UNIVERSITA' DEGLI STUDI DI PADOVA

SCUOLA DI DOTTORATO DI RICERCA IN

SCIENZE DELLE PRODUZIONI VEGETALI

INDIRIZZO AGROBIOTECNOLOGIE - CICLO XXIV

Dipartimento di Agronomia Ambientale e Produzioni Vegetali

REPRODUCTIVE BARRIERS IN CROP PLANTS: UNDERSTANDING THE GENETIC AND MOLECULAR BASES OF SELF-INCOMPATIBILITY IN OLIVE AND MALE-STERILITY IN CHICORY

Direttore della Scuola: Ch.mo Prof. Andrea Battisti

Supervisore: Ch.mo Prof. Gianni Barcaccia

Dottorando: Silvio Collani

DATA CONSEGNA TESI

31 gennaio 2012

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Silvio Collani January, 31st 2012

A copy of the thesis will be available at *http://paduaresearch.cab.unipd.it/*

Dichiarazione

Con la presente affermo che questa tesi è frutto del mio lavoro e che, per quanto io ne sia a conoscenza, non contiene materiale precedentemente pubblicato o scritto da un'altra persona né materiale che è stato utilizzato per l'ottenimento di qualunque altro titolo o diploma dell'università o altro istituto di apprendimento, a eccezione del caso in cui ciò venga riconosciuto nel testo.

Silvio Collani 31 Gennaio 2012

Una copia della tesi sarà disponibile presso *http://paduaresearch.cab.unipd.it/*

Index

Riassunto

La presente tesi è mirata allo studio di due importanti barriere riproduttive in piante di interesse agro-alimentare: auto-incompatibilità in olivo e maschio-sterilità in radicchio.

L'olivo (*Olea europaea* L.) è una della più antiche specie coltivate e, nonostante il suo grande impatto economico e culturale, pochi studi sono stati fatti fino ad ora riguardo le barriere riproduttive che incidono sulla produzione finale. Lo scopo della tesi è quello di chiarire il tipo di auto-incompatibilità in olivo fornendo evidenze cito-istologiche e molecolari. L'auto-incompatibilità è il sistema più comune adottato dalle angiosperme per evitare fenomeni di inbreeding, mantenendo in questo modo la diversità all'interno della specie. Attualmente l'olivo è classificato come specie auto-incompatibile di tipo gametofitico (GSI), ma tale classificazione è solo basata su evidenze botaniche come stigma umido e polline bi-nucleato. Approfondite analisi citologiche eseguite su oltre 34.000 granuli pollinici presenti su pistilli di cultivar auto-incompatibili e auto-compatibili non hanno prodotto risultati in accordo con la prevista GSI. Inoltre, il tentativo di isolare geni coinvolti nella GSI non ha prodotto alcun risultato. Le nostre osservazioni citologiche concordano invece con un'incompatibilità di tipo sporofitico (SSI). Esperimenti mirati ad individuare geni coinvolti nella SSI hanno permesso di isolare due geni riconducibili al determinante femminile, *OeSRK* (S-locus Receptor Kinase), due ad un enhancer della reazione d'incompatibilità, *OeSLG* (S-locus Glycoprotein) e uno al determinante maschile, *OeSCR-like* (S-locus Cysteine rich protein). Le analisi di espressione condotte mediante qRT-PCR, usando diverse combinazioni di primer, hanno permesso di evidenziare opposti modelli di espressione per i geni *OeSRK* e *OeSLG* fra le cultivar Leccino, classificata come auto-incompatibile, e Frantoio, ritenuta invece auto-compatibile. Inoltre, i full-length di *OeSRK* hanno mostrato avere una espressione pistillo-specifica, mentre quelli per *OeSCR-like* una espressione antera-specifica. Ibridazioni *in situ* hanno permesso di localizzare i trascritti di *OeSRK* e *OeSLG* nella superficie delle cellule papillari. Quindi, sulla base delle

nostre evidenze, è possibile ipotizzare che l'olivo sia caratterizzato da un nuovo tipo di auto-incompatibilità diverso da quello attualmente riportato in letteratura.

Il lavoro svolto sulla maschio-sterilità in radicchio (*Cichorium intybus* L.) è mirato alla identificazione e alla caratterizzazione di mutanti maschio-sterili. Quattro distinti mutanti, che al meglio delle nostre conoscenze rappresentano i primi mutanti maschio-sterili spontanei scoperti fino ad ora nel genere *Cichorium*, sono stati dettagliatamente caratterizzati sia per quanto riguarda la microsporogenesi e gametogenesi dal punto di vista citogenetico, sia per quanto riguarda il controllo genetico del carattere. È stato inoltre identificato un marcatore molecolare che può essere impiegato ai fini della selezione assistita dei mutanti maschio-sterili. Nel loro complesso, i dati raccolti in questo lavoro concordano con il controllo della maschio-sterilità da parte di un singolo gene nucleare che agisce in condizioni di omozigosi recessiva. Approfondite analisi citologiche hanno permesso di visualizzare che la gametogenesi maschile è arrestata allo stadio di microspora uni-nucleata. Nel dettaglio, le microspore degenerano prima del rilascio dalle tetradi e, in seguito, mostrano un collasso dell'esina. Nei mutanti è stato dimostrato che la totalità delle microspore sono raggrinzite e molto più piccole se comparate con quelle normali. È stato inoltre isolato un marcatore SSR strettamente associato con il gene responsabile della maschio-sterilità e la sequenza è stata depositata nel database NCBI (accession no. JF748831). Nel corso del 2011 sono state prodotte quattro nuove varietà ibride di radicchio "Rosso di Chioggia" che differiscono per la loro precocità, spaziando dagli 80 ai 110 giorni. Nel loro complesso, i risultati ottenuti dimostrano come mutanti maschio-sterili possano essere impiegati con profitto nella produzione in larga scala di ibridi F1 anche in quelle specie nelle quali, fino ad ora, non era ritenuta possibile la costituzione di varietà ibride. Pertanto, la scoperta di maschio-sterilità genetica spontanea, non transgenica, in radicchio apre nuove frontiere per la produzione di ibridi F1 in questa coltura ed in altre tipologie coltivate della stessa specie.

Abstract

The thesis deals with the understanding of two reproductive barriers affecting crop plants: self-incompatibility in olive and male-sterility in chicory.

Olive (*Olea europaea* L.) is one of the oldest agricultural tree crop species and, in spite of the great economical and cultural impact, a few studies have been carried out on its reproductive barriers. The aim of this research was to elucidate the self-incompatibility system in olive from cytohistological and bio-molecular standpoints. Self-incompatibility is one of the most effective systems adopted by flowering plants to prevent inbreeding, maintaining so diversity within the species. Olive is currently classified as a gametophytic self-incompatible (GSI) plant because of distinctive morphological traits, as wet-type pistil and bi-nucleated pollen. However, detailed cytological analyses of more than 34,000 pollen grains performed using pistils of self-compatible and self-incompatible cultivars under self-pollination and open-pollination conditions, were not in agreement with GSI. Furthermore, no results were achieved by molecular analyses aimed at cloning genes involved in the GSI system. Vice versa, our cytological observations were in agreement with a sporophytic self-incompatibility (SSI). The molecular attempts to isolate candidates for SSI led us to the cloning of two *OeSRK* (S-locus Receptor Kinase, the female determinant), two *OeSLG* (S-locus Glycoprotein, an enhancer of the incompatibility response) and one *OeSCR*-like (S-locus Cysteine Rich protein, the male determinant) genes. Moreover, quantitative Real-Time PCR assays, replicated using different subdomain-specific primer combinations, revealed an antagonist transcriptional expression pattern in flowers of cultivars Leccino, classified as self-incompatible cultivar, and Frantoio, classified as self-compatible cultivar, for the genes *OeSRK* and *OeSLG*. Furthermore, full-length of the *OeSRK* gene was found specific of pistils, while full-length of *SCR-like* gene was found specific of anthers. Localization of *OeSRK* and *OeSLG* transcripts, were performed by in situ hybridization and

it was found localized in the surface of papillar cells. On the whole, a new hypothesis for the genetic system controlling the self-incompatibility reaction can be postulated for olive.

Research focused on male-sterility in chicory (*Cichorium intybus* L.) deals with the discovery and genetic analysis of male sterile mutants of red chicory. Four distinct mutants, which to the best of our knowledge are the first spontaneous male sterile mutants ever discovered and described in the genus *Cichorium*, were characterized in great details for the developmental pathway of microsporogenesis and gametogenesis, and the inheritance pattern of the gene underlying the male-sterility trait. A quick molecular diagnostic assay was also developed for the early marker-assisted selection of the genotype associated to male sterile plants. Overall data clearly support a nuclear origin and a monogenic control of recessive type for the male-sterility trait in each of the red chicory mutants. Male gametogenesis was documented to arrest at the stage of uninucleate microspores. In particular, cytological observations revealed that microspores degenerate before their release from the tetrads, later showing a collapse of the exine. In the mutants, the totality of microspores proved to be shrunken and much smaller than wild-type ones. Moreover, the fine mapping of the mutant locus was attempted by molecular markers using F2 and BC1 populations segregating for male-sterility. The gene responsible for male-sterility was found tightly linked to a microsatellite of the TC/GA type whose full sequence was recently deposited in the NCBI databases under the accession no. JF748831. A molecular diagnostic assay was then developed to be profitably adopted as a tool of marker-assisted breeding and exploited for an early screening of male-sterile plants within segregating progenies stemmed from back-crosses, with a genotyping error lower than 3%. Four new hybrid varieties of radicchio *Rosso di Chioggia* with different earliness, spanning from 80 to 110 days, were bred during 2011 by crossing male sterile partially inbred clones, used as seed parents, with wild type highly inbred lines, used as pollen donors. On the whole, the constitution of F1 hybrids seems profitable in a practical breeding scheme and it is also

feasible on a large commercial scale by the selection of self-compatible genotypes, for the production of inbred lines, and the identification of male sterile genotypes, to be used as seed parents for the hybridization with unrelated pollen donors. The discovery of non-engineered male-sterility in red chicory will open new frontiers for maximizing crop productivity in this important cultivated vegetable species through the breeding of heterotic F1 hybrid varieties.

Aims of the thesis

The goal of the present thesis was to shed light on the main reproductive barriers that occur in crop plants. In particular, this research project deals with two of the main reproductive barriers adopted by flowering plants: self-incompatibility and male-sterility.

The self-incompatibility (SI) is the most widespread reproductive barrier among flowering plants and it has been supposed to occur in almost half of the species. The main SI systems currently known will be deeply reviewed as well as cases of SI breakdown in plants that are now perfectly self-compatible, such as *Arabidopsis thaliana*, and that have been used to study the molecular mechanisms downstream the incompatible reaction. In particular, SI was studied in olive (*Olea europaea* L.), which is a crop plant with a great economic impact in several countries of the Mediterranean basin and which is characterized by great yield loss due to SI. The aim was to clarify which is the SI system acting in olive and to provide cytological and molecular evidences supporting our hypothesis. To the best of our knowledge, this study represents the first one aimed at elucidating the SI system adopted by olive plants.

The other reproductive barrier investigated in this thesis was the male-sterility. In our case, male-sterility was studied in red chicory (*Chicorium intybus* L.), which is also very important as local typical product in Italy. Some spontaneous male-sterile mutants were found in cultivated lines of red chicory and characterized in great details. In particular, our goals were to understand the type and nature of male-sterility (*e.g.* cytoplasmic or nuclear), to investigate the cytogenetic feature of microsporogenesis, to genetically map the male-sterile locus and to discover the gene/s responsible for this mutation. The male-sterile mutants represent the first case of spontaneous, non-transgenic, male-sterility currently known in chicory. Besides the molecular studies aimed at understanding the cytogenetic features and molecular traits of male-sterility, our study was useful to breed the first F1 hybrid cultivar in red chicory for which a patent is currently pending.

Reproductive barriers in plants

Millions of years of evolution led flowering plants to develop a huge diversity of floral morphology and a wide range of ways of fertilization systems. Such great variety of floral morphology and fertilization systems is the result of thousands of selection events according to the environment conditions and to the pollination strategies adopted. Besides to physical evidences, such as floral morphology, the co-evolution of molecular mechanisms represented a crucial step in order to establish a wide number of reproductive strategies within flowering plants, which is considered the main goal of evolution and the training force allowing green plants to colonize the entire world, directly influencing the animal life in the Earth. In nature currently coexist both self- and cross-fertilization, which are co-evolved according to the environment, and they are characterized by both advantages and disadvantages. Cross-fertilization in flowering plants is characterized by a wide number of positive effects. In fact, two parents characterized by their own gene pool, after cross-fertilization, produce an offspring that includes a wide number of new gene combinations (*i.e.* genotypes) and new phenotypes. The wide range of new gene pools is predominantly due to segregation and recombination of homologous chromosomes during meiosis, which usually ends with the production of haploid gametes. Obligated cross-fertilization, which is achieved by several strategies aimed to inhibit self-fertilization, represents thus an enhancement of the positive effects due to the creation of new gene pools. New gene pools mean more gene combinations and, consequently, more diversity, which represents a wide source of adaptability to different and changing environmental conditions. Contrarily to cross-fertilization, self-fertilization in several species leads to negative effects often due to a worst phenotype or a lower fitness if compared with those plants originated from crossed mates. Usually the negative effects of one or more cycles of self-fertilization are pooled under the concept of inbreeding depression. However, it is worth noting that several species

do not show inbreeding depression symptoms after self-fertilization and in some cases self-fertilization is the rule (*e.g.* several *Poaceae* and *Fabaceae*). Because of the high advantage of cross-fertilization, flowering plants developed mating strategies aimed to an obligated crossing, which can be divided into two big classes: physical barriers avoiding self-fertilization and molecular mechanisms rejecting self-pollen or inhibiting normal development of sexual organs (*e.g.* functional pistils and/or anthers). Among physical barriers the most common are: i) dioecy, which is the condition characterized by diclinous, unisexual, female or male flowers, which are brought in different plants, female and male plants, such as *Ginkgo biloba*, *Taxus baccata*, genus *Cycas*, *Ilex aquifolium*, *Urtica dioica*, *Laurus nobilis*, and others; ii) dichogamy, where anther dehiscence and stigma receptivity are temporally separated and it can be divided into two classes: proterandry (usually occurring in *Asteraceae* and *Apiaceae*), when anthers are the first organ to reach maturity, or proterogyny (typical of *Brassicaceae*, *Rosaceae* and *Ranuncolaceae*), if stigma is the first organ to become receptive; iii) heteromorphic SI, typical of some species belonging to *Primulaceae*, *Rubiaceae* and *Plumbaginaceae*, consists of two or three forms of flowers characterized by the stigma and anthers at different level.

However, the most spread strategy aimed at preventing self-fertilization in flowering plants is the homomorphic self-incompatibility, which is characterized by molecular pathways rather than physical barriers. Homomorphic self-incompatibility is further divided into two different categories according to the sporophytic or gametophytic origin of the SI pollen phenotype. Homomorphic SI will be deeply described in the first chapter.

Even if the previously described systems which avoid self-fertilization are the most common in nature to prevent inbreeding depression, other methods are currently known, among which the most important are female and male sterility. Sterility of one of two sexual organs in the flower, pistils or anthers, leads to an obligate cross-fertilization. Several mutations ending with female or male sterility are usually eliminated by natural

selection, since the fitness of mutants is low compared to that of wild-type plants. However, in some cases, the sterility trait can be fixed in natural populations as reported in several taxonomic entities, such as in *Phillyrea angustifolia*, which is characterized by a high degree of androdioecious plants and belongs to the *Oleaceae* family as olive. In other cases, male or female sterility, even if not suitable in natural selection, it might be very useful for anthropic purposes. Male sterility for example, can be used as system to produce F1 hybrid varieties in those species, which are recalcitrant to this breeding strategy. Currently, two main types of male-sterility are known and they differ each other for the localization of the male-sterile gene within the cell: nuclear or cytoplasmic male-sterility. In the first case, the gene responsible for male-sterility is located in the nuclear genome, whereas in cytoplasmic male-sterility it is located in the mitochondrial genomes. As consequence of the localization of male-sterile gene some evidences can be deduced. The most obvious is that, in case of nuclear male-sterility, crossing between sterile and fertile plants leads to a segregation of the trait in the F2 progeny, whereas in case of cytoplasmic male-sterility no segregation will be observed since all progeny plants will be sterile due to maternal inheritance of the mitochondrial plastidial genome. In nuclear male-sterility, segregation ratios in F2 as well as in BC1 populations can be used to determine the number of loci involved and the recessive or dominant genetic nature of the trait. Our case-study, male-sterility in red chicory (*Cichorium intybus* L.), represented a perfect example of the analysis of F2 and BC1 segregation data to understand the number and the nature of the gene responsible for male sterility in this species.

Homomorphic self-incompatibility, hereafter called only self-incompatibility (SI), has been studied in olive (*Olea europaea* L.), which is a species characterized by several drawbacks leading to a loss of yield. It is worth noting that olive has a great economic impact in several countries, such as in Mediterranean basin, even though its importance is now widespread all over the world, including Australia and USA. A deep comprehension of the

genetic aspects of reproductive barriers leading to loss of the olive yield is crucial to increase the production in this crop tree. The biggest bottleneck of studying SI in olive is that molecular data in this species are scanty as well as for all woody plants. Moreover, the reproductive behavior of the olive tree, characterized by only one blooming per year that is scalar in the same tree and also in the same inflorescence, represents a further difficulty for any type of molecular study. The present study allowed us to collect new and robust data and to formulate a new hypothesis on SI system in olive, which is in contrast to that currently reported in literature.

Male-sterility has been investigated in chicory (*Cichorium intybus* L.), which is a species characterized by a great economic impact in some countries, particularly in Italy. The mutants characterized in the present study represent the first spontaneous male-sterile mutants ever discovered in this species. The complete absence of pollen in male sterile plants was tested by cytological analysis and the genetic control of male sterility was assessed using segregating populations. A single nuclear gene acting at the recessive state was hypothesized as the factor responsible for male sterility in the chicory mutants.

Both self-incompatibility in olive and male-sterility in chicory represent two distinct reproductive barriers that avoid self-fertilization. However, the meaning of each barrier is completely different and, under the evolutionary point of view, they are antagonist mechanisms. SI is a sexual condition reached after a long period of evolution and it is characterized by different mechanisms that are fixed in specific taxa. The creation of new gene pools from obligate cross-fertilization events is the key to SI fixation and to its widespread among flowering plants, even though several events can occur to breakdown the SI system. By contrast, male sterility that often occurs in some flowering plants across many taxa is always the result of a mutation affecting a single nuclear or cytoplasmic gene, which causes the complete lack or abortion of pollen grains or the formation of non-viable pollen grains. This condition is actually useful to understand the molecular pathways

involved in premature arrest of normal micro-sporogenesis or gametogenesis and so male-sterility represents a valuable tool to breed true F1 hybrids, but it cannot be fitted in any evolutionary context. In fact, in contrast with the self-incompatibility system that is fixed in the olive species, the mutation for male-sterility occurred accidentally in some chicory genotypes and it would be negatively selected in natural populations. In fact, male-sterile plants contribute with half amount of alleles compared to the male-fertile plants, *i.e.* maternal but not paternal gametes, and such a mutant genotype is not counterbalanced by a stronger phenotype in terms of capability to set seeds.

CHAPTER 1

SELF-INCOMPATIBILITY IN OLIVE

Self-incompatibility in olive

Introduction

Cross-fertilization in flowering plants enables to reach a higher genetic variability and consequently imparts a strong evolutionary potential. In nature, several systems leading to outbreeding have been developed during the evolution, such as dioecy, female and male sterility, dichogamy and self-incompatibility (SI). Among all the mechanisms, SI is the most widespread since it has been calculated to occur in almost half of the angiosperms. Because of its wide diffusion, SI has been hypothesized to be a key feature that allowed the diversification and dominance of the angiosperms (Whitehouse, 1950). Distinct traits may be useful to classify different SI mechanisms and they are based on floral morphology and genetic control of the pollen phenotype. According to floral morphology, SI mechanisms are divided into two big categories: heteromorphic SI, in which flowers are characterized by pistils and anthers at different level avoiding thus selfing (*e.g. Primula*), and homomorphic SI, in which selfing is inhibited by molecular pathways. On the basis of genetic control of the pollen phenotype, homomorphic SI can be further divided into two categories: gametophytic and sporophytic, that will be deeply discussed in next paragraphs. Hereafter, the use of self-incompatibility (SI) will be referred to homomorphic self-incompatibility.

The origin of SI is still matter of debate because many questions remain to be solved. First, which is the ancestral condition between self-compatibility (SC) and SI? How are the different SI strategies evolved? Did different SI systems arise from a common ancestral or several times independently during the evolution? At present, since most flowering plants are self-compatible, SC is generally considered to be primitive condition (Bateman, 1952). However, SI is quite spread among the angiosperms, since it has been found in more than 100 families and occurs in an estimated 39% of species (Igic *et al.*, 2008). In addition, due

to this broad presence among flowering plants, the origin of SI is supposed to be very ancient. It is now generally accepted the hypothesis of SI evolution from SC and, through the years, several events led to a breakdown of SI and a return back to the self-compatibility or pseudo-self-compatibility conditions. The main question still unclear is how SI is evolved and differentiated among angiosperms. Phylogenetical analyses on T2-type RNases demonstrated three different categories of these RNases, but only the class III was the one shared by the plants showing a gametophytic self-incompatibility (GSI) system (Igic *et al.*, 2001). On the basis of this finding, the authors hypothesized a common ancestor of *Asteridae* and *Rosidae*, which represent about 75% of all dicots, possessing a RNase-based GSI (Igic *et al.*, 2001), and thus GSI is classified as common ancestral system for incompatibility. Nevertheless, opposite theories hypothesize that RNase-based GSI is evolved from sporophytic self-incompatibility (SSI), which so represent the ancestral incompatibility system. The evolution of RNase-based GSI from SSI should be explained by a shift in the timing of expression of S-locus genes from prior meiosis to after meiosis (Matton *et al.*, 1994). However, it is worth noting that closely related families usually do not share the same system of SI, such as order *Solanales* characterized by families with both SSI (*Convulvolaceae*) and GSI (*Solanaceae*) and, according to this evidence, some authors affirmed that SI arose quite late in the evolution of the families and that SI is evolved independently many times during the evolution (Matton *et al.*, 1994). The occurrence of a common ancestral or an independent origin of SI systems is still matter of debate, also if it is now generally accepted that SI independently arose several times during the evolution of plant species.

Usually SI is characterized by a single locus, called S-locus, which includes at least two genes: one encoded by the maternal parent and another by the paternal parent, which are classified as female and male determinant, respectively. Allele-specific recognition and interaction between determinants trigger the incompatible response ending with the

rejection of self-pollen or self-pollen tube. The presence of one single gene encoding for the female and male determinants is the rule in SI plants, even though in some species, such as the grasses (*Poaceae*), there are two S-loci rather than one single locus (Franklin-Tong *et al.*, 2003). Both female and male determinants are usually characterized by some typical features: i) location of the S-locus in a highly divergent genomic region; ii) organ-specific expression, or preferential expression, that is pistil-specific and anther- or pollen-specific expression of female and male determinants, respectively; iii) high degree of amino acid diversity between different alleles allowing thus the S-haplotype recognition. These general rules are verified only in those cases of incompatibility for which molecular data are known, but different features cannot be excluded for other SI systems that are currently poorly understood. In fact, at present only three incompatibility systems are well known and characterized under the molecular standpoint, but other SI systems are known to be present in nature, even if the genes and molecular mechanisms involved remain still unclear. SI systems are classified into two main classes: sporophytic self-incompatibility (SSI), where the pollen SI phenotype is determined by the diploid genome of the parent, and gametophytic self-incompatibility (GSI), where the pollen SI phenotype is determined by its own haploid genome (review in Takayama *et al.*, 2005). As mentioned before, literature currently reports the existence of several incompatible systems, both SSI and GSI, but only three of them are molecularly characterized, one for SSI (*Brassicaceae*) and two for GSI (one for *Solanaceae*, *Plantaginaceae*, *Rosaceae* and one for *Papaveraceae*), for which genetic determinants and incompatible pathways are verified or hypothesized (**Table 1.1**). Actually, only for SSI of *Brassica* the pathway downstream the pollen-pistil interaction ending with the rejection of self-pollen is currently known, whereas for the other systems the hypotheses have to be demonstrated molecularly. The SSI of *Brassica* has been studied in more details because the worldwide model species *Arabidopsis thaliana*, which belongs to the *Brassicaceae* family, was a SI plant that, after independently events, loss the

SI and now it is perfectly self-compatible. Transformation with genes belonging to its closely related SI species *Arabidopsis lyrata* allowed to demonstrate that some ecotypes of *A. thaliana* returned back to the SI condition, meaning that all the genes of the incompatibility pathway were conserved in the genome. Then, the identification of genes and proteins involved in downstream pathways was easier in Arabidopsis compared to species belonging to other families. The three SI systems and causes leading to a loss of SI will be deeply discussed in next paragraphs.

Table 1.1. List of female and male determinants that are currently known to be involved in self-incompatibility reactions. Determinants are listed according to three incompatibility systems.

		Female determinant			**Male determinant**		
Family	**Type of SI**	**Gene name**	**Type of proteins**	**Size of mature proteins**	**Gene name**	**Type of proteins**	**Size of mature proteins**
Brassicaceae	SSI	**SRK**	Kinase protein	110 kDa	**SCR/SP11**	Cystein rich protein	6 kDa
Solanaceae *Rosaceae* *Plantaginaceae*	GSI	**S-RNase**	RNase (glycosilated)	30 kDa	**SLF/SFB**	F-box protein	50 kDa
Papaveraceae	GSI	**PrsS**	Small glycoprotein	17 kDa	**PrpS**	Transmembrane protein	20 kDa

Further SI systems yet unknown are not excluded and several examples are reported, particularly within SSI. In fact, because molecular studies of SSI have been largely confined to species from *Brassicaceae*, it is not yet known whether SSI, like GSI, can operate through different mechanisms (Hiscock *et al.*, 2003), as reported by several studies in two species: *Ipomoea trifida* (*Convolvulaceae*) and *Senecio squalidus* (*Asteraceae*), that show a strong SSI response but the pathways leading to avoid self-fertilization remain unknown (Kowyama *et al.*, 2000; Hiscock *et al.*, 2003; Suzuki *et al.*, 2004; Tomita *et al.*, 2004). Hiscock *et al.* (2003) reported one case of *Polemoniaceae* where in the same family there were species characterized both by SSI, *Linanthus parviflorus* (Goodwillie, 1997),

and GSI, *Phlox drummondii* (Levin, 1993). This is the only case currently known in contrast with the common situation for which species within the same family share the same incompatibility systems. However several evidences, usually ignored because they missed of molecular data supporting the observations and because they confused the distinct genetic boundary between GSI and SSI, suggest that the division between GSI and SSI is not clear and that in some cases SSI could evolve from GSI by a permanent switch in the timing of pollen S-gene expression (Hiscock *et al.*, 2003).

Breakdown of self-compatibility and its evolutionary implications

Breakdown of self-incompatibility (SI) to self-compatibility (SC) or pseudo-self-compatibility (PSC) occurred several times in several species. Mutations on S-locus, duplication events and rearrangements of the S-locus during the evolution are the main causes leading to the loss of SI. However, evolution toward self-compatibility can also be caused by the shortage of available mates and fixed in the population when the inbreeding depression due to selfing is small (Willi, 2009). One exhaustive example of breakdown of SI is *Arabidopsis thaliana*, which is predominantly SC, even though its close relative *Arabidopsis lyrata* shows SSI typical of *Brassicaceae*. Most likely SC in *A. thaliana* arose independently several times during the evolution, since SC haplotypes showed different modifications and/or rearrangement of the genes within the S-locus (Shimizu, 2008). Furthermore, transformations of SC *A. thaliana* with SRK/SCR genes from its SI *A. lyrata* showed as only some ecotypes get back to the SI condition, whereas the SC in other ecotypes was not affected, meaning that the genes belonging to the signalling pathway downstream the SRK/SCR interaction evolved in a different manner (Boggs *et al.*, 2009a). Beyond the SI restoring in *A. thaliana* through the transferring of functional SRK/SCR from *A. lyrata*, it has been demonstrated that also functional SRK/SCR genes belonging to a different genus, such as the species *Capsella grandiflora*, are able to regain SI condition in *A. thaliana* once transferred (Boggs *et al.*, 2009b). Several studies confirmed that independent events can lead to the loss of SI, and that high or low level of reduction of observed heterozygosity might be used as index to discriminate between ancient and more recent events of loss of SI, as it has been demonstrated by studies of loss of SI observed in *A. lyrata* around Great Lakes of North America (Hoebe *et al.*, 2009). Mutations of genes within the S-locus can be fixed in geographically wide samples. Theoretically, mutations disabling the male determinant are predicted to gain a relative advantage over those disabling the female determinant, because male mutations would increase through both

pollen and seeds whereas female mutations would increase only through seeds (Takayama *et al.*, 2005; Busch *et al.*, 2008). It has been reported that 95% of European *Arabidopsis thaliana* accessions share a 213 bp inversion in the SCR gene, which is responsible of the loss of SI in the species, confirming the theoretical expectations regarding the evolutionary advantage of mutations in male components (Tsuchimatsu *et al.*, 2010). Transgenic regained of SI in *A. thaliana* are now used as crucial tool to study evolutionary and mechanic features of SSI in *Brassicaceae*, allowing to understand genes involved and molecular pathways leading to the arrest of the self-pollen (Rea, 2010). Unfortunately, there are not model species available with the same features in the other incompatibility systems, in which genes and pathways downstream the incompatibility reaction for female and male determinants remain largely unclear. However, several cases of SI breakdown are also reported in *Solanaceae*, which are characterized by S-RNase-based GSI. In *Solanaceae* it has been demonstrated a direct correlation between polyploidy and loss of SI due to the formation of diploid pollen grains (Robertson, 2010). Within *Solanaceae* there are several cultivated species, such as cultivated tomato (*Lycopersicon esculentum*). Cultivated tomato is currently SC but it has been demonstrated that it evolved from SI species of the genus *Lycopersicon* after a breakdown of SI system due to loss-of-function mutations in both S-RNase, the stylar female determinant, and asparagine-rich HT protein, other stylar factor presumably involved in self-pollen tube rejection (Kondo, 2002). Loss of SI can also be due, along with deletion or mutation on S-locus, to duplication of genes within the S-locus. Studies carried out on *Petunia axillaris*, a *Solanaceae* member, showed as duplication of SLF/SFB, the male determinant, induces loss of SI (Tsukamoto *et al.*, 2005). A woody plant was also used to study loss of SI, such as a research performed on Japanese pear (*Pyrus pyrifolia* Nakai), which exhibits a S-RNase GSI. Studies conducted on one SC cultivar of Japanese pear, allowed to identify a deletion of 236 kb spanning from 48 kb upstream to 188 kb downstream of the S-RNase of one S-haplotype, which lacked 34

predicted open reading frames including the S-RNase and the pollen specific determinant (Okada *et al.*, 2008).

It has been demonstrated that loss of SI occurred several times during the evolution of angiosperms but there is no evidence that SI, once lost, has been regained back. Therefore, SI plants can be maintained because of their higher genetic diversity, if compared to the SC plants, where the lower genetic diversity can increase the extinction rate (Igic *et al.*, 2008).

