TESIS DOCTORAL

La autocompatibilidad en el almendro (*Prunus amygdalus* Batsch): estructura genética del alelo S_f y modificaciones de su expresión

Angel Vte Fernández Martí

2010

Universitat de Lleida

Departament d'Hortofruticultura Botànica i Jardineria

TESIS DOCTORAL

La autocompatibilidad en el almendro (*Prunus amygdalus* Batsch): estructura genética del alelo S_f y modificaciones de su expresión

Memoria presentada por D. **Angel Vte. Fernández Martí**, Ingeniero Agrónomo, para optar al grado de Doctor Ingeniero Agrónomo.

Directores:
Rafael Socias i Company — **Dr. Ingeniero Agrónomo**
José M. Alonso Segura — **Dr. Ingeniero Agrónomo**

Tutor:
Valero Urbina Vallejo — **Dr. Ingeniero Agrónomo**

Realizada en la Unidad de Fruticultura del CITA

Zaragoza, Mayo de 2010

cita
CENTRO DE INVESTIGACIÓN Y TECNOLOGÍA AGROALIMENTARIA DE ARAGÓN

GOBIERNO DE ARAGON

A mi familia y a Carol

AGRADECIMIENTOS

Quisiera expresar mi más sincero agradecimiento al Dr. Rafael Socias i Company, por su inmejorable dirección de esta tesis, su ánimo, apoyo y optimismo en los momentos más difíciles a lo largo de estos cuatro años. Gracias por la total confianza que has depositado en mí desde el principio, tus consejos y disposición en todo momento. Ha sido todo un lujo haberte tenido como director de esta tesis. Igualmente quería agradecerte el empujón final que me has dado para poder finalizar mi doctorado antes de que me concluyese la beca.

De igual modo quiero agradecer al Dr. José Manuel Alonso Segura su tiempo y dedicación aportado a este trabajo. Gracias igualmente por todos tus consejos y apreciaciones a lo largo de estos años.

Al Dr. Valero Urbina, por su gran amabilidad durante la tutela de esta tesis y la ayuda ofrecida para su elaboración.

Agradecer al Director del CITA de Aragón, Dr. José González Bonillo, y a la Directora de Investigación, Dra. Pilar Errea Abad el haber aceptado la realización de esta tesis en su Centro.

Al Ministerio de Educación y Ciencia (Ciencia e Innovación) así como al Fondo Social Europeo, por haber financiado este proyecto y la beca FPI que ha permitido la realización de esta tesis.

Al Dr. Ryutaro Tao, por su calurosa acogida en el laboratorio de la Universidad de Kyoto (Japón), y por todas sus enseñanzas en biología molecular y en la vida en general. Muchas gracias por todo, Tao-sensei. Igualmente quisiera agradecer a la Dra. Hisayo Yamane todos los conocimientos que me ha transmitido durante mis estancias en sucesivos años. A mi gran amigo Toshio Hanada. Ha sido un verdadero placer trabajar a tu lado, 'hasta altas horas de la madrugada', menos mal que siempre acabábamos con un buen 'shake'!!!. A Matsumoto-san y Akagi-san, por enseñarme tantas cosas de laboratorio.

Al Dr. Pere Arús, por su amable recibimiento en el IRTA de Cabrils durante mis sucesivas estancias. Al Dr. Werner Howad, gracias por enseñarme el fabuloso mundo de los 'micros' y de los 'mapas' y por haberte implicado tanto en mi tesis. A Esteve, Gisela, Andrej y, a todos los compañeros de Cabrils que me han hecho sentir como en casa.

De igual modo, quiero agradecer a la Dra. Michelle Wirthenson de la Universidad de Adelaida (Australia), con quien he compartido muy buenos momentos durante mi estancia en su departamento. A Maria Hrmova, por haberme enseñado la aplicación del 3D a nuestras proteínas.

Al Dr. Federico Dicenta, por la cesión de material vegetal.

Al personal de la unidad de Fruticultura del CITA de los que he recibido ayuda. En especial a Mariajo (gracias por todos tus consejos), A. Pina, Mayte, Olga, Julio, Sara, José Miguel, Quique, Raúl y Pepe Jaime. Gracias también a P. Tomey y a Rosa.

A Jesús y a Susana, por mantener las plantas en las mejores condiciones posibles. A P. Bergua 'la secre', por ser tan ágil, atenta y eficaz en las gestiones administrativas. Gracias a Pili, por mantener siempre el área de trabajo tan limpio y reluciente, ¡así da gusto 'currar'!. A Ossama, por sus sugerencias aportadas en esta tesis.

A todos los becarios con los que he compartido sala, y en especial a mi gran amiga Engracia. Gracias por ser como eres, por tu incondicional apoyo y sobre todo por haber demostrado siempre tu gran personalidad (y no haberte dejado nunca influir por nadie!!!).

A las compañeras de Hortícolas, por compartir laboratorio y reactivos.

A las nuevas compañeras del 'Parque', en especial a Sara y a Yoko, por la confianza que habéis depositado en mi.

A mis tíos, padrinos y primos. Gracias por haber estado ahí en todo momento. A mis abuelos, y aunque no tuve la oportunidad de conocer a uno de ellos, sé que hubiera estado muy orgulloso de saber que su nieto trabajaría en el almendro, ya que él fue un agricultor y amante de este cultivo.

A toda mi familia de Vila-Real, en especial a Montiel y José. Gracias por vuestro apoyo incondicional.

A mi hermana Loles, a mis cuñados (María, Javi y Luis), gracias por apoyarme en cada momento difícil.

A mis Padres, Mª Dolores y Angel. Gracias por haberme dado siempre todo lo que he necesitado. Gracias por vuestro continúo apoyo y confianza depositada en mí. Nunca os podré recompensar todo esto. Sois un ejemplo a seguir!.

A Carol, por ser tan especial comigo, por cuidarme, por esperarme tanto tiempo. Por todos los años vividos. Esta Tesis va por ti.

A todos vosotros que habéis confiado en mi, **GRACIAS**!

CERTIFICACIÓN DE LOS DIRECTORES DE LA TESIS

D. Rafael Socias i Company, Dr. Ingeniero Agrónomo, Jefe de la Unidad de Fruticultura del Centro de Investigación y Tecnología Agroalimentaria (CITA) de Aragón, de la Consejería de Ciencia, Tecnología y Universidad del Gobierno de Aragón, y D. José M. Alonso Segura, Dr. Ingeniero Agrónomo, Investigador Agrario del CITA de Aragón,

CERTIFICAN

Que la tesis doctoral titulada "La autocompatibilidad en el almendro (*Prunus amygdalus* Batsch): estructura genética del alelo S_f y modificaciones de su expresión" ha sido realizada íntegramente por el Ingeniero Agrónomo D. Àngel Vicente Fernández i Martí bajo nuestra dirección en dicha Unidad de Fruticultura, habiéndose cumplido todos los objetivos planteados, por lo que se autoriza su presentación.

Zaragoza, 23 de mayo de 2010

Rafael Socias i Company José M. Alonso Segura

VºBº, el Tutor

Dr. Valero Urbina Vallejo

ABSTRACT

Almond, *Prunus amygdalus* Batsch, is a diploid species of the genus *Prunus*. Although self-incompatibility (SI) is an important trait in plants from the evolutionary point of view, self-compatibility (SC) is a critical agronomic trait in order to ensure the production of an economic crop. Thus, SC has become one of the main objectives worldwide of almond breeding programmes. SC has been identified in several almond cultivars, although it happens rarely. Traditionally tests of pollen tube growth and fruit set after controlled pollinations have been carried out for determining SC cultivars in almond. However, molecular methods have been developed in the last two decades to speed up the genotype determination in cultivars and selections in order to facilitate orchard management and breeding processes. SC has been so far only related with the presence of the S_f allele, allelic with the *S* series of SI alleles and dominant over SI. As a consequence, SC/SI has always been considered as a qualitative trait, controlled by a single multi-allelic locus, called the *S*-locus.

The presence of SC of the seedling in the family 'Vivot' × 'Blanquerna' has been 25%, much less than expected, 50%. The two different versions of the S_f haplotype, one active and being SI (S_{fa}) and the other inactive thus being SC (S_{fi}), have shown the recognition of the S_{fi}-pollen tubes of 'Blanquerna' by the S_{fa}-RNase of 'Vivot'. These results suggest the presence of a modifier locus unlinked to the *S*-locus which would control the SI/SC recognition mechanism. The molecular nature of the *S*-locus has been widely studied in many species. Although the *S*-locus encodes genes determining the *S*-specificity, other non-*S*-specific genes or modifier genes are required for the *S*-RNase-based SI reaction. Two additional loci located outside the *S*-locus are here described for the first time in the rosaceous family. Thus, we may suggest that SI is a quantitative trait rather than a qualitative one.

In addition, the construction of a fosmid library in 'Vivot' and 'Blanquerna' has allowed to confirm that the alignment of their S_f-RNases and SFB_f were identical. As it was not possible to find any difference between both versions of the S_f haplotype, it was decided to further investigate about DNA methylation, which is involved in changes in phenotype or gene expression caused by mechanisms other than changes in the DNA sequence. Thus, our experiments have allowed suggesting that DNA methylation could be responsible of the activation/inactivation of the S_f haplotype. In fact, when the S_f-RNase sequence is methylated, an inhibition of the expression takes place, as it happens in 'Blanquerna'. Thus, this inactivation would be traduced into self-compatibility. In the case of S_f-RNase sequences with non-methylated cytosines, as it happens in 'Vivot', the RNase would remain active and, as a consequence, it would be self-incompatible.

On the other hand, we have constructed the three-dimensional structures of the almond S_f, S_{23} and S_8 RNases trough molecular modelling tools. The main structural difference found between all RNases was that the S_f structure showed an extended looping region. The amino acid residues forming this long loop could be prone to degradation and/or inactivation and as a consequence this S_f-RNase could be less stable and thus would allow its pollen tube growth trough its own pistil.

RESUMEN

El almendro, *Prunus amygdalus* Batsch, es una especie diploide que pertenece al género *Prunus*. Aunque la autoincompatibilidad (AI) es un carácter importante en el reino vegetal desde el punto de vista evolutivo, la autocompatibilidad (AC) es un carácter agronómico muy crítico para asegurar la producción. De este modo, la AC se ha convertido en uno de los principales objetivos de los programas de mejora de almendro en todo el mundo. Aunque la AC ha sido identificada en algunas variedades de almendro, es muy poco frecuente. Los métodos que se han utilizado para determinar la AC en el almendro son la observación de los tubos polínicos y el cuajado de frutos después de polinizaciones controladas. Sin embargo, las nuevas técnicas moleculares se han ido aplicando cada vez más durante estas dos últimas décadas. De este modo, se ha conseguido determinar el genotipo de las variedades y selecciones de una manera muy rápida y eficaz. La AC ha sido relacionada con la presencia del alelo S_f, el cual pertenece a la serie alélica *S*, y es dominante sobre los otros alelos de AI. Por ello siempre se había considerado que la AI/AC era un carácter cualitativo.

En el transcurso de esta tesis, sólo un 25% de los individuos procedentes del cruzamiento 'Vivot' × 'Blanquerna' ha sido AC, mucho menos de lo esperado, un 50%. Las dos diferentes versiones del haplotipo S_f, una activa y AI (S_{fa}) y la otra inactiva y AC (S_{fi}) han mostrado un mutuo reconocimiento, por lo que la S_{fa}-RNasa del estilo de 'Vivot' ha impedido el crecimiento del tubo polínico S_{fi} de 'Blanquerna'. Estos resultados sugieren la presencia de un locus modificador no ligado al locus *S*, el cual sería responsable del control de la AI/AC en esta población. Aunque últimamente se ha profundizado mucho en el estudio del locus *S* a nivel molecular, su mecanismo de acción sigue sin estar del todo determinado, por lo que se sugiere que otros genes externos son necesarios en la AI. En esta tesis se han localizado por primera vez en la familia de las rosáceas dos nuevos loci situados fuera del locus S, los cuales se encuentran en el grupo 6 y 8. Con estos resultados, podríamos sugerir que la AI es un carácter cuantitativo y no cualitativo, como se conocía hasta la fecha.

Por otro lado, la construcción de una librería genómica (fósmido) en 'Vivot' y 'Blanquerna' ha permitido comprobar que tanto las secuencias de las S_f-RNasas como de los SFB_f eran completamente idénticas, a pesar de tener dos expresiones tan distintas. Como no ha sido posible encontrar ninguna mutación entre ambos S_f, se ha sugerido la implicación de la metilación del ADN en esta población. La metilación puede producir cambios en los fenotipos o en la expresión de los genes sin alterar las secuencias del ADN. Los ensayos llevados a cabo en esta tesis sugieren que el ADN metilado seria el responsable de la activación y/o inactivación del haplotipo S_f. De hecho, se ha podido comprobar que cuando la secuencia de la S_f-RNasa es metilada, ésta inactiva la expresión, por lo que esta inhibición se traduciría en una expresión AC, como es el caso de 'Blanquerna'. En el caso contrario, si la secuencia de ADN no es metilada, como es el caso de 'Vivot', la RNasa permanece activa y a consecuencia la planta será AI.

Además, se ha construido la estructura en 3D de las RNasas de los alelos S_f, S_{23} y S_8 mediante herramientas bioinformáticas. La principal diferencia encontrada fue que en la estructura del S_f, había un lazo más largo que en las otras dos RNasas AI. En la bibliografía se ha descrito que los lazos grandes son susceptibles a la degradación proteólica, por lo que este mismo fenómeno podría estar pasando en estas RNasas.

RESUM

L'ametller, *Prunus amygdalus* Batsch, és una espècie diploide que pertany al gènere *Prunus*. Encara que l'autoincompatibilitat (AI) és un caràcter important des del punt de vista evolutiu en el món vegetal, l'autocompatibilitat (AC) és un caràcter agronòmic crític per a assegurar la producció. D'aquesta manera, l'AC s'ha convertit en un dels principals objectius dels programes de millora d'ametller en tot el món. L'AC ha estat identificada en diverses varietats d'ametller, encara que sol ser excepcional. Els mètodes que s'han aplicat per a determinar l'AC són l'observació dels tubs pol·línics, i el quallat de fruits després de pol·linitzacions controlades. No obstant això, les noves tècniques moleculars s'han aplicat cada vegada més al llarg de les dues últimes dècades. Amb això s'ha aconseguit de determinar el genotip de les varietats i seleccions d'una manera molt ràpida i eficaç. L'AC ha estat relacionada amb la presència de l'al·lel S_f, el qual pertany a la sèrie al·lèlica *S*, i és dominant sobre els altres al·lels d'AI. Com a conseqüència d'això, sempre s'havia considerat que l'AI/AC era un caràcter qualitatiu, controlat per un únic locus, anomenat el locus *S*.

Només un 25% dels individus procedents de l'encreuament 'Vivot' × 'Blanquerna' ha estat AC, molt menys de l'esperat, un 50%. Les dues versions diferents de l'haplotip S_f, una activa i AI (S_{fa}) i l'altra inactiva i AC (S_{fi}), han mostrat el reconeixement per part de la S_{fa}-RNasa dels pistils de 'Vivot' dels tubs pol·línics amb el S_{fi} de 'Blanquerna'. Aquests resultats suggereixen la presència d'un locus modificador no lligat al locus *S*, el qual seria el responsable del control de l'AI/AC en aquesta població. Encara que darrerament s'hagi aprofundit en l'estudi del locus *S* a nivell molecular, aquest mecanisme no està encara del tot determinat, pel que se suggereix que d'altres gens serien un requisit per a l'AI. En aquesta tesi s'ha localitzat per primera vegada en la família de les rosàcies dos nous loci, situats fora del locus *S*, els quals es troben en els grups de lligament 6 i 8. Amb aquests resultats, podríem suggerir que l'AI és un caràcter quantitatiu i no qualitatiu, com s'acceptava fins ara.

D'altra banda, la construcció d'una llibreria genòmica (fòsmid) de 'Vivot' i 'Blanquerna' ens ha permès de comprovar que tant les S_f-RNases i les SFB_f eren completament idèntiques, a pesar de tenir dues expressions tan diferents. Com no ha estat possible de trobar cap mutació entre ambdós S_f, s'ha suggerit la implicació del la metilació de l'ADN en aquesta població. La metilació pot produir canvis en els fenotips o en l'expressió dels gens, sense alterar però les seqüències de l'ADN. Els assaigs realitzats en aquesta tesi suggereixen que l'ADN metilat seria el responsable d'aquesta activació i/o inactivació de l'haplotip S_f. De fet, s'ha pogut comprovar que quan la seqüència de la S_f-RNasa és metilada, aquesta inactiva l'expressió, així que aquesta inhibició es traduiria en una expressió AC, com és el cas de 'Blanquerna'. En el cas contrari, si la seqüència d'ADN no és metilada, com és el cas de 'Vivot', la RNasa roman activa i a conseqüència la planta és AI.

A més, s'ha construït l'estructura en 3D de les RNases dels al·lels S_f, S_{23} i S_8 mitjançant tècniques bioinformàtiques. La principal diferència trobada ha estat que en l'estructura del S_f, hi havia un llaç més llarg que en les altres dues RNases AI. En la bibliografia s'ha descrit que els llaços grans són susceptibles a la degradació proteòlica, pel que aquest mateix fenomen podria passar en aquestes RNases.

ÍNDICE

Pág.

ÍNDICE DE FIGURAS

Pág.

ÍNDICE DE TABLAS

Pág.

1. INTRODUCCIÓN

1. INTRODUCCIÓN

Entre las distintas especies frutales, el almendro es el fruto seco más cultivado en la Cuenca Mediterránea y ocupa un lugar destacado en la agricultura española. Su importancia se acentúa por la valoración de terrenos marginales, así como por las industrias locales de productos derivados e incluso por motivos ecológicos. Aunque es una especie originaria de Asia Central, el almendro está bien adaptado al clima mediterráneo templado, con inviernos húmedos y veranos secos. De entre todos los frutales del género *Prunus*, el almendro ha sido el primero en florecer hasta la obtención de las variedades de floración tardía y extra-tardía y es de los últimos en madurar, con la excepción de las variedades de melocotonero de maduración muy tardía, como los del tipo Calanda.

En los últimos años, se ha conseguido en España un aumento en los rendimientos pos superficie de las plantaciones de almendro, así como un aumento en la calidad de la pepita. Esto ha sido posible gracias a la introducción de nuevas variedades como resultado de los distintos programas de mejora y a la utilización de patrones mejor adaptados a las condiciones edafoclimáticas, en especial los nuevos híbridos almendro x melocotonero, así como a una mejora de las técnicas de cultivo y a una mayor dedicación al manejo de la plantación. Actualmente, se están incorporando nuevas técnicas moleculares en los programas de mejora genética, lo que probablemente va a facilitar el proceso de selección de genotipos que posean genes de interés, lo que repercutirá en un acortamiento del periodo de obtención de las nuevas variedades y en una difusión de las mismas, más rápida. No obstante, las nuevas variedades de almendro deben estar mejor adaptadas a las condiciones de cultivo de cada zona de producción y superar la calidad de las variedades actuales, de manera que se satisfagan las demandas de la industria y, finalmente, del consumidor.

1.1 Taxonomía, origen y evolución del almendro

El almendro cultivado se encuadra botánicamente dentro de la familia Rosaceae, género *Prunus* L., subgénero *Amygdalus* (L.) Focke y especie *Prunus amygdalus* Batsch. Aunque el Comité General de Nomenclatura Botánica propuso para el almendro

el nombre científico de *Prunus dulcis* (Mill.) D.A. Webb, se acepta otra denominación taxonómica más lógica como *P. amygdalus* Batsch. Otros sinónimos muy poco utilizados son *Amydgalus communis* L. y *Prunus communis* Arcang. Todos estos nombres pueden encontrarse en la bibliografía (Socias i Company, 1998a). Cabe destacar que el almendro es una especie diploide con un numero de cromosomas 2*n*=2x=16 (Fedorov 1969).

El origen del almendro ha sido motivo de diferentes hipótesis que han provocado diversas controversias. Se ha sugerido que, al igual que los otros frutales de hueso como el melocotonero y el albaricoquero, el almendro se originó en las regiones montañosas de Asia Central y algunas zonas del Asia Occidental. Numerosos botánicos rusos desarrollaron la teoría del origen de muchas especies frutales a partir de antepasados silvestres en toda la meseta central de Asia Central (Popov et al., 1929) a partir de las hipótesis de la evolución de las especies y sus centros de origen de Vavilov (1930). Grasselly (1976), después de sus expediciones en Asia Central y las observaciones de las distintas especies espontáneas de esta región, sugirió que el almendro cultivado procede de hibridaciones naturales entre las especies silvestres *P. fenzliana* Fritsch, *P bucharica* (Korsch.) Fedtsch. y *P. kuramica* Korsch. Aunque posteriormente Ladizinsky (1999) indicó que la única y exclusiva procedencia del almendro sería *P. fenzliana,* esta hipótesis ha sido fuertemente cuestionada (Socias i Company, 2004).

La expansión del almendro cultivado se produjo gracias a la proximidad de las poblaciones silvestres con centros de civilización en las montañas de Asia Central. Desde allí, el almendro fue distribuido en forma de semilla, ya que ésta era al mismo tiempo la unidad de propagación y la parte comestible, a través de las rutas comerciales de Persia, Mesopotamia y Grecia, utilizando sobre todo la ruta de la seda (Socias i Company, 1998b). Posteriormente los navegantes griegos, fenicios y romanos distribuyeron el almendro por toda la cuenca mediterránea. También se ha sugerido que el cultivo del almendro pudo haberse originado en el Mediterráneo oriental, al igual que otros cultivos como el olivo y el granado (Zohary y Hopf, 1993), aunque en esta zona no se encuentran especies tan cercanas al almendro como las presentes en Asia Central (Grasselly, 1976; Felipe, 2000). De todas formas se han encontrado almendras en la tumba de Tutankhamón (II milenio AC) (Zohary y Hopf, 1993) y también hay

evidencias de su comercio en el Mediterráneo occidental en el siglo IV AC (Cerdá Juan, 1973).

La evolución en la cuenca mediterránea fue lenta, marcada principalmente por la selección natural de los genotipos que mostraban una mejor adaptación a las condiciones ambientales de las zonas colonizadas, así como por la selección que ejercía el hombre, que sembraba las semillas de los almendros que mostraban un mejor comportamiento agronómico y con pepitas de buen sabor. La expansión de las zonas de cultivo mediante la siembra de semillas favoreció la heterozigosis en el almendro (Grasselly, 1972). Gracias a la generalización del injerto desde finales del siglo XIX, fueron apareciendo variedades locales adaptadas a sus zonas de cultivo. Es por ello que la Cuenca Mediterránea se considera un centro secundario de domesticación de la especie (Grasselly, 1972).

El cultivo del almendro se expandió en California después de 1836, a partir de semillas de ecotipos franceses que se habían llevado durante la época colonial. Kester et al. (1990) indican cómo el conjunto de las variedades californianas proceden de descendientes seleccionados de las variedades introducidas desde Europa lo que originó ecotipos locales. La magnífica productividad y calidad de la variedad 'Nonpareil', hizo que ésta se convirtiera rápidamente en la variedad de referencia y la más cultivada, utilizándose como genitor en los programas de mejora americanos. De ahí que el conjunto de variedades americanas presente caracteres relativamente homogéneos.

Posteriormente el almendro fue llevado a Australia, África del Sur y Sudamérica, en donde su cultivo se instaló en regiones específicas con un clima mediterráneo. El tipo de cultivo de estas nuevas regiones de expansión ha sido sobre todo con influencia de variedades y sistemas de cultivo de California.

La evolución de cada grupo de variedades en zonas geográficas distantes y aisladas las unas de las otras ha contribuido a esta variabilidad, por lo que se han ido formando ***ecotipos*** muy diferentes entre sí pero que conservan algunas características comunes (Grasselly y Crossa Raynaud, 1980), las cuales se citan a continuación:

Ecotipo español: Destacan dos grupos:

Grupo a): Variedades de floración precoz y de consistencia de la cáscara dura. Destaca 'Desmayo Largueta', originaria del Valle del Ebro, de fruto elíptico-amigdaloide, de buenas capacidades de adaptación a zonas de altitud media (300-750m) y cuyas pepitas son de excelente calidad y presentación. Apta para el tostado por la facilidad del desprendimiento de la piel así como para la repostería.

Grupo b): Destaca 'Marcona', originaria de la provincia de Alicante, de frutos redondeados, cáscara dura, buena aptitud para todo uso, apreciada para turrones y para aperitivos.

Ecotipo francés: Presentan una sola flor por yema y ausencia de almendras dobles. Suelen ser de floración muy tardía, aunque su época de maduración es precoz. Las variedades más importantes son 'Aï', 'Flour en Bas' y 'Tardive de la Verdière'.

Ecotipo italiano: Variedades de floración tardía y consistencia de la cáscara dura, con elevados porcentajes de pepitas dobles. Presencia de variedades autocompatibles como 'Tuono', 'Genco' y 'Filippo Ceo'.

Ecotipo tunecino: Típico de la región de Sfax, de clima particularmente seco con inviernos suaves. Se caracterizan por sus escasas necesidades en frío invernal, lo que les confiere una floración muy temprana. Son también de maduración precoz, adaptadas a un desarrollo de la pepita anterior a los inviernos áridos propios de la zona. Destacan las variedades 'Zahaf' y 'Achaak'.

Ecotipo americano: Se caracterizan por derivar casi todas de la variedad 'Nonpareil'. Por ello el material cultivado posee características comunes como cáscara blanda y época de floración media.

1.2 Producción mundial y nacional

1.2.1 Superficie y producción mundial

El almendro es la especie de mayor importancia entre todos los frutos secos, tanto en cuanto a producción como en cuanto a los intercambios comerciales. Actualmente se cultiva en todos los países de la Cuenca Mediterránea y de Oriente Próximo, pero también en otras zonas de clima mediterráneo con inviernos suaves y veranos secos, como California y las faldas del Himalaya y regiones similares en el hemisferio sur como Chile, Argentina, Sudáfrica y Australia.

En 2008, la superficie dedicada al cultivo del almendro en el mundo fue estimada en 1.868.672 ha, de las cuales 650.000 ha se encuentran en España, representando casi un 35% del total mundial (FAO, 2010). Le siguen EEUU (368.000 ha) con un 19.7% de superficie, Irán (172.000 ha) con un 9.2% y Túnez (160.000 ha) con un 8.6%. Desgraciadamente, a pesar de la primera posición de España en cuanto a superficie dedicada al cultivo, la producción real de pepita sitúa a España en la segunda plaza, superada claramente por los Estados Unidos.

En el año 2008, la producción de almendra en cáscara fue de 1.103.733 tm para EEUU, casi 5 veces superior a la española con casi 203.500 tm (FAO, 2009), seguida de Italia con 118.723 tm, Irán con 110.000 tm y Marruecos con 86.900 tm (Fig. 1.1). Esta diferencia sería todavía más evidente si las estadísticas reflejaran la producción de almendra en pepita, que es el producto comercialmente importante, en lugar de hacerlo como almendra en cáscara, ya que impide su comparación al ser muy diferentes los rendimientos en pepita de los distintos tipos de variedades de cáscara blanda, predominantes en EEUU, respecto a las de cáscara dura, mas comunes en la cuenca Mediterránea.

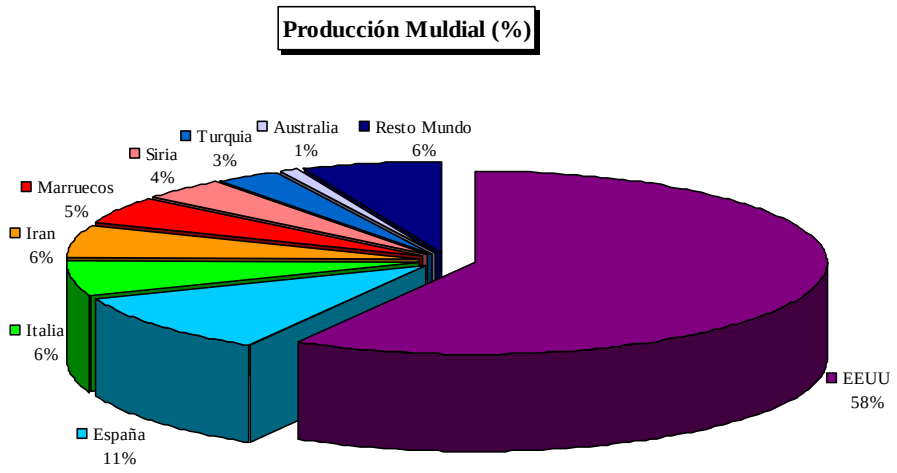


Figura 1.1: Distribución de la producción de almendra en cáscara por países en 2008.

En la figura 1.2 se puede apreciar el incremento de la producción de almendra de los dos principales países productores en los últimos 40 años. En ella se observa que el crecimiento en los Estados Unidos se ha realizado a un ritmo imparable mientras que la producción en España se mantiene prácticamente igual desde 1970.

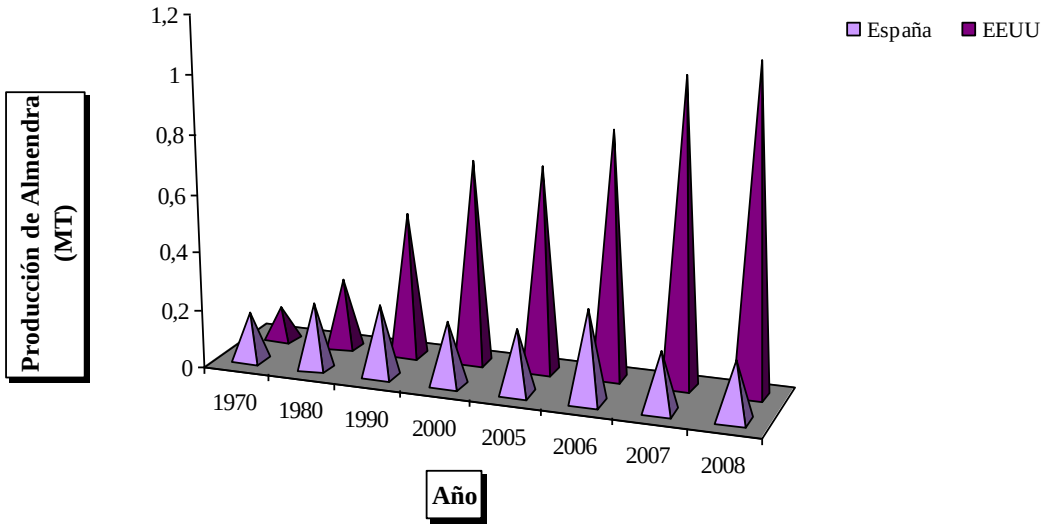


Figura 1.2: Evolución de la producción de almendra en cáscara en millones de toneladas (MT) durante las últimas 4 décadas en España y EEUU.

La gran diferencia de producción entre Estados Unidos y España se debe principalmente al diferente sistema de cultivo. En las plantaciones californianas la producción se basa en un reducido número de variedades, con mucha tecnificación y en regadío, mientras que en la zona mediterránea las plantaciones han sido predominantemente en secano, con poca o nula tecnificación, utilizando además variedades locales y de floración temprana, lo que ocasiona la pérdida ocasional de la cosecha por heladas. Debido a todos estos factores la productividad es baja, y al no ser las plantaciones rentables no se invierte en ellas, por lo que sufren un deterioro progresivo.

En California los costes de producción son altos, pero los grandes rendimientos obtenidos y el gran tamaño de las explotaciones, permiten que el cultivo sea muy rentable, ya que los precios del agua y del transporte son bajos, lo que hace que la almendra californiana domine el mercado mundial, no sólo por su volumen de producción, sino también por su precio.

1.2.2 Superficie y producción en España

Andalucía es la comunidad autónoma con la mayor superficie de almendro entre las principales zonas productoras de España, alcanzando 104.445 ha, seguida de Murcia (69.000 ha), Aragón (68.325 ha) y la Comunidad Valenciana (62.000 ha). Las comunidades con mayor producción son Murcia, Andalucía, Aragón, Comunidad Valenciana, Castilla la Mancha y Cataluña, que en conjunto representan el 90% de la producción total española (Cooperativas Agro-Alimentarias, 2009).

Aunque durante los últimos años la previsión era de un aumento del cultivo del almendro, ello no se refleja de una manera regular y constante, ya que como indican las últimas estadísticas proporcionadas por las Cooperativas Agro-alimentarias reflejadas en la figura 1.3, la producción de almendra en pepita sufre constantes altibajos de producción según el año. Ello es un reflejo de la incidencia de distintos factores adversos en la producción del almendro, especialmente los climatológicos. Destacan las heladas recurrentes del mes de marzo de 2004 (Kodad y Socias i Company, 2005), que produjeron la casi total destrucción de la cosecha de las variedades de floración

temprana y media, y también la de las de floración tardía en algunas regiones, lo que resalta la importancia de la implantación de variedades de floración muy tardía.

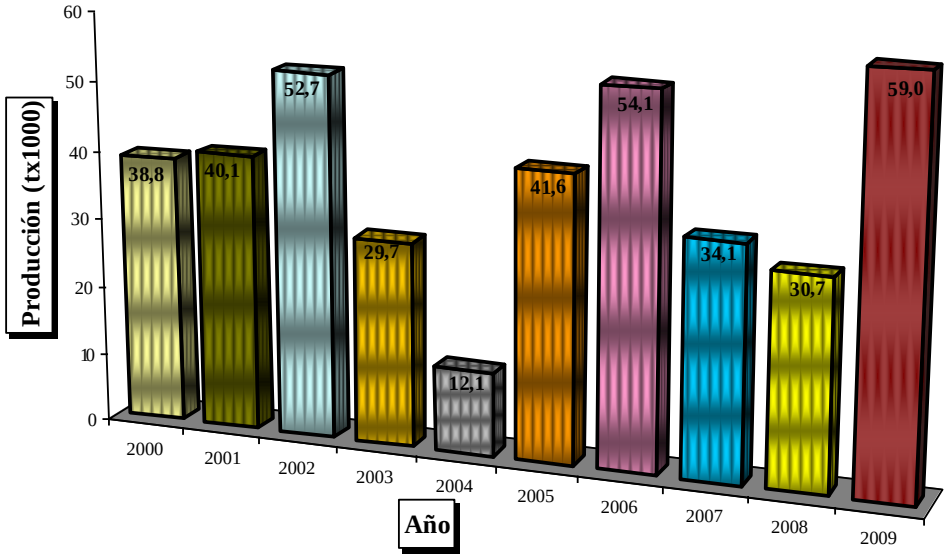


Figura 1.3: Histórico de producciones de almendra en pepita (tx1000) en España desde 2000.

Sin embargo, no sólo las heladas son un problema para la producción del almendro, sino también la sequía y las condiciones atmosféricas desfavorables para la polinización en el momento de la floración. La mayor parte de las variedades del almendro muestran autoincompatibilidad floral, por lo que requieren la polinización cruzada con otras variedades con las que sean inter-compatibles y de floración simultánea para obtener una producción comercial. Por ello estos datos sobre la variabilidad de la producción manifiestan que el cultivo del almendro en España sigue necesitando variedades que estén menos influídas por los factores climáticos limitantes de la producción. La optimización de las condiciones de cultivo, la utilización de variedades productivas de floración tardía y autocompatibles (capaces de autofecundarse) y la instalación de riego localizado, deberán tener una mayor repercusión a la hora de alcanzar rendimientos elevados en el cultivo del almendro.

1.3. Problemática del cultivo del almendro

Al igual que ocurre en otros cultivos, son varios los factores que influyen en la producción del almendro. Entre ellos destacan el material vegetal (variedades y patrones), las condiciones ambientales (climáticas y de suelo) y las técnicas de cultivo (García, 1986).

Tradicionalmente, el almendro se propagó por semilla, lo cual unido a su alogamia (necesidad de polinización cruzada), dio lugar a las notables diferencias y elevado polimorfismo que le caracteriza.

El injerto en almendro comenzó a generalizarse en España a finales del siglo XIX, dando unos resultados notables en cuanto a la productividad de la planta, calidad de los frutos así como en la resistencia o sensibilidad a agentes externos (físicos o fitopatológicos). Por ello, la elección del patrón es un aspecto de gran importancia, porque esta parte del árbol es la que debe proporcionar la conveniente adaptación a las condiciones de suelo y la que proporciona la absorción de agua y la nutrición de todo el conjunto de la planta.

Del mismo modo, la variedad va a determinar en gran medida la rentabilidad de la plantación en términos de producción, calidad del fruto y costes de cultivo. Por ello, hoy en día se procura introducir variedades que sean de floración tardía (para reducir el riesgo a heladas), autocompatibles (que permitan la obtención de una cosecha con la polinización por la propia variedad), y que a su vez tengan una alta calidad de la pepita. Por lo tanto, la elección de un patrón y de la variedad es una de las decisiones más importantes que toma un agricultor a la hora de realizar una plantación. Un error en la toma de estas decisiones puede tener consecuencias irreversibles, y cualquier intervención futura va a representar un coste adicional que va a gravar el cultivo, reduciendo por lo tanto su rentabilidad e incluso anulándola en algunos casos.

En cuanto a las técnicas de cultivo, el riego es un factor de producción que debe aplicarse siempre que sea posible. Aunque el almendro ha sido siempre considerado como una especie rústica, el déficit hídrico reduce enormemente la calidad y la producción de la cosecha. La introducción de sistemas de riego en el cultivo ha

demostrado los incrementos en la rentabilidad del mismo, sin embargo ello supone una inversión importante que muchos agricultores no pueden afrontar. También el suelo va a condicionar el crecimiento de los árboles, y en consecuencia su capacidad productiva. Así pues, aunque el almendro sea una especie relativamente poco exigente que puede crecer en una amplia gama de suelos, no tolera los suelos salinos al igual que los otros frutales.

Las condiciones medioambientales durante la floración también son determinantes en la producción. El transporte efectivo del polen se consigue mediante los insectos polinizadores, cuya mayor eficacia presentan las abejas, como en la mayoría de los frutales. Éstas desarrollan su mayor actividad cuando la temperatura ambiente se encuentra entre 15 y 26ºC, pero decrece significativamente al descender las temperaturas por debajo de los 10-12ºC y con vientos superiores a 24km/h (Tasei, 1975). Por otro lado, la incidencia de las heladas es importante en todas las zonas de cultivo del interior e intermitente también en algunas zonas de la costa, por lo que la floración tardía es un carácter que se ha revelado de gran interés, dado el peligro de las heladas para una especie de floración tan temprana como el almendro. Por ello, para evitar la pérdida de la cosecha por las heladas y una polinización deficiente, se recomienda el uso de variedades autocompatibles y de floración tardía.

1.4 Programas de mejora del almendro

Los primeros países en emprender programas de mejora del almendro fueron los Estados Unidos en 1923 (Kester y Gradziel, 1998) y la antigua Unión Soviética (URSS) en 1930 (Rikhter, 1969). Posteriormente se iniciaron programas de mejora en diferentes países mediterráneos como Francia, Italia o España, en los que se han obtenido resultados evidentes. Otros países como Grecia, Israel o Túnez han desarrollado en algún momento programas de mejora, pero sin excesiva continuidad ni obtención de resultados. Cabe igualmente destacar que Australia se ha unido recientemente a la lista de países que desarrollan programas de mejora en el almendro.

La mejora genética en España se centra en tres programas de diferentes comunidades autónomas (Cataluña, Murcia y Aragón), cuyos objetivos primordiales son comunes en todos ellos: floración tardía y autocompatibilidad.

Los trabajos sobre el almendro del CEBAS-CSIC se iniciaron en Murcia en 1971, con un trabajo de prospección de los materiales autóctonos cultivados en dicha región (Egea y García, 1975). En 1985 se empezaron a realizar cruzamientos en un programa de mejora dirigido a la obtención de nuevas variedades. De esos cruzamientos se obtuvieron en primer lugar las variedades 'Antoñeta' y 'Marta' (Egea et al., 2000). Recientemente se han obtenido dos nuevas variedades autocompatibles y de floración muy tardía: 'Penta' y 'Tardona' (Dicenta et al., 2009).

El programa de mejora del IRTA en Más Bové (Tarragona), comenzó en 1975, y el primer fruto de dicho programa fueron las variedades 'Francolí', 'Masbovera' y 'Glorieta' (Vargas y Romero, 1994). Más recientemente se ha incorporado la autocompatibilidad como objetivo de mejora y se ha registrado cuatro nuevas variedades 'Constantí', 'Marinada', 'Tarraco' y 'Vairo'. 'Tarraco' que no es autocompatible ha sido seleccionada por otras cualidades (Vargas et al., 2008).

La investigación sobre el almendro se inicio en Zaragoza en 1966 en el entonces CIDADE, actual CITA de Aragón, sentando las bases para el posterior programa de mejora (Felipe y Socias i Company, 1985). El objetivo principal de dicho programa ha sido durante todo este tiempo la selección de variedades autocompatibles, de floración tardía y de buenas características del fruto (Felipe y Socias i Company, 1985), así como la creación de nuevos patrones del género *Prunus* para el almendro (Felipe, 2002). Los resultados de estos programas se han completado con la asistencia al sector en colaboración con el Centro de Técnicas Agrarias. Las primeras variedades obtenidas fueron 'Aylés', 'Guara' y 'Moncayo' (Felipe y Socias i Company, 1987) y posteriormente 'Felisia', 'Cambra' y 'Blanquerna' (Socias i Company y Felipe, 1999), las cuales son autocompatibles y sin semillas dobles. En 2006 se obtuvieron las variedades 'Soleta' y 'Belona', caracterizadas por su elevada calidad de la pepita (Socias i Company y Felipe, 2007) y posteriormente la variedad 'Mardía', de floración extra-tardía (Socias i Company et al., 2008). Actualmente, la actividad investigadora se

sigue centrando en la obtención de nuevas variedades de floración extra-tardía, autocompatibles y con un elevado valor nutritivo de la pepita.

1.5 La autoincompatibilidad floral

La autoincompatibilidad (AI) es la incapacidad de una planta hermafrodita para producir semillas mediante su autopolinización aunque presente gametos viables. Es una estrategia reproductiva para promover la fecundación entre individuos que no estén relacionados y, por ende, es un mecanismo creador de nueva variabilidad genética y de prevención de la consanguinidad (Whitehouse, 1950).

Las primeras observaciones de los problemas de la autoincompatibilidad floral se llevaron a cabo por Darwin (1876), quien indicó que la mayoría de las plantas superiores habían desarrollado un mecanismo para prevenir la tendencia a la autofertilización. Más de 100 familias botánicas, entre las cuales se incluyen las Solanáceas, Brasicáceas, Rosáceas, Asteráceas o Fabáceas, presentan especies autoincompatibles (East, 1934). De hecho, se ha estimado que el 39% de las especies de angiospermas son autoincompatibles. Tan amplia distribución taxonómica es congruente con la existencia de varios mecanismos genéticos diferentes que regulan la AI, dependiendo de la familia considerada (de Nettancourt, 2001).

No obstante, en todos los casos se observa un mismo fenómeno: los granos de polen que llegan al estigma de la misma planta son incapaces de efectuar la fecundación de la flor ya que pueden tener lugar distintos impedimentos para ello, ya sea la falta de germinación de los granos de polen en el estigma, la detención del crecimiento de los tubos polínicos en el estilo o durante su penetración en la oosfera (de Nettancourt, 2001).

1.5.1 Tipos de mecanismos

Cuando las flores de la especie que presenta autoincompatibilidad son morfológicamente idénticas entre sí, se dice que la AI es *monomórfica*. En el caso de

que la especie presente dos o tres tipos morfológicos diferentes de flores se dice que la AI es *heteromórfica*.

1.5.1.1 Autoincompatibilidad heteromórfica

La AI *heteromórfica* ha sido mucho menos estudiada que la AI *homomórfica*. Este tipo de mecanismo se ha encontrado en 25 familias (Barrett, 1992), aunque se desconoce si la reacción de la autoincompatibilidad se produce en el estigma, el estilo o incluso en el ovario, ya que puede depender de la especie y del tipo de morfología floral, llamada *morfo*. Hasta la fecha se desconocen los genes que podrían regular este tipo de mecanismo, aunque diferentes proyectos para identificar genes específicos de cada *morfo* se han puesto en marcha en *Primula* (McCubbin et al., 2006) y *Turnera* (Shore et al., 2006).

1.5.1.2 Autoincompatibilidad monomórfica

La AI monomórfica es el mecanismo de AI mejor estudiado. Su mecanismo de actuación es el de la inhibición de la germinación o del crecimiento de los tubos polínicos en el estigma o en el estilo. El mecanismo de AI monomórfico puede ser gametofítico o esporofitico, dependiendo del comportamiento genético del polen (Kao y Tsukamoto, 2004). Hasta la fecha, cuatro de las familias que exhiben un sistema gametofítico de AI, Solanáceas (tabaco, patata, etc.), Rosáceas (manzano, peral almendro, etc.), Escrofulariáceas (*Antirrhinum*) y Papaveráceas (amapola), así como en las Brasicáceas, cuyo sistema de autoincompatibilidad es esporofítico, están siendo estudiadas extensivamente a nivel molecular. Sin duda alguna, en las especies de la familia de las Solanáceas (*Lycopersicon, Nicotiana, Petunia* y *Solanum*) se ha avanzado más que en las otras en el estudio molecular a fin de determinar el mecanismo de incompatibilidad. No obstante, en tres géneros de las Rosáceas (*Malus, Prunus* y *Pyrus*), la investigación de la incompatibilidad está ofreciendo muchas y nuevas aportaciones para poder explicar este mecanismo.

