolution of existing ones into more virulent forms, or the progressive salinisation of arable lands.

The genetic improvement of citrus focuses on the obtaining of new rootstock resistant to disease and adverse growing conditions, it aims also to diversify the limited choice of currently available rootstocks. Such improvement must be based on knowledge about the genetic control of the potential traits and the use of plant genetic resources. Citrus represents a very wide set of species, but with long juvenile periods, a high level of apomixis (nucelar embriony) and, generally high heterozygosity. This means that genetic studies on citrus are very costly, long and, therefore, scarce.

Nowadays, plant genomic research focuses mainly on enhancing the productivity, quality, and sustainability of our food production systems. However, the plant genome of any crop species is an ideal concept. In practice, we know it is variable, but to what extent does this variation affect loci controlling agronomic traits? In the case of citrus, this question is very difficult to answer as there are various species. Are there differences in the colinearity of genes, among different citrus species? could the results of a QTL analysis carried out on one species, be extrapolated to another?. Constructing genetic linkage maps would permit to compare the organization of different genomes and an efficient and continuous use of plant genetic resources to enrich the gene diversity of breeding programs, supported by markers at the selection stages.

Mapping in outbreeding heterozygous perennial crops is not as advanced as in annual crops. They require more time and more space, given the long growing cycle and large crop size. Only progeny issued from the cross between two, more or less, heterozygous parents are usually available. In this case, up to four alleles per locus may segregate and the marker phase (coupling or repulsion) can not always be deduced from parent and grandparent banding patterns.

So far 14 genetic maps have been published (Table 1) but it is difficult to compare these maps because there are few common markers. In four cases just one consensus map has been obtained for the two parents without taking into account the possible chromosomal reorganisations which have come about in the evolution of the family *Aurantioidae* (Naithani and Raghuvanshi 1958, 1963; Raghuvanshi 1962; García et al. 1999). In other cases, the genetic map has been obtained mainly with dominant markers (RAPD, ISSR,...), and therefore, although we are dealing with the same primer, we cannot be certain that a similar-sized band correspond to the same marker locus in different populations.

Reference	Family	N	cM	M	LG	Software
Jarrell et al. (1992)	Sacaton x Troyer	60	351	38	10	MAPMAKER
Durham et al. (1992)	*C. grandis* x *P.trifoliata*	65	553	52	11	MAPMAKER
	C. reticulata x *C. paradisi*	65	314	32	8	MAPMAKER
Cai et al. (1994)	*C. grandis* x *P. trifoliata*	60	1192	189	9	MAPMAKER
Luro et al. (1996)	*C. grandis*	52	600	34	7	MAPMAKER
	C. reshni x *P. trifoliata*	52	1503	95	12	MAPMAKER
Kijas et al. (1997)	Sacaton x Troyer	57	410	48	12	JOINMAP
Simone et al. (1997)	*C. aurantium*	50	1000	247	20	MAPMAKER
	C. latipes	50	600	92	12	MAPMAKER
Garcia et al. (1999)	*C. volkameriana*	80	137	45	9	JOINMAP
	P. trifoliata	80	126	38	5	JOINMAP
Cristofani et al. (1999)	*C. sunki*	80	867	63	10	MAPMAKER
	P. trifoliata	80	732	62	8	MAPMAKER
Roose et al. (2000)	Sacaton x Troyer	57	701	153	16	JOINMAP

Table 1: Linkage maps for citrus published up to the present. N: Number of individuals that form the analyzed family; cM: Total size of the linkage map. M: Number of markers. LG: Number of linkage groups forming the map.

For this reason, the first objective of the present work has been to develop new markers that are codominant and with high levels of reproducibility and polymorphism in citrus. The microsatellites are simple sequence repeats (SSR), which are very abundant in Eucariots and which display high polymorphism (Tautz and Renz, 1984; Wang et al., 1994; Russell et al., 1997). They are randomly distributed in the genome, both inside and between genes (Weber 1990). The microsatellites have a mutation rate per generation that

varies between 2.5 x 10^{-5} and 10^{-2} and are therefore the most quickly evolving DNA sequences (Kashi et al., 1990;; Weber and Wong, 1993).

As well as the microsatellites, we have also developed IRAPs (inter-retrotransposon amplified polymorphisms), given that the abundance of retrotransposon sequences in the genome in many plant species they have been used in studies of phylogeny, biodiversity and linkage (Brandes et al. 1997; Ellis et al. 1998). In the case of citrus the retrotransposon sequences corresponding to the family Ty1-*copia* are very abundant (Asíns et al, 1999) and present a greater level of polymorphism than any other type of sequence (even microsatellites) in the vegetatively propagated crop *C. clementina* (Bretó et al, 2001). Therefore we propose using these sequences in our search for polymorphism in the citrus genome, given this would supply a large number of highly reproducible markers, both simply and quickly.

Here, we report the construction of five genetic maps using three segregating populations derived from three species commonly involved in breeding programs of citrus rootstocks. Using new generated codominant markers, SSRs, has allowed a study of colinearity among genomes and the presence of lethal recessive factors. The results constitute a first contribution to comparative genomics in citrus.

MATERIALS AND METHODS

Plant material

Three segregating populations were used: family PpxPp derived by self-pollination from *Poncirus trifoliata* var. Flying Dragon (57 trees); family AxPa, derived from the cross between *Citrus aurantium* var Afin Verna and *Poncirus trifoliata* var Flying Dragon (66 hybrids); and family VxPv, derived from the cross between *Citrus volkameriana* and *Poncirus trifoliata* var. Rubidoux (80 hybrids).

The families PpxPp and VxPv were studied previously by Mestre et al. (1997) and García et al. (1999), respectively. In the present study we have increased the number of markers typed in these progenies by adding SSRs and IRAPs and we have made new linkage maps.

Molecular markers

SSRs: Two strategies were followed to obtain these markers. Microsatellite screening was carried out using the FINDPATTERNS program, GCG package (Wisconsin Package, version 8.1-OpenVMS) in all citrus sequences, included at the GenBank up to October 1999, which was searched for all possible repetitions of di-, tri-, tetra- and pentanucleotides. The sequences with microsatellites were used to design specific primers, using the PRIME program, GCG package.

The other strategy was to obtain a library of *P. trifoliata* genomic DNA with small size fragments. *P. trifoliata* DNA extracted using CsCl gradient, was digested with the *Tsp* 509I restriction enzyme, of which the 4 pb target, is compatible with the *Eco*RI target. Digestion was analysed on agarose gel, DNA fragments of between 300 and 650 pb were cut and DNA extracted using the kit "Agrose gel DNA extraction" (Boerhinger). This DNA was cloned in the vector Lambda Zap® II (Stratagene) using its EcoRI recognition sequence and the construction was packaged using the kit "Gigapack® Gold" (Stratagene).

In total, approximately 90.000 pfu were screened (80.000 with 6 probes and 10.000 with 10 probes). The probes used were: $(ATA)_8$, $(AAAT)_6$, $(CTTT)_6$, $(TTC)_8$, $(ACT)_8$, $(ATC)_8$, $(AAC)_8$, $(CT)_{10}$, $(CTC)_8$ and $(GCT)_8$, which were labeled with digoxigenin using the end labeling kit "Dig oligonucleotide 3'-End labeling" (Boerhinger). For screening, phage plaques were transferred to a nylon membrane (Hybond-N, Amersham), then they were denaturalized with 0.5 M NaOH, 1.5 M NaCl and fixed for 2 hours at 80 ºC. The filters were hybridized

using mixtures of different probes in SSPE 5x, SDS 1% at 55ºC and rinsed twice with SSPE 2x, SDS 0.1% at 50ºC and rinsed twice again with SSPE 0.5x, SDS 0.1% also at 50ºC. Hybridization signal was detected using CSPD. All positives were screened twice to reduce the false positives and then, the positive clones were converted in pBluescript® plasmids by excision *in vivo* (following the Stratagene protocol). These clones were sequenced and PRIME program was used to design specific primers.

The SSRs obtained from the Net (with the prefix CR), those obtained from the DNA library (with the prefix CL) and those described by Kijas et al. (1997) were amplified and analyzed as in Ruiz et al (2000). Primer sequences were reported elsewhere (Ruiz et al. 1999).

IRAPs: 8 primers were designed from 4 citrus sequences having homology with the retrotranscriptase domain of the *Copya* family: the sequences were CL3, CL5 and CL6 from *C. clementina*, and SI4 from *Citrus sinensis* (EMBL database accession numbers: CCL131363, CCL131362, CCL131364 and CSI131367). The primers were denoted R or F (reverse or forward) for each sequence. The primers were used, alone or in pairs, to amplify intertransposonic sequences. IRAPs were amplified and analyzed as in Bretó et al. (2001)

Within the AxP family, as well as SSRs and IRAPs, we also analyzed some RAPDs and SCARs, as in García et al. (1999) and resistance analogues (AR), which are sequences amplified with primers designed using consensus sequences of resistance genes (Mago et al. 1999).

Linkage analysis

JOINMAP 2.0 (Stam 1993, Stam and Van Ooijen 1995) with a linkage criteria of LOD 6, in general, recombination fraction of 0.5 and Kosambi mapping function was used for linkage analysis. The population was analyzed as the

"Cross pollinator" population type with no previous knowledge of the linkage phase of the markers. A study was also made of linkage groups obtained according to a lowering in the criterion LOD = 6.0, at intervals of 0.5 units, while maintaining marker order within each group.

Nomenclature of the linkage groups has been carried out along the following lines: the first letter indicates which parent it belongs to, V (*C. volkameriana*), A (*C. aurantium*) and P (*P. trifoliata*). In the *P. trifoliata* maps each linkage group has two letters, the second indicates the other (female) parent of family, v (family VxPv), a (family AxPa) and p (family PpxPp). All the groups also have a number; if this is a Roman number (I, II, III,...), it indicates that no homology has been found with respect to linkage groups of other maps. If, by contrast, the linkage group is numbered with Arabic numbers, this indicates that this group presents two or more markers that are common to another linkage group of another map. A continuation of number indicates, in brackets, the map in which there exists a homologous or homeologous group.

The MAPMAKER program (Lander et al. 1987) was also used with a LOD = 3.0, to compare the results obtained with this program with the maps created with JOINMAP. The MapChart program (Voorrips 2001) was used to draw the linkage maps.

RESULTS

Molecular markers

In the database a total of 1477 citrus sequences were located, of which 61 contained microsatellites, but specific primers could only be designed to amplify these microsatellites in 27 sequences. The majority of the sequences found in the GenBank were cDNAs sequenced by Hisada et al. (1997) from *C. unshiu* ESTs. These sequences were short, approximately 300 pb, thus, in some

of the microsatellites found in them, specific primers could not be designed because there was not enough space between the microsatellite and the end of the sequence.

On the other hand, after the first DNA library screening, 81 positive clones, of which in a second check screening, carried out to eliminate false positives, 9 were confirmed as carrying one microsatellite. These 9 clones were sequenced, and primers could be designed for 6 of them.

As an initial test we analyzed the microsatellites using some individuals belonging to the AxPa family and observed that for some microsatellites other products of amplification appeared in addition to the product of the expected size, so corresponding to other loci. We decided not to modify the conditions, nor eliminate these microsatellites, given that, the extra-products did not interfere with the main one in the genetic interpretation and contribute more polymorphic markers to be included in the genetic maps.

		Markers	Linked markers	Linkage groups	cM	SSR_S	$IRAP_S$	$RAPD_S$	$RFLP_S$	Isoenz.	Others
Fam AxPa	Pa	48	40	6	275.2	23 (73.3%)	15	5			5
	A	120	104	18	441.5	54 (84.4%)	53	8			5
Fam VxPv	Pv	73	43	10	341.9	30 (80.6%)	11	21	6	3	2
	V	97	79	11	460.1	51 (94.3%)	8	22	11	4	1
Fam PpxPp	Pp	66	43	11	269.7	29 (79.3%)	26		2	5	4

Table 2: Number and type of marker in each map. Pa (map of *P. trifoliata* belonging to the family AxPa), A (map of *C. aurantium* belonging to the family AxPa), Pv (map of *P. trifoliata* belonging to the family VxPv), V (map of *C. volkameriana* belonging to the family VxPv) and Pp (map of *P. trifoliata* belonging to the family PpxPp). Together with the number of microsatellites (SSR) it indicates the heterozygosity calculated for each of the maps using this kind of marker.

Table 2 shows the number and type of polymorphic marker used to construct the map of each parent. Heterozygosity for microsatellites (Table 2) varies between 73.3 % in *P. trifoliata,* in the AxPa family, and 94.3 % in *C. volkameriana,* in the VxPv family. To obtain these values, only the main microsatellite bands have been used, the polymorphism of which was produced mainly by a change in the number of repetition units (Litt and Lutty 1989; Tautz 1989, Weber and May 1989, Morgante and Olivieri 1993).

Heterozygosity for IRAPs in the AxPa family is presented in Table 3. It has been calculated using only the combination of primers based on only one retrotransposon, CL3, CL5, CL6 or SI4 (i.e. primer combinations 3F3R, 5F5R, 6F6R and 4F4R, respectively). Similarly, the minimum number of copies of each retotransposon in the genome of each parental species has been estimated by the number of bands (Table 3). This varies between 4, for the retrotransposon SI4 in *P. trifoliata* and 26 for CL6, also in *P. trifoliata.*

RT sequence	Parental	Bands	Polymorphisms	Heterozigosity (%)	
CL3	*C. aurantium*	25	12	48,0	33,3
	P. trifoliata	20	3	15,0	
CL5	*C. aurantium*	16	10	62,5	42,9
	P. trifoliata	19	5	26,3	
CL6	*C. aurantium*	22	8	36,4	20,8
	P. trifoliata	26	2	7,7	
Si4	*C. aurantium*	11	3	27,3	26,7
	P. trifoliata	4	1	25,0	
Total	*C. aurantium*	74	33	44,6	30,8
	P. trifoliata	69	11	15,9	

Table 3: Number of bands, number of polimorphic bands and heterozygosity values in the family AxPa for each of the parents and retrotransposon (RT) sequences used in the primer design. Heterozygosity has been estimated as: number of polymorfic bands divided by the total number of bands.

Linkage maps

In each of the families used in the study various types of molecular markers have been analysed, as shown in Table 2. The SSR-type markers have been analyzed in all 3 populations in order to compare the different maps.

Both in the *C. volkameriana* x *P. trifoliata* family and the *C. aurantium* x *P. trifoliata* family, the parental *Citrus* species are heterozygous in more loci than *P. trifoliata* for the analyzed markers, therefore marker density in the *Citrus* spp. maps is greater than those obtained in the *Poncirus trifoliata* maps.

Distortion has been observed in the segregation of the markers in all the parents. In the AxPa family, *C. aurantium* has 23.3% and *P. trifoliata* 22.9%. In the PpxPp family, *P. trifoliata* displays distortion on segregation in 39.4% of the markers. In the VxPv family, *C. volkameriana* presents a distortion of 28.9%, while *P. trifoliata* var Rubidoux presents distortion in 39.7% of the markers.