S-RNase-based Gametophytic Self-Incompatibility

S-RNase-based gametophytic self-incompatibility (S-RNase-based GSI), also termed *Solanaceae*-type GSI, is one of the three known and characterized incompatibility systems, and it is the most widespread over the flowering plants. The SI phenotype in GSI is determined by its own haploid genome, differing from the SSI where the pollen SI phenotype is determined by the diploid genome of its parent (Takayama *et al.*, 2005) S-RNase-based GSI is one of the two divergent gametophytic mechanisms molecularly characterized so far and evidences suggest the involvement of RNA and protein degradation in this system (Franklin-Tong *et al.*, 2003; McClure, 2004; Takayama *et al.*, 2005; Newbigin, 2008; Zhang *et al.*, 2009), differing from the other type, *Papaveraceae*-type GSI, where instead the interaction between proteins encoded from female and male determinants lead to Ca^{2+}-depending signaling network that ultimately results in a programmed cell death (PCD) of incompatible pollens (Takayama *et al.*, 2005; Wheeler *et al.*, 2009; Zhang *et al.*, 2009). It is worth noting that, even though the first taxa studied for this system was *Solanaceae* family, S-RNase-based GSI is also found in other taxa, such as *Rosaceae* and *Plantaginaceae*, and it is not excluded that other families can share this system but they are still uncharacterized. The rejection of self-pollen tubes in S-RNase-based GSI occurs at different levels of transmitting tissues of the style always ending with the arrest of growing self-pollen tubes (**Figure 1.1**). The arrest of self-pollen tubes in styles allows to easily recognize species showing S-RNase-based GSI from the other well-studied SI systems, *Papaveraceae*-type GSI and SSI, which are instead characterized by the rejection of self-pollen at the stigma surface. For example, an aniline blue staining, which is specific for callose, the polysaccharide typical of growing pollen tubes, might be used as fast and cheap screening to test the presence or absence of pollen tubes through the style and so to discriminate among SI systems.

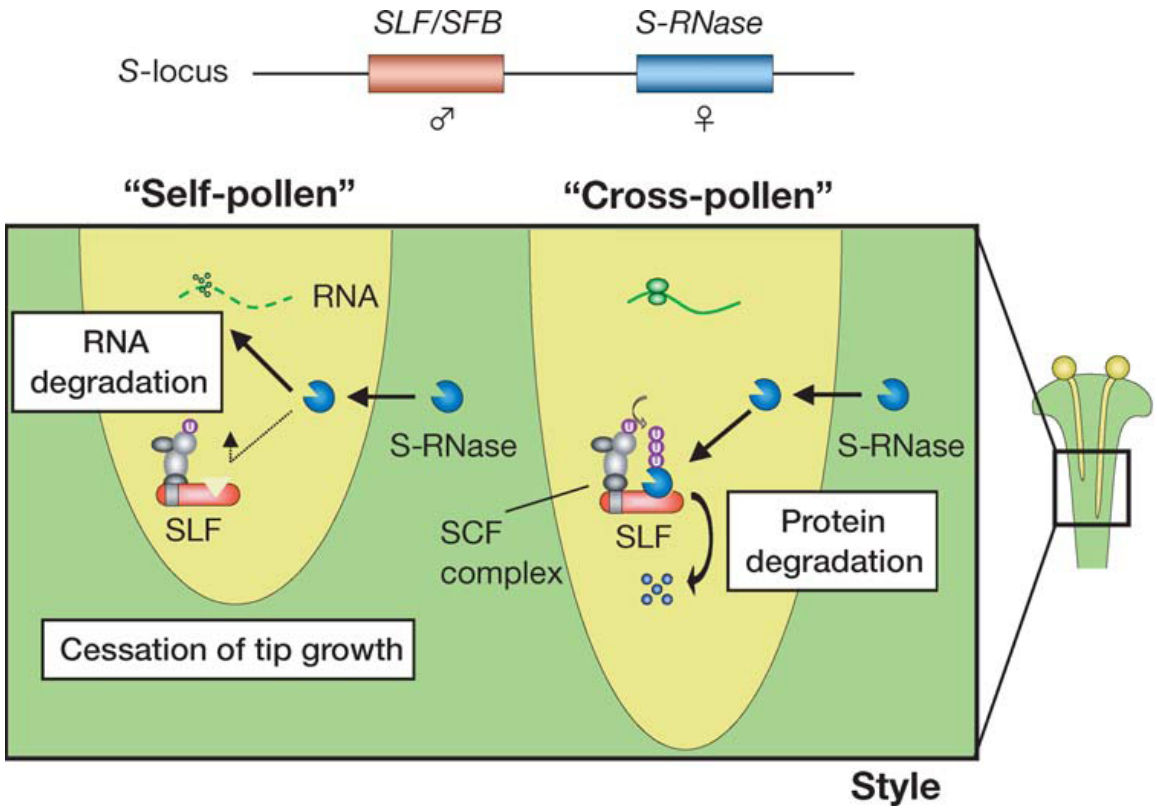


Figure 1.1. Molecular model of self-incompatibility response in *Solanaceae*, *Rosaceae* and *Plantaginaceae* (S-RNase-based GSI) families (from Takayama *et al.*, 2005).

S-locus in S-RNase-based GSI includes at least two genes: *S-RNase* and *SLF/SFB*, the female and male determinants, respectively. S-RNase is the first characterized at the end of 80's, whereas the male determinant was discovered almost 15 years later. The finding of the male determinant allowed to hypothesize the "inhibitors model", where SLF/SFB transcribed in growing pollen tubes recruits cross-S-RNases, but not self-S-RNases, encoded instead through the style and transferred in high quantity inside the growing pollen tubes, and this recruitment leads to the ubiquitination and subsequent degradation of cross-S-RNases, avoiding thus the citotoxic activity of S-RNase (reviewed in Takayama *et al.*, 2005; Zhang *et al.*, 2009). Even if this model seems quite robust, some other aspects remain unclear. In fact, evidences supporting the hypothesis are verified by *in vitro* assays but *in vivo* experiments are needed to get further confirms, in particular to shed light on

mechanisms explaining the competitive interaction between self- and cross-pollen tubes, since both of them can interact with the same S-RNase.

It has been demonstrated that the S-locus is usually localized in heterochromatic regions where the recombination is suppressed, even though in *Solanaceae*, *Rosaceae* and *Plantaginaceae* that share the same incompatibility system the S-locus is apparently localized in variable regions of the genomes, including a centromeric region in *Solanaceae* and a distal part of the chromosome in *Plantaginaceae*. Consequently, these findings lead to hypothesize a role of the chromosomal localization in the maintenance of S-haplotypes (Yang *et al.*, 2007).

Female determinant

Female determinant was firstly identified in the self-incompatible *Nicotian alata* and it consisted on a small stylar glycoprotein of about 30 kDa, that co-segregate with the S-haplotype in genetic crosses and have a ribonuclease activity, for which was then called S-RNase (McClure et al, 1989). This finding allowed the identification and cloning of many related S-RNases from other *Solanaceae* members as well as from other families. Overall sequence data revealed that S-locus proteins contain a region homologous to the catalytic domain of the fungal T2-type ribonucleases (Takayama *et al.*, 2005). Furthermore, phylogenetic analyses on T2-type RNases demonstrated the existence of three different categories of T2-type RNases, but the class III was the only one shared by all plants having an S-RNase-based GSI system (Igic *et al.*, 2001). The S-RNases are expressed exclusively in the pistil and their role played in SI was directly confirmed by gain- and loss-of-function experiments (Lee *et al.*, 1994; Murfett *et al.*, 1994). Despite the highly divergent amino acid sequence identity, ranging from 38% to 98% (McCubbin *et al.*, 2000), S-RNases from *Solanaceae* contain a number of conserved regions. In particular, five highly conserved regions were found, named C1 through C5, being these regions similar to those from

Rosaceae and *Plantaginaceae*, except for the forth region that is absent (Takayama *et al.*, 2005). The crystal structure of two S-RNase from *Nicotiana alata* and *Pyrus pyryfolia* has been determined and it consists on eight helices and seven β-strands with two hypervariable regions, called HVa and HVb, geometrically closed to one another and exposed to the molecular surface (Ida *et al.*, 2001; Matsuura *et al.*, 2001). The S-RNase in self-pollen tubes acts as citotoxin and it is directly involved in degradation of RNA and growth inhibition, but the exact mechanism of how this specific inhibition is achieved still remains largely unclear (reviewed in Zhang *et al.*, 2009). After the isolation of the male determinant, SLF/SFB (S-locus F-box in *Solanaceae* and *Plantaginaceae* or S-haplotype-specific F-box in *Rosaceae*, see following paragraph), the role of S-RNase became clear. The SLF/SFB gene encodes for a protein that, after the interaction with SSK1 (SLF-interacting SKP1-like1) and CUL1-like protein, forms the SCF complex that would be responsible for S-RNase ubiquination (Zhang *et al.*, 2009). Since ubiquinated proteins are usually delivered to the 26S proteasome allowing their degradation, the mechanisms of how S-RNase acts in rejection of self-pollen tubes might be easily explained, but this could make sense only after an allelic-specific interaction between SLF and S-RNases. Nevertheless, no allelic-specific interaction was detected between SLF and S-RNase (Qiao, 2004a; Hua *et al.*, 2006) and so the understanding of the exact mechanism remains still unclear. Although there is no allelic-specific interaction between the two determinants, *in vitro* assays demonstrated that SLF and S-RNase can physically interact and that the interaction between S-RNase and non-self SLF was much stronger than that between S-RNase and self SLF (Hua *et al.*, 2006). Further evidences supporting this observation were provided by the identification of three different functional domains, named FD1 through FD3, within SLF proteins of *Petunia inflata* (Hua *et al.*, 2006), where FD2 was sufficient to interact with S-RNase with high affinity, but this interaction can be weakened by FD1, FD3 or both, which are characterized by one variable region each and this might explain the S-

haplotype specificity. In spite of these *in vitro* experiments, *in vivo* evidences are needed to confirm the exact role of S-RNase in the incompatible response and to understand how the rejecting of self-pollen tubes occurs (Zhang *et al.*, 2009).

Male determinant

The pollen S-determinant was isolated in each of the three S-RNase-based GSI family and named SLF (S-locus F-box) in *Solanaceae* (Sijacic, 2004; Wang *et al.*, 2004) and in *Plantaginaceae* (Lai *et al.*, 2002; Qiao, 2004b), whereas it was called SFB (S-haplotype specific F-box) in *Rosaceae* (Entani *et al.*, 2003; Ushijima *et al.*, 2003). A good candidate for the male determinant in SI response has to be characterized by some typically features, such as localization in highly divergent genomic regions of the S-locus, demonstration of S-haplotype specific diversity and anther-specific expression (Takayama *et al.*, 2005). A candidate fulfilling all the features needed for a pollen S-determinant was isolated in *Prunus mume* by Entani *et al.* (2003) and additional analyses allowed to identify two hypervariable regions, HVa and HVb, at the C terminus (Ushijima *et al.*, 2003; Kao *et al.*, 2004), which are included one in FD1 and the other in FD3, two out of three conserved domains that, along with FD2, characterized SLF proteins (Hua *et al.*, 2006). Hypervariable regions should explain allelic-specific nature of male determinants and the allelic-specific recognition between SLF and S-RNase, as previously described.

Papaveraceae-type GSI

At present, two highly divergent GSI systems have been characterized: S-RNase-based type and *Papaveraceae*-type, whose occurrence has been demonstrated only in the *Papaveraceae* family. Although this system is gametophytic as the S-RNase-based method and the S-haplotype of pollen is determined by its haploid S-genotype, the two pathways leading to the arrest of self-pollen tubes growth are completely different (reviewed in Takayama *et al.*, 2005; Wheeler *et al.*, 2009; Zhang *et al.*, 2009). The first macroscopic difference is the point where the interaction between determinants takes place: while in S-RNase-based GSI the incompatible reaction occurs at the different level of the style, in *Papaveraceae*-type GSI the incompatible reaction takes place in the stigma surface. Consequently, the growth of emerging pollen tubes is blocked before they penetrate the stigmatic surface (**Figure 1.2**).

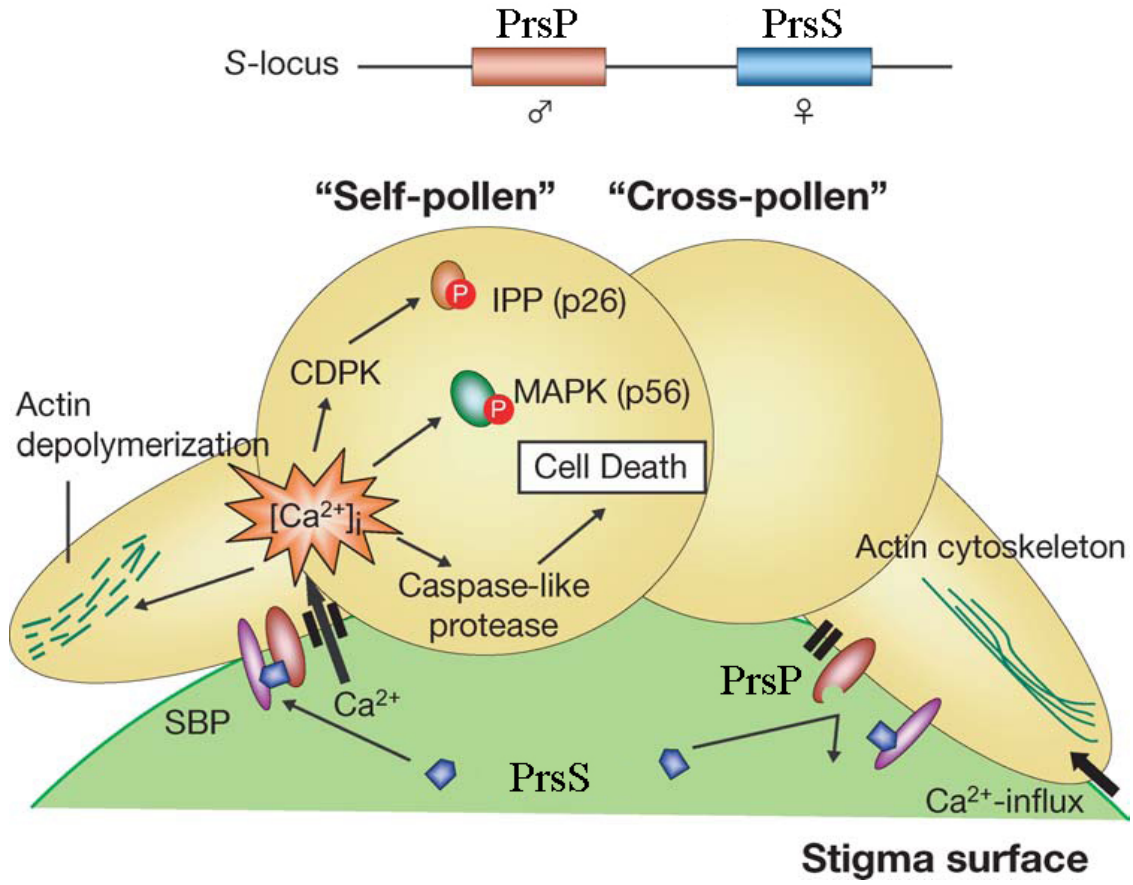


Figure 1.2. Molecular model of the self-incompatibility response in the *Papaveraceae* (*Papaveraceae*-type GSI) family (modified from Takayama *et al.*, 2005).

The predicted pathway leading to the arrest of growth of self-pollen tubes is completely different from that of the S-RNase-based GSI system. In *Papaveraceae*, the interaction between compatible determinants triggers a Ca^{2+}-dependent signalling cascade in incompatible pollen tubes allowing a rapid increase in cytosolic free Ca^{2+} (Franklin-Tong, 1999). It has been supposed that the increase in cytosolic free Ca^{2+} acts as starter of an intracellular signalling network with several targets, resulting in the rapid inhibition of pollen tubes tip growth and death of incompatible pollen grains (Franklin-Tong et al, 2003). Both female and male determinants are now known as well as their main features.

Female determinant

The first gene candidate as female determinant in SI of *Papaver rhoeas* was cloned in the early 90's (Foote *et al.*, 1994), when researchers isolated one S-gene with no significant similarity to any previously reported gene sequence, including S-genes from other species. This gene was a single-copy gene and *in vitro* assays showed that the isolated S-gene had S-specific pollen inhibitory activity (Foote *et al.*, 1994). Since no homology with any other gene was found, they also hypothesized a new SI system occurring in *Papaver*. We will refer to this gene as PrsS, for *Papaver rhoeas* stigma S determinant, according to the recent classification given by Wheeler *et al.* (2009) after the discovery of the S-pollen determinant, called PrpS (*Papaver rhoeas* pollen S determinant). PrsS encode for a small (ca. 17 kDa) secreted protein expressed in the stigma (Franklin-Tong *et al.*, 2003). PrsS proteins are highly polymorphic and share between 51.3% and 63.7% amino acid sequence identity and, in contrast to the S-determinants of *Brassicaceae* and *Solanaceae*, amino acid sequence variation is not found in hypervariable blocks, but rather throughout the S-proteins (Takayama *et al.*, 2005). The backbone peptide (ca. 15 kDa) is polymorphic throughout the S-pollen protein, which has four conserved cysteine residues (Higashiyama,

2010) and a predicted conserved secondary structure that is comprised of six β-strands and two α-helices connected by seven hydrophilic surface loops (Takayama *et al.*, 2005).

Male determinant

The finding of the male determinant in *Papaveraceae* represents the major advance in the understanding of the *Papaver* SI system (Wheeler *et al.*, 2009). Before the recent discovery, the pollen S gene was predicted to be a receptor located at the pollen plasma membrane (Franklin-Tong *et al.*, 2003) since particular residues located in hydrophilic surface loops of PrsS are crucial for the recognition of S-pollen and thus was though that PrsS might act as signaling ligand (Jordan *et al.*, 1999; Kakeda *et al.*, 1998). A putative pollen receptor has been identified in *Papaver*, called SBP (S-protein binding protein), that specifically binds stigmatic PrsS proteins but not in a haplotype-specific manner, and for this reason it was proposed that it might act as an accessory receptor rather than as the real male determinant (Franklin-Tong *et al.*, 2003). Only recently Wheeler *et al.* (2009) cloned the real male determinant in *Papaveraceae*, termed PrpS. PrpS is a single copy gene linked to the pistil PrsS gene and encode for a novel 20 kDa protein (Wheeler *et al.*, 2009). Furthermore PrpS is associated with the plasma membrane, confirming the previous prediction of a male determinant encoding for a transmembrane protein (Wheeler *et al.*, 2009).

Sporophytic self-incompatibility

Sporophytic self-incompatibility (SSI) is one out of the three incompatibility systems for which molecular determinants are known. Differing from S-RNase-based GSI and *Papaveraceae*-type GSI systems, the SSI is a sporophytic system and the pollen SI phenotype is determined by the diploid genome of the parental plant (Takayama *et al.*, 2005). SSI was initially discovered in *Brassicaceae* and the inhibition to the self-fertilization is reached by avoiding pollen hydration or by inducing a rapid arrest of the pollen tube growth at the stigma surface (Takayama *et al.*, 2005). Self-pollen rejection is localized to the site of pollen-stigma contact and usually a single stigma epidermal cell, papilla cell, can inhibit a self pollen grain and, at the same time, it can allow the germination and growth of non-self pollen tubes (Tantikanjana *et al.*, 2010). The trigger of the incompatibility response is achieved by the interaction of two highly polymorphic genes encoding for the female determinant, S-locus receptor kinase (SRK), and the male determinant, S-locus cystein rich protein (SCR) (Takayama *et al.*, 2001; Kachroo *et al.*, 2002; Takayama *et al.*, 2005; Zhang *et al.*, 2009; Tantikanjana *et al.*, 2010). A further S-locus protein encoding for a secreted glycoprotein (S-locus glycoprotein, SLG), which is the first protein belonging to the S-locus isolated from *Brassica* (Nasrallah *et al.*, 1985), could amplify the SSI response though its presence is not required to trigger the SI reaction (Takayama *et al.*, 2001; Kachroo *et al.*, 2002). All of these genes are tightly linked in the same locus and are typical of each S-haplotype. The interaction between female and male determinants occurs at the plasma membrane, even though studies on the sub-cellular localization of one SRK of *Brassica* showed that only a small fraction localized to the plasma membrane whereas the majority of the protein was found in endosomes (Ivanov *et al.*, 2009). So far, the SSI is the best-known incompatibility system since the involved proteins and molecular pathways have been already uncovered and experimentally validated. One accepted model predicts the interaction of more membrane-anchored

proteins, two SRK and one MLPK (M-locus protein kinase), and this interaction triggers a phosphorilation and ubiquination cascades ending with the negative regulation of another protein, Exo70A1, which is essential for proper germination of pollen grains (reviewed in Zhang *et al.*, 2009; Tantikanjana *et al.*, 2010). MPLK proteins are known to include two distinct forms in *Brassica rapa* both of which localize in the plasma membrane and interact with SRK, and they are crucial for the incompatibility response since their inactivation leads to a complete breakdown of SI reaction (Murase *et al.*, 2004; Kakita *et al.*, 2007). Phoshorylated SRK and MLPK interact and phosphorylate the ARC1 protein (Armadillo Repeat-Contain protein 1), which is a U-box E3 ubiquitin ligase (Gu *et al.*, 1998; Stone *et al.*, 1999; Stone *et al.*, 2003). Recently it was proved that Exo70A1, a putative component of the exocyst complex involved in polarized secretion in yeast and animals, plays a crucial role in SSI. Expression reduction of Exo70A1 disrupts compatible pollen tube growth, whereas its overexpression leads to a partial breakdown of SI (Samuel *et al.*, 2009), leading to suppose that Exo70A1 is required for compatible pollination and the ARC1-mediated degradation of Exo70A1 leads to self-pollen rejection by inhibiting polarized secretion of compatibility factors (reviewed also in Tantikanjana *et al.*, 2010) (**Figure 1.3**).

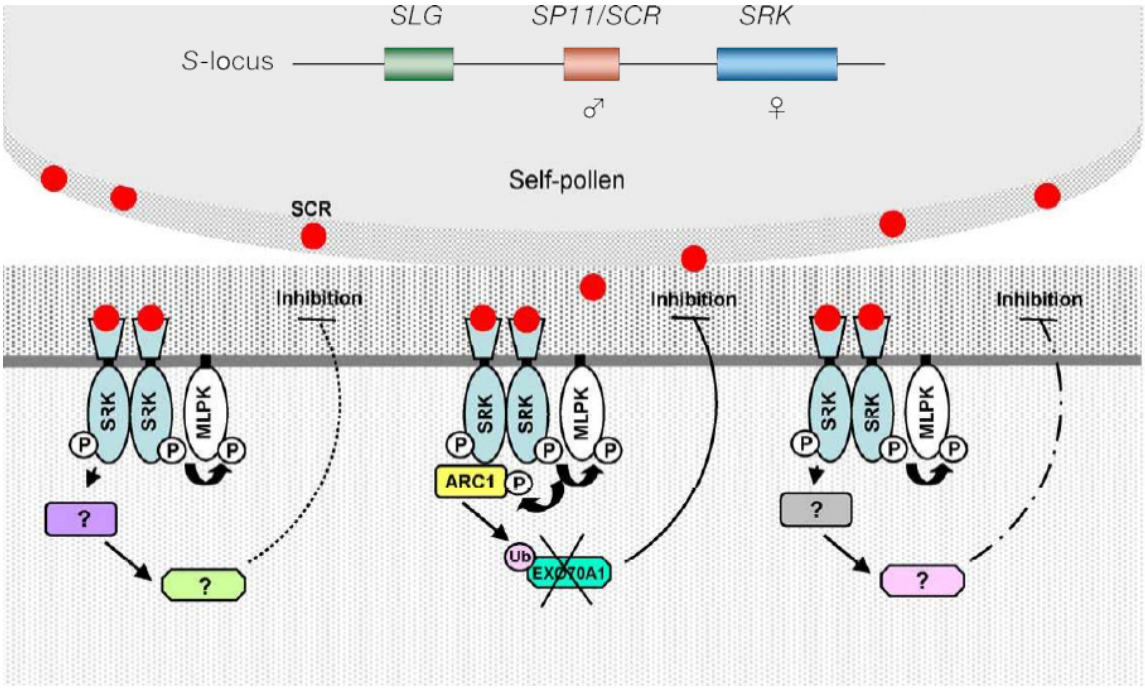


Figure 1.3. Molecular model of the self-incompatibility response in the *Brassicaceae* (SSI) family (from Tantikanjana *et al.*, 2010).

It is worth noting that the SSI genes as SRK, MLPK and ARC1 are predominantly expressed in the stigma, whereas Exo70A1 is widespread along all the tissues, leading to the evidence that the SSI response is the result of an interaction between stigma-specific and ubiquitous proteins (Tantikanjana *et al.*, 2010). However, recent studies carried out on transgenic SI lines of *A. thaliana* reported that, differently to what reported for *Brassica*, in *A. thaliana* inactivation of the MLPK ortholog (AtAPK1) and overexpression of Exo70A1 neither abolish nor weaken the SI response(Kitashiba, 2011). Even though this model is accepted and demonstrated, at least for *Brassica*, several evidences allow to hypothesize the existence of ARC1/Exo70A1-independent signalling pathways that use currently unknown components and that may have redundant or overlapping roles with the ARC1-mediated pathway, or that may inhibit other essential steps preventing thus the self-pollination (Tantikanjana *et al.*, 2010).

Female determinant

The first S-locus gene isolated in *Brassica* was SLG, an S-locus glycoprotein, which encodes for a 50-60 kDa glycoprotein with several N-linked oligosaccharides and 12 conserved cysteine residues (Nasrallah *et al.*, 1987; Takayama *et al.*, 1987). This finding led to the isolation of a second S-locus gene: SRK, S-locus receptor kinase (Stein, 1991), which consists of an SLG-like extracellular domain (S-domain), a transmembrane domain, and an intracellular serine/threonine kinase domain (Takayama *et al.*, 2005). SLG and SRK show all the features that would be expected for the female determinant of SI, such as: i) predominant expression in the papillae of stigma into direct contact with pollen; ii) expression that occurs just prior to flower opening and coincides with the timing of SI acquisition by the stigma; and iii) exhibition of allelelic sequence diversity among S-haplotypes (Takayama *et al.*, 2005). However, only SRK is the true female determinant, whereas SLG can play as enhancer of the incompatible response in some S-haplotypes, but not in all, being not essential in SI reaction (Takasaki *et al.*, 2000; Silva, 2001). The first studies aimed to study the localization of the SRK protein within the cell were performed using transgenic lines of Tobacco and showed as SRK was first glycosilated and then targeted to the plasma membrane (Stein, 1996). Recently, protein localization studies conducted by means of specific antibody against SRK in *Brassica oleracea* showed as SRK proteins localized preferentially to cytoplasmic compartments, such as endosomes, and to a lower extent to the plasma membrane (Ivanov *et al.*, 2009). These results seem not in agreement with the expected localization of SRK in plasma membrane for the direct receptor-ligand interaction with the male determinant (SCR). The same authors showed, on one hand, that the amount of SRK protein present in the plasma membrane was not sufficient to trigger the incompatible response (Ivanov *et al.*, 2009) but, on other hand, it was proved that MLPK protein, which is essential for a correct rejection of self-pollen, is located in the plasma membrane and it plays a crucial role in SI rejection interacting with

SRK (Murase *et al.*, 2004; Kakita *et al.*, 2007). For these reasons the authors hypothesized that the SRK within the endosomes might be activated by the receptor-ligand reaction in the plasma membrane, leading to a localized incompatible reaction in a specific part of the stigmatic cell and leaving the same cell free to interact with other self- and cross-pollen grains. This hypothesis would justify the very particular feature of the SI signalling, where a single papilla cell can react simultaneously to self- and cross-pollen grains (Ivanov *et al.*, 2009).

Male determinant

Similarly to the other SI systems, the discovery of the male determinant took place several years after the finding of the female determinant. Once it was proved that the biological activity responsible for SI was included in a small protein fraction (<10 kDa) of the pollen coat (Stephenson *et al.*, 1997), parallel researches led to the isolation of the male determinant. Two independent and different genetic approaches were followed and both achieved the same results. Sequencing of the entire S-locus containing SLG/SRK and the use of fluorescent differential display, allowed to isolate the male determinant, which was characterized by two exons divided by one intron of 4,1 kb in length. This determinant, which was termed SCR (S-locus cysteine rich) or SP11 (S-locus protein 11), was attributable to a single locus whose gene had an anther-specific expression (Schopfer, 1999). Analyses of SCR/SP11 in further S-haplotypes confirmed its role as male determinant, such as S-haplotype-specific sequence polymorphism, localization in the proximal flanking region of the SRK gene, expression in the tapetum of the anther, and recombinant SCR/SP11 lead to the inhibition of hydration of the cross-pollen (Takayama *et al.*, 2000). SCR/SP11 encodes for secreted, small, basic cysteine rich proteins, shows an extensive haplotype-associated polymorphism in which the alleles share a relatively conserved signal sequence but have mature proteins that are highly variable (from 19,5%

up to 94% amino acid identity), suggesting a strong positive selection for diversification (Takayama *et al.*, 2005). Despite of its highly variable amino acid identity, few residues are highly conserved among S-haplotypes, such as eight cysteine residues (termed C1 through C8), a glycine residue between C1 and C2 and an aromatic residue between C3 and C4 (Takayama *et al.*, 2000; Watanabe *et al.*, 2000; Shiba, 2002). The tertiary structure of S_8-haplotype was determined and it consists of a compact fold, with the N and C termini located near each other; it folds into an α/β sandwich structure comprising a layer of a twisted three-stranded β-sheet backed by another layer formed by an α-helix with loops flanking it; the layers are united by two disulfide bonds between cysteine residues 16-35 and 24-45, another disulfide bond between Cys-33 and Cys-47 for the stabilization of the β-sheet and the fourth disulfide bond between Cys-6 and Cys-50 bundles the N- and C-terminal segments of this protein (Mishima *et al.*, 2003). Interestingly, *in situ* hybridization analyses performed in anthers have showed that SCR/SP11 is expressed both sporophytically in tapetal cells and gametophytically in microspores (Takayama *et al.*, 2000), even though in same S-haplotypes SCR/SP11 was detected only in the tapetum cell layer of anthers (Shiba, 2002). The expression of the SCR/SP11 in the tapetum is expected and it is in agreement with a SSI system, because it justifies the sporophytical nature of this incompatibility.

Self-incompatibility in olive

Olive (*Olea europaea* L.) is one of the oldest agriculture crops and it has been in cultivation for thousands of years. It was originated in the eastern Mediterranean region and then it has spread to many other parts of the world. Spain, Italy and Greece are the major producers and consumers of olive products in the world, although both production and consumption are increasing in areas outside the Mediterranean basin (Guerin *et al.*, 2007). In spite of its great economical importance, at the present, very few studies have been carried out on biological or physiological aspects leading to yield loss, especially under the molecular standpoint. In fact, like many other hermaphrodite plants, olive is characterized by plentiful flowering followed by poor fruit set and consequently low yield (Stephenson, 1981; Sutherland, 1987; Burd, 1998). The olive tree has both hermaphrodite and male flowers because a high incidence of female sterility, due to ovary abortion, may involve most of the flowers as demonstrated in some cultivars, like for instance the Italian cultivar Dolce Agogia (data not shown), and also confirmed by recent studies (Reale *et al.*, 2006). The high rate of ovary abortion and pistil desiccation in olive flowers has as direct consequence a higher number of male flowers and this evidence shares several analogies with the high incidence of androdioecious flowers in *Phillyrea angustifolia*, which is also a member of *Oleaceae* family (Vassiliadis *et al.*, 2000; Saumitou-Laprade *et al.*, 2010). The high incidence of male flowers in a crop plant as olive, means severe yield loss, since the absence of functional female organs does not allow the production of fruits. The imbalance between female and male flowers is not the only reason influencing fruit production, since besides this feature other reproductive barriers, such as male sterility and self-incompatibility, may occur in olive. As a matter of fact, in addition to the high incidence of ovary abortion, olive is also characterized by pollen sterility. Besnard *et al.* (2000) reported several cases of cytoplasmic male sterility in olive, also if they do not exclude that nuclear male sterility could also occur in olive. Further evidences of several male-sterile

phenotypes, differing in the stage of pollen abortion, have been reported for some of the most common cultivars. However, even if female and male sterility can negatively affect the final yield, self-incompatibility is the most important reproductive barrier influencing olive yield. Despite of the widespread occurrence of self-incompatibility in many olive cultivars, to the best of our knowledge, scanty cytological and no molecular evidences are reported in literature for this species.