1.5.1.2.1. Autoincompatibilidad esporofítica (AIE)

En el sistema esporofítico (AIE) el control genético de la AI puede ser dialélico o multialélico en un locus (*S*) o en varios loci (*S* y *Z*). En este sistema, a diferencia de la AI gametofítica, la reacción de incompatibilidad está determinada por los dos alelos *S* presentes en el esporofito. La AIE se ha identificado en las familias de las Brasicáceas, Asteráceas, Betuláceas y Esterculiáceas, aunque solamente se ha descrito con detalle a nivel molecular en *Brassica* (Brasicáceas). En dicho mecanismo, el fenotipo del polen viene determinado por el genotipo diploide de la antera. En *Brassica*, la cubierta del polen lleva los productos génicos de los dos alelos *S*. Estos productos son proteínas pequeñas, ricas en cisteína. El gen que codifica estas proteínas se denomina SCR (*S* cysteine-rich protein) y se expresa en el tapetum de la antera (esporofito) así como también en la microspora y en el polen, o sea, el gametofito (Schopfer et al., 1999; Takayama et al., 2000). El determinante femenino de la respuesta AI en *Brassica* es una proteína transmembranal conocida como SRK (*S*-receptor kinase), la cual presenta un dominio quinasa intracelular y un dominio extracelular variable. SRK se expresa en el estigma y funciona como un receptor de de la proteína SCR del polen. Por otro lado, otra proteína estigmática denominada SLG (*S* locus glycoprotein) presenta una secuencia similar a SRK y parece funcionar como un 'co-receptor' del determinante masculino, amplificando la respuesta de autoincompatibilidad.

1.5.1.2.2. Autoincompatibilidad gametofitica (AIG)

El control genético responsable de este sistema de AI viene determinado por múltiples alelos en un solo locus llamado locus *S*, el cual contiene los genes que controlan las especifidades del polen y del pistilo. De este modo, se emplea el término "haplotipo" para describir variantes del locus *S*, mientras que el término "alelo" se usa para describir variantes de un gen en el locus *S*. La reacción de AI se rige por la coincidencia o diferencia de los alelos *S* en el grano de polen y en el pistilo, que como tejido materno y diploide tiene los dos alelos de la planta, mientras que el grano de polen, como célula haploide, contiene solamente un alelo. Desde el punto de vista molecular, cada alelo del locus *S* está compuesto por al menos dos genes, el gen del pistilo y el del polen.

El gen del pistilo

El locus *S* produce en los pistilos una glicoproteína básica con actividad ribonucleasa (McClure et al., 1989), la cual es responsable del rechazo del polen en el estilo, ya que las *S*-RNasas actúan como citotoxinas, por lo que degradan el ARN ribosómico del tubo polínico e inhiben su crecimiento (McClure et al., 1990; Royo et al., 1994). Esta proteína *S* fue secuenciada en primer lugar en Solanáceas por Ioerger et al. (1991), revelando la presencia de 5 regiones conservadas (C1-C5) y dos regiones hipervariables (HVa y HVb). Más tarde también se secuenció en el género *Prunus*, diferenciándose de las Solanáceas solamente en la región RC4, la cual es específica de las Rosáceas (Ushijima et al. 1998), así como en la ausencia de una de las dos regiones hipervariables (Fig. 1.4).

La presencia de distintas regiones conservadas hizo posible clonar y obtener fragmentos de cDNA de las proteínas *S* a través de PCR. Este descubrimiento fue realmente de gran utilidad, ya que ha permitido obtener los diferentes genotipos de distintas variedades (Tao et al., 1999a).

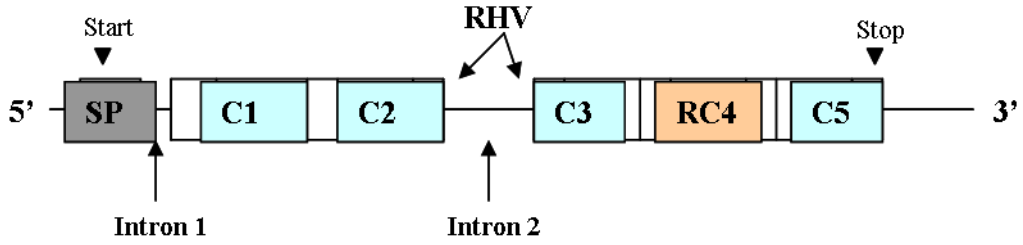


Figura 1.4: Representación esquemática de la estructura de la *S*-RNasa de *Prunus*.

Merece especial atención la aplicación estas técnicas moleculares en distintas especies de las Rosáceas, ya que gracias a estos avances y a partir de una pequeña cantidad de ADN obtenida desde el primer año de vida del árbol se ha podido conocer su genotipo y diseñar los cruzamientos para obtener descendencias con los caracteres de interés por el mejorador.

El gen del polen

En los últimos años, la mayoría de las investigaciones se han orientado a la identificación de los genes expresados en el polen. Estos genes fueron finalmente identificados gracias a la secuenciación a gran escala del DNA genómico en las zonas próximas a las codificadas por las *S*-RNasas.

El primer gen candidato prometedor se obtuvo a partir de un análisis de secuencia de ADN de una región de 64kb que contenía la S_2-RNasa de *Antirrhinum hispanicum* (Lai et al., 2002). Esta región contenía un gen F-box que se expresaba específicamente en el polen, el cual denominaron como SLF (*S*-Locus F-box).

Posteriormente, experimentos similares se llevaron a cabo en distintas especies de Rosáceas (Entani et al., 2003; Ushijima et al., 2003) y Solanáceas (Sijacic et al., 2004). El grupo del Dr. Entani analizó una región genómica de 60 kb alrededor del locus *S* en la variedad 'Nanko' de albaricoquero japonés (*Prunus mume* (Sieb.) Sieb. et Zucc.) y encontraron que la *S*-RNasa estaba presente en una región que contenía igualmente un gen F-box.

Por la misma fecha se secuenció completamente un segmento de unos 70 kb del locus *S* de almendro. En este estudio se construyó una librería genómica del tipo fósmido en las variedades 'Nonpareil', 'Sauret' y 'Monterrey'. Esta región contenía dos genes F-box que se expresaban en el polen, uno de los cuales mostraba un alto nivel de polimorfismos específicos en cada haplotipo *S* (Ushijima et al., 2003), de ahí su nombre de SFB (*S*-haplotype-specific F-box). Respecto al otro gen, se descubrió posteriormente que presentaba un bajo nivel de polimorfismo y que igualmente se expresaba en el pistilo, por lo que se descartó como posible gen candidato del polen. Por lo tanto, en el género de *Prunus*, el gen SFB fue asignado como el gen candidato del polen.

Estos genes codifican para proteínas con el motivo 'caja-F' (F-box), por lo que se sugiere que están implicados en el proceso de ubiquitinación (ubiquitin-mediated protein degradation), el cual ocurre en el proteasoma 26, que es el complejo macromolecular encargado de la degradación de proteínas. Por lo tanto, la proteina SLF podría ser una ligasa tipo E3, la cual formaría parte del complejo multiproteico SCF

(contiene las proteínas Skp1, Cullin y F-box) que es un intermediario de la ubiquitinación de proteínas. Los experimentos realizados por Qiao et al. (2004) en *Antirrhinum* demostraron que la ubiquitinación de las *S*-RNasas a través de la degradación de proteínas ubiquitina/26S proteasoma es la responsable del crecimiento de los tubos polínicos en las plantas compatibles.

Para confirmar estos resultados, los investigadores realizaron pruebas de germinación de polen in vitro, usando un medio de germinación del polen que a su vez contenía extractos de pistilos de plantas compatibles e incompatibles. Los resultados obtenidos sugirieron que en las plantas compatibles, el crecimiento del tubo polínico requiere la ubiquitinación de la actividad *S*-RNasa a través de la ubiquitina/26S proteasoma. Sin embargo, dichos experimentos todavía no se han realizado en el género *Prunus* para determinar si dicho complejo actúa de igual manera en los árboles frutales.

La estructura del gen SFB (Fig. 1.5) presenta un motivo F-box, dos regiones variables (V1 y V2) y dos regiones hipervariables (HVa y HVb), las cuales son posibles sitios de interacción con las *S*-RNasas (Ikeda et al., 2004).

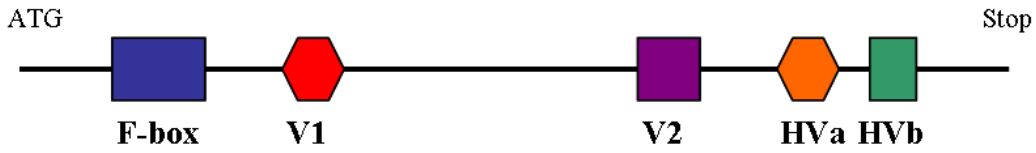


Figura1. 5: Representación esquemática de la estructura del SFB de almendro.

Por otro lado se ha seleccionado el gen del *S* locus F-box (SLF) como el mejor gen candidato en la familia de las Solanáceas y Plantagináceas gracias a unos experimentos de transformación genética en *P. inflata* (Sijacic et al., 2004) y *P. hybrida* (Qiao et al., 2004).

Recientemente, múltiples genes F-box, llamados *S* locus F-box brothers (SFBBs), se han identificado como los posibles candidatos del alelo *S* del polen en *Malus* y *Pyrus* (Cheng et al. 2006; Sassa et al. 2007). Sin embargo, todavía no se

conoce con exactitud cuál de ellos podría ser el gen determinante de la actividad del polen en estas dos especies.

Los análisis filogenéticos sugieren que todos los SFB/SLF de *Antirrhinum*, *Petunia* y *Prunus* pertenecen a un mismo grupo filogenético ya que comparten los cuatro dominios conservados (Wang et al., 2004), lo cual implica que el antepasado común de muchas dicotiledóneas poseen un sistema de GSI basado en el complejo *S*-RNasa/F-box. Sin embargo, todos los SFB de *Prunus* podrían formar una subdivisión distinta a la del SFL de *Antirrhinum* y *Petunia*, sugiriendo que probablemente hayan divergido más pronto durante la evolución del mecanismo de AI (Matsumoto et al., 2008; Ushijima et al., 2004).

Otros genes que modularían la AI

Aunque el locus *S* codifica los determinantes para el haplotypo *S*, hay evidencia de la existencia de otros genes no ligados al locus *S*, llamados genes modificadores, que son necesarios en el mecanismo de AI. En este campo se están produciendo actualmente avances significativos.

Factores modificadores del pistilo

McClure et al. (1999) identificaron un gen llamado HT-B, el cual es una proteína de pequeño peso molecular con una secuencia rica en asparagina y que se expresaba en el pistilo de las plantas autoincompatibles de *N. alata*, pero sorprendentemente no se expresaba en las autocompatibles de *N. plumbaginifolia*. También se encontraron otros genes homólogos de este gen en otros géneros de las Solanáceas, como *Lycopersicon* y *Solanum* (Kondo et al., 2002). En estos estudios se demostró que tampoco había expresión del gen HT-B en las plantas AC. Estos resultados sugieren que la proteína HT-B está implicada en la respuesta de AI, aunque se desconoce cuál es su verdadera función.

El segundo factor externo que podría estar implicado en el mecanismo de AI se conoce con el nombre de proteína 120K, que es una glicoproteina de 120 kDa de peso molecular y muy abundante en la matriz extracelular del estilo (Lind et al., 1996). Dicha

proteína se une a la *S*-RNasa, inhibiendo la expresión en las plantas AC (Cruz-García et al., 2005). Sin embargo, al igual que ocurre con el gen HT-B, se desconoce por el momento su función exacta.

Factores modificadores del polen

Las proteínas F-box sirven como adaptadores que unen las proteínas específicas de substrato al complejo SCF. Dicho complejo contiene las proteínas Skp1 (S-phase kinase-associated protein 1), Cullin1 y F-box, el cual funciona como una E3 Ubiquitina ligasa en el sistema del proteoseoma 26S-Ubiquitinación (UPS). Por ello, se sugirió que cuando el SLF está implicado en el mecanismo de AI, también participaría en el complejo SCF mediante la ubiquitinación de la *S*-RNasa. El reciente hallazgo de la proteína AhSSK1 (*Antirrhinum hispanicum* SLF-interacting Skp1-like1) pone de manifiesto que este gen podría actuar igualmente como un adaptador que conectase el SLF a la proteína Cullin1, las cuales podrían tener una gran responsabilidad en la ubiquitinación de la *S*-RNasa (Huang et al., 2006). Más recientemente se ha determinado otra nueva proteína equivalente a AhSSK1, llamada PhSSK1 (*Petunia hybrida* SLF-interacting Skp1-like1) la cual se expresa específicamente en el polen y actúa también como un adaptador en el complejo SCF (Zhao et al., 2010). En este estudio se sugiere que el determinante *S* del polen puede contribuir más bien a inhibir antes que a proteger la *S*-RNasa en la familia de las solanáceas. No obstante, al igual que ocurre con los factores externos del pistilo, se requieren estudios más precisos con el fin de deducir conclusiones más relevantes para entender el mecanismo de AI en las plantas.

Modelos bioquímicos que explicarían la autoincompatibilidad gametofítica

Se ha propuesto varios modelos en los últimos años para explicar el probable mecanismo de autoincompatibilidad en el mundo vegetal.

En primer lugar se propusieron los modelos del receptor y del inhibidor (Kao y McCubbin 1996). El modelo receptor propone que el producto *S* del polen permite la entrada de las *S*-RNasas con un haplotipo idéntico al suyo. Sin embargo en el modelo del inhibidor, los productos del gen *S* del polen son inhibidores de la actividad RNasa

de las proteínas estilares de genotipo diferente al propio y, por consiguiente, únicamente las proteínas *S* del mismo genotipo funcionarían en el interior del tubo polínico, inhibiendo su crecimiento.

Posteriormente, Qiao et al. (2004) propusieron el modelo de degradación de proteínas, por el proteosoma *26* (Fig. 1.6). El proteosoma, también llamado proteasoma, es un complejo proteico presente en todas las células eucariotas que se encarga de realizar la degradación de proteínas no necesarias o dañadas. En las células eucariotas los proteosomas se encuentran en el núcleo y en el citoplasma y representan un importante mecanismo por el cual las células controlan la concentración de proteínas mediante la degradación de las mismas. Las proteínas a degradar se marcan mediante una pequeña proteína llamada ubiquitina. Cuando se unen muchas moléculas de ubiquitina, por medio de la ubiquitina ligasa, dan lugar a una cadena poliubiquitínica que permite al proteosoma identificar y degradar la proteína.

En este modelo propuesto por el grupo del Dr. Qiao (2004), la compatibilidad es el resultado de la degradación a través de la ubiquitina de las *S*-RNasas con haplotipos distintos al genotipo del polen. Por lo tanto, cuando una *S*-RNasa es ubiquitinada (degradada), se produce el crecimiento del tubo polínico y por lo tanto la polinización será compatible. Mientras que si una *S*-RNasa no es ubiquitinada, ésta no se degradará, el tubo polínico se detendrá y por lo tanto la planta será AI.

Recientemente, Goldraij et al. (2006) han propuesto el modelo de compartimentación, el cual intenta explicar por qué las *S*-RNasas provocan un efecto de toxicidad en los tubos polínicos. En dicho modelo se plantea que las *S*-RNasas, la proteína HT-B y la glicoproteína 120K formarían una barrera de rechazo que el polen debería superar para llegar al ovario. Este grupo de investigación mostró que la *S*-RNasa es "secuestrada" en el interior de un compartimiento de la vacuola. En el caso de incompatibilidad, la proteína HT-B permanecería estable cuando la *S*-RNasa reconoce al producto S del polen. Por lo tanto la *S*-RNasa emplearía su efecto citotóxico provocando la inhibición del crecimiento del tubo polínico. Mientras que en el caso de una polinización compatible, la proteína HT-B se degradaría, la *S*-RNasa estaría compartimentalizada y el RNA permanecería estable siendo la polinización compatible (Goldraij et al., 2006; McClure and Frankling-Tong 2006).

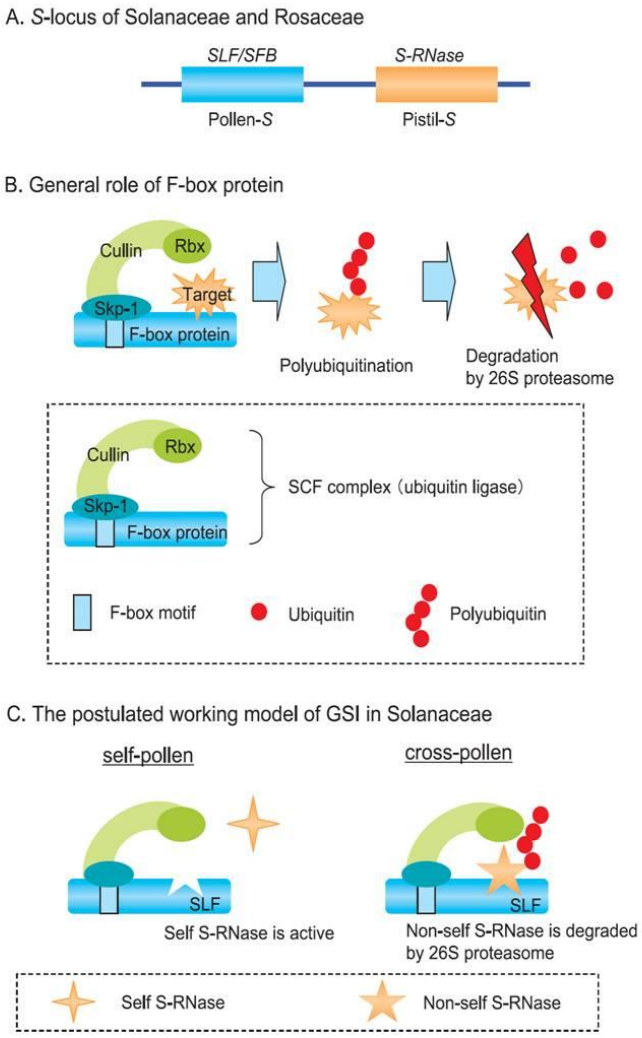


Figura 1.6: Representación gráfica del modelo de autoincompatibilidad gametofítica (AIG). El locus *S* contiene al menos 2 genes, la *S*-RNasa y el SLF/SFB (a). El gen del polen interactúa con otras proteínas, que a su vez se integran en un complejo enzimático, el cual es conocido como el complejo SCF (b). Modelo propuesto de AIG en solanáceas. Representación adaptada de Yamane et al. (2009).

Aunque desde los primeros estudios se ha asumido que el factor determinante del polen en *Prunus* tenía la misma función que el de las Solanáceas (Entani et al., 2003; Ushijima et al., 2003), una serie de recientes estudios sugieren que el gen SFB de esta especie podría actuar de una manera diferente al del resto de familias, ya que presenta solamente el dominio necesario para la interacción del haplotipo especifico *S*, y no el dominio de interacción general como ocurre en Solanáceas y Plantagináceas (Hauck et al., 2006; Sonneveld et al., 2005). Por ello, varios autores han discutido si los modelos previamente descritos en Solanáceas y Plantagináceas se ajustan o no al

reconocimiento molecular entre las *S*-RNasas y el SFB en las Rosáceas (Matsumoto et al., 2008; Sonnelved et al., 2005).

1.6 La biología floral del almendro

1.6.1 La autoincompatibilidad en el almendro

Aunque a finales del siglo XIX y principios del XX se realizaron algunas observaciones sobre la polinización del almendro (revisado en Socias i Company, 1990), los primeros ensayos de polinización se llevaron a cabo en California por Tufts (1919) con 17 variedades, ampliados por Tufts y Philp (1922) con cuatro más. Su conclusión fue que las 21 variedades de almendro californianas ensayadas eran AI, aunque en algunas autopolinizaciones se obtuvo un pequeño cuajado que no se puede considerar económicamente importante. Estas conclusiones se confirmaron en los sucesivos estudios sobre la polinización del almendro en diversos países, como España, Rusia, Alemania, Sudáfrica, Australia, etc. en los años 1930 (Socias i Company, 1990), hasta el descubrimiento de la AC en Portugal por Almeida (1945).

El almendro, al igual que el cerezo, el ciruelo, el manzano y el peral, es una especie diploide con un sistema de AI tipo gametofitico (Socias i Company et al, 1976). La AI en el almendro se manifiesta mediante la detención del crecimiento de los tubos polínicos incompatibles en el estilo (Socias i Company et al., 1976). La AI es una ventaja para muchas especies, pero para el cultivo del almendro es un inconveniente, porque la parte comercial es una semilla que sólo se obtiene si se produce la fertilización del óvulo después de una polinización efectiva.

1.6.2 La autocompatibilidad en el almendro

La primera mención a la AC en el almendro tuvo lugar por Almeida (1945) después del estudio de 23 variedades en Portugal. Su conclusión fue que estas variedades mostraban distintos niveles de cuajado, dividiéndolas en cuatro grupos, desde las absolutamente AI hasta las altamente AC, como ‘José Dias’ y ‘Duro Italiano’. Sin embargo, este estudio pasó desapercibido y no fue hasta los años 1970 cuando se dio

importancia a la AC, especialmente después del descubrimiento de un grupo de variedades AC en la población italiana de la Apulia, como 'Filippo Ceo', 'Tuono' o 'Genco' (Grasselly y Olivier, 1976).

A partir de este momento se iniciaron los primeros estudios sobre este carácter tan interesante desde el punto productivo en el almendro, asegurándose su transmisión a la descendencia (Socias i Company y Felipe, 1977). Los primeros resultados indicaron que la expresión de la AC era dominante sobre la AI, tanto entre las descendencias de dos variedades AC como en las descendencias de una variedad AC y otra AI (Socias i Company, 1984). Así se crearon las bases para los diversos programas de mejora genética que ya estaban en curso o que se iniciaron en estos años. Rápidamente se propuso que la AC se debía a la presencia del alelo S_f, posiblemente alélico con los alelos de la AI (Socias i Company, 1984).

1.6.3 La interincompatibilidad en el almendro

La interincompatibilidad es la imposibilidad de producir semillas después del cruzamiento entre dos plantas fértiles. No es frecuente en almendro, ya que sólo se ha hallado entre cultivares relacionados genéticamente, pertenecientes a la misma población o programa de cruzamientos (Socias i Company, 1990).

El primer caso de interincompatibilidad se descubrió ya en el primer estudio sobre la polinización del almendro, ya que Tufts (1919) señaló la interincompatibilidad de 'Nonpareil' con 'I.X.L.' y de 'Languedoc' con 'Texas'. Así mismo, los primeros intentos de establecer los grupos de interincompatibilidad (CIG) se llevaron a cabo en las variedades californianas mediante la realización de polinizaciones controladas en campo (revisado en Kester y Asay, 1975). Sin embargo, debido a los posibles datos erráticos causados por las condiciones climáticas, se empezó a determinar la interincompatibilidad mediante microscopía de fluorescencia después de su aplicación en el almendro por Socias i Company et al. (1976). Posteriormente, Kester et al. (1994) determinó seis CIG mediante polinizaciones controladas en campo y laboratorio usando 27 variedades californianas.

Mas recientemente, la identificación de las ribonucleasas estilares ha permitido determinar numerosos alelos *S* de variedades europeas y americanas, así como ampliar los CIG (Bošković et al., 1997, 1999 y 2003). Posteriormente, los marcadores de ADN también han servido para la identificación de alelos *S* de más de 180 variedades americanas y europeas (Ma y Oliveira, 2001; López et al., Ortega et al., 2006; Kodad et al., 2008a; Halász et al., 2009) ampliando actualmente los grupos de interincompatibilidad hasta 19 (Ortega et al., 2006; Kodad et al., 2008b).

1.6.4 Estrategias para determinar la autocompatibilidad en el almendro

Las primeras determinaciones de la AI en el almendro, así como de la interincompatibilidad entre variedades, se realizaron mediante polinizaciones controladas en campo (Tufts, 1919; Tufts y Philp, 1922). Así mismo el cuajado fue el criterio utilizado para las primeras determinaciones de la AC (Almeida, 1945).

El crecimiento de los tubos polínicos es una indicación clara de la compatibilidad de una polinización y por ello se ha utilizado repetidamente en la determinación de la AC desde las primeras observaciones en el almendro por Socias i Company et al. (1976). Las flores utilizadas para el examen de los tubos polínicos se pueden mantener en diferentes ambientes, ya sobre las ramas originales o en el laboratorio, pues se obtienen los mismos resultados inequívocos (Socias i Company, 2001).

Los estudios del crecimiento de los tubos polínicos se han asociado con los resultados de cuajados después de polinizaciones artificiales (Ben Njima y Socias i Company, 1995; Kodad y Socias i Company, 2006; Socias i Company y Felipe, 1987), pues se han obtenido resultados concordantes, aunque la depresión por consanguinidad puede afectar la expresión de la autocompatibilidad tanto en el cuajado como en el crecimiento de los tubos polínicos (Alonso and Socias i Company, 2005).

Posteriormente, la identificación de las ribonucleasas estilares a través de la técnica NepHGE, ha permitido determinar los genotipos *S* en las distintas variedades de almendro, especialmente después de observar que las variedades AC no tenían actividad

RNasa y presentaban solamente una banda en el zimograma en lugar de dos como en las variedades AI (Bošković et al., 1999).

Por último, el desarrollo de marcadores de ADN también ha servido para la identificación del alelo S_f, tanto por cebadores conservados como específicos (Tao et al., 1999a; Tamura et al., 2000; Channuntapipat et al., 2001, 2003; Sonneveld et al., 2003; Sutherland et al., 2004; Vaughan et al., 2006).

1.6.5 Caracterización molecular de genotipos S de almendro

Los nuevos marcadores de ADN son, a fecha de hoy, la herramienta elegida por los mejoradores para obtener la información fiable sobre la variación genética y así poder identificar los alelos del locus *S*. Los marcadores moleculares presentan como características principales la rapidez de la obtención de los resultados, el coste relativamente barato y sobre todo la posibilidad de aplicarlos rutinariamente y de fácil desarrollo. Por ello, varios autores han diseñado un gran número de cebadores que son capaces de amplificar las regiones conservadas de las *S*-RNasas así como el gen *S* del polen de almendro. Estos marcadores son aplicables tanto al gen *S* del pistilo como al del polen y han sido diseñados por varios grupos de investigación (Tao et al., 1999a; Tamura et al., 2000; Channuntapipat et al., 2001, 2003; Sonneveld et al., 2003; Sutherland et al., 2004; Vaughan et al., 2006).

Los primeros cebadores diseñados en *Prunus* fueron los obtenidos por Tao et al. (1999a). Estos marcadores, aunque fueron específicamente diseñados para cubrir las regiones conservadas C2-C4 de cerezo, también sirvieron para amplificar distintas especies del género *Prunus*, entre ellas la del almendro. Paralelamente, Tamura et al. (2000) diseñaron unos cebadores a partir de secuencias de ADN de almendro, por lo que fueron muy eficaces a la hora de caracterizar los genes del locus *S*. Estos cebadores llamados AS1II/AmyC5R fueron diseñados para amplificar desde la región C1 de las RNasas hasta la C5.

Posteriormente, Channuntapipat et al. (2001) diseñaron otros nuevos marcadores con el mismo objetivo, amplificar desde la región C1 hasta la C5 (ConF/ConR). Por último, Sonneveld et al. (2003) y Sutherland et al. (2004) diseñaron otro set de

cebadores que amplificaban igualmente desde la región C1 (PaCons1-F) hasta la región C5 (EMPC5consRD). Estos marcadores son capaces de amplificar un gran rango de alelos y funcionan también en distintas especies del género *Prunus*.

A parte de los cebadores diseñados en las regiones conservadas, Channuntapipat et al. (2003) han diseñado marcadores específicos a partir de las secuencias de los intrones para muchos de los alelos del locus *S*. Además de la gama de cebadores disponibles para las *S*-RNasas, hay que añadir los nuevos cebadores diseñados en el gen del polen SFB por Vaughan et al. (2006), los cuales sirven igualmente para identificar y caracterizar los genotipos del locus *S*.

1.6.6 Hipótesis sobre el origen de la autocompatibilidad

Aunque a fecha de hoy, se desconoce el verdadero origen del alelo S_f de autocompatibilidad en el almendro, se han expuesto diversas hipótesis que tratan de explicarlo. La primera, propuesta por Grasselly y Olivier (1976), sugiere que el alelo S_f pudo aparecer como consecuencia de una mutación natural en el locus *S*, y que los agricultores la podrían haber mantenido gracias a la selección del material para su cultivo, principalmente en las poblaciones naturales de la región italiana de la Apulia.

Posteriormente varios autores propusieron la hipótesis de que el alelo S_f aparecería en el almendro a partir de una formación espontánea de híbridos interespecíficos de la especie cultivada con la especie silvestre *P. webbii* (Spach) Vierh. (Godini, 1979; Reina et al., 1985; Socias i Company, 1984 y 1990).

Por último, Bošković et al. (2007) propusieron que la presencia del alelo autocompatible S_f se produciría como consecuencia de la mutación de una arginina por una histidina en la región C2 de la *S*-RNasa. Sin embargo, este grupo de investigación tuvo que publicar una rectificación reconociendo un error en la secuenciación de este alelo, por lo que esta hipótesis queda completamente invalidada.

1.7 Antecedentes previos relacionados con este estudio

En el 2004 se realizaron las primeras observaciones microscópicas de los individuos procedentes de un cruzamiento realizado en el CITA de Aragón entre 'Vivot' y 'Blanquerna'. Este cruzamiento se realizó porque el parental femenino 'Vivot' presenta muy buenas condiciones por su facilidad de descascarado y por su excelente comportamiento en pastelería y repostería. Por otro lado, el parental masculino 'Blanquerna', procedente del programa de mejora del CITA, es autocompatible y de excelente calidad de fruto, tanto por sus características físicas como organolépticas. Ambos parentales son de floración media, por lo que el objetivo del cruzamiento era la obtención de variedades de floración media, autocompatibles, de excelente calidad y adaptadas a las zonas de cultivo litorales o medias.

Como 'Vivot' era consideraba una variedad autoincompatible, como su comportamiento siempre ha indicado, los resultados que se esperaban eran que al menos el 50% de la población fuese autocompatible. Sin embargo, las observaciones preliminares indicaron que muchos de los plantones obtenidos eran autoincompatibles. Por ello, se decidió realizar durante varios años observaciones microscópicas en todos los individuos para poder confirmar los resultados previos obtenidos, y también estudiar intensamente esta población usando el mayor número posible de técnicas moleculares a fin de poder encontrar una explicación a este resultado inesperado.

1.8 Objetivos de la tesis

El objetivo principal de esta memoria realizada en la Unidad de Fruticultura del CITA de Aragón, es el estudio de la autocompatibilidad en el almendro, usando la población 'Vivot' × 'Blanquerna'. Para esta investigación se han empleado tanto técnicas de determinación agronómica para la AC como técnicas moleculares.

Así pues, el objetivo general se abordará a partir de los siguientes sub-objetivos:

Objetivo 1. Determinación de las proporciones fenotípicas de la autocompatibilidad en toda la descendencia. Para ello se realizaron observaciones de

llegada o no de tubos polínicos a la base del estilo mediante un microscopio de luz ultravioleta. Igualmente se determinaron los niveles de cuajado en los individuos seleccionados.

Objetivo 2. Identificación en toda la población de los genotipos del locus *S* mediante el uso de las *S*-RNasas estilares (NEpHGE) y marcadores moleculares específicos y conservados mediante PCR.

Objetivo 3. Correlación de las proporciones fenotípicas y genotípicas de la descendencia e interpretación y búsqueda de factores modificadores implicados en el mecanismo de autoincompatibilidad.

Objetivo 4. Caracterización molecular del locus *S* (*S*-RNasa y SFB) de 'Vivot' y 'Blanquerna' mediante la construcción de una librería genómica (fósmido), así como el análisis de los diferentes niveles de expresión mediante PCR a tiempo real.

Objetivo 5. Construcción de un mapa genético de 'V×B' mediante el uso de microsatélites (SSR) y marcadores específicos del locus *S* para la búsqueda de genes externos al locus *S* responsables de la autocompatibilidad.

Objetivo 6. Identificación de factores epigenéticos posiblemente responsables de la expresión diferencial de la AI/AC en el almendro con el fin de determinar si el principal mecanismo epigenético, metilación, es el responsable del silenciamiento del alelo S_f en las variedades compatibles de almendro.

Objetivo 7. Determinación de las estructuras proteicas en 3D de los diferentes alelos encontrados en la población estudiada mediante herramientas de modelización molecular, a fin de observar sus posibles diferencias estructurales.

2. MATERIAL VEGETAL

2. MATERIAL VEGETAL

El material vegetal empleado en este estudio se encuentra localizado en las parcelas de almendros de la Unidad de Fruticultura del Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA). Las fincas experimentales se encuentran en Zaragoza (Fig. 2.1), a una altitud de 220 m sobre el nivel del mar, con un clima mediterráneo continental semidesértico, que es el propio de la depresión del Ebro. Los veranos son muy calurosos superando frecuentemente los 30ºC, mientras que los inviernos son frescos, siendo normales las heladas y las nieblas.

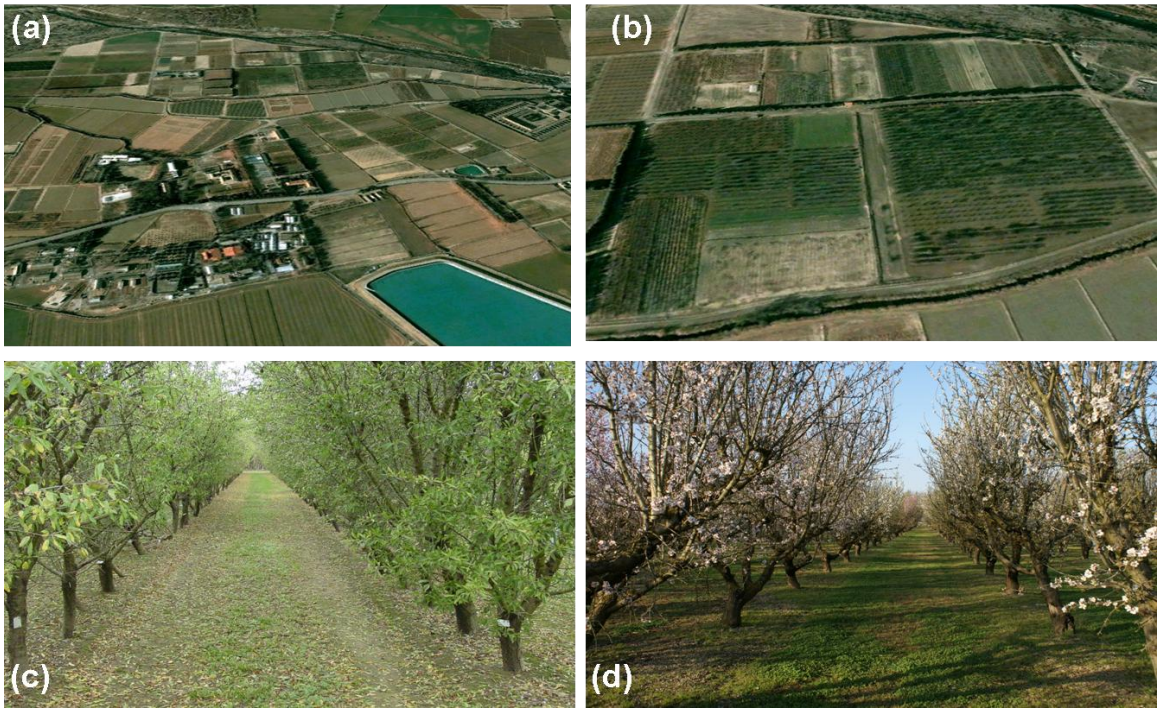


Figura 2.1: Vista aérea de las instalaciones del CITA (a). Fotografía aérea de la parcela experimental de almendro donde se realizaron los ensayos (b). Almendros provenientes del cruzamiento entre 'Vivot' y 'Blanquerna' (c). Banco de germoplasma de almendro en el CITA (d).

En este estudio se ha empleado inicialmente una población de 84 individuos procedentes del cruzamiento 'Vivot' x 'Blanquerna' ('VxB'). Sin embargo, mediante técnicas moleculares se detectó posteriormente que siete de estos 84 individuos no pertenecían realmente a la población 'VxB', por lo que probablemente procedían de una contaminación por polen extraño. Debido a esta contaminación, se decidió la eliminación de estos plantones y estudiar a fondo solamente los 77 individuos restantes.

Este cruzamiento se llevó a cabo en el año 2000, sin embargo no fue hasta el 2004 cuando empezaron a tomarse los primeros datos fenotípicos.

Estas plantas se mantienen según las técnicas normales de cultivo aplicadas a todas las parcelas de almendro del CITA.

El parental femenino **'Vivot'** es una variedad local procedente de la isla de Mallorca, siendo actualmente la variedad más plantada en la isla (Fig. 2.2). Su floración es precoz y su maduración tardía. Presenta un descascarado bueno y es usado principalmente para pastelería y repostería. Al inicio de este trabajo, 'Vivot' estaba descrita como una variedad AI (Rubí, 1980), aunque posteriormente se determinó que su genotipo era S_fS_{23}. Sin embargo, en el desarrollo de esta tesis se ha comprobado que se trata de una variedad AI a pesar de poseer el alelo S_f.

Figura 2.2: Frutos de los dos parentales utilizados en este estudio.

El parental masculino **'Blanquerna'** procede de una semilla obtenida a través de una polinización libre de 'Genco'. 'Blanquerna' es una variedad autocompatible, usada frecuentemente en el programa de Mejora Genética de la Unidad de Fruticultura del CITA (Fig. 2.2). La época de floración es ligeramente posterior a 'Marcona', aunque se puede utilizar eficazmente como polinizador suyo. Su época de maduración, sin embargo, es anterior a la de 'Marcona'. Cabe resaltar que es una variedad muy

productiva y que la pepita es de sabor agradable y de muy buena apariencia. El genotipo *S* determinado en esta variedad fue $S_f S_8$ (Fernández i Martí et al., 2009b).

3. PHENOTYPE CHARACTERIZATION OF SELF-COMPATIBILITY

3. PHENOTYPE CHARACTERIZATION OF SELF-COMPATIBILITY

3.1 Introduction

Almond, with very few exceptions, is a self-incompatible (SI) species (Socias and Company, 1974). Although self-compatibility (SC) was discovered in almond as early as in 1945 by Almeida, attention was not paid to the issue until the 1970s (Grasselly and Olivier, 1976). After SC was again identified in almond when the first studies on the transmission of SC to the offspring were conducted by Socias i Company and Felipe (1977).

Several approaches have been used to asses the level self-compatibility, such as observation of pollen tube growth after artificial pollinations (Socias i Company et al., 1976), by bagging branches at bloom (Graselly and Olivier, 1976) or more recently by molecular techniques. This last information, however, is only genetic, not horticultural, and the final evaluation of SC of a cultivar is its productivity under field conditions.

Set-fruitfulness is the most natural approach to asses the level of self-compatibility. However, climatic conditions during bloom may not be favourable, including frosts causing different levels of damage in buds and flowers (Kodad and Socias i Company, 2005), resulting in final fruit sets not clearly reflecting the level of SC. In addition, emasculation and pollination must be done in the open air, where weather conditions affect undertaking these processes because temperatures are usually very low at almond blooming time and winds might difficult the labour. Thus, fruit set determination in the field should be reserved to the final steps of SC evaluation in elite selections.

Pollen tube growth studies have been often associated with fruit setting following artificial pollination (Ben Njima and Socias i Company, 1995; Socias y Company and Felipe, 1987), giving concordant results. Thus, pollen tube growth has been commonly used in compatibility determinations since is a good indicator of the compatibility of the pollinations. However, the same problems arisen for fruit set evaluation are found when working in the field for pollen tube growth. A possible way

to avoid these weather limitations is through emasculation of the flowers in laboratory and pollination with their own or different pollen. Self-compatibility is then determined by observing pollen tube growth in a fluorescence microscopy.

Thus, the main objective of this chapter was to assess the phenotypic expression of self-compatibility following the progression of pollen tubes through the pistils in the progeny coming from the cross 'Vivot' × 'Blanquerna' and evaluate some selected individuals by fruit set after bagging in the field.

3.2 Methods

3.2.1 Pollen tube growth and fruit set

SI/SC was assayed in the parents as well as in the progeny by the microscopic observation of pollen tube growth over three consecutive years. At least 12 flower buds from each seedling at stage D (Felipe 1977) were collected from field, emasculated in the laboratory and placed in a tray with tap-water (Fig. 3.1), allowing the contact of the flower peduncles with the tray water to prevent dehydration. Pollen viability was also determined by observing the germination of pollen grains in vitro. As viability was good in all cases, the results are not included.

On the other hand, the anthers were removed from the emasculated flowers and allowed to dry for 48h (Fig. 3.1). This pollen was used to self-pollinate the pistils in the tray. Ninety-six hours after pollination the pistils were collected, placed in vials containing 5 ml of a 5% solution of Na_2SO_3, and autoclaved for 12 min at 1.2 kg cm^{-2} in order to soften the tissues. The samples were maintained at 4ºC until observation, when they were stained with 0.1% (v/v) aniline blue in 0.1 N potassium phosphate as a specific stain for callose (Linskens and Esser, 1957).

The pistils were prepared according to the method described by Socias i Company (1979), dissecting the outer part of the pistil and leaving only the transmission tissue trough which pollen tube grow, This growth was observed in the pistils under an

Olympus BH2 microscope with UV epifluorescence using a BP 355/425 exciter filter and a LP 470 barrier filter.

Figure 3.1: Emasculated flowers of 'Vivot' to be pollinated with 'Blanquerna' pollen (a). Bagging of several branches in order to determine the fruit set in some seedlings studied (b). Emasculated pistils placed on trays to be self-pollinated for further microscopic observation (c). Pollen grains placed in different trays before use them for pollination (d). Fruit set of bagged 'Blanquerna' (e).

For fruit determination, a branch containing at least 100 flower buds was bagged before bloom (Socias i Company et al., 2005) in the field in order to asses the level of self-fertilization of the parents and eight target seedlings. Fruits were counted three months after bloom and fruit set was determined as the number of fruit in relation to the number of bagged flowers.

3.2.2 Criterion for determining SC/SI

A genotype was considered SI when no pollen tubes reached the base of the style. Otherwise a genotype was considered SC when pollen tubes reached the style base of at least 8 of 12 styles each year of observation, repeating the assays during three different years. No seedlings with intermediate results were observed.

Following the criterion of Grasselly et al. (1981), fruit set was ranged as: a) fruit sets lower than 2% of the initial bud number: self-incompatible; b) between 2% and 5%: low self-compatible; c) between 5% and 10%: self-compatible; d) higher than 10%: highly self-compatible.

3.3 Results and Discussion

Artificial pollination in the lab was always effective, as the number of pollen grains on the stigmata was over 40 in all cases (Fig. 3.2a), a number considered enough for achieving fertilization in almond (Weinbaum et al., 1986). The observation of pollen tube growth after self-pollination allowed classifying the phenotype of the parents and the seedlings as SC or SI. The microscopic observations of the female parent 'Vivot' allowed to confirm that is a clearly SI cultivar because of showing the arrest of its pollen tubes in the middle third of the style (Fig. 3.2b), whereas the other parent 'Blanquerna' was confirmed as being clearly SC (Fig. 3.2c).

From the progeny of 77 individuals, only 19 (24.7%) showed pollen tubes at the base of the style, whereas in the other 58 (75.3%) pollen tubes did not reach the ovary and stopped in the middle third of the style, showing the characteristic arrest of pollen tube growth as the typical SI response (Table 3.1).