		Gametic Segregation			Zygotic Segregation		
LG	Marker	Expected a:b	Observed a:b	χ^2 1 d.f.	Expected aa:ab:bb	Observed aa:ab:bb	χ^2 1 d.f.
P_p3	gp47	51:51	33:69	12,71 **	5,3:22,3:23,3	3:27:21	2,24
	TAA 27	57:57	78:36	15,47 **	26,7:24,6:5,7	26:26:5	0,17
P_p4	1.40_190	40:40	53:27	8,45 **	17,6:17,9:4,6	18:17:5	0,10
P_p5	TAA 41	56:56	65:47	2,89	18,9:27,3:9,9	15:35:6	4,49 *
	CR 25	55:55	60:50	0,91	16,4:27,3:11,4	12:36:7	5,63 *
P_p7	pg 52	54:54	68:40	7,26 **	21,4:25,2:7,4	19:30:5	1,97
	CR 7	56:56	41:71	8,04 **	7,5:26,0:22,5	4:33:19	4,07 *
	CR 17	55:55	69:41	7,13 **	21,6:25,7:7,6	18:33:4	4,41 *
	CR 18	55:55	69:41	7,13 **	21,6:25,7:7,6	18:33:4	4,41 *
Unlinked	cG18	49:49	35:63	8,00 **	6,3:22,5:20,3	6:23:20	0,02

Table 5: Type of segregation displayed by the different codominant markers in the family PpxPp. The expected zygotic frequencies have been calculated using the observed gametal frequencies, for this reason only 1 degree of freedom is used. Significance level: *: 0.05 and **: 0.005

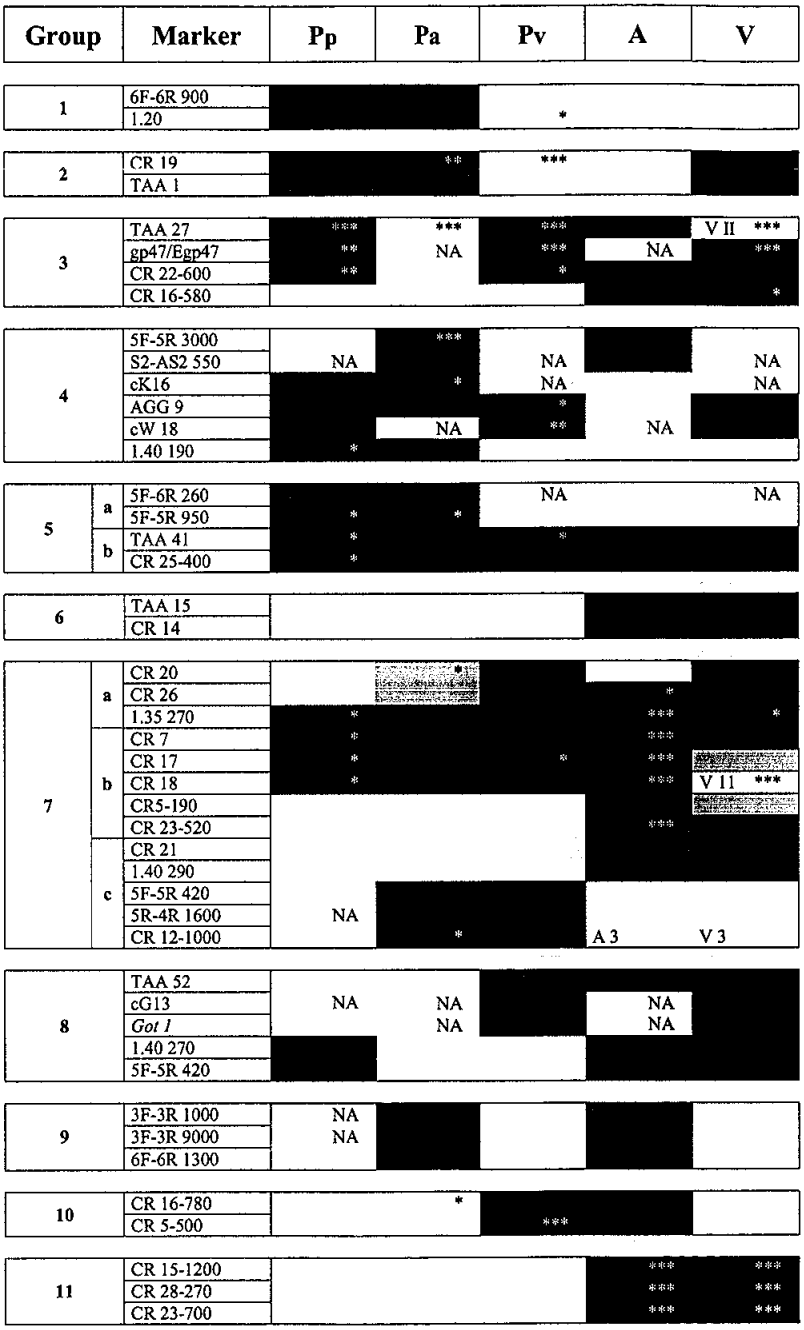


Table 4: The black cells show the markers that are common to different maps and that maintain the same order. The grey cells indicate a change in the order of the markers. If a marker has change from linkage group, it is indicated by putting the group in which it is found. The asterisks show a distortion in the segregation of the loci. Significance level: *: 0.05, **: 0.005 and ***: <0.0005

Table 4 indicates the distortion on segregation found in some markers that integrate the maps. This distortion is that obtained directly for the outcome of the JMSLA program, in JOINMAP. In the PpxPp family, the origins of distortion could be studied, seeing that on having codominant markers and just one parent it is possible to determine if the distortion is gametal or zygotic. The zygotic frequencies have been calculated using the gametal frequencies observed. As shown in Table 5, the linkage groups P_P 3, P_P 4 and the unlinked marker cG18, present a distortion of the gametal segregation. Markers at group P_P 5 has zygotic distortion and those at P_P 7 displays significant differences with respect to both gametal and zygotic segregations. Both in P_P 7 and P_P 5b the zygotic distortion comes from the low number of homozygous individuals observed contrary to expected. This would indicate the presence of recessive lethal factors in the genomic regions where a decrease in the frequency of the homozygous progeny has been detected.

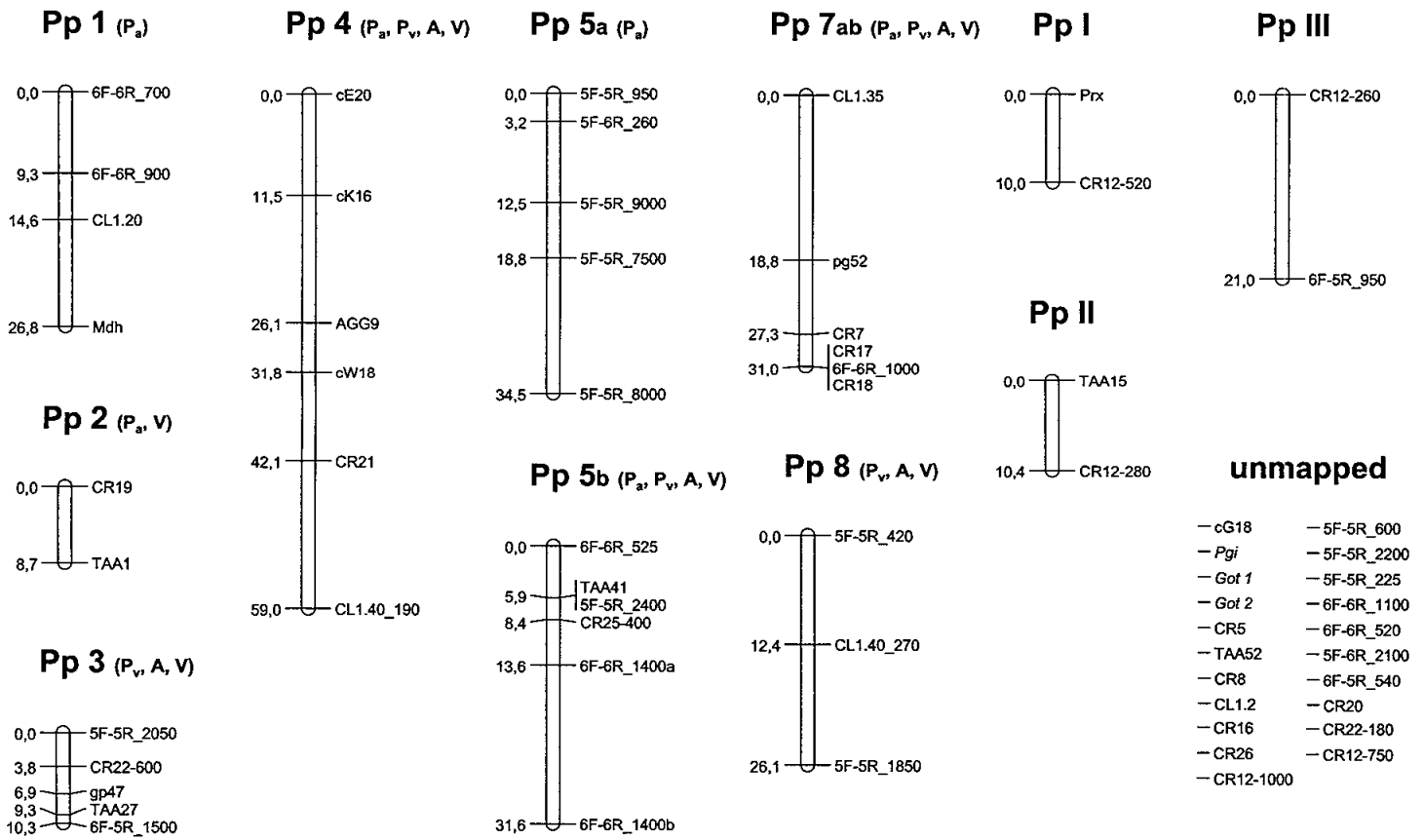
A genetic map has been established for each of the parents of each family from the linkage analysis of the markers obtained (Figure 1 a-e).

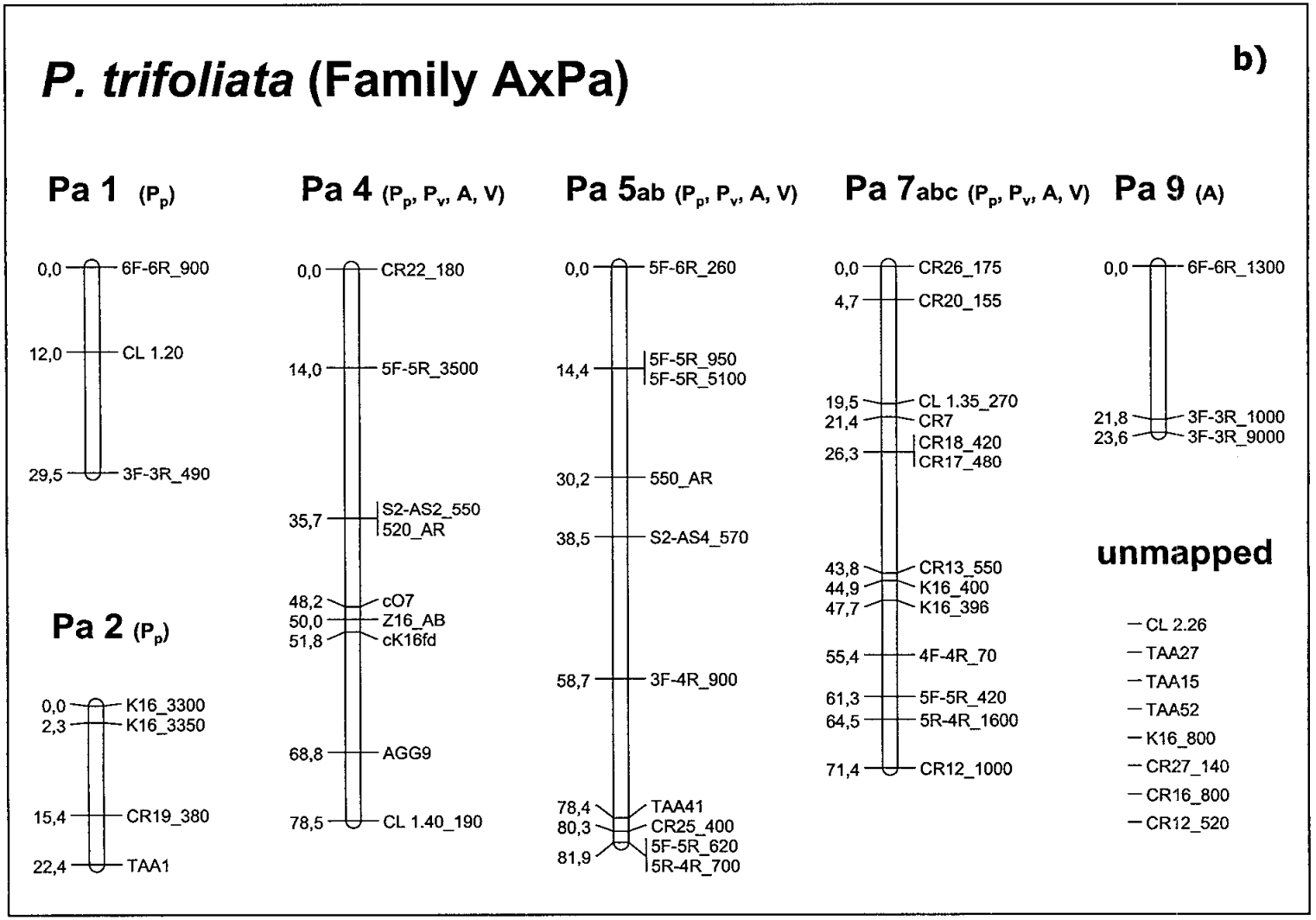
Table 4 shows a summary of the comparative analysis between the genomes of the different parental genotypes, in which one can see the markers that are common to the different linkage groups and whether the order of the groups is maintained (black boxes) or there are reorganizations.

Figure 1: Linkage maps obtained for each of the parents. a) *P. trifoliata* var Flying Dragon from the family PpxPp. b) *P. trifoliata* var Flying Dragon from the family AxPa. c) *P. trifoliata* var Rubidoux from the family VxPv. d) *C. aurantium* from the family AxPa. e) *C. volkameriana* from the family VxPv.