Under the morphological point of view olive flowers are adapted to either self- and cross-pollination. In some flowers, anthers are close enough to the stigma so that when they dehisce, the pollen grains fall on the stigma and self-pollination occurs. At the same time, there are flowers where the filaments are flattened so that the anthers spread away from the stigma, thus favoring cross-pollination (Guerin *et al.*, 2007). This might partially explain the existence of both reproductive features in olive. However, despite the first idea that classifies the olive as self-compatible or partially self-incompatible species (Fernàndez-Bolanòs *et al.*, 1969), recent studies using paternity tests confirmed a high degree of self-incompatibility in olive (Mookerjee *et al.*, 2005; Diaz *et al.*, 2006). Furthermore, Serrano *et al.* (2008) showed that when stigma exudates are present, usually anthers are still turgid and intact, explaining as a differential developmental timing may favor cross-pollination in olive, although they did not exclude the possibility of self-pollination between different flowers of the same tree or self-pollination within the same flower at a later moment.

Compatibility tests were carried out in several olive cultivars but results were quite contradictory (Bartolini S, 1995; Cuevas J, 1997), varying particularly between researchers, years, and locations (Cuevas J *et al.*, 2001; Lavee *et al.*, 2002) and, as consequence, the classification of cultivars as self-compatible or self-incompatible is currently very confused. Furthermore, it has been proved that also environmental conditions and type of orchards may also affect the self-incompatibility reaction and the fruit set. Particularly, pollination stage is a critical period because influenced by a wide range of factors. For successful

cross-pollination is required an adequate amount of compatible pollen when flowers are in full bloom and this condition can be reached only if compatible cultivars are blooming simultaneously. Besides the blooming time, the length of flowering can also play a crucial role in cross-pollination and it is variable according to cultivars and years (Ghrisi *et al.*, 1999). Lavee *et al.* (2002) studied 36 olive cultivars over a period of 12 years confirming that the length of flowering depends on climatic conditions. At the end, environmental conditions are directly involved in blooming time and flowering length, which directly influence amount of pollen dispersal and rate of compatible crosses and, consequently, the final yield.

Olive can be considered a pollen-limited plant, where a plant is considered pollen-limited if additional pollen increases fruit or seed production (Larson *et al.*, 2000). In general olive plants produce a great quantity of pollen, as expected from anemophilous plants, and the overall amount is controlled by genetic factors and also influenced by environmental conditions, as demonstrated by Reale *et al.* (2006) in studies performed over a 2-year period. A direct correlation between quantity of pollen released during flowering and production of olives was also confirmed by previous studies (Besselat B, 1992; Fornaciari *et al.*, 1998; Minero *et al.*, 1998). In particular, it was observed that the concentration of pollen during the second ten days of flowering is directly correlated with the final production of fruits (Orlandi *et al.*, 2005). During the progamic phase, pistil tissues can recognize male gametes and discriminate between desirable and undesirable pollen grains (De Graaf BHJ *et al.*, 2001), accepting those from the appropriate species and rejecting pollen from unrelated species or from the same plant through self-incompatibility. Characteristics of pollen and the timing of their landing on stigmas represent a crucial step in accepting or rejecting pollen grains, influencing so the fate of pollen-pistil interactions. Besides these observations, the correct growth of pollen tubes represents a further step that is affected by internal and external factors, such as cell-cell signalling pathways and

environmental conditions, respectively. Hartmann *et al.* (1966) observed that some cultivars had particular temperature requirements for optimum pollen tube growth, and (Bradley MV *et al.*, 1961) found that the effect of temperature on pollen tube growth is dependent on cultivar combination. Furthermore, low temperatures have been found to reduce pollen tube growth, resulting in pollen tubes unable to reach the embryo sac before its degradation (Martin, 1994). On the other hand, relatively high temperatures result in faster growth of pollen tubes (Griggs *et al.*, 1975), although Fernandez-Escobar *et al.* (1983) and Cuevas *et al.* (1994) observed better growth of pollen tubes at 25°C than at 30-35°C. It seems that different environments also influence compatibility among cultivars and too high temperatures have a negative effect on pollination (Androulakis *et al.*, 1990). Environmental factors have also been implicated in fruit set rates (Ghrisi *et al.* 1999) and in the formation of perfect flowers (Lavee *et al.*, 2002). On the whole, these observations contribute to a confused classification of olive cultivars as self-compatible or self-incompatible and the little knowledge represents a severe drawback to the study of self-incompatibility in this species.

Olive is currently classified as S-RNase-based gametophyitic self-incompatible plant, but the classification is based only in botanical features shared with those plants characterized by S-RNase-based GSI, such as wet stigma and binucleate pollen grain. However, scanty cytological observations and no genetic-molecular evidences hamper the understanding of this system. Moreover, our preliminary cyto-histological and bio-molecular observations were not in agreement with S-RNase-based GSI system, leading us to formulate the hypothesis that SSI might occur in olive. Moreover, our new hypothesis seems in agreement with that reported for *Phillyrea angustifolia* (Saumitou-Laprade *et al.*, 2010). *Phillyrea angustifolia* is a member of *Oleaceae* family and, usually, species within the same family share the same type of incompatibility, although in *Polemoniaceae* family there is one case of coexistence of SSI and GSI in the same family (reviewed by Hiscock *et*

al., 2003), where *Linanthus parviflorus* shows SSI (Goodwillie, 1997), whereas in *Phlox drummondii* occurs GSI (Levin, 1993). Following different approaches, Saumitou-Laprade *et al.* (2010) also hypothesized a SSI system occurring in *Phillyrea* because on the basis of genetic analyses they found only two groups of incompatibility, and this finding can be explained only by SSI since with GSI at least 3 groups of incompatibility are required.

To the best of our knowledge, this research is the first study aimed at isolating and cloning genes responsible for, or involved in the SI system in olive, and at testing their expression in different tissues of the plant and in specific organs of the flower. All cytological, genetic and molecular results are presented and critically discussed.

Olive flowers and pollen features

Olive flowers are organized in paniculate inflorescences usually of 10-16 flowers, reaching in rare case 40 flowers. The number of flowers per inflorescence can change among cultivars as well as within the same cultivar, since the number of flowers is directly correlated with nutritional and climatic conditions (Lavee *et al.*, 1996). The individual branches of inflorescences contain from 1 to 4 flowers on short peduncles (Martin, 1990; Rapoport, 1998). Flowers can be bisexual (perfect) or male (staminate), and the proportions are usually depending upon cultivar, growing conditions, "on" or "off" condition, and position on tree. Dimassi *et al.* (1999) recorded a greater proportion of perfect flowers in the middle of inflorescences located in the middle of flowering shoots on the southern (sunny) side of trees, the most favorable location for the growth of individual shoots on trees in the northern hemisphere. Staminate flowers desiccate before than perfect flowers (Rapoport *et al.*, 1991). A perfect flower measures approximately 4-5 mm length, is white in color and it has four sepals joined to constitute a gamosepalous calyx, and four only basal fused petals to constitute a dialypetalous corolla. The staminate flower is instead characterized by a male part well developed and a pistil degenerated. The pistil is composed of a bilobed light-green stigma, a short cream-colored style and a round green ovary flanked by two yellow, turgid and intact anthers (Serrano *et al.*, 2008). It contains four ovules, two in each of two locules (Rapoport, 1998), which are short lived. Pollen is produced in abundance over about 5 days and individual stigmas remain receptive for about 2 days. Flowering in individual trees lasts for about 10 days and in orchards for about 20 up to 30 days. Olive is a wind-pollinated plant but the pollination is hindered by strong winds, rain and high temperature (Connor *et al.*, 2005). After compatible pollination several pollen tubes are growing in the stigma, but usually only one reaches the ovary and only one of the four ovules is fertilized (Rapaport, 1998). The amount and the vigor of pollen are important factors to increase the fertilization rate, especially when the environmental conditions are

suboptimal for ovule fertility and pollen dispersal (Connor *et al.*, 2005). It has been demonstrated that pollen tubes are more vigorous after cross-pollination (Fernandez-Escobar *et al.*, 1983; Cuevas *et al.*, 2001; Ghrisi *et al.*, 1999). We also demonstrated that in the cultivars Leccino, Moraiolo, Dolce Agogia and Frantoio, the percentage of germination of pollen grains is higher after cross-pollination than self-pollination.

Olive is a pollen-limited plant because additional pollen usually increases the fruit production (Larson *et al.*, 2000). Both quantity and quality as well as viability and germinability of pollen grains can vary among cultivars and synchronization between pollen dispersal and stigma receptivity is crucial to reach good fruit yields. The olive pollen has been studied under several standpoints, such as quantity and quality in different cultivars, relationships between amount of pollen and number of flowers, subcellular features and allergenic properties. Studies of pollen performed on five Italian cultivars (Dolce Agogia, Leccino, Frantoio, Moraiolo and San Felice) over two years have demonstrated that the cultivar Dolce Agogia is the one characterized by the strongest precocity and that cultivars Leccino and Frantoio can be considered the best pollinators in terms of percentage of pollen viability and germinability, whereas Dolce Agogia can also be considered a good pollinator in terms of quantity, but it scored the lowest levels of pollen viability and germinability and the highest percentage of aborted flowers and ovaries (Reale *et al.*, 2006).

Studies on olive and grapevine cultivars of the Mediterranean basin revealed a direct correlation among number of flowers, amount of released pollen grains and yield of fruits (Besselat B, 1992; Fornaciari *et al.*, 1998; Minero *et al.*, 1998). Crucial for olive pollination in orchard is the amount of the pollen released from anthers during the second ten days of the flowering period that seems directly involved in fruit set (Orlandi *et al.*, 2005). The presence of staminate flowers in olive may enhance plant fitness, providing an increase in the number of pollen grains available to achieve fertilization (Cuevas *et al.*, 2004). A

correct choice of cultivars within the orchard, such as good pollinators, represents a crucial step to achieve good olive yields.

Studies focused on free and conjugated ubiquitin in olive, which is a small protein usually involved in protein degradation during the late stage of olive pollen development and during *in vitro* pollen hydration, reported a reduction in free ubiquitin level during pollen maturation. Since this reduction does not overlap with a parallel reduction of the corresponding transcripts, as happens in maize, the existence of an active post-transcriptional regulatory mechanism during pollen development seems to occur in olive (Alché *et al.*, 2000). Concentration and localization of lipid bodies in pollen grains have also been studied and it has been suggested that they can represent a significant storage of material used for pollen germination and for early growth of pollen tubes, and they can also play a role in an early polarization and hydration avoiding the loss of water due to the hydrophobic features of lipids (Rodriguez-Garcia *et al.*, 2003).

Specific proteins that cover the exin of pollen grains, named oleosins, are translated both in the tapetal cells and pollen grains, and they represent the source of numerous allergenic reactions in humans. However, some proteins that cover the exin of pollen grains are also involved in cell-cell signalling pathways activating the receptor-ligand reaction, such as in some self-incompatibility systems, in which the male determinant are carried on the pollen coat.

Reproductive cycle of olive

The reproductive cycle of olive is carried out over two years and a single plant is characterized by floral buds of the current year and fruits developed from flowers of the previous year. The presence of the current year's fruit and seed can inhibit the induction of buds during the summer. This inhibition of floral induction by fruit and seed growth might also partially explain the alternative bearing that is characteristic of olive, in which years of intense fruiting ("on") tend to be followed by years of restricted flowering ("off") (Rallo, 1998). However, the interaction between environmental conditions and endogenous pathways over the period ranging from induction in summer to flowering in spring is still unclear. Flowering in olive occurs on buds formed in the leaf axils on shoots produced in the previous year and the mature flower is achieve by three steps: induction, initiation, and differentiation (Connor *et al.*, 2005).

During the induction a change in gene expression and in cytoplasmic content induces the bud to produce a flower. Floral induction occurs in mid-summer (7 to 8 weeks after full bloom) around the time of pit hardening (endocarp sclerification) of the current season's fruit. Unfortunately, no studies on genes involved in floral transition are currently available in olive. However, floral induction seems influenced by compounds released by developing fruits and seeds that are translocated to the buds, detectable by histochemical techniques (Stutte *et al.*, 1986; Rallo *et al.*, 1991; Fernandez-Escobar *et al.*, 1992; Lavee *et al.*, 1996; Fabbri *et al.*, 2000).

The initiation of floral structures required a period of vernalization with temperatures below 7°C allowing thus the release from dormancy (Rallo *et al.*, 1991; Fabbri *et al.*, 2000). Floral initiation can be recognized soon after bud burst about two months prior to flowering in late spring (de la Rosa *et al.*, 2000; Rapoport, 1998). To date, no genes involved in vernalization in olive are yet available and only studies on low molecular mass compounds, such as polyamines and gibberellins are present in literature. It has been demonstrated that

polyamines like spermidine, spermine and their diamine precursor, and putrescine, which are universally distributed in response to biotic and abiotic stresses, flowering fruit development, *in vitro* embryogenesis, morphogenesis and other cellular processes like RNA and DNA replication and protein synthesis (Liu *et al.*, 2006), play a role in floral initiation in olive (Rugini *et al.*, 1985; Ianotta *et al.*, 1996; Pan *et al.*, 1999; Malik *et al.*, 2007). Furthermore, an increased amount of gibberellins and a relative lower amount of acidic inhibitors are detected in flower buds shortly before the floral initiation and it is supposed directly connected with the cold period (Badr *et al.*, 1970). However, the chilling requirement is not absolute, as reported for olives grown in subtropical locations (Ayerza *et al.*, 2001). According to the authors the initiation of floral structures can be also achieved by water stress during the winter, which is thought to play an equivalent role of vernalization.

Differentiation of floral primordia to mature floral organs allows to reach the mature flower consist on inflorescences characterized by a variable number of flowers. Environmental conditions during differentiation play an important role in floral morphology, including number of flowers per inflorescence and the proportion of staminate flowers (Rallo *et al.*, 1981; Rapoport *et al.*, 1991).

Identification of conserved miRNAs in olive

MicroRNAs (miRNAs) represent one of the three major classes of endogenous small RNAs in plants, along with transacting siRNAs and heterochromatic siRNAs, which are involved in both transcriptional and post-transcriptional gene regulation (reviewed by Chen, 2009). The miRNAs are small, noncoding endogenous RNA molecules that regulate gene expression in plants and animals. Biogenesis in plants of these RNAs of about 18-22-nucleotides in length is a multi-step process, which includes the cooperation of several proteins. A plant *MIR* gene is transcribed by RNA polymerase II and transcripts are usually subjected to maturation processes as a normal mRNA, such as splicing, 5'-capping and poly-adenylation, ending with a primary miRNA (pri-miRNA). The pri-miRNA is processed into the stem loop precursor (pre-miRNA) by a DICER-like protein (DCL1 in *Arabidopsis*) and other proteins, such as SERRATE (SE) and HYPONASTIC LEAVES 1 (HYL1). The pre-miRNA is further processed by the same protein complex into a duplex of miRNA and miRNA*. The miRNA-miRNA* duplex is then methylated on the 2'-OH of the 3' terminal nucleotides by HUA ENHANCER 1 protein (HEN). One strand of the duplex is loaded into an AGO1-containing miRISC complex and it is exported from the nucleus to the cytoplasm through export factors, including HASTY. In the cytoplasm, the action takes place by cleavage or translation arrest of target mRNAs. Moreover, in the cytoplasm the SDN1 family of exonucleases may degrade single-strand miRNAs to limit their steady-state levels (reviewed by Chen, 2009).

During the last decade, the crucial role played by miRNAs in gene regulation has been an attraction for many researchers. Furthermore, new massive genomes and ESTs sequencing technologies enabled to collect huge amounts of data that represent a great source to search small RNAs. The parallelism between new massive sequencing technologies and their decreasing costs allowed to apply these methodologies also to those species that so far were not considered as good candidates for research. Among these species there are several crop

plants, such as olive (*Olea europaea* L.). To date, despite its high impact in some local and regional economies, a very few studies on biomolecular aspects have been carried out in olive and information is scanty in literature and gene aspects. The understanding of metabolic pathways and their regulatory processes represents a great deal for those species in which several problems occur influencing negatively the final yield. Olive belongs to *Oleaceae* family, a taxa containing several genera with economic importance (*Olea*), forestry importance (*Fraxinus, Phyllirea* and *Ligustrum*) and ornamental importance (*Siringa, Jasminum* and *Forstizia*).

Recently it has been demonstrated that small RNAs may interact with SI genes determining allele dominance effects and loss of SI. In SSI of *Brassicaceae*, the recessive allele of male determinant *SP11* can be methylated, inhibiting its expression. It has been proved that inverted genomic sequences similar to the recessive *SP11* promoter in the flanking regions of dominant *SP11* alleles, which were specifically expressed in the anther tapetum and processed into 24-nucleotide RNA (called *SP11* methylation inducer, *Smi*), were the responsible of the methylation (Tarutani *et al.*, 2010) Finnegan *et al.*, 2011). A further small RNA is supposed to be the cause of loss of SRK function in *Arabidopsis thaliana*, leading to the breakdown of SI in this species (Sunkar *et al.*, 2004)(Lu *et al.*, 2006).

Our goal was to investigate miRNAs in olive flowers by screening the four 454 pyrosequencing cDNA libraries available, searching for putative miRNAs involved in the regulation of olive SI-related genes. The olive miRNAs collected were also used to obtain a state-of-art of the conservation of different miRNA families among different species.

To the best of our knowledge, this study is the first aimed at isolating and characterizing the miRNAs in the *Oleaceae* family and it represents the basis for further researches in other species within the same family.

Materials and methods

Plant materials

Plant materials, such as flowers, fruits, leaves and roots, were collected from both experimental fields and greenhouses of CNR at Perugia (Umbria, Italy) through different years, from 2006 to 2011, according to the experiments. Plant materials for RNA extraction were immediately immerged in liquid nitrogen and then stocked at -80°C until use. Plant materials for DNA extraction were collected and stored at -20°C. Plant materials for both cytological analyses by means of aniline blue staining and for *in situ* hybridization were collected and immediately immerged in proper fixative solutions, according to the protocols listed below. Plant materials were collected from both self-compatible, cv. Frantoio and cv. Kalamata, and self-incompatible cultivars, cv. Leccino, cv. Moraiolo and cv. Dolce Agogia. Plant materials were collected after both self-pollination and cross-fertilization according to the experiments.

Aniline blue staining

Aniline blue staining was performed in order to study the behaviour of pollen grains and pollen tubes in self- and cross-pollinated pistils of one self-compatible (Frantoio) and three self-incompatible (Leccino, Moraiolo, Dolce Agogia) cultivars. Flowers were fixated in Carnoy solution (absolute ethanol and glacial acetic acid 3:1), stained by means of a solution 0.1% aniline blue in phosphate buffer and 50% glycerol and then visualized under epi-fluorescence microscope. To get statistical validation, aniline blue staining was also applied to about 34,000 pollen grains and the germination rates recorded for each cultivar by considering six different classes: a) pollen grains not germinated; b) pollen grains germinated on the papillae cells; c) pollen grains germinated on the stigma surface; d) pollen tubes reaching the proximal part of the style; e) pollen tubes reaching the middle part

of the style; f) pollen tubes reaching the distal part of the style, after both self- and cross-fertilization.

Aniline blue staining was also performed in transversal sections of both self- and cross-pollinated pistils of cv. Frantoio and cv. Leccino.

Iodine-Potassium-Iodide (IKI) staining

Transversal sections of styles of self-incompatible cv. Leccino and self-compatible Frantoio after self-pollination were stained with a Iodine-Potassium-Iodide (IKI) solution, which is able to mark starch content within the cell. The length of the starch molecules determines the color of the reaction: red color for shorter molecules and dark blue color for the longer molecules. The staining was used as indirectly marker of growth of pollen tubes through the style, since the starch might be consumed by the heterotrophic growth of compatible pollen tubes. The IKI solution is prepared by first dissolving 2 g of KI in 100 ml of water and then adding 0.2 g of iodine into the KI solution.

Scansion electronic microscopy (SEM)

SEM analyses were performed on the incompatible cultivar Leccino after self-fertilization and the attention was focused on the stigma. Samples were first fixed in 2,5% gluteraldehyde solution at 4°C for 4 hours and then washed with distillated water 5 times for 30 minutes each. Samples were then fixed in 2% osmium tetroxide solution at 4 °C for 4 hours and washed with distillated water 5 times for 30 minutes each. Specimens were dehydrated through washes in ethanol at serial concentrations (25%, 50%, 75%, 95%, 100%) for 10 minutes each step and specimens were mounted with silver paste and coated with gold alloy. Samples were visualized under the Scansion Electron Microscope (SEM).

Isolation of DNA and RNA samples, and synthesis of cDNAs

Molecular analyses were performed using total RNA extracted from flowers, along with other tissues such as fruits, leaves and roots, of cultivars Leccino (self-incompatible) and Frantoio (self-compatible) at different developmental stages by RNeasy Plant Mini Kit (Qiagen) and treated by DNAse (Qiagen) according to the manufacturer's protocols. RNA samples were then run in agarose gel to test the integrity and quantified using nano-drop spectrophotometer. cDNAs were produced from total RNA through RT-PCR and, according to the experiments, it was used SuperScript VILO cDNA Synthesis Kit (Invitrogen), which is characterized by random primers, and SuperScript III (Invitrogen) using both Oligo_dT and gene-specific primer as starter of the reaction.

DNA samples were extracted from cultivars Leccino, Frantoio, Kalamata, Moraiolo and Dolce Agogia using GenElute Plant Genomic DNA Miniprep Kit (SIGMA) according to the manufacturer's protocol.

Isolation of the SI determinants

Female and male determinants of S-RNase-based GSI and SSI systems, typical of the *Solanaceae* and *Brassicaceae*, respectively, were searched by means of PCR using degenerated primers. Genes belong to S-RNase-based GSI, such as *S-RNase* (S-locus RNase) and *SLF*/*SFB* (S-locus F-box protein), female and male determinants respectively, belong to different taxa, *Solanaceae*, *Rosaceae* and *Plantaginaceae*, were recovered from NCBI nucleotide database and used to produce a multiple alignment. Conserved region within the multiple alignments were used to design degenerate primers. Several amplifications and several conditions were attempted in olive cDNAs and the amplicons were sub-cloned using the TOPO-TA Cloning Kit (Invitrogen) according to the manufacturer's protocol and sequenced. No sequences with significant homology with genes involved in S-RNase-based GSI were retrieved. After that, the analyses were focused

on *SRK* (S-locus receptor kinase), *SLG* (S-locus glycoprotein) and *SCR* (S-locus cysteine rich protein), female determinant, enhancer of the incompatibility response and male determinant of the SSI, respectively. Using a similar S-RNase-based GSI approach, several sequences encoding for S-locus genes belong to *Brassicaceae* family were used to produce multiple alignment and degenerate primers were designed on conserved region. Amplifications, sub-cloning and sequencing were carried out and allowed to retrieve sequences with high homology with searched genes, *SRK* and *SLG*, which were then used as starter for RACE experiments in order to obtain the full-lengths. Approach based on degenerate primers did not yield results for the male determinant, *SCR*, because of its high nucleotidic diversity also within the same species. Different approaches were so used to isolate putative candidates. Flower-specific 454 pyrosequencing libraries of cDNAs of olive flowers were screened for the identification of *SCR*-like sequences, but, despite the 464969 accessions available, no sequences having reference with SCR were found. After this preliminary screening, other three approaches were attempted. The first approach was to search for sequences encoding for protein with similar amino acidic composition, such as defensin protein, which exhibit 8 conserved cysteins as SCR, even though their distribution within the protein is slightly different. This approach enabled to isolate 3 accessions. The second approach consisted on searching for sequences according to the following parameters: i) belonging to pollen-pistil interaction pathways in different species; ii) anther-specific expression; iii) differentially expressed between the incompatible cv. Leccino and the compatible cv. Frantoio. This approach allowed to isolate further 10 candidates. The third approach was instead attempted by screening the flower-specific 454 pyrosequencing libraries searching for *SCR*-like sequences using an algorithm built *ad hoc* according to the guidelines reported in literature for cysteine-rich proteins and tertiary structure of one SCR of *Brassica* (Schopfer *et al.*, 1999; Vanoosthuyse *et al.*, 2001; Mishima *et al.*, 2003; Higashiyama, 2010), and it was designed after dataset assembling and protein multi-

alignment. 454 pyrosequencing libraries were first assembled and the assembling of the original 464,969 ESTs produced 94,546 accessions, including isotigs, contigs and singlets. The assembled dataset was translated according to the 6 reading frame and the search for the cysteine rich pattern was carried out using aminoacidic sequences. In order to determine the cysteine rich pattern, the alignment of 43 sequences characterizing the SCR protein, http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=148474, was performed (**Figure 1.4**).

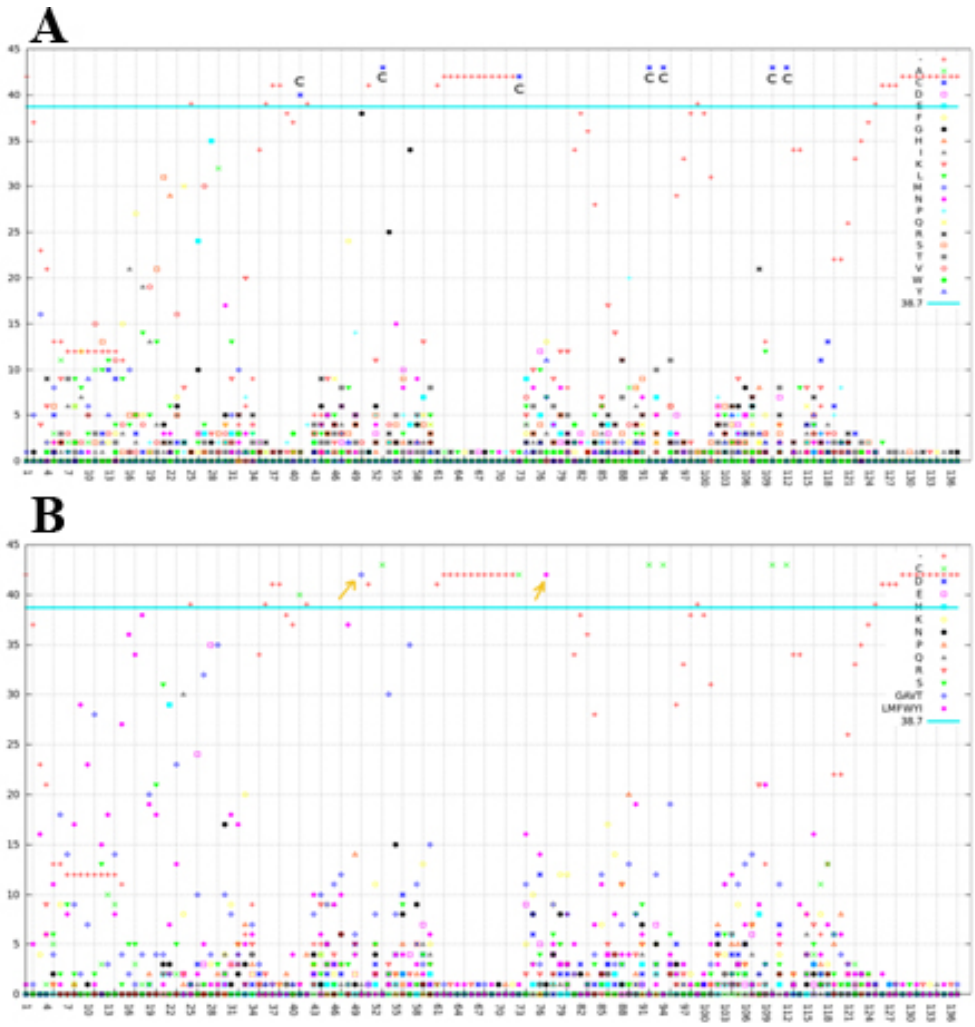


Figure 1.4. A: Cysteine pattern retrieved by alignement of 43 sequences belong to SCR family and considering only those conserved in more than 90% of accessions. B: amino acids pooled together to decrease the stringency of the final algorithm.

Conserved amino acids in more than 90% of the sequence (≥39 sequences) were used as guideline pattern. The conserved pattern consisted on 7 cysteine (**Figure 1.4A**) and other

amino acids that are grouped to reduce the stringency of the algorithm, such as GAVT (small amino acids) and LMFWYI (hydrophobic and some aromatic amino acids) as reported in **Figure 1.4B**. The final pattern was constituted by the sequence expression:

C-x(3-10)-[GAVT]-x(1,2)-C-x(5,19)-C-x(3)-[LMFWYI]-x(3,15)-C-x-C-x(3,16)-C-x-C-x(1,6)-C

where C is a cysteine residue, x is a random amino acid, (N,N) is the minimum and maximum number of random amino acids, [?] indicates the presence of one of the amino acids listed within the brackets. This approach allowed to recovered 28 accessions that were then used to build a tree using the PHYLM algorithm and using the SCR of that the tridimensional structure is known (Mishima *et al.*, 2003) as reference. The tree was then used to choose 10 sequences as representative samples (**Figure 1.5**).

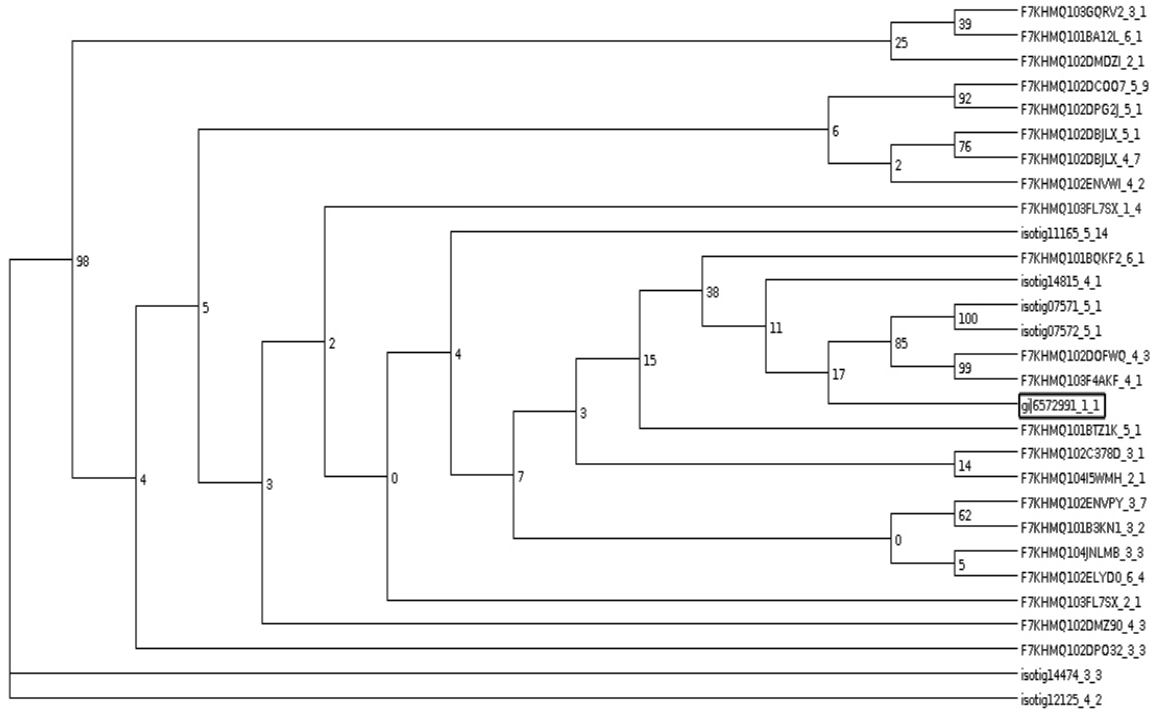


Figure 1.5. Tree built by the algorithm PHYLM with 100 bootstrap replica performed on 28 accessions retrieved by the algorithm built *ad hoc*. The accession in the square indicates the accession number of one SCR from *Brassica rapa* of that the tridimensional structure is known.

Accessions retrieved by the screening of 454 pyrosequencing libraries of cDNAs from olive flowers, 23 in total, were then analyzed by means of quantitative Real-Time PCR in pistils and anthers of both cultivars Leccino and Frantoio.

Expression analyses

Expression analyses of *SLG*, *SRK* and *SCR*-like genes were performed using mRNA isolated from pistils and anthers by means of quantitative Real-Time PCR using different subdomain-specific primer combinations. Pistils were collected at four developmental stages: T0 (opening flower), and T2, T3 and T4 (2, 3 and 4 days after flower opening, respectively). Closed (A1), dehiscent (A2) and senescent (A3) anthers were collected throughout bloom. All samples were amplified in the same conditions and normalized by ΔΔCt method to the *Olea europaea* Elongation Factor α1 gene.