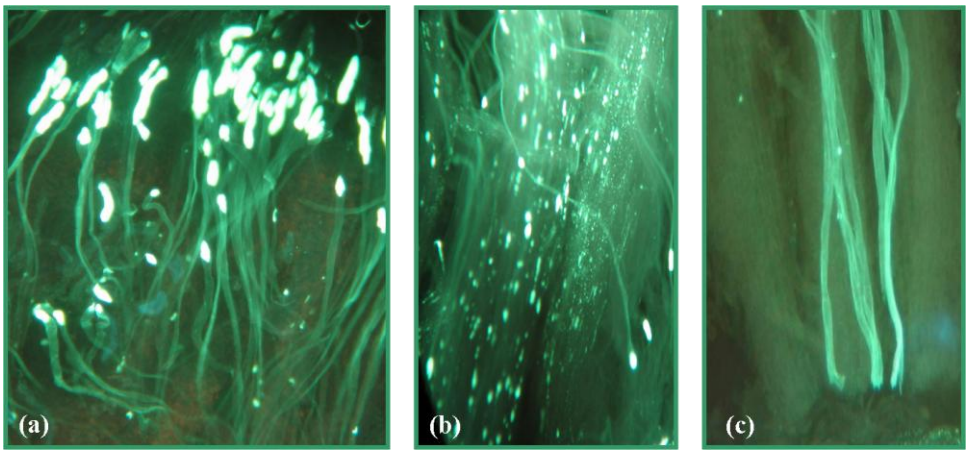


Figure 3.2: Germinated pollen grains on the stigma surface of a self-pollinated pistil (a). Microscopic observation of pollen tube growth after self-pollination of 'Vivot' (b) and 'Blanquerna' (c), showing the arrest of tube growth in SI 'Vivot' and the arrival of pollen tubes at the style base in SC 'Blanquerna'.

Table 3.1: Number of pollen tubes reaching the base of the style and the phenotype for each seedling.

Seedling	Total	Phenotype	Seedling	Total	Phenotype
P1-1	0 of 18	SI	**P1-48**	24 of 29	SC
P1-3	0 of 16	SI	**P1-49**	15 of 30	SC
P1-5	6 of 22	SI	**P1-50**	6 of 22	SI
P1-8	13 of 30	SI	**P1-51**	21 of 24	SC
P1-10	6 of 22	SI	**P1-52**	9 of 25	SI
P1-11	29 of 34	SC	**P1-53**	10 of 28	SI
P1-12	6 of 22	SI	**P1-54**	8 of 25	SI
P1-13	23 of 27	SC	**P1-55**	7 of 19	SI
P1-14	6 of 22	SI	**P1-56**	26 of 30	SC
P1-15	20 of 24	SI	**P1-57**	25 of 30	SC
P1-16	6 of 22	SI	**P1-58**	11 of 30	SI
P1-17	7 of 22	SI	**P1-59**	6 of 22	SI
P1-18	22 of 30	SC	**P1-60**	12 of 30	SI
P1-19	6 of 22	SI	**P1-61**	12 of 30	SI
P1-20	18 of 20	SC	**P1-62**	6 of 22	SI
P1-21	12 of 28	SI	**P1-63**	11 of 26	SI
P1-22	12 of 36	SI	**P1-64**	11 of 28	SI
P1-23	6 of 24	SI	**P1-65**	6 of 22	SI
P1-24	10 of 26	SI	**P1-66**	27 of 30	SC
P1-25	19 of 24	SC	**P1-67**	12 of 26	SI
P1-26	23 of 30	SC	**P1-68**	13 of 30	SI
P1-27	6 of 22	SI	**P1-69**	6 of 22	SI
P1-28	15 of 18	SC	**P1-71**	6 of 22	SI

P1-30	6 of 22	SI	**P1-72**	9 of 24	SI
P1-31	6 of 22	SI	**P1-73**	6 of 22	SI
P1-32	11 of 26	SI	**P2-1**	10 of 30	SI
P1-33	11 of 26	SI	**P2-2**	11 of 24	SI
P1-35	6 of 22	SI	**P2-3**	6 of 22	SI
P1-37	21 of 24	SC	**P2-4**	6 of 22	SI
P1-38	11 of 30	SI	**P2-5**	28 of 30	SC
P1-39	6 of 22	SI	**P2-6**	6 of 22	SI
P1-40	12 of 30	SI	**P2-8**	6 of 22	SI
P1-41	11 of 28	SI	**P2-9**	26 of 30	SC
P1-42	6 of 22	SI	**P2-10**	25 of 29	SC
P1-43	8 of 26	SI	**P2-11**	6 of 22	SI
P1-44	6 of 22	SI	**P2-12**	6 of 22	SI
P1-45	22 of 26	SC	**P2-14**	27 of 30	SC
P1-46	6 of 22	SI	**P2-15**	14 of 30	SI
P1-47	6 of 22	SI	**Blanquerna**	27 of 30	SC
			Vivot	5 of 30	SI

In order to confirm the results obtained by microscopic observation, fruit set determinations were also carried out in the parents and in eight selected individuals, five previously considered as SC and three as SI (Table 3.2).

Table 3.2: Fruit sets in bagged branches of the parents and eight target individuals of the almond population 'Vivot' × 'Blanquerna'

Individual	**Phenotype**	**Fruit set (%)**
Vivot (♀ parent)	SI	1.32
Blanquerna (♂ parent)	SC	20.8
P1-45	SC	4.6
P2-5	SC	5.7
P2-9	SC	10.0
P1-28	SC	5.0
P1-51	SC	3.9
P1-5	SI	0
P1-53	SI	1
P1-71	SI	0

Sets after bagging confirmed the SI of 'Vivot', with a percentage of 1.32; however, in the case of 'Blanquerna' results showed its condition of a highly SC cultivar, with a percentage higher then 10%.

The same approach confirmed also the SI phenotype of the seedlings previously identified as SI by pollen tube growth (Table 3.2). Among the SC seedlings, only one (P-2-9) was highly SC according to Grasselly et al. (1981) with a fruit set of 10%, whereas the others may only be considered as SC or partially SC because fruit sets were lower. However, as environment and flower morphology may strongly affect fruit set in bagged branches (Socias i Company et al. 2005), the sets obtained show the evident ability of the flowers of these seedlings to set fruit with their own pollen.

Taking into account that 'Blanquerna' is a SC cultivar and 'Vivot' a SI one, independently of its genotype, at least half of the progeny would have been to show a SC behaviour following a Mendelian distribution of SC. However, this hypothesis does not fit with the results observed during the three consecutive years of microscopic observation of pollen tube growth, with only 24% of SC seedlings. This irregular behaviour could be due to a deficient transmission of the SC allele of 'Blanquerna' to the progeny resulting in a lower than expected ratio of SC seedlings.

Once the phenotypic data was determined in all progeny, the next objective was the determination of the genotype of each seedling by using several molecular techniques. For such objective, firstly the *S*-RNases of all individuals were determined through NEpHGE analysis. In a second step, *S*-allele determination was confirmed by using *S*-allele-specific PCR. Thus, the combination of these different approaches may shed light on of this unexpected behaviour in the individuals studied.

4. GENOTYPE DETERMINATION OF SELF-COMPATIBILITY

4. GENOTYPE DETERMINATION OF SELF-COMPATIBILITY

4.1 Introduction

Incompatibility in almond is of the gametophytic type (GSI). As already described, this system is usually determined by a set of different alleles (the *S* alleles) at a single locus (the *S* locus). Almond has been found to be one of the most polymorphic cultivated species within the *Prunus* genus (Byrne, 1990; Fernández i Martí et al., 200b; Socias i Company et al., 1992), and also the *S* alleles of almond have shown a high diversity (López et al., 2004; Ortega et al., 2005; Kodad et al., 2008a).

Almond breeding programs are pursuing the development of SC cultivars. Thus, the knowledge of the *S*-genotype is very important both for designing crosses for breeding and for selecting the progeny of these crosses (Batlle et al., 1997). In addition, the correct assignment of cultivars to the corresponding cross-incompatibility group (CGI), formed by cultivars possessing the same SI genotype, is essential for orchard design, because the cultivars of the same orchard must belong to different CGIs in order to ensure their efficient reciprocal pollination. The first attempt to establish CIGs was achieved by Kester et al. (1994) who named the first six groups in Californian cultivars by test crossing in the field and in the laboratory. Nowadays, more than 180 almond cultivars have been *S*-genotyped (Ma y Oliveira, 2001; López et al., Ortega et al., 2006; Kodad et al., 2008a; Halász et al., 2009) and 19 CIG have been established (Ortega et al., 2006; Kodad et al., 2008b).

Two different ways have been applied in order to determine the *S* genotype of almond cultivars and progenies. The first method has been the establishment of the cross-compatibility or -incompatibility of a given cross by artificial pollinations carried out in the laboratory, by observing the pollen tube growth following hand pollination of the flowers, or in the field, by recording fruit set following controlled pollinations. However, the determination of *S*-genotype by using these techniques is laborious and time-consuming, as it happens with SC/SI determination as described in the previous section. To overcome these disadvantages, different techniques of molecular biology

have been developed and used routinely for the study of SI in Rosaceae allowing the determination of the *S*-alleles in many almond cultivars and, consequently, the different CGI groups.

In a first step, Tao et al. (1997) identified the *S*-RNases associated with GSI in 13 almond cultivars by using isoelectric focusing (IEF) and bidimensional electrophoresis (2D-PAGE). At the same time, other *S* alleles were detected by using zymograms to determine the presence of stylar ribonucleases (Bošković et al., 1997; 1999). In that study, the different ribonuclease bands were identified when a stylar extract was subjected to a non-equilibrium pH gradient electrofocusing (NEpHGE), thus allowing the separation of the RNase proteins after staining.

This last method of *S* allele determination has been widely applied in many almond cultivars (Duval et al., 1998; Certal et al., 2002; Ortega and Dicenta, 2003; Alonso, 2004; Lopez et al., 2004) as well as in other *Prunus* species such apricot (Burgos et al., 1998) and sweet cherry (Bošković and Tobutt, 1996).

When applying this technique to SC almond cultivars it was found than only one band was identified, thus concluding that the S_f allele, which is considered to be the allele conferring SC in almond (Graselly and Olivier, 1976; Socias i Company and Felipe, 1988), lacks RNase activity. As a consequence, it was concluded that the presence of a single band would indicate a SC genotype. However, this is not always the case because this technique is not very sensitive and does not distinguish alleles with the same isoelectric point (Bošković et al., 2003). Consequently, the presence of a single band may give a negative SC assignment, being this method only fully reliable for the genotype determination in a progeny where the genotypes of the parents are known.

More recently, the sensitivity of allele determination has been increased by using PCR primers to amplify *S* alleles by the PCR reaction. Thus, the nucleotide sequences of several cDNAs deposited in the public databases have served to design conserved and specific primers (Tao et al., 1997; Ushijima et al., 1998; Ma and Oliviera, 2001; Tamura et al., 2000; Channuntapipat et al., 2001, etc.). PCR-based strategies have been developed for the confirmation of the identity of many *S*-alleles and the identification of

new ones (Channuntapipat et al., 2001; López et al., 2006; Ortega et al., 2006; Kodad et al., 2008b; Halász et al., 2009).

The knowledge of the *S*-alleles in almond cultivars is an information directly useful for pomologists and breeders when designing a commercial orchard or planning crosses for breeding new cultivars. Thus, the aim of this chapter was the determination of the *S*-genotype of the parents and the 79 seedlings by two different ways, by NEpHGE and by PCR using both conserved and specific primers, in order to relate the genotype and the phenotype of all individuals and to ascertain the nature of this unexpected phenotype behaviour.

4.2 Methods

4.2.1 DNA Extraction

The preparation of highly pure DNA is essential for DNA-based molecular techniques. In plant species the extraction is hampered by secondary metabolites including carbohydrates, proteins and polyphenolic compounds. For this study the CTAB method has been applied, following the procedure reported by Doyle and Doyle (1987).

CTAB (cetyltrimethylammonium bromide) is a non-ionic detergent that is a major component of the initial extraction buffer. It is used to bind to the DNA and to prevent the binding of DNA to proteins while limiting DNA degradation (Gregory, 2004). Other chemicals present in most extraction buffers include PVP (polyvinyl-pyrrolidone), and β-mercaptoethanol, included to prevent oxidation of the DNA by polyphenolic compounds.

After extraction, DNA samples were quantified using a NANODROP spectrophotometer ND-1000® at the absorption ratio $A_{260/280}$ for quantification and quality control, then diluted to a final stock concentration of 20 ng/µl with TE buffer (10mM Tris, 1mM EDTA pH 8.0) and kept at -20ºC for subsequent PCR amplification.

4.2.2 Protein analysis

The *S*-genotypes were determined by studying the stylar RNases in polyacrylamide gels after non-equilibrium pH gradient electrofocusing (NEpHGE). After separation, the gels were stained for RNase activity. A total of 30 styles from each individual were collected and frozen at -80ºC until use.

The stylar proteins were extracted and separated electrophoretically on poliacrylamide gels containing 1.2% Pharmalyte, pH 6.7-7.7 and 4% Pharmalyte of pH 3 to 10 using NEpHGE, following the conditions described in Bošković et al. (1997). The running comprised 1 h at 150 V, 1h at 300 V and 3 h at 400 V.

After electrophoresis, the gels were removed from the trays and incubated in the buffer at 37ºC during 20 min. After incubation, the solution was changed by another containing RNA and placed also at 37ºC during 40 min. Then, the RNA solution was changed again for a new one. The trays were then incubated again during 20 min.

Finally, the gels were slid into a new tray and dipped into a solution of 7% acetic acid during 3 min (changing it 2 or 3 times) and then placed 90 s in a new tray containing toluidine blue. The gels were the rinsed in distillate water several times. Gels stained with toluidine blue were scanned and analysed.

All protein analysis of each individual were carried out twice for confirmation of the *S*-genotypes.

4.2.3 Identification of S-RNases by PCR

In a first step, the *S*-RNase allele identification of the parents, 'Vivot' and 'Blanquerna', was carried out by PCR using the primer combination PaConsI-F (Sonneveld et al. 2003)/EMPC5consRD (Sutherland et al. 2004) because these primers have been designed to cover the region between the signal peptide and the C5 conserved region. The same primer set was also applied in the determination of the *S*-alleles in the progeny.

In a second step, PCR-based-specific-primers were utilised in order to confirm the genotypes of the seedlings. Hence, the specific primers for the S_f, S_8 and S_{23} developed by Channuntapipat et al. (2003) were used in both parents and all the progeny.

PCR analysis was performed in a volume of 25 µl containing 1x PCR buffer (Invitrogen, Barcelona, Spain), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 µM of each primer and 1 Unit of Platinum® Taq DNA Polymerase (Invitrogen). The amplification reaction for the primer set PaConsI-F/EMPC5consRD was 2 min at 94ºC for denaturation, then 35 cycles of 1 min at 94ºC, 2 min at 50ºC and 4 min at 68ºC followed by a final extension of 10 min at 68ºC (Ortega et al., 2006). Concerning the S_f primers, the PCR program was 3 min at 95ºC for denaturation, then 34 cycles of 30 s at 95ºC, 45 s at 60ºC and 1 min at 72ºC followed by a final extension of 10 min at 72ºC. For the S_{23} and S_8 alleles, the PCR program was the same as for S_f, but with an annealing temperature of 65ºC.

All PCR analysis of each individual were carried out twice for confirmation of the *S*-genotypes

4.2.4 Statistical analysis

SC/SI transmission was tested for different hypothesis by the Chi square (X^2) test of goodness-of-fit using Yates' (1934) correction for one degree of freedom (Table 5.2, in the next chapter).

4.3 Results

4.3.1 S-genotype of the parents

The results obtained by NEpHGE revealed that 'Blanquerna' only produced one band, presumably corresponding to S_8, whereas the other allele, S_f, as expected, did not show RNase activity and did not produce any band (Bošković et al. 1999). Conversely, 'Vivot', produced two different bands (Fig. 4.1), one attributed to S_{23} and the other to

the active S_f (S_{fa}; Kodad et al., 2009) although, as it will be mentioned below, this allele should theoretically lack RNase activity (Bošković et al. 1999).

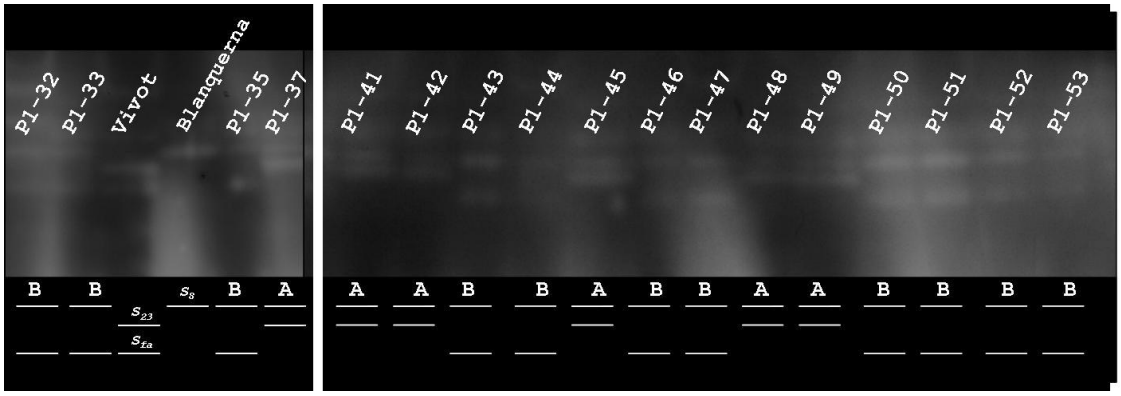


Figure 4.1: Stylar protein electrophoresis (NEpHGE) of 'Vivot', 'Blanquerna' and some seedlings of their progeny showing the two band patterns identified corresponding to the two possible genotypes: S_8S_{23} (A) and S_8S_{fa} (B).

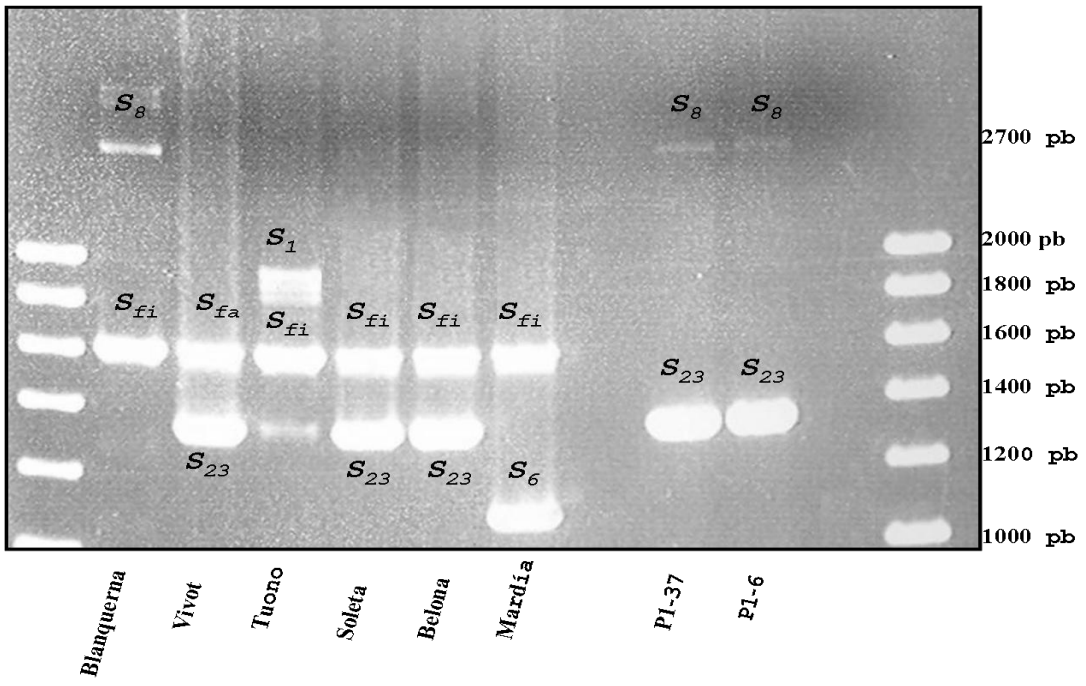


Figure 4.2: PCR amplification of some reference cultivars and two seedlings from the progeny of 'Vivot' × 'Blanquerna' by using the primer combination PaConsI/EMPC5R (Ortega et al., 2006). The two seedlings and 'Blanquerna' show a faint band for the S_8 allele (as mentioned in the text), with a band size of ≈ 2700 pb. Although a third faint band was also observed in 'Tuono', it was probably due to a contamination because later determinations did not show this band.

The PCR amplification by using the conserved primers set PaConsI-F/EMPC5consRD showed two bands for 'Vivot', as expected in any almond cultivar, one corresponding to S_f and the second to S_{23}. In 'Blanquerna', however, only a normal band was obtained, that of S_f, whereas a faint band appeared for a size of ±2700 bp. This second band, corresponding to S_8, showed a very low signal (Fig. 4.2), even sometimes undetectable. A similar situation was also reported by Channuntapipat et al. (2003) when observing that the presence of either S_1 or S_7 masked the amplification of S_8 by PCR when using conserved primers.

4.3.2 S-genotype of the progeny

The zymograms of the *S*-RNases of the progeny by NEpHGE showed two different kinds of band patterns (Fig. 4.1), both showing two different *S*-RNases. By comparing with the *S*-RNase bands of the parents, the bands of the first pattern (A) were attributed to the alleles S_8 and S_{23} and those of the second pattern (B) to S_8 and the active S_f (S_{fa}). Surprisingly, the band corresponding to the S_8 allele present in 'Blanquerna' was also present in all the individuals of the progeny, whereas the inactive S_f (S_{fi}) allele was not present in any plant because all individuals showed two bands. The distribution of genotypes (Table 5.2, in the next chapter) showed that 44.2% of the plants had the S_8S_{23} genotype whereas 55.8% presented the genotype S_8S_{fa}. In two seedlings, P-1-15 and P-1-40, the identification of the genotype by NEpHGE was unsuccessful in spite of repeated attempts.

PCR analysis of the progeny by using the same primer set as for 'Blanquerna' and 'Vivot' (PaConsI-F/EMPC5consRD) showed the amplification of only one band in all seedlings instead of two as expected, belonging either to S_f or to S_{23}. The other allele in all individuals would correspond to S_8 and would have been inherited from 'Blanquerna', but, as in the parent, this allele was undetectable in many cases and gave a very low signal in others. As just mentioned, the amplification of this allele is masked by the presence of some other haplotypes by using conserved primers such as S_1 or S_7 (Channuntapipat et al., 2003). The same phenomenon could take place in this progeny, where amplification of S_8 could be obstructed whenever S_{fa} or S_{23} are also amplified, both in 'Blanquerna', as just described, and in its progeny (Fig. 4.3).

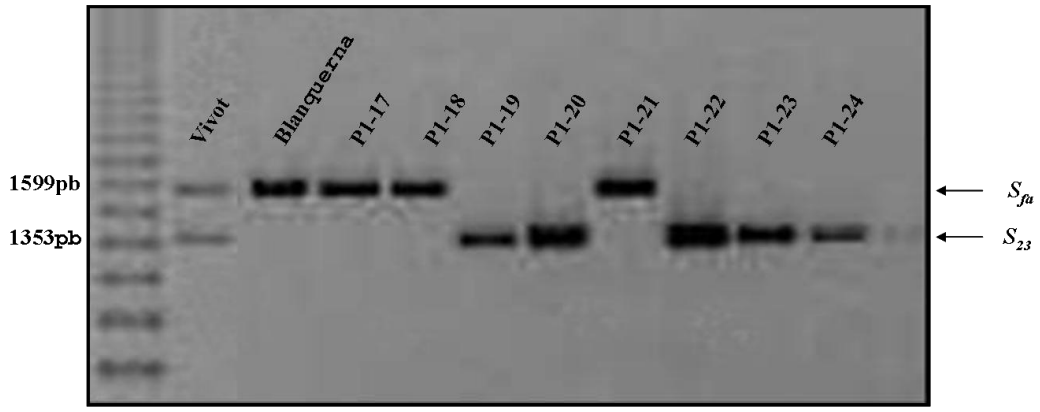


Figure 4.3: PCR amplification by the primer combination PaConsI/EMPC5R (Ortega et al., 2006) of some individuals from the cross 'Vivot' × 'Blanquerna' and the parents showing that only one band was identified in each, S_{fa} or S_{23}, without amplification of the S_8 allele. The band showed in 'Blanquena' is S_{fi}.

For this reason, a specific S_8 primer (Chanuntapipat et al. 2003) was tested in all individuals, being the presence of S_8 confirmed in all the progeny and the parent 'Blanquerna', but not in 'Vivot' (Fig. 4.4). In addition, in order to confirm also the presence of the other alleles in all the progeny, the specific primers for S_{23} and S_f were used in all seedlings. Thus, the determination by PCR allowed to identify that 34 over 77 individuals had the genotype $S_{23}S_8$, corresponding to nearly 45%, whereas the number of seedling having the other genotype, S_fS_8, was of 43 (55%) (Table 5.2, in the next chapter).

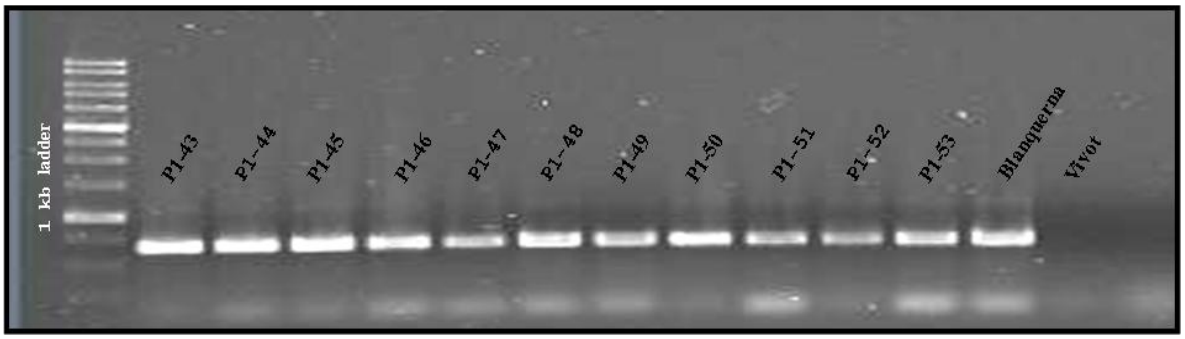


Figure 4.4: PCR amplification of the S_8 allele in some individuals from the cross 'Vivot' × 'Blanquerna' and the parents by using the S_8 specific primer (Channuntapipat et al., 2003).

4.4 Discussion

If the *S* alleles had followed a Mendelian transmission, the progeny 'Vivot' ($S_{23}S_{fa}$) × 'Blanquerna' (S_8S_{fi}) should have the segregation ratio of 1 $S_{23}S_8$: 1 $S_{23}S_{fi}$: 1 $S_{fa}S_8$: 1 $S_{fa}S_{fi}$. However, the molecular results showed that only two genotypes were found in the individuals of this progeny, S_8S_{23} and S_8S_{fa}, and the theoretical segregation was modified to a 1:1 ratio, typical of semi-compatible crosses (Dicenta and García, 1993).

This genotypic ratio might be explained by the recognition and arrest of the inactive S_{fi}-pollen tubes of 'Blanquerna' by the S_{fa}-RNase of 'Vivot' (Fig. 4.5). This hypothesis is strengthened by the lack of the $S_{fa}S_{fi}$ and $S_{23}S_{fi}$ genotypes in the progeny, because all the individuals possessing the S_f allele showed RNase activity, thus showing the presence of the active S_{fa} allele. This S_{fa} must have been inherited from the mother parent, 'Vivot', because the pollen parent, 'Blanquerna', possesses the inactive S_f allele, S_{fi}. Consequently, the pollen parent has only transmitted one of the two possible *S* alleles to the progeny, S_8, and the cross-pollination between the two parents has only been semi-compatible. The active S_{fa} allele from 'Vivot' pistils had thus recognized the inactive S_{fi} allele from 'Blanquerna' pollen, hindering its growth through the 'Vivot' pistils and avoiding the transmission of the inactive S_f allele to the progeny.

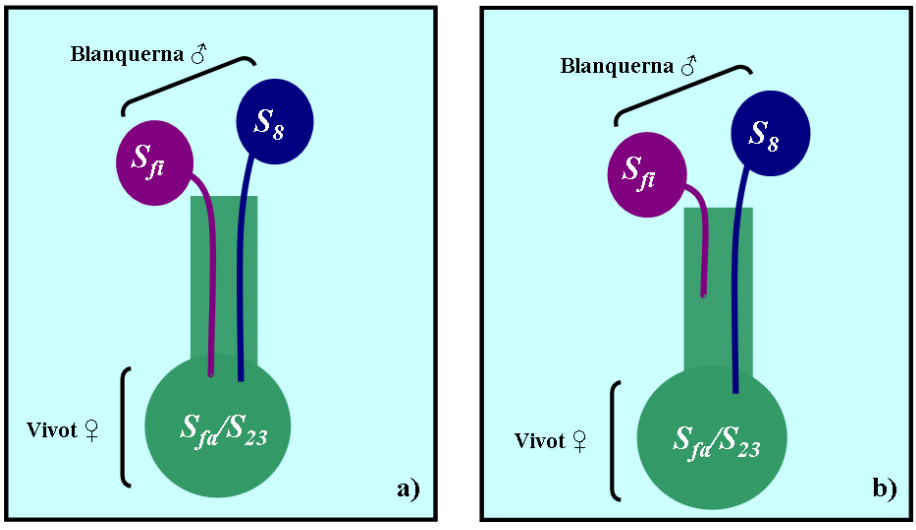


Figure 4.5: Expected pollen tube growth after 'Vivot' pollination by 'Blanquerna' pollen (a). Arrest of the S_{fi} pollen tube growth trough the 'Vivot' pistils (b). Thus, the two only possible genotypes found in the progeny were S_8S_{fa} and S_8S_{23}.

'Vivot' and 'Blanquerna' have been shown to share an identical *S*-locus for the S_f haplotype (chapter 6). This identity has been fully functional in recognizing the *S* alleles in the pollination of 'Vivot' pistils by 'Blanquerna' pollen and confirmed the full activity of S_{fa}, not only in producing a S_f-RNase, but also in recognising an identical allele, even one possessing a different phenotypical expression.

A similar behaviour has also been observed by Bošković et al. (2007) when the pollen of the SC almond 'Falsa Barese' (S_1S_f) failed to pollinate the SI 'Fra Giulio Grande' (S_1S_{30}). Although they incorrectly identified as S_{30} the second allele of 'Fra Giulio Grande', both cultivars had the same S_1S_f genotype, although S_f was in its active form in 'Fra Giulio Grande'. Therefore, the pistils of 'Fra Giulio Grande' recognised both alleles of 'Falsa Barese' pollen as their own, thus rendering this cross-pollination incompatible because of allele identity.

Due to this incompatible recognition, only the S_8 pollen of 'Blanquerna' is able to grow down in the pistils of 'Vivot' giving in the progeny an allele segregation of S_8S_{fa} and S_8S_{23} with an expected distribution 1:1 (Table 5.3, in the next chapter).

Thus, in order to elucidate the possible mechanisms acting in this unexpected reaction, several approaches were undertaken and developed in the following chapters to advance with alternative hypothesis and to create new proposals concerning SI/SC transmission and reaction in almond, as well as in other Rosaceae species.

5. PRESENCE OF A MODIFIER LOCUS FOR SELF-COMPATIBILITY

5. PRESENCE OF A MODIFIER LOCUS FOR SELF-COMPATIBILITY

5.1 Introduction

Almond SI is controlled by a single locus, the *S*-locus (Socias i Company, 1984) which has been shown to contain two components, one specifically expressed in the style (stylar-*S*), and the other in the pollen (pollen-*S*). The stylar-*S* component encodes an allelic series of stylar glycoproteines with ribonuclease activity, *S*-RNases (Tao et al., 1997), responsible for the inhibition of pollen tube growth through the degradation of the pollen RNA, since *S*-RNases are thought to function as specific cytotoxins (McClure et al. 1999). The pollen-*S* component has been recently identified in *Prunus* as the *S*-locus F-box gene (SFB), being tightly linked to the *S*-RNase gene. This pollen-*S* component has been identified in the *S* locus of several *Prunus* species, including almond (Ushijima et al. 2003), Japanese apricot, *P. mume* (Sieb.) Sieb et Zucc. (Entani et al. 2003; Yamane et al. 2003a), apricot, *P. armeniaca* L. (Romero et al. 2004), sweet cherry, *P. avium* L., and sour cherry, *P. cerasus* L. (Yamane et al. 2003b), and Japanese plum, *P. salicina* Lindl. (Watari et al., 2007).

Natural or artificially induced self-compatible (SC) mutants in SI species have been used to study the molecular basis of the SI mechanism. In both cases, SI may result from a loss of function of the factors in the *S* locus (Tao et al. 2007) or from components external to the *S* locus (de Nettancourt 2001).

Stylar-part mutations (SPM) in the *S*-locus have been reported in several families with gametophytic SI (GSI) such as Solanaceae (Royo et al. 1994) and Rosaceae (Sassa et al. 1997), revealing that ribonuclease activity of the *S*-RNases is required to inhibit pollen tube growth. The expression of SC in almond has been attributed to the presence of the S_f allele, which is dominant over the other *S* alleles (Socias i Company 1984). This S_f allele is thought to lack RNase activity (Bošković et al. 1999), possibly because of the low transcriptional level of the S_f-RNase (Hanada et al. 2009).

SC could also be due to a mutation in the F-box gene (pollen-part mutation or PPM). To date, numerous PPMs in *Prunus* have been shown to be caused by mutations in the *SFB*, such as the S_3' and S_4' haplotypes in sweet cherry, which were produced by X-ray irradiation (Lewis and Crowe 1954) and appeared to be mutated from the original *SFBs* (Ushijima et al. 2004; Sonneveld et al. 2005). The *SFB* is completely deleted in S_3', while S_4' has a frame shift mutation. The *SFB* in the SC S_f haplotype of Japanese apricot was supposed to be a PPM and has been shown to have an insertion in the middle of the coding region (Ushijima et al. 2004). The SC Italian sweet cherry 'Kronio' has a PPM because the S_5' haplotype encodes a truncated *SFB* (Marchese et al. 2007).

Breakdown of SI has also been associated with mutations affecting a modifier locus unlinked to the *S*-locus. Tsukamoto et al. (2003) have described that the breakdown of SI in a natural population of *Petunia axillaris* (Lam.) B.S.P. subsp. *axillaris* was due to a modifier locus which is unlinked to the *S*-locus and suppresses the expression of an *S*-RNase gene. In Rosaceae, Wünsch and Hormaza (2004) have also reported that the pollen *S* component of the SC sweet cherry 'Cristobalina' is affected by a factor unlinked to the *S* locus. Similarly, the presence of a mutation in a modifier locus has been reported to affect the function of the *S*-pollen factor in the SC apricot 'Canino' (Vilanova et al. 2006). And more recently, the presence of non-*S*-specific factors affecting the SI reaction has also been observed in European pear (*Pyrus communis* L.; Moriya et al. 2009).

Most almond breeding programmes aim to produce SC cultivars that would be suitable for monocultivar orchards, and less dependent on the activity of bees for pollination (Socias i Company 1990). SC in almond breeding programmes has mostly derived from 'Tuono', a cultivar from the Italian region of Apulia (Socias i Company 2002). However, it would be interesting to search for different sources of SC in almond in order to avoid inbreeding depression in the breeding progenies (Alonso and Socias i Company 2005).

Recently the S_f allele has been identified by PCR analysis in several local cultivars from the island of Majorca (Kodad et al. 2008b), including 'Vivot' and 'Ponç', which were supposed to be self-incompatible (Rubí 1980). As already stated, the

progeny obtained from the cross 'Vivot' ($S_{23}S_f$) × 'Blanquerna' (S_8S_f) showed an unexpected self-incompatible phenotype in most of the seedlings. Furthermore, both 'Ponç', with a genotype $S_{27}S_f$, and 'Vivot' had shown a SI behaviour (Fernández i Martí et al., 2008, 2009a; Kodad et al. 2009), thus raising the possibility that the expression of SC in almond may not only be due to the presence of the S_f allele, which in 'Ponç' and 'Vivot' showed RNase activity, being considered an active S_f allele and called S_{fa}.

Due to the unexpected SI behaviour of most seedlings of the progeny 'Vivot' × 'Blanquerna', the aim of this chapter was to ascertain the nature of the strange pollination behaviour of this progeny.

5.2 Materials and Methods

In this chapter we ha used a multilevel approach, including the observation of pollen tube growth, fruit sets, protein analysis and PCR analysis which have already been detailed in the previous chapters.

5.3 Results

All results have already been described in the previous chapter. In the present one all results are shown in Table 5.1 for comparison, including phenotypes and genotypes of the parents and the progeny.

5.4 Discussion

According to the *S*-genotype distribution in the progeny (S_8S_{fa} or S_8S_{23}), all the seedlings would have been SI because none of them possesses the S_{fi} allele presumably conferring SC. However, a SC phenotype was observed in 19 of the 77 plants, nearly 25%. These SC individuals included plants of the two possible genotypes because 14 of them showed the S_8S_{23} genotype (73.7%) and five the S_8S_{fa} genotype (26.3%). Consequently, their SC cannot be due to the S_f allele but to an additional genetic factor

affecting the SI mechanism which until now had not been observed because it could be linked to the S_f allele conferring SC in almond.

Table 5.1: Phenotype and genotype determination in all seedlings studied in this work

Individual	Phenotype [x]	NepHGe	PCR
Vivot (♀ parent)	SI	$S_{23}S_{fa}$	$S_{23}S_{fa}$
Blanquerna (♂ parent)	SC	S_8S_{fi}	S_8S_{fi}
P1-1	SI	S_8S_{23}	S_8S_{23}
P1-3	SI	S_8S_{fa}	S_8S_{fa}
P1-5	SI	S_8S_{fa}	S_8S_{fa}
P1-8	SI	S_8S_{23}	S_8S_{23}
P1-10	SI	S_8S_{23}	S_8S_{23}
P1-11	SC	S_8S_{23}	S_8S_{23}
P1-12	SI	S_8S_{fa}	S_8S_{fa}
P1-13	SC	S_8S_{fa}	S_8S_{fa}
P1-14	SI	S_8S_{fa}	S_8S_{fa}
P1-15	SI	- [y]	S_8S_{23}
P1-16	SI	S_8S_{fa}	S_8S_{fa}
P1-17	SI	S_8S_{fa}	S_8S_{fa}
P1-18	SC	S_8S_{fa}	S_8S_{fa}
P1-19	SI	S_8S_{23}	S_8S_{23}
P1-20	SC	S_8S_{23}	S_8S_{23}
P1-21	SI	S_8S_{fa}	S_8S_{fa}
P1-22	SI	S_8S_{23}	S_8S_{23}
P1-23	SI	S_8S_{23}	S_8S_{23}
P1-24	SI	S_8S_{23}	S_8S_{23}
P1-25	SC	S_8S_{23}	S_8S_{23}
P1-26	SC	S_8S_{23}	S_8S_{23}
P1-27	SI	S_8S_{fa}	S_8S_{fa}
P1-28	SC	S_8S_{fa}	S_8S_{fa}
P1-30	SI	S_8S_{fa}	S_8S_{fa}
P1-31	SI	S_8S_{fa}	S_8S_{fa}
P1-32	SI	S_8S_{fa}	S_8S_{fa}
P1-33	SI	S_8S_{fa}	S_8S_{fa}
P1-35	SI	S_8S_{fa}	S_8S_{fa}
P1-37	SC	S_8S_{23}	S_8S_{23}
P1-38	SI	S_8S_{fa}	S_8S_{fa}
P1-39	SI	S_8S_{23}	S_8S_{23}
P1-40	SI	- [y]	S_8S_{23}
P1-41	SI	S_8S_{23}	S_8S_{23}
P1-42	SI	S_8S_{23}	S_8S_{23}
P1-43	SI	S_8S_{fa}	S_8S_{fa}
P1-44	SI	S_8S_{fa}	S_8S_{fa}
P1-45	SC	S_8S_{23}	S_8S_{23}
P1-46	SI	S_8S_{fa}	S_8S_{fa}
P1-47	SI	S_8S_{fa}	S_8S_{fa}
P1-48	SC	S_8S_{23}	S_8S_{23}

P1-49	SC	S_8S_{23}	S_8S_{23}
P1-50	SI	S_8S_{fa}	S_8S_{fa}
P1-51	SC	S_8S_{fa}	S_8S_{fa}
P1-52	SI	S_8S_{fa}	S_8S_{fa}
P1-53	SI	S_8S_{fa}	S_8S_{fa}
P1-54	SI	S_8S_{fa}	S_8S_{fa}
P1-55	SI	S_8S_{fa}	S_8S_{fa}
P1-56	SC	S_8S_{23}	S_8S_{23}
P1-57	SC	S_8S_{23}	S_8S_{23}
P1-58	SI	S_8S_{23}	S_8S_{23}
P1-59	SI	S_8S_{fa}	S_8S_{fa}
P1-60	SI	S_8S_{fa}	S_8S_{fa}
P1-61	SI	S_8S_{23}	S_8S_{23}
P1-62	SI	S_8S_{fa}	S_8S_{fa}
P1-63	SI	S_8S_{fa}	S_8S_{fa}
P1-64	SI	S_8S_{fa}	S_8S_{fa}
P1-65	SI	S_8S_{fa}	S_8S_{fa}
P1-66	SC	S_8S_{fa}	S_8S_{fa}
P1-67	SI	S_8S_{23}	S_8S_{23}
P1-68	SI	S_8S_{23}	S_8S_{23}
P1-69	SI	S_8S_{fa}	S_8S_{fa}
P1-71	SI	S_8S_{fa}	S_8S_{fa}
P1-72	SI	S_8S_{23}	S_8S_{23}
P1-73	SI	S_8S_{23}	S_8S_{23}
P2-1	SI	S_8S_{fa}	S_8S_{fa}
P2-2	SI	S_8S_{fa}	S_8S_{fa}
P2-3	SI	S_8S_{fa}	S_8S_{fa}
P2-4	SI	S_8S_{fa}	S_8S_{fa}
P2-5	SC	S_8S_{23}	S_8S_{23}
P2-6	SI	S_8S_{fa}	S_8S_{fa}
P2-8	SI	S_8S_{23}	S_8S_{23}
P2-9	SC	S_8S_{23}	S_8S_{23}
P2-10	SC	S_8S_{23}	S_8S_{23}
P2-11	SI	S_8S_{fa}	S_8S_{fa}
P2-12	SI	S_8S_{fa}	S_8S_{fa}
P2-14	SC	S_8S_{23}	S_8S_{23}
P2-15	SI	S_8S_{23}	S_8S_{23}

[x] SC: self-compatible; SI: self-incompatible

[y] Doubtful genotype

During nearly 40 years some major genes outside the *S* locus have been considered to contribute to the control of SI (Thompson and Taylor 1971; Hinata et al. 1995; Tsukamoto et al. 2003). The combination of our results, including test crosses, stylar RNase analysis and PCR allele identification, may suggest that a modifier locus located outside the *S* complex may be the responsible for the genetic control of SC/SI in our progeny. The parents had not shown any part mutation, either in the stylar or in the

pollen components of the S_f-haplotype, since no differences were found between their sequences (Chapter 6). Thus, their different behaviour must be due to a factor external to the S_f-haplotype, being a modifier gene a good candidate to explain their different phenotypic expression.

The presence of SC in almost 25% of the progeny, independently of the presence and activity of two *S*-RNases, suggests a model involving a modifier locus unlinked to the *S*-locus and controlling the SI/SC recognition mechanism. The ratio of 1:3 for SC:SI seedlings (Table 5.2 and 5.3) suggests that this modifier allele should be recessive and that it should work in the style. If this modifier gene had been dominant or co-dominant or had worked in the pollen, 50% or more than 50% of the plants should have been SC.

Table 5.2: Analysis of the ratio of phenotypes and genotypes in the 'Vivot' × 'Blanquerna' almond progeny as observed in Table 5.1

Technique	Nº of progeny observed	Expected ratio	Nº of observed genot/phenot	X^2	α
NEpHGE	77	1 S_8S_{fa}:1 S_8S_{23}	43:34	0.831	0.361
PCR	77	1 S_8S_{fa}:1 S_8S_{23}	43:34	0.83	0.363
Microscopy	77	1 SC : 1 SI	19:60	18.75	1.5×10^{-5}
Microscopy	77	3 SC :1 SI	19:60	0.0042	0.94

The term M for "Modifier" is being proposed for this model. The genotypes of the parents would be $S_{fi}S_8$/Mm for 'Blanquerna' and $S_{fa}S_{23}$/Mm for 'Vivot'. Since the S_{fi} pollen from 'Blanquerna' could not grow down in the S_{fa} pistils of 'Vivot', the segregation in the progeny would be: 1 S_8S_{fa}/MM: 2 S_8S_{fa}/Mm: 1 S_8S_{fa}/mm: 1 S_8S_{23}/MM: 2 S_8S_{23}/Mm: 1 S_8S_{23}/mm, with an overall ratio of 1:2:1 for MM:Mm:mm. If SC would be due to the presence of "mm", only 25% of the offspring should be SC, which is very close to the distribution obtained from microscopic observations (24.7%).

A similar model was proposed by Tsukamoto et al. (2003) in a population of *Petunia axillaris* containing mostly SI plants, where the breakdown of SI was caused by a modifier locus suppressing the expression of the *S*-RNases by controlling the transcription of the style function of some alleles, such as S_{13} but not of S_1 or S_{15}, showing the dominance of the "M" allele over the "m" allele.