P. trifoliata (Family PpxPp)

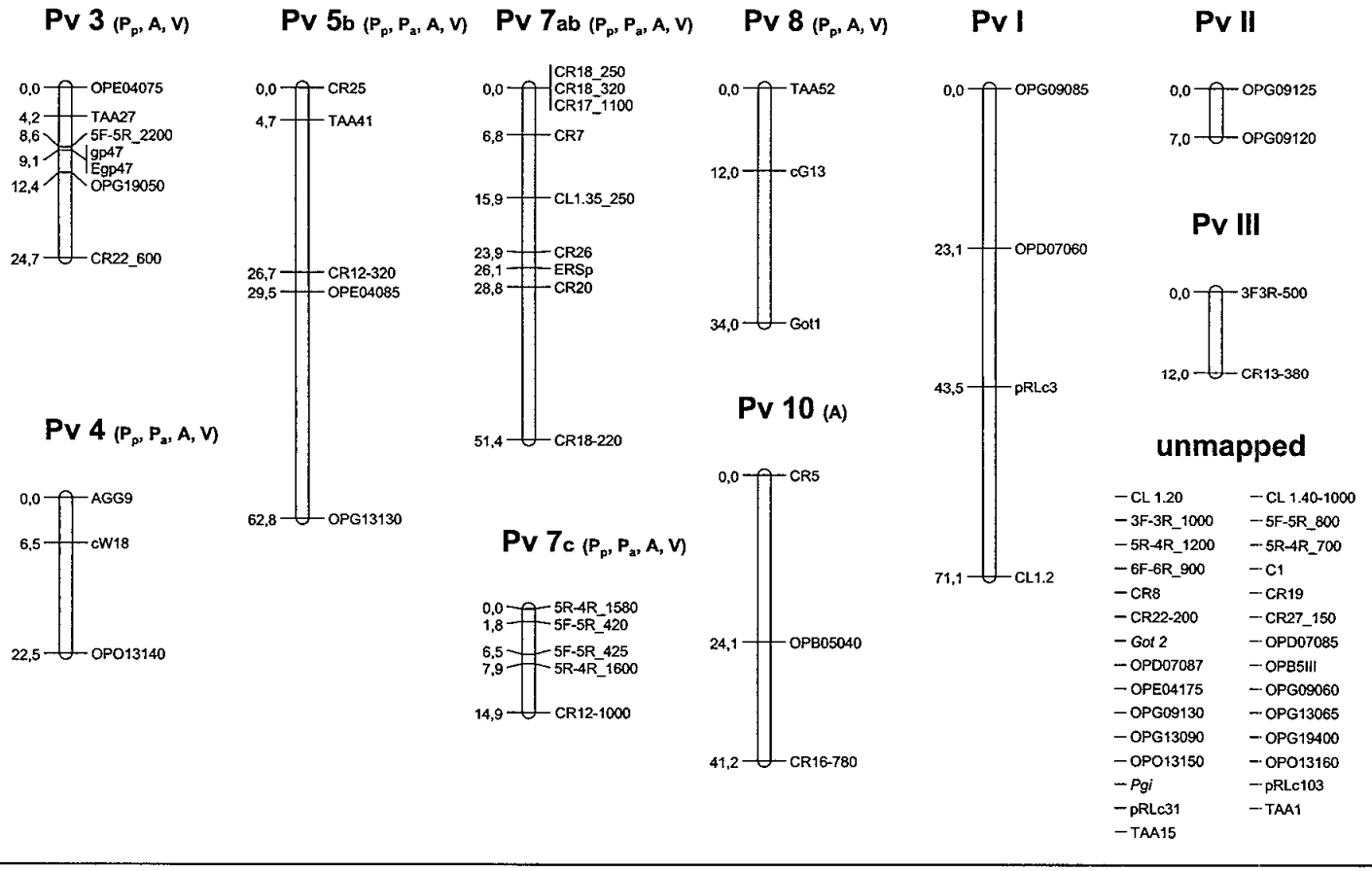
a)

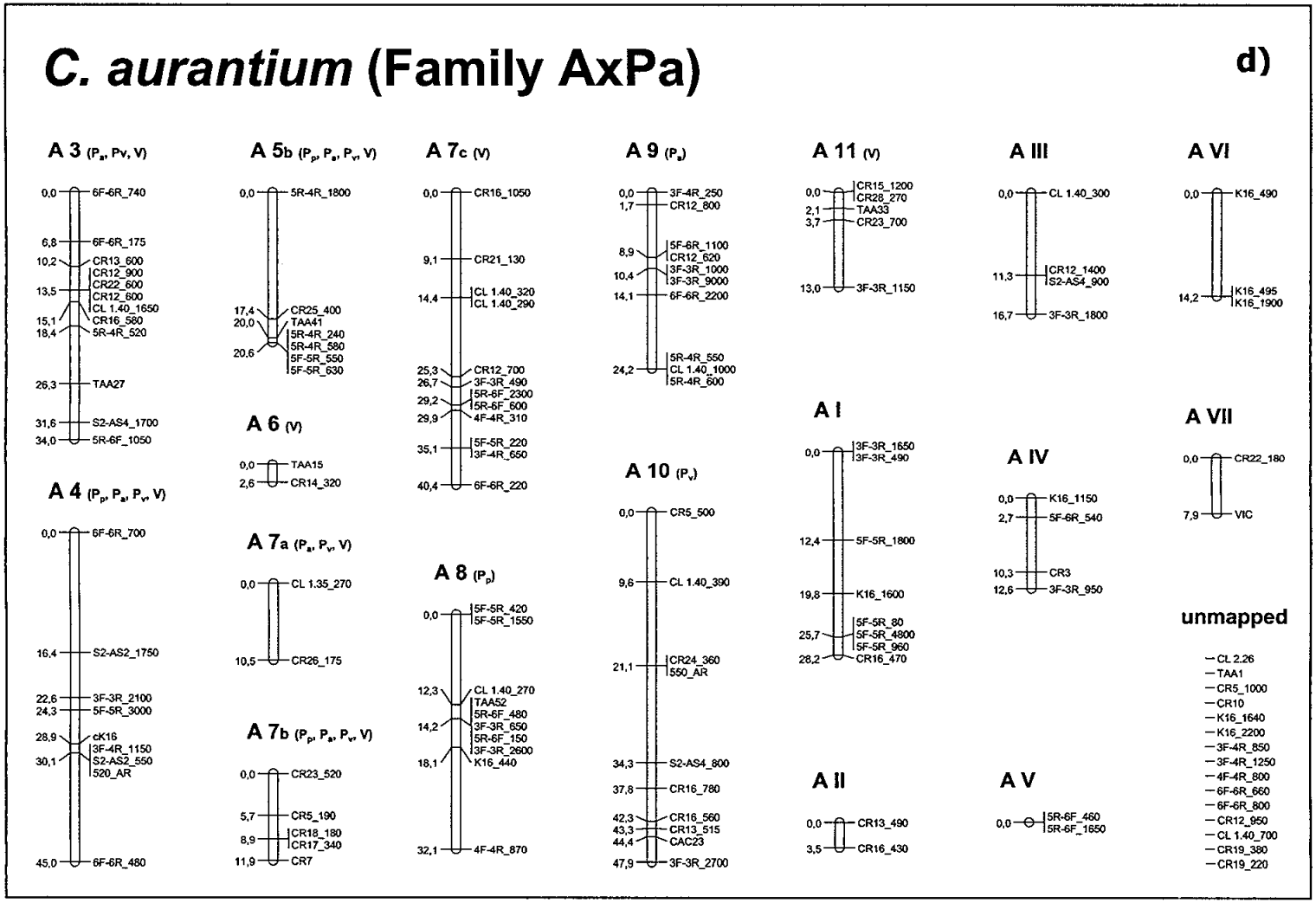



P. trifoliata (Family AxPa)
b)
Pa 1 (Pp)
0,0 6F-6R_900
12,0 CL 1.20
29,5 3F-3R_490
Pa 2 (Pp)
0,0 K16_3300
2,3 K16_3350
15,4 CR19_380
22,4 TAA1
Pa 4 (Pp, Pv, A, V)
0,0 CR22_180
14,0 5F-5R_3500
35,7 S2-AS2_550
520_AR
48,2 cO7
50,0 Z16_AB
51,8 cK16fd
68,8 AGG9
78,5 CL 1.40_190
Pa 5ab (Pp, Pv, A, V)
0,0 5F-6R_260
14,4 5F-5R_950
5F-5R_5100
30,2 550_AR
38,5 S2-AS4_570
58,7 3F-4R_900
78,4 TAA41
80,3 CR25_400
5F-5R_620
81,9 5R-4R_700
Pa 7abc (Pp, Pv, A, V)
0,0 CR26_175
4,7 CR20_155
19,5 CL 1.35_270
21,4 CR7
26,3 CR18_420
CR17_480
43,8 CR13_550
44,9 K16_400
47,7 K16_396
55,4 4F-4R_70
61,3 5F-5R_420
64,5 5R-4R_1600
71,4 CR12_1000
Pa 9 (A)
0,0 6F-6R_1300
21,8 3F-3R_1000
23,6 3F-3R_9000
unmapped
CL 2.26
TAA27
TAA15
TAA52
K16_800
CR27_140
CR16_800
CR12_520

P. trifoliata (Family VxPv)

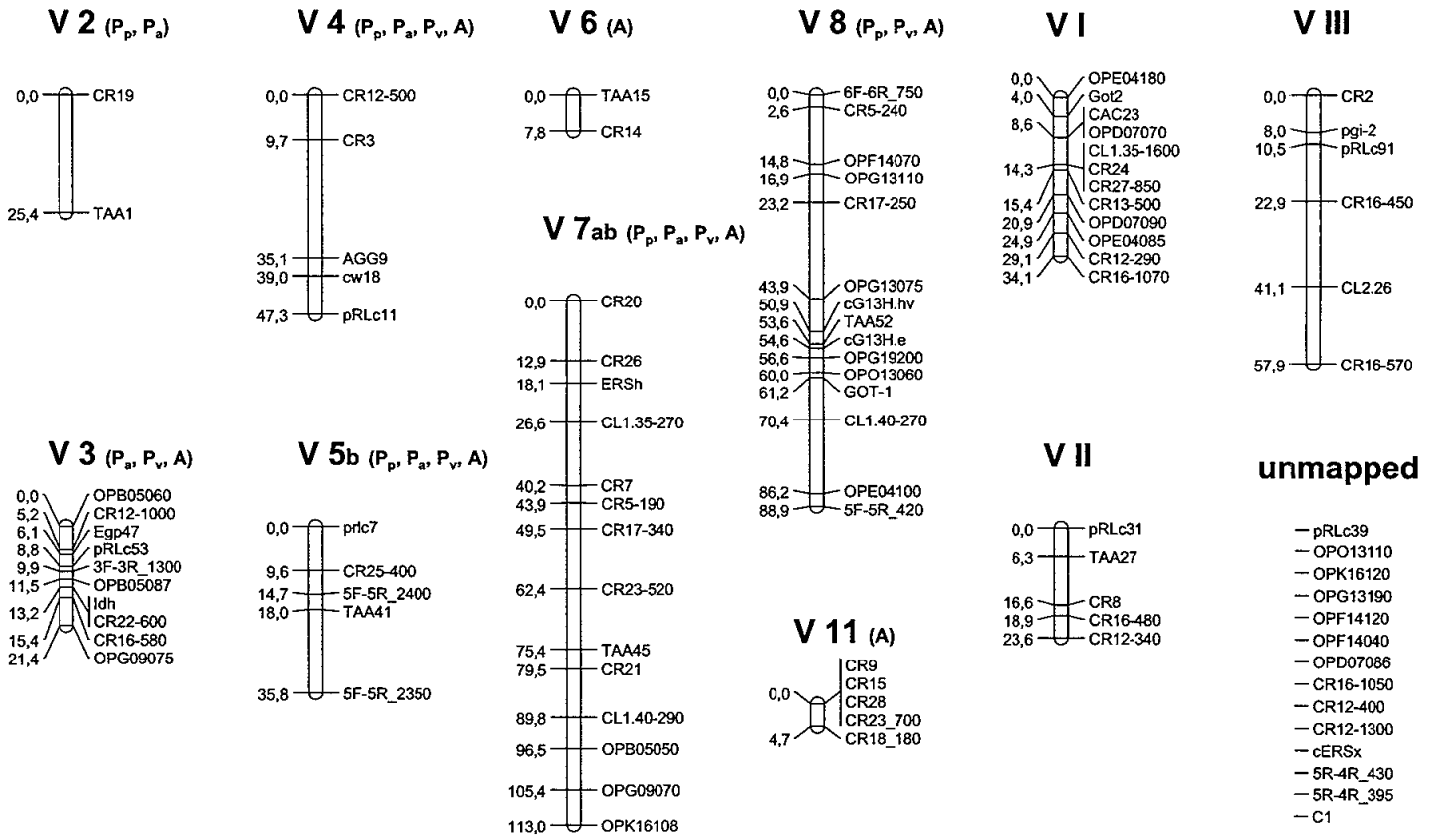
c)




C. aurantium (Family AxPa)
d)
A 3 (Pa, Pv, V)
0,0 6F-6R_740
6,8 6F-6R_175
10,2 CR13_600
CR12_900
CR22_600
13,5 CR12_600
CL 1.40_1650
15,1 CR16_580
18,4 5R-4R_520
26,3 TAA27
31,6 S2-AS4_1700
34,0 5R-6F_1050
A 4 (Pp, Pa, Pv, V)
0,0 6F-6R_700
16,4 S2-AS2_1750
22,6 3F-3R_2100
24,3 5F-5R_3000
28,9 cK16
30,1 3F-4R_1150
S2-AS2_550
520_AR
45,0 6F-6R_480
A 5b (Pp, Pa, Pv, V)
0,0 5R-4R_1800
17,4 CR25_400
20,0 TAA41
5R-4R_240
5R-4R_580
20,6 5F-5R_550
5F-5R_630
A 6 (V)
0,0 TAA15
2,6 CR14_320
A 7a (Pa, Pv, V)
0,0 CL 1.35_270
10,5 CR26_175
A 7b (Pp, Pa, Pv, V)
0,0 CR23_520
5,7 CR5_190
8,9 CR18_180
CR17_340
11,9 CR7
A 7c (V)
0,0 CR16_1050
9,1 CR21_130
14,4 CL 1.40_320
CL 1.40_290
25,3 CR12_700
26,7 3F-3R_490
29,2 5R-6F_2300
5R-6F_600
29,9 4F-4R_310
35,1 5F-5R_220
3F-4R_650
40,4 6F-6R_220
A 8 (Pp)
0,0 5F-5R_420
5F-5R_1550
12,3 CL 1.40_270
TAA52
14,2 5R-6F_480
3F-3R_650
5R-6F_150
18,1 3F-3R_2600
K16_440
32,1 4F-4R_870
A 9 (Pa)
0,0 3F-4R_250
1,7 CR12_800
8,9 5F-6R_1100
CR12_620
10,4 3F-3R_1000
3F-3R_9000
14,1 6F-6R_2200
24,2 5R-4R_550
CL 1.40_1000
5R-4R_600
A 10 (Pv)
0,0 CR5_500
9,6 CL 1.40_390
21,1 CR24_360
550_AR
34,3 S2-AS4_800
37,8 CR16_780
42,3 CR16_560
43,3 CR13_515
44,4 CAC23
47,9 3F-3R_2700
A 11 (V)
0,0 CR15_1200
CR28_270
2,1 TAA33
3,7 CR23_700
13,0 3F-3R_1150
A I
0,0 3F-3R_1650
3F-3R_490
12,4 5F-5R_1800
19,8 K16_1600
25,7 5F-5R_80
5F-5R_4800
5F-5R_960
28,2 CR16_470
A II
0,0 CR13_490
3,5 CR16_430
A III
0,0 CL 1.40_300
11,3 CR12_1400
S2-AS4_900
16,7 3F-3R_1800
A IV
0,0 K16_1150
2,7 5F-6R_540
10,3 CR3
12,6 3F-3R_950
A V
0,0 5R-6F_460
5R-6F_1650
A VI
0,0 K16_490
14,2 K16_495
K16_1900
A VII
0,0 CR22_180
7,9 VIC
unmapped
CL 2.26
TAA1
CR5_1000
CR10
K16_1640
K16_2200
3F-4R_850
3F-4R_1250
4F-4R_800
6F-6R_660
6F-6R_800
CR12_950
CL 1.40_700
CR19_380
CR19_220

C. volkameriana (Family VxPv)

e)



DISCUSSION

Molecular markers

Comparing the two methods used to develop microsatellites, screening on the Net was less time-consuming and less demanding on resources. By using this method 27 new SSRs have been developed. The same method has been used in other plant species such as *Arabidopsis thaliana* (Bell and Ecker, 1994), maize (Senior and Heun, 1993) and sorghum (Brown et al. 1996). The efficacy of this method greatly depends on the number of sequences present in the data base of the species under analysis. For instance, only 2 microsatellites were obtained for sorghum because there were only 45 sequences in the data base. The cDNA sequencing project by Hisada et al. (1997) has increased considerably the number of citrus sequences in the data base in recent years. Doubtless genome projects will greatly speed up obtaining SRR markers using bioinformatics.