Quantitative Real-Time experiments were performed using a StepOne 96-wells machine (Applied Biosystems) and SYBR Green Master mix (Applied Biosystems). From two to three biological repetitions were performed according to the experiment and three technical repetitions were performed each time.

***In situ* hybridization**

Flower samples were collected at five different developmental stage, named S1 through S5, for both cultivars Leccino and Frantoio. Stage 1 corresponds to 10 days before anthesis, stage 2 to 5 days before anthesis, stage 3 to flowering day, stage 4 to 3 days after flowering, stage 5 to 15 days after flowering. Samples were fixated by using freshly ice cold FAA (50 ethanol : 35 RNase free water : 10 37% formaldehyde : 5 glacial acetic acid) and vacuum infiltrated four times for 15 minutes each. Samples were then transferred in a rocking platform at 4°C for 12 hours and rinsed with 50% ethanol. Tissues were then dehydrated through washes in ethanol at 50%, 85%, 90%, 100%, for 2 hour each, kept overnight in

new fresh 100% ethanol and then cleared in citrisolve (50 ethanol : 50 citrisolve followed by 100% citrisolve; 2 hours each step at room temperature). Tissues were then covered with enough citrisolve and the beaker was filled with Paraplat Plus chips and incubated overnight at 60°C. Over the subsequent two days the solution in the beaker was changed with fresh molten Paraplast Plus 3 times a day. Tissues were then embedded in Paraplast Plus and stored at 4°C until use. Fragments of *OeSRK*, *OeSRK-like*, *OeSLG* and *OeSRC-like_12* genes were cloned from olive cDNAs by using different subdomain primers as reported in **Table 1.2**.

Table 1.2. Primer sequences used to design probes against SI-related genes. Amplicons size and size of the two sets of probes, unhydrolyzed and hydrolyzed, were also reported.

Gene	Primer name	Primer sequence	Amplicon size	Unhydrolyzed probe	Hydrolyzed probe
OeSRK	OeSRK_for	GTGGTTGATGGTCACTTCTCAG	406 bp	406 bp	200 bp
	OeSRK_rev	CATTGGCAGAACTTACCGTAG			
OeSRK-like	OeSRK-like_for	GCAATGGATGGCCTATTCTCTA	406 bp	406 bp	200 bp
	OeSRK-like_rev	TCACAGTGCAGGATTCATCCT			
OeSLG	OeSLG_for	CCATTAGCACAACTCAGAGCCTCA	422 bp	422 bp	200 bp
	OeSLG_rev	TGCAGCATTCCAACCGAAGTTC			
OeSCR-like_12	OeSCR-like_for	AAGAAATGGCAGGTCGTCCGTA	280 bp	280 bp	200 bp
	OeSCR-like_rev	GGTCCGCCTTCCTTAGTCTTCC			

Amplicons were cloned using TOPO-TA Cloning Kit and TOP10 *E. coli* competent cells according to the manufacturer's protocol and plated in 50 mg/L kanamycin selective LB plates. Single colonies were tested by PCR to test the effective presence of the insert and grown overnight in liquid 50mg/L kanamycin LB medium. Plasmids were recovered by midi-prep kit (QIAGEN) according to the manufacturer's procedures, the quality was tested on agarose gel and quantified using nano-drop spectrophotometer. Vectors containing our fragments were linearized using two different digestion enzymes, *Spe*I and *Not*I (New England Biolabs), and two sets of digoxigenin-labeled probes were transcribed in both anti-

sense and sense orientation using T7 and T3 RNA polymerase (ROCHE) and then treated by DNase to eliminate the plasmid residues. RNA probes were then run in agarose gel to test the efficiency of RNA polymerases and the absence of plasmid DNA. One set of probes was alkaline digested in order to generate smaller probes of 200 bp while the other set full-length ranged from 300 to 500 bp (**Table 1.2**). Probes were then precipitated, quantified by digoxigenin and stored at -20°C until use. Embedded samples were sectioned at 8 µm using a microtome and then transferred in positive charged slides (Probe On Plus, Fisherbrand) at 40°C for 4 hours. RNA *in situ* hybridization was carried out according to the protocol in (Kramer, 2005). Slides were transferred in a rack and washes twice in citrisolve for 10 minutes each. Samples were then rehydrated by washes in serial ethanol dilutions (100%, 95%, 85%, 70%, 50%, and 30%) and left for 2 minutes in 1X PBS buffer. Specimens were then treated by Pronase enzyme for 20 minutes at 37°C and washed in 1X Glycine/PBS buffer. Samples were then washed in 4% Paraformaldehyde buffer for 20 minutes, washed in 1X PBS buffer, treated by Acetic Anhydride for 10 minutes and washed again in 1X PBS buffer. Samples were then dehydrated by serial ethanol dilutions (30%, 50%, 70%, 85%, 95% and 100%). After that, slides were left to dry vertically for about 10 minutes and then probes, along with HYB solution, were applied after a preheating at 80°C and left overnight in hybridization oven at 38°C-44°C according to different probes. The day after probes and HYB solution were removed and samples washed twice in 0,2X SSC buffer at 55°C for 1 hour each, washed twice in 1X NTE buffer at 37°C for 5 min each and then treated by RNase A for 20 minutes at 37°C. After that slides were washed again twice in 1X NTE buffer for 5 minutes each, washed in 0.2X SSC at 55°C for 1 hour and then washed in 1X PBS at room temperature for 5 minutes. Slides were then moved from rack to a plastic container and treated with 1% Blocking solution (Roche) for 45 minutes and then Anti-Dig antibody was applied to the slides and left 2 hours in a humid chamber. After that, samples were washed four times in 1% BSA buffer for 15 minutes each and then washed in G3

buffer for 10 minutes. At the end NBT/BCIP color substrate (Roche) was added to the slides, which were kept for 1-3 days in a humid chamber and placed in dark, until the signal was clear. Results were visualized on a Zeiss AxioImager Z2 microscope.

Identification of conserved miRNAs in olive and their target prediction

All plant miRNA sequences, a total of 3,938 mature sequences belong to 51 species, were downloaded from the miRNA database (release 17, April 2011: www.mirbase.org) and they were used as input file to screen four 454 pyrosequencing libraries of cDNAs from olive flowers consist of 464,969 accessions. The screening was performed using the perl script MIRcheck (Jones-Rhoades *et al.*, 2004). MIRcheck is a perl script designed to identify RNA sequences with secondary structures similar to plant miRNAs. Three subsequent steps were followed. First, the 3938 plant miRNAs were used as query to screen the libraries looking for similar sequences. The maximum mismatch admitted was 4 nucleotides but neither insertions nor deletions were allowed. Second, the putative miRNAs retrieved were used as known region and 200 bp downstream and 200 bp upstream were recovered from the original sequence. Third, segments of about 300-400 bp were then used to predict the pre-miRNAs. All putative miRNAs and their pre-miRNAs retrieved by MIRcheck scripts were then tested one by one using one on-line software which predicts the secondary structure and provides an image of the folded stem-loop structures (http://mfold.rna.albany.edu) (Zuker, 2003). The pre-miRNAs resulted from this last analysis were accepted or declined according to Meyers *et al.* (2008): i) the miRNA and miRNA* are derived from the opposite stem-arms such that they form a duplex with two nucleotide, 3' overhangs; ii) base-pairing between the miRNA and the other arm of the hairpin, which includes the miRNA*, is extensive such that there are typically four or fewer mismatches miRNA bases; iii) asymmetric bulges are minimal in size (one or two bases) and frequency (typically one or less), especially within the miRNA/miRNA* duplex. A

further selection was carried out by accepting only those sequences showing a ΔG (free energy) more than 14 kj.

A further systematic approach was attempted to shed light the conservation of miRNA families through different species. All the miRNA families present in two or more families were pooled together and the conservation of miRNAs through families was provided.

Targets for olive miRNAs were predicted using an on-line prediction web-site (http://plantgrn.noble.org/psRNATarget) (Dai *et al.*, 2011), using olive miRNAs as query and *Arabidopsis* genome as target database. The maximum expected mismatches of pairing between miRNAs and targets was set at 3 and no insertions and/or deletions was admitted.

Homology between olive miRNAs and our SI genes, *OeSRK*, *OeSLG* and *SCR-like*, was also investigated using MIRcheck scripts previously described, using two different screenings of both cDNA and genomic sequences. The first screening was performed setting the perl script as 4 0 0 (mismatches, insertion and deletion, respectively) between miRNA and target gene, the second was instead set as 2 2 2.

Results

Aniline blue staining

In order to mark callose plaques of pollen tubes growing through the transmitting tissues of the style, aniline blue staining was performed as first comparative analysis of pistils from cv. Leccino and cv. Frantoio 48 hours after self-pollination. This analysis revealed marked differences for the absence and presence of pollen tubes within the transmitting tissue of the style and the ovary (**Figure 1.6**). In the self-incompatible cv. Leccino no growing pollen tubes were visualized in any of the pistils investigated. Germinating pollen grains were visualized on the stigma of cultivar Leccino, but no one neither penetrated the stigma surface nor grew into the style. Vice versa, the self-compatible cv. Frantoio showed several pollen grains germinating on the stigma and pollen tubes growing into the style.

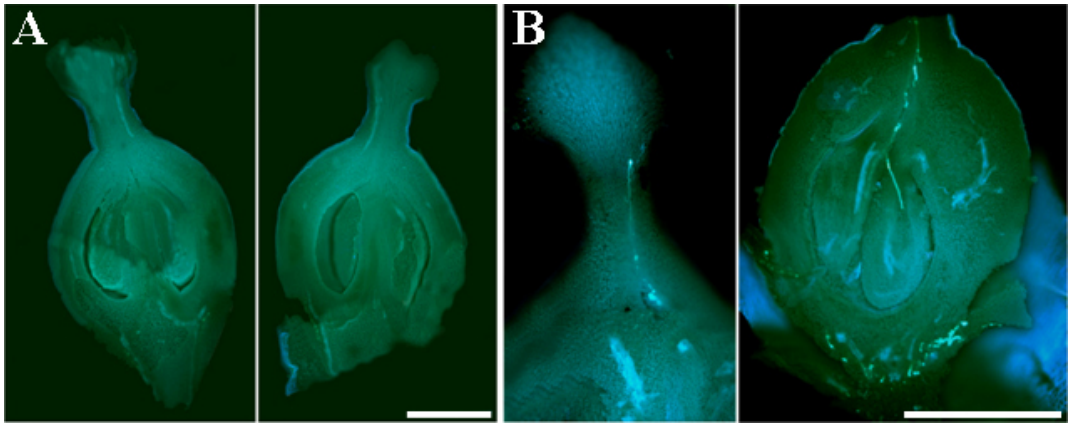


Figure 1.6. Aniline blue staining in pistils of cv. Leccino (SI) and cv. Frantoio (SC) 48 hours after self-pollination. The cultivar Leccino (A) did not reveal any pollen tube through the style, whereas a well-marked pollen tube is visualized under the stigma as well as through the style in the cultivar Frantoio (B). Bars= 0,5 mm.

With the aim to confirm the first screening and to get a more robust set of data, aniline blue staining assays were performed again in a large number of flowers collected from four different cultivars: three self-incompatible (Leccino, Moraiolo, Dolce Agogia) cultivars and one self-compatible (Frantoio) cultivar adopted as positive control. All analyses were performed after both self- and cross-pollination. After self-pollination, no pollen tubes were

visualized through the style in incompatible cultivars (Leccino, Moraiolo and Dolce Agogia), whereas in the style of the compatible cultivar Frantoio several pollen tubes were marked (**Figure 1.7**, panel A). By contrast, pollen tubes were discovered in all cultivars after cross-pollination, with no difference between compatible and incompatible cultivars (**Figure 1.7**, panel B). This last set of findings fully validated the results obtained with the first aniline blue staining performed in the preliminary screening based on two cultivars only after self-pollination.

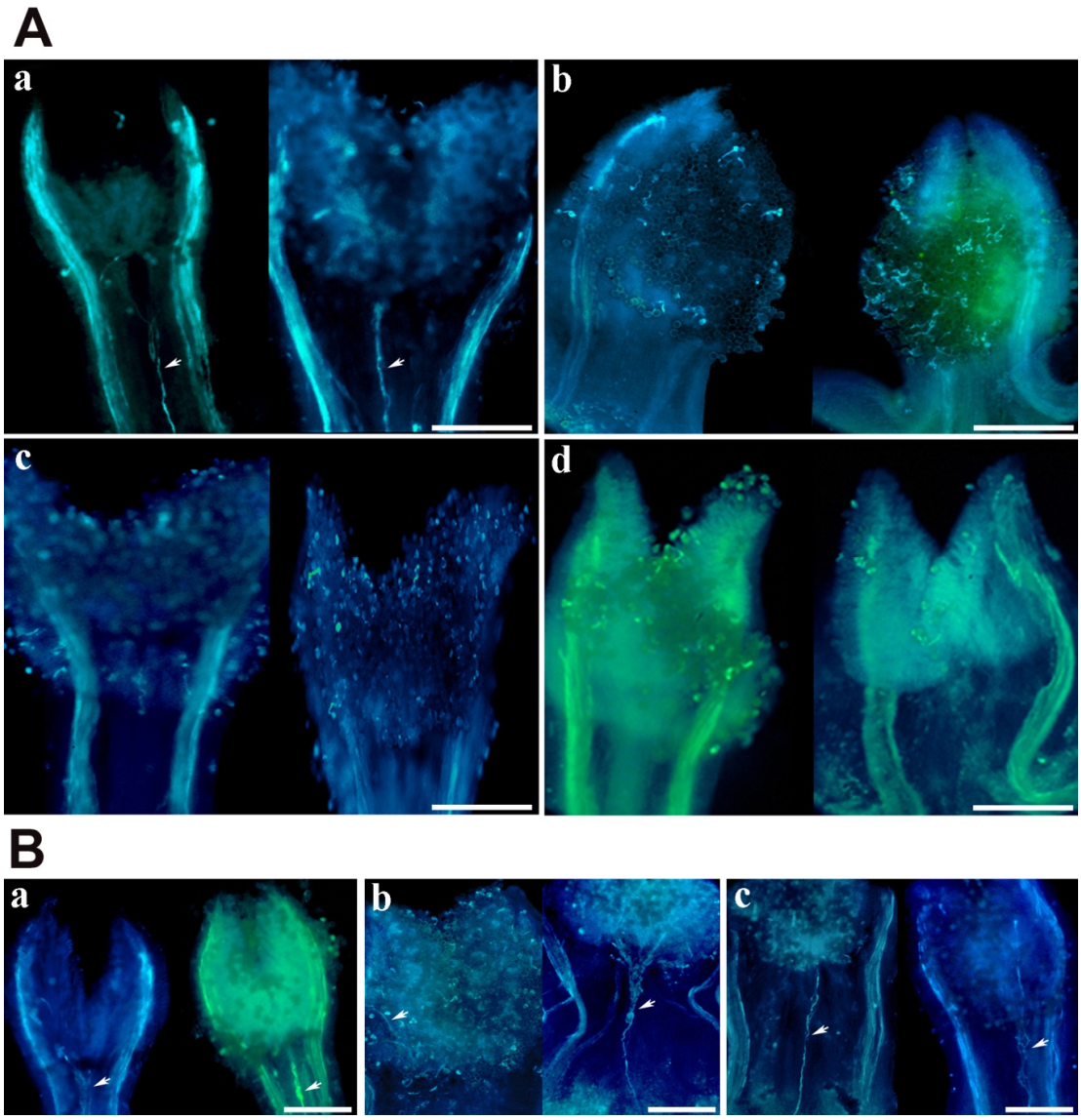


Figure 1.7. Cytological analysis of olive pollinated pistils. Panel A: Aniline blue staining of self-pollinated pistils in the self-compatible cultivar Frantoio (a) and in the self-incompatible cultivars Leccino (b), Moraiolo (c), Dolce Agogia (d). Panel B: Aniline blue staining of open-pollinated pistils in the self-incompatible cultivars Leccino (a), Moraiolo (b) and Dolce Agogia (c). White arrows indicate stained pollen tubes. Bars= 100 µm.

Cytological observations by means of aniline blue staining were also performed on transversal sections of pollinated pistils of both cultivars Leccino and Frantoio after both self- and cross-pollination (**Figure 1.8**). On the whole, final results were in agreement with preliminary results, in fact no pollen tubes were visualized in cv. Leccino after self-pollination, whereas several pollen tubes were visualized at different levels of the style in cv. Frantoio after self-pollination, even though usually only one was able to reach the ovary. It is worth mentioning that in olive the ovary is characterized by four ovules, but normally only one is fertilized and a single olive fruit is produced for each flower.

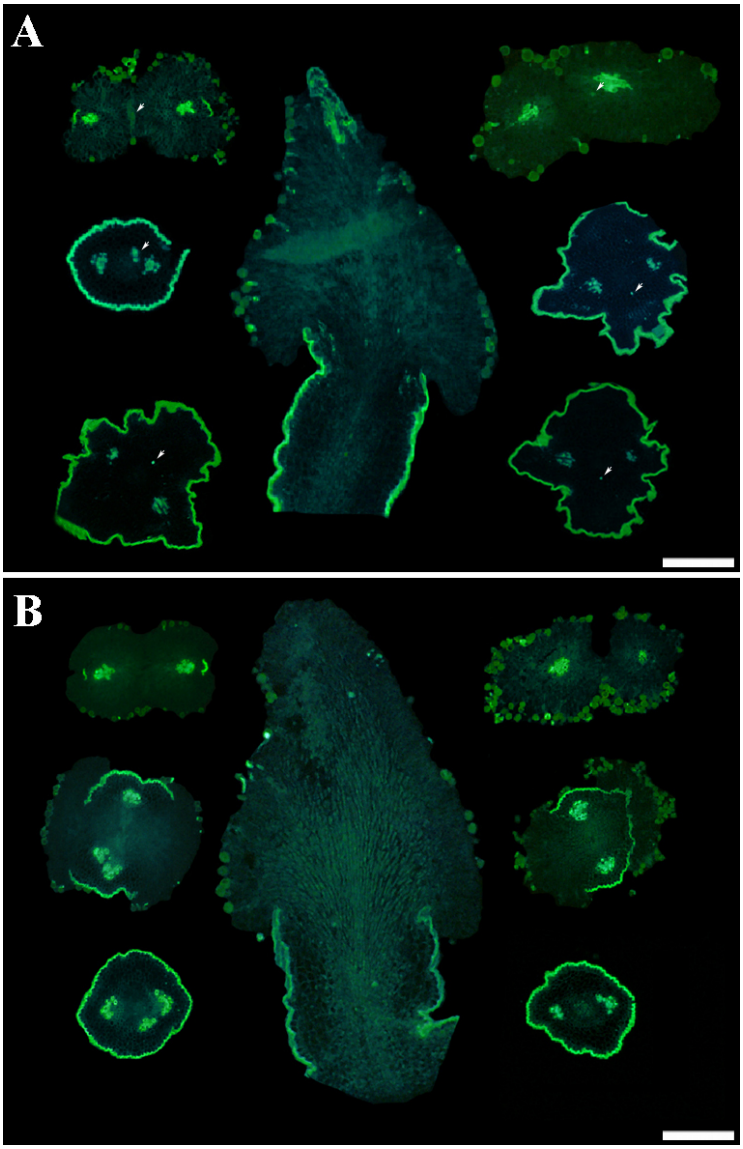


Figure 1.8. Aniline blue staining of transversal sections of self-pollinated pistils. A: self-compatible cultivar Frantoio. B: self-incompatible cultivar Leccino. White arrows indicate stained pollen tubes. Bars= 100 µm.

A direct correlation between aniline blue stained sections of the medial part of style of both cultivars Leccino and Frantoio after self-pollination highlighted the absence and the presence of pollen tubes in the self-incompatible cv. Leccino and in the self-compatible cv. Frantoio, respectively (**Figure 1.9**).

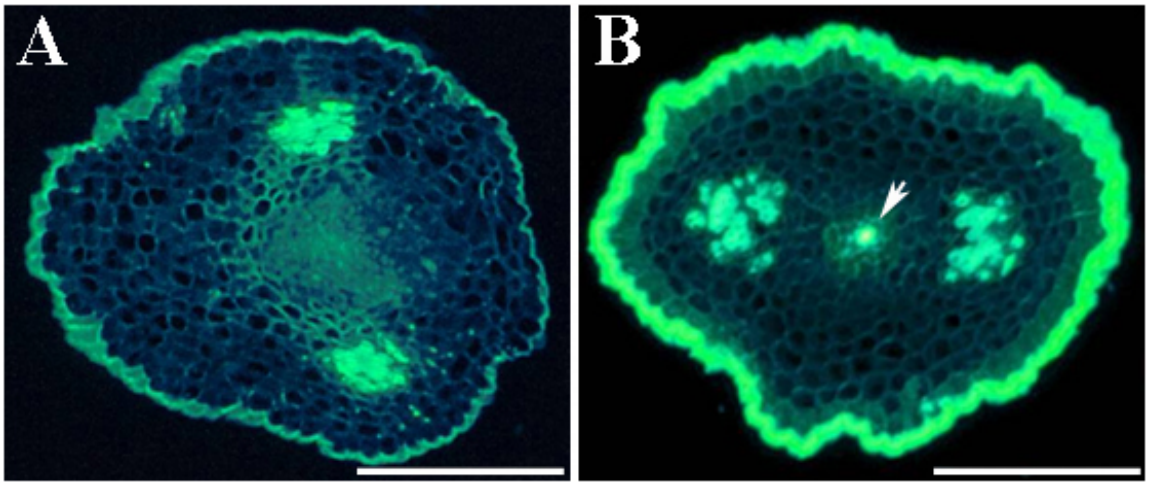


Figure 1.9. Aniline blue staining of transversal sections of self-pollinated pistils at higher magnification. A: self-incompatible cultivar Leccino. B: self-compatible cultivar Frantoio. White arrow indicates stained pollen tubes. Bars= 100 µm.

A further analysis of stigmas of the cv. Leccino after self-pollination showed that several pollen grains were not germinated in the papillae and that few pollen tubes were not penetrated through the papillae cells (**Figure 1.10**).

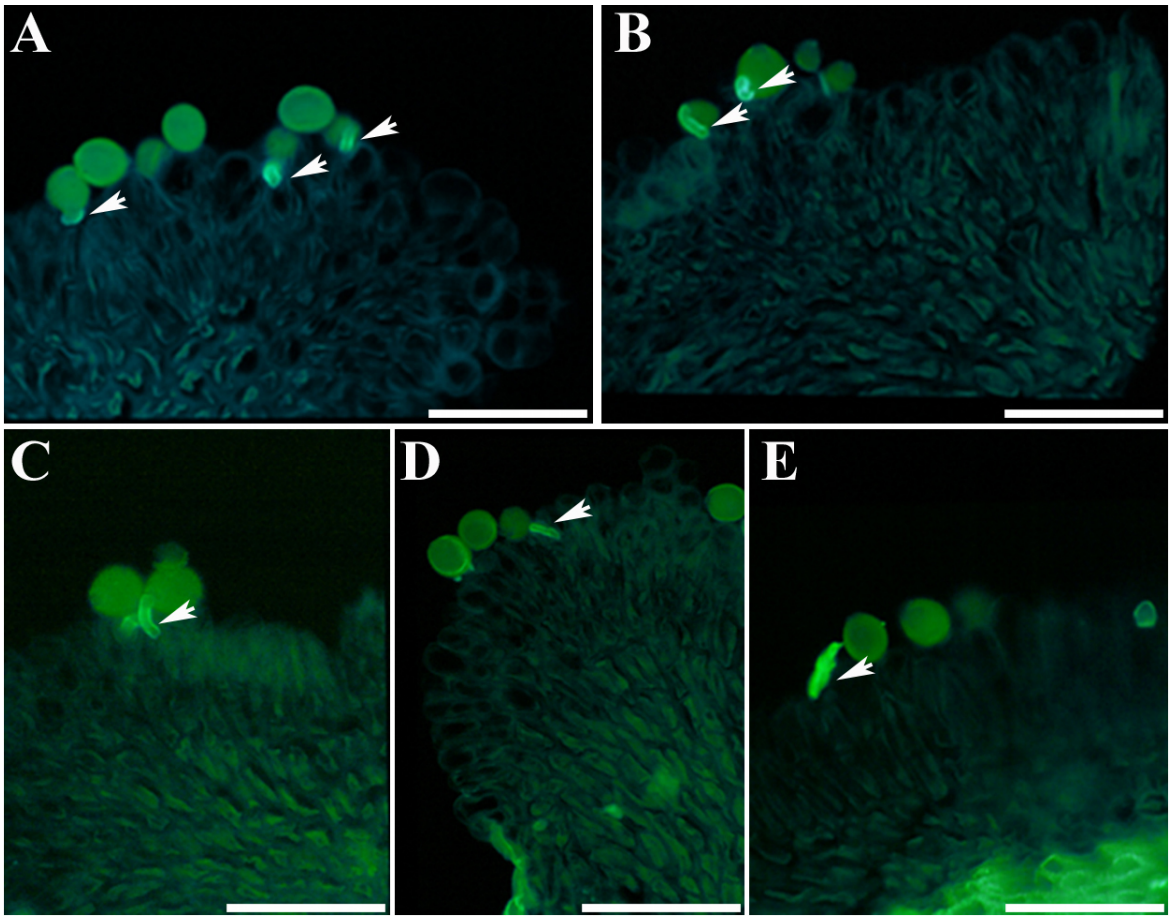


Figure 1.10. Aniline blue staining of self-pollinated pistils in the self-incompatible cultivar Leccino. A-E: few pollen grains were germinated on papillae cells and emerging pollen tubes did not penetrate the stigma surface. White arrows indicate emerging pollen tubes. Bars= 100 µm.

By contrast, under open-pollination conditions, the cv. Leccino showed several pollen grains germinated in the stigma and pollen tubes growing from the stigma to the basal part of the ovary (**Figure 1.11**).

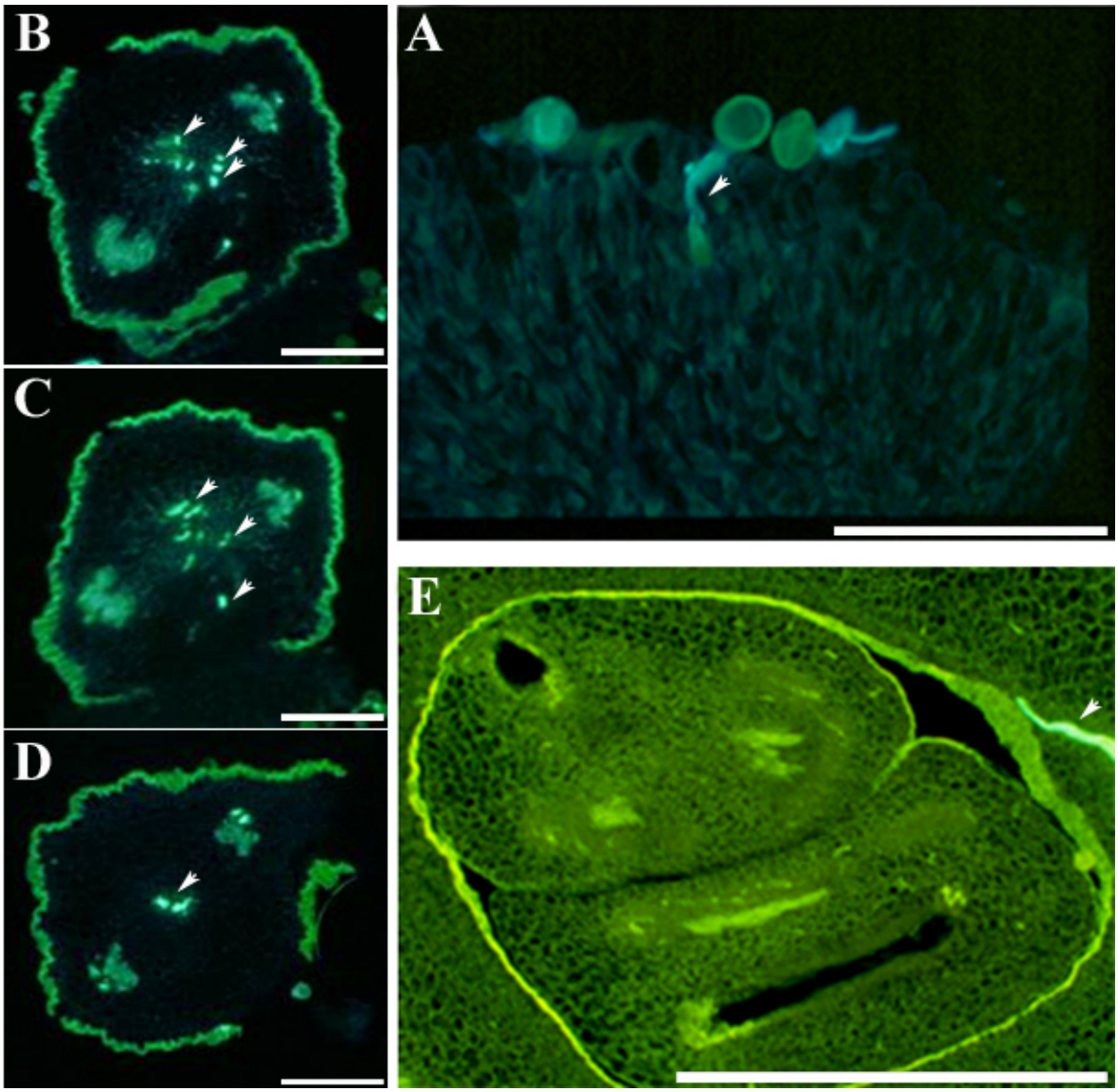


Figure 1.11. Aniline blue staining of open-pollinated pistils in the self-incompatible cultivar Leccino. A: Pollen tubes growing through papillae cells. B-D: several pollen tubes were found throughout the pistil. E: one pollen tube reached the ovules. White arrows indicate stained pollen tubes. Bars= 100 µm.

To get a robust statistical support, further analyses of pollen grains on pollinated pistils were performed and the percentage of germination was computed for each cultivar. Almost 34,000 pollen grains of the same four cultivars (Leccino, Moraiolo, Dolce Agogia and Frantoio) were analyzed and divided into six different classes: a) pollen grains germinated on the stigma surface; b) pollen grains germinated on papillae; c) pollen tubes reaching the proximal part of the style; d) pollen tubes reaching the middle part of the style; e) pollen

tubes reaching the distal part of the style; f) not germinated pollen grains. Replicated experiments were performed after both self- and cross-pollination (**Figure 1.12**).