Table 5.3: Predicted *S*-genotypes and percentage of SC/SI seedlings

S_8S_{23}	S_8S_{23}	S_8S_{23}	/	S_8S_{fa}	S_8S_{fa}	S_8S_{fa}
MM	Mm	mm	/	MM	Mm	mm
1	2	1		1	2	1
25%	50%	25%		25%	50%	25%
SI	SI	SC*		SI	SI	SC*

* The ratio 1:3 for SC/SI suggests that the modifier allele should be recessive and that it should work in the style.

Several external factors affecting SI have been reported in other *Prunus* species (Wunsch and Hormaza 2004; Vilanova et al. 2006; Moriya et al., 2009). Different types of modifier "factors" have been suggested to take part in the breakdown of SI, including an activator for the production of an active suppressor of the *S*-RNase expression and also the opposite, a non-functional activator and even a null allele (Tsukamoto et al. 2003).

The modifier locus affecting S_f in almond may encode an activator of the expression of the SI mechanism because the almond genotypes carrying the S_{fa} (or the mislabelled S_{30}) allele, such as 'Vivot, 'Ponç', 'Cinquanta Vignali' and 'Fra Giulio Grande', show RNase activity and are phenotypically SI in spite of the homology between the S_{fa} (or S_{30}) and the S_{fi} sequences. Thus, SC could be due to the loss of this activator function, suppressing the transcription of the S_{fa}-RNase alelle to become the S_{fi}-RNase allele. But, in the other side, plants with the SI genotypes S_8S_{23} or S_8S_{fa} but phenotypically SC, must carry a suppressor for the expression of the gene. Thus, both an activator and a suppressor may be involved in this population.

Also other genes not related to the *S* locus may be suggested as responsible for SC in our plants and in those reported to carry the S_{fa} allele. Since SC was found in 14 plants with the S_8S_{23} and in five with the S_8S_{fa} genotype, SC may have two different kinds of expression in almond, showing that the presence of the S_f allele is not always linked to SC.

Further steps are needed to elucidate the relationship of the modifier gene presumably present in this progeny with the change from S_{fa} to S_{fi}.

6. THE DOUBLE EXPRESSION OF THE S_f HAPLOTYPE

6. THE DOUBLE EXPRESSION OF THE S_f HAPLOTYPE

6.1 Introduction

Self-compatibility has been naturally found in several almond cultivars, although it happens rarely. The presence of the S_f haplotype has been always related to SC which is dominant over SI (Socias i Company 1984). As SC is a useful agronomic trait, it has been included as one of the main objectives in most almond breeding programmes (Socias i Company 1990).

It has been recently found that the S_f allele shows two different phenotypic expressions, one SI and the other SC, as shown by pollination tests and by NEpHGE analysis (Fernández i Martí et al. 2008; Kodad et al. 2010). This discrepancy has been noticed in genotypes from different growing regions (Bošković et al. 2007; Kodad et al. 2009; Martínez-García 2009). However, the nucleotide alignment in the five conserved domains (C1, C2, C3, RC4 and C5) and the hyper-variable region is the same in all the genotypes having the S_f allele, independently of their phenotypic expression.

In addition, the reciprocal recognition of both forms of the S_f allele has been described in the previous chapeters, as shown by the arrest of the SC S_f pollen tubes of 'Blanquerna' by the SI S_f -RNase of 'Vivot'. This fact indicates that the active S_f allele of 'Vivot' pistils has recognised the inactive S_f allele from 'Blanquerna' pollen, preventing the transmission of the inactive S_f allele (S_{fi}) to the offspring. As a consequence, it was suggested that the breakdown of SI in almond could be caused by a modifier gene or by an alteration located outside the coding region of the S_f gene.

Cloning and sequencing of the S_f haplotype by traditional TA cloning only gives a partial sequence of the coding region (Bošković et al. 2007; Kodad et al. 2009). As a consequence, the construction of a genomic library to find out if any difference between the two versions of the S_f haplotype may indicate if they differ in the transcription factor or if this difference is located outside the *S* locus. Thus, it was decided to construct a fosmid library because this kind of genomic library has several advantages over the others. Firstly, a fosmid library is easier to handle than a BAC library due to its small insert size (ca 50 kb); consequently, in many comparative genomic studies, the region of

interest is not so large as that covered by a fosmid clone, as it happens in the case of the *S* locus in almond. Secondly, even if the fosmid clones are of similar size than the cosmid clones, fosmids are more stable than cosmids (Kim et al. 1992), and also more suitable for rapidly creating genomic or chromosome-specific libraries (Gingrich et al. 1996).

Hence, the main objective of this chapter was the characterisation of the *S* locus of 'Vivot' and 'Blanquerna' by constructing a fosmid library, in order to find out any substitution in their genomic sequences or to locate the external factor responsible for the breakdown of SI in almond. On the other hand, we were also interested in measuring the expression level of the S_{fa} and S_{fi} RNases trough quantitative PCR.

6.2 Materials and methods

6.2.1 Plant material

For these experiments we have used pistils, pollen grains and young leaves from 'Vivot' and 'Blanquerna', as well as from 'Tuono' as a control.

6.2.2 Isolation of genomic DNA

DNA was isolated from young leaves using a Nucleon™ PhytoPure™ Genomic DNA Extraction Kit (GE Healthcare, Piscataway, NJ, USA) following the manufacturer's protocol. This kit was used because a large amount of DNA of high quality is required for constructing a genomic library. After DNA extraction, the samples were treated with RNase, in order to inhibit RNases. The DNA solution was then further purified by phenol/chloroform extraction and the DNA was precipitated with ethanol to be finally dissolved in TE. DNA quantification was performed by comparison with a λ DNA marker (Promega, Madison, WI, USA).

6.2.3 RT-PCR analyses and nested PCR

Total RNA was extracted from pollen grains and styles following the cold phenol SDS method described by Tao et al. (1999b). One microgram of total RNA treated with DNase I (Invitrogen, Tokyo, Japan) was used for first strand cDNA synthesis by SuperScriptIII RT (Invitrogen). The *S*-RNase and SFB were amplified using the primer combination PruT2/rtPd_S_fR, PruC2/PruC4 and Pd_SFB$_f$F1/Pd_SFB$_f$R1 (Hanada et al., 2009). Expression of the *actin* gene was used as a control with the primers *Act*F1/*Act*R1 (Ushijima et al. 2003) following the PCR conditions described by Yamane et al. (2003a).

The PCR reaction mixture contained 1X *ExTaq* buffer, 250 µM of each dNTP, 200 nM of each primer, 20 ng of template cDNA and 0.5 units of *ExTaq* polymerase (TaKaRa Bio, Shiga, Japan) in a 25 µl reaction volume.

The PCR reaction was run with the following program: 3 min of denaturation at 94ºC; 35 cycles of 94ºC for 1 min, 58ºC for 1 min and 72ºC for 1 min 30 s, and a final extension of 7 min at 72ºC. Nested PCR was used in order to confirm the transcription of the S_f-RNase gene in 'Vivot'. For this, 2 µl of the RT-PCR product was used as a template and re-amplified using the combination primers PruC2/rtPd_S_fR. PCR was performed using a program of 35 cycles of 94ºC for 1 min, 60ºC for 1 min and 72ºC for 1 min 30 s, with an initial denaturation step at 94ºC for 3 min and a final extension step at 72ºC for 7 min.

6.2.4 Expression analysis by Real Time PCR of the stylar product

Quantative RT-PCR was performed with a LightCycler® 480 Real-Time PCR System (Roche). The PCR mixture (10 µl) contained 1 µl cDNA, which contained cDNA equivalent to the amount synthesized from 0.025 µg total RNA, 0.20 µM of each primer, 10 µM TaqMan® probe and 5 µl Probe Master 2x (Roche). All reactions were compared with a SyBr Green Supermix (Roche) using as internal control the housekeeping gene (*Actin*). Final concentration of the *Actin* primer was 0.5 µM and for the S_f, S_8 and S_{23} primers were 0.9 µM. Specific primers and probes for S_f, S_8 and S_{23}

were specifically designed for this study. Sequences of probes and primers designed in this work are listed in Table 6.1.

Table 6.1: Probe and primer sequences for the S_f, S_8 and S_{23} almond RNases designed in this work to be used in quantitative rt-PCR

Type	Name	Sequence 5' → 3'	Target allele
Probe	S_f-RNase	TAG TGG TGA TGG ATC TTA	S_f
Probe	S_8-RNase	TGA TCC AGC AAT GCC TA	S_8
Probe	S_{23}-RNase	TTG CAA ACA GGC TAA GAA	S_{23}
Primer	S_f-F	CTT GTT CTT GCT TTT GCT TTC TTC	S_f
Primer	S_f-R	ATT TAC AAT TTG TGC AAC AAT GGC	
Primer	S_8-F	ACC CAC CCT TCG TTG CAA A	S_8
Primer	S_8-R	CAA CTC TCA GTT GTT ACA TGA AGT GGT	
Primer	S_{23}-F	CTC GGA CAT AAT ATC ACC CAT TAA AG	S_{23}
Primer	S_{23}-R	TGT TGT TAC ATG GAA GTG GTA TTT TGT TAT TG	

Melt curve analysis was performed for SYBR® green reaction after each run to determine the specificity of the amplification of *Actin*. A standard dilution curve with known amounts of cDNA for each gene was added and used for estimating the quantity of each gene, being each sample analysed thrice. The PCR reaction was run with the following program: 2 min of pre-incubation at 95ºC; amplification with 45 cycles of 95ºC for 30 s, 57ºC for 30 s and 72ºC for 15 s. The gene quantification method was based on the relative expression of the target versus the reference gene, *Actin*.

6.2.5 Construction and screening of a genomic library

For cloning the *S*-locus of 'Vivot' and 'Blanquerna' fosmid libraries were constructed from genomic DNA using the CopyControl Fosmid Library Production Kit (Epicentre, Madison, WI, USA) following the procedure described by Ikeda et al. (2004). Libraries were screened for the presence of *S*-RNase and SFB by using DIG-dUTP-labelled probes. After screening, positives clones were isolated and the presence of the *S*-RNase and the SFB was checked by PCR using conserved primers. A genomic clone that contained the *S* locus of the *S* haplotypes (S_f, S_{23} and S_8) were selected and used for further studies. The nucleotide sequences of the clones were determined by primer walking and sequenced in a CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA). The deduced amino acids of the *S*-RNases and SFBs were aligned using Clustal X (Thompson et al., 1997)

6.2.6 Analysis of the physical distance between S-RNase and SFB

Fosmid clones were used as template for PCR analysis of the physical distance between the two genes (*S*-RNase and SFB). The physical distance was determined by using the primer combinations Pru-C2/SFBC5F, Pru-C2/rtPD_SFB$_f$_F and rtPd_S_f_F/SFBC5F (Table 6.2; Hanada et al., 2009) and the Expand Long Template PCR System (Roche, Nutley, NJ, USA) following the procedure described by Watari et al. (2007).

Each reaction consisted of 1x Expand Long Template buffer, 3.5mM of each dNTP, 2.75 mM $MgCl_2$, 300 nM of each primer and 25 ng of fosmid DNA as a template, and one unit of Expand Long Template Enzyme mix in a 15 µl reaction volume.

Long PCR was performed using a program with an initial denaturing step at 94ºC for 2 min, 10 cycles of 94ºC for 10 s, 58ºC for 30 s, 68ºC for 15 min, 20 cycles of 94ºC for 10 s, 58ºC for 30 s, 68ºC for 15 min with an extension of an extra 20 s after each cycle, and final extension at 68ºC for 7 min. Primer combinations used in this study are shown in Figure 6.1. Amplified fragment were separated by 1% agarose gel electrophoresis and visualized with ethidium bromide under ultraviolet light.

6.2.7 DNA gel blot analysis

Six µg of genomic DNA of 'Vivot', 'Blanquerna' and 'Tuono' were digested with two restriction enzymes (*Eco*RI and *Hin*dIII), electrophoresed on 0.6% agarose gels and blotted onto a nylon membrane. The membrane was probed with DIG-dUTP-labeled DNA fragment and washed under high-stringency conditions (twice at room temperature in 2X SSC, 0.1% SDS for 5 min and twice at 68ºC in 0.1X SSC, 0.1% SDS). Hybridisation signals were visualized as described in Tao et al. (1999a).

The DIG-dUTP-labelled probe was synthesized by PCR using the sweet cherry S_6-RNase cDNA with the Pru-C2 and Pru-C4R primer set (Tao et al. 1999a) following the conditions described by Watari et al. (2007). For SFB, the DIG-labelled probe was synthesized using the sweet cherry SFB_6 and the primer combination SFB-C1F/SFB-C5R (Yamane et al. 2003b).

6.3 Results

6.3.1 Characterization of the S-locus in 'Vivot' and 'Blanquerna'

In spite of the different phenotypes of 'Vivot' (SI) and 'Blanquerna' (SC), both cultivars gave positive amplifications by the specific S_f primer. Since the partial sequence obtained from both cultivars showed the same nucleotide alignment (Fernández i Martí et al. 2008), a fosmid library was constructed in order to clarify the possible location and/or the reason of their different S_f expression. Genomic clones containing the *S* locus from the S_f and S_8 haplotypes in 'Blanquerna' and the S_f and S_{23} haplotypes in 'Vivot' were successfully cloned.

Using the genomic clones, a fragment region of about 4.1 kb containing the entire S_f-RNase coding sequence with 5'- and 3'-flanking regions was sequenced. There appeared to be the putative TATA box and the IB-like motif in the 5'-flanking region upstream of the start codon of S_f-RNase of both cultivars. This motif has been reported to be necessary for *S*-RNase expression in potato (Ficker et al. 1998) and is located within the CR (conserved region situated upstream from the putative TATA box).

The amino acid alignment of the S_f-RNase and SFB$_f$ from both 'Blanquerna' and 'Vivot' did not show any difference (Fig. 6.1), even in the region upstream the start codon (TATA box and IB-like motif). Thus, the S_f haplotypes found in both cultivars are exactly the same and do not show any mutation, neither in the stylar nor in the pollen part. As it has been previously mentioned, the only difference is at the phenotypic level, through microscopic observations, fruit set after bagging, as well as through the activity of the *S*-RNases (Bošković et al. 2007; Fernández i Martí et al. 2008; Kodad et al. 2009).

Additionally, the DNA sequences obtained in this work has been deposited in the EMBL Nucleotide Sequence Database. The 5.8-kb and 2.3-kb regions that include respectively the S_8-RNase (AB481108) and *SFB$_8$* (AB480747) genes from the other *S*-haplotype of 'Blanquerna' were also determined. Furthermore, the 3.2-kb region that includes the S_{23}-RNase (AB488496) gene from 'Vivot' was determined.

6.3.2 Physical distance between S_f-RNase and SFB$_f$

Based on the length of the amplified products, the estimated physical distance between the *S*-RNase and the SFB was determined by long-PCR which yielded three different length of fragments (A: 7.6 kb, B: 7.3 kb and C: 6.3 kb). Apparently, there is a distance of about 6 kb between the 2 genes (Fig. 6.1), similar to that reported by Hanada et al. (2009).

Table 6.2: DNA sequences of oligonucleotide primers used in this study (Hanada et al., 2009)

Primer Name	Sequence
Pru-T2	5'-TST TST TGS TTT TGC TTT CTT C-3'
rtPd_SbR	5'-TCT CTT CAG TTT GAC TCG CAT TTT-3'
rtPd_SfR	5'-GTC GTC GCT TGG CTC TTA GG-3'
Pd_SFBb_F1	5'-CAC CAA GTT TTG AAT GTC AGG TT-3'
Pd_SFBb_R1	5'-CAC TAG GAT ATA GGT CAA GAA GCA-3'
Pd_SFBf_F1	5'-CAC CCA AAT TTT GAA CGA AAA GAC-3'
Pd_SFBf_R1	5'-TTC CTC CTC ACA AGA ATA AAA CTT-3'

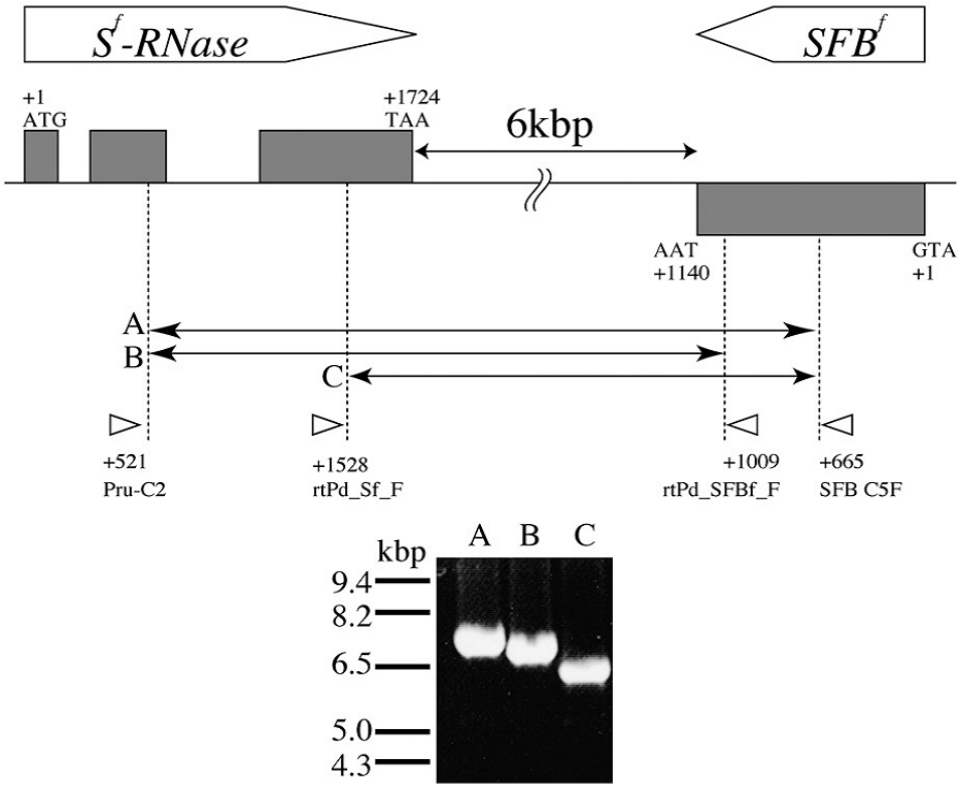


Figure 6.1: Estimated physical distance between the *S*-RNase and the SFB, adapted from Hanada et al. (2009). Primer combinations are as follow: (A) Pru-C2/SFB C5F, (B) Pru-C2/rtPd_SFBf_F and (C) rtPd_Sf_F/SFB C5F

6.3.3 DNA blot analysis

Genomic DNA blot analyses with *S*-RNase and *SFB* probes using several enzymes (*Eco*RI and *Hin*dIII) were carried out to confirm the *S*-RNases and SFBs of 'Vivot' and 'Blanquerna' identity. Southern blot analysis gave the same haplotype-specific band size for S_f in 'Vivot' and in 'Blanquerna'. The other *S* haplotype band in 'Vivot' was identified as S_{23}. The second *S* haplotype band of 'Blanquerna' was also hybridized and considered to be S_8. The SFB probes also revealed that the S_f signals appeared in the same position in 'Blanquerna' as in 'Vivot'. The analyses were also conducted with 'Tuono' for comparison obtaining the expected results (Fig 6.2).

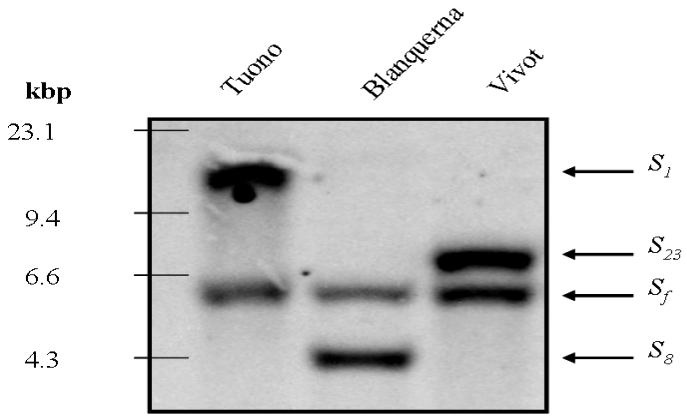


Figure 6.2: Southern blot analysis of genomic DNA of 'Vivot', 'Blanquerna' and 'Tuono' digested with *Eco*RI restriction enzyme.

6.3.4 RT-PCR and RNase activity analyses

In order to confirm the activity in the styles of both parents, pistil RNA was used to synthesize cDNA trough reverse transcription. With the primary RT-PCR (Fig. 6.3b) and by Nested PCR (Fig. 6.3c), the S_f allele produced mRNA for an active S_f-RNase in 'Vivot' styles. On the contrary, the absence of S_f-RNase activity in the pistils of 'Blanquerna' appeared to be from the lack of transcription, as suggested by the NEpHGE results in Chapter 4 (Fig. 6.3d).

For the pollen part, allele-specific RT-PCR showed that pollen SFB_f from 'Vivot' and 'Blanquerna' were detected from pollen cDNA in both cultivars (Fig. 6.3e). *Actin* specific primer set was used for reference control in genomic DNA and cDNA from style and pollen of 'Blanquerna', 'Vivot' and 'Tuono' (Fig. 6.3a).

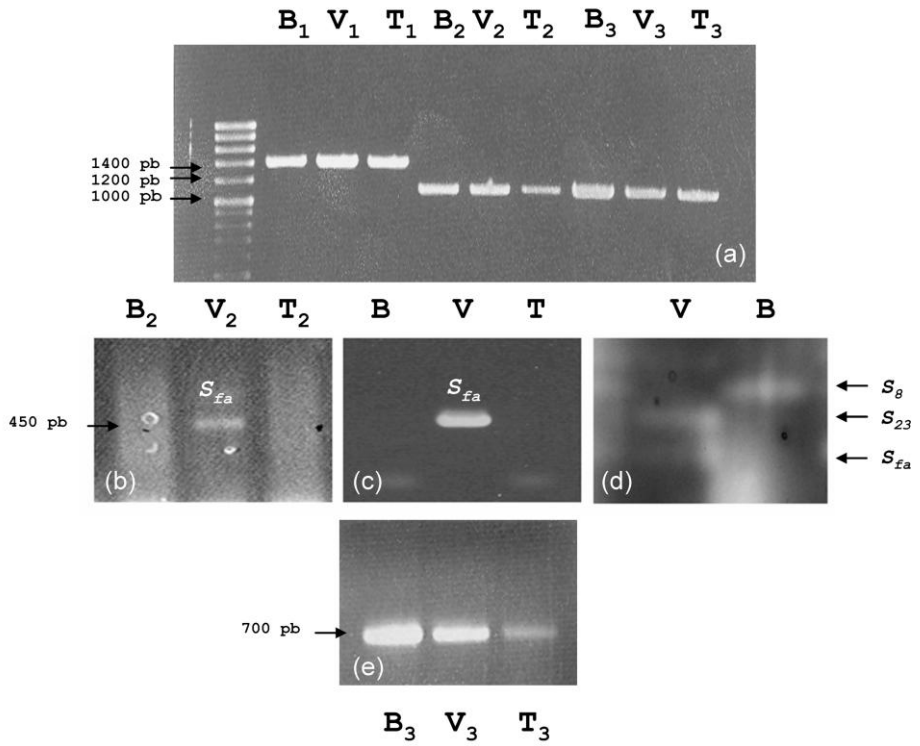


Figure 6.3: Expression analysis of *S*-RNases and SFBs in almond style and pollen. **(a);** Actin amplification from RNA obtained from pistil (2) and pollen (3) of 'Blanquerna'(B), 'Vivot' (V) and 'Tuono' (T). PCR was also performed with genomic DNA (1) of 'Blanquerna', 'Vivot' and 'Tuono'. **(b)**; cDNA from pistil (2) of 'Blanquerna', 'Vivot' and 'Tuono' using the primer combination PruT2/rtPd_S_fR (Hanada et al. 2009) with a size of 603 pb. **(c);** Nested PCR with the cDNA of the three samples analysed with the combination primer PruC2/rtPd_S_fR with a size of 433 pb. **(d)**; NEpHGE (Bošković et al. 1997) of 'Vivot' showing the two different bands (S_{23} and S_{fa}) and of 'Blanquerna' with only one band (S_8). **(e);** cDNA from pollen (3) of 'Blanquerna', 'Vivot' and 'Tuono' using the primer combination Pd_SFB*f*F1/Pd_SFB*f*R1 (Hanada et al. 2009) with a fragment size of around 700 pb.

6.3.5 Expression level of the two different versions of the S_f-RNase

Transcripts of the S_{fa}-RNase gene were present in the style tissue of 'Vivot', whereas our real time PCR showed the absence of S_{fi}-RNase transcripts in the style of the SC cultivar 'Blanquerna', confirming the previous indication that no S_f-RNase activity was found in the style of SC cultivars (Bošković et al., 1997). These results agree with those obtained by rt-PCR and NEpHGE analysis, where a transcript product was detected in the SI cultivar possessing the S_f haplotype but not in the SC one. Thus, the detected expression in the pistil tissue of the S_{fa} gene indicates that this gene is probably functional, contrary to S_{fi} (Fig. 6.4).

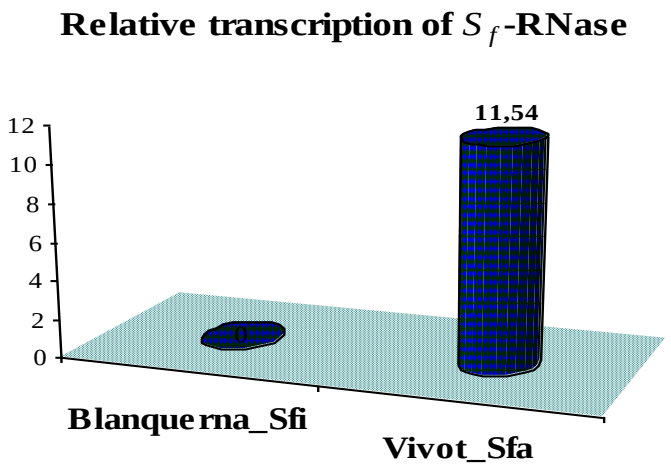


Figure 6.4: Quantitative RNase expression using real-time PCR. Relative transcript level of the pistil tissue of the S_{fa} (Vivot) and S_{fi} (Blanquerna) genes are compared to the housekeeping gene *Actin*

6.4. Discussion

The recent progress of molecular techniques has allowed increasing the reliability of sequencing and fragment size determination in genomic studies. As a consequence, several sequences for the active and inactive S_f-RNase have already been published. For the inactive S_f haplotype (S_{fi}), several sequences with some mutations in the C2 and C5 regions have been reported, thus raising the question if these different sequences for the same allele may indicate different sources for SI or if mistakes could have been taken place in their sequencing (Fig. 6.5).

The S_f-RNase sequences here obtained for 'Vivot' (AB467370) and 'Blanquerna' (AB467371) were identical. A conserved histidine residue in the C2 region, which is essential for the structure and the function of the *S*-RNase, and a phenylalanine residue in place of a leucine in the C5 region were identified in these S_f-RNases when our sequences were compared with those already published. In relation to the SFB_f, no alterations in the amino acid alignment were obtained when comparing with others SFB sequences (Fig. 6.5).

The 'Tuono' S_f-RNase sequence by Bošković et al. (2007) (AM690356) is 98% identical to ours, differing by the presence of an arginine residue instead of a histidine in the C2 region and a leucine residue in place of a phenylalanine in the C5 region. Bošković et al. (2007) suggested that this mutation in the C2 region represented the wild type progenitor allele from which S_f could be derived, thus being the cause of inactivation and loss of the SI function. However, the same research group published later a 'note added in proof' in which they recognised that a missequencing took place in the S_f-RNase sequence. As a consequence, the reasoning on a possible mutation was incorrect, thus invalidating the conclusion about the possible 'wild type' of the S_f haplotype.

On the other hand , the S_f-RNase sequence by Chanuntapipat et al. (2001) from two SC genotypes, selection IRTA-12-2 (AY291117) and 'Lauranne' (AF510414), was 99% identical to ours, confirming the presence of a conserved histidine residue instead of an arginine in the C2 region, as it happens in 'Vivot' and 'Blanquerna'. However, these two sequences show the presence of a leucine residue in place of a phenylalanine in the C5 region, contrary to what happens in 'Vivot' and 'Blanquerna'.

Recently, Hanada et al. (2009) have characterized the S_f-RNase and SFB_f genes from the SC cultivar 'Lauranne' by constructing a genomic fosmid library. The S_f-RNase sequence obtained (AB433984) is 100% identical to the S_f found in 'Blanquerna' and 'Vivot'. These authors also revealed the presence of a histidine residue in place of arginine in the C2 region and of a phenylalanine residue instead of leucine in the C5 region. These results were additionally corroborated by the same research group in the S_f-RNase sequence from two different sources of 'Tuono', one collected from the germplasm bank collections of CITA in Spain and the other from the UCD collection in the USA (Hanada et al. 2009).

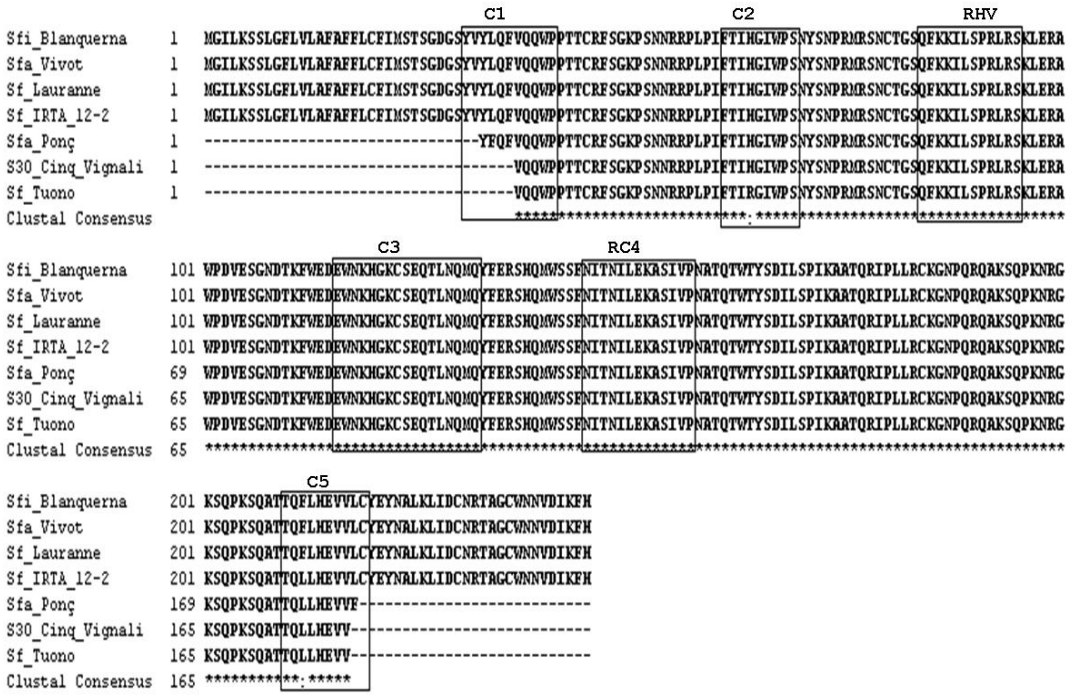


Figure 6.5: The deduced amino acid sequence alignment for the *Prunus amygdalus* Batsch S_{fi} (inactive S_f), S_{fa} (active S_f), S_f and S_{30} RNases. The five conserved regions, C1, C2, C3, RC4 and C5 and the hypervariable region (RHV) reported in the Rosaceous *S*-RNases (Ushijima et al. 1998) are boxed. EMBL/DDJB/GenBank accession numbers are as follow: S_{fi}_Blanquerna (AB467371) (this work), S_{fa}_Vivot (AB467370) (this work), S_f_Lauranne (AB433984) (Hanada et al. 2009), S_f_IRTA-12-2 (AY291117) (Channuntapipat et al. 2003), S_{fa}_Ponç (EU293146) (Kodad et al. 2009), S_{30}_Cinquanta Vignali (CAM84225) (Bošković et al. 2007) and S_f_Tuono (AM690356) (Bošković et al. 2007).

When the sequences obtained by Bošković et al. (2007) for the putative S_{30} allele in the SI cultivars 'Cinquanta Vignali', 'Fra Giulio Grande' and 'Santoro' (CAM84225, CAM84229 and CAM84189 respectively) were compared with our sequences, only one mutation was found, a phenylalanine substitution in place of the reported leucine in the C5 region, thus these sequences are 99% similar to ours.

On the other hand, the S_f-RNase sequence from 'Ponç' (EU293146) (Kodad et al. 2009) also presents a histidine residue in place of arginine in the C2 region. Finally, the presence of a histidine instead of arginine in the C2 region and a phenylalanine in place of leucine in the C5 region has also been noticed by Martínez-García (2009).

As proposed by Hanada et al. (2009), the replacement of the phenylalanine in the C5 region could be due to the fact that of this region is usually used as a primer when cloning the S_f-RNase fragment. As a consequence, sequences already published could be mistakenly reported. However, it is proposed that this replacement does not affect the functionality of the S_f-RNase because the fully functional S_2-RNase of the S_2 haplotype of apricot has a phenylalanine residue at that position (Romero et al., 2004).

The SFB_f sequences obtained in 'Blanquerna' (AB480705) and 'Vivot' (AB480704) are 100% identical to those reported by Hanada et al. (2009) (AB361036), Bošković et al. (2007) (AM711126) and Kodad et al. (2009) (EU310402) and also 100% identical to the SFB_{30} sequences reported by Bošković et al. (2007) (AM711127 and AM711125; Fig. 6.6).

The construction of a fosmid library in 'Vivot' and 'Blanquerna' has allowed to confirm that the alignment of their 5'-flanking regions were also the same. The 1900 bp obtained from the active and inactive S_f upstream sequences, including the TATA-Box and IB-like motif, revealed that there is not any *cis*-element for pistil-expression that differs in the two cultivars. Additionally, the two sequences were aligned with the sweet cherry S_6-RNase (Ishizaka et al. 2003) and the apricot S_4-RNase sequences (Romero et al. 2004). These results confirmed that the *S*-locus genomic structure in almond is very similar to those reported in other *Prunus* species sharing common structures and features (Fig. 6.7).

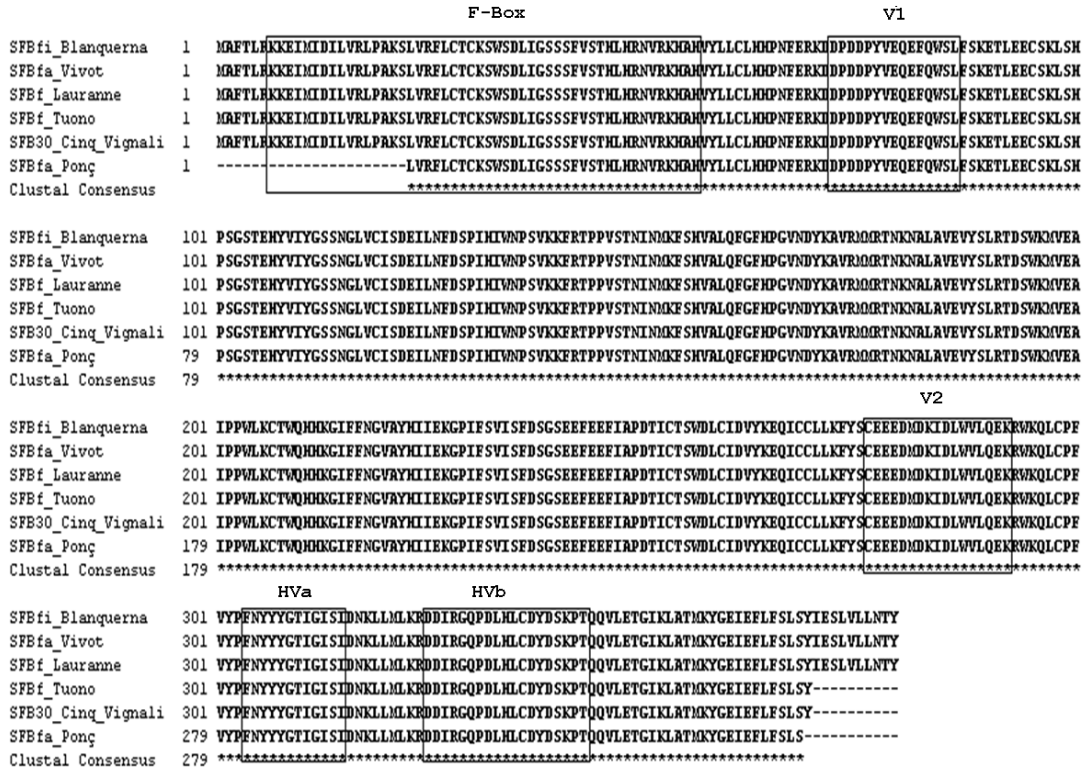


Figure 6.6: The deduced amino acid sequence alignment for the *Prunus amygdalus* Batsch SFB_{fi} (inactive S_f), SFB_{fa} (active S_f), SFB_f and SFB_{30}. The position of the F-box motif, V1, V2, HVa and HVb (Ikeda et al. 2004) are boxed. EMBL/DDJB/GenBank accession numbers are as follow: SFB_{fi}_Blanquerna (AB480705) (this work), SFB_{fa}_Vivot (AB480704) (this work), SFB_f_Lauranne (AB361036) (Hanada et al. 2009), SFB_{fa}_Ponç (EU310402) (Kodad et al. 2009), SFB_{30}_Cinquanta Vignali (AM711125) (Bošković et al. 2007) and SFB_f_Tuono (AM711126) (Bošković et al. 2007).

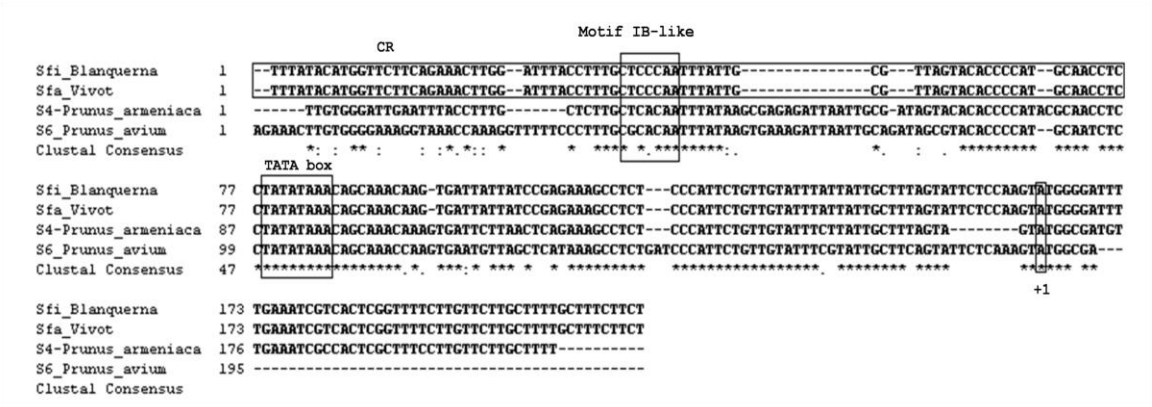


Figure 6.7: Alignment of the 5´-flanking regions from sweet cherry (*Prunus avium* L.) S_6-RNase (Ishizaka et al. 2003; AB075597), apricot (*Prunus armeniaca* L.) S_4-SRNase (Romero et al. 2004; AY587564) and almond (*Prunus amygdalus* Batsch) S_{fi}-RNase (AB467371) and S_{fa}-RNase (AB467370). Nucleotides corresponding to the conserved regions (CR), putative TATA box and the motif IB-like are labelled and boxed. The first nucleotide of the translational start codon is numbered as (+1).

All the approaches undertaken to ascertain the nature and/or location of a possible factor or gene affecting the differential expression of the S_f allele in almond have not shown any difference in the *S* locus between the genotypes showing this different expression. The construction of a fosmid library has also shown that the transcription factor is identical in the two genotypes, suggesting that an external factor may take place somewhere else in the almond genome, but not in the nearest region to the *S*-locus.

These results from both a SI and a SC cultivar provide evidence that the understanding of the SC/SI mechanisms operating in the GSI systems are not yet explained, as confirmed by the double expression of the S_f haplotype. Thus, further approaches are needed in order to analyse the precise nature of this differential expression of the almond S_f haplotype. The following chapters will deal with different possible approaches to this controversial subject.

7. IDENTIFICATION OF QTLs ASSOCIATED WITH SELF-COMPATIBILITY IN ALMOND

7. IDENTIFICATION OF QTLs ASSOCIATED WITH SELF-COMPATIBILITY IN ALMOND

7.1 Introduction

Although SC in almond has been considered a qualitative trait, controlled by a single multi-allelic locus (Socias i Company, 1984), called the *S*-locus (de Nettancourt 1977), a quantitative approach was suggested in order to find out any possible action of other loci outside the *S*-locus.

SSR (simple-sequence repeat or microsatellite) markers have become a very useful tool for constructing linkage maps and locating genes controlling phenotypic characters. Consequently, the development of markers associated with a trait may improve the speed and precision of breeding programmes with the aim of selecting for this trait by marker-assisted selection. More than 10 molecular genetic maps have been constructed for different *Prunus* species (http://www.bioinfo.wsu.edu/gdr). Among these maps, that obtained by Joobeur et al. (1998) from the cross 'Texas' almond × 'Earlygold' peach [*P. persica* (L.) Batsch] ('T × E') is considered the reference *Prunus* map. A total of 827 markers covering a total distance of 524 cM have been placed in this map (Shulaev et al., 2008). Additionally, the high level of synteny between the genome of the different *Prunus* species (Arús et al., 2005), contributed to the identification of 28 major genes affecting agronomic traits and more than 20 QTLs (Dirlewanger et al., 2004).

Among the 28 major genes located in the *Prunus* genome, the *S* gene was mapped for the first time using an almond progeny from the cross 'Ferragnès' × 'Tuono' and found to be located in the distal part of group 6 (Ballester et al. 1998). The same position was also determined with SSR markers by Sánchez-Pérez et al. (2007) in the progeny from the cross R1000 x 'Desmayo Largueta'.

Despite the presence of two phenotypic expressions of the S_f allele, all SC genotypes so far identified possessed this allele in its inactive form, S_{fi}. However, as

shown in Chapter 5, some genotypes without the S_f allele have been recently identified as highly SC after pollen tube growth and fruit set trials, thus suggesting that additional external factors or other genes not related to the *S* locus, possibly modifier genes located outside the *S* complex, may be responsible for the genetic control of SC/SI in almond.

As a consequence, our goal was to better understand the SI complex in almond and to know which additional external factors may be able to modulate the SI mechanisms. A possible strategy for ascertaining the genes involved in the SI reaction was considered to be the identification of QTLs through the development of linkage maps using SSRs, which would facilitate the selection for this trait. These kinds of markers are, up to date, the preferred tool for genetic mapping, cultivar identification and marker assisted selection in higher plants. SSRs are easily transferable to a large number of *Prunus* species (Wünsch, 2009), extremely polymorphic (Gupta et al., 1996), codominant, and abundant in almond, which is one of the most polymorphic cultivated species within the *Prunus* genus (Byrne, 1990; Fernández i Martí et al., 2009b; Socias i Company and Felipe, 1992).

Taking into account the unexpected behaviour in the SC/SI expression of the progeny 'Vivot' × 'Blanquerna' ('V × B'), the objective of this chapter was the construction of a genetic linkage map of the 'V × B' progeny to test the hypothesis that genes external to the *S* locus and segregating in this progeny would confer SC to individuals predicted to be SI. With this aim a QTL approach has been undertaken using a combination of SSRs and *S*-locus markers.

7.2 Methods

7.2.1 Screening for self-compatibility

As described in previous chapters, SI/SC was assayed in the progeny by microscopic observation of pollen tube growth during three consecutive years, because pollen tube growth has been often associated with fruit setting following artificial pollinations, giving concordant results (Ben Njima and Socias i Company, 1995). Additionally, the *S*-genotypes were determined by NEpHGE and *S*-allele-specific PCR.

The criteria adopted for phenotypic assignment was also applied for the QTL analysis. Hence, the individuals showing the arrest of pollen tube growth were considered as 0, whereas the seedlings with pollen tubes reaching the style base at least in eight out of 12 pistils were classified as 1.

7.2.2 DNA marker genotyping, genetic mapping and QTL analysis

A total of 102 SSR markers previously described in other *Prunus* species (Table 7.1) were tested in the 'Vivot' and 'Blanquerna' almond progeny to identify polymorphic markers between the two parents, pursuing a good coverage of the *Prunus* bin mapping 'T × E' (Howad et al., 2005). From those, 52 SSRs were selected because of their polymorphism in one or both parents and distribution over the 'T × E' *Prunus* reference map (Table 7.2). Additionally, several specific primers for the *S*-locus (PaConsIF/EMPC5R; S_fF/S_fR; S_{23}F/S_{23}R; S_8F/S_8R) were included in the analysis (Channuntapipat et al. 2003; Sonneveld et al. 2003; Sutherland et al. 2004).