Library screening is much more costly and requires its construction, screening and sequencing of positive clones, which is the starting point for the method based on bioinformatic. However, if one wishes to obtain SSRs in lesser studied species, for which no sequence is available in the data bases, this is the method one must follow although it is also possible to try microsatellites designed for closely related species (Brown et al. 1996).

Microsatellite alleles may present the same size but different nucleotide sequences; this is called homoplasy (Estoup et al. 1995; Anger and Benatchez 1997). To get around this lack of information one can use the heteroduplex fragments (Perez et al. 1999), which are two double-chain DNA molecules, each formed by a combination of the two DNA chains corresponding to the two different alleles of the heterozygous individual. On forming a heteroduplex, like this, regions improperly paired are produced which makes the molecule migrate

more slowly in an electrophoresis (White et al. 1992). The presence of heteroduplex allows to differentiate two or more alleles that have the same electrophoretirc mobility (Figure 2) given that, for each combination of alleles a different pair of heteroduplex are produced and therefore the genotype of the individual can be safely assigned.

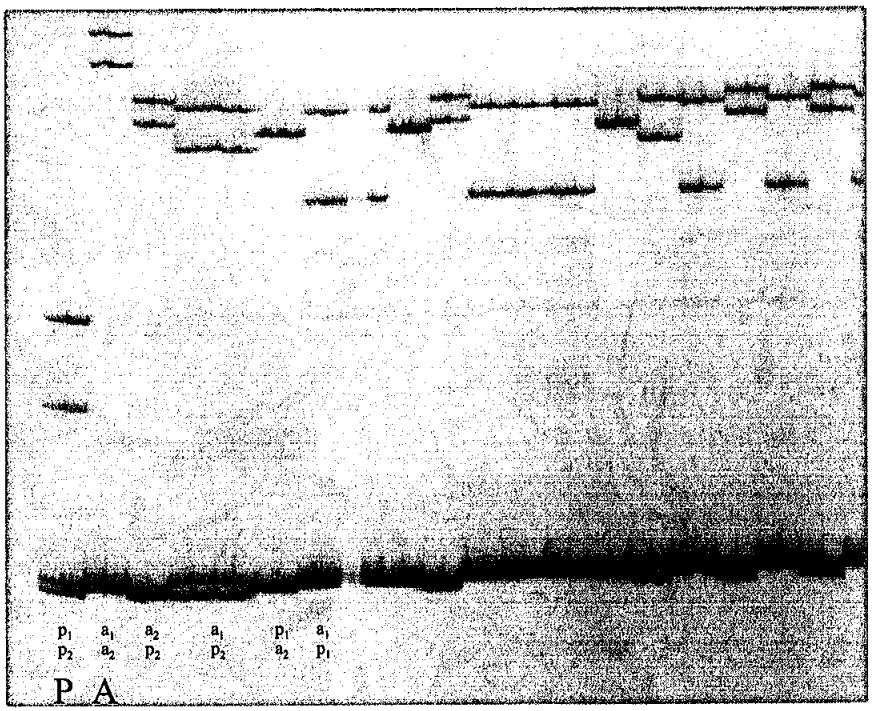


Figure 2: Silver stained gel electrophoresis of the marker CR 25 from hybrids belonging to the family AxPa. a_1 and a_2 corresponding to the alleles of *C. aurantium* (A) and p_1 and p_2 are the alleles of *P. trifoliata* (P). The 4 possible genotypes are indicated by the different combinations of heteroduplex bands (upper bands).

The heterozygosity calculated with the microsatellites is specific to this type of sequence, seeing that they evolve much more quickly than the rest of the genome (Russell et al. 1997) in the presence of sexual reproduction (Bretó et al, 2001). *P. trifoliata* has lower heterozygosity, between 73.3% and 80.6 %, than *Citrus* species, which have values of between 84.4% and 94.3%.

As well as the microsatellites, some SCARs, RFLPs, and isoenzymes have been used to compare linkage maps. They have also been used to compare IRAPs, given that in some cases, observing band size and relative intensity, one can determine whether two bands of different families correspond to the same locus. The IRAPs are good markers to saturate linkage maps because with a few primer combinations one can obtain many markers with high repeatability and more or less random distribution.

The most IRAPs have been analysed in the AxPa family, and the *C. aurantium* map is the most saturated by this type of marker. In this map one can observe, as in other plant species, that the retrotransposons are distributed throughout the genome (Brandes et al. 1997; Waught et al. 1997). However, some zones seem to be enriched. These regions might be hot insertion points (Ananiev et al. 1998). In the remaining maps, although marker density is lower, this distribution pattern can also be observed.

To estimate the minimum number of existing copies of each of the retrotransposons, in the AxPa family we counted the number of bands that are amplified when direct and reverse primers of the same RT sequence are combined in a PCR reaction. This number could be underestimated because various amplifications might have the same size or because retrotransposon copies exist that are isolated in the genome and are not far enough from each other to permit adequate amplification. On the other hand, one could also think it is overestimated, seeing that there are various copies in tandem, the number of amplifications is greater than the number of retrotransposon copies, however, if we focus on where the polymorphic bands map (IRAPs) we can observe that, although tandem copies must exist, this is not the most frequent case.

With the estimated number of copies and the number of polymorphic bands of each primer combination, we calculate the heterozygosity for IRAPs.

This heterozygosity should not be taken as the general heterozygosity of the genome given it is specific for each of the different retrotransposons and each species. The average heterozygosity for the IRAPs is between 44.6 % for *C. aurantium* and 16 % for *P. trifoliata* (table 3). That is to say, quite a lot lower than the heterozygosity found for SSRs, even though it follows the same trend, i.e., greater in *C. aurantium* than in *P. trifoliata*.

For the two primers used corresponding to the CL5 sequence, both *C. aurantium* and *P. trifoliata* give a maximum % of heterozygosity, 62.5 and 26.3, respectively. Specifically, a large part of the polymorphisms observed within the species *C. clementina*, closely related with *C. aurantium* (Herrero et al 1996), corresponds to IRAPs based on this CL5 sequence (Bretó et al, 2001).

Construction and comparison of linkage maps

The genome size of citrus was estimated, using the MAPMAKER program (Lander et al. 1987), at between 1500 and 1700 cM (Jarrell et al. 1992). If this were true, our maps, made using JOINMAP would cover between 30.6-27%, of the *C. volkameriana* map, and 18-16% of the *P. trifoliata* map in the PpxPp family.

MAPMAKER does not accept the data with mixed segregation phases, thus it is not suitable to analyze our families. Despite this, an analysis was made with the AxPa family in which one could observe closely linked markers, but in different phases, they appear in two different groups although they follow the same order as that found using JOINMAP. We also observed that those groups containing the same markers and following the same order were 25% longer, as average, when using MAPMAKER. This coincides with what Mestre et al. (1997) observed for the linkage group P_P 4 belonging to the PpxPp family. According to Cai et al. (1994) the total size of a linkage group can reduce by at least 50% using JOINMAP instead of MAPMAKER, thus the genome coverage we have

achieved must be underestimated since linkage maps have been obtained using the JOINMAP program.

A factor that may contribute to differences in colinearity between maps is differences in the LOD criteria. Changes in the order of the markers on a lowering of LOD score have been observed (Figure 3). Seeing that for future genetic analyses, correct ordering is of great importance, we preferred to construct maps at high LOD scores, even though this means that fewer markers fall within the linkage groups and therefore the total length of the genome covered in cM is apparently lower.

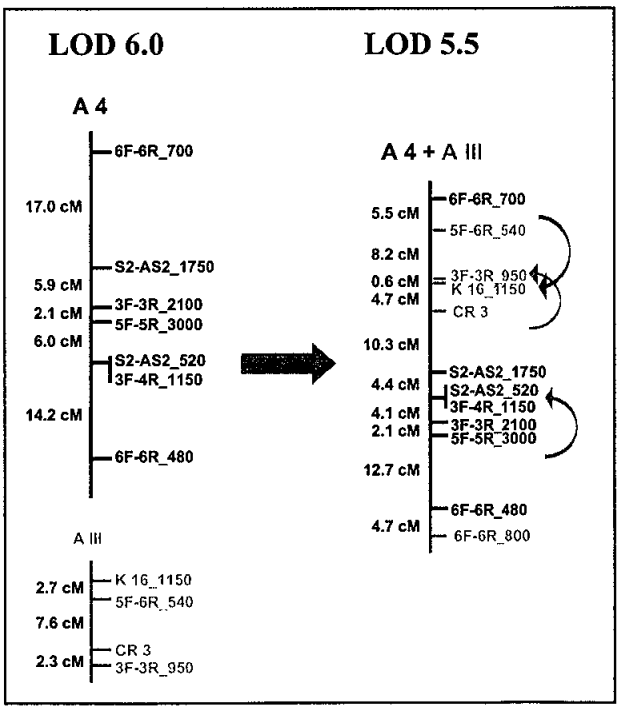


Figure 3: Linkage groups A4 and AIII of *C. aurantium*, obtained at LOD=6.0, which link togeter on lowering LOD to 5.5. Arrows are used to show the site at which the changes in marker ordering took place.

Differences in the distances between two markers in different maps are sometimes remarkable. Differences in the distance could be due to the sample of individuals or to the variation between genotypes in their recombination frequency. Concerning the applicability of genetic maps on QTL analysis and

MAS, differences in the order of markers have worse consequences than difference in their distances.

On comparing the maps for *P. trifoliata* var. "Flying Dragon", Pp and Pa (Table 4), differences mostly arise from the fact that a few markers have not been analyzed in the three families studied. Also, the absence of one marker in a linkage group may prevent the next marker from joining this group, given that the distance can be too large for the next marker to join the group at LOD 6, leaving it, therefore, unlinked.

On comparing the maps of the two varieties of *Poncirus trifoliata,* "Flying Dragon" and "Rubidoux", we observe a possible reorganization in the map of "Flying Dragon" (Pa) in the linkage group 7, compared to the other maps (Pv, A and V). This reorganization, consists in the inversion of a zone of the genome which comprises the markers CR 20 and CR 26. This cannot be observed in the Pp map, seeing that the marker CR 26 displays great distortion in its segregation (presence versus absence) in this family, which is due to a selection against homozygous for absence of band. On the other hand, CR 20 remains unlinked. This locus on the Pa map is to be found at a distance of 14.8 cM from the next one CL 1.35-250, therefore, as it is quite a large distance, it is normal that it does not retain linkage between the marker CR 20 and the group Pp 7. Another possible reorganization affect group 2, "Flying Dragon" presents markers CR 19 and TAA 1 linked with a distance of between 7.1 (Pa 2) and 8.7 (Pp 2) cM, while in "Rubidoux" these two markers do not join, remaining unlinked.

"Flying Dragon" is a dwarfing rootstock for citrus and according to Cheng and Roose (1995) this characteristic is controlled by just one dominant gene. These same authors suggest that "Flying Dragon" must have originated as a mutation from a small-flowered, non-dwarfing genotype, like "Rubidoux", given that in the 40 marker loci (isoenzymes and RFLPs) they analyzed, no difference

was found. Surprisingly, despite this similarity in marker genotype their genetic maps display differences in the lineal ordering of markers in groups 7 and 2. It is possible that some of these reorganizations could be responsible for its being a dwarfing rootstock.

When we compare the maps of *Citrus* with those of *Poncirus* we observe that both *Citrus* maps have zones that are common but at the same time differentiate them from the maps of the genus *Poncirus*, as for instance the complete groups 11 and 6, and parts of group 7 (from CR 23-520 to CL 1.40-290). This is not due, in principle to deletions, but to the fact that *Poncirus trifoliata* is homozygous for the majority of the markers of these groups.

A reorganization between the maps of *P. trifoliata* and those of *Citrus* involves marker CR 12-1000 in group 7c which is in group 3 in both the *Citrus* maps.

In groups 3 Pp and 3 Pv of Flying Dragon and Rubidoux, respectively, markers TAA 27 and gp47 are linked in the same group, by contrast, in the map of *C. volkameriana* the marker TAA 27 is situated in group II. Surprisingly, TAA 27 locates at group A 3 in the *C. aurantium* map where it does not present segregation distortion. In fact, this is the only species where TAA 27 does not present segregation distortion.

In *C. volkameriana* another reorganization has occurred between the markers CR 17 and CR 5-190 in group V 7, these two markers are inverted and CR 18, which is between them in other maps, is to be found in group V 11 here.

Other factor contributing differences in the order of markers is the addition of new markers to a previous map, i.e. when a previously constructed map is enriched with new markers. This would be the situation on comparing the maps obtained in the family VxPv with those previously published by this group (García et al 1999). Some of the markers have changed their position within the

same linkage group V 3, the RFLP pRLc53 was previously located between the markers *Idh* and RAPD OPG09075 and, in the group V I, the RAPD OPE04180, which is at one of the ends, was found between CAC 23 and OPD07090. In the Pv map there are also two changes when compared to the map of *P. trifoliata* drawn up by García et al. (1999); in Pv I the marker OPD07060 was at the end of the group, leaving pRLc3 in the center and, in group Pv 3, the markers TAA 27 and Egp47/gp47 have changed position with respect to the markers OPE04075 and OPG19050. In other words, adding new markers to the maps, which is habitual practice in research teams, can bring about the change in order of some of the previously mapped markers. If the order of the markers in some regions of the maps is so fragile, one must be cautious interpreting QTLs detected at those regions because subsequent MAS schemes may fail.

We can also compare our maps with some of those published previously by other teams (Table 1). One is a research team at the University of California, which has some common microsatellites, isoenzymes and RFLPs. Three versions of this map have been published due to successive updates: Jarrell et al. (1992), Kijas et al. (1997) and Roose et al. (2000). The one by Kijas et al. (1997) is the same as that by Jarrell et al. (1992) but with some extra markers (SSRs), and a marker (gp47) disappears only in group H. Some changes in marker order can also be observed. Three of our linkage groups have at least two markers in common with the map by Kijas et al. (1997). Groups V 8 and Pv 8 having markers TAA 52 and *Got 1* must be the same as group A. These markers are also maintained in the map by Roose et al. (2000). The third linkage group corresponds to group D by Kijas et al. (1997) which has the markers *Idh*, pRLc53, TAA 27 and pRLc11^{s}. In our maps, the first two markers are in group V 3, but TAA 27 and pRLc11 are found in different groups in the *C. volkameriana* map, while in *C. aurantium*, TAA 27 locates at group A 3. Markers at

homeologous group 3 present frequently segregation distortions (Tables 4 and 5). This may affect their location making it less precise, more variable.

We have identified here some factors that affect the ordering of makers and make comparative genomic difficult. These are: the addition of new markers, the chosen LOD criterion, the linkage phases, segregation distorsions, and the work with consensus maps when the parents have undergone chromosomal reorganizations. Genetic linkage maps play a prominent role in many areas of genetics: QTL analysis, map-based cloning of genes, marker-assisted breeding and, recently, comparative genomics. Therefore, tools are urgently required for establishing the quality of the data and the maps produced. Although a novel combination of techniques that establishes posterior intervals to the location of markers has recently been reported by Jansen et al (2001) no software that provides such tools is available yet.