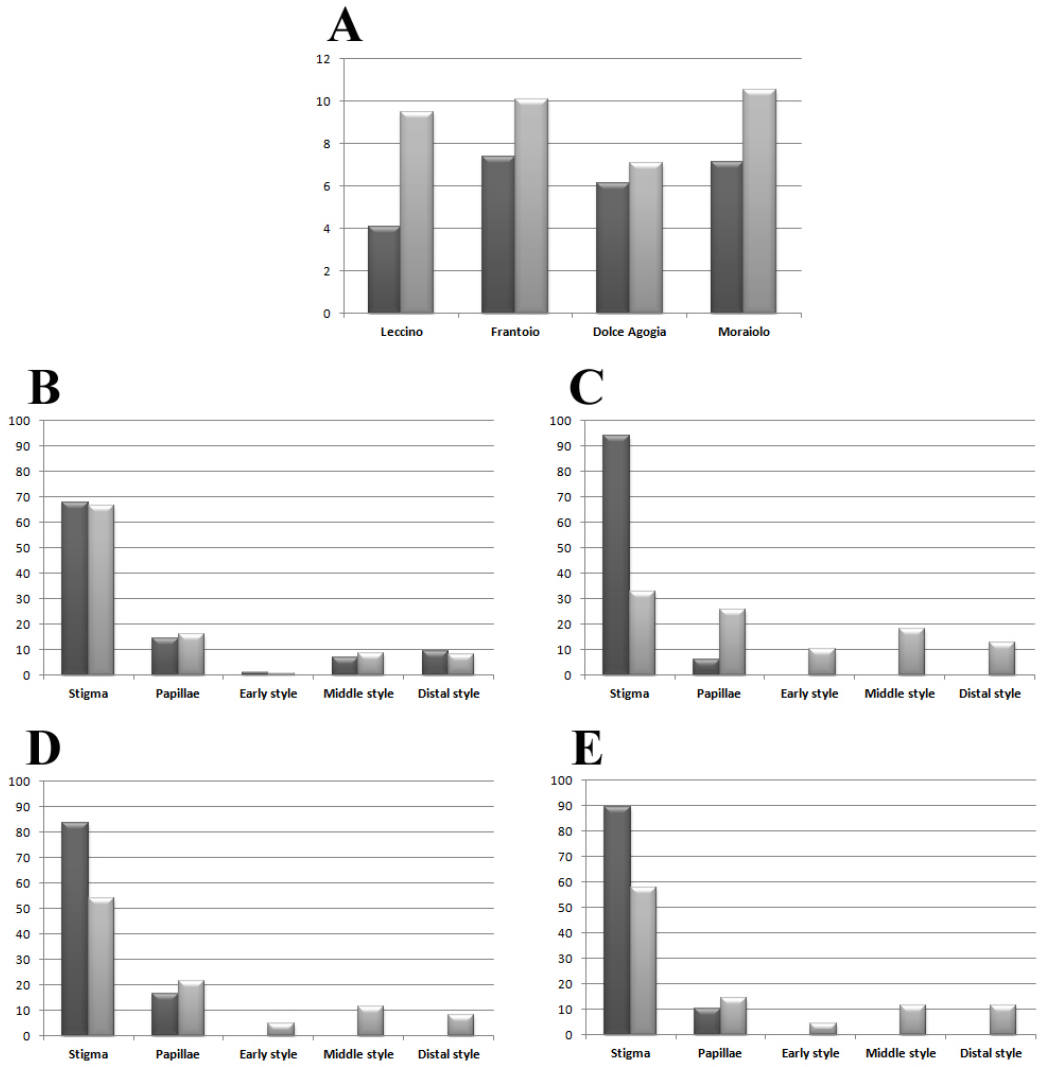


Figure 1.12. Results on pollen grain germination and pollen tube growth. A: percentage of germinated pollen grains of the olive cultivars Leccino, Frantoio, Dolce Agogia and Moraiolo. B-E: Percentage of pollen tubes found in different tissues of the pistil in cv. Frantoio (A), cv. Leccino (B), cv. Dolce Agogia (C) and cv. Moraiolo (D). Dark grey indicates the percentage of germinated pollen grains after self-pollination and light grey after open-pollination.

On the whole, the percentage of pollen germination was quite low, ranging between 4% and 6% after self-pollination, and 6% and 10% after cross-pollination (**Table 1.3**). Cross-pollination positively correlated with percentage of pollen germination, being almost twice than that calculated for self-pollination. Among germinated pollen grains after self-pollination, only the self-compatible cultivar Frantoio showed pollen tubes growing

through the transmitting tissues of the style (classes d, e and f), whereas in the other self-incompatible cultivars, Leccino, Moraiolo and Dolce Agogia, no pollen tubes were visualized over the stigma surface.

Table 1.3. Percentage of pollen germination after both self-pollination and open-pollination tested in four olive cultivars.

	Percentage of germination (%)	
Cultivar	**self-pollination**	**open-pollination**
Frantoio	7.38	10.10
Leccino	4.07	9.47
Dolce Agogia	6.12	7.09
Moraiolo	7.14	10.53

By contrast, under cross-pollination conditions, the self-incompatible cultivars Leccino, Moraiolo and Dolce Agogia showed pollen tubes growing through their styles as well as the self-compatible cultivar Frantoio. Moreover, no significant differences were observed between self-compatible and self-incompatible cultivars (**Figure 1.12**). Values of germinated pollen divided into the 5 classes are reported in **Table 1.4**.

Table 1.4. Percentage of germinated pollen divided into the 5 classes after self-pollination and open-pollination in four olive cultivars. Self-poll.: self pollination. Open-poll.: open pollination.

	Stigma		**Papillae**		**Early style**		**Middle style**		**Distal style**	
	Self poll.	**Open poll.**	**Self poll.**	**Open poll.**	**Self poll.**	**Open poll.**	**Self poll.**	**Open poll.**	**Self poll.**	**Open poll.**
Frantoio	68.00	66.67	14.50	15.94	1.00	0.72	7.00	8.70	9.50	7.97
Leccino	94.00	32.94	6.00	25.88	0.00	10.20	0.00	18.04	0.00	12.94
D. Agogia	83.61	54.10	16.39	21.31	0.00	4.92	0.00	11.48	0.00	8.20
Moraiolo	89.62	57.82	10.38	14.63	0.00	4.42	0.00	11.56	0.00	11.56

Iodine-potassium-iodide staining

A staining of pollinated pistils using an iodine-potassium-iodide (IKI) solution, which is able to mark starch grains within the cells, was performed in transversal sections of the style of both cultivars Leccino and Frantoio after self-pollination (**Figure 1.13**).

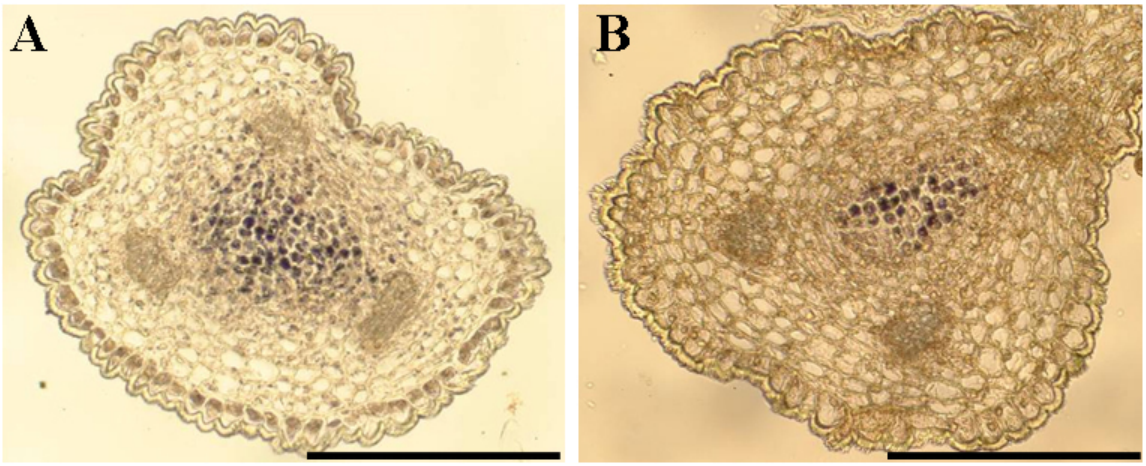


Figure 1.13. Iodine-Potassium-iodide (IKI) staining of transversal sections of self pollinated pistils. A: self-incompatible cultivar Leccino. B: self-compatible cultivar Frantoio. Bars= 100 µm.

Results showed that the self-compatible cultivar Frantoio was characterized by low amount of starch in transmitting tissues, whereas in the self-incompatible cultivar Leccino the amount of starch was higher. This might be due to the growth of self-pollen tubes in the cv. Frantoio but not in the cv. Leccino, because the heterotrophic growth of self-pollen tubes into the styles of the cv. Frantoio consumed the starch available in the cells.

SEM analyses

Further evidences of the inability of pollen tubes emerging from the low amount of germinated pollen grains in the cultivar Leccino after self-pollination were gathered by means of SEM analysis. Results showed that few pollen grains were germinated and that the few emerging pollen tubes were not able to penetrate into the stigma surface (**Figure 1.14**).

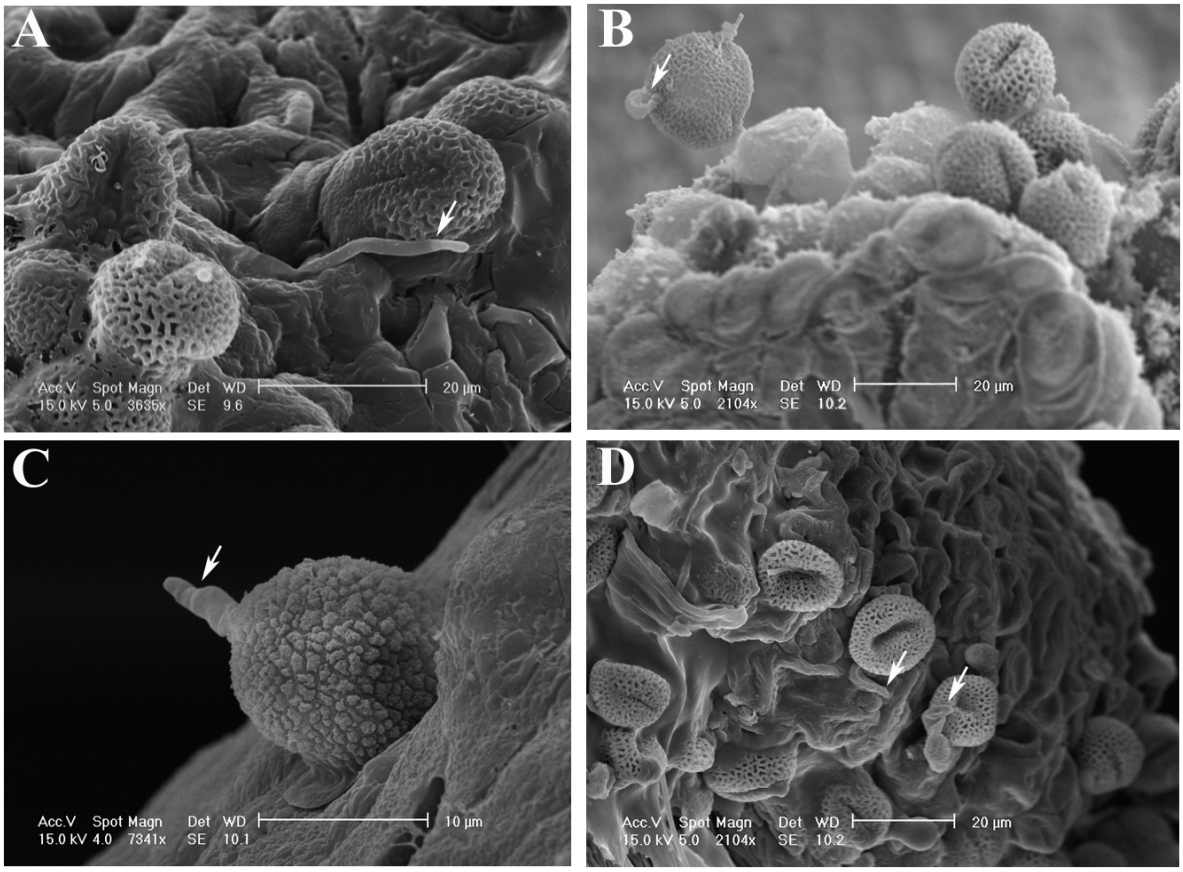


Figure 1.14. SEM images of self-pollinated stigmas in the self-incompatible cultivar Leccino. A-D: several emerging pollen tubes (white arrows) were found growing in the opposite site to the stigma surface.

On the whole, data deriving from cytological analyses were fully in agreement with a SSI system, because during S-RNase-based GSI the incompatibility response is triggered by a stylar reaction. Here we demonstrated that no self-pollen tubes are detectable in the styles of self-incompatible cultivars after self-pollination.

OeSLG and *OeSRK* gene expression analysis

Expression analysis of the olive *OeSLG* and *OeSRK* genes was carried out by quantitative Real-Time PCR experiments in both Leccino and Frantoio cultivars. The two predicted genes for *OeSLG* were too much similar at the nucleotide sequence level and despite several attempts no primers discriminating each other could be designed. We hypothesized that expression data here presented are the addition of the expression of both *OeSLG*-like

genes. Moreover, the two *OeSRK*-like genes isolated showed a low similarity each other and no problems were experienced to test separately their expression levels and patterns. They were called *OeSRK* and *OeSRK*-like.

Experimental data were analyzed by ΔΔCt method using the Elongation Factor α1 as housekeeping gene. Normalized values were compared to mark different expression levels among different tissues and through different developmental stages.

A first set of expression patterns was performed in the self-incompatible cultivar Leccino and in the self-compatible cultivar Frantoio. By means of qRT-PCR, the three genes *OeSLG*, *OeSRK-like* and *OeSRK* were tested among different tissues through different developmental stages, such as open and closed flowers, fruits setting and veraison, and roots (**Figure 1.15**). *OeSLG* in cv. Leccino was preferentially expressed at the stage of closed flowers, whereas in cv. Frantoio it was slightly more expressed at the stage of fruits veraison. *OeSRK-like* in cv. Leccino was preferentially expressed at the stage of closed flowers and in roots, whereas in cv. Frantoio it was more expressed at the stage of open flowers and in roots. *OeSRK* in cv. Leccino was preferentially expressed at the stage of fruits veraison and closed flowers, whereas in cv. Frantoio it was more expressed at the stage of fruits veraison.

Furthermore, for closed and open flowers, three replicated analyses were carried out in both cultivars in order to get a more robust set of data (**Figure 1.16**). All three genes showed similar patterns within the same cultivar but opposite patterns between different cultivars. In fact, *OeSLG*, *OeSRK-like* and *OeSRK* expression levels were decreasing from closed flowers to open flowers in cv. Leccino, whereas in cv. Frantoio the expression levels were increasing from closed flowers to open flowers.

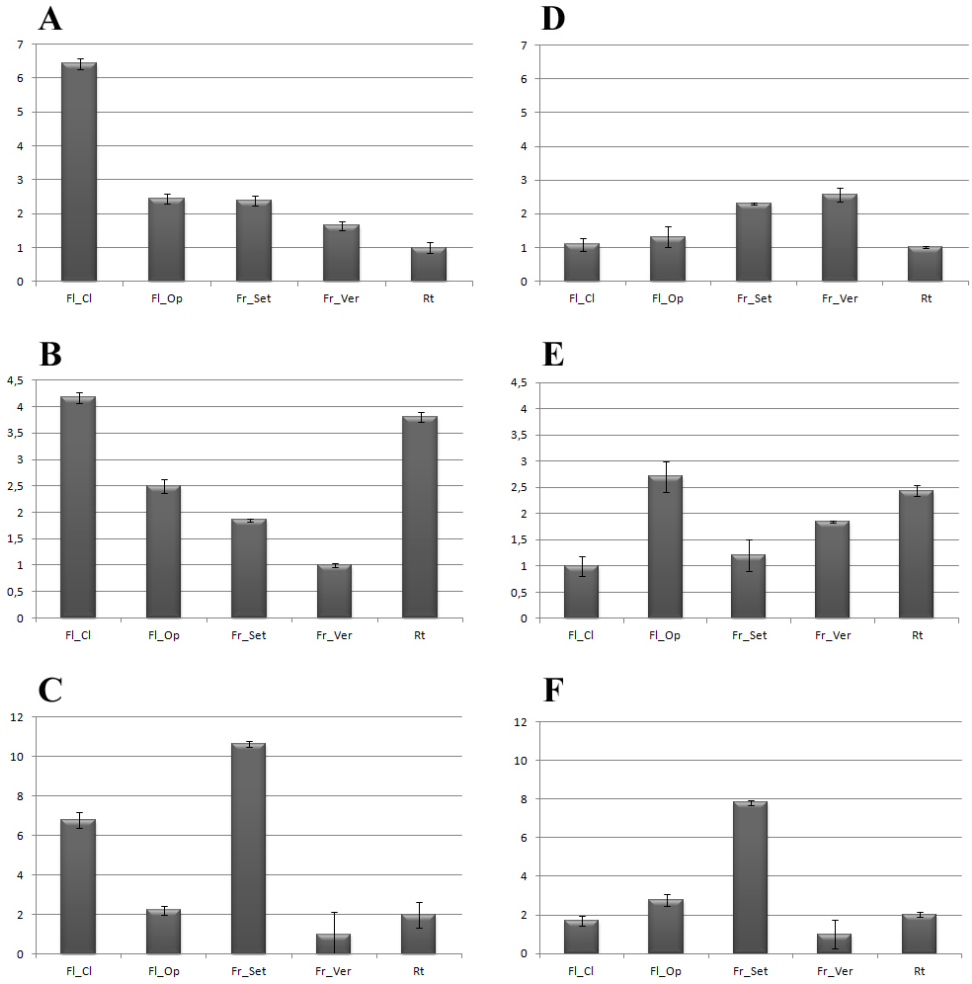


Figure 1.15. Gene expression levels tested by qRT-PCR in different tissues and through different developmental stages. Fl_Cl: closed flowers; Fl_Op: open flowers; Fr_Set: fruits setting; Fr_Ver: fruits veraison; Rt: roots. A-C: cv. Leccino. D-F: cv. Frantoio. A, D: *OeSLG*. B, E: *OeSRK-like*. C, F: *OeSRK*. Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

Considering that Leccino is a self-incompatible cultivar and Frantoio is a self-compatible cultivar, results were in agreement with those expected by putative genes involved in self-incompatibility. In fact, just before the flowering the whole set of transcripts, codifying for the proteins involved in self-incompatibility mechanism, should be present in the flower tissue, whereas in the late flowering the amount of transcripts should decrease, since the self-incompatibility reaction has already taken place.

This first set of expression levels and patterns of *OeSLG*, *OeSRK-like* and *OeSRK* genes was used as guideline to perform new experiments based on an investigation of gene expression in specific tissues, such as pistils and anthers, in a separate manner.

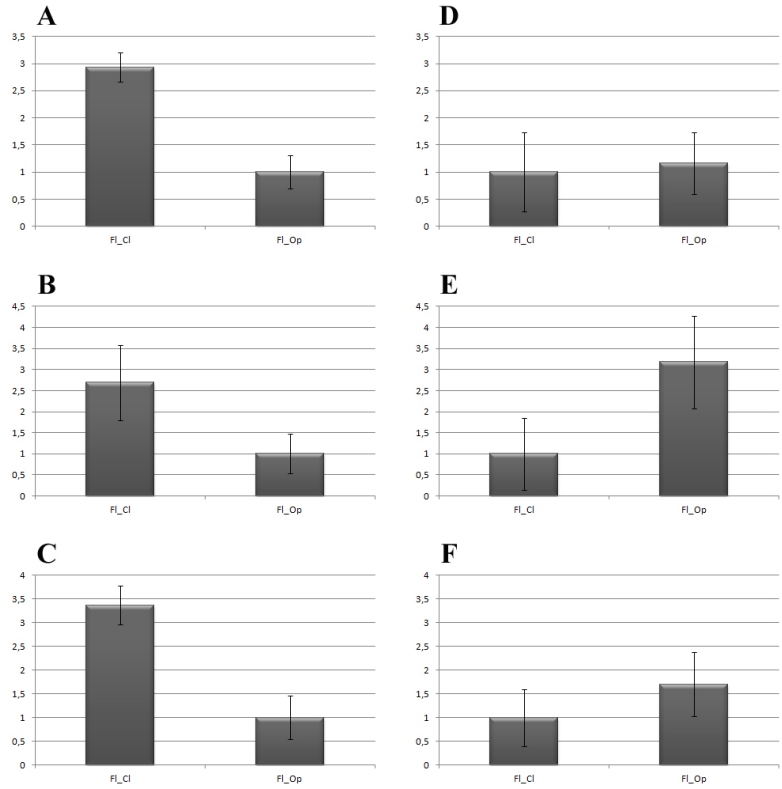


Figure 1.16. Results of qRT-PCR replicated experiments in flowers. Fl_Cl: closed flowers; Fl_Op: open flowers. A-C: cv. Leccino. D-F: cv. Frantoio. A, D: *OeSLG*. B, E: *OeSRK-like*. C, F: *OeSRK*. Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

Furthermore, since the germination of pollen grains in olive was found to occur 48-72 hours after landed on the stigma, in order to finely mark difference between the two cultivars, the expression analyses were replicated in pistils and anthers collected throughout several days after flowering in the case of pistils and throughout different developmental stages in the case of anthers. The pistils were collected at four developmental stages: T0 (flowering day), T2, T3 and T4 (2, 3 and 4 days after flowering, respectively). The anthers were collected at three different developmental stages: 1) closed flower with yellow anthers; 2) open flower with yellow anthers; 3) flower with dehisced anthers, named A1 through A3, respectively. Interestingly, qRT-PCR expression analyses showed as *OeSRK* in pistils of cv. Frantoio was poorly expressed at T0 and T2, and no expression was visualized at T3 and T4 (**Figure 1.17**). Similarly, no expression was found in anthers (**Figure 1.18**). By contrast, in cv. Leccino *OeSRK* expression showed an increase from T0 through T4 in

pistils (**Figure 1.17**) and a decrease from A1 through A3 in anthers (**Figure 1.18**). The *OeSLG* gene also showed an opposite trend between cv. Leccino and cv. Frantoio. In pistils of cv. Frantoio was more expressed at later stages (T3 and T4), whereas in cv. Leccino the expression was higher at earlier stages (T0 and T2) (**Figure 1.17**). In anthers it was more expressed in cv. Frantoio at all stages, which was characterized by an increase from A1 to A2 and then a decrease at the A3 (**Figure 1.18**). In cv. Leccino expression of *OeSLG* was slightly decreasing from A1 through A3 (**Figure 1.18**). The *OeSRK*-like expression in cv. Leccino rapidly increased from T0 to T2 and then decreased from T3 to T4, whereas in cv. Frantoio was decreasing from T0 to T2 and then increasing through T4 (**Figure 1.17**). The *OeSRK*-like expression in anthers of both cultivars showed trends very similar to those recorded for *OeSLG* (**Figure 1.18**).

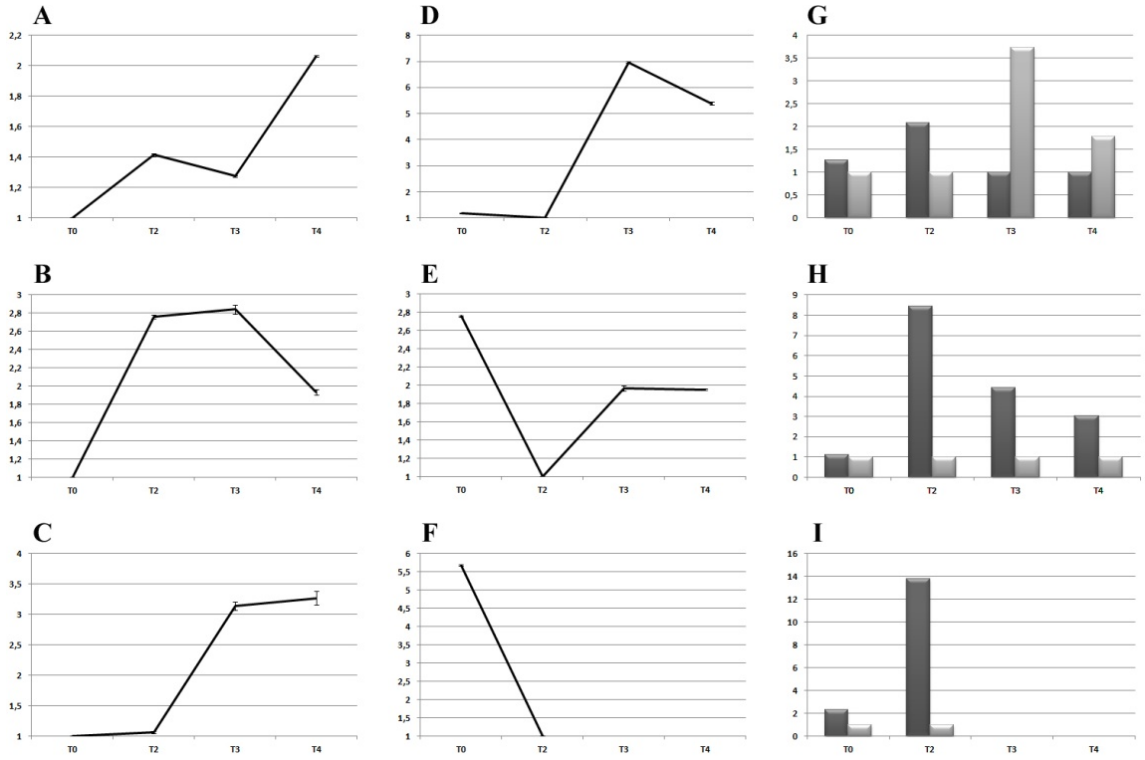


Figure 1.17. Expression patterns of the *OeSLG* (A, D, G), *OeSRK-like* (B, E, H) and *OeSRK* (C, F, I) genes as assessed by qRT-PCR analysis in pistils. A-C: self-incompatible cultivar Leccino. D-F: self-compatible cultivar Frantoio. A-F: expression patterns through four developmental stages of pistils (T0, T2, T3 and T4, meaning flowering day, 2,3 and 4 days after flowering, respectively). G-I: relative expression patterns between cultivars Leccino (dark grey) and Frantoio (light grey). Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

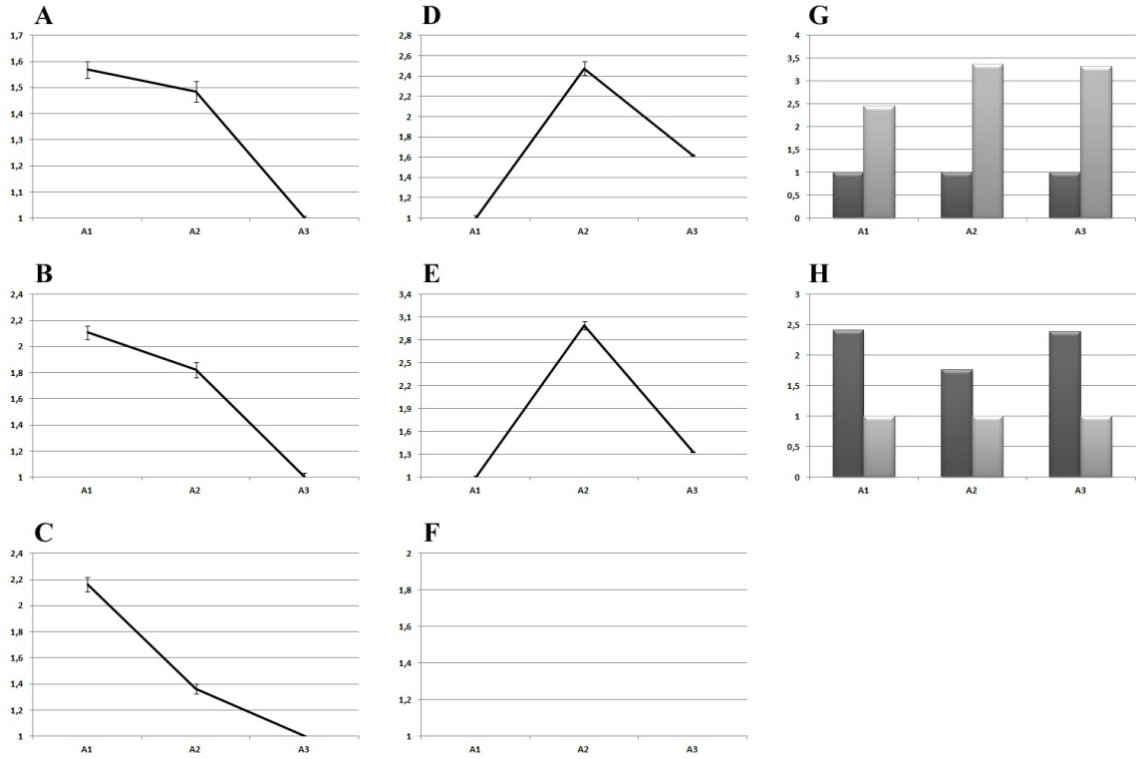


Figure 1.18. Expression patterns of the *OeSLG* (A, D, G), *OeSRK-like* (B, E, H) and *OeSRK* (C, F) genes as assessed by qRT-PCR analysis in anthers. A-C: self-incompatible cultivar Leccino. D-F: self-compatible cultivar Frantoio. A-F: expression pattern through four developmental stages of anthers (A1, A2, A3 meaning closed, dehiscent and senescent anthers, respectively). G-H: relative expression patterns between cultivars Leccino (dark grey) and Frantoio (light grey). Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

Male determinants: search for a good candidate and gene expression analyses

Searching of SCR-relative sequences using multi-alignments of heterologous sequences having the *SCR* gene as reference sequence and subsequent approach based on degenerated primers did not produce results. Distinct 454 pyrosequencing libraries of cDNAs from olive flowers were screened by different approached searching for genes with homologies to *SCR* of *Brassica* as well as for genes known to play a role in pollen-stigma interaction. The first approach aimed at isolating *SCR*-like genes did not reveal results. The screening looking for defensin protein genes, which are characterized by 8 conserved cysteine residues as *SCR* even though their distribution within the protein is slightly different from those of *SCR* (Higashiyama, 2010), allowed to isolate three genes that differed each other for few SNPs. The low nucleotide difference did not allow to analyze separately these genes and they were treated as unique sequence named *OeSCR-like_1*. qRT-PCR analyses aimed at testing the anther-specific expression, which is expected from a good candidate as male determinant, are so presumably the addition of the expression of the genes isolated. Expression analyses showed that in pistils of cv. Leccino the expression increased from T0 to T3 as well as in cv. Frantoio, even though the amount of transcripts was lower than in cv. Leccino. In anthers expression patterns were opposite: cv. Leccino was characterized by a decrease of the expression from stage A1 through stage A3, whereas in cv. Frantoio there was an increase from stage A1 through stage A3 (**Figure 1.19**). However, the expression in both cultivars Leccino and Frantoio was about 5 times higher in pistils than in anthers and, for this reason, we discarded these genes as male determinant candidates.

The second approach was to screen distinct 454 pyrosequencing libraries of cDNAs from olive flowers looking for genes, which are known to be specifically expressed in anthers and involved in pollen-pistil interaction. A total of 10 candidates were isolated and analyzed by qRT-PCR through three developmental stages, A1 through A3, of anthers in both cultivars Leccino and Frantoio (**Figure 1.20**).

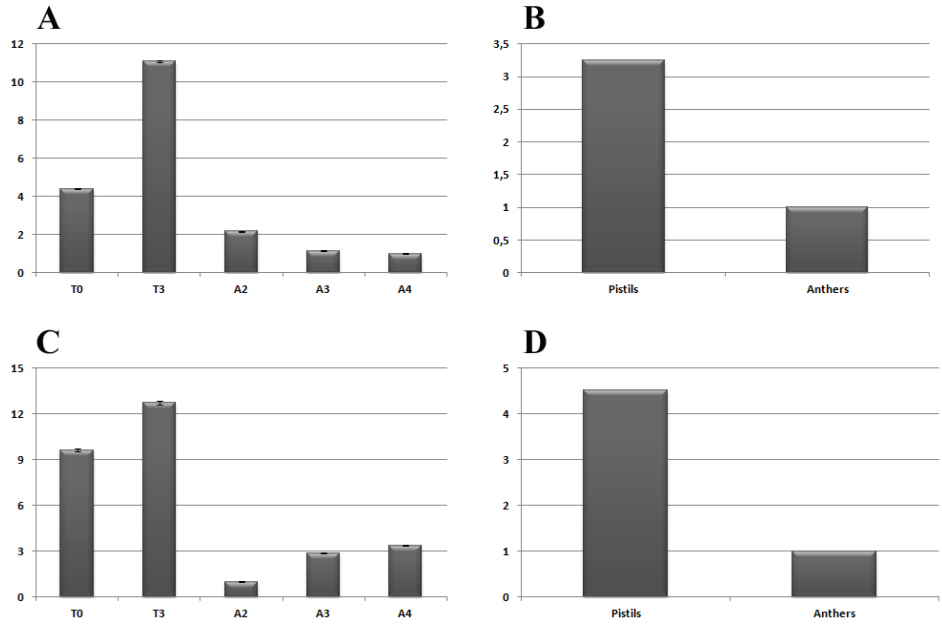


Figure 1.19. A, C: *OeSCR-like_1* (defensin proteins) expression patterns through different developmental stages of pistils (T0 and T3, flowering day and 3 days after flowering respectively) and of anthers (A2, A3 and A4, for further details see figure 33) in cultivars Leccino (A) and Frantoio (C). B, D: Comparison of expression in pistils and anthers in cultivars Leccino (B) and Frantoio (D). Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

In addition, qRT-PCR analyses were further performed in pistils and anthers to test the tissue-specificity expression. All genes confirmed their expected anther-specific expression even though some of them were anther-specific in only one cultivar. However, even if the amount of transcripts in pistils were low, it is worth noting that 8 of the 10 genes analyzed (*OebGlu1*, *OeGlu2*, *OeOle6*, *OeOle16*, *OeLat52*, *OePME1*, *OePME2* and *OeNTP*) had a similar expression pattern through the four developmental stages of pistils in both cultivars. In fact, in cv. Leccino the expression increased from T0 to T2 and then decreases through T4, while in cv. Frantoio the expression rapidly increased from T0 to T2, then it remained constant at the stage T3 and, lastly, rapidly increased to T4. These 8 genes also shared the expression pattern through the three developmental stages of anthers, but no differences between cultivars were revealed. In anthers the expression increased from A1 to A2 and then decreased to A3. Similar expression pattern for anthers was also observed for *OeDRP*, even though the gene did not share the same expression pattern in pistils. Interestingly, the

only gene analyzed that differed for the expression pattern in anthers was *OeOleosin1*, which exhibited a constant increase from A1 through A3 in both cultivars Leccino and Frantoio.