PCR reactions were performed in a 10 μL volume and the reaction mixture contained 1× PCR buffer (Invitrogen, Barcelona, Spain), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 μM of each primer, one unit of Taq DNA Polymerase (Invitrogen) and 20 ng of genomic DNA. The cycling parameters consisted in a denaturation during 1 min at 94 ºC, 35 cycles of 15 s at 94 ºC, 15 s for the corresponding annealing temperatures and 1 min at 72 ºC, followed by a final extension of 2 min at 72 ºC .

The PCR reactions were carried out in a 96-well block Thermal cycler (Applied Biosystems, Madrid, Spain). PCR products were detected using an ABI PRISM 3130 Genetic Analyzer and GeneMapper v3.7 analysis software (Applied Biosystems). For capillary electrophoresis detection, forward SSR primers were labelled with 5'-fluorescence dyes PET, NED, VIC and 6-FAM and the size standard used in the sequencer was Gene Scan ™ 500 Liz ® (Applied Biosystems).

The maps of the parents were constructed using MapMaker/EXP 3.0 (Lander et al., 1987). Linkage groups were obtained at a logarithm of –odd (LOD) score threshold

Table 7.1: SSRs used in the identification of QTLs in the almond cross 'Vivot' × 'Blanquerna'.

Species of origin	SSR name	Reference	N° of SSRs tested	N° of SSRs amplified	N° of SSRs mapped	N° of loci mapped	% of total SSRs placed in the 'V × B' map
Peach	BPPCT	Dirlewanger et al., 2002	21	19	14	15	27
Peach	CPPCT	Aranzana et al., 2002	28	27	13	13	25
Japanese Plum	CPSCT	Mnejja et al., 2004	6	6	6	6	12
Almond	EPDCU	Howad et al., 2005	6	6	2	2	4
Peach	EPPCU	Howad et al., 2005	9	9	1	1	2
Peach	PCHGMS/Ma0	Sosinski et al., 2000; Yamamoto et al., 2002	5	5	2	2	4
Peach	UDP	Cipriani et al., 1999; Testolin et al., 2000	17	17	9	9	17
Cherry	Others	Cantini et al., 2001; Downey and Iezzoni, 2000	10	8	5	5	9
-	Total	-	102	97	52	53	100

of 4.0 followed by the "order" or "compare" commands. The "error detection" command was used to identify possible scoring errors. Kosambi's mapping function was used to translate recombination frequencies into map distances (Kosambi, 1944). The resulting linkage maps were drawn using FitMaps v. 1.0 (Graziano and Arús, 2002). Additionally, Joinmap 3.0 (Van Ooijen and Voorrips, 2001) was used in this study to construct an integrated map using the 53 loci. Markers were grouped using a minimum LOD ≥ 3.

Quantitative trait loci effects were detected using the simple interval mapping procedure of MapQTL 4.0 (Van Ooijen et al., 2002). QTLs with a LOD score equal or higher than 2.0 were declared significant.

7.3 Results

7.3.1 Polymorphism, segregation and transferability of SSR markers

The selected primer pairs amplified fragments of the 52 SSRs in the range of the expected size for each microsatellite. All SSRs, except five in 'Vivot' and eight in 'Blanquerna', showed polymorphism in both parents. The genotypes of these 52 SSRs were compatible with the origin of all individuals studied in the progeny of 'Vivot' × 'Blanquerna'. Most markers segregated according to the expected ratios with significant χ^2 values (Table 7.2).

The BPPCT and CPPCT series of primers developed in peach (Aranzana et al., 2002; Dirlewanger et al., 2002) showed the highest percentage of usefulness in the 'V × B' map (0.27 and 0.25, respectively), followed by the UDP series (Cipriani et al., 1999; Testolin et al., 2000) with 0.17, as a result of their high polymorphism and their distribution over the whole 'T × E' map. The CPSCT markers developed in Japanese plum (*P. salicina* Lindl.) showed a high polymorphism and transferability to almond, with six of the SSRs tested placed in the map. For the 10 SSRs isolated from cherry (*P. avium* L.), eight were amplified, although finally only five (PceGa25, PceGa34, Pms02, Pms40 and Ps12e2) were placed in the map.

Table 7.2: SSRs and *S*-locus markers used in the identification of QTLs in the almond cross 'Vivot' × 'Blanquerna' with their position in each linkage group of the map, the expected segregation rate in the progeny, the observed segregation and probability

Marker	LG	Segregation Type	Segregation Observed	χ^2	α
UDP96-018	1	1:2:1	19:33:19	0,35	0,84
CPPCT026	1	1:1:1:1	21: 20: 17: 15	0,68	0,88
EPPCU1090	1	1:1:1:1	17: 22: 14: 14	2,55	0,46
BPPCT020_A	1	1:1:1:1	22: 23: 18: 14	2,63	0,45
BPPCT020_B	1	1:1:1:1	16: 23: 12: 26	6,37	0,05
CPPCT053	1	1:1:1:1	15: 24: 16: 18	2,67	0,44
BPPCT011	1	1:1[z]	33:39	0,50	0,48
CPPCT042	1	1:1[z]	34:38	0,22	0,64
BPPCT028	1	1:1:1:1	17: 16: 24: 19	2,00	0,57
CPPCT044	2	1:1[z]	40:35	0,33	0,56
UDP98-025	2	1:1[y]	40:37	0,17	0,73
BPPCT002	2	1:1:1:1	9: 21: 18: 23	6,46	0,09
UDP96-013	2	1:1:1:1	25: 22: 14: 15	4,52	0,21
CPSCT021	2	1:1:1:1	12: 18: 21: 25	4,74	0,19
PceGa34	2	1:1:1:1	12: 23: 21: 15	4,43	0,22
BPPCT007	3	1:1:1:1	27: 23: 15: 12	7,52	0,06
BPPCT039	3	1:1:1:1	22: 19: 17: 17	0,89	0,82
EPDCU3083	3	1:1[z]	34:32	0,06	0,80
UDP96-008	3	1:1:1:1	22: 21: 17: 16	1,37	0,71
BPPCT010	4	1:1:1:1	11: 22: 25: 17	6,00	0,11
CPPCT005	4	1:1:1:1	15: 19: 27: 14	5,60	0,13
PMS40	4	1:1:1:1	11: 22: 24: 16	5,74	0,12
UDP96-003	4	1:1:1:1	18: 23: 21: 14	2,42	0,50
CPPCT046	4	1:1:1:1	21: 13: 19: 23	2,95	0,40
PCHGMS55	4	1:1:1:1	10: 17: 21: 18	3,94	0,27
PS12e2	4	1:1:1:1	22: 21: 16: 17	1,37	0,71
CPPCT040	5	1:1:1:1	19: 23: 12: 20	3,50	0,32
UDP97-401	5	1:1:1:1	9: 20: 24: 21	6,97	0,07
BPPCT017	5	1:2:1	14:35:28	5,72	0,06
BPPCT037	5	1:1[z]	45:32	2,19	0,14
CPSCT006	5	1:1[y]	37:36	0,01	0,90
CPSCT022	5	1:1[y]	40:34	0,49	0,49
PceGA25	5	1:1[y]	33:35	0,06	0,80
BPPCT014	5	1:1[y]	38:31	0,71	0,40
CPPCT008	6	1:1[y]	34:37	0,13	0,72

BPPCT008	6	1:1:1:1	10: 27: 22: 18	8,00	0,05
CPSCT012	6	1:1:1:1	12: 29: 23: 12	9,69	0,02
BPPCT025	6	1:1:1:1	8: 26: 26: 14	11,50	0,01
Ma40	6	1:1[y]	37:34	0,12	0,72
UDP98-412	6	1:1[y]	36:40	0,21	0,64
CPPCT021	6	1:1[z]	37:39	0,05	0,81
S-locus	6	1:1	33:42	1,08	0,30
CPPCT022	7	1:1:1:1	18: 23: 17: 18	1,16	0,76
CPPCT039	7	1:1:1:1	14: 13: 22: 17	2,97	0,39
CPSCT004	7	1:2:1	16:39:19	0,46	0,79
UDP98-408	7	1:1:1:1	14: 15: 12: 17	0,90	0,82
CPPCT033	7	1:1:1:1	22: 21: 14: 19	2,00	0,57
PMS02	7	1:1:1:1	20: 20: 19: 11	3,26	0,35
EPDCU3392	7	1:1:1:1	17: 30: 11: 16	10,60	0,01
CPSCT018	8	1:1[y]	31:44	2,25	0,13
BPPCT006	8	1:1[y]	43:33	1,30	0,25
CPPCT006	8	1:1:1:1	18: 25: 14: 17	3,50	0,32
UDP98-409	8	1:1:1:1	10: 25: 17: 13	7,80	0,05

[z] Locus heterozygous in 'Blanquerna'

[y] Locus heterozygous in 'Vivot'

All markers amplified a single locus, with the exception of BPPCT020, which amplified two loci, one in G1 and the other in G6. The total number of loci mapped and scored in the 'V × B' map was 53. The heterozygosity observed was slightly higher in 'Vivot' (0.9) than in 'Blanquerna' (0.85).

7.3.2 Map construction

Segregation data for the SSRs obtained from the progeny of 77 seedlings were used to generate the parental maps. A total of 47 loci for 'Vivot' and 44 for 'Blanquerna' were utilised for constructing their respective maps covering the eight linkage groups. More than 66% of the loci were fully informative (1:1:1:1 segregation), 6% segregated as 1:2:1, whereas for the other markers (26%) one of the parent was monomorphic and segregated 1:1. All SSRs were positioned in the eight linkage groups (G1 to G8) of the common map with an average of 6.4 markers per linkage group, ranging from 4 (G3 and G8) to 8 (G5 and G6). The maps of the two parents were connected using 31 markers as anchors (Fig. 7.1). Similarly, a map was constructed for

the progeny, spanning 451 cM in eight linkage groups (Fig. 7.2), similar to other almond maps published (Arús et al., 2005), covering 86.4% of the length of the *Prunus* reference map (Dirlewanger et al., 2004).

All markers used in this study were placed in accordance with the *Prunus* 'T × E' reference map, with the exception of locus CPPCT021 in the G6, which was located 10 cM before locus UDP98-412 in the 'V × B' map. Slight order differences have also been found when comparing other *Prunus* crosses (Arús et al., 2005; Clarke et al., 2009; Verde et al., 2005), a possible indication of recombination between co-linear markers. The incompatibility *S* locus was mapped in G6 (41 cM) and was flanked by the SSR markers Ma40 (at 38 cM) and CPPCT021 (at 45 cM), according to the other *Prunus* maps (Arús et al., 1999; Clarke et al., 2009; Dirlewanger et al., 2004; Sánchez-Pérez et al., 2007).

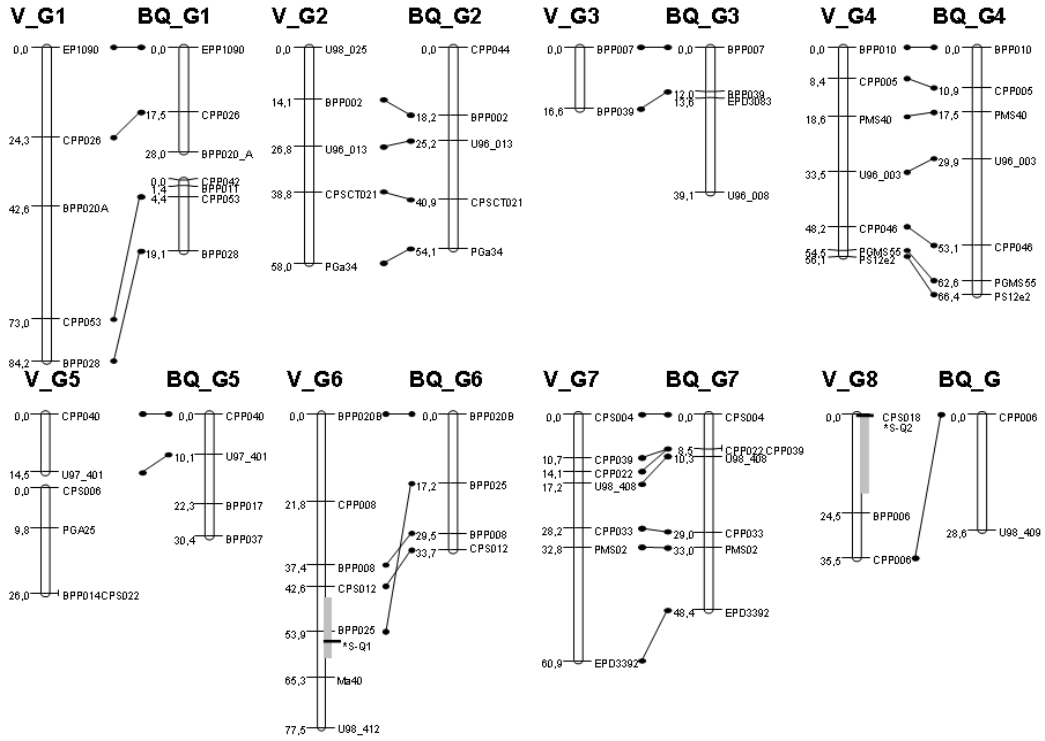


Fig. 7.1: Anchor markers of the molecular linkage maps of 'Vivot' and 'Blanquerna' constructed using MAPCHART V. 2.1 (Voorrips 2002). Two QTLs associated with SC in 'Vivot' are shown in G6 and G8. QTL positions are indicated as vertical bars, the position of the black horizontal line indicates the maximum LOD score. QTL length is plotted from LOD±1 with respect to the maximum LOD score.

7.3.3 Identification of QTLs for self-compatibility

Once the linkage map was constructed, QTLs associated with SC were identified by interval mapping when their LOD score was higher than 2. In the map of the male parent 'Blanquerna', no QTLs were detected in any of the eight linkage groups, however, in the case of the female parent 'Vivot', two significant QTLs, *S*-Q_1 and *S*-Q_2, were identified, the first located in the middle of the G6 and the second at the beginning of G8. The percentage of variability explained by *S*-Q_1 was 41.6 and by *S*-Q_2 was 23.8. *S*-Q_1 was placed around locus BPPCT025 in G6 and showed the highest LOD value for this area (6.20). *S*-Q_2 was located in G8 at position 0.0 cM, with a LOD value of 4.49 and close to CPSCT018, the first SSR marker both in the 'T × E' and the 'V × B' maps (Fig. 7.2).

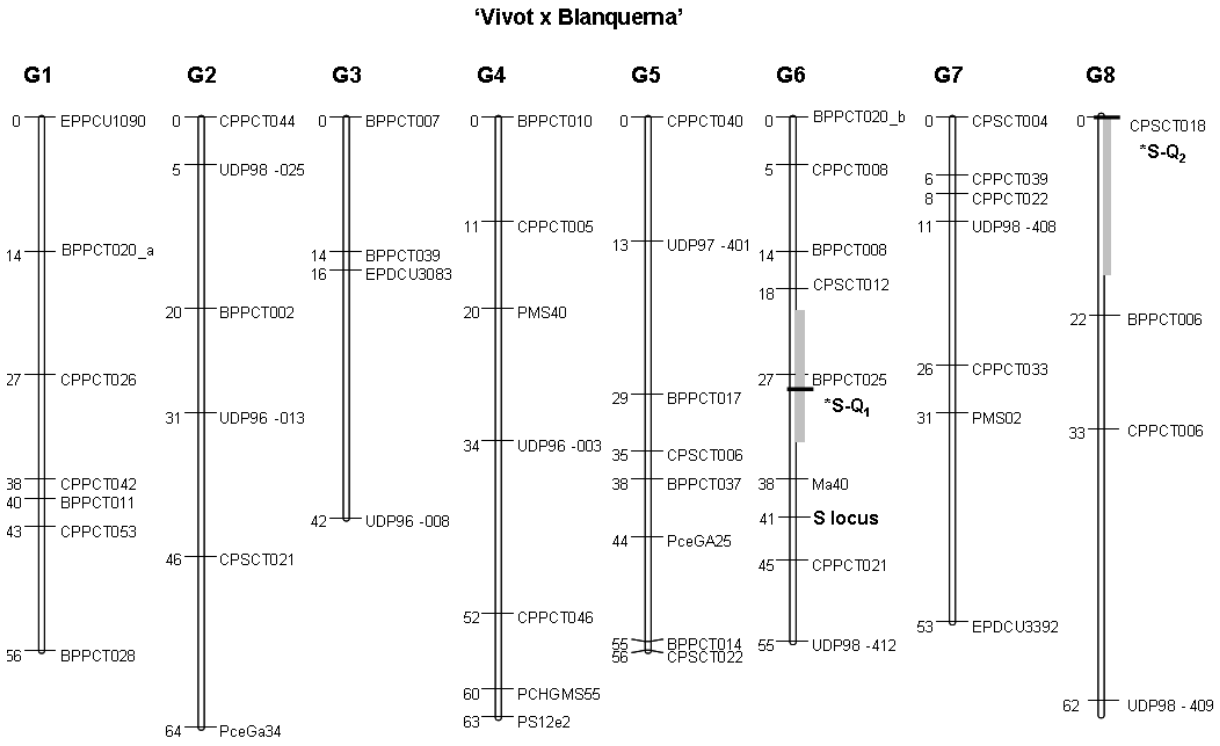


Fig. 7.2: Integrated map of 'Vivot' × 'Blanquerna' containing 53 loci, including the two QTLs associated with SC identified on G6 and G8. The map was constructed using JOINMAP V. 3.0 software

These two QTLs were examined in the progeny studied in order to check their effect (Table 7.3). For the marker CPSCT018, the alleles found in the parents were 159/170 for 'Vivot' and 159/159 for 'Blanquerna, thus the 159 allele coming from 'Vivot' was found in 17 out of the 19 SC genotypes. Concerning the marker BPPCT025, the alleles identified in 'Vivot' were 180/189 and 170/180 for 'Blanquerna'. The allele 180 coming also from 'Vivot' was present in 18 plants out of the 19 SC genotypes. As a result of their combination, 16 of the 19 SC plants in our

Table 7.3: Phenotype, *S*-allele genotype and SSR allele segregation in the parents and in the SC progeny of 'Vivot' and 'Blanquerna'

Genotype	Phenotype	*S*-locus	CPSCT018	BPPCT025
Vivot ♀	SI	$S_{23}S_{fa}$	159/170	180/189
Blanquerna ♂	SC	S_8S_{fi}	159/159	180/170
P1-11	SC	S_8S_{23}	159/159	180/189
P1-26	SC	S_8S_{23}	159/159	180/189
P1-28	SC	S_8S_{fa}	159/159	180/189
P1-37	SC	S_8S_{23}	159/159	180/189
P1-48	SC	S_8S_{23}	159/159	180/189
P1-49	SC	S_8S_{23}	159/159	180/189
P1-56	SC	S_8S_{23}	159/159	180/189
P1-57	SC	S_8S_{23}	159/159	180/189
P2-10	SC	S_8S_{23}	159/159	180/189
P2-14	SC	S_8S_{23}	159/159	180/189
P2-5	SC	S_8S_{23}	159/159	180/189
P2-9	SC	S_8S_{23}	159/159	180/189
P1-13	SC	S_8S_{fa}	159/159	170/180
P1-18	SC	S_8S_{fa}	159/159	170/180
P1-51	SC	S_8S_{fa}	159/159	170/180
P1-66	SC	S_8S_{fa}	159/159	170/180
P1-20	SC	S_8S_{23}	159/159	170/189
P1-25	SC	S_8S_{23}	159/170	180/189
P1-45	SC	S_8S_{23}	159/170	180/189

progeny showed the presence of the two alleles, with a probability of nearly 85% of identifying SC genotypes.

7.4 Discussion

During the last two decades many efforts have been devoted to elucidate the molecular, biochemical and cellular processes involved in the SI reactions. Because of the complexity of the SI locus, a wide array of changes has been described in the GSI system, the one found in almond. These changes include mutations in the *S*-alleles themselves, in genes linked to the *S*-locus, in genes involved in the downstream rejection pathway and in unlinked genes modifying the expression of the *S*-alleles (Good-Avila et al., 2008). Although *S*-alleles have been considered to settle a clear specificity, other genes seem to be required for the SI/SC response.

In the Solanaceae, another family with a GSI system, several external factors unlinked to the *S*-locus and responsible for breakdown of SI have been already described (Cruz-Garcia et al., 2003; Tsukamoto et al., 2003). In *Nicotiana alata* Link & Otto the products of several unlinked genes were involved in the formation of a multi-protein complex, which represents the *S*-RNase in its active form (Cruz Garcia et al., 2003). In *Petunia axillaris* (Lam.) B.S.P. a modifier locus unlinked to the *S*-locus has been described to suppress the expression of an *S*-RNase gene (Tsukamoto et al. 2003). In addition, two non-haplotype-specific pistil factors, HT-B and 120K, required for pollen rejection and located outside the *S* locus, were identified in *Nicotiana* and *Petunia* (Goldraij et al., 2006; Hua and Kao, 2006). It seems that these factors are required for the SI reaction in pistils because transformants suppressing either the 120K or the HT-B gene expression do not reject the *S*-specific pollen (McClure et al., 1999).

Within the Rosaceae family, an external factor as possible cause of breakdown of SI has been identified in three *Prunus* and one *Pyrus* species. Wünsch and Hormaza (2004) reported that the pollen *S* component of the SC sweet cherry 'Cristobalina' was affected by a factor unlinked to the *S* locus. Similarly, the presence of a mutation in a modifier locus has been reported to affect the function of the *S*-pollen factor in apricot (*P. armeniaca* L., Vilanova et al., 2006). In this thesis, we have also proposed the

presence of an external factor affecting the SI mechanism in almond. The presence of non-*S*-specific factors affecting the SI reaction have also recently been observed in European pear (*Pyrus communis* L.; Moriya et al., 2009).

Several models and possible candidate genes modulating the SI systems have been proposed in the recent years in *Petunia*, *Antirrhinum* and *Arabidopsis*. However, no models or genes have been yet proposed in any Rosaceae species. This may be due, on one hand, to that the ability of breeders and researchers is somehow limited because working with fruit tree species implies more time, more space and more speculation than working with annual species (Socias i Company, 1998b). On the other hand, the molecular recognition mechanism of the *S*-RNase in Solanaceae and Plantaginaceae would be different to that of Rosaceae (Sonneveld et al., 2005). As a consequence, so far only one approach has focused on the search of unlinked genes related to SI in *Prunus*, based on the bulk segregant analysis in a progeny of the SC sweet cherry 'Cristobalina' by using SSR and AFLP markers. Nevertheless, it was concluded that no relevant polymorphism between both bulks was found (Cachi and Wünsch, 2009).

In our case, however, we have reached the goal of locating the possible linkage groups where the modifier gene/s might be placed, after developing a genetic map using SSR markers and undertaking a QTL mapping analysis for this purpose. The SSRs used were developed in other species but were transportable across *Prunus* species and allowed to cover the 'V × B' map with an average of 6.4 SSR/linkage group. In specific areas of the genome it has been more difficult to find polymorphic markers in 'Blanquerna' than in 'Vivot', resulting in a less saturated map of 'Blanquerna'. This fact may be due to the SC condition of 'Blanquerna', which could have some chromosomal regions identical by descent.

Two QTLs were found to be related to the breakdown of SI in almond, BPPCT025, placed in the middle of G6, and CPSCT018 placed at the beginning of G8. In fact, these two QTLs appeared to affect only the female parent 'Vivot', which possesses the SI version of the S_f-haplotype. Thus, we may suggest that SI in almond is controlled by a major gene (*S*-locus) and two QTLs (*S*-Q_1 and *S*-Q_2) only in the cultivars possessing the S_{fa} allele, whereas in the plants with the inactive S_f allele, SI is only controlled by the *S*-locus. In the case of progenies derived from cultivars

possessing the 'inactive' S_f allele, the specific-S_f primer pair developed in almond by Channuntapipat et al. (2001) would be strongly recommended for an efficient selection for SC. However, in the case of progenies showing the presence of the active S_f, the combined use of the BPPCT025 and CPSCT018 SSR markers would be very effective when discriminating between SC and SI seedlings.

These results pose a question about the monogenic nature of SI/SC in almond, so far considered as a qualitative trait (Socias i Company, 1984). It seems that in many species, including Rosaceae, SI might act as a quantitative trait rather than as a qualitative one, as suggested by Good-Avila et al. (2008). Thus, the knowledge provided by molecular techniques has allowed confirming the hypothesis proposed by de Nettancourt (1977) about the possible interaction of multiple unlinked modifiers genes involved in the SI complex, defining this phenomenon as pseudo-self-compatibility (PSC).

Socias i Company (1990) suggested that almond is a SI species with a genetic background of PSC as indicated by the small self set observed in some cultivars, independently of the fact that these sets may be caused by foreign pollen contamination (Martínez-García, 2009). Over this background, only one S_f allele could break the SI system, but probably interacting with this background of PSC as shown by the effect of these two QTLs. Since the first work by Almeida (1945), all range of fruit sets has been described when self-pollinating almond cultivars (Socias i Company, 1990). From the agronomical point of view fruit set has been considered the main evaluation criterion for SC selection in almond (Kodad and Socias i Company, 2008a). The effect of the S_f allele and the two QTLs may explain this wide range of fruit sets, independently of the changing year effect on these sets (Socias i Company et al., 2005).

The pollen parent 'Blanquerna' only transmitted one of the two possible *S* alleles to the progeny, being S_8 the only allele able to grow through 'Vivot' pistils. Thus, the *S*-genotype distribution in the progeny (S_8S_{fa} or S_8S_{23}) would have been SI, because none of the progeny possesses the S_{fi} allele (inactive S_f), presumably conferring SC. However, an SC phenotype was observed in 19 of 77 plants, including 14 S_8S_{23} and five S_8S_{fa} and corresponding to nearly 25%. Our results suggest that both QTLs are not only affecting the plants with the S_f allele (the five S_8S_{fa} plants), but also to the plants of genotype

S_8S_{23}, not possessing the S_f allele. Thus, the stylar modifier factors would act in all genotypes, independently if they carry or not the S_f allele. As a consequence, SC in this progeny would be controlled by two recessive complementary loci (S-Q_1 and S-Q_2), with dependence to the *S*-locus. Thus, the allele combination CPSCT018-159 and BPPCT025-180 give a probability of 84.2% of identifying SC genotypes.

Despite the effect of these two QTLs, the way of recognition of the S_{fi} allele by S_{fa}, thus hindering the growth of the S_{fi} pollen in pistils possessing the S_{fa} allele, still remains unclear. At the origin, both versions of the S_f haplotype must likely come from a common ancestor, having evolved differently depending on the environmental and genetic characteristics of each separated population, leading to the accumulation of different modifier genes, as the two QTLs described in this work may be. Socias i Company et al. (1999) also described the effect of a major gene affected by modifier genes for blooming time in almond, suggesting that these modifier genes were presumably different depending on the geographical origin of the genotypes showing their effect.

Until recently, the movement of almond genotypes from one region to the other was limited. This fact, together with the traditional growing management in most almond growing regions, the consideration as a marginal fruit tree and the SI present in most cultivars, has produced a very polymorphic species (Fernández i Martí et al., 2009b; Socias i Company and Felipe, 1992). However, the advances in almond breeding, including as parents genotypes from different origins (Socias i Company et al., 2010), but also with a repeated utilization of related parents in the breeding programmes (Socias i Company, 2002), may produce the interaction of genes so far isolated and the apparition of inbreeding depression (Alonso and Socias i Company, 2007). Although the two forms of the S_f allele may coincide in some Mediterranean regions, their interaction could not be noticed until the molecular approaches have been applied to their study. In addition, the incompatibility between different versions of the S_f haplotype may be beneficial from the evolutionary point of view, but harmful in the hybrid progeny for the commercial production of almond.

As a result of this approach, two additional loci located outside the *S*-locus as other possible genes involved in the SI mechanism are here described for the first time in the Rosaceae family.

8. DNA METHYLATION AS A POSSIBLE INACTIVATION OF THE S_f-RNASE

8. DNA METHYLATION AS A POSSIBLE INACTIVATION OF THE S_f-RNASE

8.1 Introduction

Once clearly established that the almond S_f haplotype shows two different phenotypic expressions, one active (S_{fa}) and the other inactive (S_{fi}), several approaches have been undertaken in order to find out any modification and/or INDELS (insertion and deletion) in their respective sequences, although both versions have shown a comprehensive genetic identity. The results from the genomic library obtained in this study have not only confirmed that Vivot' and 'Blanquerna' were completely identical in the coding region (C1 to C5), but also that the alignment of their 5'-flanking regions were also the same. The TATA-Box and IB-like motif found in the active and inactive S_f haplotypes revealed that there was not any *cis*-element for pistil-expression that differed in the two cultivars.

On the other hand, as reported in the previous chapter, other loci could be affecting the breakdown of self-compatibility in almond. However, in spite of having located two QTLs related to this complex, other factors or mechanisms could be involved in the turnover of the expression of the S_f haplotype, either from SC to SI or the opposite, from SI to SC. The direction of this turnover is yet unclear because up to date is unknown which version is the original one.

The absence of any nucleotide sequence difference at the S_f locus indicates that a possible mutation might be caused by an epigenetic change in this region of the genome. An essential property of eukaryotic cells is the ability to establish heritable patterns of gene silencing without alteration in the DNA sequence. This mechanism is known as 'epigenetic', based on the study of changes in phenotype or gene expression caused by mechanisms other than changes in the DNA sequence (Jeanish and Bird, 2003). Thus, the same gene sequence can be either well-expressed or transcriptionally silent depending on whether it lies in euchromatin or heterochromatin (Bender, 2004).

The epigenetic process involves DNA methylation, which is a major modification of eukaryotic genomes that affect gene expression. Epigenetic research uses a wide range of molecular biological techniques to deepen our understanding of epigenetic phenomena, including fluorescent in situ hybridization and methylation-sensitive PCR sequencing.

Heterochromatin domains are determined in part by methylation of cytosines at the 5 position of the pyrimidine ring (5-Me-C). This modification is catalyzed by cytosine methyltransferase enzymes using S-adenosyl methionine (AdoMet) as the methyl group donor, which is added to cytosines in appropriate regions of the genome after DNA replication (Fig. 8.1). The modified DNA strand can be amplified using Polymerase Chain Reaction (PCR) and either sequenced directly, or cloned and sequenced to give methylation data from single DNA molecules.

The 5-Me-C ring was reported for the first time in 1925 by Johnson and Coghill from DNA tubercle bacilli. Later, Holliday and Pugh (1975) were the first to propose that methylation in mammalian DNA might have an important role in the regulation of gene expression. Their theories provided a model for epigenetic inheritance of a given pattern of DNA methylation and proposed a role for DNA methylation in the specific control of gene expression in given cell types. It was also proposed to account for those cases, such as X chromosome inactivation in female eutherian mammals, where only one of the two homologous genes in a diploid cell is active, whilst the other is inactive (Riggs, 1975).

More recently, the role of aberrant DNA methylation in several diseases has been the focus of much scientific interest. It has been shown that methylation is a common event in the progression of cancer (Herman et al., 1994). In particular, aberrant methylation in many cases has been associated with the loss of expression of the so-called tumour suppressor gene. To date, methylation has been associated with over a hundred genes in cancer. In addition, methylation of specific genes in cancerous cell may provide excellent early markers for cancer diagnosis (Millar et al., 1999).

In plants, significant progress has been made in understanding DNA methylation, specifically in the model plant, *Arabidopsis thaliana* (Luo and Preuss, 2003). However,

up to date, DNA methylation modifications have only been reported in a few plant species.

As mentioned above, DNA methylation has been implicated in the silencing of transposable elements and genes, affecting thus its functions in gene silencing and transcription (Fig. 8.2). The first relationship between gene transcription and DNA methylation was explored by methylating DNA in vitro, transfecting it into mouse cells, and assaying expression by RNA levels or activity of the enzyme encoded by the gene (Kruczek and Doerfler, 1983). Further studies showed that methylated genes were not expressed, and in vitro methylation of specific sites indicated that methylation in the *5'* portion of gene or *3'* portion could be inhibitory to gene expression (Busslinger et al., 1983).

Although methylation can occur in the *cis* or *trans* regions (Zilberman and Henikoff, 2005), it has been noticed that in many genes methylation in the *5'* promoter region has the most profound effect on its expression (Siegfried et al., 1999). This has been later confirmed by Soppe et al. (2000) when reporting that inhibition of gene expression by *5'* methylation was more obvious and constant than in the *3'* flanking sequences.

Cytosine DNA methylation, the chemical modification of cytosine bases with a methyl group, has long been known in plant genomic DNA sequences. Most DNA methylation in plants and animals occurs on the CG dinucleotide context (Bird et al., 1987). However, the high amount of 5-methylcytosine found in some plant species has suggested that methylation is not restricted to the CG sequence context and has led to the discovery that cytosine is also methylated in the CNG group, where N is any nucleotide (Gruenbaum et al., 1981).

The methods most commonly used for detecting the cytosine methylation sites are two: (1) restriction enzyme digestion of methylated DNA, and (2) DNA bisulphite modification method.

In the first method, the enzymes utilized are sensitive to whether the DNA sequence at the restriction site is methylated or not. These enzymes have become a

major tool in the identification of methylated DNA loci (Cubas et al., 1999). An example is the enzyme McrBC, which is commonly used in these kind of experiments and has the property of cutting methylated DNA (recognition site $R^mC(N)_{55-103}R^mC$) (Lippman et al., 2003).

On the other hand, a more recent and accurate development has been trough 'bisulphite sequencing', which allows the site-specific identification of methylated cytosines on an amplified sequence (Martin et al., 2009). Bisulphite sequencing thus results from the bisulphite treatment of DNA to determine its pattern of methylation. The treatment of DNA with bisulphite uses a straightforward procedure based on a novel desulfonation method to efficiently convert non-methylated cytosine (C) to uracil (U). Thus, the bisulphite treatment introduces specific changes in the DNA sequence which depends on the methylation status of individual cytosine residues, yielding single-nucleotide resolution information about the methylation status of a segment of DNA. Methylated C residues are protected from the conversion, but non-methylated C residues are detected by the cytosine to thymine transition in the treated sequence.

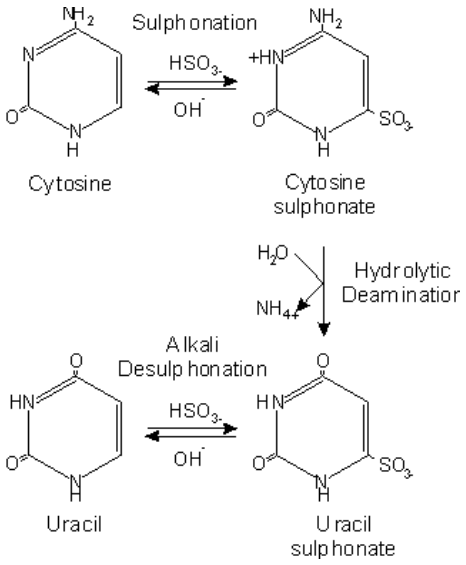


Figure 8.1: The DNA Bisulphite Modification kit (Human Genetic Signatures).

The objective of this treatment is therefore reduced to differentiating between single nucleotide polymorphisms (SNP) in cytosines and thymines resulting from bisulphite conversion.

Thus, the main objective of this chapter was to find out if any 'epigenetic' mechanism could be responsible for the differential phenotypic expression of self-(in)compatibility in almond cultivars.

8.2 Material and methods

Several SI/SC almond cultivars have been included in the methylation analysis, mainly the two parents of the population studied in this thesis, 'Blanquerna' as reference for a SC cultivars with the S_f haplotype ('S_8S_{fi}'), and 'Vivot' as reference for a SI cultivar with the same S_f haplotype ($S_{23}S_{fa}$). In addition, two other cultivars were included as controls, 'Soleta', another SC cultivars with the S_f haplotype ('$S_{23}S_{fi}$'), and 'Ponç', another SI cultivar with the S_f haplotype ($S_{27}S_{fa}$).

For a refinement of our results, we have additionally included two homozygous selections possessing the S_f haplotype. The first one is 'A2-199', coming from self-pollination of 'C-1322', a seedling from the cross 'Tuono' × 'Genco' (Dicenta, unpublished), belonging to the CEBAS-CSIC breeding program (Murcia). The other S_f homozygous selection, 'M-2-16', comes from the CITA breeding program at Zaragoza, being the result of self-pollinating of selection 'M-2-2', a seedling from the cross 'Tuono' × 'Ferragnès'.

Total DNA was extracted from styles at stage D (Felipe, 1977) using the CTAB method, following the procedure reported by Doyle and Doyle (1987). For a further purification of the stylar DNA, the cold phenol method was additionally used in order to reduce the amount of proteins and polysaccharides. DNA samples were then quantified by NANODROP and by gel electrophoresis with a λ DNA marker (Takara, Otsu, Japan).

Once the DNA was extracted, it was treated by using the DNA bisulphite modification treatment 'MeathylEasy' following the protocol indicated by the manufacture (Human Genetic Signatures, Brisbane, Australia).

This kit has been designed to convert efficiently cytosine to uracil or thymine and to reduce DNA degradation and loss without decreasing the conversion of C to U or T residues. The converted DNA is then able to be used for PCR amplification, cloning, sequencing and restriction digest assays. Methylated cytosines (Cs) are identified by treating genomic DNA with sodium bisulphite, which converts non-methylated C to U or T.

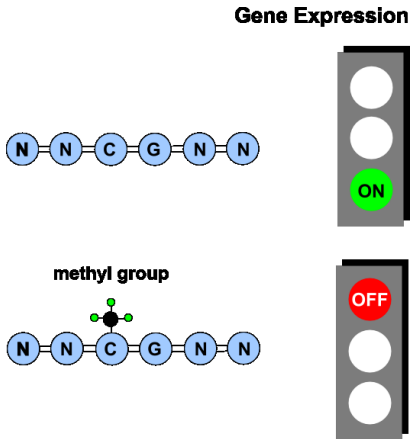


Figure 8.2: Methylation of C residues producing selective gene inactivation (OFF), and non-methylated C, producing gene activation (ON).

For PCR amplification, specific primers for methylation covering a maximum length of 250 bp available in our S_f upstream region were designed in this work. It was decided to design primers for covering short sequences following the manufacture recommendation since amplification with short fragments may give an exact methylation patterns. In total, 14 different primers were designed, located all of them in the 5' flanking region of the S_f-RNase (Table 8.1). They have been designed taken into account that after bisulphite treatment cytosine nucleotides were converted to uracil or thymine.

After treatment of the DNA with the bisulphite method, all possible primer combinations previously designed were checked by PCR. After this verification, the best five amplifications were selected and used for performing the analysis and obtaining their sequences.

Although PCR products could have been directly sequenced, it was preferred to clone the PCR products prior to sequencing for getting an adequate sensitivity. Thus, the 5 different PCR products were cloned using the PGEM-T-Easy Vector System (Promega, Madison, WI, USA). A total of four clones for each primer combination selected (five) of the 6 different DNA were used for sequencing, representing a total of 120 sequences. Clones were additionally purified using ExoSAP-IT (USB, Cleveland, OH, USA) and sequenced by the DTCS Quick Start kit and CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA, USA).

Table 8.1: Sequences of the specific primers for methylation designed for covering the upstream region of the S_f-RNase

Name	Sequence 5' → 3'	Reference
F1	GGTTTTGTTTTTTGTTTTTTTGTATTTGGT	This work
F2	TTAAATAATATTAAGATAAATTTGTAAATAAAGG	
F3	TTATTTGTTTAGGAAGTGTTAATTATAGGTAGTTGTT	
F4	AGTTGTAAGTTTGAATAAATAAAATATAATAGAAGTT	
F5	AAATTGTAAAATTAATTGAGAGATAAAAAATAGAGTTG	
F6	TGTTAGTATATTTTATGTAATTTTTTATATAAATAGTAAAT	
F7	TGTGGAAGTTTAGTAATTAAATTATGGTTTATTTT	
R1	AAATTATAATATTATTTAACCCCTTAACAATTTAATTATA	
R2	ACCACTAATACTCATAATAAAACACAAAAAAAAAA	
R3	ACTCTATTTTTTATCATCTCAATTAATTTTACAATTT	
R4	ATATACTACACAATAAATTAAAAACAAAAAAAAATCCA	
R5	AAAATAAACCATAATTTAATTAATAAAATTCCACA	
R6	TATTTATTCAAACTTACAACTTTCAAACCAAAAATTTAT	
R7	AAACAAAAAAACAAAAAACAAAACCTAAAAC	

All unmethylated sites should be displayed as thymines in the resulting amplified sequence. Concerning the methylated sequence, all cytosines should not be converted to any nucleotide and remain as cytosine in the resulting amplified sequence of the sense strand, and as guanine in the amplified antisense strand.

8.3 Results

After cloning and sequencing the PCR products, it was concluded that the bisulphite treatment worked properly in all DNAs tested because the treatment was able to eliminate all Cs that were not present as 5mCs by replacement with Ts, thus creating a nearly C-less sequence of mostly 3-base-DNA having predominantly A, G and T.

In the sequences shown in Figure 8.3, the presence of many cytosines can be observed in the upstream region of genomic DNA of 'Blanquerna' obtained from young leaves. After sequencing three clones of 'Blanquerna' with the kit, almost all cytosines were converted to thymines, as indicated in the manufacture protocol. However, only a cytosine residue was not able to be converted to thymine in the same position for the three clones.

After confirming the efficacy of the treatment, every nucleotide in the short sequences was meticulously checked in order to find out any possible cytosine responsible for methylation.

```
Bq6-1.B01_091119202H      ---------------------------------GTAAATAAAGGGGTATAT 18
Bq6-2.B02_09111921YW_     ---------------------------------GTAAATAAAGGGGTATAT 18
Bq6-3.F01_09112518EQ_     ---------------------------------GTAAATAAAGGGGTATAT 18
Bq_Sf_Upstream            AAAAAGAACCCAAACAACACCAAGACAAAGCCGCAAACAAAGGGGCACAC 350
                                                            * *** ******* * *

Bq6-1.B01_091119202H      AAAGTGAGAGGGATTTAAAGAGATTTAGGGTAATAAAAAATTAAAGCAGA 68
Bq6-2.B02_09111921YW_     AAAGTGAGAGGGATTTAAAGAGATTTAGGGTAATAAAAAATTAAAGCAGA 68
Bq6-3.F01_09112518EQ_     AAAGTGAGAGGGATTTAAAGAGATTTAGGGTAATAAAAAATTAAAGCAGA 68
Bq_Sf_Upstream            AAAGTGAGAGGGATTTAAAGAGATTTAGGGCAACAAAAAACTAAAGCAGA 400
                          ******************************** ** ****** ********

Bq6-1.B01_091119202H      GAGTATAATTAAATTGTTAAGGGGTTAAATAATATTATAATTT 111
Bq6-2.B02_09111921YW_     GAGTATAATTAAATTGTTAAGGGGTTAAATAATATTATAATTT 111
Bq6-3.F01_09112518EQ_     GAGTATAATTAAATTGTTAAGGGGTTAAATAATATTATAATTT 111
Bq_Sf_Upstream            GAGCACAACTAAACCGTCAAGGGGTCAAACAACACCACAACTT 443
                          *** * ** ****  ** ******* *** ** *  * ** **
```

Figure 8.3: Nucleotide sequences of 3 different clones of 'Blanquerna' showing that the sequences from pistilar DNA are the same as from genomic DNA from leaf. The red circle shows where the methylation cytosine residue appeared.

Among the 5 primer combinations used for cloning and sequencing the DNAs after treatment with the bisulphite kit, only a cytosine methylation residue could be

detected in the region flanked by the primer combination F2/R1. The length of this primer region was around 120 bp, located approximately 650 bp upstream to the start codon.

The analysis with the pistil DNA samples of the SI cultivars 'Vivot' and 'Ponç' was also carried out. In these cultivars, any conversion of the thymine nucleotide to cytosine could be detected in all the four clones used for each cultivar (Fig. 8.4).

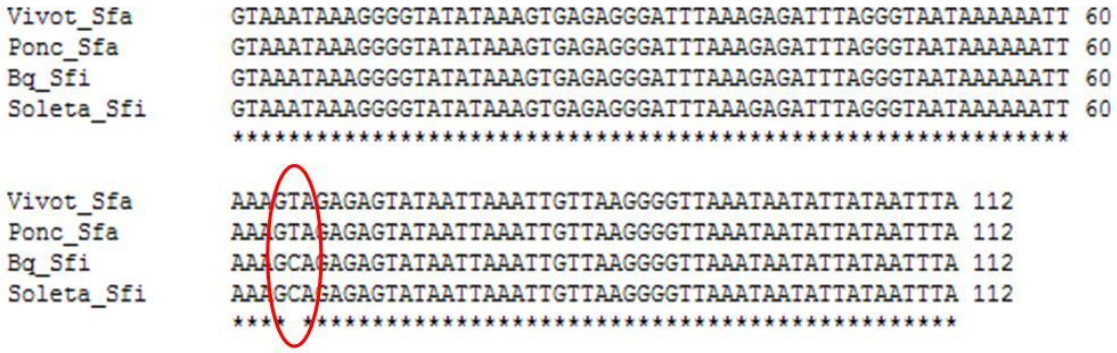


Figure 8.4: Nucleotide sequences of the SC cultivars 'Blanquerna' and 'Soleta' and the SI cultivars 'Vivot' and 'Ponç'. The red circle shows that the sequences from pistilar DNA in the SC cultivars are exactly the same as from genomic DNA but with the exception of the conversion of the cytosines to thymines. Red circle shows where the methylation redidues happens.