ACKNOWLEDGMENTS

This work was supported in part by grants from IVIA (CR), Conselleria de Cultura, Educació i Ciència (GV-3114/95), INIA (USDA-MAPA 5) and INIA (SC99-047). The authors thank Drs. E. Carbonell and F. Gonzalez for helpful advices and Dr. J. Forner for providing the VxPv family.

THE DIVERSIFICATION OF *Citrus clementina* Hort. ex Tan., A VEGETATIVELY PROPAGATED CROP SPECIES

Mª PAZ BRETÓ, CARLOS RUIZ, JOSÉ ANTONIO PINA y Mª JOSÉ ASÍNS

Instituto Valenciano de Investigaciones Agrarias.
Ctra. Moncada-Náquera, Km 4.5, 46113 Moncada, Valencia.

Molecular Phylogenetics and Evolution **21(2):285-293**

ABSTRACT

Clementines, due to their high quality, are one of the most important cultivated citrus mandarins. As in the case of sweet orange and satsuma mandarins, genetic variability within this species is minimal when analysed by molecular markers, since the existing varieties have not been obtained through hybridisation, but through the selection of spontaneous mutations affecting traits of agronomic interest. This would explain, at least in part, the greater diversity for agronomic traits when compared to the variability for molecular markers. Another possible (non-exclusive) reason is that the types of molecular marker used are not focused on the kind of molecular change mainly involved in the origination of new clementine cultivars; i.e. are all sources of variation equally involved in the diversification of these plants? To answer this question different kinds of markers based on primers of random sequence, simple sequence repeats and retrotransposon sequences that may reveal point mutations, somatic recombination and transposon activity, respectively, were used to compare the level of variability among 24 clementine varieties. Their ISSR, RAPD and AFLP analysis provided only two polymorphic bands, distinguishing just two varieties. No variability was found by SSRs; i.e. no new allele arising through somatic recombination was detected. Instead, the amplification of sequences adjacent to retrotransposons yielded a higher number of polymorphisms (14.6 % versus 2.4 % for the previous mentioned marker types). Two geographical distant groups, one from North Africa and the other from Spain, have evolved in agreement with polymorphisms based on IRAP markers anchored to, at least, two different *Copia-like* retrotransposon sequences. Therefore, this study suggests that the DNA of this type of mobile elements is evolving faster than the DNA of other markers in this clonal lineage.

Key Words: Retrotransposons, DNA fingerprinting, IRAPs, DNA methylation, fruit breeding

INTRODUCTION

Citrus is one of the most important fruit crops in the world; however, the genetic variability within each cultivated species is very narrow (Herrero et al. 1996). All cultivars of clementine mandarins (*Citrus clementina* Hort. ex Tan.) derive from a single plant, likely the product of an uncontrolled cross, identified in an orphanage of Oran (Algeria) by the Father Clement Rodier in 1890 (Trabut, 1902), which was propagated and called "Ordinary Clementine". It is known in Spain as cultivar "Fina" since its introduction in the early 1900s, becoming a very important crop in the second half of the century because its excellent organoleptic and easy-peeling qualities. Similar to sweet oranges (*C. sinensis* (L.) Osb.) and satsuma mandarins (*C. unshiu* (Mak.) Marc.), clementines are vegetatively propagated and the new cultivars are obtained after careful selection of spontaneous somatic mutations. Bud mutations arise often in citrus (Raghuvanshi 1962; Cameron and Frost 1968) and are generally detected by the growers themselves in branches of trees showing altered horticultural traits, such as maturity and flowering time, or fruit characteristics. Contrasting with this diversity for agronomic traits, very low genetic variability has been found in cultivated citrus by use of isozymes, RFLPs, RAPDs and ISSRs (Herrero et al. 1996; Fang and Roose 1997; Kijas et al. 1997; Federici et al. 1998; Coletta-Filho et al. 1998; Iannelli et al. 1998, Protopapadakis and Papanikolaou 1999). RAPDs and ISSRs are efficient marker systems in cultivar identification, therefore the scarce variability found can not be attributed to the type of marker. Another possibility is that those types of marker focus mainly on the detection of point mutations, insertions and deletions in the targeted loci, and these loci are not necessarily involved in the agronomic traits that differentiate the cultivars.

There are other mechanisms that might be a source of somatic genetic variation such as somatic recombination, DNA methylation and transposon activity. The most suitable type of marker to check somatic recombination as a

source of somatic variation would be one based on satellite DNA, such as microsatellites, since processes of unequal exchange control the variation in number of repeats (Harding et al. 1992). Additionally, microsatellites are hypervariable DNA sequences. Variation in DNA methylation has been proposed as a mechanism that may explain the wide range of changes that can occur after tissue-culture. DNA methylation provides one mechanism for altering the expression pattern of the gene by affecting the local structure of the chromatin (Lewis and Bird 1991). Whereas hypermethylation in general is correlated with gene inactivity (Matzke and Matzke 1991), hypomethylation at specific sites has been correlated with active transcription (Jones and Buckley 1990). Changes in DNA methylation have been found in callus and regenerated plants of maize (Brown et al. 1991) and tomato (Smulders et al. 1995). DNA methylation has dual roles in defense against invading DNA and transposable elements and in gene regulation (Finnegan et al. 1998). Transposon sequences have been found methylated in citrus species, including *C. clementina* (Asíns et al. 1999), however, no differences at the methylation pattern of these sequences were found among its cultivars.

Retrotransposable elements, particularly the Ty1-*copia* class, are abundant in all the studied plant genomes, and generally dispersed along all the chromosomes (Voytas et al. 1992; Flavell et al. 1992; Brandes et al. 1997). Since retrotransposons do not excise from their insertion sites, they are stable in the genome and behave like mendelian loci. And so, sequences belonging to the *copia* family of retrotransposons have been used in several species for phylogenetic, biodiversity and mapping applications (Heslop-Harrison et al. 1997; Brandes et al. 1997; Ellis et al. 1998; Wang et al. 1999). Asíns et al. (1999) identified several sequences in *Citrus* showing homology to elements of the Ty1-*copia* class. Therefore, it is possible to amplify sequences contiguous to genomic regions where an insertion event has taken place by using specific primers

derived from the citrus elements. Two techniques have been reported recently, the IRAP technique (Inter-Retrotransposon Amplified Polymorphisms, Kalendar et al. 1999) and a modification of the AFLP protocols that involves transposon sequences, named S-SAP or Transposon Display (Waugh et al. 1997; Van der Broek et al. 1998). They are based on the use of primers facing outward from the transposon sequences and the amplification of part of the flanking sequence up to another retrotransposon or to a restriction site.

Investigations on the sources of variation that are operating in the origin of bud mutations are also useful because they could allow the development of new methods in citrus breeding and would make cultivar identification easier by focusing on a type DNA. Identification of scion cultivars in nursery situations is particularly difficult because some cultivars are distinguishable only by fruit traits, and citrus trees usually do not bear fruit until 3-4 years after planting. The ability to identify citrus cultivars using a small amount of leaf or other vegetative tissue would be helpful in protecting the rights of citrus breeders, growers and nurseries.

The main objective of the present study is to compare the level of polymorphism within a specifically selected collection of clementine cultivars by using a wide array of DNA marker types (RAPD, ISSR, AFLP, microsatellites, IRAP) corresponding to different sources of variation: point mutations, DNA methylation, somatic recombination and transposon activity.

MATERIAL AND METHODS

Plant material

Twenty four accessions of *C. clementina* were analyzed by DNA markers (Table 1). The trees belong to the Citrus Germplasm Bank at I.V.I.A. These entries were chosen because they represent both the most distinctively different

cultivars among the important ones grown in Spain, and entries whose field characterisation is not clear. For example, "Orogrande" is agronomically very similar to "Clemenules" and "Caffin" to "Oronules". Some varieties originated in Morocco were also included.

	RAPDs		Retrotransposons					
			IRAPs					S-SAPs
Accessions	I6-600	E15-1600	5F5R-700	5R4R-810	5R4R-800	3F3R-360	3F4R-900	1EA5R-200
Oroval	0	0	0	2	1	0	1	1
Hernandina (R)	0	0	1	1	2	0	1	1
Tomatera	0	0	1	1	2	0	1	1
Reina	0	0	1	1	2	0	1	1
Clemenules	0	0	1	1	2	0	1	1
Fina	0	0	1	1	2	0	1	1
Esbal (R)	0	0	1	1	2	0	1	1
Bruno	0	0	1	1	2	0	1	1
Arrufatina (R)	0	1	1	1	2	0	1	1
Orogrande	0	0	1	1	2	0	1	1
Clementard	0	0	1	1	2	0	1	1
Clemencira	0	0	1	1	2	0	1	1
Marisol	0	0	0	2	1	0	1	1
Fina-Bechi	0	0	1	1	2	0	1	1
Oronules (R)	0	0	1	1	2	0	1	1
Sra-91	0	0	1	1	2	0	1	1
Bekria	0	0	1	1	2	2	1	1
Nour	0	0	1	1	2	0	1	1
Ma-3	0	0	1	1	2	2	1	1
Caffin	0	0	1	1	2	1	1	1
Carte-Noir	0	0	1	1	2	0	1	1
Loretina (R, E)	1	0	0	2	0	0	0	0
Clemenpons (R, E)	0	0	1	1	2	0	1	1
Mioro (R, E)	0	0	1	1	2	0	1	1

Table 1: Accessions and their phenotypes for consistent polymorphic markers: 0 stands for absence of a band, and 1 and 2 for presence (indicating two clearly different intensities). R: registered variety, E: only a grower is allowed to multiply the variety. Moroccan accessions in italics.

Marker analysis

At least two different DNA extractions from 1 g of leaf tissue per cultivar were obtained following the methods of Dellaporta et al. (1983) with minor modifications. All PCR reactions were performed in duplicate and from different DNA extractions.

RAPDs: 26 primers (10mers from Operon Technologies Inc.) were used by themselves or in pairs for PCR reactions. The 25 μl amplification reactions consisted of 150 ng of template DNA, 1x supplied reaction buffer, 1.5 mM $MgCl_2$, 100 μM each dNTP, 0.2 μM primer and 1 u Taq (EcoTaq, Ecogen). Each reaction was overlaid with 25 ul of mineral oil, and amplified in a MJ Research thermal cycler. The cycling profile was: 95°C, 5 min; 45 cycles of 95°C (1 min), 44°C (1 min), a slope of +0.3°C per sec, and 72°C (2 min), and a final extension at 72°C (8 min); except for primers OPN07 and OPN14, where thermocycling was according to Coletta-Filho et al. (1998). PCR products were mixed with 5X loading buffer (50% v/v Glycerol, 1x TAE, 10% v/v saturated Bromophenol Blue, 0.2% w/v Xylene Cyanole), and visualised by agarose-ethydium bromide (1.2 % agarose-TAE gels) or by polyacrylamide-silver staining: sequencing-type 10% polyacrylamide gels, with a 5% stacking phase in 1x TBE and fixed to the glass. Electrophoresis conditions were 20W, 30 min, and 40 W until the bromophenol-blue dye reached the bottom of the gel, in an S2 Sequencing System (Gibco BRL/Life Technologies). Silver staining was performed according to Ruiz et al. (2000).

ISSRs: 16 of the primers for inter-simple sequence repeat analysis used by Fang and Roose (1997) to differentiate citrus varieties were screened. Only 12 yielded good fingerprinting patterns in clementines. All of them were primers anchored at the 3' end, and part of UBC primer Kit #9 from the Biotechnology Laboratory, University of British Columbia, Canada. The amplification mix was similar to the one described for RAPDs, but 25 ng of DNA were used as template, and it contained 2% Formamide and 1μM primer. Two different PCR programs were used: for (AG) and (GA) repeats it consisted of 94°C, 7 min; 2 cycles of 94°C (1 min), 35°C (1 min), a slope of +0.3°C per sec, and 72°C (2 min); 35 cycles of 94°C (1 min), 45°C (1 min), a slope of +0.3°C per sec, and 72°C (2 min), and a final extension at 72°C (7 min). For (CA), (AC) and (GT) repeats the

cycling profile was 94°C, 7 min; 40 cycles of 94°C (1 min), 52°C (1 min), a slope of +0.3°C per sec, and 72°C (1 min), and a final extension at 72°C (7 min). PCR products were electrophoresed and revealed as described for RAPDs (polyacrylamide gels).

SSRs: Nine primer pairs for citrus microsatellites described by Kijas et al. (1997) and 13 additional primer pairs developed by the group (to be published elsewhere) were analysed according to Ruiz et al. (2000). Some of the PCR reactions were also performed on DpnII- or Sau3AI-digested DNA: 300 ng of genomic DNA were incubated overnight with 0.6 u of the corresponding restriction enzyme in a 5 μl reaction, at 37 °C.

AFLPs: DNA was digested with EcoRI and MseI, and 8 primer combinations (with 3 selective nucleotides) were analysed following the standard procedures described by Vos et al. (1995). Bands were visualised by ^{33}P-labeling the EcoRI-directed primer, and exposing the dried gel to Phosphor Imager screens (Molecular Dynamics) for 24 h, or to Biomax MR films (Kodak) for 2 days.

IRAPs: 8 primers were designed from 4 citrus sequences having homology with the RT domain of the *copya* family: the sequences were CL3, CL5 and CL6 from clementines, and SI4 from *Citrus sinensis* (EMBL database accession numbers: CCL131363, CCL131362, CCL131364 and CSI131367). The primers were denoted R or F (reverse or forward) for each sequence. The primers were used, alone or in pairs, to amplify intertransposonic sequences. In some cases, DNA was digested with the restriction enzymes BamHI, NlaIII and Sau3AI, as described for SSRs. PCR conditions, electrophoresis and visualisation were similar to those described for RAPDs, but 300 ng of DNA and 0.12 μM of each primer were used instead for amplification reactions.

S-SAPs: 9 amplifications were carried out using one AFLP primer and one of the following retrotransposon primers, labelled with ^{33}P: the above-mentioned clementine specific primer CL5R, and two degenerated primers for the *copia*-RT domain, RetroQ and RetroY (Asíns et al. 1999). Template DNA was digested with EcoRI and MseI, and AFLPs protocols were followed.

Specific information on sequences of the primers used is available upon request.

Cluster analysis

A dendrogram based on hypothesized genotypes (homozygous for presence of band, 2; heterozygous,1; and homozygous for absence of band, 0) was obtained using the chord distance (Cavalli-Sforza and Edwards 1967) and the UPGMA (unweighted pair-group method with arithmetic average; Sneath and Sokal 1973) method of aggregation.

RESULTS

Table 2 shows which of the analysed marker types resulted in consistent polymorphisms. Ten polymorphic bands were detected in the initial screening that were not so when DNA from two different extractions per accession were used in the reactions. This lack of repeatability happened for some AFLP bands (2) and mostly when digested DNA was used in SSRs (3), IRAPs (3) and Transposon Display analyses (2) (Figure 1). None of these pseudo-polymorphisms were used to study the genetic relationships among cultivars.

In general, every RAPD, ISSR and AFLP reaction, resolved in polyacrylamide gels, allowed the visualisation of 15 to 45 bands. Only two of all those bands, the RAPDs I6-600 and E15-1600, were consistently polymorphic when scored in all 24 varieties (Figure 1), and they distinguished the varieties "Loretina" and "Arrufatina", respectively (Table 1).

Marker Types	Primers -	Primers +	Consistent Polymorphisms
RAPDs	26		2
ISSRs	12		0
SSRs	22	15 *	0
AFLPs		8 *	0
IRAPs	28	4 *	5
S-SAPs		9 *	1

Table 2: Type and number of markers, and consistent polymorphims. Number of primers or primer combinations with (+) or without (-) a previous DNA digestion by an endonuclease restriction enzyme. * refers to the presence of markers where the polymorphisms were not consistent through repetitions.