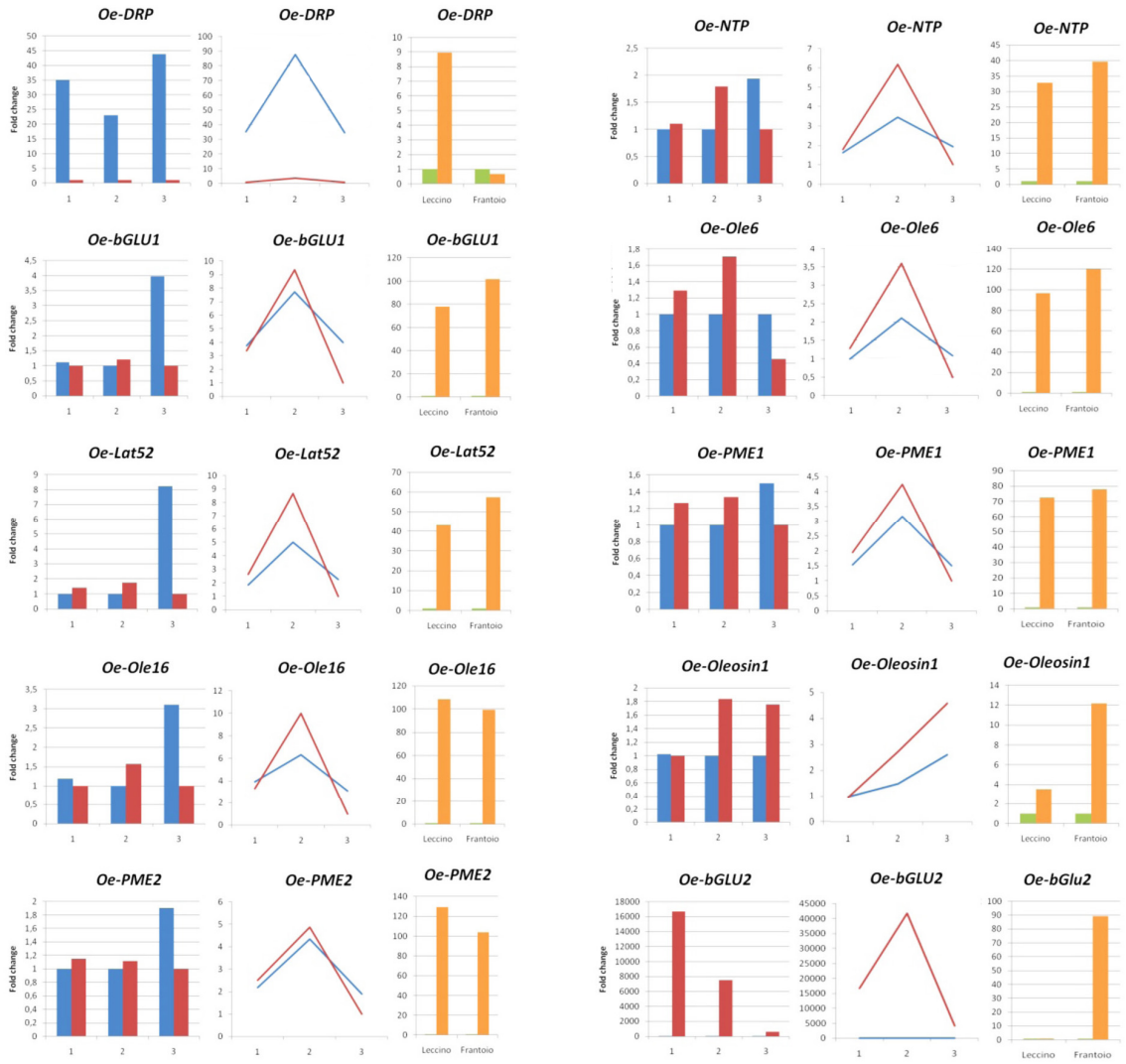


Figure 1.20. Gene expression levels assessed by qRT-PCRs for 10 genes known to be anther-specific expressed and to play a role in pollen-pistil interactions in other species. Blue: cv. Leccino. Red: cv. Frantoio. Histograms: comparison of gene expression values between cultivars. Lines: comparison of gene expression values within the same cultivar. Green: gene expression in pistils. Orange: gene expression in anthers. Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

These anther-specific genes represent a valid tool to study and characterize pollen-pistil interactions in olive, but, since we were wondering for a gene encoding for a cysteine rich

protein and since among these 10 anther-specific genes, no one showed cysteine residues conserved within the predicted amino acidic sequence, we performed a third screen of four distinct 454 pyrosequencing libraries. This screening was carried out using an algorithm built *ad hoc* and able to read all six reading frames of the 464,969 accessions and to isolate those with a cysteine rich structure.

This approach allowed to isolate 28 clones which were used to build a tree using the PHYLM algorithm and using one SCR nucleotide accession of *Brassica* as reference (**Figure 1.5**).

Ten of the 28 cDNA clones isolated using the algorithm were chosen as representative sequences and were analyzed by qRT-PCR. Expression analyses were performed in cv. Leccino in two development stages of pistils, T0 and T3, and in four developmental stages of anthers, from A1 to A4 (**Figure 1.21**).

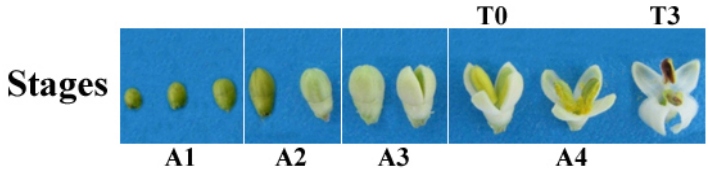


Figure 1.21. Olive flowers at different developmental stages used for gene expression analysis of the *OeSCR-like* genes having reference with the cysteine rich pattern built *ad hoc*. T0: flowering day. T3: 3 days after flowering. A1: immature anthers. A2: closed flowers with mature anthers. A3: open flowers with open anthers. A4: open flowers with dehiscent and senescent anthers.

Only 1 of the 10 cysteine-rich clones analyzed was anther-specific: *OeSCR-like_12*. The gene encoded for a 97 amino acid predicted protein, characterized by 12 residues of cysteine, having reference with proteins involved in gibberellins response (**Figure 1.22**). Gene expression analysis by qRT-PCR was repeated in both cultivars, Leccino and Frantoio, and the first result was confirmed (**Figure 1.23**). Expression patterns displayed a peak at the stage A2 and then decreased throughout the stage A4.

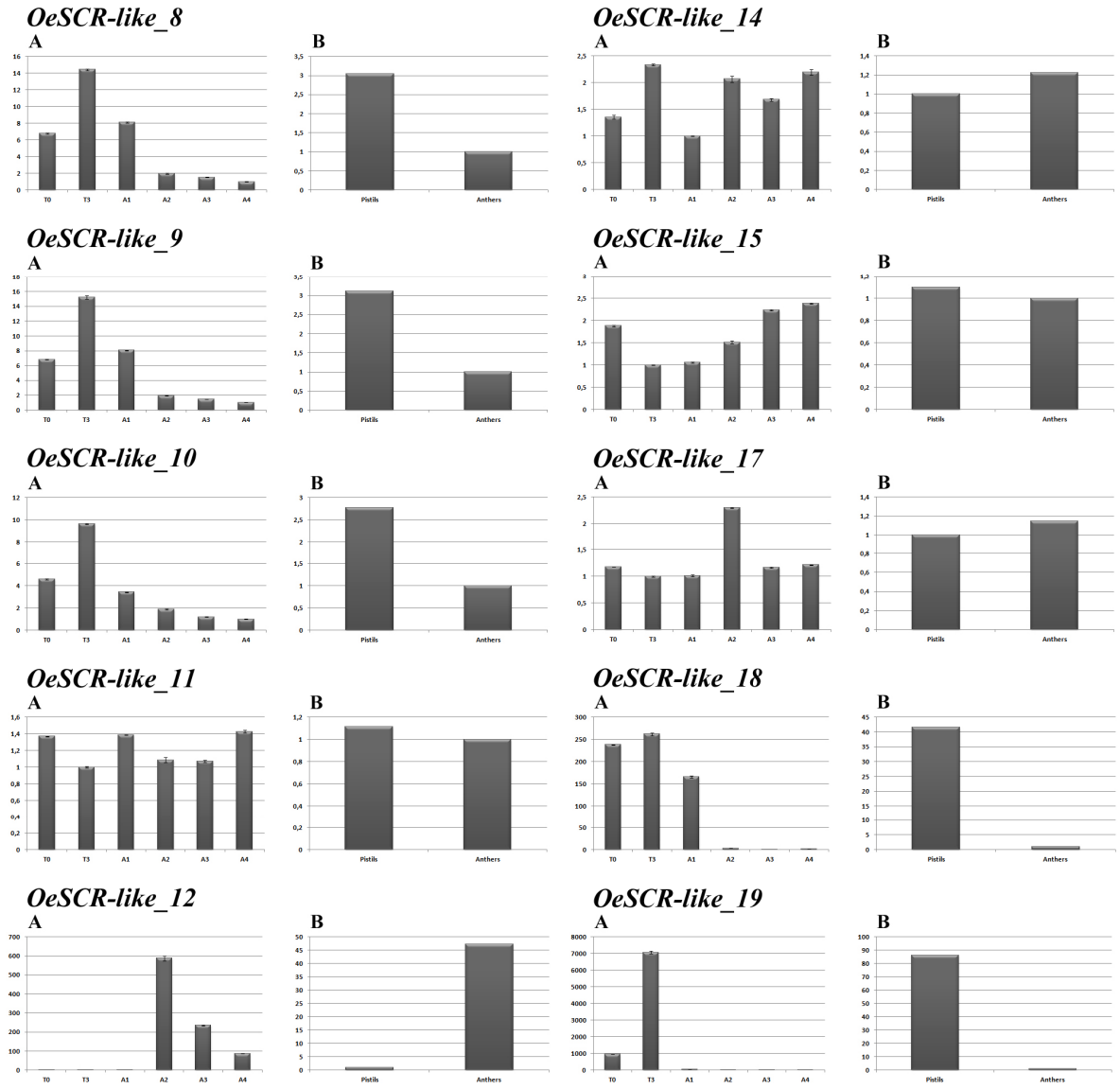


Figure 1.22. Gene expression levels assessed by qRT-PCRs for the 10 genes chosen as representative sample among the 28 accessions having reference with the cysteine-rich pattern of *SCR* of *Brassica* isolated from distinct 454 pyrosequencing libraries. A: gene expression patterns in pistils (T0 and T3, flowering day and 3 days after flowering respectively) and in anthers (A1, A2, A3 and A4, for further details see figure 1.21). B: Comparison of the expression levels between pistils and anthers. Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1

An anther-specific expression was confirmed for both cultivars. *OeSCR-like_12* was chosen as the best male determinant candidate and it was used, along with *OeSLG*, *OeSRK* and *OeSRK-like*, for further analyses, such as *in situ* hybridization and co-segregation analyses to test, respectively, the localization of transcripts and the association between putative SI genes..

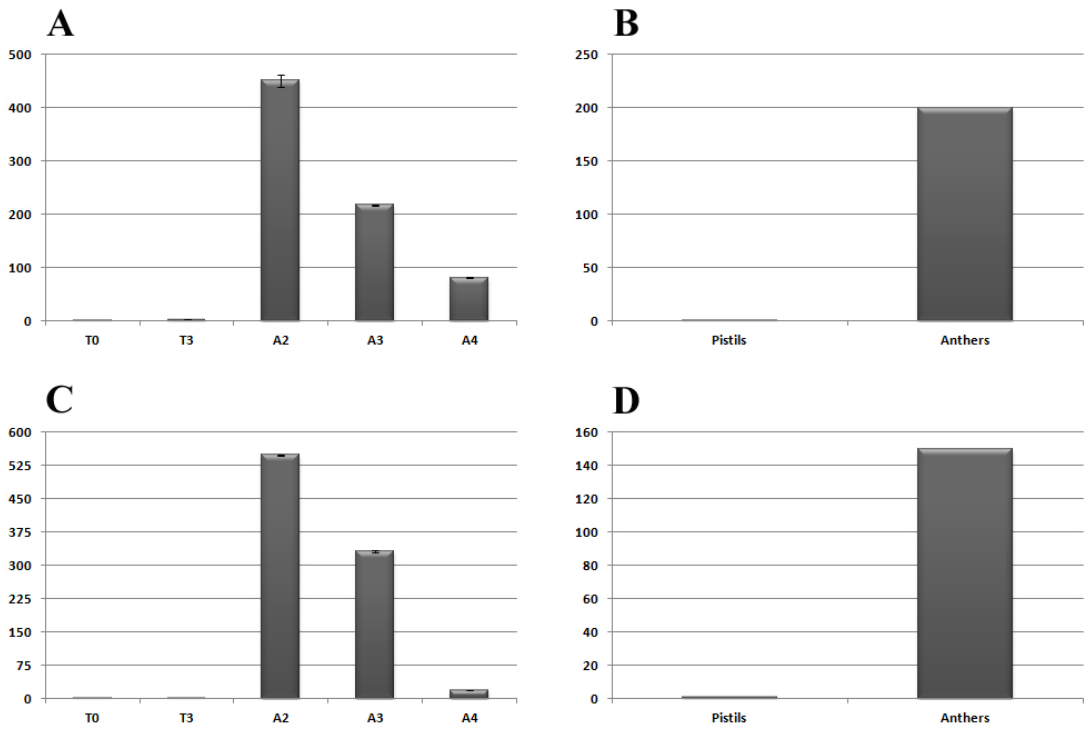


Figure 1.23. *OeSCR-like_12* replicated experiments. Expression patterns through different developmental stages of pistils (T0 and T3, flowering day and 3 days after flowering respectively) and of anthers (A2, A3 and A4, for further details see figure 1.21) in cultivars Leccino (A) and Frantoio (C). Comparison of expression in pistils and anthers in cultivars Leccino (B) and Frantoio (D). Ordinate values indicate the fold change of normalized values calculated according to the minimum value for each graph which was assessed with the value 1.

All the experiments aimed at finding a good male determinant candidate in olive allowed to isolate 3 genes encoding for defensin protein but with no anther-specific expression, 10 genes specifically expressed in anthers and putatively involved in anther-pistil interaction, but without any similarity with the amino acid structure of *SCR* of *Brassica*, and 10 genes having reference with the cysteine-rich pattern of *SCR* among which one was specifically expressed in anthers. All these genes are now available for further experiments.

In situ hybridization analyses

In situ hybridization analyses were performed using flowers from cultivars Leccino and Frantoio in order to visualize the localization of transcripts of the candidate SI-related genes. Two *SRK-like* genes, named *OeSRK1* and *OeSRK-like*, one *SLG* gene, named *OeSLG*, and one *SCR-like* gene, named *OeSCR-like_12*, were tested in flowers at different developmental stages. The olive flower is characterized by hard tissues, particularly if compared to other species, and this might be the cause of wax-inclusion problems found during our experiments. Severe problems were found in anthers, which are characterized by the presence of air at the mature stage, whereas no problems were experienced in pistils. *In sity* hybridizations with *OeSRK* and *OeSLG* probes showed a signal localized in the surface of papillae cells of the stigma, whereas hybridizations using *OeSRK-like* and *OeSCR-like_12* probes did not reveal any signal (**Figure 1.24**). It is worth noting that, even though it was not possible to test the putative male determinant *OeSCR-like_12* in anthers, the absence of signal in pistils might be considered a good negative control. Negative control hybridizations without probe were also performed and no signals were recovered, indicating that the signal obtained with *OeSRK* and *OeSLG* probes was reliable. Signals of the putative female determinant, *OeSRK*, and its putative enhancer, *OeSLG*, in the surface of papillae was in agreement with the hypothesis of a role played by these genes in the SI response.

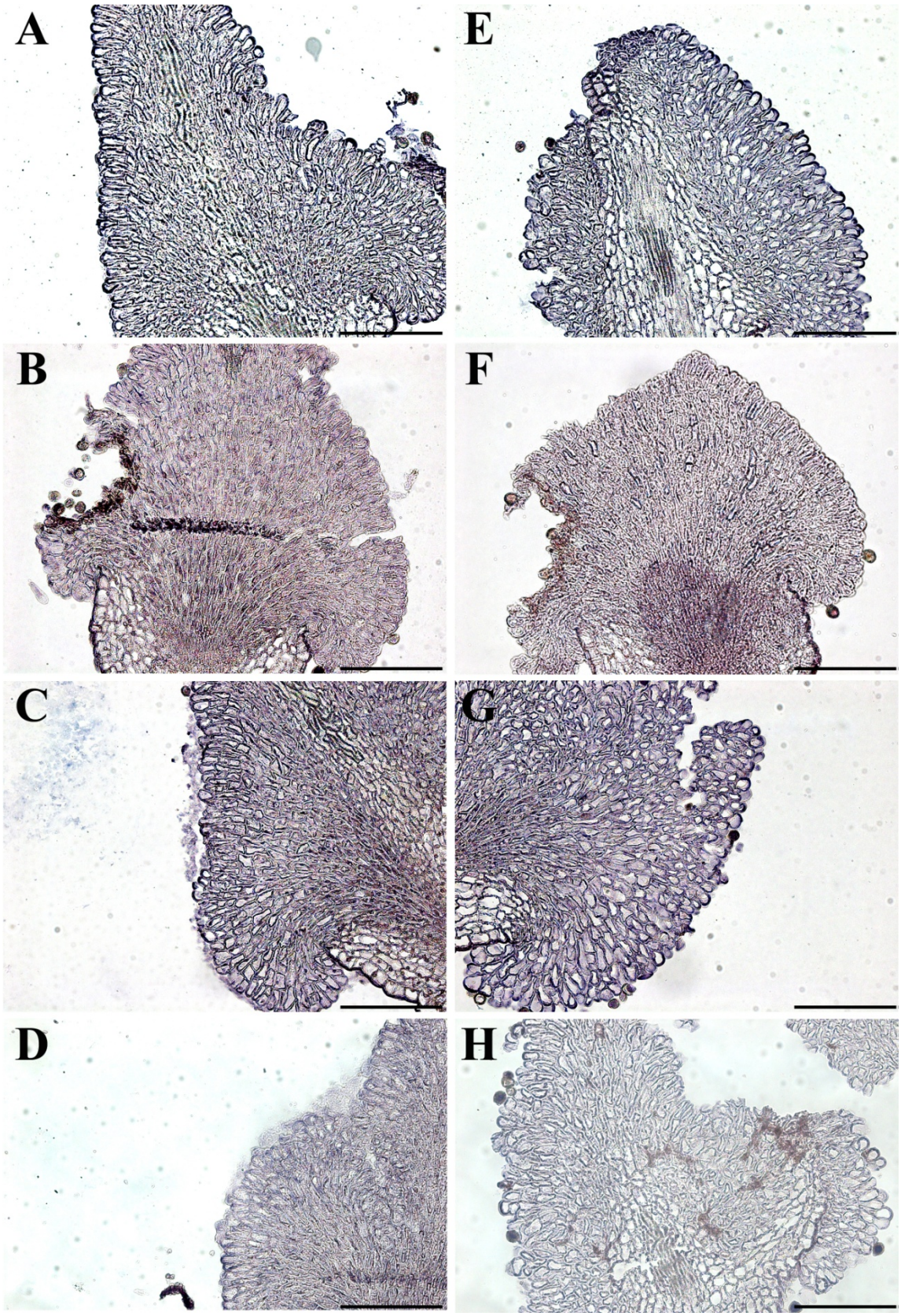


Figure 1.24. Localization of expression domains for *OeSRK* (A, E), *OeSRK-like* (B, F), *OeSLG* (C, G) and *OeSCR-like_12* (D, H) genes as assessed by *in situ* hybridization analysis in stigmas. A-D: self-compatible cultivar Leccino. E-F: self-compatible cultivar Frantoio. Bars= 100 µm.

Identification of conserved miRNAs in olive and their target prediction

After the first screening of 464,969 accessions, belonging to four distinct libraries of olive flower cDNAs, aimed at identifying putative miRNAs with a maximum of 4 mismatches comparing to those currently available in the database (3,938 miRNAs belonging to 51 families), as many as 200,709 miRNA candidates were isolated. The screening to predict folding and secondary structure led to the identification of 1,490 putative miRNAs and their related pre-miRNAs. Prediction of secondary structures according to the guidelines reported in Meyers *et al.* (2008) and with a ΔG value more than 14 kj allowed us the isolation of 204 definitive miRNA candidates (**Figure 1.25** and **Appendix 1.2**).

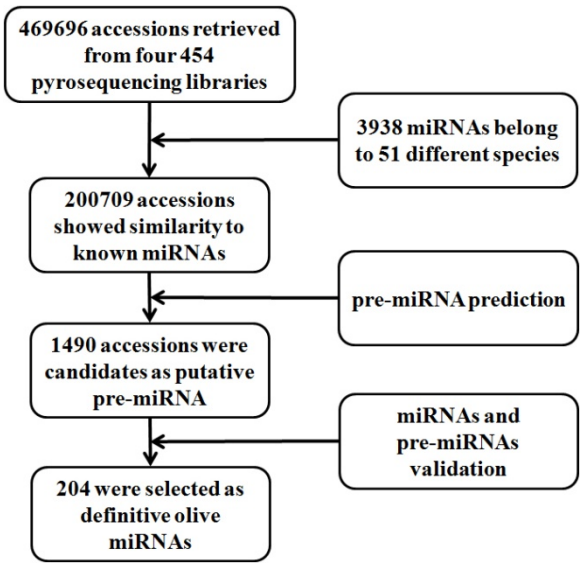


Figure 1.25. Multi-step process followed to obtain the overall 204 olive miRNAs and their pre-miRNAs.

Specifically, 104 accessions having reference with 104 different miRNAs families were extracted and, since the screening was carried out in four libraries that differed each other for cultivar genotypes and developmental stages, a further 100 miRNAs were recovered as being characterized by small differences in the miRNAs sequence as well as differences in the nucleotide composition of their pre-miRNA sequences. All stem-loop structures of pre-miRNAs were represented by images and their values of ΔG were reported (**Appendix**

1.3). Size of miRNAs was ranging between 18 and 22 nucleotides. In details the 18nt miRNAs were 9, the 19nt were 22, the 20nt were 56, the 21nt were 92, the 22nt were 25 and the average value was of 20.5 nucleotides. Size of pre-miRNAs was ranging from 60 to 206 nucleotides, with an average value of 100.69 nucleotides.

Conservation analysis of miRNAs families throughout different species was carried out using all the miRNA families present in two or more species for a total of 204 miRNA conserved families. Results showed that the families currently more conserved throughout the species were miR156, miR167, miR159 being represented in 28 species followed by miR159 and miR166 in 26 species and by miR395 in 25 families. The presence in more than 20 species was achieved for miR160, miR169, miR172, miR319 and miR398. Representation in more than 15 species were found for miR164, miR168, miR399, miR395, miR162, miR408, miR390, miR393, miR397 and miR482. Lastly, the presence in more than 10 species was found for miR384 and miR827 (**Table 1.5**). A graphical representation was reported in **Figure 1.26**.

Table 1.5. Summary of miRNA families conserved in two or more species. miRNA: name of the miRNA family; N: progressive number referred to figure ???; NS: number of families exhibiting that specific miRNA.

miRNA	**N**	**NS**	**miRNA**	**N**	**NS**	**miRNA**	**N**	**NS**	**miRNA**	**N**	**NS**
miR156	**1**	**28**	miR530	52	6	miR1023	103	2	miR2093	154	2
miR157	2	5	miR535	53	8	miR1027	104	2	miR2109	155	2
miR158	3	3	miR536	54	2	miR1038	105	2	miR2111	156	8
miR159	4	26	miR771	55	2	miR1039	106	2	**miR2112**	**157**	**2**
miR160	5	23	miR773	56	2	miR1040	107	2	miR2118	158	4
miR161	6	4	miR774	57	2	miR1042	108	2	miR2119	159	3
miR162	7	17	miR781	58	3	**miR1044**	**109**	**2**	miR2275	160	2
miR163	8	2	miR783	59	3	miR1052	110	2	miR2592	161	2
miR164	9	19	miR813	60	2	miR1056	111	2	miR2595	162	2
miR165	10	2	**miR816**	**61**	**2**	miR1063	112	2	miR2612	163	2
miR166	11	26	miR821	62	2	miR1067	113	2	miR2628	164	2
miR167	12	28	miR822	63	2	miR1097	114	2	miR2630	165	2
miR168	**13**	**19**	miR823	64	2	miR1119	115	2	miR2633	166	2
miR169	14	22	miR824	65	5	miR1120	116	2	mIR2641	167	2
miR170	15	2	miR825	66	3	miR1122	117	4	miR2651	168	2
miR171	16	28	miR827	67	10	miR1126	118	2	**miR2654**	**169**	**2**
miR172	17	20	miR828	68	3	miR1127	119	2	miR2657	170	2
miR173	18	2	miR829	69	2	miR1128	120	2	miR2868	171	2
miR319	19	20	miR831	70	2	**miR1134**	**121**	**2**	miR2873	172	2
miR390	20	16	miR833	71	2	miR1135	122	2	miR2911	173	2
miR391	21	2	miR834	72	3	miR1139	123	2	miR2913	174	2
miR393	22	15	**miR835**	**73**	**2**	miR1311	124	2	miR2919	175	2
miR394	23	14	miR837	74	3	miR1313	125	2	miR2923	176	2
miR395	24	18	miR838	75	3	miR1320	126	2	miR2925	177	2
miR396	**25**	**25**	miR839	76	2	miR1428	127	2	miR2928	178	2
miR397	26	15	miR840	77	2	miR1432	128	4	miR2937	179	2
miR398	27	20	miR841	78	2	miR1433	129	2	miR2950	180	3
miR399	28	19	miR842	79	2	miR1435	130	3	**miR3434**	**181**	**3**
miR400	29	2	miR844	80	2	miR1436	131	2	miR3438	182	2
miR402	30	3	miR845	81	3	miR1444	132	2	miR3440	183	2
miR403	31	7	miR846	82	2	**miR1450**	**133**	**2**	miR3447	184	2
miR408	32	17	miR847	83	3	miR1507	134	4	miR3476	185	2
miR413	33	2	miR848	84	2	miR1508	135	3	miR3511	186	2
miR414	34	4	**miR851**	**85**	**2**	miR1509	136	3	miR3513	187	2
miR415	35	2	miR852	86	2	miR1510	137	4	miR3706	188	2
miR416	36	2	miR853	87	2	miR1511	138	2	miR3712	189	2
miR417	**37**	**2**	miR854	88	2	miR1514	139	2	miR3950	190	2
miR418	38	2	miR856	89	2	miR1515	140	2	miR3953	191	2
miR419	39	3	miR857	90	4	miR1518	141	2	miR4221	192	3
miR420	40	2	miR858	91	2	miR1521	142	2	**miR4227**	**193**	**2**
miR426	41	2	miR859	92	2	miR1523	143	2	miR4228	194	2
miR437	42	6	miR860	93	2	miR1527	144	2	miR4239	195	2
miR444	43	6	miR861	94	2	**miR1533**	**145**	**2**	miR4240	196	2
miR447	44	2	miR862	95	2	miR1535	146	2	miR4243	197	2
miR472	45	5	miR863	96	2	miR1852	147	2	miR4350	198	2
miR473	46	3	**miR868**	**97**	**2**	miR1862	148	2	miR5021	199	2
miR477	47	6	miR869	98	2	miR1863	149	2	miR5138	200	2
miR479	48	4	miR948	99	2	miR1878	150	2	miR5140	201	2
miR482	**49**	**15**	miR950	100	3	miR1887	151	2	miR5142	202	2
miR528	50	5	miR951	101	2	miR1888	152	2	miR5202	203	2
miR529	51	8	miR952	102	2	miR2083	153	2	miR5301	204	2

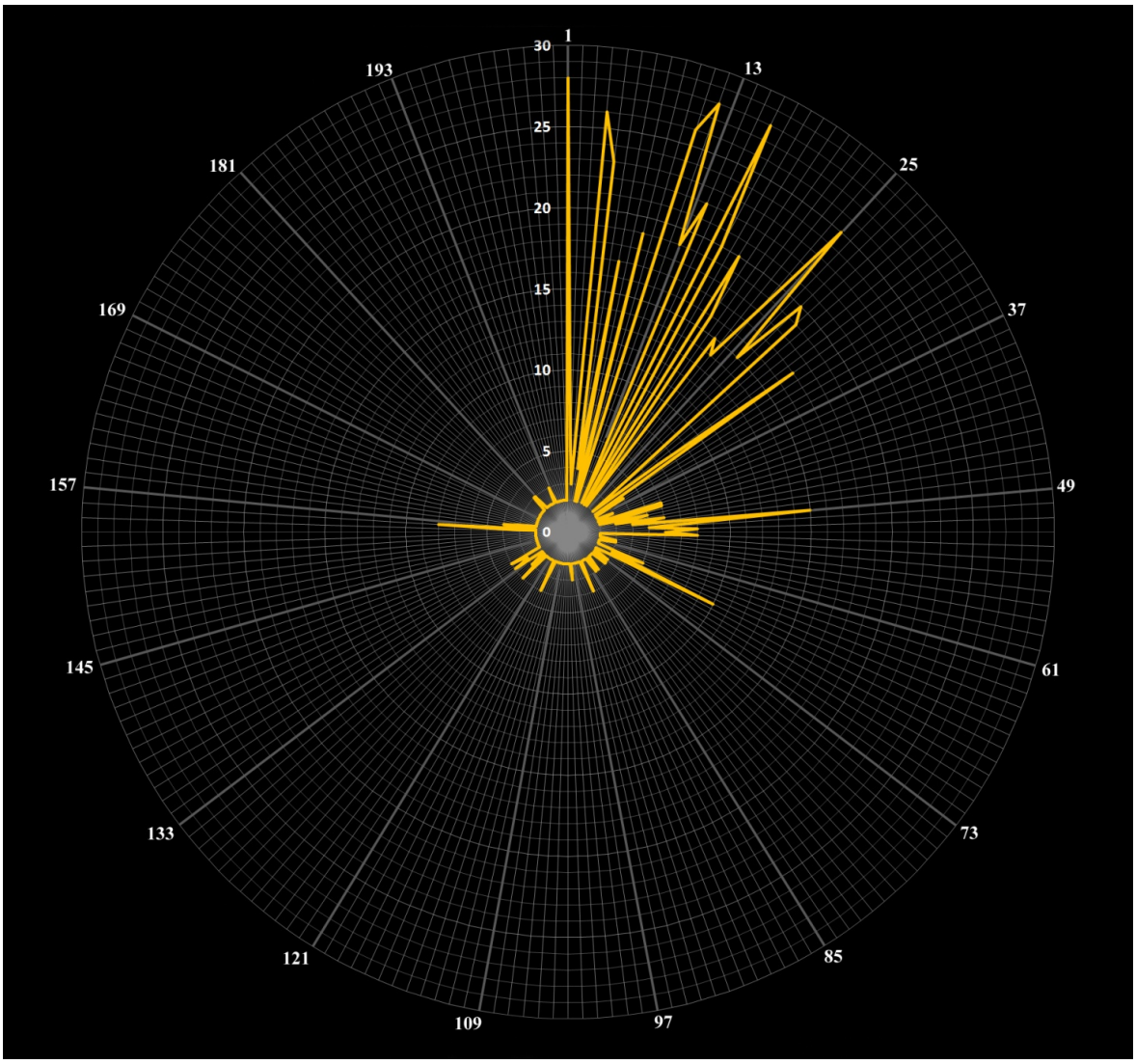


Figure 1.26. Graph representing the conservation of miRNA families throughout different species. Radial axes are the progressive number referring to specific miRNA. Concentric axes are the number of families which exhibits that specific miRNA. Some numbers are reported to facilitate the reading of the graph. For further details see table ???.