On the other hand, a cytosine residue was found in the sequences of the SC cultivars after bisulphite treatment. This methylation cytosine was detected in the form CNG, where the N nucleotide was adenine (Fig. 8.5). In fact, both cultivars showed the same methylation residue in the same position of the sequence. In order to confirm this methylation, the same procedure was carried out as described above including cloning of the PCR product until sequencing, in the homozygous S_f selections 'A2-199' and 'M2-16'. In both cases, also a cytosine residue was detected in the *5'* upstream region covered by the primer set F2/R1.

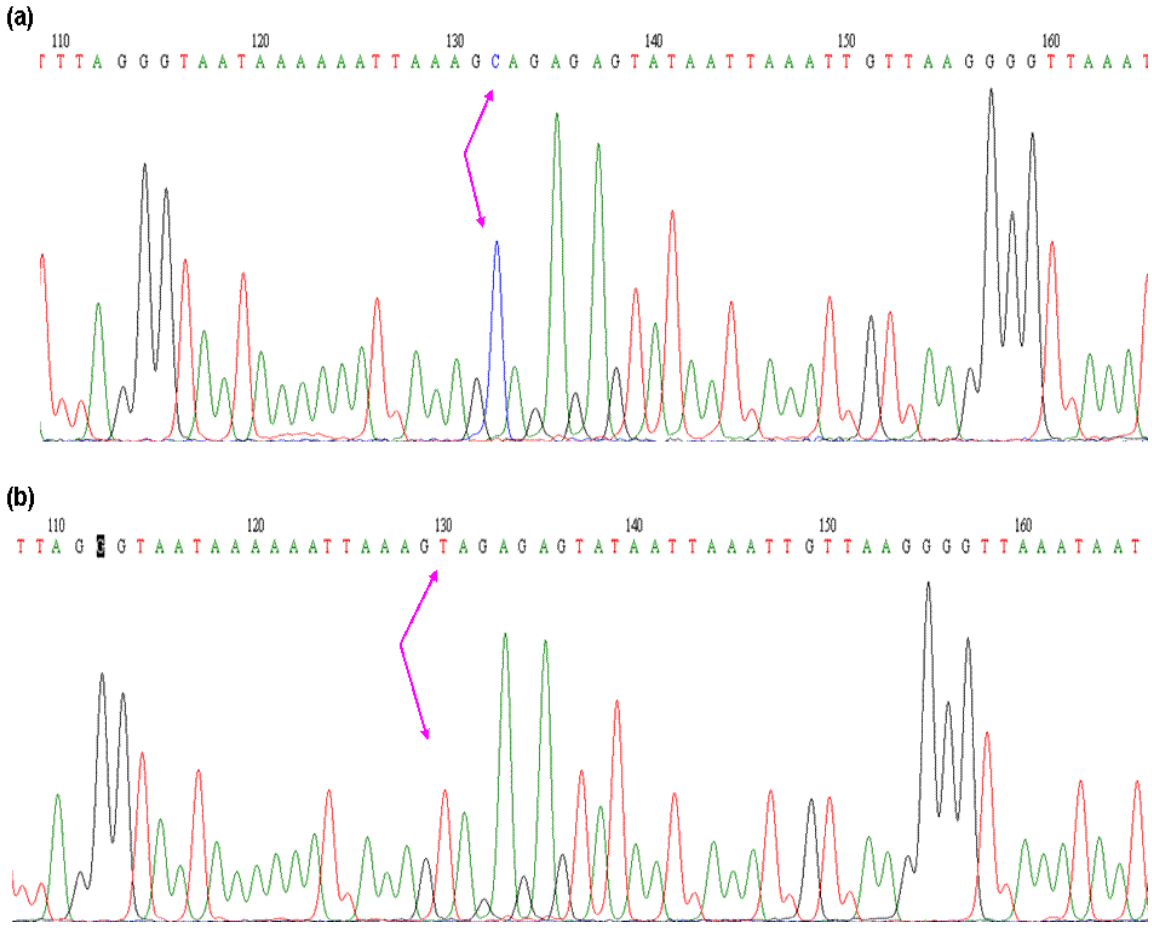


Figure 8.5: DNA sequence from methylated DNA of the almond SC 'Blanquerna' (a) and from the SI cultivar 'Vivot', which sequence was not-methylated after bisulphite conversion (b). Pink arrows show where methylated DNA produced its conversion or not to cytosine.

8.4 Discussion

It has been well established that DNA methylation occurs generally in differentiated tissues. In previous studies, DNA methylation was described to take place in specific organs (Gehring and Henikoff, 2007; Shiba et al. 2006). In *Brassica rapa* (Shiba et al., 2006), DNA was extracted from the anther tapetum, in order to examine the genomic region of the SP11 (gene responsible for pollen specificaty in Sporophytic SI), in which suppression did not take place. Methylation level was compared in several tissues such as leaves, pollen and petals, but methylation was specific for the anther tapetum and not for the other tissues. Thus, in our case it was decided to use DNA from

pistil because the gene whose action we were interested in is expressed in the pistils through two different phenotypic expressions.

Genomic sequencing trough the bisulphite method allows the determination of the methylation status of every cytosine residue. This information is particularly relevant to studies of methylation in developmental processes. Thus, the bisulphite genomic sequencing method was applied in our plants to determine the exact methylation patterns of the S_f-RNase gene in a very small amount of genomic DNA.

Methylation of cytosine is associated with gene silencing, and genes with 5-methylcytosine in their promoter region are usually transcriptionally silent (Jones and Takai, 2001). Thus, DNA methylation does not change the DNA sequence, but does change its expression level, referred as an epigenetic change. Bender (2004) proposed that if the absence of RNase transcripts in the specific tissues is detected, the detection of DNA methylation could be an effective approach for determining this epigenetic change.

In our experiment, this hypothesis may have been confirmed, because a change of DNA methylation appeared only in the self-compatible cultivars. In fact, previous studies have shown that the S_{fi} gene lacks RNase activity (Bošković et al., 1997), thus its expression is inhibited. These results may allow to link SI to DNA methylation.

DNA methylation could be responsible of the activation/inactivation of the S_f haplotype in almond. When the S_f-RNase sequence is methylated, an inhibition of the expression takes place, as it happens in 'Blanquerna' and 'Soleta'. Thus, this inactivation would be traduced to SC. In the case of S_f-RNase sequences with non-methylated cytosines, as it happens in 'Vivot' and 'Ponç', the RNase would remain active and, as a consequence, it would be SI.

At the origen, the S_f haplotype, would be in its active form, thus conferring AI as it happens with the other *S* haplotypes. However, due to this epigenetic change and consequently due to DNA methylation, the same haplotype could have appeared in a new form, the SC one because methylation could affect the activity of this allele. The same phenomenon could probably have taken place in other Rosaceae species, such as

sweet cherry, apricot and pear, although for the moment any information is available on the fact that methylation of the SI alleles of these species could have happened in order to change their activity towards SC.

The development of new markers able to distinguish the two different forms of the same allele would be a magnificent tool. Thus further studies are required in order to better understand this SI mechanism.

8.5 Conclusion

DNA methylation has been widely recognized to suppress gene expression as a regulatory factor (Jones and Takai, 2001). The range of different 'anomalous' phenomena has provided a handful of possibilities for many different researchers with access to molecular analysis. All these attempts to unravel each fascinating 'exception' to be accepted as a model are now providing a clearer picture on how an epigenetic system operates in plants. Importantly, a huge body of work is beginning to show that these are not separate 'anomalies' but rather the different faces of the same underlying system. Thus, these findings suggest that the methylated *5'* region of the S_f-RNase is a crucial site for *S*-RNase expression of S_f in the style.

9. MOLECULAR MODELLING OF *S*-RNASES INVOLVED IN SELF-INCOMPATIBILITY

9. MOLECULAR MODELLING OF *S*-RNASES INVOLVED IN SELF-INCOMPATIBILITY

9.1 Introduction

In spite of the knowledge on the genetic structure of the female and male determinants of SI, as intensely described in the previous chapters, the nature of their interaction remains unclear. Thus, the knowledge of the three dimensional (3D) structure of the proteins involved in the SI reaction may shed light in elucidating the recognition mechanism of Rosaceous *S*-RNases at the molecular levels to understand how these proteins mediate the GSI function to fulfil their biological roles.

Proteins are built from sequences of amino acid residues, ordered by the corresponding genetic code. The linked amino acid residues bond in space to form a 3D structure. The knowledge of the 3D structure has been useful in order to understand how some proteins work and which molecular mechanisms underpin their function. As a consequence, this approach has been considered useful in the general frame of studying the mechanism of GSI in almonds.

Protein structure can be determined experimentally using X-ray crystallography, nuclear magnetic resonance spectroscopy and cryoelectron microscopy, but these approaches are time-consuming (Ida et al., 2001). Consequently, predictive computer molecular modelling has been considered a useful alternative. This approach has challenged biochemists, computational chemists, physicists, mathematicians and computer scientists for decades. Molecular modelling may be defined as the science and art of studying molecular structure and function through model building and computation.

Protein structures are guided by two sets on principles operating on vastly different time scales. The first set of principles is defined by the laws of physics, while the second set is directed by the theory of evolution. Each of these two sets of principles

has led to the development of predictive methods for building 3D protein models (Hrmova and Fincher, 2009).

Currently, one of the most popular comparative modelling programs is MODELLER, which was designed by Andrej Sali and Thomas Blundell in Cambridge, UK (Sali and Blundell, 1993). MODELLER is a computer program that models 3D structures of proteins and their assemblies by satisfaction of spatial restraints. The user provides an alignment of a sequence to be modelled with related 3D structures already known and MODELLER will automatically calculate a model. The array of 10-50 models produced by MODELLER can be evaluated through another set of computer programs that assess the spatial quality and the energy profiles of protein models.

Thus, after selecting the best model, it needs to be put into perspective with a biological function and tested to see if it is helpful in proposing a useful hypothesis in biology.

Thus, the objective of this chapter was to identify the three-dimensional structures of the almond *S*-RNases and SFBs trough molecular modelling tools and to investigate the link between the 3D structures and the SI mechanism.

9.2 Material and Methods

S_f-RNase (AB467371), S_{23}-RNase (AB488496) and S_8-RNase (AB481108) sequences obtained from the almond cultivars ‘Blanquerna’ and ‘Vivot’ were used for constructing their respective 3D models.

The modelling procedure begun with the alignment of the sequence to be modelled (target) with related known 3D structure (template) derived from the Protein Data Bank (PDB) using programs such as Fasta and Blast (EMBL nucleotide database). Among the possible templates, the one to be selected has to show the highest identity with the target, higher than 35%. The first step of the experimental approach requires the identification of the 3D template and then we used the coordinates of this template protein as a basis for further modelling.

In our case, the best candidate template selected was the RNase MC1 mutant with accession number 1J1G (Numata et al., 2003). The adjustment between our sequences and the template was performed manually in order to minimize the number of gaps and insertions/deletions (INDELS). The frame of the 3D model was generated by using MODELLER 9v5. A total of 40 models were constructed for each *S*-RNase. The 4 models with the lowest value of the Modeller objective function were chosen for further refinement. Energy function was evaluated through the software PROSAIIv3, which is also a powerful tool in protein structure research. This program serves to indicate the quality of protein structures. It calculates a score for the input-structure of the best models selected.

On the other hand, stereochemical quality and overall G-factors of the protein structure selected was calculated for the four candidates by the Ramachandran plot analysis in the PROCHECK software. This software serves for assessing the stereochemical quality of a given protein structure, and for comparing the residue-by-residue geometry of a set of closely-related structures. The models with lower number of amino acid residues in disallowed regions were selected as the most suitable models (Figs. 9.1 and 9.2).

A Ramachandran plot (also known as Ramachandran map or a Ramachandran diagram) provided by PROCHECK visualises dihedral angles ψ against φ of amino acid residues in a protein structure. It shows the possible conformation of ψ and φ angles for a polypeptide (Ramachandran et al., 1963).

For a further loop refinement of the 3D structures, 40 new models of the previously best model were generated taken as a reference the one previously selected. The same steps as described above were followed, selecting the best four models according to their lowest values of the Modeller objective function, and then selecting the 'best of the best' from the results obtained by PROSAIIv3 and PROCHECK.

Finally, the molecular graphics of the models were generated with PYMOL, which is a molecular visualization system and a 3D editor (http://www.pymol.org).

9.3 Results and Discussion

The 3D models of the S_f, S_{23} and the S_8-RNases were generated and compared. Their 3D structures were based on the template structure RNase MC1 mutant (N71S), because the identity between the template and the target sequences were 42%. However, SFB_f, SFB_8 and SFB_{23} structures could not be determined because no sequence identity higher than 30% with templates was found in the Protein Data Bank (PDB).

Protein structures are built up by combinations of secondary structural elements, α-helices and β-strands that are connected by loops. These structural elements form the core regions (the inside of the molecule) and are connected by loop regions on the protein surface with surface exposed α-helices and β-strands.

Concerning our *S*-RNases, the overall dimensions of the molecules were approximately 40 Å x 50 Å x 30 Å. Their structure belonged to the α and β class, with six α helices and six β strands connected by loops. The folding topologies of its main chains were very similar to the topologies of the RNase T_2 family enzymes.

Ramachandran plot statistics for the S_f, S_8 and S_{23} RNases showed than 97% amino acid residues were positioned in the 'allowed' regions. In fact, when structures place 95-97% or more of the amino acid residues in the 'allowed' positions, they are considered to be reliable in modelling experiments, and this indicates how well the structures fit with the expected main chain length and torsion angle distributions (Laskowski et al., 1993; Kleywegt and Jones, 1996).

Among the four models selected for the further refinement, only the best model was chosen based on the Ramachandran plot statistics. As shown in Fig. 9.1, in the model sfi1_BL00040001 (a), all residues were positioned in the allowed region (red arrow), whereas in the model sfi1_BL00010001, 1.6% of the residues were in the disallowed region (green arrow). Thus, the model BL00040001 has been selected as the best model to be analysed.

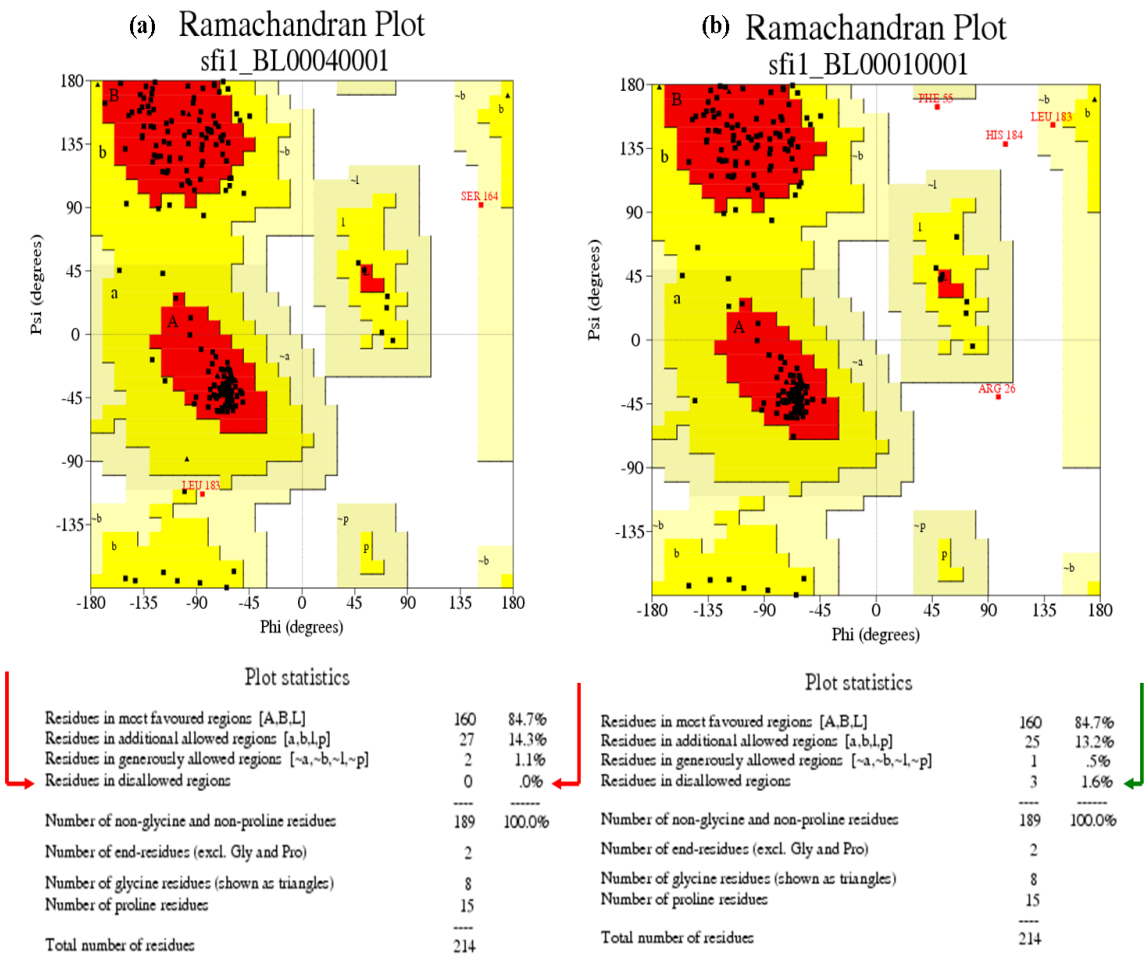


Fig 9.1: The Ramachandran plot of sfi1_BL00040001 (a), and sfi1_BL00010001 (b).

Higher numbers of residues in the disallowed region reflect a distorted geometry in the models, because there are higher proportions of residues falling outside the limits of main chain bond length and torsion angles of the small molecule library (Engh and Huber, 1991). Thus, the results indicate that our models were well defined and had a high quality of protein structure.

When the three S_f, S_8 and S_{23} *S*-RNases were superpositioned, the S_f-RNase structure had a loop region longer than in the S_8 and S_{23} models. This loop, shown in Figs. 9.2 and 9.3, contained the amino acid residues CKG NPQ RQA KSQ PKN RGK SQP KSQ ATT QFL, which were placed between the conserved domains RC4 and C5. Trough the software PYMOL it has been possible to visualise, which amino acid residues comprised α-helices, β-strands and loops (Fig. 9.2).

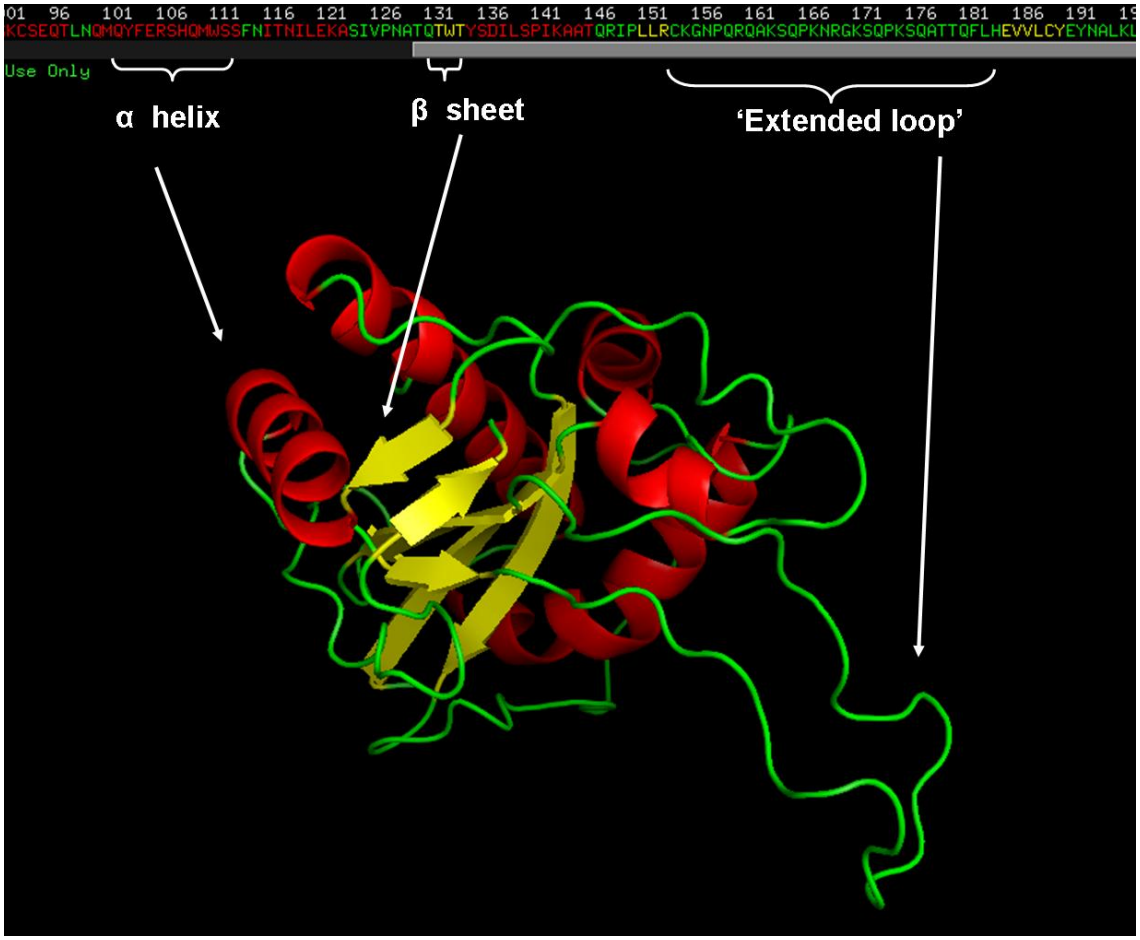


Fig. 9.2: 3D model of the S_f-RNase showing its α-helices (red), β-sheets (yellow) in both senses (parallel and antiparallel) and loops (green). In the upper part of the figure the respective residues comprising of α-helix, β-sheet and loops are coloured in red, yellow and green.

It is known that in 3D structures, loops serve to interconnect α-helices and β-strands and that longer loops are in many cases susceptible to proteolytic degradation (Branden and Tooze, 1998).

As the main structural difference found between the S_f, S_8 and S_{23} resides is in this 'extended looping region', it could be possible that the amino acid residues forming this long loop presented in the S_f structure could be prone to degradation and/or inactivation and as a consequence this *S*-RNase could be less stable and thus would allow its pollen tube growth trough its own pistil.

```
Blanquerna_S8-RNase     MATLRQSFAFLVLAFAFFLCFIMST---GSYVYFQFVQQWPPTTCRLSSK-PSNQHRPLQ 56
Vivot_S23_RNase         MAVWKSSPAFLVLAFALFLCFIMST---GSYVYFQFVQQWPPTNCRVRIKRPCPNPRPLQ 57
Sfi_Blanquerna          MGILKSSLGFLVLAFAFFLCFIMSTSGDGSYVYLQFVQQWPPTTCRFSGK-PSNNRRPLP 59
                        *.  :.* .*******:********   *****:*********.**.  * *. : ***

Blanquerna_S8-RNase     RFTIHGLWPSNYSNPRKPSNCNGSQFNFMKVYPQLRIKLKRSWPDVEGGNDIKFWEGEWN 116
Vivot_S23_RNase         YFTIHGLWPSNYSNPTKPSKCTGPKFDARKVSPKMRIKLKISWPDVESGNDIRFWEGEWN 117
Sfi_Blanquerna          IFTIHGIWPSNYSNPRMRSNCTGSQFK-KILSPRLRSKLERAWPDVESGNDIKFWEDEWN 118
                         *****:********   *:*.*.:*.   : *::* **: :*****.****:***.***

Blanquerna_S8-RNase     KHGTCSERTLNQMQYFEVSHAMWRSYNITNILKDAHIVPNPTQRWKYSDIVSPIKTATGR 176
Vivot_S23_RNase         KHGTCSRERLNQMQYFERSHDMWLSYNITEILKNASIVPNATQKWSYSDIISPIKAATGS 177
Sfi_Blanquerna          KHGKCSEQTLNQMQYFERSHQMWSSFNITNILEKASIVPNATQTWTYSDILSPIKAATQR 178
                        ***.**.. ******** ** ** *:***:**:.* ****.** *.****:****:**

Blanquerna_S8-RNase     TPTLRCKTDP-----AMPNN-----------SQLLHEVVFCYGYNAKLHIDCNRIAGCRN 220
Vivot_S23_RNase         TPLLRCK---------QAKN-----------TLLLHEVVFCYEYDALKQIDCNRIAGCGN 217
Sfi_Blanquerna          IPLLRCKGNPQRQAKSQPKNRGKSQPKSQATTQFLHEVVLCYEYNALKLIDCNRIAGCWN 238
                         * ****          .:*          : :* ****:** *:*   ********* *

Blanquerna_S8-RNase     HIDILFQ 227
Vivot_S23_RNase         QQAISFQ 224
Sfi_Blanquerna          NVDIKFH 245
                        :  * *:
```

Figure 9.3: Multiple sequence alignment of the sequences of S_f, S_8 and S_{23} RNases, indicating the amino acid residues that belong to the 'extended loop' in the S_f 3D structure (underlined in blue).

Additionally, the 3D models of S_f, S_8 and S_{23} RNases were compared with that of another rosaceous species, the *Pyrus pyrifolia* S_3-RNase (Matsuura et al., 2001). The S_3-RNase structure was consistent with the S_8 and S_{23} SI RNase models. As the S_8 and S_{23} models did not contain this long loop, we may suggest that the amino acid residues found in the SC RNase (Fig. 9.3) could be responsible for formation of a long loop between the conserved domains RC4 and C5. Thus, the main structural differences between the SI and SC RNases reside in this region.

We have commented in previous chapters that the S_f haplotype has two different forms, one conferring SC and the other SI despite its identical sequence. Also epigenetic factors would be probably responsible for turnover of the expression from SI to SC in plants. Thus, with our results we cannot confirm that the presence of the long loop in the S_f may cause SI in almond. However, since it seems that all self-incompatible RNases studied here (excluding the S_{fa}) do not contain the extended loop (Fig. 9.4), we may suggest that the main structural difference between the SI RNases and SC RNases reside in this 'extended loop region'.

Further studies are required to ascertain the possible role of this loop in SI and to find out if a possible association between the extended loop and its biological function is involved in the SI mechanism of plants.

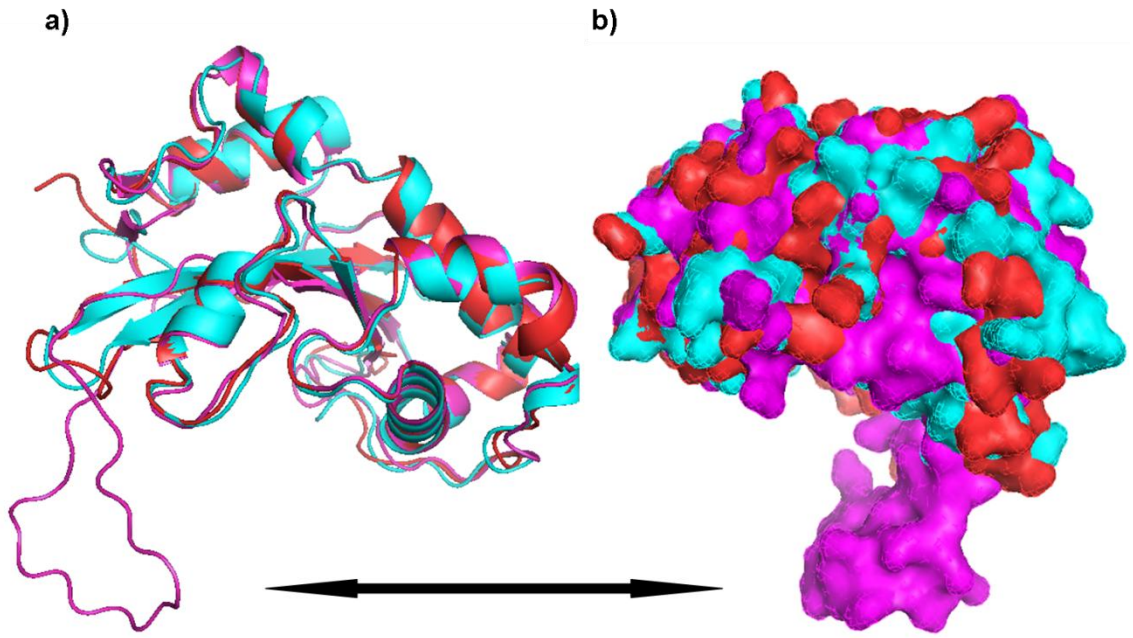


Figure 9.4: Ribbon diagram (a) and surface representation (b) of modelled structure of almond S_f-RNase (pink), S_8-RNase (blue) and S_{23}-RNase (red), showing secondary structural elements. Black arrow indicates the 'long loop' found in the S_f model.

10. DISCUSIÓN GENERAL

10. DISCUSIÓN GENERAL

El objetivo de partida de esta tesis fue profundizar en el estudio de las causas del inesperado comportamiento fenotípico observado en las primeras autopolinizaciones de los individuos obtenidos en el cruzamiento de 'Vivot' × 'Blanquerna'.

La observación microscópica del crecimiento de los tubos polínicos indicó que no se producía la proporción esperada de 1:1 en la compatibilidad de la autopolinización de esta población, a pesar de la presencia del alelo S_f, responsable de la AC, en el parental masculino.

Por ello, se decidió realizar durante dos años consecutivos más observaciones microscópicas en todos los individuos y así poder confirmar los resultados previos obtenidos. Igualmente se decidió aplicar las nuevas técnicas moleculares a esta población para profundizar en el conocimiento genético y molecular de los mecanismos que pueden estar actuando.

10.1 El fenotipo de la descendencia

Las autopolinizaciones realizadas en laboratorio en las variedades 'Vivot' y 'Blanquerna' mostraron que 'Vivot' es claramente AI y 'Blanquerna' es AC. Estos resultados fueron posteriormente corroborados en campo mediante el embolsamiento de las ramas.

Durante 3 años consecutivos las observaciones microscópicas del crecimiento de los tubos polínicos en toda la población, mostraron que solamente 19 plantas del total de 77 eran AC, lo que representa casi un 25%, mientras que las otras 58 eran AI (casi el 75%).

Teniendo en cuenta que 'Blanquerna' es una variedad AC y 'Vivot' AI (independientemente de su genotipo), el 50% de los individuos deberían haber sido AC

siguiendo una distribución mendeliana (dicenta y García, 1993). Sin embargo, los resultados obtenidos no se ajustaron a estas proporciones, lo que hace suponer que en este cruzamiento se ha producido una anomalía en la trasmisión del alelo responsable de la AC de 'Blanquerna' a la descendencia, que ha producido una desviación de las proporciones esperadas hacia una mayor presencia de individuos AI de la esperada en la descendencia.

10.2 El genotipo de la descendencia

Los genotipos *S* de los padres se determinaron mediante dos técnicas moleculares (NEpHGE y PCR), concluyendo que 'Vivot' poseía el genotipo S_fS_{23}, con la presencia del alelo S_f a pesar de ser AI, mientras que el genotipo de 'Blanquerna' era S_fS_8.

Teniendo en cuenta estos genotipos y siguiendo una trasmisión mendeliana, la descendencia debería tener una distribución en la segregación de 1 $S_{23}S_8$: 1 $S_{23}S_f$: 1 S_fS_8: 1 S_fS_f. Sin embargo, los datos moleculares mostraron únicamente dos genotipos en la descendencia, $S_{23}S_8$ y S_fS_8, por lo cual la segregación real fue de 1: 1.

Parece ser que algunas variedades descritas recientemente como AI a pesar de la presencia del alelo *Sf*, como 'Ponç', 'Alzina' y 'Garondès' (Kodad et al., 2009; 2010), presentan una versión activa del alelo S_f (S_{fa}). Todo parece indicar que el parental utilizado en esta población, 'Vivot', posee igualmente esta versión activa del alelo S_f. Como consecuencia de estas observaciones, se ha adoptado la designación S_{fa} para la versión activa del alelo S_f y su expresión AI, y la designación S_{fi} para la versión inactiva de expresión AC.

Por lo tanto, con esta diferenciación de las dos versiones del alelo S_f (activo e inactivo), la distribución de genotipos si no se produjeran reacciones de incompatibilidad sería 1 $S_{23}S_8$: 1 $S_{23}S_{fi}$: 1 $S_{fa}S_8$: 1 $S_{fa}S_{fi}$. Sin embargo, sólo se identificaron dos genotipos, $S_{23}S_8$ y $S_{fa}S_8$, lo que implica que los tubos polínicos de 'Blanquerna' que expresan el alelo S_{fi} de han sido reconocidos y degradados por la S_{fa}-RNasa de 'Vivot', por lo que no ha podido producirse el crecimiento del polen S_{fi} a

través de los pistilos de 'Vivot' ($S_{23}S_{fa}$). Como consecuencia de ello, sólo el alelo S_8 de 'Blanquerna' ha podido trasmitirse a la descendencia, en la que no se ha encontrado ningún genotipo $S_{fa}S_{fi}$ o $S_{23}S_{fi}$.

Estos resultados reflejan la completa actividad del S_{fa}, no solamente en cuanto a la producción de la *S*-RNasa, como indican los resultados de NEpHGE, sino también en su reconocimiento de su homólogo S_{fi}.

10.3 Correlación entre fenotipos y genotipos

Los resultados fenotípicos y genotípicos obtenidos en este trabajo han resultado ser una desviación de lo esperado en el momento de la creación de esta familia.

Según los genotipos identificados, todos los árboles deberían haber sido autoincompatibles ($S_{fa}S_8$ y $S_{23}S_8$), puesto que ninguno de ellos poseía el alelo S_{fi}. Sin embargo, casi un 25% de ellos, han sido caracterizados como AC, incluidos algunos con el genotipo S_8S_{23}.

Este fenómeno sugiere la posibilidad de que otro gen o factor ajeno al locus *S* podría estar interfiriendo en la familia estudiada. En otras familias y especies vegetales se han descrito varios factores externos o genes modificadores que pueden afectar a la compatibilidad de la polinización (Thompson y Taylor, 1971; Hinata et al., 1995; Tsukamoto et al., 2003).

Parece que uno de estos genes podría actuar en esta población. Este gen sería posiblemente recesivo, ya que la proporción de plantas AC:AI es de 1:3. Si este modificador hubiese sido dominante o codominante, al menos el 50% de las plantas deberían haber sido AC, pero no ha sido el caso.

El modelo propuesto en esta tesis para poder confirmar la presencia de genes externos ha sido el siguiente: si los genotipos de los padres son $S_{fa}S_{23}$/Mm para 'Vivot' y $S_{fi}S_8$/Mm para 'Blanquerna' y como indican los genotipos, el alelo S_{fi} de 'Blanquerna' es incapaz de crecer en los pistilos de 'Vivot', la segregación de la población debería ser

1 S_8S_{fa}/MM: 2 S_8S_{fa}/Mm: 1 S_8S_{fa}/mm: 1 S_8S_{23}/MM: 2 S_8S_{23}/Mm: 1 S_8S_{23}/mm, con una proporción de 1:2:1 para MM:Mm:mm.

En el caso de que la AC se debiese a la presencia del genotipo "mm", solamente el 25% de la población debería ser AC. Esta distribución es, pues, muy cercana a la distribución obtenida en las observaciones microscópicas (24.7%).

En otras especies se han propuesto modelos similares, como es el caso de *Petunia axillaris* (Tsukamoto et al., 2003). Los resultados obtenidos en la presente memoria permiten concluir que la presencia del alelo S_f en el almendro no indica siempre la presencia de AC. De igual modo, se puede sugerir que otros genes, aun desconocidos y posiblemente localizados fuera del locus S, podrían interferir en el mecanismo de AI en el reino vegetal.

10.4 Caracterización del locus S de 'Vivot' y 'Blanquerna'

Como todas las secuencias del alelo S_f, independientemente de su expresión, depositadas hasta ahora en las bases de datos son completamente iguales (Bošković et al., 2007; Kodak et al., 2009; 2010), uno de los primeros objetivos de esta memoria fue la de construir una librería genómica del tipo fósmido, con el fin de comprobar si las diferentes expresiones eran debidas a posibles mutaciones localizadas en la región 'upstream' de las *S*-RNasas. Sin embargo, los resultados de este estudio han mostrado que las secuencias de ambas versiones eran idénticas al 100%, al igual que las secuencias del factor del polen, el SFB.

Posteriormente, después de haber comprobado que el alelo S_f producía mRNA para la S_f-RNasa activa en los pistilos de 'Vivot', se quiso cuantificar y comprobar el nivel de expresión mediante PCR a tiempo real. Estos resultados confirmaron que se producían transcritos de la S_{fa}-RNasa en el tejido pistilar de 'Vivot', mientras que estaban ausentes en los pistilos de 'Blanquerna', confirmando las indicaciones previas de la no actividad del S_f en las variedades AC de almendro (Bošković et al., 1997). Por lo tanto, la expresión del S_{fa} detectada en el pistilo indica que este alelo es funcional y AI, mientras que la otra versión es inactiva y AC.

Todos estos resultados refuerzan la hipótesis de que algún factor externo, cuya localización en el genoma no ha sido por el momento descrita en ninguna otra especie de rosáceas, podría estar interfiriendo en el mecanismo de AI.

10.5 Identificación de QTLs asociados a la AC en la población estudiada

Como se ha descrito a lo largo de esta memoria, la AI requiere de otros genes externos al locus *S* para su completa funcionalidad. Este tipo de factores se han identificados en especies como *Nicotiana alata* Link & Otto o *Petunia axillaris* (Lam.) B.S.P. Sin embargo, en el caso de la familia de las rosáceas, a pesar de la publicación de varias propuestas sobre la presencia de genes modificadores en el mecanismo de la AI, no se han localizado estos posibles genes en el genoma de *Prunus* o *Pyrus* (Moriya et al., 2009; Vilanova et al., 2006; Wünsch and Hormaza, 2004).

El análisis de QTLs llevado a cabo mediante el empleo de marcadores microsatélites y específicos para el locus *S*, ha permitido determinar en el genoma de *Prunus* los grupos de ligamiento en los que se encuentran dos QTLs implicados en la AC. El primer QTL se ha localizado en la zona central del grupo 6, mientras que el segundo QTL se ha posicionado al principio del grupo 8. Estos resultados permiten sugerir que la AC en el almendro es controlada por un gen mayor (*S*-locus) y dos QTLs (*S*-Q_1 y *S*-Q_2) en las plantas que poseen la versión activa del alelo S_f, mientras que en las plantas con la versión inactiva (S_{fi}), la AC seria controlada únicamente por el locus *S*. Por lo tanto, usando el marcador especifico para el alelo S_f, diseñado por Channuntapipat et al. (2001), sería suficiente para la efectiva selección de los descendientes de variedades AC que presenten este alelo. Sin embargo, para los individuos procedentes de cruzamientos con el alelo S_f en su versión activa, se recomendaría la combinación de los SSR BPPCT025 y CPSCT018 para la discriminación de plantas AC.

10.6 Metilación de ADN en 'Vivot' y 'Blanquerna'

A pesar de haber localizado dos QTLs relacionados con la expresión de la AC, otros factores o mecanismos podrían estar igualmente implicados en las dos diferentes expresiones del haplotipo S_f en el almendro. Por ello hay que tener en cuenta la existencia de mecanismos capaces de cambiar los fenotipos de los genes sin que se produzcan cambios en las secuencias del ADN (Jeanish and Bird, 2003). Estas modificaciones son conocidas como factores epigenéticos. Entre ellos se encuentra la metilación del ADN, el cual es fundamental en la regulación del silenciamiento de los genes, y que puede provocar alteraciones en la transcripción genética sin necesidad de que se produzca una alteración en la secuencia del ADN, siendo uno de los mecanismos responsables de la plasticidad fenotípica.

Un método eficaz para determinar si el ADN está metilado o no, es el tratamiento del ADN con bisulfito (Martin et al., 2009). Como la metilacion de las citosinas está asociada al silenciamiento de los genes, este método ha sido diseñado propiamente para poder convertir químicamente las citosinas en timinas únicamente en las zonas donde no se produce ninguna metilacion. Sin embargo, este kit es incapaz de convertirlas a timinas en las zonas donde sí se haya producido la metilación.

En este estudio se ha podido determinar que las variedades AC y poseedoras de la versión inactiva del alelo S_f, presentan un cambio de ADN metilado en las secuencias. Se ha comprobado que la metilacion de una citosina ha tenido lugar en la región 'upstream' usando la combinación de cebadores F2/R1 diseñados a partir de la secuencia del alelo S_f. De este modo, se ha observado que cuando la secuencia de la S_f-RNasa está metilada, se produce una inhibición de la expresión, de manera que esta inactivación se traduce en una expresión AC ('Blanquerna' y 'Soleta'), mientras que en el caso contrario, cuando el alelo S_f no está metilado, como ocurre en 'Vivot' y 'Ponç', la RNasa permanecería activa y por lo tanto sería AI.

10.7 Estructura 3D de los alelos S_f, S_8 y S_{23}

En esta memoria se ha determinado igualmente la estructura 3D de las RNasas (S_f, S_8 y S_{23}). Las dimensiones de la molécula han sido aproximadamente de 40 x 50 x 30 Å en todos los casos. Las estructuras proteicas obtenidas en este estudio pertenecen a la clase α y β, con seis hélices y seis hebras beta.

Sin embargo, cuando se han superpuesto todas las estructuras, la de la S_f-RNasa presentó una región con un lazo mayor que las de las otras dos, ambas de alelos AI. La función del lazo es conectar las hélices α a las hebras β. Branden y Tooze (1998) sugirieron que los lazos grandes son susceptibles a la degradación proteólica.

Por lo tanto, la región larga del lazo presente en la estructura 3D del S_f podría ser propensa a la degradación e inactivación y, como consecuencia de ello, permitiría el crecimiento del tubo polínico a través de su propio pistilo. Este aspecto, sin embargo, requiere una mayor profundización.

10.8 La autocompatibilidad en el almendro

Desde la primera determinación de la AC en el almendro (Almeida, 1945) se han utilizado varias técnicas para evaluar su nivel de expresión, como el conteo de frutos después del embolsamiento de ramas, el crecimiento de tubos polínicos y, más recientemente, la identificación del alelo S_f mediante marcadores moleculares y secuenciación. Sin embargo, ninguno de ellos ha dado resultados totalmente seguros, ya que todos tienen ventajas y desventajas.

Los datos que se obtienen en campo a través de los embolsamientos y las polinizaciones artificiales, a pesar de ser el método más natural para determinar el verdadero nivel de AI, presentan una desventaja muy importante, como son las condiciones atmosféricas. Por lo tanto, las emasculaciones y polinizaciones que se realizan al aire libre, pueden estar negativamente influidas por heladas y lluvias. Así pues, se recomienda aplicar estos métodos para determinar la AC al final del proceso de selección, cuando las plantas presenten indicios fundados para su selección final.

Por otro lado, la determinación de la AC por medio del crecimiento de los tubos polínicos es un método muy fiable y usado frecuentemente en los programas de mejora, ya que se ha encontrado una correlación entre el crecimiento de los tubos polínicos y el cuajado de una selección autopolinizada (Ben Njima and Socias i Company, 1995; Kodad and Socias i Company, 2006; Socias i Company y Felipe, 1987).

Los recientes avances en genética a través del empleo de las técnicas moleculares, y especialmente en la determinación del alelo S_f han tenido una gran repercusión en los programas de mejora en el almendro. Sin embargo, al igual que los otros métodos, estas técnicas presentan limitaciones. La teoría de que todos los genotipos que tuviesen el alelo S_f deberían ser AC o que la presencia del haplotipo S_f en el almendro estuviese siempre asociada a la AC se muestra que no es totalmente válida. La nueva versión autoincompatible de este alelo esta siendo recientemente objeto de estudio en muchos centros de investigación. Aunque la AI en el almendro había sido siempre asociada a la presencia del alelo S_f, el cual es dominante sobre los otros haplotipos del locus S (Socias i Company 1984), los nuevos resultados ponen en duda esta hipótesis e introducen nuevos aspectos en el examen del origen de este haplotipo.