When using the specific *copia*-derived primers, 5 intertransposonic bands and 1 transposon-AFLP band were consistently polymorphic (Figure 1). The polymorphisms were obtained using the primers CL5F, CL5R, SI4R, CL3F and CL3R; i.e. primers anchored to 3 out of the 4 RT sequences. As shown in Table 2, the three first primers (RT sequences CL5 and SI4) were involved in polymorphisms that distinguished the related cultivars "Oroval", "Marisol" and "Loretina", and the primers CL3F and CL3R (RT sequence CL3) in the differentiation of North African cultivars "Bekria", "MA-3" and "Caffin".

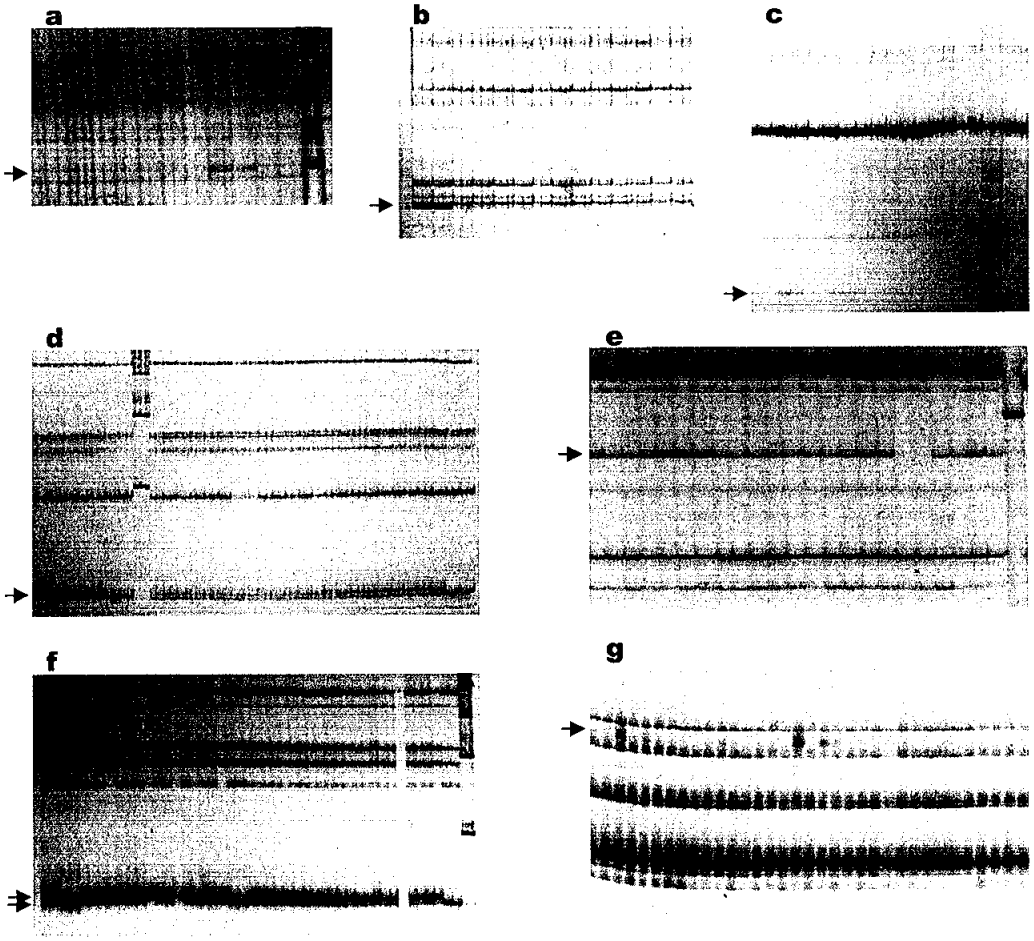


Figure 1.- Electrophoretic profiles of polymorphic markers. Two adjacent lanes per accession; (a) and (b) RAPDs with primers E15 and I6, respectively; From (c) to (f) IRAPs with primers 3F3R, 5F5R, 3F4R and 5R4R, respectively. Arrows on (a) to (f) point at polymorphic bands. Arrows on (g) point at a non-repeatable AFLP polymorphism.

A possible genetic interpretation of differences concerning band intensities (2, 1 and 0 in Table 1) is to consider them as the number of alleles for the presence of band per locus, allowing to calculate a genetic distance to quantify relatedness. The genetic relationships among the studied accessions, assuming this genetic interpretation, are shown in Figure 2. Since most of the varieties were not discriminated by any marker, they clustered in a main group (the "Fina" group); but "Arrufatina" was separated from it by the RAPD marker

E15-1600, and there were two more groupings, one formed by "Oroval", "Marisol" and "Loretina" (characterised by several polymorphic markers), and the other by "Bekria", "MA-3" and "Caffin" (characterised by the marker 3F3R-360).

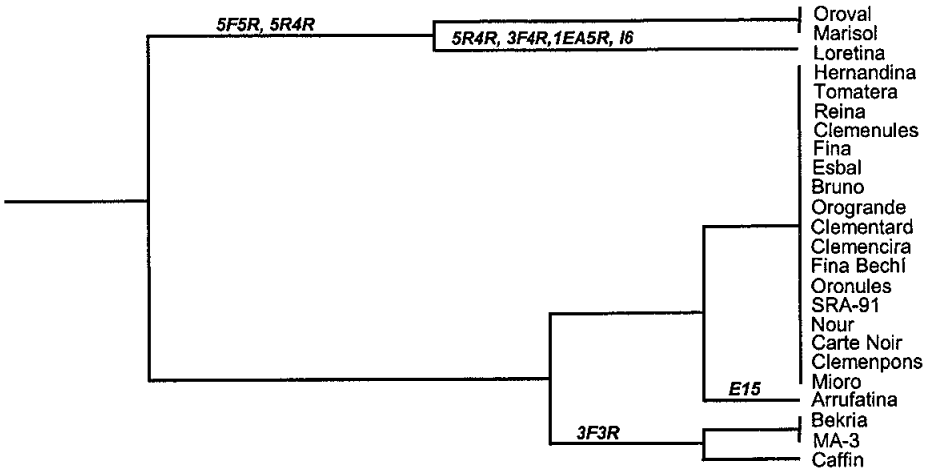


Figure 2: Dendrogram displaying the genetic distances among varieties. The codes on the arms refer to the marker loci involved in the differentiation.

Other genetic interpretation of band intensities is in terms of number of fragments of a certain size. Then 0 would mean no fragment; 1, a fragment and 2, two fragments of the same size but differing in sequence (and locus). Therefore considering this alternative genetic interpretation two loci have to be considered for bands 5R4R-810 (*5R4Rc* and *5R4Rd);* 5R4R-800 (*5R4Ra* and *5R4Rb*) and 3F3R-360 (*3F3Ra* and *3F3Rb*). This genetic interpretation is assumed in Table 3 to construct the ordered allelic profiles of cultivars. It should be noted that the groupings are the same than those shown in figure 2 and they agree with the decreasing tendency of the number of IRAP bands from a group of North African cultivars (putative origin of Spanish clementines) to the most recent one, "Loretina". This genetic interpretation also allows us to draw the phylogenetic diagram shown in figure 3 where discriminative mutations, most of them towards the absence (or decreasing intensity) of IRAP bands, are indicated.

Accessions	*I6*	*E15*	*5F5R*	*5F4Rc*	*5F4Rd*	*5F4Ra*	*5F4Rb*	*3F3Ra*	*3F3Rb*	*3F4R*	*1EA5R*	IRAPs
Bekria MA-3	0	0	*	*	0	*	*	*	*	*	*	5
Caffin	0	0	*	*	0	*	*	*	0	*	*	5
Fina Group	0	0	*	*	0	*	*	**0**	0	*	*	4
Arrufatina	0	*	*	*	0	*	*	**0**	0	*	*	4
Oroval Marisol	0	0	**0**	*	*	*	**0**	0	0	*	*	3
Loretina	*	0	0	*	*	**0**	0	0	0	**0**	**0**	1

Table 3.- Ordered allelic profiles of cultivars after interpretation of differences in intensity in terms of number of fragments of a certain size. Then, two loci are considered for bands 5R4R-810 (*5R4Rc* and *5R4Rd);* 5R4R-800 (*5R4Ra* and *5R4Rb*) and 3F3R-360 (*3F3Ra* and *3F3Rb*).* stands for presence of allele leading to band and 0 for an alternative allele, or mutation, leading to absence of band. Changing alleles from a group to the next (below) are indicated in bold characters. The total number of IRAP bands present in the cultivar (or group of cultivars) is indicated in the last column. Note the decreasing tendency of this number from the North African cultivars down to the most recent one, "Loretina".

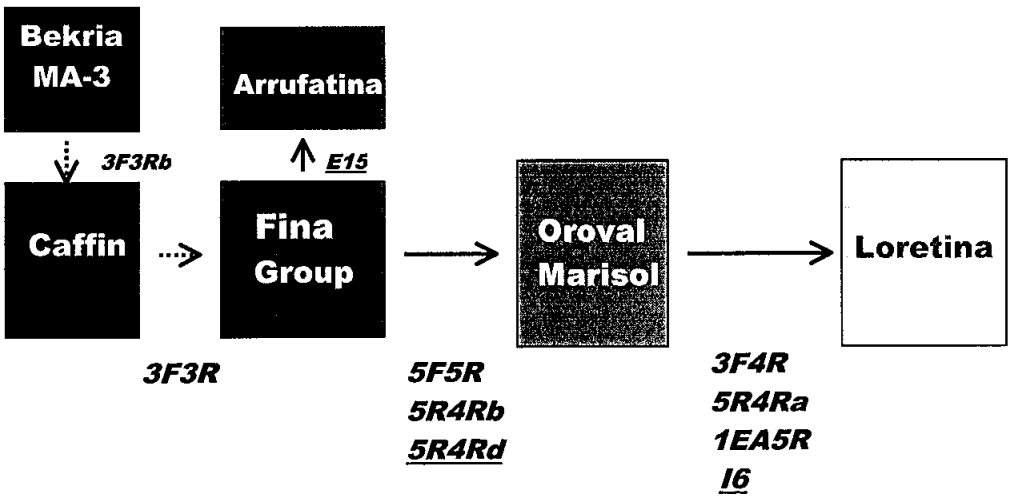


Figure 3.- Phylogenetic diagram based on the allelic profiles of Table 3. It shows the accumulation of mutations from the putative origin of Spanish clementine cultivars down to the most recent one, "Loretina". The black intesity of boxes correspond to the number of IRAP bands. Solid arrows mean known origin; discontinuos arrows mean hypothetized origin. Underlined alleles correspond to the presence of a new fragment, the rest correspond to mutations leading to absence of band.

DISCUSSION

Reliability, reproducibility and repeatability are essential for a technique to be used in fingerprinting, and especially for varietal certification. We have observed several polymorphisms that are not detected when a second DNA extraction from the same clone is analysed. Somatic instability in DNA fingerprints has been reported between branches of a tree for a dispersed middle repetitive sequence in citrus (Orford et al. 1995), for organ-specific AFLPs in wheat (Donini et al. 1997), and for RAPDs of leaves subjected to different physiological stages (Bitonti et al. 1996). The inconsistent polymorphisms in our SSR, AFLP, IRAP and S-SAP analyses appeared when the template DNA was digested. There are two possible explanations for this lack of repeatability among DNA extractions for makers that involve a step of DNA restricton: inhibitors of the restriction enzyme co-extracted with the DNA or differences in the methylation state of the DNA among branches in a given plant, since restriction enzymes such as Eco RI, BamHI and Sau3AI are sensitive to methylation. DNA methylation occurs at a varying degree in plant DNA. Methylation patterns are dynamic and change during cell differentiation, both temporally and spatially, presumably as a mechanism for the regulation of gene expression (Brown 1989; Finnegan et al. 1998). Usually, phylogenetic and variability studies do not include replicates of the DNA extractions; we recommend verification of the repeatability of the polymorphisms by analyzing several DNA extractions of each individual instead of the same DNA extraction several times, specially when DNA restrictions are involved in the type of marker.

We have observed a higher yield of polymorphisms when using markers derived from 4 retrotransposon sequences than when using markers derived from "random" and microsatellite primers: only two primers in the second group, out of the 70 primers or primer-combinations studied (2.4 %), differentiated consistently 2 varieties. On the contrary, 6 polymorphisms were detected out of

the 41 reactions performed with the *copia*-related primers (14.6%). Similar results have been obtained in barley and pea with other members of the *Copia* class of retrotransposons: Waugh et al. (1997) and Ellis et al. (1998) found that S-SAPs reveal a higher level of polymorphism than AFLPs, and Kalendar et al. (1999) detected more intraespecific variability with IRAPs than with ISSRs.

No variability was found in clementines with ISSRs, SSRs or AFLPs. It is particularly striking that AFLPs, believed to cover a high proportion of the genome because of the high number of bands scored in each analysis, failed to detect any variability in this study. On the other hand, although it has been concluded that ISSR markers represent a high potential for the characterisation of citrus cultivars and they have been successfully used to differentiate sweet orange varieties (Fang and Roose 1996; 1997), they did not reveal any variability among the clementine varieties, in spite of somatic mutations being the origin of the new cultivars in both species. The genetic variability among sweet orange cultivars must be much larger than among clementines because the culture of sweet orange is more widely extended (Brasil, USA, China and Spain, mainly) and older than the culture of clementines, mostly restricted to Spain and some other Mediterranean countries.

The case for SSRs is different: microsatellite sequences evolve very fast and they have the highest frequency of polymorphism when compared to other marker types, given the high number of alleles that can be found at a single locus (i.e., up to 37 in barley and 36 in kiwi (Saghai-Maroof et al. 1994; Weising et al. 1996). But the lack of variability detected in clementines by SSRs can be explained by the origin of the accessions used in this study, the selection of fortuitous somatic mutations. Therefore, the cause of so many alleles per microsatellite locus in some cultivated species must be the unequal, or asymmetric, meiotic recombination within sexual reproduction, and not somatic

recombination. Under such conditions, SSRs are not superior to RAPDs or other marker systems for genotype identification but, instead, highly inefficient.

The use of specific primers to amplify intertransposonic sequences has been very productive; in addition, IRAP reactions with 2 such primers yielded a high number of bands (several of them being polymorphic). When comparing polymorphisms detected by RAPD versus IRAP analysis, it is noticeable that IRAPs usually affect several closely related cultivars simultaneously while RAPDs affect one cultivar at a time. The analysis of genetic relationships among clementine accessions (Figures 2 and 3), mostly based in the polymorphisms associated with *copia*-like retrotransposons, generates groupings that agree with what is known about their origin. The cultivar "Fina" seems to be the original one from which the rest of Spanish accessions have derived by spontaneous mutations (Bono, 1996). A sequence of mutation events produced the lineage: "Fina" to "Oroval" to "Marisol" to "Loretina". The 3 last cultivars share some features that differentiate them from "Fina": earlyness in harvest time, up-right growing habit and presence of thorns; "Loretina", in its turn, differs from "Oroval" and "Marisol" in being less prone to yielding puffy mandarines and in the proliferation of latent buds at the grafting point. The known origin of these clementine cultivars agrees with the phylogenetic diagram depicted in Figure 3, which is based on the accumulation of mutations. This accumulation would explain the reduction or loss of IRAP bands from the oldest Spanish cultivar "Fina" (or even from a group of cultivars from North Africa, where "Fina" originated) to the most recent one, Loretina. Except for change at *5R4Rd,* only mutations involving RAPD loci correspond to a new band.