Target prediction of olive miRNAs against *Arabidopsis* genome allowed to annotate 186 out of 204 miRNAs and only the remaining 18 miRNAs did not reveal any target (**Table 1.7**). Among the multiple targets predicted for each miRNA, only the first option was reported, which was the best option considering "Exp" and "UPE" values. "Exp" value is the number of mismatches between miRNA and target sequences, and the lowest is the value of this number the higher is the probability of interaction. "UPE" value is the maximum energy to unpair the target site, which is the energy required to open (unpair) secondary structure around target site on target mRNA, thus the less energy means the more possibility that miRNA is able to bind the target mRNA.

Results provided by the two screenings aimed to predict miRNAs that might interact with our SI genes, *OeSRK*, *OeSLG* and *OeSCR-like*, using different stringent conditions were pooled together. **Table 1.6** shows that 38 miRNAs were predicted to interact with *OeSRK*, 12 with *OeSLG* and 19 with *OeSCR-like*.

Table 1.6. miRNAs predicted to interact with the SI genes *OeSRK*, *OeSLG* and *OeSCR-like*.

miRNA	SI gene	miRNA	SI gene	miRNA	SI gene	miRNA	SI gene
oeu-miR414e	SRK	oeu-miR1533e	SRK	oeu-miR863	SLG	oeu-miR156a	SCR
oeu-miR477a	SRK	oeu-miR1533f	SRK	oeu-miR1582	SLG	oeu-miR863	SCR
oeu-miR482j	SRK	oeu-miR1533g	SRK	oeu-miR2628	SLG	oeu-miR1056	SCR
oeu-miR529a	SRK	oeu-miR1533h	SRK	oeu-miR3434a	SLG	oeu-miR1122	SCR
oeu-miR529b	SRK	oeu-miR1535d	SRK	oeu-miR5021d	SLG	oeu-miR1533b	SCR
oeu-miR816	SRK	oeu-miR2093a	SRK	oeu-miR5021f	SLG	oeu-miR1533c	SCR
oeu-miR838a	SRK	oeu-miR2628	SRK	oeu-miR5140a	SLG	oeu-miR1533d	SCR
oeu-miR838b	SRK	oeu-miR2928c	SRK	oeu-miR5140c	SLG	oeu-miR1533e	SCR
oeu-miR847	SRK	oeu-miR3434a	SRK	oeu-miR5140d	SLG	oeu-miR1533f	SCR
oeu-miR1044a	SRK	oeu-miR3513	SRK	oeu-miR5140f	SLG	oeu-miR1533h	SCR
oeu-miR1044b	SRK	oeu-miR3706c	SRK	oeu-miR5140g	SLG	oeu-miR2923	SCR
oeu-miR1097	SRK	oeu-miR5021e	SRK	oeu-miR5142	SLG	oeu-miR2925d	SCR
oeu-miR1320	SRK	oeu-miR5021f	SRK			oeu-miR2925e	SCR
oeu-miR1435b	SRK	oeu-miR5138	SRK			oeu-miR5021b	SCR
oeu-miR1521	SRK	oeu-miR5140a	SRK			oeu-miR5021d	SCR
oeu-miR1533a	SRK	oeu-miR5140b	SRK			oeu-miR5021f	SCR
oeu-miR1533b	SRK	oeu-miR5140c	SRK			oeu-miR5140f	SCR
oeu-miR1533c	SRK	oeu-miR5140f	SRK			oeu-miR863	SCR
oeu-miR1533d	SRK	oeu-miR5140g	SRK			oeu-miR1056	SCR

All of these predicted cases of interaction have to be demonstrated by further studies.

Table 1.7. miRNAs prediction target against *Arabidopsis* genome. miRNAAcc: name of olive miRNAs. Target_acc: accession number of target gene having reference with Arabidopsis genome. Exp.: mismatches of pairing between miRNAs and targets. UPE: maximum energy to unpair the target site, which is the energy required to open (unpair) secondary structure around target site on target mRNA, thus the less energy means the more possibility that miRNA is able to contact target mRNA. Target start/end: position of the target sequence within the target mRNA. miRNA_aligned_fragment: olive miRNA sequences having homology with target mRNA. Target_aligned_fragment: target mRNA sequences having references with olive miRNAs. Inhibition: predicted method of inactivation target mRNAs, cleavage or translation of inhibition. Target_Desc.: description of target mRNAs.