Hasta la fecha hay varias hipótesis que tratan de explicar la AC en el almendro. Según Grasselly y Olivier (1976), la autocompatibilidad pudo aparecer como consecuencia de una mutación natural en el locus *S*, manteniéndose posteriormente gracias a la selección de los agricultores para su cultivo. Otros autores sugieren que la autocompatibilidad aparecería en almendro a partir de una formación espontánea de híbridos interespecíficos entre la especie cultivada y la especie silvestre *P. webbii* (Spach) Vierh. (Godini, 1979; Reina et al., 1985; Socias i Company, 1984 y 1990). Recientemente, Bošković et al. (2007) sugirieron que la presencia del alelo de autocompatibilidad S_f, se produciría como consecuencia de la mutación puntual de una arginina por una histidina en la región C2 de la *S*-RNasa. Sin embargo, estos autores reconocieron posteriormente un error de secuenciación, por lo que esta hipótesis queda completamente invalidada.

Los resultados obtenidos en esta tesis indican que la actual hipótesis de una naturaleza monogénica de la AC/AI en el almendro, considerada como un carácter

cualitativo (Socias i Company, 1984), puede ser cuestionada. Los otros dos genes encontrados en el grupo 6 y 8 indican que la AI en almendro y posiblemente en las plantas en general es un carácter cuantitativo (Good-Avila et al., 2008). Esta hipótesis que ha sido realzada en este trabajo fue propuesta por de Nettancourt (1977), el cual interaccionaba al mecanismo de AI con otros genes modificadores no ligados al locus *S*. Este fenómeno se conoce como pseudo-autocompatibilidad (PSC), el cual fue propuesto por primera vez en el almendro por Socias i Company (1990).

Por otro lado, todo parece indicar que otros factores epigenéticos podrían influir en el mecanismo de AI de las plantas. Parece ser que la metilacion podría ser responsable de las dos versiones del alelo S_f en el almendro. En primer lugar, el haplotipo S_f habría surgido en la forma activa, siendo AI (al igual que ocurre en los otros haplotipos). Sin embargo, debido a los cambios epigenéticos y a consecuencia de la metilación del ADN, el mismo haplotipo habría sufrido una transformación dando lugar a la nueva forma, la AC. Ello invalida la hipótesis de Bošković et al. (2007) y conjuga la de Grasselly and Olivier (1976) de una mutación, con la de Godini (1979), Reina et al. (1985) y Socias i Company (1984 y 1990) de la transmisión desde *P. webbii*, en cuanto esta mutación pudo haber ocurrido en esta última especie y pudo transmitirse al almendro cultivado, especialmente en la zona italiana de la Apulia.

11. CONCLUSIONES GENERALES

11. CONCLUSIONES GENERALES

1. La distribución fenotípica de la autocompatibilidad observada en la familia 'Vivot' (AI) × 'Blanquerna' (AC) no coincide con ninguna de las posibilidades descritas hasta ahora, tanto para el caso que las variedades no compartan ningún alelo de autoincompatibilidad (50% AC / 50% AI) como para el caso de que compartan un alelo de autoincompatibilidad (100% AC), lo que implica la presencia en esta población de algún factor distorsionante en la transmisión de la autocompatibilidad .

2. 'Vivot' es una variedad AI, como ha mostrado la parada de los tubos polínicos en la mitad del estilo en flores autopolinizadas en laboratorio y la ausencia de frutos cuajados en ramas embolsadas en campo, pero sin embargo presenta el alelo S_f, tradicionalmente ligado a la presencia de la AC.

3. La distribución genotípica en la descendencia 'Vivot' ($S_{23}S_{fa}$) × 'Blanquerna' (S_8S_{fi}) no fue la esperada 1 $S_{23}S_8$:1 $S_{23}S_{fi}$:1 $S_{fa}S_8$:1 $S_{fa}S_{fi}$, sino que pasó a ser 1 S_8S_{23}:1 S_8S_{fa}, y, segregación que sólo se explica si el alelo S_f de 'Vivot' se expresa en el pistilo como un alelo de autoincompatibilidad que detiene el crecimiento de los tubos polínicos con genotipo S_f.

4. La expresión del alelo S_f en el pistilo de 'Vivot' es una S-RNasa activa que produce una reacción de incompatibilidad con los tubos polínicos S_{fi}. Se propone la denominación de este haplotipo del alelo S_f como S_{fa}, frente al alelo que mantiene la autocompatiblidad que se denominará S_{fi}.

5. Se ha propuesto un modelo en el que un gen modificador estaría involucrado en esta familia. Dicho gen seria recesivo y actuaría únicamente en el pistilo.

6. La construcción del mapa genético de 'V×B' ha permitido localizar dos QTLs implicados en el mecanismo de AI, los cuales han sido localizados en los grupos de ligamiento 6 y 8.

7. Una selección asistida por medio de los marcadores SSR BPPCT025 y CPSCT018 daría una probabilidad del 85% para discriminar las plantas AC en aquellos individuos que provengan de un cruzamiento en el que algún parental tuviese el alelo S_{fa}.

8. Cuando la secuencia 'upstream' de la S_f-RNasa presenta alguna citosina, indica que ese alelo es metilado, por lo tanto se produce una inhibición de la expresión, traduciéndose en una inactivación del gen, siendo pues AC (como es el caso de 'Blanquerna'). En el caso de que todas las citosinas se conviertan a timinas, después del tratamiento con el bisulfito, indicaría que la secuencia no es metilada, por lo tanto la RNasa permanecería activa, y la planta seria AI (es el caso de 'Vivot').

9. Las estructuras tridimensionales de las RNasas (S_f, S_8 y S_{23}) forman parte de las de la familia de las RNasas T_2. Todas las estructuras determinadas presentan seis hélices α y seis laminas β.

10. La estructura 3D de la S_f-RNasa presenta un lazo muy superior al de las otras dos RNasas AI, por lo que los amino ácidos de ese lazo, situados entre las regiones conservadas RC4 y C5, podrían formar el grupo de residuos que provoca esa diferencia estructural en las RNasas AI e AC.

11. Todos estos resultados apuntan a que la AC en el almendro puede estar controlada por un complejo de loci con un diferente nivel de expresión. Probablemente el locus *S* sea el de mayor efecto, mientras que los otros modificarían en mayor o menor nivel su expresión. En esta tesis se ha avanzado por primera vez en una especie del género *Prunus* en aspectos como la metilación y la estructural tridimensional de las proteínas, campos en los que necesariamente se deberá profundizar.

12. BIBLIOGRAFÍA

12. BIBLIOGRAFÍA

Almeida CR Marques de (1945). Ácêrca da incompatibilidade na amendoeira. An Inst. Agron. Lisb. 15:1-186.

Alonso JM (2004). Efecto de la consanguinidad sobre la expresión y la transmisión de la autocompatibilidad en el almendro. Tesis Doctoral, Univ. Lleida.

Alonso JM and Socias i Company R (2005). Self-incompatibility expression in self-compatible almond genotypes may be due to inbreeding. J Amer Soc Hort Sci 130:865-869.

Alonso JM and Socias i Company R (2007). Negative inbreeding effects in tree fruit breeding: self-compatibility transmission in almond. Theor Appl Genet. 115:151-158.

Aranzana MJ, Garcia-Mas J, Carbó J, Arús P (2002). Development and variability analisis of microsatellites markers in peach. Plant Breed. 121:87-92.

Arús P, Ballester J, Jáuregui B, Joobeur T, Truco MJ, Vicente MC de (1999). The European *Prunus* mapping project: update of marker development in almond. Acta Hort. 484:331-336.

Arús P, Yamamoto T, Dirlewanger E, Abbott AG (2005). Synteny in the Rosaceae. Plant Breed Rev. 27:175-211.

Ballester J, Bošković R, Batlle I, Arús P, Vargas FJ, Vicente MC de (1998). Location of the self-compatibility gene in the almond linkage map. Plant Breed 117:69-72.

Barrett SCH (1992). The evolution of plant sexual diversity. Nat. Rev. Genet 3:274-284.

Batlle I, Ballester J, Bošković R, Romero MA, Tobutt K and Vargas FJ (1997). Use of stylar ribonucleases in almond breeding to design crosses and select self-compatible seedlings. FAO-CIHEAM Network on Nuts, Nucis Newsletter, 4: 5-6.

Ben Njima N, and Socias i Company R (1995). Characterization of some self-compatible almonds. Pollen tube growth. HortScience 30(2):318-320.

Bender J (2004). DNA Methylation and epigenetics. Annu. Rev. Plant Biol. 55:41–68.

Bird AP (1987). CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3. 342-347.

Bošković R. and Tobutt KR (1996).Correlation of stylar ribonuclease zymograms with incompatibility alleles in sweet cherry. Euphytica 90: 245-250.

Bošković R, Tobutt KR, Batlle I, Duval H (1997). Correlation of ribonuclease zymograms and incompatibility genotypes in almond. Euphytica 97: 167-176.

Bošković R, Tobutt KR, Duval H, Batlle I, Dicenta F and Vargas FJ (1999). A stylar ribonuclease assay to detect self-compatible seedlings in almond progenies. Theor. Appl. Genet. 99:800-810.

Bošković R, Tobutt KR, Batlle I, Duval H, Martinez-Gómez P and Gradziel TM (2003). Stylar ribonucleases in almond: correlation with and prediction of incompatibility genotypes. Plant Breeding 122:70-76.

Bošković R, Tobutt, KR, Ortega E, Sutherland BG and. Godini A (2007). Self-(in)compatibility of the almonds *P. dulcis* and *P. webbii*: detection and cloning of 'wild-type S_f' and new self-compatibility alleles encoding inactive S-RNases. Mol Genet Genomics. 278, 665-676.

Branden C, and Tooze J (1998). Introduction to Protein Structure, 2nd Edition, Garland Publishing, New York, 410 p.

Burgos L, Pérez-Tornero O, Ballester . and Olmas O (1998). Detection and inheritance of stylar ribonucleases associated with self-incompatibility alleles in apricot. Sex. Plant Reprod. 11:153-158.

Busslinger M, Hurst J and Flavell RA (1983). DNA methylation and the regulation of globin gene expression. Cell 34 197–206.

Byrne DH (1990). Isozyme variability in four diploid stone fruits compared with other woody perennial plants. J Hered 81:68-71.

Cachi A and Wünsch A (2009). Bulked segregant analysis for the identification of molecular markers linked to self-compatibility in 'Cristobalina' sweet cherry. Acta Hort. 814:395-400.

Cantini C, Iezzoni AF, Lamboy WF, Boritzki M and Struss D (2001). DNA fingerprinting of tetraploid cherry germplasm using SSR. J Am Soc Hort Sci. 126:205-209.

Cerdá Juan D (1973). Economía antigua de Mallorca. En: J. Mascaró Pasarius (ed) Historia de Mallorca, vol 1. J. Mascaró Pasarius, Palma de Mallorca, pp 193-200.

Certal AC, Almeida RB, Boskovik R, Oliveira MM and Feijo JA (2002). Structural and molecular analysis of self-incompatibility in almond (*Prunus dulcis*). Sex. Plant Reprod. 15:13-20.

Channuntapipat C, Sedgley M and Collins G (2001). Sequences of the cDNAs and genomic GNAs encoding the S_1, S_7, S_8 and S_f alleles from almond, *Prunus dulcis*. Theor Appl Genet. 103:1115-1122.

Channuntapipat C, Wirthensohn M, Ramesh SA, Batlle I, Arús P, Sedgley M and Collins G (2003). Identification of incompatibility genotypes in almond using specific primers based on the introns of the S-alleles. Plant Breed. 122:164-168.

Cheng JH, Han ZH, Xu XF and Li TZ (2006). Isolation and identification of the pollen expressed polymorphic F-box genes linked to the *S*-locus in apple (*Malus* x *domestica*). Sex Plant Reprod. 19: 175-183.

Cipriani G, Lot G, Huang HG, Marrazzo MT, Peterlunger E, Testolin R (1999). AC/GT and AG/CT microsatellite repeats in peach (*Prunus persica* (L.) Batsch): isolation, characterization and cross-species amplification in *Prunus*. Theor Appl Gen. 99:65-72.

Clarke JB, Sargent DJ, Bošković R, Belaj A, Tobutt KR (2009). A cherry map from the inter-specific cross *Prunus avium* 'Napoleon' x *P. nipponica* based on microsatellites, gene-specific and isoenzyme markers. Tree Gen Genomes. 5:41-51.

Cooperativas Agro-Alimentarias (2009). http://www.agro-alimentarias.coop.

Cruz-Garcia F, Hanckock CN, McClure B (2003). *S*-RNase complexes and pollen rejection. J Exp Bot. 54:123-130.

Cruz-Garcia F, Hancock N, Kim D, McClure B (2005). Stylar glycoproteins bind to S-RNase in vitro. Plant J. 42:295–304.

Cubas P, Vincent C and Coen E (1999). An epigenetic mutation responsible for natural variation in floral symmetry, Nature. 401: 157–161.

Darwin CR (1876). The effect of cross and self-fertilization in the vegetable kingdom. Appleton, New York, 482 pp.

de Nettancourt D (1977). Incompatibility in angiosperms. Monographs of Theoretical and Applied Genetics 3, Springer-Verlag, Berlin, Heidelberg-NY.

de Nettancourt D (2001). Incompatibility and incongruity in wild and cultivated plants. Springer, Berlin-Heidelberg-NY.

Dicenta F and García JE (1993). Inheritance of self-compatibility in almond. Heredity. 70:313–317.

Dicenta F, Ortega E, Martinez-Gomez P, Sanchez-Perez R, Gambin, M and Egea J (2009). Penta and Tardona: two new extra-late flowering self-compatible almond cultivars. Acta Horticulturae: 814:189-192.

Dirlewanger E, Cosson P, Travaud M, Aranzana MJ, Poizat C, Zanetto A, Arús P, Laigret F (2002). Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet. 105:127-138.

Dirlewanger E, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P (2004). Comparative mapping and marker assisted selection in Rosaceae fruit crops. Proc Natl Acad Sci USA. 101:9891–9896.

Downey LD, Iezzoni AF (2000). Polimorphic DNA markers in cherry are identified using sequences from sweet cherry, peach and sour cherry. J Am Soc Hort Sci. 125:76-80.

Doyle JJ and Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.*, 19: 11-15.

Duval H, Batlle I, Bošković R and Tobutt KR (1998). Confirmation of incompatibility alleles in almond cultivars predicted by analysis of stylar ribonucleases. Cah. Options Méditerr. 33:25-27.

East EM (1934). Norms of pollen tube growth in incompatible matings of self-sterile plants. Proc. Natl. Acad. Sci. USA. 20: 225-230.

Egea J and Garcia JE (1975). Lignes de recherché actuelle sur l'amandier au CEBAS Murcie. Objectifs et premieres observations. 2º Coloquio GREMPA. Montpellier. 1-11.

Egea J, Dicenta F, Berenguer T and Garcia JE (2000). 'Antoñeta' and 'Marta' almonds. Hortscience. 35:1358-1359.

Engh R.A and Huber R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. Acta Cryst. A, 47, 392-400.

Entani T, Iwano M, Shiba H, Che FS, Isogai A, Takayama S (2003). Comparative analysis of the self-incompatibility (*S*-)locus region of *Prunus mune*: Identification of a pollen-expressed F-box gene with allelic diversity. Genes Cells 8:203-213.

FAO (2009 and 2010). Los datos de FAOSTAT, http://faostat.fao.org.

Fedorov A A (1969). Chromosome numbers of flowering plants. Leningrad.

Felipe AJ (2000). El almendro: El material vegetal. Integrum, Lérida. 461pp.

Felipe AJ (1977). Almendro: estados fenológicos. Inf. Técn. Econ. Agrar. 8(27):8-9.

Felipe AJ and Socias i Company R (1985). L'amelioration genetique de l'amandier à Saragosse. Options Méditerranéennes. Seri etudes. 1, 9-14.

Felipe AJ and Socias i Company R (1987). 'Aylés', 'Guara' and 'Moncayo' almonds. Horscience. 22: 961-962.

Fernández i Martí A, Alonso JM, Kodad O, Socias i Company R (2008). Differential phenotypic expression of self-compatibility in almond. 4th International Rosaceae Genomics Conference (RGC4). Pucon, Chile. March 2008.

Fernández i Martí A, Hanada T, Alonso JM, Yamane H, Tao R, Socias i Company R (2009a). A modifier locus affecting the expression of the *S*-RNase gene could be the cause of breakdown of self-incompatibility in almond. Sex Plant Reproduction 22:179-186.

Fernández i Martí A, Alonso JM, Espiau MT, Rubio-Cabetas MJ, Socias i Company R (2009b). Genetic diversity in Spanish and foreign almond germplasm assessed by molecular characterization with SSRs. J Am Soc Hort Sci. 134: 535-542.

Fernández i Martí A, Hanada T, Alonso JM, Yamane H, Tao R, Socias i Company R (2010). The almond S_f haplotype shows a double expression despite its comprehensive genetic identity (in press).

Ficker M, Kirch HH, Eijlander R, Jacobsen E and Thompson RD (1998). Multiple elements of the S_2-RNase promoter from potato (*Solanum tuberosum* L.) are required for cell type-specific expression in transgenic potato and tobacco. Mol. Gen. Genetic 257: 132-142.

Gehring M and Henikoff S (2007). DNA methylation dynamics in plant genomes. Biochimica et Biophysica Acta. 1769: 276–286.

Gingrich JC, Boehrer DM, Garnes JA, Johnson W, Wong BS, Bergmann A, Eveleth GG, Langlois RG and Carrano AV (1996). Construction and characterization of human chromosome 2-specific cosmid, fosmid and PAC clone libraries. Genomics. 32: 65-74.

Godini A (1979). Ipotesi sulla comparsa dell´autocompatibilità nel mandorlo. Riv. Sci. Tecn. Agrar. 19(2/3):3-10.

Goldraij A, Kondo K, Lee CB, Hancock CN, Sivaguru M, Vazquez-Santana S, Kim S, Phillips TE, Cruz-Garcia F and. McClure B (2006).

Compartmentalization of *S*-RNase and HT-B degradation in self-incompatible *Nicotiana*. Nature. 439 : 805-810.

Good-Avila SV, Mena-Ali JI, Stephenson AG (2008). Genetic and environmental causes and evolutionary consequences of variations in self-fertility in self-incompatible species. In VE Franklin-Tong (Ed) Self-incompatibility in flowering plants. Springer, Berlin, pp 33-47.

Grasselly, C (1972). L'amandier; caractères morphologiques et physiologiques des variétés, modalité de leurs transmissions chez les hybrids de première generation. Thesis Univ. Bordeaux.

Grasselly C (1976). Origine et évolution de l'amandier cultivé. Options Mediterr. 32: 45-50.

Grasselly C and Olivier G (1976). Mise en évidence de quelques types autocompatibles parmi les cultivars d´amandier (*P. amygdalus* Batsch) de la population des Pouilles. Ann. Amélior. Plant. 26:107-113.

Grasselly C and Crossa Raynaud P (1980). L'amandier. G.P. Maisonneuve et Larose, Paris. 446pp.

Grasselly C, Crossa- Raynaud P, Olivier G and Gall H (1981). Transmission du caractère d´autocompatibilité chez l´amandier (*Amygdalus communis*). Options Méditerr. CIHEAM/IAMZ. 81/I:71-75.

Graziano E and Arús P (2002). FITMAPS and SHOWMAP: Two programs for graphical comparison and plotting of genetic maps. J Hered. 93:225-227.

Gregory D (2004). The development of a genetic linkage map for almond based on molecular and agronomic markers. PhD thesis. The University of Adelaide, Australia.

Gruenbaum Y, Naveh-Many T, Cedar H, Razin A (1981). Sequence specificity of methylation in higher plant DNA. Nature. 292:860–862.

Gupta PK, Balyan HS, Sharma PC, Ramesh B (1996). Microsatellites in plants: a new class of molecular markers. Curr Sci.70:45-54.

Halasz J, Fodor A, Pedryc A and Hegedus A (2009). S-genotyping of Eastern European almond cultivars: identification and characterization of new (S36–S39) self-incompatibility ribonuclease alleles. Plant breeding. 129: 227-232.

Hanada T, Fukuta K, Yamane H, Esumi T, Gradziel TM, Dandekar AM, Fernández i Martí A, Alonso JM, Socias i Company R and Tao R (2009).

Cloning and characterization of self-compatible S_f haplotype in almond (*Prunus dulcis* (Mill.) D.A. Webb.). HortScience. 44: 609-613.

Hauck NR, Yamane H, Tao R, Iezzoni A (2006). Accumulation of nonfunctional *S*-haplotypes results in the breakdown of gametophytic self-incompatibility in tetraploid *Prunus*. Genetics. 172:1191–1198.

Herman JG, Latif F, Weng Y, Lerman MI, Zbar B, Liu S, Samid, D, Duan DS, Gnarra JR and Linehan WM (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci U S A. 91(21):9700-9704.

Hinata K, Watanabe M, Yamakawa S, Satta Y, Isogai A (1995). Evolutionary aspects of the *S*-related genes of the *Brassica* self-incompatibility system: synonymous and nonsynonymous base substitutions. Genetics. 140:1099-1104.

Holliday R and Pugh J E (1975). DNA modification mechanisms and gene activity during development. Science. 187, 226–232.

Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005). Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. Genetics. 171:1305-1309.

Hrmova M and Fincher GB (2009). Functional genomics and structural biology in the definition of gene function. In: Plant Genomics, Humana Press Inc., Totowa, NJ, USA (Somers D, Langridge P, Gustafson P**,** eds). Methods in Molecular Biology. 513: 199-227. Invited review.

Hua Z and Kao T (2006). Identification and characterization of components of a putative *Petunia S*-locus F-Box-containing E3 ligase complex involved in *S*-RNase-based self-incompatibility. Plant Cell. 18:2531-2553.

Huang J, Zhao L, Yang Q. and Xue Y (2006). AhSSK1, a novel SKP1-like protein that interacts with the S-locus F-box protein SLF. Plant J. 46:780–793.

Ida K, Norioka S, Yamamoto M, Kumasaka T, Yamashita E, Newbigin E, Clarke AE, Sakiyama F and Sato M. (2001). The 1.55 A resolution structure of Nicotiana alata S_{F11}-RNase associated with gametophytic self-incompatibility. J Mol Biol. 314:103-12.

Ikeda K, Igic B, Ushijima K, Yamane H, Hauck NR, Nakano R, Sassa H, Iezzoni AF, Kohn JR and Tao R (2004). Primary structural features of the *S* haplotype-specific F-box protein, SFB, in *Prunus*. Sex. Plant Reprod. 16, 235-243.

Ishizaka T, Nakano H, Suzuki T and Kitashiba H (2003). Characterization of the *S*-RNase promoters from sweet cherry (*Prunus avium L.*). Genes Genet. Syst. 78, 191-194.

Jeanish R and Bird A (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33:S245–S254.

Johnson TB and Coghill RD (1925). Research on pyrimidines, C111. The discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus. Journal of Am. Chem. Soc. 47: 2838-2844.

Jones PA and Takai D (2001). The role of DNA methylation in mammalian epigenetics. Science. 293:1068–1070.

Joobeur T, Viruel MA, Vicente MC de, Jáuregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arús P (1998). Construction of a saturated linkage map for *Prunus* using an almond x peach F_2 progeny. Theor Appl Genet. 97:1034-1041.

Kao TH and McCubbin A G (1996). How flowering plants discriminate between self and non-self pollen to prevent inbreeding. Proc. Natl. Acad. Sci. USA 93, 12059-12065.

Kao TH and T Tsukamoto (2004). The molecular and genetic bases of *S*-RNase based self-incompatibility. Plant Cell 16, suppl., 72-83.

Kester DE and Asay RN (1975). Almonds. In: Advances in fruit breeding. Editores: janick J y Moore JN. Purdue University Press, West Lafayette. Indiana. Pp. 387-419.

Kester DE, Gradziel, TM and Grasselly C (1990). Almond. P. 699-758. In: j.N. Moore and J.R. Ballington, jr. (eds). Genetic resources of temperate fruit and nut crops. Int. Hort. Sci., Wageningen, The Netherlands.

Kester DE, Gradziel TM and Micke WC (1994). Identifying pollen incompatibility groups in California almond cultivars. J. Amer. Soc. Hort. Sci. 119 (1):106-109.

Kester DE and Gradziel TM (1998). The university of California almond breeding programme: I. Historical aspects. FAO-Nucis-Newsletter, 7:8-10.

Kim UJ, Shizuya H, de Jong PJ, Birren B and Simon MI (1992). Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res. 20: 1083-1085.

Kleywegt GJ and Jones TA (1996). Phi/Psi-chology: Ramachandran revisited. Structure 4: 1395-1400.

Kodad O y Socias i Company R (2005). Daños diferenciales por heladas en flores y frutos y criterios de selección para la tolerancia a heladas en el almendro. Inf. Técn. Econ. Agrar. 101:349-365

Kodad O y Socias i Company R (2006). Pollen source effect on pollen tube growth in advanced self-compatible almond selections (Prunus amygdalus Batsch). Adv. Hort. Sci. 20:256-261.

Kodad O and Socias i Company R (2008 Fruit set evaluation for self-compatibility selection in almond. Scientia Hort. 118:260-265.

Kodad O, Alonso JM, Sánchez A, Oliveira MM and. Socias i Company R (2008a). Evaluation of genetic diversity of *S*-alleles in an almond germplasm collection. J. Hort. Sci. Biotechnol. 83: 603-608.

Kodad O, Sánchez A, Saibo N, Oliveira MM and Socias i Company R (2008b). Identification and characterization of new S-alleles associated with self-incompatibility in almond. Plant Breed. 127(6): 632-638.

Kodad O, Socias i Company R, Sánchez A and Oliveira MM (2009). The expression of self-compatibility in almond may not only be due to the presence of the S_f allele. J Am Soc Hort Sci. 134:221-227.

Kodad O, Alonso JM, Fernández i Martí A, Oliveira MM and Socias i Company R (2010). Molecular and physiological identification of new *S*-alleles associated with self-(in)compatibility in local Spanish almond cultivars. Scientia Hortic. 123, 308-311.

Kondo K, Yamamoto M, Itahashi R, Sato T, Egashira H, Hattori T and Kowyama Y (2002). Insights into the evolution of self-compatibility in Lycopersicon from a study of stylar factors. Plant J. 30, 143–153.

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

Kruczek I and Doerfler W (1983). Expression of the chloramphenicol acetyltransferase gene in mammalian cells under the control of adenovirus type 12 promoters: effect of promoter methylation on gene expression, Proc. Natl.Acad. Sci. U. S. A. 80: 7586–7590.

Ladizinsky G (1999). On the origin of almond. Genet. Resour. Crop Evol. 46: 143-147.

Lai Z, Ma W, Han B, Liang L, Zhang Y, Hong G, Xue Y (2002). An F-box gene linked to self-incompatibility (*S*) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol Biol. 50:29-42.

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

Laskowski RA, McArthur MW, Moss DS and Thornton JM (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J Appl. Cryst. 26, 283-291.

Lewis CI and Crowe LK (1954). The induction of self-fertility in tree fruits. J Hort Sci. 29:220-225.

Lind JL, Bönig I, Clarke AE and Anderson MA (1996). A style-specific 120kDa glycoprotein enters pollen tube of *Nicotiana alata in vivo*. Sex. Plant Reprod. 9: 75-86.

Linskens HF and Esser K (1957). Über eine spezifische Anfärbung der Pollenschläuche im Griffel und die Zahl der Kallosepfropfen nach Selbstung und Fremdung. Naturwissenschaften. 44:16-21.

Lippman ZB, May C, Yordan T, Singer R and Martienssen R (2003). Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification, PLoS Biol. 1-E67.

López M, Mnejja M, Rovira M, Collins G, Vargas FJ, Arús P and Batlle I (2004). Self-incompatibility genotypes in almond reevaluated by PCR, stylar ribonucleases, sequencing analysis and controlled pollinations. Theor Appl Genet. 109: 954–964.

López M, Vargas FJ and Batlle I (2006). Self-(in)compatibility almond genotypes: A review. Euphytica. 150: 1–16.

Luo S and Preuss D (2003). Strand-biased DNA methylation associated with centromeric regions in Arabidopsis. Proc Natl Acad Sci. 100(19):11133-11138.

Ma RC and Oliveira MM (2001). Molecular cloning of the self-compatibility genes S1 and S3 from almond (*Prunus dulcis* cv. Ferragnès). Sex Plant Reprod. 14:163-167.

Marchese A, Bošković RI, Caruso T, Raimondo A, Cutuli M, Tobutt K (2007). A new self-compatibility haplotype in the sweet cherry 'Kronio', S_5', attributable to a pollen-part mutation in the *SFB* gene. J Exper Bot. 58.4347-4356.

Martin A, Troadec C, Boualem A, Rajab M, Fernández R, Morin H, Pitrat M, Dogimont C and Bendahmane A (2009). A transposon-induced epigenetic change leads to sex determination in melon. Nature. 461: 1135-1138.

Martínez García PJ (2009). Mejora genética del almendro [*Prunus dulcis* (Miller) D.A. Webb]: Aspectos agronómicos y moleculares de la incompatibilidad floral y su incidencia sobre la fructificación. PhD Thesis, Univ. Murcia, Spain

Matsumoto D, Yamane H and Tao R (2008). Characterization of *SLFL1*, a pollen-expressed F-box gene located in the *Prunus S* locus. Sex.Plant Reprod. 21: 113-121.

Matsuura T, Sakai H, Unno M, Ida K, Sato M, Sakiyama F, Norioka S (2001). Crystal structure at 1.5-A resolution of *Pyrus pyrifolia* pistil ribonuclease responsible for gametophytic self-incompatibility. J Biol Chem. 276(48):45261-9.

McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE (1989). Style self incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature. 342: 955-957.

McClure BA, Gray JE, Anderson MA, Clarke AE (1990). Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature. 347:757-760.

McClure BA, Mou B, Canevascini S, Bernatzky R (1999). A small asparagine-rich protein required for *S*-allele specific pollen rejection in *Nicotiana*. Proc Natl Sci USA. 96:13548-13553.

McClure BA and Franklin-Tong V (2006). Gametophytic self-incompatibility: understanding the cellular mechanisms involved in 'self' pollen tube inhibition. Planta. 224:233–245.

McCubbin AG, Lee C and Hetrick A (2006). Identification of genes showing differential expression between morphs in developing flowers of *Primula vulgaris*. Sex Plant Reprod. 19: 63-72.

Millar DS, Ow KK, Paul CL, Russell PJ, Molloy PL and Clark SJ (1999). Detailed methylation analysis of the glutathione *S*-transferase pi (GSTP1) gene in prostate cancer. Oncogene. 18:1313–1324.

Mnejja M, García-Más M, Howad W, Badenes ML, Arús P (2004). Simple sequence repeat (SSR) markers of Japanese plum (*Prunus salicina* Lindl.) are highly polymorphic and transferable to peach and almond. Mol Ecol Notes. 4:163-166.

Moriya Y, Okada K, Yamamoto K, Iwanami H, Bessho H, Takasaki-Yakuda T (2009). Characterisation of partial self-compatibility in the European pear cultivar, 'Grand Champion'. J Hort Sci Biotechnol. 84:77-82.

Numata T, Suzuki A, Kakuta Y, Kimura K, Yao M, Tanaka I, Yoshida Y, Ueda T, Kimura M (2003). Crystal structures of the ribonuclease MC1 mutants N71T and N71S in complex with 5'-GMP: structural basis for alterations in substrate specificity. Biochemistry. 13:42 (18): 5270-8.

Ortega E and Dicenta F (2003). Inheritance of self-compatibility in almond: breeding strategies to assure self-compatibility in the progeny. Theor. Appl. Genet. 106:904-911.

Ortega E, Sutherland B, Dicenta F, Bošković R and Tobutt KR (2005). Determination of incompatibility genotypes in almond using first and second intron consensus primers: detection of new *S* alleles and correlation of reported *S* genotypes. Plant Breeding. 124: 188-196.

Ortega E, Bošković R, Sargent DJ and Tobutt KR (2006). Analysis of *S*-RNase alleles of almond (*Prunus dulcis*): characterization of new sequences, resolution of synonyms and evidence of intragenic recombination. Mol. Gen. Genomics. 276: 413-426.

Popov MG, Kostina KF and Poyarkova AI (1929). Árboles y arbustos silvestres de Asia Central (en ruso). Tr. Prikl. Bot. Genet. Selek. 2: 241-483.

Procheck: http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html

Prosa2003: https://prosa.services.came.sbg.ac.at/download/download.php

PyMol: http://pymol.sourceforge.net

Qiao H, Wang H, Zhao L, Zhou, Huang J, Zhang Y and Xue Y (2004). The F-Box protein AhSLF-S_2 physically interacts with *S*-RNases that may be inhibited by the Ubiquitin/26Proteasome pathway of protein degradation during compatible pollination in *Antirrhinum*. The Plant Cell. 16: 582-595.

Ramachandran GN, Ramakrishnan C and Sasisekharan V (1963). Stereochemistry of polypeptide chain configurations. J Mol Biol. 1963 7:95-9.

Reina A, Giorgio V and Godini A (1985). Autres types autocompatibles parmi la population d´amandier des Pouilles. Options Méditerr. CIHEAM/IAMZ 85/I:25-29.

Riggs AD (1975). Cytogenet. Cell Genet. 14(1): 9–25.

Rikhter AA (1969). Formas y métodos de mejora del almendro. Tr. Gos. Nikit. Bot Sad 43:81-94.

Romero C, Vilanova S, Burgos L, Martinez-Calvo J, Vicente M, Llácer G and Badenes ML (2004). Analysis of the *S*-locus structure in *Prunus armeniaca* L. Identification of S-haplotype specific *S*-RNase and F-box genes. Plant Mol. Biol. 56, 145-157.

Royo J, Kunz C, Kowyama Y, Anderson M, Clarke AE and Newbigin E (1994). Loss of a histidine residue at the active site of S-locus ribonuclease is associated with self-compatibility in *Lycopersicon peruvianum*. Proc. Natl. Acad. Sci. USA. 91: 6511-6514.

Rubí V (1980). El almendro. Delegación de Baleares, Ministerio de Agricultura, Palma de Mallorca, Spain.

Sali A. and Blundell TL (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234: 779–815.

Sánchez-Pérez R, Howad W, Dicenta F, Arús P, Martínez-Gómez P (2007). Mapping major genes and quantitative trait loci controlling agronomic traits in almond. Plant Breed. 126:310-318.

Sassa H, Hirano H. Nishio T, Koba T (1997). Style-specific self-incompatibility mutation caused by deletion of the *S*-RNase gene in Japanese pear (*Pyrus serotina*). Plant J. 12:223–227.

Sassa H, Kakui H, Miyamoto, Suzuki Y, Hanada T, Ushijima K, Kusaba M., Hirano H and Koba T (2007). *S* locus F-box brothers: Multiple and pollen specific F-box genes with *S* haplotype-specific polymorphisms in apple and japanese pear. Genetics. 175: 1869-1881.

Schopfer CR, Nasrallah ME and Nasrallah JB (1999). The male determinant of self-incompatibility in *Brassica*. Science: 266: 1697–1700.

Shore JS, Arbo MM and Fernández A (2006). Breeding system variation, genetics and evolution in the Turneraceae. New Phytol. 171: 539-551.

Shiba H, Kakizaki T, Iwano M, Tarutani Y, Watanabe M, Isogai A and Takayama S (2006). Dominance relationships between self-incompatibility alleles controlled by DNA methylation, Nature Genet. 38:297–299.

Shulaev V, Korban SS, Sosinski B, Abbot A, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arús P, Dandekar AM, Lewers K, Brown SK, Davis TM, Gardiner

SE, Potter D, Veilleux RE (2008). Multiple models for Rosaceae genomics. Plant Physiol 147:985-1003.

Siegfried ZS Eden M, Mendelsohn X, Feng BZ, Tsuberi H and Cedar H (1999). DNA methylation represses transcription in vivo, Nat. Genet. 22: 203–206.

Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S and Kao TH (2004). Identification of the pollen determinant of *S*-RNase-mediated self-incompatibility. Nature. 429: 302-305.

Socias i Company R (1974). Effect of temperature and genotype on pollen tube growth of some self-compatible almond selection derived from peach x almond hybrydation. Thesis, University of California, Davis.

Socias i Company R (1978). La autocompatibilidad en la mejora del almendro. I Congreso Internacional de Almendra y Avellana, Reus, 25-28 octubre 1976, 513-522.

Socias i Company R (1979). Aportación a las técnicas de observación de tubos polínicos. Caso del almendro. An. Inst. Nac. Invest. Agrar. Ser. Prod. Veg. 10:233-236.

Socias i Company R (1984). A genetic approach to the transmission of self-compatibility in almond. Options Méditerr. CIHEAM/IAMZ 84/II: 123-128.

Socias i Company R (1990). Breeding self-compatible almonds. Plant Breed. Rev. 8: 313-338.

Socias i Company R (1998a). La taxonomie de l'amandier. Cah. Options Méditerr. 33:91-93.

Socias i Company R (1998b). Fruit tree genetics at a turning point: the almond example. Theor Appl Genet. 96:588-601.

Socias i Company R (2001). Avances recientes en la autocompatibilidad del almendro. Inf. Técn. Econ. Agrar. 97V(3):215-225.

Socias i Company R (2002). Latest advances in almond self-compatibility. Acta Hort. 591:205-212.

Socias i Company R (2004). The contribution of *Prunus webbii* to almond evolution. Plant Genet. Resour. Newsl. 14:9-13.

Socias i Company R, Kester DE and Bradley MV (1976). Effects of temperature and genotype on pollen tube growth of some self-incompatible and self-compatible almond cultivars. J. Amer. Soc. Hort. Sci. 101:490-493.

Socias i Company R and Felipe AJ (1977). La polinización del almendro. CRIDA 03, Zaragoza, 29 pp.

Socias i Company R and Felipe AJ (1987). Pollen tube growth and fruit set in a self-compatible almond selection. HortScience. 22:113-116.

Socias i Company R and Felipe AJ (1988). Self-compatibility in almond: transmission and recent advances in breeding. Acta Hort. 224:307-317.

Socias i Company R and Felipe AJ (1992). Almond: a diverse germplasm. HortScience. 27:717-718..

Socias i Company R and Felipe AJ (1999). Recent advances in breeding for autogamy in almond. Cah. Options Méditerr. 56: 65-69.

Socias i Company R, Felipe AJ and Gómez Aparisi J (1999). A major gene for flowering time in almond. Plant Breeding. 118: 443-448.

Socias i Company R, Gómez Aparisi J and Alonso JM (2005). Year and enclosure effects on fruit set in an autogamous almond. Scientia Horticulturae. 104: 369-377.

Socias i Company R and Felipe AJ (2007). 'Belona' and 'Soleta', two new autogamous almonds. Nucis. 13: 12-15.

Socias i Company R, Kodad O, Alonso JM and Felipe AJ (2008). 'Mardía' almond. HortScience. 43: 2240-2242.

Socias i Company R, Alonso JM, Kodad O, Gradziel TM (2010). Almond [*Prunus dulcis* syn. *P. amygdalus*]. In: ML Badenes and D Byrne (Eds) Fruit Breeding, Springer Verlag, Berlin (in press).

Sonneveld T, Tobutt KF and Robbins TP (2003). Allele-specific PCR detection of sweet cherry self-incompatibility (*S*) alleles S_1 to S_{16} using consensus and allele-specific primers. Theor. Appl. Genet. 107: 1059-1070.

Sonneveld T, Tobbut KR, Vaughan SP. and Robbins TP (2005). Loss of pollen-*S* function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of an *S* haplotype-specific F-box gene. Plant Cell. 17:35-51.

Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJ (2000). The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. Mol. Cell. 6: 791–802.

Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000). Characterization of microsatellite markers in peach (*Prunus persica* (L.) Batsch). Theor Appl Genet. 101:421-428.

Sutherland BG, Robbins TP and Tobutt KR (2004). Primers amplifying a range of *Prunus S*-alleles. Plant Breed. 123: 582-584.

Takayama S, Shiba H, Iwano M, Shimosato H, Che FS, Kai N, Watanabe M, Suzuki G, Hinata K and. Isogai A (2000).The pollen determinant of self-incompatibility in *Brassica campestris*." Proc. Natl Acad. Sci USA. 97: 1920–1925.

Tamura M, Ushijima K, Sassa H, Hirano H, Tao R, Gradziel TM and Dandekar AM (2000) Identification of self-incompatibility genotypes of almond by allele-specific PCR analysis. Theor. Appl. Genet. 101:344-349.

Tao R, Yamane H, Sassa H, Mori H, Gradziel TM, Dandekar AM, and Sugiura A (1997). Identification of stylar RNases associated with gametophytic self-incompatibility in almond (*Prunus dulcis*). Plant Cell Physiol. 38: 304-311.

Tao R., Yamane H, Sugiura A, Murayama H, Sassa H and Mori H (1999a). Molecular typing of S alleles trough identification, characterization and cDNA cloning for s-RNases in sweet cherry. J. Amer. Soc. Hort. Sci. 124: 224-233.

Tao R, Yamane H and Gugiura A (1999b). Cloning and nucleotide sequence of cDNAs encoding S_1- and S_4-RNases from sweet cherry (*Punus avium* L.) Plant Physiol. 120: 1207-1213.

Tao R, Watari A, Hanada T, Habu T, Yaegaki H, Yamaguchi M, Yamane H (2007). Self-compatible peach (*Prunus persica*) has mutant versions of the *S* haplotypes found in self-incompatible *Prunus* species. Plant Mol Biol. 63:109-123.

Tasei JN (1975). La pollinisation des arbres fruitiers. Arboric. Fruit. 261 : 14-15.

Testolin R, Marrazzo T, Cipriani G, Quarta R, Verde I, Dettori T, Pancaldi M, Sansavini S (2000). Microsatellite DNA in peach (*Prunus persica* (L.) Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. Genome. 43:512- 520.

Thompson KF, Taylor JP (1971). Self-compatibility in kale. Heredity 27:459-471.

Thompson JD, Gibson TJ Plewniak F and Higgins DG (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882.

Tsukamoto T, Ando T, Kokubun H, Watanabe H, Sato T, Masada M, Marchesi E, Kao TH (2003). Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by a modifier locus that suppresses the expression of an *S*-RNase gene. Sex Plant Reprod. 15:255-263.

Tufts WP (1919). Almond pollination. Calif. Agric. Bull. 306, 32 pp.

Tufts WP and Philp GL (1922). Almond pollination. Calif. Agric. Bull. 346, 35 pp.

Ushijima K, Sasss H, Tao R, Yamane H, Dankenkar AM, Gradziel TM., Hirano H (1998). Cloning and characterization of cDNAs encoding S-RNases from almond (Prunus dulcis): primary structural features and sequence diversity of the S-RNases in Rosaceae. Mol.Gen. Genet. 260: 261-268.

Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R and Hirano H (2003). Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. Plant Cell 15: 771-781.

Ushijima K, Yamane H, Watari A, Kakehi E, Ikeda K, Hauck NR, Iezzoni AF and Tao R (2004). The S haplotype-specific F-box protein gene, SFB, is defective in self-compatible haplotypes pf *Prunus avium* and *Prunus mume*. The Plant Journal 39: 573-586.

Van Ooijen JW, Voorrips RE (2001). JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Reasearch International, Wageningen.

Van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2002). MapQTL 4.0, software for the calculation of QTL positions on genetic maps. Plant Reasearch International, Wageningen.

Vargas FJ y Romero MA (1994). 'Masbovera', 'Glorieta' and 'Francolí', threenew almond varieties from IRTA. Acta Hort. 373:75-82.

Vargas FJ, Romero MA, Clavé J, Vergés J, Santos J and Batlle I (2008). 'Vayro', 'Marinada', 'Constantí, and 'Tarraco' almonds. HortScience 43:535-537.

Vavilov NI (1930). Wild progenitors of the fruit trees of Turkistan and Caucasus and the problem of the origin og fruit trees. IX International Horticultural Congress report and proceedings. London August 7-15, 1930. Pp 271-286.

Verde I, Lauria M, Dettori T, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo MT, Cipriani G, Hartings H, Testolin R, Abbott AG, Motto M, Quarta R (2005). Microsatellite and AFLP markers in the *Prunus persica* (L.)

Batsch x *P. ferganensis* BC_1 linkage map: saturation and coverage improvement. Theor Appl Genet. 111:1013-1021.

Vilanova S, Badenes ML, Burgos L, Martínez-Calvo J, Llácer G, Romero C (2006). Self-compatibility of two apricot selections is associated with two pollen part mutations of different nature. Plant Physiol.142:629-641.

Voorrips RE (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 93: 77-78.

Wang Y, Tsukamoto T, Yi K, Wang X, Huang S, McCubbin AG and Kao T (2004). Chromosome walking in the *Petunia inflata* self-incompatibility (*S*-) locus and gene identification in an 881-kb contig containing S_2-RNase. Plant Molecular Biology. 54: 727-742.

Watari A, Hanada T, Yamane H, Esumi T, Tao R, Yaegaki H, Yamaguchi M, Beppu B and Kataoka I (2007). A low transcriptional level of S^e*-RNase* in the S^e-haplotype confers self-compatibility in Japanese plum. J. Amer. Soc. Hortic. Sci. 132, 396-406.

Weinbaum SA, Polito VS and Kester DE (1986). Pollen retention following natural self-pollination in peach, almond, and peach × almond hybrids. Euphytica 35:193-200.