The greatest amount of discriminative changes among the clementine cultivars studied correspond to mutations at retrotransposon (RT) sequences (Figure 3). Retrotransposons are a class of dispersed middle repetitive sequences and have contributed to the genetic diversity of their host species. Three

mechanisms are involved: transposition, homologous recombination between retrotransposons or LTRs, and frequent mutation of methylated cytosine to thymine in retrotransposon sequences. DNA methylation is a frequently reported mechanism controlling transposable elements (Finnegan et al. 1998). A higher mutation rate at RT sequences (probably because they become methylated) and their abundace in the citrus genome (Asíns et al. 1999), would explain why IRAP markers generate more polymorphisms than the other types of markers in *C. clementina*. Therefore, IRAP markers must be considered to be a fast and efficient way of obtaining polymorphisms for identification and linkage studies in this species. The citrus genome contains many retrotransposons of the *copia* and other families that we have not used in this study but that might explain, at the molecular level, the variation not accounted for yet among the clementine accessions.

The process of differentiation in clementines is fairly recent (i.e., "Arrufatina" appeared in 1968, "Oronules" in 1970; "Esbal" in 1966), and some genotypes are highly unstable, such as "Clemenules", which appeared in 1953 and produced the new varieties "Arrufatina", "Clemenpons" and "Mioro" since. Our results suggest that changes at *copia-like* retrotranposons (three out of 4 sequences checked) have facilitated the diversification of clementines. Most plant retrotransposons do not transpose under normal circumstances, but their proliferation can be set off by stress conditions, such as pathogen infection, wounding and cell culture (Moreau-Mihri et al. 1996; Hirochika. 1997; Takeda et al. 1998; Mihri et al. 1999). It is remarkable that a stress factor, which is by definition a selective force reducing variability, allows, at the same time, the increase of genetic variability by activating the mechanism of transposition. Nevertheless, this is not the end of retrotransposons as source of genetic variability because a higher mutation rate, likely due to their methylation, would explain both the high intraespecific heterogeneity of these sequences in *Citrus*

and other genera (Flavell et al. 1992; Brandes et al. 1997; Asíns et al. 1999) and the higher polymorphism of IRAPs versus other markers found in *C. clementina*, a vegetatively propagated crop species. If these phenomena and their agronomic implications were understood, a new and useful source of variation would be available for citrus breeding.

AKNOWLEDGEMENTS

We are grateful to Drs. Sergio Bramardi and Emilio Carbonell, for the statistical analysis and helpful comments and to Dr. Luis Navarro for allowing us the use of the Citrus Germplasm Bank at IVIA. This work was supported by research grants INIA (SC95-034), Ministerio de Educación y Ciencia (MPB) and IVIA (CR) .

6. RESUMEN

Los cítricos son las especies frutales más importantes a escala mundial, siendo actualmente España el cuarto país productor. Esto contrasta con la escasez de estudios genéticos para su mejora, lo cual es debido a que presentan varios problemas, entre ellos su compleja biología reproductiva. Muchas especies de cítricos son apomícticas y sus semillas producen embriones nucelares que limitan el desarrollo de embriones zigóticos y por tanto, la construcción de familias segregantes. En los programas de mejora de cítricos y en los viveros es muy importante distinguir entre individuos zigóticos y nucelares para poder eliminar los genotipos no deseados. Normalmente, los isoenzimas se han empleado para determinar el origen genético de las plantas jóvenes. En el primer trabajo se propone el uso de marcadores microsatélites como una metodología alternativa y los comparamos con los isoenzimáticos para la distinción de plántulas zigóticas. Con objeto de facilitar el análisis de microsatélites se desarrolló también un protocolo rápido de extracción de ADN y un método de alta resolución y fácil revelado sin utilizar fluorocromos ni radioisótopos.

Se emplearon dos familias segregantes: una derivada de un cruce interespecífico entre *Poncirus trifoliata* (L.) Raf. Var. "Flying Dragon" y el tangor "Ortanique" (*Citrus reticulata* (Blanco) x *Citrus sinensis* (L.) Osb.), familia PxO, la otra familia analizada fue obtenida por autofecundación de *Fortunella crasifolia* Swing., familia Fc.

En el cribado de las plántulas PxO únicamente fue necesaria la utilización de un marcador microsatélite, ya que con éste se pueden distinguir los 4 alelos presentes en la población. En esta misma familia se analizaron otros microsatélites al igual que varios isoenzimas, y con todos se obtuvieron los mismos resultados, el 87% de los individuos de la familia son zigóticos. En familias obtenidas por cruzamiento entre especies muy distantes filogenéticamente, cualquier marcador codominante suele ser útil, sin embargo,

cuando los parentales están genéticamente más relacionados, es más difícil encontrar variabilidad en cada locus. Por esto, la eficiencia de los microsatélites aumenta respecto a la de isoenzimas, debido a que presentan mayor grado de polimorfismos. Un caso extremo es la familia Fc, que esta formada por 106 individuos y fue generada por autofecundación. Se analizaron 5 microsatélites y 5 isoenzimas, y se encontraron 6 individuos zigóticos empleando los microsatélites, mientras que con los isoenzimas sólo se pudieron detectar 3 de los 6 observados con los microsatélites. Por tanto, la conclusión de este trabajo es que el empleo de microsatélites presenta mayor eficiencia que el de marcadores isoenzimáticos para identificar plántulas de origen sexual en los cítricos o, en general, plantas de origen desconocido.

Una vez comprobada la gran utilidad que tienen los microsatélites en la mejora genética por su nivel de polimorfismo, rapidez, sencillez y gran repetibilidad que poseen, nos planteamos la obtención de nuevos microsatélites de cítricos. Con este objetivo se emplearon dos estrategias: la construcción de una genoteca de ADN genómico de *P. trifoliata* con fragmentos de pequeño tamaño, y la estrategia bioinformática consistente en cribar microsatélites empleando el programa FINDPATTERNS en todas las secuencias de cítricos incluidas en GenBank, en las cuales se buscaron todas las posibles repeticiones de di-, tri-, tetra- y pentanucleótidos.

En total se obtuvieron 33 nuevos microsatélites, 6 en el cribado de la genoteca y el resto a partir de las secuencias de la base de datos. Estos nuevos marcadores se emplearon para la realización de mapas genéticos en tres familias segregantes:

- Autofecundación de *Poncirus trifoliata* (var. Flying Dragon), Familia PpxPp
- *Citrus aurantium* x *Poncirus trifoliata* (var. Flying Dragon), Familia AxPa
- *Citrus volkameriana* x *Poncirus trifoliata* (var. Rubidoux), Familia VxPv

Se construyó un mapa para cada uno de los parentales de cada familia, que poseen entre 48 y 120 marcadores, dependiendo de la heterozigosis de cada parental. El análisis comparativo entre genomas fue posible gracias al uso de los marcadores SSR principalmente, observándose varias reorganizaciones entre las distintas especies, así como entre las dos variedades de *Poncirus trifoliata*. Se observó que el orden de los marcadores en el mapa puede variar al bajar el LOD. Puesto que para futuros análisis genéticos lo más importante es que la ordenación sea la correcta, se prefirió construir los mapas a LOD alto a costa de dejar marcadores fuera de los grupos de ligamiento.

Para la realización de los mapas de ligamiento también se usaron otros tipos de marcadores como los IRAPs, que aunque son menos informativos porque no son codominates, producen gran número de polimorfismos muy repetitivos con pocas combinaciones de cebadores y su distribución en el genoma es, en general, aleatoria.

El análisis de segregación a nivel genómico puso de manifiesto la distorsión en la segregación en ciertas zonas concretas, que corresponderían a problemas gaméticos y/o zigóticos, es decir, presencia de factores letales.

Una vez obtenidos los nuevos marcadores SSRs e IRAPs en cítricos, y sabiendo su localización genómica, nos propusimos utilizarlos para investigar sobre el origen de la variación molecular en el grupo de los mandarinos clementinos, debido a su importancia económica y por tratarse de un reto, ya que es una especie que se propaga vegetativamente, por lo que la variabilidad genética es muy limitada. En general, las nuevas variedades son detectadas por los propios citricultores mediante la identificación de mutaciones somáticas que afectan caracteres agronómicos, tales como la época de recolección.

Se emplearon distintos tipos de marcadores con objeto de comparar su nivel de polimorfismo en una colección de 24 variedades. No se observó ningún

polimorfismo al emplear SSR, por lo que la recombinación somática no debe ser una fuente importante de variabilidad en esta especie. Cuando se emplearon ISSRs, RAPDs y AFLPs, sólo un 2.4% de las combinaciones de cebadores produjeron polimorfismos, en comparación con los generados al emplear IRAPs (14.6%). Además, el diagrama evolutivo basado en estos últimos polimorfismos se ajusta bastante bien al origen conocido de algunas variedades.

Por tanto, se concluye que los cambios en el ADN de los retrotransposones y las zonas adyacentes son más frecuentes que en el resto del ADN estudiado. Esto sugiere que los factores que provoquen la actividad de retrotransposones, como condiciones bióticas o abióticas de estrés, pueden ser una fuente importante de variabilidad genética a utilizar en programas de mejora de especies de propagación vegetativa, donde abunden los elementos de este tipo, como es el caso de los naranjos navel o los mandarinos clementinos.

7. CONCLUSIONES

- La selección de individuos zigóticos mediante marcadores es más eficiente cuando se emplean microsatélites que al emplear isoenzimas, especialmente cuando se trata de una población derivada de autofecundación y el nivel de heterozigosis es bajo.

- La obtención de microsatélites por el método de cribado en las bases de datos de secuencias es más rentable. En total se han desarrollado y mapeado 33 nuevos microsatélites. Estos marcadores codominantes han permitido realizar un análisis genómico comparado.

- El porcentaje de marcadores en heterozigosis en *P. trifoliata* es mucho menor que el observado en especies de *Citrus*. Así mismo, la heterozigosis para marcadores SSR es mucho mayor que para IRAP, lo que apoyaría la hipótesis de una mayor tasa de cambio en SSR frente a IRAP.

- Las secuencias de retrotransposones se distribuyen aleatoriamente en el genoma, aunque existen regiones cromosómicas donde son especialmente abundantes. El número de copias depende de la secuencia de retrotransposón y de la especie estudiada.

- Se produce una distorsión en la segregación de los marcadores entre un 22.9% y un 39.7%, según la especie que se esté analizando. En la familia obtenida por autofecundación de *P. trifoliata* se ha estudiado el origen de la distorsión y se han observado grupos de ligamiento con distorsión gamética, con distorsión zigótica y con ambas. La existencia de distorsión zigótica sugiere la presencia de factores letales recesivos en el genoma.

- Se han observado diferentes reorganizaciones al comparar los mapas genéticos de las tres especies, así como al comparar la dos variedades de *P. trifoliata*. Estas reorganizaciones deben ser tenidas en cuenta en la

aplicabilidad, a través de especies, de los análisis genéticos basados en mapas de ligamiento.

- Existen factores que afectan a la robustez de los mapas de ligamiento, tales como el criterio LOD elegido, la presencia de factores letales, el tipo de marcador empleado y la fase de ligamiento de los marcadores.

- La amplificación de secuencias adyacentes a retrotransposones del grupo *copia* en clementinos produce una proporción mucho mayor de polimorfismos que el uso de marcadores basados en microsatélites o cebadores de secuencia aleatoria.

- El análisis filogenético basado principalmente en los polimorfismos asociados a retrotransposones da lugar a agrupamientos que coinciden con la historia conocida sobre cómo se originaron algunas de las variedades de clementinos.

- Los factores que provoquen la actividad de retrotransposones tales como condiciones bióticas y abióticas de estrés, pueden utilizarse como una fuente importante de variabilidad genética en programas de mejora de especies de propagación vegetativa cuyo genoma contenga numerosas copias de estos elementos, como son los principales grupos de variedades de cítricos.

8. BIBLIOGRAFÍA

Allard RW (1960) Principles of plant breeding. John Willey and Sons, New York

Ananiev EV, Phillips RL, Rines HW (1998) Complex structure of knob DNA on maize chromosome 9: retrotransposon invasion into heterochromatin. Genetics 149:2025-2037

Anderson CM, Castle WS, Moore GA (1991) Isozymic identification of zygotic seedlings in Swingle citrumelo *Citrus paradisi* x *Poncirus trifoliata* nursery and field populations. J Amer Soc Hort Sci 116:322-326

Angers B, Bernatchez L (1997) Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. Mol Biol Evol 14:230-238

Arus P, Tanksley SD, Orton TJ, Jones RA (1982) Electroforetic variability as a tool for determining seed purity and for breeding hybrid varieties of *Brassica oleracea*. Euphytica 31:417-428

Ashari S, Aspinall D, Sedgley M (1988) Discrimination of zygotic and nucellar seedlings of five polyembryonic citrus rootstocks by isozyme analysis and seedling morphology. J Hort Sci 63:695-703

Asíns MJ, Herrero R, Navarro L (1995) Factors affecting tree isozyme-gene expression. Theor Appl Genet 90: 892-898

Asíns MJ, Mestre PF, Herrero R, Navarro L, Carbonell EA (1998) Molecular markers: a continously growing biotechnology area to help citrus improvement. Fruits-Paris 53:293-302

Asíns MJ, Monforte AJ, Mestre PF, Carbonell EA (1999) *Citrus* and *Prunus copia*-like retrotransposons. Theor Appl Genet 99:503-510

Beidler LL, Hilliard PR, Rill RL (1982) Ultrasensitive staining of nucleic acids with silver. Anal Biochem 126:374-380

Bell CS, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of arabidopsis. Genomics 19:137-144

Bitonti MB, Cozza R, Wang G, Ruffini-Castiglione M, Mazzuca S, Castiglione S, Sala F, Innocenti AM (1996) Nuclear and genomic changes in floating and submerged buds and leaves of heterophyllous waterchestnut (*Trapa natans*). Physio Plant 97:21-27

Bono R (1996) Variedades cultivadas. Comunitat Valenciana Agraria 4:281-302

Brandes A, Heslop-Harrison JS, Kamm A, Kubis S, Doudrick RL, Schmidt T (1997) Comparative analysis of the chromosomal and genomic organization of *Ty1-copia*-like retrotransposons in pteridophytes, gymnosperms and angiosperms. Plant Mol Biol 33:11-21

Bretó MP, Ruiz C, Pina JA, Asíns MJ (2001) The diversification of *Citrus clementina* Hort. x Tan., a vegetatively propagated crop species. Molecular Phylogenetics and Evolution. In press

Brown PTH (1989) DNA methylation in plants and its role in tissue culture. Genome 31:717-729

Brown PTH, Göbel E, Lörz H (1991) RFLP analysis of Zea mays callus cultures and their regenerated plants. Theor Appl Genet 81:227-232

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, González-Candelas F, Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor Appl Genet 93:190-198

Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in *Citrus* using random amplified plymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. Theor Appl Genet 89:606-614

Cameron JW (1979) Sexual and nucellar embriony in F1 hybrids and advanced crosses of *Citrus* and *Poncirus*. J Am Soc Hort Sci 104:408-410

Cameron JW, Frost HB (1968) Genetics, breeding and nucellar embryony. In: Reuthen W, Webber HJ and Batchelor LD (eds). The citrus industry, vol. 2. University of California Press, Berkeley, California, pp 325-370

Cameron JW, Soost RK (1969) Citrus. In: Ferweda, F.D. and F. de Witt (Eds) Outlines of perennial crop breeding in the tropics. 129-162. Wageningen, H. Veenman and Zonen

Cavally-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Evolution 21:550-570

Cheng FS, Roose ML (1995) Origin and inheritance of dwarfing by *Citrus* rootstock *Poncirus trifoliata* "Flying Dragon". J Amer Soc Hort Sci 120:286-291

Coletta-Filho HD, Machado MA, Targon MLPN, Moreira MCPQDG, Pompeu J Jr (1998) Analysis of the genetic diversity among mandarins (*Citrus* ssp.) using RAPD markers. Euphytica 102:133-139

Cristofani M, Machado MA, Grattapaglia D (1999) Genetic linkage maps of *Citrus sunki* Hort. ex. Tan. and *Poncirus trifoliata* (L.) Raf. and mapping of citrus tristeza virus resistance gene. Euphytica 109:25-32

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version ii. Plant Mol Biol Rep 1:19-21

Donini P, Elias ML, Bougourd SM, Koebner RMD (1997) AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. Genome 40:521-526

Durham RE, Liou PC, Gmitter FG Jr, Moore GA (1992) Linkage of restriction fragment length polymorphisms and isozymes in *Citrus*. Theor Appl Genet 84:39-48

Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambose MJ (1998) Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. Mol. Gen. Genet. 260:9-19

Estoup A, Tailliez C, Cornuet JM, Solignac M (1995) Size homoplasy and mutational processes of interrupted microsatellites in two species, *Apis mellifera* and *Bombus terrestris* (Apidae). Mol. Biol. Evol. 12:1074-1084

Fang DQ, Roose ML (1996) Fingerprinting citrus cultivars with Inter-SSR markers. Proc Int Soc Citriculture 1:185-188

Fang DQ, Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theor Appl Genet 95:408-417

FAO (Food an Agriculture Organization) (2001) http://www.fao.org

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. Theor Appl Genet 96:812-822

Finnegan EJ, Genger RK, Peacock WJ, Dennis ES (1998) DNA methylation in plants. Annu Rev Plant Physiol Plant Mol Biol 49:223-247

Flavell AJ, Dunbar E, Anderson R, Pearce SR Hartley R, Kumar A (1992) *Ty1-copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. Nucleic Acids Res 20:3639-3644

Frankel OH (1958) The dynamics of plant breeding. Journal of the Australian Institute of Agricultural Science 24:112-123.

Frost HB (1943) Genetics and breeding. In: Webber, H.J. and L.D. Batchelor (Eds) The citrus industry. Vol.I, 817-913. Univ California Press, Berkeley, Los Angeles

Frost HB, Soots RK (1968) Seed reproduction development of gametes and embryos. In: Reuther, W., H.J. Webber ans L.D. Batchelor (Eds) The citrus industry. Vol II, 327-371. Univ California, Berkeley

García R, Asíns MJ, Forner J, Carbonell EA (1999) Genetic analysis of apomixis in *Citrus* and *Poncirus* by molecular markers. Theor Appl Genet 99:511-518

Gottlieb LD (1986) The genetic basis of plant form. Philos Trans R Soc Lond 313:197-208.

Guerra MS (1984) Cytogenetics of Rutaceae. II. Nuclear DNA content. Caryologia 37:219-226

Haldane JBS (1919) The combination of linkage values, and the calculation of distance between the loci of linked factors. J Genetics 8:299-300

Harding RM, Boyce AJ, Clegg JB (1992) The evolution of tandemly repetitive DNA: recombination rules. Genetics 132:847-859

Hederson ST, Petes TD (1992) Instability of simple sequence DNA in *Saccharomyces cerevisiae*. Mol Cell Biol 12:2749-2757

Herrero R, Asíns MJ, Carbonell EA, Navarro L (1996) Genetic diversity in the orange subfamily Aurantoideae. I. Intrspecies and intragenus genetic variability. Theor Appl Genet 92:599-609

Heslop-Harrison JS, Brandes A, Taketa S, Schmidt T, Vershinin AV, Alkhimova EG, Kamm A, Doudrick RL, Schwarzacher T, Katsiotis A, Kubis S, Kumar A, Pearce SR, Flavell AJ, Harrison GE (1997) The chromosomal distributions

of Ty1-*copia* group retrotransposable elements in higher plants and their implications for genome evolution. Genetica 100:197-204

Hirochika H (1997) Retrotransposons of rice: their regulation and use for genome analysis. Plant Molec Biol 35:231-240

Hisada S, Akihama T, Endo T, Moriguchi T, Omura M (1997) Expressed sequence tags of *Citrus* fruit during rapid cell development phase. Journal of the American Society for Horticultural Science 122:808-812

Iannelli D, Cottone C, Viscardi M, D'Apice L, Capparelli R, Boselli M (1998) Identification of genotypes of lemon by flow cytometry and RAPD markers. Intnt J Plant Sci 159:864-869

Iglesias L, Lima H, Simon JP (1974) Isozyme identification of zygotic and nucellar seedlings in *Citrus*. J Hered 65:81-84.

Jarrell DC, Roose ML, Traugh SN, Kupper RS (1992) A genetic map of citrus based on segregation of isozymes and RFLPs in an intergeneric cross. Theor Appl Genet 84:49-56

Jones AP, Buckley JD (1990) The role of DNA methylation in cancer. Adv Cancer Res 54:1-23

Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A. (1999) IRAP and REMAP: two new retrotransposon–based DNA fingerprinting techniques. Theor Appl Genet 98:704-711

Karhu A, Hulme P, Karjalainen M, Karvonen P, Karkkainen K, Neale D, Savolainen O (1996) Do molecular markers reflect patterns of differentiation in adaptative traits of conifers? Theor Appl Genet 93:215-221

Kashi Y, Tikochinsky Y, Genislav E, Iragi F, Nave A, Beckmann JS, Gruenbaum Y, Soller M (1990) Large restriction fragments containing poly-TG are highly polymorphic in a variety of vertebrates. Nucleic Acids Res 18:1129-1132

Khan IA, Roose ML (1988). Frequency and characteristics of nucellar and zygotic seedlings in three cultivars of trifoliate orange. J Am Soc Hort Sci 113:105-110.

Kijas JMH, Thomas MR, Fowler JCS, Roose ML (1997) Integration of trinucleotide microsatellites into a linkage map of *Citrus*. Theor Appl Genet 94:701-706

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eusgen 12:172-175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181

Lewis J, Bird A (1991) DNA methylation and chromatin structure. *FEBS Lett.* 285:155-159

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397-401

Luro F, Laigret F, Lorieux M, Ollitrault P (1996) Citrus genome mapping with molecular markers: Two maps obtained by segregation analysis of progeny of one intergeneric cross. Proc Int Soc Citriculture 2:862-866

Mago R, Nair S, Mohan M (1999) Resistance gene analogues from rice: cloning, sequencing and mapping. Theor Appl Genet 99:50-57

MAPA (Ministerio de Agricultura Pesca y Alimentación) (2001) http://www.mapya.es

Matzke MA, Matzke AJM (1991) Differential inactivation and methylation of a transgene in plants by two suppressor loci containing hologous sequences. Plant Mol Biol 16:821-830

Mestre PF, Asíns MJ, Pina JA, Carbonell EA, Navarro L (1997) Molecular markers flanking citrus tristeza virus resistance gene from *Poncirus trifoliata* (L.) Raf. Theor Appl Genet 94:458-464

Michelmore R (1995) Molecular approaches to manipulation of disease resistance genes. Annu Rev Phytopathol 15:393-427

Mihri C, de Wit PJGM, Grandbastien MA (1999) Activation of the promoter of the Tnt1 retrotrnasposon in tomato after inoculation with the fungal pathogen *Cladosporium fulvum*. Molec Plant Microbe Interact 12:592-603

Moore GA, Castle WS (1988) Morphology and isozymic analysis of open-pollinated Citrus rootstock populations. J Hered 79:59-63.

Moreau-Mhiri C, Morel JB, Audeon C, Ferault M, Grandbastien MA, Lucas H (1996) Regulation of expression of the tobacco Tnt1 retrotransposon in heterologous species following pathogen-related stresses. Plant-Journal 9:409-419

Morgante M, Olivieri AM (1993) PCR-amplified microsatellite as markers in plant genetics. The Plant J 3:175-182

Mullis K, Fallona S, Schraf S, Saiki R, Horn G, Erlich H (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. Cold Spring Harbor Symp Quant Biol 51:263-273

Naithani SP, Raghuvanshi SS (1958) Cytogenetical studies in the genus *Citrus*. Nature 181:1406-1407

Orford SJ, Steele Scott N, Timmis JN (1995) A hypervariable middle repetitive DNA sequence from citrus. Theor Appl Genet 91:1248-1252

Parfitt DE, Arulsekar S, Ramming, DW (1985) Identification of Plum x Peach hybrids by isozyme analysis. Hort Science 20:246-248.

Perez JA, Maca N, Larruga JM (1999) Expanding informativeness of microsatellite motifs through the analysis of heteroduplexes: a case applied to *Solanum tuberosum*. Theor Appl Genet 99:481-486

Pina JA, Sevillano JJ, Polo A (2000) Plantas de viveros de cítricos en la Comunidad Valenciana. Comunitat Valenciana Agraria 15:23-32

Protopapadakis E, Papanikolaou X (1999) Use of four isoenzymatic systems in lemon and lemon-like citrus cultivars to detect their genetic diversity. J Hort Sci Biotechn 74:26-29

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends in Genetics 9(8):275-280

Raghuvanshi SS (1962) Cytogenetical studies in the genus *Citrus* IV. Evolution in genus *Citrus*. Cytologia 27:172-188

Raybould AF, Mogg RJ, Aldam C, Gilddon CJ, Thorpe RS,Clarke RT (1998) The genetic structure of sea beet (*Beta vulgaris* ssp maritima) populations. III. Detection of isolation by distance at microsatellite loci. Heredity 80:127-132

Roose ML, Fang D, Cheng FS, Tayyar RI, Federici CT, Kupper RS (2000) Mapping the citrus genome. Proc. 1st Internt. Citrus Biotech Symp. Israel.

Roose ML, Traugh SN (1988) Identification and performance of citrus trees on nucellar and zygotic rootstocks. J Am Soc Hort Sci 113:100-105

Ruiz C, Bretó MP, Asíns MJ (2000) A quick methodology to identify sexual seedlings in citrus breeding programs using SSR markers. Euphytica 112:89-94

Ruiz C, García MR, Gonzalez F, Asíns MJ (1999) Estudio comparativo de la organización genómica de *Poncirus trifoliata*, *Citrus volkameriana* y *Citrus aurantium.* Proc. II Congreso de la Sociedad Española de Genética. Diputación Provincial de A Coruña, A Coruña, Spain

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor Appl Genet 95:714-722

Saghai-Maroof MA, Biyashev GP, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc Natl Acad Sci USA 91:5466-5470

Senior ML, Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome 36:884-889

Simone M de, Russo MP, Puelo G, Marsan PA, Lorenzoni C, Marocco A, Recupero GR (1998) Construction of genetic maps for Citrus aurantium and C. latipes based on AFLP, RAPD and RFLP markers. Fruits 53:383-390

Smulders MJM, Rus-Kortekaas W, Vosman B (1995) Tissue culture-induced DNA methylation polymorphisms in repetitive DNA of tomato calli and regenerated plants. Theor Appl Genet 91:1257-1264

Sneath PHA, Soakal RR (1973) Numerical taxonomy. WH Freeman and Co. San Francisco (USA)

Soltis DE, Soltis PS (1989) Isozymes in plant biology. Advances in plant sciences series. Vol.4 Dudley TR (Gral ed) Dioscorides Press. Portland, Oregon

Spiegel-Roy P, Bardi A, Shani A (1977) Peroxidase isozymes as a tool for early separation of nucellar and zygotic Citrus seedlings. Proc Int Soc Citriculture 2:619-624

Stam P (1993) Constructionof integrated genetic linkage maps by means of a new computer package:JOINMAP. Plant 5:739-744

Stam P, Van Ooijen JW (1995) JOINMAP (tm) version 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Takeda S, Sugimoto K, Otsuki H, Hirochika H (1998) Transcriptional activation of the tobacco retrotransposon Tto1 by wounding and methyl jasmonate. Plant Molec Biol 36:365-376

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. Bio Technology 7:257-264

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucl Acid Res 17:6463-6471

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 17:6463-6471

Torres AM (1983) Fruit trees. In: Isozymes in plant genetics breeding. Tanksley SD and Orton TJ (eds). Elsevier Science Publishers, Amsterdam, pp. 401-421

Torres AM, Soost RK, Diedenhofen V (1978) Leaf isozymes as genetic markers in citrus. Amer J Bot 65:869-881

Torres AM, Soost RK, Mau-Lastovicka T (1982) Citrus isozymes. genetics and distinguishing nucellar from zygotic seedlings. J Hered 73:335-339

Trabut L (1902) L'hybridation des Citrus: una nouvelle tangerine "La Clémentine". Revue Horticole 74:232-234

Van Der Broeck D, Maes T, Sauer M, Zethof J, De Keukeleire P, D'Hauw M, Van Montagu M, Gerats T (1998) Transposon Display identifies individual transposable elements in high copy number lines. Plant Journal 13:121-129

Voorrips RE (2001) MapChart version 2.0: Windows software for the graphical presentation of linkage maps and QTLs. Plant Researc International, Wageningen, The Netherlands

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407-4414

Voytas DF, Cummings MP, Konieczny AK, Ausubel FM, Rodermel SR (1992) *Copia*-like retrotransposons are ubiquitous among plants. Proc Natl Acad SciUSA 89:7124-7128

Wang S, Liu N, Peng K, Zhang Q (1999) The distribution and copy number of *copia*-like retrotransposons in rice (*Oryza sativa* L.) and their implications in the organization and evolution of the rice genome. Proc Natl Acad Sci USA 96:6824-6828

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88:1-6

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W (1997) Genetic distribution of Bare-1-like retrotransposable elements in

the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet 253:687-694

Webber HJ (1932) Variation in citrus seedlings and their relation to rootstock selection. Hilgardia 7:1-79

Weber J, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the Polymerase Chain Reaction. Am J Hum Genet 44: 388-396

Weber JL (1990) Informativeness of human (dC-dA)n * (dG-dT)n polymorphisms. Genomics 7:524-530

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123-1128

Weising K, Fung RWM, Keiling J, Atkinson RG and Gardner RC (1996) Cloning and characterization of microsatellite repeats from *Actinidia chinensis*. Molec Breeding 2:117-131

White G, Powell W (1997) Isolation and characterisation of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. Molec Ecology 6:851-860

White MB, Carvalho M, Derse D, O'Brien SJ, Dean M (1992) Detecting single base substitutions as heteroduplex polymorphisms. Genomics 12:301-306

Xiang C, Roose ML (1988) Frecuency and characteristics of nucellar and zygotic seedlings in 12 citrus rootstocks. Sci Hort 37:47-59

Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around the *Tm*-2 locus of tomato during backcross breeding. Theor Appl Genet 77:353-359

Zamir D, Tanksley SD, Jones RA (1981) Genetic analysis of the origin of plants regenerated from anther tissues of *Lycopersicon esculentum*. Plant Sci Lett 21:223-227