Whitehouse HLK (1950). Multiple-allelomorph incompatibility of pollen and style in the evolution of the angiosperms. Annals of Botany. 14: 199-216.

Wünsch A and Hormaza JI (2004). Genetic and molecular analysis in Cristobalina sweet cherry, a spontaneous self-compatible mutant. Sex Plant Reprod. 17:203-210.

Wünsch A (2009). Cross-transferable polymorphic SSR loci in *Prunus* species. Scientia Hort. 120:348-352.

Yamamoto T, Mochida K, Imai T, Shi IZ, Ogiwara I, Hayashi T (2002). Microsatellite markers in peach (*Prunus persica* (L.) Batsch) derived from an enriched genomic library and cDNA libraries. Mol Ecol Notes. 2:298-302.

Yamane H, Ushijima K, Sassa H, Tao R (2003a). The use of *S* haplotype-specific F-box protein gene, *SFB*, as a molecular marker for *S*-haplotypes and self-compatibility in Japanese apricot (*Prunus mume*). Theor Appl Genet .107:1357-1361.

Yamane H, Ikeda K, Ushijima K, Sassa H, Tao R (2003b). A pollen-expressed gene for a novel protein with an F-motif that is very tightly linked to a gene for *S*-

RNase in two species of cherry, *Prunus cerasus* and *P. avium*. Plant Cell Physiol. 44:764-769.

Yamane H and R Tao (2009). Molecular Basis of Self-(in)compatibility and Current Status of *S*-genotyping in Rosaceous Fruit Trees. J. Japan. Soc. Hort. Sci. 78 (2): 137–157.

Yates F (1934). Contingency tables involving small numbers and the Chi-squared test. J R Stat Soc Suppl. 1:217-235.

Zhao L, Huang J, Zhao Z, Li Q, Sims TL and Xue Y (2010). The Skp1-like protein SSK1 is required for cross-pollen compatibility in *S*-RNase-based self-incompatibility. The Plant Journal. 62(1):52-6

Zilberman D and Henikoff S (2005). Epigenetic inheritance in Arabidopsis- selective silence. Curr. Opin. Genet. Dev. 15: 557-62.

Zohary D and Hopf M (1993). Domestication of plants in the old world. Clarendon Press, Oxford, UK.

13. ANEXOS

13. ANEXOS

13.1 Artículos publicados en revistan internacionales durante el transcurso de esta tesis

Sex Plant Reprod
DOI 10.1007/s00497-009-0102-7

ORIGINAL ARTICLE

A modifier locus affecting the expression of the *S*-RNase gene could be the cause of breakdown of self-incompatibility in almond

Angel Fernández i Martí · Toshio Hanada · José M. Alonso · Hisayo Yamane · Ryutaro Tao · Rafel Socias i Company

Received: 11 May 2009 / Accepted: 3 June 2009

Abstract Self-compatibility has become the primary objective of most almond (*Prunus amygdalus* Batsch) breeding programmes in order to avoid the problems related to the gametophytic self-incompatibility system present in almond. The progeny of the cross 'Vivot' ($S_{23}S_{fa}$) × 'Blanquerna' (S_8S_{fi}) was studied because both cultivars share the same S_f allele but have a different phenotypic expression: active (S_{fa}) in 'Vivot' and inactive (S_{fi}) in 'Blanquerna'. In addition, the microscopic observation of pollen tube growth after self-pollination over several years showed an unexpected self-incompatible behaviour in most seedlings of this cross. The genotypes of this progeny showed that the S_{fi} pollen from 'Blanquerna' was not able to grow down the pistils of 'Vivot' harbouring the S_{fa} allele, confirming the active function of this allele against the inactive form of the same allele, S_{fi}. As self-compatibility was observed in some S_8S_{23} and S_8S_{fa} individuals of this progeny, the S_f haplotype may not always be linked to the expression and transmission of self-compatibility in almond, suggesting that a modifier locus may be involved in the mechanism of self-incompatibility in plants.

Communicated by Teh-hui Kao.

A. Fernández i Martí · J. M. Alonso · R. Socias i Company (✉)
Unidad de Fruticultura, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Av. Montañana 930, 50059 Zaragoza, Spain
e-mail: rsocias@aragon.es

T. Hanada · H. Yamane · R. Tao
Laboratory of Pomology, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

Published online: 17 June 2009

Keywords *Prunus amygdalus* Batsch · Self-(in)compatibility · *S* locus · Modifier gene · Transmission

Introduction

Most almond (*Prunus amygdalus* Batsch) cultivars are self-incompatible (Socias i Company 1990), thus cross-pollination is required to ensure a crop. Self-incompatibility (SI) is a cell-to-cell recognition mechanism in flowering plants allowing a pistil to distinguish self from non-self pollen, and plays an important role in promoting genetic diversity through outbreeding. This trait is an important evolutionary mechanism maintaining genetic variability amongst plant populations. In almond, SI is controlled by a single locus, the *S* locus (Socias i Company 1984) which contains two components, one specifically expressed in the style (stylar-*S*), and the other in the pollen (pollen-*S*). The stylar-*S* component encodes an allelic series of stylar glycoproteins with ribonuclease activity, called *S*-RNases (Tao et al. 1997), which are responsible for the inhibition of pollen tube growth through the degradation of the pollen RNA, since *S*-RNases are thought to function as specific cytotoxins (McClure et al. 1999).

The pollen-*S* component has recently been identified as the *S* locus F-box gene, SLF (Lai et al. 2002; Hua and Kao 2006) in Solanaceae and Plantaginaceae. In *Prunus* a good candidate gene for the pollen-*S* component has been named the *S*-haplotype-specific F-box gene (*SFB*), as it is tightly linked to the *S*-RNase gene. This pollen-*S* component has been identified in the *S* locus of several *Prunus* species, including almond (Ushijima et al. 2003), Japanese apricot, *P. mume* (Sieb.) Sieb and Zucc. (Entani et al. 2003; Yamane et al. 2003a), apricot, *P. armeniaca* L. (Romero

J. Amer. Soc. Hort. Sci. 134(5):535–542. 2009.

Genetic Diversity in Spanish and Foreign Almond Germplasm Assessed by Molecular Characterization with Simple Sequence Repeats

Àngel Fernández i Martí, José M. Alonso, María T. Espiau, María J. Rubio-Cabetas, and Rafel Socias i Company[1]

Unidad de Fruticultura, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Av. Montañana 930, 50059, Zaragoza, Spain

Additional index words. *Prunus amygdalus*, *Prunus dulcis*, identification, geographical distribution, similarity, genealogy

Abstract. **Genetic diversity of the Spanish national almond (*Prunus amygdalus* Batsch) collection was characterized with 19 simple sequence repeat (SSR) markers selected because of their polymorphism in almond and other *Prunus* L. species. A total of 93 almond genotypes, including 63 Spanish cultivars from different growing regions, as well as some international cultivars and breeding releases were analyzed. All primers produced a successful amplification, giving a total of 323 fragments in the genotypes studied, with an average of 17 alleles per SSR, ranging from 4 (EPDCU5100) to 33 (BPPCT038). Allele size ranged from 88 bp at locus PMS40 to 260 bp at locus CPPCT022. The heterozygosity observed (0.72) was much higher not only than in other *Prunus* species, but also than in other almond pools already studied. The dendrogram generated using the variability observed classified most of the genotypes according to their geographical origin, confirming the particular evolution of different almond ecotypes. The SSR markers have consequently shown their usefulness for cultivar identification in almond, for establishing the genetic closeness among its cultivars, and for establishing genealogical relationships.**

Almond {*Prunus amygdalus* [syn. *P. dulcis* (Mill.) D.A. Webb]} is a species in the Rosaceae family with a genome 2n = 16. The almond is the most important tree nut crop in terms of production and is one of the most polymorphic fruit species (Kester et al., 1991; Socias i Company and Felipe, 1992). This high polymorphism may be due to its self-incompatibility (Socias i Company, 1998) and to the utilization of open-pollinated seedlings in traditional almond culture (Grasselly, 1972; Rikhter, 1972). The almond originated in central Asia (Grasselly, 1976) and is probably the oldest tree nut crop to be domesticated, possibly during the third millennium BCE (Spiegel-Roy, 1986). Over several centuries, almond spread from its center or origin toward the Mediterranean Basin, and was introduced by the Phoenicians, Greeks, and Romans into the different Mediterranean regions, expanding from Greece and the Balkans to Spain and Portugal. In all the Mediterranean Basin, both in the northern and southern shores, almond production became concentrated in specific areas, mostly with a traditional cultural system, adapted to the drought-resistant and frost-sensitive characteristics of the available almond germplasm, resulting as a consequence in the emergence of adapted land races associated with specific production areas (Grasselly and Crossa-Raynaud, 1980). Although seedling propagation resulted in the proliferation of a large number of highly variable local genotypes, their origin from a restricted germplasm often limited their genetic diversity, as shown in most islands of the Mediterranean Sea. Thus, the Mediterranean area is considered as a secondary source of domestication for almond (Felipe, 2000; Kester et al., 1991).

The Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA) almond collection (Espiau et al., 2002) was established in the late 1960s by A.J. Felipe, first by gathering cultivars produced by the Spanish nurseries. At that time, the number of nurseries was important and most of them produced a very high number of cultivars, thus allowing a significant number of accessions to be included. At the same time, a continuous collection effort was undertaken, with expeditions into most of the Spanish growing regions, paying particular attention to such geographically isolated areas as the Balearic and the Canary Islands. Furthermore, an interchange of plant material was maintained with different research centers of all almond-producing regions, allowing the introduction of a large number of foreign cultivars, both traditional and releases from the breeding programs. Especially significant was the number of accessions provided by C. Grasselly from Institut National de la Recherche Agronomique (INRA) in France. As a result of all these introductions, the CITA almond collection shows a very large variability, reflecting the wide genetic diversity of its accessions from all over the world (Socias i Company and Felipe, 1992). Taking this variability into account, this collection was designed as a reference for the Group de Recherches et d'Études Méditerranéen pour l'Amandier (GREMPA), being also the almond reference collection for the Spanish Plant Genetic Resources Network and for the Spanish and the European Plant Variety Offices. This collection was the initial basis for almond studies in Spain, including breeding, pollen compatibility, cultivar description (Felipe, 2002), chilling and heat requirements (Alonso et al., 2005), and *S*-genotype identification (Kodad et al., 2008).

Received for publication 24 July 2009. Accepted for publication 8 Sept. 2009. This research was funded by the grants INIA-RF2008-00027, CICYT AGL2007-65853-C02-02, AGRI GEN RES 870/2004 068 (SAFENUT), and the Research Group A12 of Aragón. A. Fernández i Martí acknowledges a scholarship cofunded by the Spanish 'Ministerio de Educación y Ciencia' and the European Social Fund (FSE), under projects AGL 2004-06674-C02-01 and BES-2006-12621.
Helpful comments by P. Arús, W. Howad and E. Collell (IRTA-Cabrils) are highly appreciated. We recognize the magnificent task of Dr. Antonio J. Felipe in assembling the Spanish almond germplasm collection.
[1]Corresponding author. E-mail: rsocias@aragon.es.

HortScience 44(3):609–613. 2009.

Cloning and Characterization of a Self-compatible S^f Haplotype in Almond [*Prunus dulcis* (Mill.) D.A. Webb. syn. *P. amygdalus* Batsch] to Resolve Previous Confusion in Its S^f-*RNase* Sequence

Toshio Hanada, Kyoko Fukuta, Hisayo Yamane, Tomoya Esumi, and Ryutaro Tao[1]
Graduate School of Agriculture, Kyoto University, Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan

Thomas M. Gradziel and Abhaya M. Dandekar
Department of Plant Science, University of California, Davis, CA 95616

Ángel Fernández i Martí, José M. Alonso, and Rafel Socias i Company
Unidad de Fruticultura, CITA de Aragón, Av. Montañana 930, 50059 Zaragoza, Spain

Additional index words. self-incompatibility, *S* haplotype, *S-RNase*, *SFB*

Abstract. **Most of the self-compatible (SC) cultivars of almond [*Prunus dulcis* (Mill.) D.A. Webb. syn. *P. amygdalus* Batsch] have the S^f haplotype. In this study, we cloned and characterized the *S* locus region of the S^f haplotype of SC 'Lauranne'. The relative transcriptional orientation of SFB^f and S^f-*RNase* and the physical distance between them are similar to those of other functional self-incompatible (SI) *S* haplotypes of *Prunus*, indicating that the genomic structure of the SC S^f haplotype appears to be intact. Although there is no apparent mutation in the coding sequence of SFB^f, the S^f-*RNase* sequence in this study and previously reported S^f-*RNase* sequences show discrepancies. First, as opposed to previous indications, the 'Lauranne' S^f-*RNase* sequence encodes a histidine residue in place of a previously reported arginine residue in the conserved C2 region of *Prunus* S-RNase. Direct sequencing of the polymerase chain reaction products from the S^f-*RNase* of 'Tuono' confirmed that 'Tuono' S^f-*RNase* also encodes the histidine residue. We found another difference in the 'Lauranne' S^f-*RNase* sequence and other reported S^f-*RNase* sequences. Namely, 'Lauranne' S^f-*RNase* encodes a phenylalanine residue in place of a previously reported leucine residue in the conserved C5 region of *Prunus* S-RNase. This is also the case for 'Tuono' S^f-*RNase*. Expression analysis of S^f-*RNase* and SFB^f by reverse transcriptase–polymerase chain reaction showed that S^f-*RNase* transcripts were barely detectable in pistil, whereas SFB^f transcripts were accumulated at a similar level to the level that was observed with *SFB* of other functional SI *S* haplotypes of almond. We discuss the possible molecular mechanisms of SC observed with the S^f haplotype with special references to the expression of S^f-*RNase*.**

Most of the fruit tree species in the genus *Prunus* (Rosaceae), including almond [*Prunus dulcis* (Mill.) D.A. Webb.], exhibit the S-RNase-based gametophytic self-incompatibility (GSI) system (de Nettancourt, 2001; Yamane and Tao, 2009). The GSI reaction in *Prunus* is controlled by the *S* locus, which contains the style-specific ribonuclease gene (*S-RNase*) (Tao et al., 1997; Ushijima et al., 1998) and the pollen-specific F-box protein gene (*SFB*) (Ushijima et al., 2003; Yamane et al., 2003) for the pistil and pollen specificities, respectively. Because the pistil and pollen determinant genes have been identified, the variants of the *S-RNase* and *SFB* combinations are called *S* haplotypes and the variants of a given *S* locus gene are called pistil and pollen *S* alleles.

Almond cultivars are largely self-incompatible (SI) (Tufts and Philip, 1922), although self-compatible (SC) cultivars exist (Socias i Company, 1990). Grasselly and Olivier (1976) reported that several SC cultivars such as Tuono, Filippo Ceo, Occhiorosso, and Genco were found among the almond population of the Italian region of Puglia, where *P. webbii* (Spach) Vierh. grows wild. Thus, most of the SC almond selections were considered to be derived from interspecific hybridization with SC *P. webbii* (Socias i Company, 2004). As a consequence, the SC S^f haplotype of 'Tuono' has been long thought to be derived from *P. webbii*. However, a recent finding of the almond SI S^{30} haplotype, a putative wild-type S^f haplotype, poses a question about this hypothesis (Bošković et al., 2007). The origin and molecular basis of SC in S^f haplotype is intriguing for future SC breeding programs in almond.

We previously sequenced the partial S^f-*RNase* sequence that was flanked by the Pru-C2 and Pru-C5 primer sequences that were designed from the conserved C2 and C5 regions, respectively, of rosaceous S-RNase (Tao et al., 1999; Ushijima et al., 1998) and the full coding sequence for SFB^f (Hanada et al., 2009). No substantial differences or defects in the deduced amino acid sequence were found in the partial S^f-*RNase* and full-length SFB^f sequences. Although our full-length SFB^f sequence completely matches the partial SFB^f sequences that were reported by other research groups, partial S^f-*RNase* sequences that were posted on the public database by several research groups, including our group, show differences (Barckley et al., 2006; Bošković et al., 2007; Channuntapipat et al., 2001; Ma and Oliveira, 2002). Although it appeared later that the S^f-*RNase* sequence reported by Barckley et al. (2006) could be from misannotated S^1-*RNase* (synonymous to S^b-*RNase*) because their S^f-*RNase* sequence shows 100% match to S^1-*RNase* of Tuono (S^1S^f), there are still some minor differences among the S^f-*RNase* sequences reported by the different research groups. This led us to thoroughly reinvestigate the *S* locus of the S^f haplotype in almond. We, therefore, isolated fosmid clones that contained the S^f locus of the SC 'Lauranne' to determine the full coding sequence of the S^f-*RNase* and the structure of the *S* locus region of the S^f haplotype. Furthermore, we conducted expression analysis of *S* locus genes by reverse transcriptase–polymerase chain reaction (RT-PCR) and discuss the possible molecular mechanisms of SC observed with the S^f haplotype with special references to the expression of S^f-*RNase*.

Materials and Methods

Plant materials. Young leaves, pistils, and pollen grains were collected from three almond cultivars, Lauranne (S^3S^f), Ferragnès (S^1S^3), and Tuono (S^1S^f), that were grown at the CITA de Aragón in Zaragoza, Spain. 'Lauranne' is from 'Ferragnès' × 'Tuono', and the S^f and S^3 haplotypes in 'Lauranne' are estimated to be from 'Tuono' and 'Ferragnès', respectively. We also used young leaf samples from 'Tuono' that was grown at the University of California at Davis, CA. Leaf samples were collected, frozen in liquid nitrogen, lyophilized, and stored at –20 °C with desiccant until used.

Isolation of DNA. Genomic DNA was isolated from young leaves using the Nucleon PhytoPure® plant and fungal DNA extraction kit (GE Healthcare, Piscataway, NJ) with

Received for publication 29 Jan. 2009. Accepted for publication 10 Feb. 2009.
This work was supported by a Grant-in-Aid (no. 20248004) for Scientific Research (A) to R.T.
[1]To whom reprint requests should be addressed; e-mail rtao@kais.kyoto-u.ac.jp.

Scientia Horticulturae 123 (2010) 308–311

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Molecular and physiological identification of new *S*-alleles associated with self-(in)compatibility in local Spanish almond cultivars

Ossama Kodad [a,1], José M. Alonso [a], Angel Fernández i Martí [a], M. Margarida Oliveira [b,c], Rafel Socias i Company [a,*]

[a] Unidad de Fruticultura, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Av. Montañana 930, 50059 Zaragoza, Spain
[b] Instituto de Tecnologia Química e Biológica (ITQB), Instituto de Biologia Experimental e Tecnológica (IBET), Quinta do Marquês, 2784-505 Oeiras, Portugal
[c] Universidade de Lisboa, Faculdade de Ciências, Dep. de Biologia Vegetal, Campo Grande, 1749-016 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 26 May 2009
Received in revised form 31 August 2009
Accepted 16 September 2009

Keywords:
Prunus amygdalus
Self-incompatibility
S-alleles
Cross-incompatibility groups

ABSTRACT

The *S*-allele characterisation of 'Alzina' and 'Garondès', two local almond cultivars from the island of Majorca, by a multidimensional approach has allowed the confirmation of the presence of the S_f-allele and the identification of a new allele not previously described in almond, S_{36}. When these cultivars were phenotypically evaluated, both showed a self-incompatible phenotype and were cross-incompatible, as assessed by artificial self- and cross-pollinations and fruit sets after field pollinations, confirming that their S_f-allele is in its active form, S_6. Thus a new CGI group in almond is proposed and named XXVIII. These results confirm the wide diversity of *S*-alleles in almond both at genotypic and phenotypic levels, as well as their similarity with the *S*-alleles from other close *Prunus* species. This similarity suggests the possibility of allele introgression between species or allele identity by descent from a common ancestor.

1. Introduction

Most almond [*Prunus amygdalus* Batsch syn. *Prunus dulcis* (Mill.) D.A. Webb] cultivars are, with few exceptions, self-incompatible (Socias i Company, 1990). Gametophytic self-incompatibility (GSI) is a widespread mechanism in flowering plants that prevents self-fertilization and promotes outcrossing (de Nettancourt, 2001). The GSI mechanism, which is found in the Solanaceae, Rosaceae and Scrophulariaceae has an *S*-RNase as the pistil *S*-component and an F-box protein as the pollen *S*-component (Kao and Tsukamato, 2004). In an incompatible situation, the pistil RNases degrade the pollen RNA, thereby preventing pollen tube growth.

Previously to the introduction of new cultivars in the years 1970s, traditional almond growing in Spain was based in two main cultivars, 'Desmayo Largueta' and 'Marcona', but with a large number of local cultivars normally grown only in a reduced region and rarely expanding outside their original area (Felipe, 2000). The wealth of local cultivars had been due to the initial almond propagation by seeds in the past, ensuing a close relationship among many of them and in the high frequency of a reduced number of *S*-alleles, as it has been described in populations showing a close relationship among their individuals (Kester et al., 1994; Lansari and Lakhal, 2001). As a consequence, cases of cross-incompatibility have been found in several local cultivars (Felipe, 2000; Martínez-García et al., 2009), resulting in reduced production. Therefore, the knowledge of the *S*-genotype in almond cultivars is useful for orchard design by growers to ensure cross-pollination and reach a commercial crop level, as well as for parental choice in breeding programmes.

The *S* locus shows a high diversity in almond and more than 37 alleles have been identified (Kester et al., 1994; López et al., 2004; Ortega et al., 2006; Halász et al., 2008; Kodad et al., 2008a,b). The first attempt to establish cross-incompatibility groups (CIG) was achieved by Kester et al. (1994) in Californian cultivars by test crosses in the field, allowing the identification of six CIGs with allele assignation to 32 cultivars. Ortega et al. (2006) updated the table to 18 CIGs including 71 cultivars. More recently, five new alleles were identified (Kodad et al., 2008b) and 16 alleles were assigned to 29 local Spanish cultivars (Kodad et al., 2008a). Thus, the table given by Ortega et al. (2006) needed to be updated up to 25 CIGs with 108 cultivars (Kodad and Socias i Company, 2009).

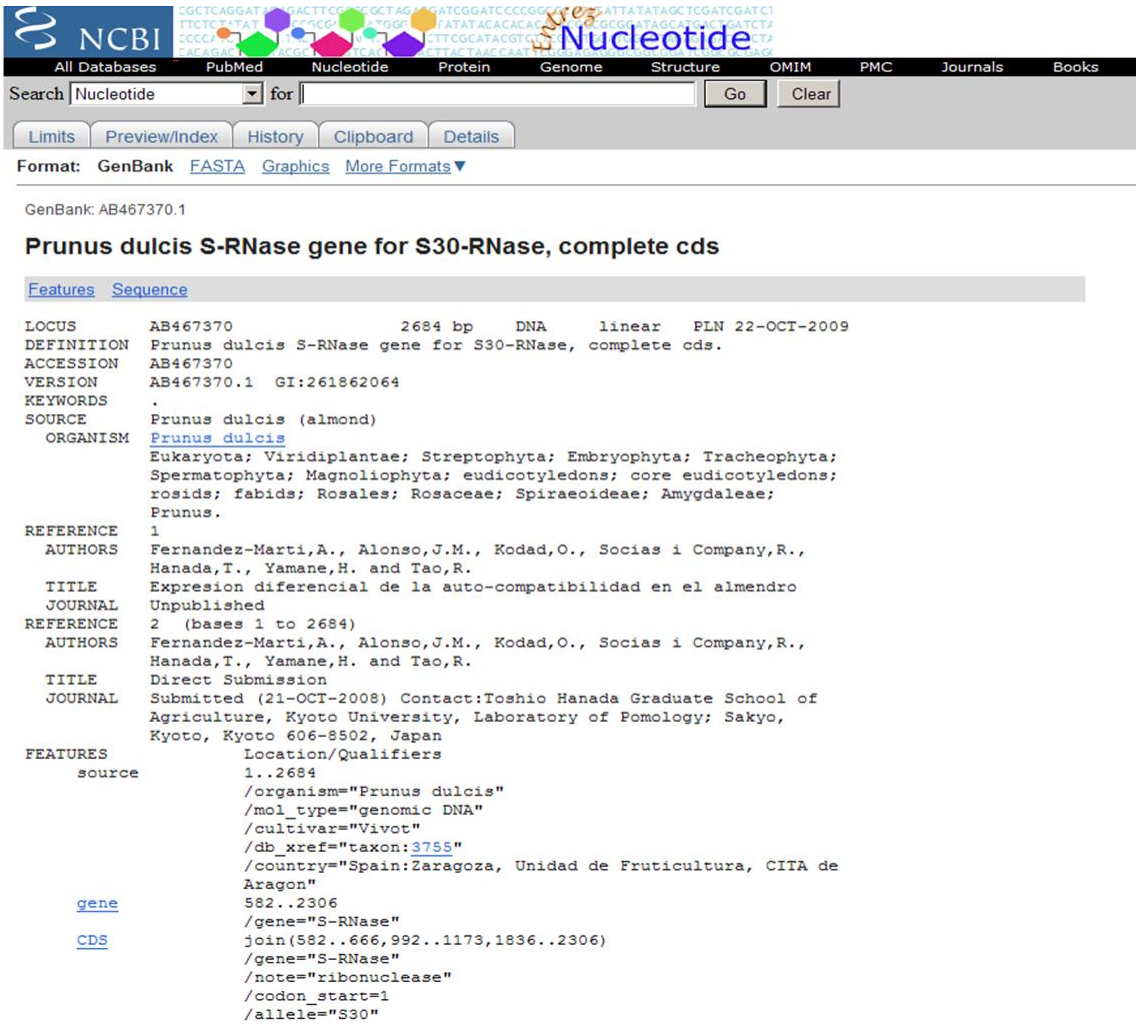
S-genotype determination was initially carried out in field test crosses, although this method is particularly difficult because the results may vary depending on the climatic conditions at blooming. Molecular advances in the study of SI in Rosaceae (McClure et al., 1989) led to the use of non-equilibrium pH gradient electrofocusing (NEpHGE) for *S*-allele identification in almond (Bošković et al., 2003). However, this technique is not very sensitive and does not distinguish alleles with the same isoelectric point (PI) (Bošković et al., 2003), although it is useful to confirm the

* Corresponding author. Tel.: +34 976 716313; fax: +34 976 716335.
E-mail address: rsocias@aragon.es (R. Socias i Company).
[1] Present address: Département d'Arboriculture, Ecole Nationale D'Agriculture, BP. S40, Meknès, Morocco.

0304-4238/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.scienta.2009.09.010

13.2 Accesiones del GenBank de las $S_{fa/i}$-Nasas y SFB_f de 'Blanquerna' y 'Vivot'

13.2.1 Accesion del GenBank de la S_{fa}-RNase de 'Vivot'



NCBI Nucleotide

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Nucleotide for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

Format: GenBank FASTA Graphics More Formats▼

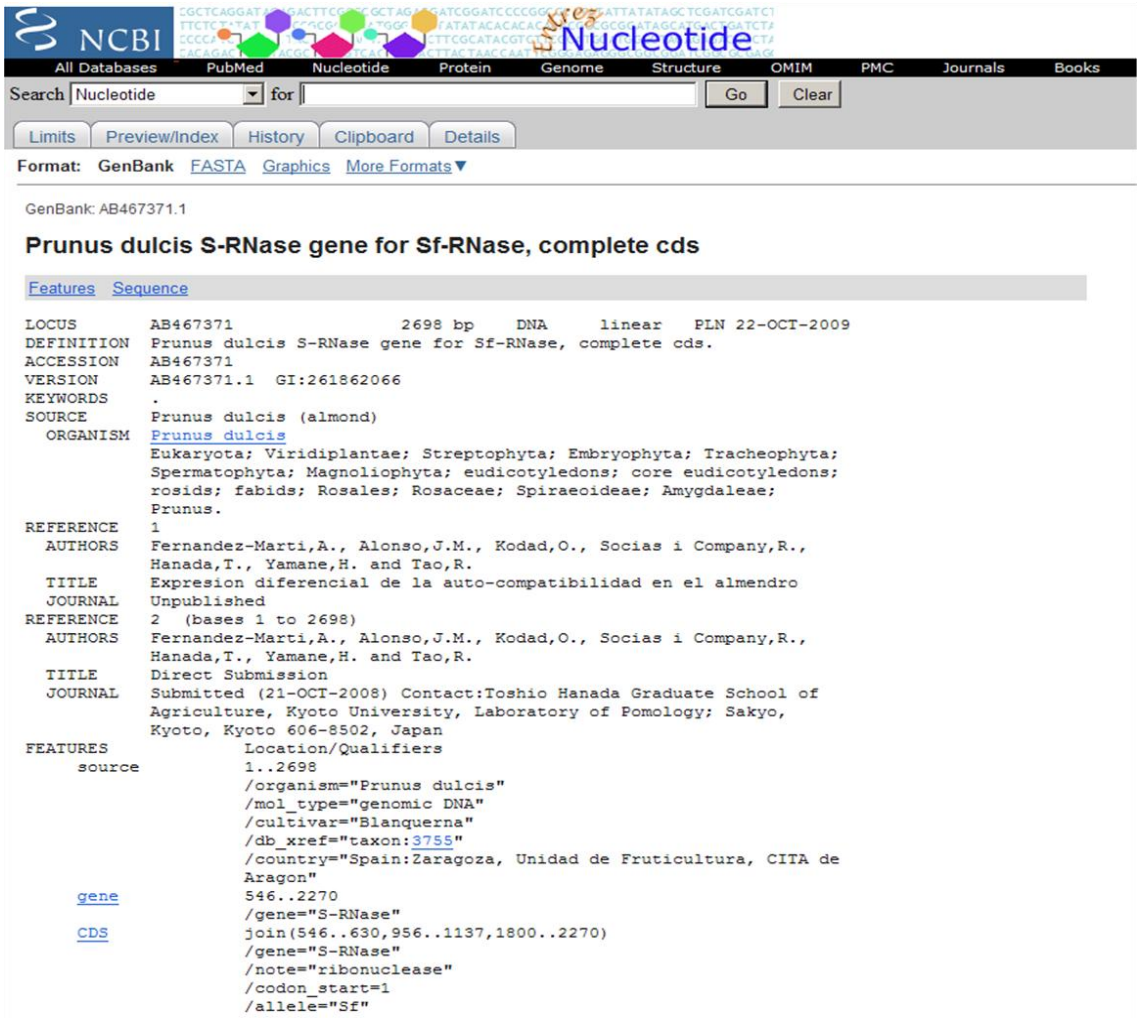
GenBank: AB467370.1

Prunus dulcis S-RNase gene for S30-RNase, complete cds

Features Sequence

```
LOCUS       AB467370               2684 bp    DNA     linear   PLN 22-OCT-2009
DEFINITION  Prunus dulcis S-RNase gene for S30-RNase, complete cds.
ACCESSION   AB467370
VERSION     AB467370.1  GI:261862064
KEYWORDS    .
SOURCE      Prunus dulcis (almond)
  ORGANISM  Prunus dulcis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; fabids; Rosales; Rosaceae; Spiraeoideae; Amygdaleae;
            Prunus.
REFERENCE   1
  AUTHORS   Fernandez-Marti,A., Alonso,J.M., Kodad,O., Socias i Company,R.,
            Hanada,T., Yamane,H. and Tao,R.
  TITLE     Expresion diferencial de la auto-compatibilidad en el almendro
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2684)
  AUTHORS   Fernandez-Marti,A., Alonso,J.M., Kodad,O., Socias i Company,R.,
            Hanada,T., Yamane,H. and Tao,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2008) Contact:Toshio Hanada Graduate School of
            Agriculture, Kyoto University, Laboratory of Pomology; Sakyo,
            Kyoto, Kyoto 606-8502, Japan
FEATURES             Location/Qualifiers
     source          1..2684
                     /organism="Prunus dulcis"
                     /mol_type="genomic DNA"
                     /cultivar="Vivot"
                     /db_xref="taxon:3755"
                     /country="Spain:Zaragoza, Unidad de Fruticultura, CITA de
                     Aragon"
     gene            582..2306
                     /gene="S-RNase"
     CDS             join(582..666,992..1173,1836..2306)
                     /gene="S-RNase"
                     /note="ribonuclease"
                     /codon_start=1
                     /allele="S30"
```

13.2.2 Accesion del GenBank de la S_{fi}-RNase de 'Blanquerna'



NCBI Nucleotide

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Nucleotide for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

Format: GenBank FASTA Graphics More Formats▼

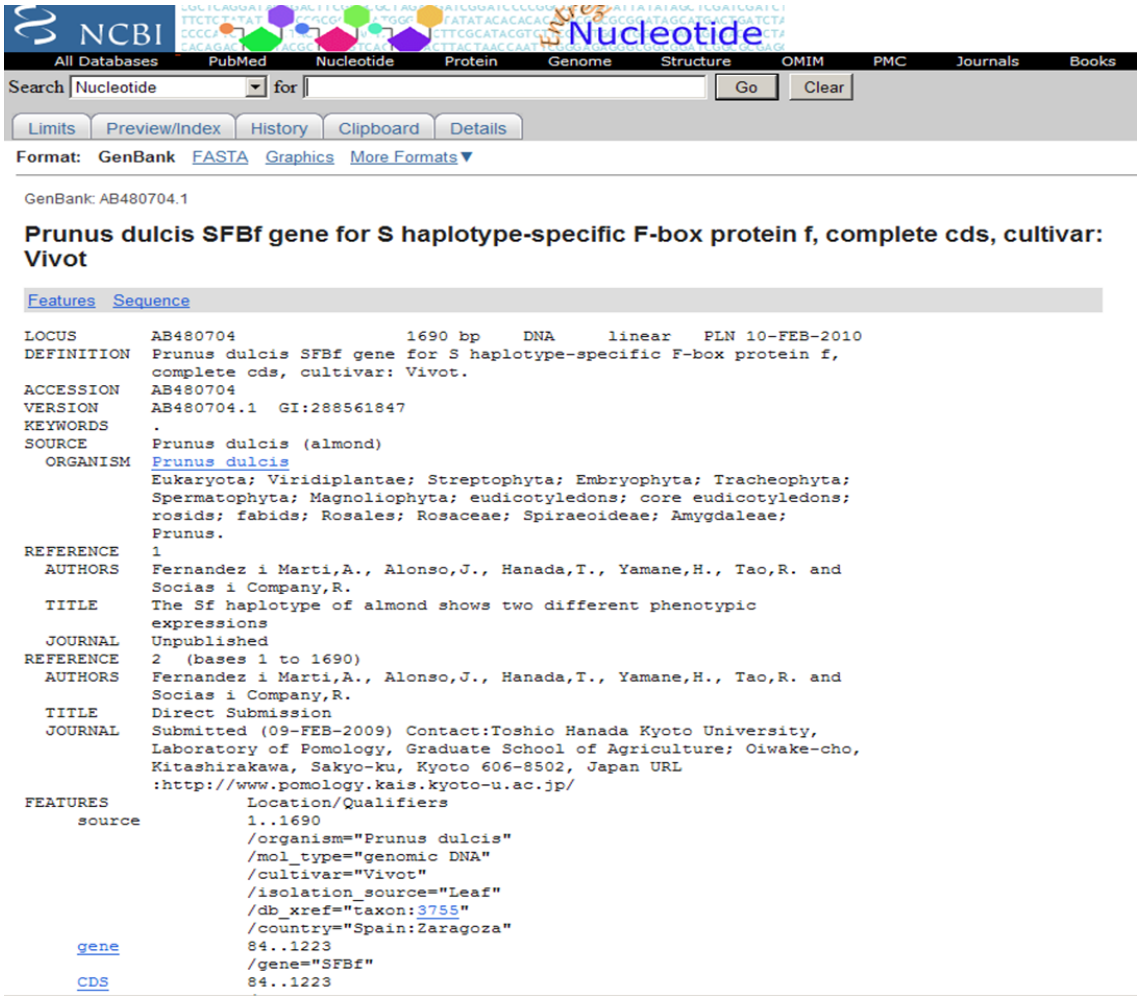
GenBank: AB467371.1

Prunus dulcis S-RNase gene for Sf-RNase, complete cds

Features Sequence

```
LOCUS       AB467371               2698 bp    DNA     linear   PLN 22-OCT-2009
DEFINITION  Prunus dulcis S-RNase gene for Sf-RNase, complete cds.
ACCESSION   AB467371
VERSION     AB467371.1  GI:261862066
KEYWORDS    .
SOURCE      Prunus dulcis (almond)
  ORGANISM  Prunus dulcis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; fabids; Rosales; Rosaceae; Spiraeoideae; Amygdaleae;
            Prunus.
REFERENCE   1
  AUTHORS   Fernandez-Marti,A., Alonso,J.M., Kodad,O., Socias i Company,R.,
            Hanada,T., Yamane,H. and Tao,R.
  TITLE     Expresion diferencial de la auto-compatibilidad en el almendro
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2698)
  AUTHORS   Fernandez-Marti,A., Alonso,J.M., Kodad,O., Socias i Company,R.,
            Hanada,T., Yamane,H. and Tao,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2008) Contact:Toshio Hanada Graduate School of
            Agriculture, Kyoto University, Laboratory of Pomology; Sakyo,
            Kyoto, Kyoto 606-8502, Japan
FEATURES             Location/Qualifiers
     source          1..2698
                     /organism="Prunus dulcis"
                     /mol_type="genomic DNA"
                     /cultivar="Blanquerna"
                     /db_xref="taxon:3755"
                     /country="Spain:Zaragoza, Unidad de Fruticultura, CITA de
                     Aragon"
     gene            546..2270
                     /gene="S-RNase"
     CDS             join(546..630,956..1137,1800..2270)
                     /gene="S-RNase"
                     /note="ribonuclease"
                     /codon_start=1
                     /allele="Sf"
```

13.2.3 Accesion del GenBank del SFB_f de 'Vivot'



NCBI Nucleotide

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Nucleotide for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

Format: GenBank FASTA Graphics More Formats▼

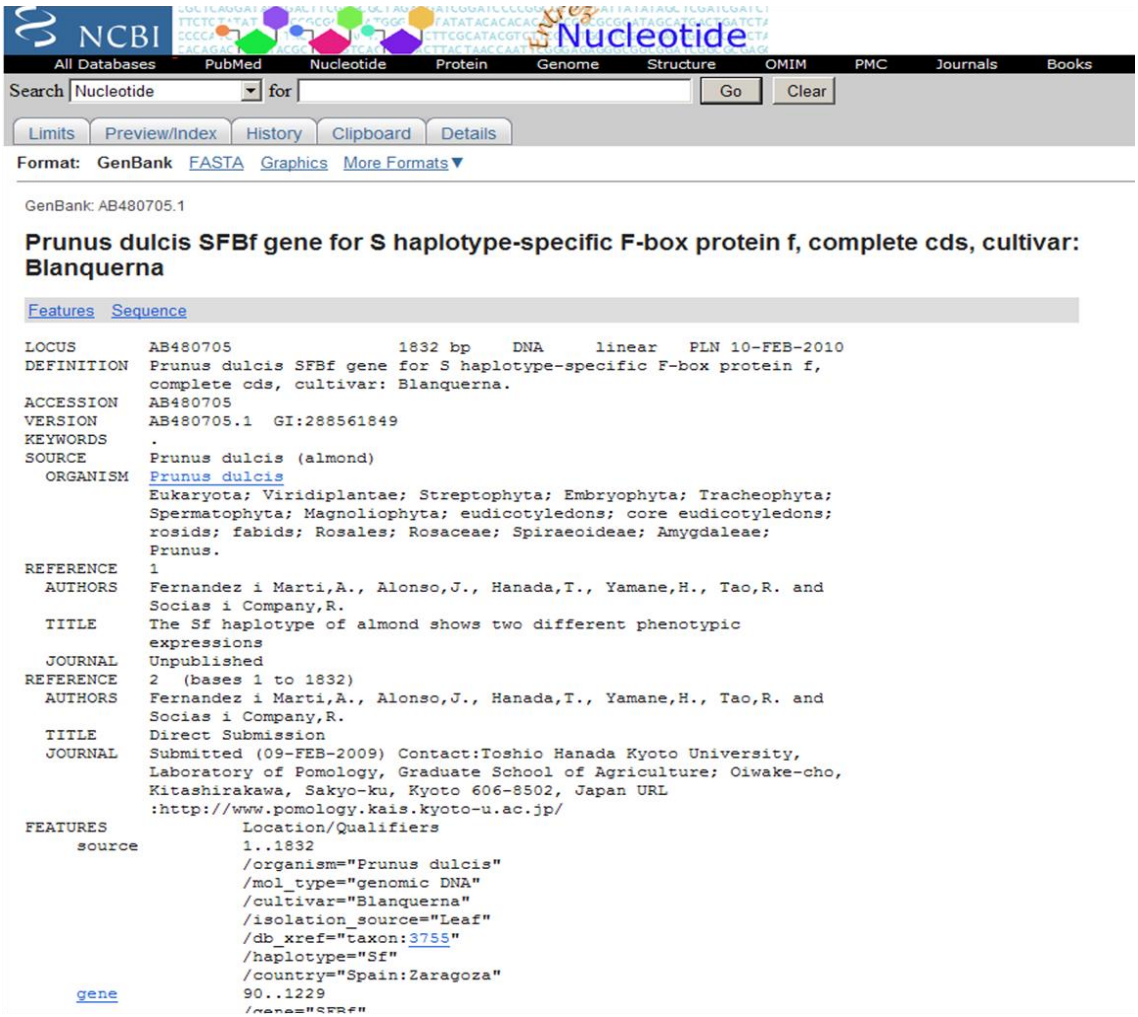
GenBank: AB480704.1

Prunus dulcis SFBf gene for S haplotype-specific F-box protein f, complete cds, cultivar: Vivot

Features Sequence

```
LOCUS       AB480704               1690 bp    DNA     linear   PLN 10-FEB-2010
DEFINITION  Prunus dulcis SFBf gene for S haplotype-specific F-box protein f,
            complete cds, cultivar: Vivot.
ACCESSION   AB480704
VERSION     AB480704.1  GI:288561847
KEYWORDS    .
SOURCE      Prunus dulcis (almond)
  ORGANISM  Prunus dulcis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; fabids; Rosales; Rosaceae; Spiraeoideae; Amygdaleae;
            Prunus.
REFERENCE   1
  AUTHORS   Fernandez i Marti,A., Alonso,J., Hanada,T., Yamane,H., Tao,R. and
            Socias i Company,R.
  TITLE     The Sf haplotype of almond shows two different phenotypic
            expressions
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1690)
  AUTHORS   Fernandez i Marti,A., Alonso,J., Hanada,T., Yamane,H., Tao,R. and
            Socias i Company,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-2009) Contact:Toshio Hanada Kyoto University,
            Laboratory of Pomology, Graduate School of Agriculture; Oiwake-cho,
            Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan URL
            :http://www.pomology.kais.kyoto-u.ac.jp/
FEATURES             Location/Qualifiers
     source          1..1690
                     /organism="Prunus dulcis"
                     /mol_type="genomic DNA"
                     /cultivar="Vivot"
                     /isolation_source="Leaf"
                     /db_xref="taxon:3755"
                     /country="Spain:Zaragoza"
     gene            84..1223
                     /gene="SFBf"
     CDS             84..1223
```

13.2.4 Accesion del GenBank del SFB_f de 'Blanquerna'



NCBI Nucleotide

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard Details

Format: GenBank FASTA Graphics More Formats▼

GenBank: AB480705.1

Prunus dulcis SFBf gene for S haplotype-specific F-box protein f, complete cds, cultivar: Blanquerna

Features Sequence

```
LOCUS       AB480705               1832 bp    DNA     linear   PLN 10-FEB-2010
DEFINITION  Prunus dulcis SFBf gene for S haplotype-specific F-box protein f,
            complete cds, cultivar: Blanquerna.
ACCESSION   AB480705
VERSION     AB480705.1  GI:288561849
KEYWORDS    .
SOURCE      Prunus dulcis (almond)
  ORGANISM  Prunus dulcis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; fabids; Rosales; Rosaceae; Spiraeoideae; Amygdaleae;
            Prunus.
REFERENCE   1
  AUTHORS   Fernandez i Marti,A., Alonso,J., Hanada,T., Yamane,H., Tao,R. and
            Socias i Company,R.
  TITLE     The Sf haplotype of almond shows two different phenotypic
            expressions
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1832)
  AUTHORS   Fernandez i Marti,A., Alonso,J., Hanada,T., Yamane,H., Tao,R. and
            Socias i Company,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-2009) Contact:Toshio Hanada Kyoto University,
            Laboratory of Pomology, Graduate School of Agriculture; Oiwake-cho,
            Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan URL
            :http://www.pomology.kais.kyoto-u.ac.jp/
FEATURES             Location/Qualifiers
     source          1..1832
                     /organism="Prunus dulcis"
                     /mol_type="genomic DNA"
                     /cultivar="Blanquerna"
                     /isolation_source="Leaf"
                     /db_xref="taxon:3755"
                     /haplotype="Sf"
                     /country="Spain:Zaragoza"
     gene            90..1229
                     /gene="SFBf"
```