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR156a	AT1G27370.1	0.5	12.274	2368	2388	UUGACAGAAGAGAGAGAGCAU	GUGCUCUCUCUCUUCUGUCAA	Cleavage	| Symbols: | squamosa promoter-binding protein-like 10 (SPL10) | chr1:9505201-9508280 REVERSE [PFAM] 1806-2036 PF03110.7 SBP domain;
Oeu-miR156b	AT3G17820.1	3.0	19.515	733	752	CAGAAGGUAGCGAGCAGCUC	GGGUUGCUCGAUACCUUCUC	Translation	| Symbols: ATGSKB6, GLN1.3, GLN1;3 | ATGSKB6; copper ion binding / glutamate-ammonia ligase | chr3:6097414-6099595 FORWARD [PFAM]
Oeu-miR156c	AT3G17820.1	3.0	19.515	733	752	CAGAAGGUAGCGAGCAGCUC	GGGUUGCUCGAUACCUUCUC	Translation	| Symbols: ATGSKB6, GLN1.3, GLN1;3 | ATGSKB6; copper ion binding / glutamate-ammonia ligase | chr3:6097414-6099595 FORWARD [PFAM]
Oeu-miR158	AT2G11390.1	3.0	12.221	762	781	UGCAAAAUGUAGACAGAGUA	UGUUUUGUUUAUAUUUUGUA	Cleavage	| Symbols: | transposable element gene | chr2:4536804-4537705 FORWARD [PFAM]
Oeu-miR159a	AT5G67090.1	2.5	19.507	1155	1175	AUUGGAUUGAAGGGAGCUCUC	GAGAGUUCCCUUCGGUUCAGU	Cleavage	| Symbols: | subtilase family protein | chr5:26774018-26776391 REVERSE [PFAM]
Oeu-miR159b	AT3G21470.1	3.0	16.936	672	691	UUGUAUUGAAGAGAGCAUCA	UGAUGCUUUCUUUAAUAUGC	Cleavage	| Symbols: | pentatricopeptide (PPR) repeat-containing protein | chr3:7563608-7565248 FORWARD [PFAM]
Oeu-miR159c	None								
Oeu-miR161	None								
Oeu-miR162a	None								
Oeu-miR162b	AT2G35540.1	3.0	10.935	591	611	UCAAUAGACCUCUGCCUCCUG	UAGGAGGAAGAGGCCUGUUGA	Cleavage	| Symbols: | DNAJ heat shock N-terminal domain-containing protein | chr2:14927158-14928930 FORWARD [PFAM]
Oeu-miR167a	AT1G35840.1	3.0	12.254	3972	3992	UGAAUGUGCCAGCAAGAUAUU	GAUAUCUAGCUGGCACAAUCA	Cleavage	| Symbols: | transposable element gene | chr1:13324970-13329493 FORWARD [PFAM]
Oeu-miR167b	AT1G35840.1	3.0	12.254	3972	3992	UGAAUGUGCCAGCAAGAUAUU	GAUAUCUAGCUGGCACAAUCA	Cleavage	| Symbols: | transposable element gene | chr1:13324970-13329493 FORWARD [PFAM]
Oeu-miR167c	AT1G35840.1	3.0	12.254	3972	3992	UGAAUGUGCCAGCAAGAUAUU	GAUAUCUAGCUGGCACAAUCA	Cleavage	| Symbols: | transposable element gene | chr1:13324970-13329493 FORWARD [PFAM]
Oeu-miR169	AT3G20910.1	2.0	15.333	1035	1054	AGCCAAGAAUGACUUGCCGA	GCGGCAAUUCAUUCUUGGCU	Cleavage	| Symbols: NF-YA9 | NF-YA9 (NUCLEAR FACTOR Y, SUBUNIT A9); specific transcriptional repressor/ transcription factor | chr3:7326382-7328576 FORWARD [PFAM] 609-776 PF02045.8 CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B;
Oeu-miR171a	AT5G59400.2	2.5	17.1	32	52	UCAUUGAGACGAGGCCUUGUC	GAUAAGGCCUCGUCUCAUUGG	Cleavage	| Symbols: | unknown protein | chr5:23957745-23959632 FORWARD [PFAM]
Oeu-miR171b	AT5G59400.2	2.5	17.1	32	52	UCAUUGAGACGAGGCCUUGUC	GAUAAGGCCUCGUCUCAUUGG	Cleavage	| Symbols: | unknown protein | chr5:23957745-23959632 FORWARD [PFAM]
Oeu-miR171c	AT1G63330.1	3.0	12.417	487	507	AUGAGCCGAAUCAAUAUCACU	GGUGAUAUUGAUUUGGCUUUU	Cleavage	| Symbols: | pentatricopeptide (PPR) repeat-containing protein | chr1:23489628-23491519 FORWARD [PFAM]
Oeu-miR171d	AT3G60630.1	1.0	14.373	1041	1061	UGAGCCGCGCCAAUAUCACUU	AAGGGAUAUUGGCGCGGCUCA	Cleavage	| Symbols: | scarecrow transcription factor family protein | chr3:22410365-22412608 REVERSE [PFAM]
Oeu-miR172a	AT4G16670.1	2.5	16.79	835	855	UCAAUCCUGAUGCAGCUGCAG	CUGCAGCUGCAACAGCAUUGA	Translation	| Symbols: | phosphoinositide binding | chr4:9385219-9387754 FORWARD [PFAM] 1023-1337 PF08458.3 Plant pleckstrin homology-like region;
Oeu-miR172b	AT4G16670.1	2.5	16.79	835	855	UCAAUCCUGAUGCAGCUGCAG	CUGCAGCUGCAACAGCAUUGA	Translation	| Symbols: | phosphoinositide binding | chr4:9385219-9387754 FORWARD [PFAM] 1023-1337 PF08458.3 Plant pleckstrin homology-like region;
Oeu-miR394	AT1G42320.1	3.0	19.946	1055	1074	UGGUCAUUCUGUCCAGUUCC	AGAGCUGGACAGAAUGGUCG	Cleavage	| Symbols: | transposable element gene | chr1:15800929-15805502 FORWARD [PFAM]
Oeu-miR395a	AT4G03415.1	3.0	17.649	617	637	CUGAAGUGUCUGGAGGAAAGC	UUUUUCUUUCAGACGCUUCAG	Cleavage	| Symbols: | catalytic/ protein serine/threonine phosphatase | chr4:1503708-1506500 REVERSE [PFAM]
Oeu-miR395b	AT4G03415.1	3.0	17.649	617	637	CUGAAGUGUCUGGAGGAAAGC	UUUUUCUUUCAGACGCUUCAG	Cleavage	| Symbols: | catalytic/ protein serine/threonine phosphatase | chr4:1503708-1506500 REVERSE [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR395c	AT5G43780.1	0.5	11.92	395	414	UGAAGUGUUUGGGGGAACUC	GAGUUCCUCCAAACACUUCA	Cleavage	| Symbols: APS4 | APS4; sulfate adenylyltransferase (ATP) | chr5:17589394-17591535 REVERSE [PFAM]
Oeu-miR395d	AT5G43780.1	1.5	11.571	395	415	CUGAAGUGUUUGGGGGAACUC	GAGUUCCUCCAAACACUUCAU	Cleavage	| Symbols: APS4 | APS4; sulfate adenylyltransferase (ATP) | chr5:17589394-17591535 REVERSE [PFAM]
Oeu-miR395e	AT5G43780.1	2.0	11.571	395	415	CUGAAGUGUUUGGGGGAGCUC	GAGUUCCUCCAAACACUUCAU	Cleavage	| Symbols: APS4 | APS4; sulfate adenylyltransferase (ATP) | chr5:17589394-17591535 REVERSE [PFAM]
Oeu-miR396	AT4G16830.2	2.0	14.51	1000	1020	UCUCACAGCUUUCUUGGCCUU	AAGGCUAAGAAGGCUGUGAGC	Cleavage	| Symbols: | nuclear RNA-binding protein (RGGA) | chr4:9470409-9472594 FORWARD [PFAM]
Oeu-miR402	None								
Oeu-miR414a	AT5G57685.1	0.0	14.158	425	445	UCUUCAUCGUCAUCAUCGUCA	UGACGAUGAUGACGAUGAAGA	Cleavage	| Symbols: AtGDU3 | AtGDU3 (Arabidopsis thaliana GLUTAMINE DUMPER 3) | chr5:23366004-23366867 REVERSE [PFAM]
Oeu-miR414b	AT3G18100.2	0.5	18.696	150	170	UCAUCUUCUUCGUCAUUGUCA	UGACGAUGACGAAGAAGAUGA	Cleavage	| Symbols: MYB4R1 | MYB4R1 (myb domain protein 4R1); transcription factor | chr3:6200578-6204583 FORWARD [PFAM]
Oeu-miR414c	AT5G19900.1	1.5	11.494	780	800	UCAUUUUUAUCAUCCUCUUCA	UGAGGAGGAUGAUGAAGAUGA	Cleavage	| Symbols: | PRLI-interacting factor, putative | chr5:6728160-6730045 REVERSE [PFAM]
Oeu-miR414d	AT3G19510.1	1.5	18.829	1263	1283	UCAUAUUCUUCAUCAUCAGCA	UUCUGAUGAUGAAGAGUAUGA	Cleavage	| Symbols: HAT3.1 | HAT3.1; DNA binding / sequence-specific DNA binding / transcription activator/ transcription factor | chr3:6762764-6766588 REVERSE [PFAM]
Oeu-miR414e	AT1G13220.2	0.5	23.816	3048	3068	UCAUCUUCAUCAGCAGCUGCA	UGCAGCUGCUGGUGAAGAUGA	Cleavage	| Symbols: LINC2 | LINC2 (LITTLE NUCLEI2) | chr1:4515699-4520284 FORWARD [PFAM]
Oeu-miR437	AT1G17147.1	2.0	12.072	455	475	AAAAAGAAAGAAGUUUAACUU	AAGGUGAACUUCUUUUUUUUU	Cleavage	| Symbols: | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: VQ (InterPro:IPR008889); BEST Arabidopsis thaliana protein match is: VQ motif-containing protein (TAIR:AT1G78410.1); Has 18 Blast hits to 18 proteins in 5 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 18; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). | chr1:5863431-5864009 REVERSE [PFAM]
Oeu-miR444a	AT3G47760.1	2.5	17.149	1183	1203	AGCUGGUGCUGUUUCAAGCUU	AAGAUUGAGACAGCAUCAGUU	Cleavage	| Symbols: ATATH4, ATH4 | ATATH4; ATPase, coupled to transmembrane movement of substances / transporter | chr3:17611787-17616639 FORWARD [PFAM]
Oeu-miR444b	AT3G42798.1	3.0	11.255	45	65	UGCAGAUGAUUUGUCAAGCUU	AAGUUUGAAAAAUCGUUUGCG	Cleavage	| Symbols: | transposable element gene | chr3:14901653-14902694 FORWARD [PFAM]
Oeu-miR472a	AT5G43730.1	3.0	17.179	623	644	CAUUUCCUAUUCCUCCCAUACC	GGUAUGGGGGGAAUAGGAAAAA	Cleavage	| Symbols: | disease resistance protein (CC-NBS-LRR class), putative | chr5:17560179-17562929 FORWARD [PFAM]
Oeu-miR472b	AT5G43730.1	3.0	17.179	623	644	CAUUUCCUAUUCCUCCCAUACC	GGUAUGGGGGGAAUAGGAAAAA	Cleavage	| Symbols: | disease resistance protein (CC-NBS-LRR class), putative | chr5:17560179-17562929 FORWARD [PFAM]
Oeu-miR472c	AT3G46730.1	2.0	20.722	574	594	AUUUCCUAUACCUCCCAUACC	GGUAUGGGAGGUCUAGGAAAA	Translation	| Symbols: | disease resistance protein (CC-NBS class), putative | chr3:17213069-17215612 REVERSE [PFAM]
Oeu-miR473	AT5G39990.1	2.5	19.519	1413	1432	ACUCAUUCUCACGGCUUCCA	UGGAAGCCAUGAGAAUGGGA	Cleavage	| Symbols: | glycosyltransferase family 14 protein / core-2/I-branching enzyme family protein | chr5:16004281-16006740 FORWARD [PFAM]
Oeu-miR477a	AT5G38880.1	1.5	15.513	848	867	UUCUUCUUUAGAGGCUUCCA	UGGAGGUCUCUAGAGAAGAA	Cleavage	| Symbols: | unknown protein | chr5:15563306-15569194 REVERSE [PFAM]
Oeu-miR477b	AT1G01590.1	2.5	10.063	881	901	AGAUGCCCUUGUGGAAGACGA	UCGUCUUCCACGUCGGCAUCU	Translation	| Symbols: FRO1, ATFRO1 | FRO1 (FERRIC REDUCTION OXIDASE 1); ferric-chelate reductase | chr1:214229-217304 FORWARD [PFAM]
Oeu-miR477c	AT1G24310.1	2.0	15.069	1426	1445	UUCUCUUUCAAAGGCUUGAA	UUAAAGCCUUUGAAAGAGAC	Cleavage	| Symbols: | unknown protein | chr1:8624061-8626405 FORWARD [PFAM]
Oeu-miR482a	AT3G46530.1	2.5	16.423	666	687	UCUUUCCGAGUCCUCCCAUUCC	GGUAUGGGAGGCCUUGGAAAGA	Translation	| Symbols: RPP13 | RPP13 (RECOGNITION OF PERONOSPORA PARASITICA 13); ATP binding | chr3:17130367-17133338 REVERSE [PFAM]
Oeu-miR482b	AT3G46530.1	2.5	16.423	666	687	UCUUUCCGAGUCCUCCCAUUCC	GGUAUGGGAGGCCUUGGAAAGA	Translation	| Symbols: RPP13 | RPP13 (RECOGNITION OF PERONOSPORA PARASITICA 13); ATP binding | chr3:17130367-17133338 REVERSE [PFAM]
Oeu-miR482c	AT5G43730.1	3.0	17.179	623	644	CAUUUCCUAUUCCUCCCAUACC	GGUAUGGGGGGAAUAGGAAAAA	Cleavage	| Symbols: | disease resistance protein (CC-NBS-LRR class), putative | chr5:17560179-17562929 FORWARD [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR482d	AT5G43730.1	3.0	17.179	623	644	CAUUUCCUAUUCCUCCCAUACC	GGUAUGGGGGGAAUAGGAAAAA	Cleavage	| Symbols: | disease resistance protein (CC-NBS-LRR class), putative | chr5:17560179-17562929 FORWARD [PFAM]
Oeu-miR482e	None								
Oeu-miR482f	AT3G62130.1	2.0	21.605	445	464	UCUUUCUUGUACCUCCCUUC	GAAGGGAAGUAUAAGAAGGA	Cleavage	| Symbols: | epimerase-related | chr3:23004931-23007006 FORWARD [PFAM]
Oeu-miR482g	AT5G50760.1	2.0	16.586	107	126	UCUUCGUUGUUUCUCCCUUU	GAAGAGAGAAACAAUGAAGA	Cleavage	| Symbols: | auxin-responsive family protein | chr5:20644661-20645469 FORWARD [PFAM]
Oeu-miR482h	AT1G18670.1	0.5	10.007	1771	1790	UCUUCCUUUUUUUUCUCAUU	AAUGAGAAAGAAAAGGAAGA	Cleavage	| Symbols: IBS1 | IBS1 (IMPAIRED IN BABA-INDUCED STERILITY 1); ATP binding / kinase/ protein kinase/ protein serine/threonine kinase/ protein tyrosine kinase | chr1:6426898-6430696 REVERSE [PFAM]
Oeu-miR482i	AT1G28050.1	1.5	21.686	499	518	UCUUCAUUCUUCCUCCCUUG	CAAGGGAGGAAAGAUGAAGA	Translation	| Symbols: | zinc finger (B-box type) family protein | chr1:9775528-9777810 REVERSE [PFAM] 1306-1440 PF06203.7 CCT motif;
Oeu-miR482j	AT3G23109.1	1.5	9.712	133	152	UCUUCUUCGUUCUUCUCAUU	GAUGAGAAGAGCGAGGAAGA	Cleavage	| Symbols: | Pseudogene of AT1G61410; tolA protein-related | chr3:8226791-8227003 FORWARD [PFAM]
Oeu-miR529a	AT3G29310.1	2.0	15.352	1495	1514	CUAUUCCCUUUCUCUUCUUC	GGAGAAGAGAAAGGGAAUGU	Cleavage	| Symbols: | calmodulin-binding protein-related | chr3:11249652-11251820 FORWARD [PFAM]
Oeu-miR529b	AT3G29310.1	2.0	15.352	1495	1514	CUAUUCCCUUUCUCUUCUUC	GGAGAAGAGAAAGGGAAUGU	Cleavage	| Symbols: | calmodulin-binding protein-related | chr3:11249652-11251820 FORWARD [PFAM]
Oeu-miR530a	AT3G01390.1	3.0	17.627	278	297	UACAUUAGCACCUGAGUCUC	GAGACUCUGGUGCGAAUGUG	Cleavage	| Symbols: VMA10, AVMA10 | VMA10 (VACUOLAR MEMBRANE ATPASE 10); hydrogen ion transporting ATP synthase, rotational mechanism | chr3:149999-151539 REVERSE [PFAM]
Oeu-miR530b	AT2G32310.2	3.0	9.21	838	857	UGCAUUUGGUGCUGCAACUU	AAGUUGUACCACCAAAUGAA	Cleavage	| Symbols: | unknown protein | chr2:13724394-13726207 FORWARD [PFAM]
Oeu-miR781	AT1G34260.1	2.5	14.124	1919	1939	UUAGAUUUUUCUGGUUAAAUA	UGUUUAAUCAGAGAAAUCUAU	Cleavage	| Symbols: | phosphatidylinositol-4-phosphate 5-kinase family protein | chr1:12485967-12492169 FORWARD [PFAM]
Oeu-miR783a	AT4G02560.1	3.0	16.768	1876	1896	AUGCUUUGUGGGCUCAUUUUC	GAAAAUGAGUUCACAAACUAU	Cleavage	| Symbols: LD | LD (luminidependens); transcription factor | chr4:1123490-1128421 REVERSE [PFAM]
Oeu-miR783b	AT4G07703.1	2.0	13.995	3065	3085	AUGGUUGGCAGCUUCAUUUUC	GAAAAUGGAGUUGCUAAUCAU	Cleavage	| Symbols: | transposable element gene | chr4:4489426-4492749 FORWARD [PFAM]
Oeu-miR813a	AT3G20580.1	2.5	18.889	674	695	GGUUUUUGGCAUGGGUUUGACC	GGUCACGCCCAUGCCAAGAACU	Cleavage	| Symbols: COBL10 | COBL10 (COBRA-LIKE PROTEIN 10 PRECURSOR) | chr3:7187856-7190562 REVERSE [PFAM]
Oeu-miR813b	AT3G20580.1	2.5	18.889	674	695	GGUUUUUGGCAUGGGUUUGACC	GGUCACGCCCAUGCCAAGAACU	Cleavage	| Symbols: COBL10 | COBL10 (COBRA-LIKE PROTEIN 10 PRECURSOR) | chr3:7187856-7190562 REVERSE [PFAM]
Oeu-miR813c	AT3G20580.1	2.5	18.889	674	695	GGUUUUUGGCAUGGGUUUGACC	GGUCACGCCCAUGCCAAGAACU	Cleavage	| Symbols: COBL10 | COBL10 (COBRA-LIKE PROTEIN 10 PRECURSOR) | chr3:7187856-7190562 REVERSE [PFAM]
Oeu-miR816	AT4G26400.1	3.0	13.009	1427	1446	GUCAGAUAUUAUGCUACAAC	UUUGUAUUAUAAUAUCUGAU	Cleavage	| Symbols: | zinc finger (C3HC4-type RING finger) family protein | chr4:13344808-13346597 REVERSE [PFAM]
Oeu-miR825	AT2G01930.1	2.5	16.644	886	906	UUCUCAAGAACUUUUAUGAAC	GUUUAAGAAGGUUCUUGAGAA	Cleavage	| Symbols: BPC1, BBR, ATBPC1 | BPC1 (BASIC PENTACYSTEINE1); DNA binding / specific transcriptional repressor/ transcription factor | chr2:427121-428839 REVERSE [PFAM]
Oeu-miR834	AT4G19390.1	3.0	18.226	93	113	UGGUAGCAUUAGCGUUGGUGA	CCAUUAACGCUAAUGCUAUCG	Cleavage	| Symbols: | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised conserved protein UCP022348 (InterPro:IPR016804), Uncharacterised protein family UPF0114 (InterPro:IPR005134); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G13720.1); Has 553 Blast hits to 553 proteins in 219 species: Archae - 18; Bacteria - 407; Metazoa - 0; Fungi - 0; Plants - 54; Viruses - 0; Other Eukaryotes - 74 (source: NCBI BLink). | chr4:10574768-10576428 REVERSE [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR837a	AT4G01070.1	1.0	2.802	55	75	CCUUUUUUCUUUUUUUUUUUA	CAAAAAAAAAAAGAAAAAAGG	Cleavage	| Symbols: GT72B1, UGT72B1 | GT72B1; UDP-glucosyltransferase/ UDP-glycosyltransferase/ transferase, transferring glycosyl groups | chr4:461592-463449 REVERSE [PFAM]
Oeu-miR837b	AT3G09360.1	2.0	3.888	1843	1863	UUCAUUUUCUUGUUCUUUUCG	GGAAGAGAACAAGGAAAUGAA	Cleavage	| Symbols: | RNA polymerase II transcription factor/ protein binding / transcription activator/ transcription regulator/ translation initiation factor/ zinc ion binding | chr3:2873449-2878591 FORWARD [PFAM]
Oeu-miR837c	AT3G08640.1	2.0	10.56	1272	1292	AACUGAAAAAGAAAAUGAUGA	UUGUUGUUUUCUUUUUCAGUU	Cleavage	| Symbols: | alphavirus core protein family | chr3:2622880-2624308 FORWARD [PFAM]
Oeu-miR838a	AT5G65630.1	0.0	14.489	2259	2279	UUUUCUUCUUCUUCUUCUCCU	AGGAGAAGAAGAAGAAGAAAA	Cleavage	| Symbols: GTE7 | GTE7 (Global transcription factor group E 7); DNA binding | chr5:26225832-26228449 REVERSE [PFAM]
Oeu-miR838b	AT2G01720.1	1.0	16.255	409	429	UUUUCUUCUUCUUCUUGCCAA	CUGGCAAGAAGAAGAAGAAAA	Cleavage	| Symbols: | ribophorin I family protein | chr2:317030-320264 REVERSE [PFAM]
Oeu-miR847	AT1G05894.1	1.5	21.468	94	114	UCAUUCUUCUUCUUCUUCAUG	CAUGAGGAAGAAGAAGAAUGC	Cleavage	| Symbols: | unknown protein | chr1:1784803-1786292 FORWARD [PFAM]
Oeu-miR854	AT1G31810.1	1.0	8.774	2186	2206	GAGGAGGAUGUGGAGGUGGAG	CUCCACCUCCACCUCCUCCUC	Translation	| Symbols: | actin binding | chr1:11399607-11406259 REVERSE [PFAM]
Oeu-miR857	AT2G48140.1	2.0	8.838	712	732	UUUUUUUUGUUGAAGUUGUGU	ACAUGAUUUCAGCAAAAAAAA	Cleavage	| Symbols: EDA4 | EDA4 (embryo sac development arrest 4); lipid binding | chr2:19686409-19687664 FORWARD [PFAM]
Oeu-miR863	AT1G09330.1	2.5	20.262	142	162	UUGAGAGCAAGAAGAAGUAAU	AUUACUUCUUCUUUUUCUCGG	Cleavage	| Symbols: | unknown protein | chr1:3012762-3015091 REVERSE [PFAM]
Oeu-miR948	AT1G59610.1	2.0	21.706	2917	2936	CCAGGCUGUGUUGGAGCUGG	CCAGCUCCGCCACAGUCUGG	Translation	| Symbols: ADL3, CF1 | ADL3 (ARABIDOPSIS DYNAMIN-LIKE 3); GTPase | chr1:21893221-21901086 FORWARD [PFAM]
Oeu-miR950	AT1G36390.1	3.0	23.179	753	772	UCAGCUCCUCGGUGGUUGUG	GAUAAUAACCGAGGAGCUGA	Cleavage	| Symbols: | co-chaperone grpE family protein | chr1:13701668-13703572 REVERSE [PFAM]
Oeu-miR952	AT1G39830.1	2.5	9.679	5851	5870	AAGUGAGAAUGAAAAUGGUG	CAUCGUUUUCAUUUUUAUUU	Cleavage	| Symbols: | transposable element gene | chr1:14840744-14846755 REVERSE [PFAM]
Oeu-miR1023	None								
Oeu-miR1027	AT5G13800.2	1.5	20.216	1186	1206	UUUUUAUCUUCUUUUCCUAUA	UAUGGGGAAAGAAGAUAAAGA	Cleavage	| Symbols: | hydrolase, alpha/beta fold family protein | chr5:4451837-4454576 REVERSE [PFAM]
Oeu-miR1038	None								
Oeu-miR1039	AT2G26910.1	3.0	17.455	2798	2818	UCUUUGGGUCAUUGUGUGCUG	CAGCAGACAUUGACUCAGAGA	Cleavage	| Symbols: PDR4, ATPDR4 | PDR4 (PLEIOTROPIC DRUG RESISTANCE 4); ATPase, coupled to transmembrane movement of substances | chr2:11481623-11488186 FORWARD [PFAM] 2110-2307 PF08370.4 Plant PDR ABC transporter associated;
Oeu-miR1040	None								
Oeu-miR1042	None								
Oeu-miR1044a	AT2G06410.1	3.0	15.682	879	898	UAGUAGAGCAUAUAUGUAUU	AAUAUAUACAUGCUUUGCUG	Cleavage	| Symbols: | transposable element gene | chr2:2531435-2534389 REVERSE [PFAM]
Oeu-miR1044b	AT2G06410.1	3.0	15.682	879	898	UAGUAGAGCAUAUAUGUAUU	AAUAUAUACAUGCUUUGCUG	Cleavage	| Symbols: | transposable element gene | chr2:2531435-2534389 REVERSE [PFAM]
Oeu-miR1052a	AT1G02730.1	3.0	15.544	754	774	UUGCUUUCAUUAAUUGUGAUA	UAUCACAAGUGGUGGAGGCAA	Cleavage	| Symbols: ATCSLD5, CSLD5 | ATCSLD5; 1,4-beta-D-xylan synthase/ cellulose synthase | chr1:594573-598657 REVERSE [PFAM] 1388-3706 PF03552.7 Cellulose synthase;
Oeu-miR1052b	AT1G02730.1	3.0	15.544	754	774	UUGCUUUCAUUAAUUGUGAUA	UAUCACAAGUGGUGGAGGCAA	Cleavage	| Symbols: ATCSLD5, CSLD5 | ATCSLD5; 1,4-beta-D-xylan synthase/ cellulose synthase | chr1:594573-598657 REVERSE [PFAM] 1388-3706 PF03552.7 Cellulose synthase;
Oeu-miR1056	AT1G58330.1	3.0	18.009	409	429	UGGAUCUUUGCAUCUAAUCUC	GAGAUUUGAUGCGAAGGUUCG	Cleavage	| Symbols: ZW2 | ZW2 | chr1:21641130-21642041 FORWARD [PFAM]
Oeu-miR1063	None								
Oeu-miR1067	AT5G04540.1	3.0	15.448	369	389	GAAUACUGAAGUUUGAUAUCA	UGUUAUCAAAUUUCCGUAUUC	Cleavage	| Symbols: | phosphatase/ protein tyrosine phosphatase | chr5:1296442-1302444 REVERSE [PFAM]
Oeu-miR1097	None								
Oeu-miR1122	AT4G03827.1	0.5	19.8	10	29	UACAUACAUACGUAUAUAUA	UGUAUAUACGUAUGUAUGUA	Cleavage	| Symbols: | unknown protein | chr4:1784402-1784530 REVERSE [PFAM]
Oeu-miR1313	AT1G32740.1	2.5	15.676	1003	1023	UACCAAUGAAAUUUCUGUUAG	CAAAUAUAAAUUUCAUUGGUA	Cleavage	| Symbols: | protein binding / zinc ion binding | chr1:11844877-11846347 FORWARD [PFAM]
Oeu-miR1320	AT4G15590.1	2.0	11.843	5814	5834	UGUUAAAUUGAUUUGUUUCAA	UUAAAGCAAAUCAAUUUAAUA	Cleavage	| Symbols: | transposable element gene | chr4:8900245-8907255 REVERSE [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR1428	AT1G60830.1	2.5	18.69	274	295	AAUACACAGCCCCUAUUUGGUG	AACCAAGUAGUGGCUGUGUAUU	Cleavage	| Symbols: | U2 snRNP auxiliary factor large subunit, putative | chr1:22395058-22395603 REVERSE [PFAM]
Oeu-miR1435a	AT2G42060.1	2.5	23.632	812	831	UUUCUGAAGUUCAACUUGUU	AACAGCUUGAGCUUCAGAAG	Cleavage	| Symbols: | CHP-rich zinc finger protein, putative | chr2:17548095-17549069 FORWARD [PFAM]
Oeu-miR1435b	AT2G19910.1	1.5	14.313	2903	2922	UUUCUUGAGGUAAAAUUUUC	GCAAAUUUUACCUUAAGAAA	Cleavage	| Symbols: | RNA-dependent RNA polymerase family protein | chr2:8595820-8600757 REVERSE [PFAM]
Oeu-miR1444	AT2G30570.1	3.0	15.134	588	607	UGCACUUUUGGUUAAUGUUC	GAACAUGAAUCAAAAGUGUU	Cleavage	| Symbols: PSBW | PSBW (PHOTOSYSTEM II REACTION CENTER W) | chr2:13019028-13020194 REVERSE [PFAM]
Oeu-miR1450a	AT4G19770.1	2.5	24.442	122	143	UUCAGUGGCUCGGUGAAGUUGC	GCGAUUUCGCUGAGCUACUGAA	Cleavage	| Symbols: | glycosyl hydrolase family 18 protein | chr4:10753310-10754181 REVERSE [PFAM]
Oeu-miR1450b	AT4G19770.1	2.5	24.442	122	143	UUCAGUGGCUCGGUGAAGUUGC	GCGAUUUCGCUGAGCUACUGAA	Cleavage	| Symbols: | glycosyl hydrolase family 18 protein | chr4:10753310-10754181 REVERSE [PFAM]
Oeu-miR1450c	AT4G19770.1	2.5	24.442	122	143	UUCAGUGGCUCGGUGAAGUUGC	GCGAUUUCGCUGAGCUACUGAA	Cleavage	| Symbols: | glycosyl hydrolase family 18 protein | chr4:10753310-10754181 REVERSE [PFAM]
Oeu-miR1450d	AT4G19770.1	2.5	24.442	122	143	UUCAGUGGCUCGGUGAAGUUGC	GCGAUUUCGCUGAGCUACUGAA	Cleavage	| Symbols: | glycosyl hydrolase family 18 protein | chr4:10753310-10754181 REVERSE [PFAM]
Oeu-miR1450e	AT4G19770.1	2.5	24.442	122	143	UUCAGUGGCUCGGUGAAGUUGC	GCGAUUUCGCUGAGCUACUGAA	Cleavage	| Symbols: | glycosyl hydrolase family 18 protein | chr4:10753310-10754181 REVERSE [PFAM]
Oeu-miR1450f	AT4G19770.1	2.5	24.442	122	143	UUCAGUGGCUCGGUGAAGUUGC	GCGAUUUCGCUGAGCUACUGAA	Cleavage	| Symbols: | glycosyl hydrolase family 18 protein | chr4:10753310-10754181 REVERSE [PFAM]
Oeu-miR1508a	AT3G15020.1	3.0	19.903	697	717	GAGAAAGGGGAGGAGCAGUCG	CGAUUCUUCCUCUCUUUUCUC	Cleavage	| Symbols: | malate dehydrogenase (NAD), mitochondrial, putative | chr3:5056068-5058248 FORWARD [PFAM]
Oeu-miR1508b	AT3G15020.1	3.0	19.903	697	717	GAGAAAGGGGAGGAGCAGUCG	CGAUUCUUCCUCUCUUUUCUC	Cleavage	| Symbols: | malate dehydrogenase (NAD), mitochondrial, putative | chr3:5056068-5058248 FORWARD [PFAM]
Oeu-miR1510	AT4G31050.1	2.5	17.043	969	989	UUUUGUUUUGUCAAUUCCACC	GGUGGUAUUGGUAGAACAAAA	Cleavage	| Symbols: | lipoyltransferase (LIP2p) | chr4:15114250-15115633 FORWARD [PFAM]
Oeu-miR1511	AT5G37800.1	3.0	19.613	442	461	AACCAGUCUCUGUUGCCAUU	CAUGGAAAUGGAGACUGGUU	Cleavage	| Symbols: | basic helix-loop-helix (bHLH) family protein | chr5:15036197-15037574 FORWARD [PFAM]
Oeu-miR1518	AT5G47800.1	3.0	8.011	221	241	UGUGUUGAAAAGUUGAUAUGA	UCAUAUCAACCUUUGAGCACA	Translation	| Symbols: | phototropic-responsive NPH3 family protein | chr5:19353676-19356300 FORWARD [PFAM]
Oeu-miR1521	AT5G12430.1	2.5	9.784	1668	1687	CAGUUAAUGGAGAAGGUUGA	ACAACAUUCUCCGUUAACUG	Cleavage	| Symbols: | DNAJ heat shock N-terminal domain-containing protein | chr5:4027979-4034086 REVERSE [PFAM]
Oeu-miR1523	AT1G53020.1	2.5	20.349	676	695	AUGGGAGAAUUCUGAGCUGA	UCAGUUCAGGAUUCUUCCAA	Cleavage	| Symbols: UBC26, PFU3 | UBC26 (UBIQUITIN-CONJUGATING ENZYME 26); ubiquitin-protein ligase | chr1:19751372-19759474 REVERSE [PFAM]
Oeu-miR1527	AT4G06628.1	3.0	19.619	2467	2486	UAAUUCAGCCUUAGAAAAGC	UCUUUUCUAAGGUUGAUUUA	Cleavage	| Symbols: | transposable element gene | chr4:3749370-3753281 REVERSE [PFAM]
Oeu-miR1533a	AT3G16240.1	1.5	12.129	972	990	AGGAUAAAAAUUAUAACGA	UCGUUGUAAUUUUUAUCCA	Cleavage	| Symbols: DELTA-TIP, TIP2;1, DELTA-TIP1, AQP1, ATTIP2;1 | DELTA-TIP; ammonia transporter/ methylammonium transmembrane transporter/ water channel | chr3:5505404-5507050 FORWARD [PFAM]
Oeu-miR1533b	AT1G53165.1	1.0	1.902	2576	2594	AGAAAAAGAAAAAUAAUGA	UCAUUAUUUUUGUUUUUCU	Cleavage	| Symbols: ATMAP4K ALPHA1 | ATMAP4K ALPHA1; ATP binding / kinase/ protein kinase/ protein serine/threonine kinase/ protein tyrosine kinase | chr1:19813804-19819564 FORWARD [PFAM]
Oeu-miR1533c	AT2G45050.1	1.5	7.605	910	928	AAAACAAAAAGAAUAGUGA	UCACUCUUUUUUUUGUUUU	Cleavage	| Symbols: | zinc finger (GATA type) family protein | chr2:18582872-18584083 FORWARD [PFAM]
Oeu-miR1533d	AT2G47730.1	2.0	8.334	1631	1649	AUGAUAUAAAUUAUAUUGA	UCAGUAUUAUUUAUGUCAU	Cleavage	| Symbols: ATGSTF8, ATGSTF5, GST6, GSTF8 | ATGSTF8 (ARABIDOPSIS THALIANA GLUTATHIONE S-TRANSFERASE PHI 8); glutathione binding / glutathione transferase | chr2:19556326-19559463 FORWARD [PFAM]
Oeu-miR1533e	AT5G49460.1	1.5	2.823	2207	2225	AUAAUUGAAAAAAUAAGGA	UUCUUCUUUUUUCAAUUAU	Cleavage	| Symbols: ACLB-2 | ACLB-2 (ATP CITRATE LYASE SUBUNIT B 2); ATP citrate synthase | chr5:20054523-20058464 FORWARD [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR1533f	AT1G08160.1	1.5	1.298	890	908	AGAAUAAAAAAUAUAAUUA	UUAUUAUAUUUUUUAUUUU	Cleavage	| Symbols: | harpin-induced protein-related / HIN1-related / harpin-responsive protein-related | chr1:2559487-2560403 REVERSE [PFAM]
Oeu-miR1533g	AT3G05930.1	1.5	12.961	792	810	AUAAUAAAGAAAAUAAUGG	UCAUUCUUUUCUUUAUUAU	Cleavage	| Symbols: GLP8 | GLP8 (GERMIN-LIKE PROTEIN 8); manganese ion binding / nutrient reservoir | chr3:1770337-1771409 FORWARD [PFAM]
Oeu-miR1533h	AT5G23270.1	1.5	16.789	100	118	AUAAUGAAAAUAAAAAAAA	UUUUUUUUGUUUUCAUUAC	Cleavage	| Symbols: STP11, ATSTP11 | STP11 (SUGAR TRANSPORTER 11); carbohydrate transmembrane transporter/ sugar:hydrogen symporter | chr5:7838906-7841026 FORWARD [PFAM]
Oeu-miR1535a	AT1G05640.1	2.0	14.499	984	1002	UUUGUUUGUGGUGGUGUGG	GCACACUGCCACAAACAAA	Cleavage	| Symbols: | ankyrin repeat family protein | chr1:1687436-1689501 REVERSE [PFAM]
Oeu-miR1535b	AT2G30362.1	1.5	5.445	29	47	UUUGUUUGUGGUGGUGUAG	CUCCACCACCACAAGCAAA	Cleavage	| Symbols: | other RNA | chr2:12937351-12939638 FORWARD [PFAM]
Oeu-miR1535c	AT4G03935.1	2.5	10.645	82	100	AUUGAUUGUGGGGAUGACU	AGUCAUCUUCACCAUCAAU	Cleavage	| Symbols: | other RNA | chr4:1869052-1874687 REVERSE [PFAM]
Oeu-miR1535d	AT1G78550.1	2.0	16.353	952	970	CCUGUUUGCUGUGAUCUCU	AGAGAUCAUAACAAACGGG	Translation	| Symbols: | oxidoreductase, 2OG-Fe(II) oxygenase family protein | chr1:29544927-29546487 REVERSE [PFAM]
Oeu-miR1852	AT5G35930.1	3.0	16.66	1014	1034	AGAUGGAUUCAGGAAGCAUGU	ACAUGUUUUCUGAAUCUAUAU	Cleavage	| Symbols: | AMP-dependent synthetase and ligase family protein | chr5:14066992-14074751 REVERSE [PFAM]
Oeu-miR1862a	AT5G41980.1	2.5	15.269	1255	1274	UGAAGGUUGGUGUAUUUUGG	CCAAAAUAUGUCAACCUUCU	Cleavage	| Symbols: | unknown protein | chr5:16793614-16795006 FORWARD [PFAM]
Oeu-miR1862b	AT5G44190.1	2.5	21.298	1363	1382	UUGAGGUUGUUUAAUUUUGA	UCGAAGUUAAGCAACUUUAA	Cleavage	| Symbols: GLK2, ATGLK2, GPRI2 | GLK2 (GOLDEN2-LIKE 2); DNA binding / transcription factor/ transcription regulator | chr5:17798369-17800824 FORWARD [PFAM]
Oeu-miR1888	AT4G35000.1	2.5	15.627	148	168	UUGUUUAAGAUUUGUGAAGAU	UUCUUCACAAAUCCUAAACGA	Cleavage	| Symbols: APX3 | APX3 (ASCORBATE PEROXIDASE 3); L-ascorbate peroxidase | chr4:16664822-16667718 REVERSE [PFAM]
Oeu-miR2083	AT4G02010.1	2.5	19.822	454	474	ACAUUGACGAGUGAAACCAGA	UGUGGUUUCUCUCGUCAGUGU	Cleavage	| Symbols: | protein kinase family protein | chr4:881090-885399 FORWARD [PFAM]
Oeu-miR2093a	AT1G65320.1	1.5	16.923	271	290	UUGCUUUAAUUUGAAGAAGA	UCUUCUUCAAAUUGAAGUAG	Cleavage	| Symbols: | CBS domain-containing protein | chr1:24259888-24262951 REVERSE [PFAM]
Oeu-miR2093b	AT3G49800.1	3.0	12.252	1551	1570	GUGCAUUUCUUGGAAGGGCA	UGCCUUUUCAAAAGAUGCAU	Translation	| Symbols: | BSD domain-containing protein | chr3:18471625-18473743 REVERSE [PFAM]
Oeu-miR2592	AT1G70060.1	2.0	17.804	1940	1960	GAUUCUUUGAGUCAUGUUGUU	AACAGCAGGAUUCAAAGAAUC	Cleavage	| Symbols: SNL4 | SNL4 (SIN3-LIKE 4) | chr1:26383789-26389916 FORWARD [PFAM] 1393-1665 PF08295.5 Histone deacetylase (HDAC) interacting;
Oeu-miR2595a	AT5G67540.2	2.5	8.755	360	380	UUUAUUUUCCUUUUUAUGUCU	AGACAUAAAGAGGAUAAUGAA	Cleavage	| Symbols: | glycosyl hydrolase family protein 43 | chr5:26944028-26946365 REVERSE [PFAM]
Oeu-miR2595b	AT4G00891.1	2.5	11.076	127	147	UACAUUUUCUUGUAUAUCUCU	AAAGAUGUACAAGAAAAUGUC	Cleavage	| Symbols: | pseudogene of APUM24 (RNA binding) | chr4:377481-377908 REVERSE [PFAM]
Oeu-miR2612	AT5G23405.1	3.0	22.994	723	743	UGUUGGUGUCAACUCGUACGG	CCAUACGAGUUGACAUUAGUA	Cleavage	| Symbols: | high mobility group (HMG1/2) family protein | chr5:7882848-7884353 REVERSE [PFAM]
Oeu-miR2628	AT1G47380.1	1.0	6.987	1827	1846	CAAGAAAAAAUGAAGAGGAA	UUCCUUUUCAUUUUUUUUUG	Cleavage	| Symbols: | protein phosphatase 2C-related / PP2C-related | chr1:17372537-17376075 REVERSE [PFAM]
Oeu-miR2630	AT2G23360.1	3.0	16.896	440	460	UAGUUUUGUUCCUUGCUAAUU	UGUUAGCAAAGAACAAAACUG	Cleavage	| Symbols: | transport protein-related | chr2:9949456-9952845 FORWARD [PFAM]
Oeu-miR2633	AT4G07454.1	2.5	14.183	48	68	UGAAAUUUUGCUAAAGAUUUA	UGAAUUGUUAGCAGAAUUUCA	Cleavage	| Symbols: | transposable element gene | chr4:4235774-4238371 REVERSE [PFAM]
Oeu-miR2641	AT5G53350.1	2.5	16.911	650	669	GUUCGAACCUUUACAUGUAU	AGAUAUGUAAGGGUUUGAAC	Cleavage	| Symbols: CLPX | CLPX; ATP binding / ATPase/ nucleoside-triphosphatase/ nucleotide binding / protein binding | chr5:21643898-21647707 FORWARD [PFAM]
Oeu-miR2651	AT5G35520.1	2.5	12.361	124	144	UUUGAGAGGUAUGGGUGCAUU	GAUGCAUCCAAAUCUCUCAAA	Translation	| Symbols: MIS12, ATMIS12 | MIS12 (MINICHROMOSOME INSTABILITY 12 (MIS12)-LIKE) | chr5:13701396-13703489 FORWARD [PFAM]
Oeu-miR2654	AT4G07740.1	2.0	20.768	528	547	AUUCAGAGAGAAAGUAUCCG	CGGAUACUUUCGCUCUGAAA	Translation	| Symbols: | unknown protein | chr4:4539638-4540682 FORWARD [PFAM]
Oeu-miR2657	AT3G30405.1	2.5	16.639	2166	2187	UGGUUUUUCUUCUAUUUUGUUG	CAAUGAAGUAGAAGAAGAAUCA	Cleavage	| Symbols: | transposable element gene | chr3:12019426-12025026 FORWARD [PFAM]
Oeu-miR2868	AT3G06340.2	1.5	19.305	1011	1030	UUUGUGUUUUGUAGUAGAAC	GUUCUACUGCAAAACACAAU	Cleavage	| Symbols: | DNAJ heat shock N-terminal domain-containing protein | chr3:1920207-1923311 REVERSE [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR2873	AT5G10490.1	3.0	19.965	1229	1249	GUGGACUUGUGGUUUGAUAUG	CAUCCCAAACCAUAAGUUCAC	Cleavage	| Symbols: MSL2 | MSL2 (MscS-LIKE 2); ion channel | chr5:3300225-3304761 REVERSE [PFAM]
Oeu-miR2911a	AT4G20700.1	3.0	23.008	147	166	CGGCCGGGGACGGACUGGGA	UCUCGGCUCGUCCCCGGUCG	Cleavage	| Symbols: | CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1204 (InterPro:IPR009596); BEST Arabidopsis thaliana protein match is: RNA binding / RNA-directed DNA polymerase (TAIR:AT4G20520.1); Has 123 Blast hits to 123 proteins in 28 species: Archae - 2; Bacteria - 16; Metazoa - 37; Fungi - 4; Plants - 62; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). | chr4:11098270-11100055 FORWARD [PFAM] 907-1578 PF06721.4 Protein of unknown function (DUF1204);
Oeu-miR2911b	None								
Oeu-miR2913	AT4G25770.1	3.0	9.052	157	177	CAGGUCGGGGAUUGAAAGUUG	CAAAUCUCAAUCCCCGAUUUG	Cleavage	| Symbols: | unknown protein | chr4:13117525-13120114 REVERSE [PFAM]
Oeu-miR2919	AT1G15420.1	2.0	6.771	55	73	GCGGGGGGGGAGGGAAGGA	UCUUUCUCUCUCUCCCCGC	Cleavage	| Symbols: | unknown protein | chr1:5301613-5303390 REVERSE [PFAM]
Oeu-miR2923	AT2G26720.1	1.5	7.013	592	613	AGACAAAAAAAUAAAAUAAAAAA	UUUUUUAUUUUAUUUUUUAUCU	Cleavage	| Symbols: | plastocyanin-like domain-containing protein / mavicyanin, putative | chr2:11384782-11385402 FORWARD [PFAM]
Oeu-miR2925a	None								
Oeu-miR2925b	AT2G29760.1	3.0	17.244	1526	1544	GGGCGGCCGGGAGCAUCGU	UUGAUGUUCUCGGCCGCUC	Cleavage	| Symbols: | pentatricopeptide (PPR) repeat-containing protein | chr2:12712884-12715852 FORWARD [PFAM]
Oeu-miR2925c	AT3G21295.1	3.0	17.644	47	65	GGGCUGCCUCCGGCUUCGU	UUGAAGCCGGAGGAAGCCC	Cleavage	| Symbols: | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: PWWP (InterPro:IPR000313); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G51745.1); Has 428 Blast hits to 314 proteins in 77 species: Archae - 0; Bacteria - 111; Metazoa - 87; Fungi - 55; Plants - 63; Viruses - 0; Other Eukaryotes - 112 (source: NCBI BLink). | chr3:7489236-7492602 FORWARD [PFAM]
Oeu-miR2925d	AT3G45020.1	2.0	24.568	236	254	UGGUGGCCGCCGCCUUAGU	GCUAAGGCGGCGGCUAUCG	Cleavage	| Symbols: | 50S ribosomal protein-related | chr3:16468961-16470153 REVERSE [PFAM]
Oeu-miR2925e	AT3G45020.1	2.0	24.568	236	254	UGGUGGCCGCCGCCUUAGU	GCUAAGGCGGCGGCUAUCG	Cleavage	| Symbols: | 50S ribosomal protein-related | chr3:16468961-16470153 REVERSE [PFAM]
Oeu-miR2928a	AT1G26330.1	2.0	15.581	3060	3079	AAGAUGAGGACAUUUUGGUC	UACCAAAAUGUCCUCAUCUC	Cleavage	| Symbols: | DNA binding | chr1:9107519-9113724 FORWARD [PFAM]
Oeu-miR2928b	AT1G26330.1	2.0	15.581	3060	3079	AAGAUGAGGACAUUUUGGUC	UACCAAAAUGUCCUCAUCUC	Cleavage	| Symbols: | DNA binding | chr1:9107519-9113724 FORWARD [PFAM]
Oeu-miR2928c	AT5G57130.1	3.0	14.64	1499	1518	GAGAUGACGAGAUUUUAUUG	CAAUGAGAUUUCUUCAUCUU	Cleavage	| Symbols: | protein binding | chr5:23144980-23149568 FORWARD [PFAM]
Oeu-miR2937a	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937b	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937c	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937d	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937e	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937f	AT3G29200.1	3.0	3.293	57	77	AUGAGAGGUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937g	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937h	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]
Oeu-miR2937i	AT3G29200.1	2.5	3.293	57	77	AUGAGAGAUGUUGAAGGGGGC	AUCCCUUUUAACAUCUCUCAU	Cleavage	| Symbols: CM1, ATCM1 | CM1 (CHORISMATE MUTASE 1); L-ascorbate peroxidase/ chorismate mutase | chr3:11164393-11166416 REVERSE [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR2950a	AT3G05520.1	3.0	15.585	1051	1071	AUCCAUCUCUUGCAUAAUGGA	UCCAUUAUGUAAGUGAAGGAU	Cleavage	| Symbols: | F-actin capping protein alpha subunit family protein | chr3:1598547-1601310 FORWARD [PFAM]
Oeu-miR2950b	None								
Oeu-miR3434a	AT4G19000.1	3.0	21.713	811	830	AAUUGGUUGAUUGUAUCAUU	AAUGAUAUAAUCAACAAAUG	Cleavage	| Symbols: | IWS1 C-terminus family protein | chr4:10405503-10407339 FORWARD [PFAM]
Oeu-miR3434b	AT3G14980.1	2.5	14.243	62	81	UCUUGGCUUUUUCUAUGAUU	AAUUAUAGAAAAAGCUAAAA	Cleavage	| Symbols: | PHD finger transcription factor, putative | chr3:5039931-5045403 REVERSE [PFAM]
Oeu-miR3438	AT2G42740.1	2.0	18.125	70	90	UCGAAAGCUUCUUCUCUGACG	CGUCGGAGAAGAAGCUAUCGA	Cleavage	| Symbols: RPL16A | RPL16A; structural constituent of ribosome | chr2:17791729-17793082 FORWARD [PFAM]
Oeu-miR3447	AT2G31400.1	2.5	15.587	3006	3026	UUUGAGUUGGUUUGUUACAAA	UUUGUAACAAAGCAACUCGAU	Translation	| Symbols: | pentatricopeptide (PPR) repeat-containing protein | chr2:13387004-13390706 REVERSE [PFAM]
Oeu-miR3476	AT2G32990.1	3.0	16.587	714	734	UGAUGUGGGUUUGUUGGAUGA	CCGUUCAACAAAUCCACAUUA	Cleavage	| Symbols: AtGH9B8 | AtGH9B8 (Arabidopsis thaliana glycosyl hydrolase 9B8); catalytic/ hydrolase, hydrolyzing O-glycosyl compounds | chr2:14003250-14006017 FORWARD [PFAM]
Oeu-miR3511	None								
Oeu-miR3513	AT4G10540.1	3.0	16.007	420	440	UGAAUUUCUGAGUUUGCAAUG	CAUUGUAGACUCAGGAGUUUG	Cleavage	| Symbols: | subtilase family protein | chr4:6512515-6515743 REVERSE [PFAM]
Oeu-miR3706a	AT1G03060.1	3.0	19.054	4635	4654	UUUCGGAGAGAUGGCUGAGU	AUUUAGCCAUCUCUCUGAGC	Cleavage	| Symbols: | WD-40 repeat family protein / beige-related | chr1:712473-726891 REVERSE [PFAM]
Oeu-miR3706b	AT1G03060.1	3.0	19.054	4635	4654	UUUCGGAGAGAUGGCUGAGU	AUUUAGCCAUCUCUCUGAGC	Cleavage	| Symbols: | WD-40 repeat family protein / beige-related | chr1:712473-726891 REVERSE [PFAM]
Oeu-miR3706c	AT1G36225.1	2.0	13.465	929	948	UUGUGGAAAAAUGGAGAAGA	UCUUCUCCAUUCUUCCACAU	Translation	| Symbols: | transposable element gene | chr1:13606640-13607930 FORWARD [PFAM]
Oeu-miR3712	None								
Oeu-miR3950	AT2G27285.1	2.5	10.715	714	734	UUUUUUGGUAACAUUAUUUCU	AGAGAUGAUGUUACUAAGAAG	Cleavage	| Symbols: | unknown protein | chr2:11675960-11678118 REVERSE [PFAM] 375-746 PF09745.2 Coiled-coil domain-containing protein 55 (DUF2040);
Oeu-miR3953	AT5G26160.1	3.0	13.602	2162	2182	UUCAGUUCCGGAAGCCGUAGA	UUUGCAGCUUCUGGAAUUGAA	Cleavage	| Symbols: | unknown protein | chr5:9142865-9146525 FORWARD [PFAM]
Oeu-miR4221	AT1G23550.1	3.0	4.129	1081	1102	UUUUUUUCUGUUUAAUUCGUGC	UCACGAAUUAAAAAAAAAAAAA	Translation	| Symbols: SRO2 | SRO2 (SIMILAR TO RCD ONE 2); NAD+ ADP-ribosyltransferase | chr1:8350793-8352791 FORWARD [PFAM]
Oeu-miR4350a	AT5G16180.1	3.0	18.166	2008	2029	UCAAGGGAUUUUGUGUCAUUGG	CUGAUGGCACAGAAUCUUUUGA	Cleavage	| Symbols: CRS1, ATCRS1 | CRS1 (ARABIDOPSIS ORTHOLOG OF MAIZE CHLOROPLAST SPLICING FACTOR CRS1); RNA splicing factor, transesterification mechanism | chr5:5279884-5282898 FORWARD [PFAM]
Oeu-miR4350b	None								
Oeu-miR5021a	AT4G36890.1	0.5	4.249	338	357	AGAGAAGAGGAGGAAGAAGA	UCUUCUUCUUCCUCUUCUCU	Cleavage	| Symbols: IRX14 | IRX14 (irregular xylem 14); transferase, transferring glycosyl groups / xylosyltransferase | chr4:17379175-17381780 REVERSE [PFAM]
Oeu-miR5021b	AT5G59510.1	2.0	7.702	492	511	CUAGAAAAAUAAGAAGAAAA	UUUUUUUCUUAUUUUUUUAC	Cleavage	| Symbols: RTFL5, DVL18 | RTFL5 (ROTUNDIFOLIA LIKE 5) | chr5:23990074-23990696 FORWARD [PFAM]
Oeu-miR5021c	AT3G10380.1	1.5	23.214	79	98	AGAGAAGACGAAGAAGAAAG	UUUUCUUCUUCGUCUUCUCC	Cleavage	| Symbols: SEC8, ATSEC8 | SEC8 (SUBUNIT OF EXOCYST COMPLEX 8) | chr3:3219523-3228933 REVERSE [PFAM]
Oeu-miR5021d	AT1G09330.1	2.5	18.848	142	161	UGAGAGCAAGAAGAAGUAAU	AUUACUUCUUCUUUUUCUCG	Cleavage	| Symbols: | unknown protein | chr1:3012762-3015091 REVERSE [PFAM]
Oeu-miR5021e	AT1G16850.1	0.5	11.822	92	111	UGACACAAAGAGGAAGAAAA	UUUUCUUCCUCUUUGUGUUA	Cleavage	| Symbols: | unknown protein | chr1:5764807-5765651 REVERSE [PFAM]
Oeu-miR5021f	AT2G24130.1	0.0	10.613	118	137	GAAGAAGAAGAAGAAGAAAU	AUUUCUUCUUCUUCUUCUUC	Cleavage	| Symbols: | leucine-rich repeat transmembrane protein kinase, putative | chr2:10258148-10261220 FORWARD [PFAM]
Oeu-miR5138	AT3G17290.1	3.0	15.325	536	553	GAAAAUCAUCAGGCGCUC	AAGCGUUUGAUGAUUUUA	Cleavage	| Symbols: | transposable element gene | chr3:5905052-5906716 REVERSE [PFAM]
Oeu-miR5140a	AT4G12740.1	2.0	12.67	1380	1397	GCUCUUGAAGCUAUGGUG	CACUAUAGUUUCAAGAGA	Cleavage	| Symbols: | adenine-DNA glycosylase-related / MYH-related | chr4:7494578-7497596 REVERSE [PFAM]
Oeu-miR5140b	AT2G41740.1	0.5	23.872	2176	2193	GCUGGUGAAAUUUUGUUU	AAACAGAAUUUCACCAGC	Cleavage	| Symbols: VLN2, ATVLN2 | VLN2 (VILLIN 2); actin binding | chr2:17410715-17417740 REVERSE [PFAM]

Table 3 continue

miRNAAcc.	Target_Acc.	Exp.	UPE	Target start	Target end	miRNA_aligned_fragment	Target_aligned_fragment	Inhibition	Target_Desc.
Oeu-miR5140c	AT3G13170.1	1.5	11.035	789	806	GCUCGUAAUGAUUUGGUA	UAUUGAAUCAUUACGAGC	Cleavage	| Symbols: ATSPO11-1 | ATSPO11-1; ATP binding / DNA binding / DNA topoisomerase (ATP-hydrolyzing)/ catalytic | chr3:4231560-4234192 REVERSE [PFAM]
Oeu-miR5140d	AT3G47500.1	0.5	22.902	1008	1025	GCGGGUGAAAACCUGGUG	CACCAGGUUUUUACCCGC	Cleavage	| Symbols: CDF3 | CDF3 (CYCLING DOF FACTOR 3); DNA binding / protein binding / transcription factor | chr3:17504000-17506058 REVERSE [PFAM] 407-595 PF02701.8 Dof domain, zinc finger;
Oeu-miR5140e	AT1G35230.1	2.0	16.697	248	265	GCUGCUGGAGAUUCAGUG	UACUGAAUCUCCACCAGC	Cleavage	| Symbols: AGP5 | AGP5 (ARABINOGALACTAN-PROTEIN 5) | chr1:12917149-12917763 FORWARD [PFAM]
Oeu-miR5140f	AT2G45245.1	1.5	14.866	156	173	GUUGUUGAAGAUAUAGUG	CCCUGUAUCUUCAACAAC	Cleavage	| Symbols: | other RNA | chr2:18659422-18660655 REVERSE [PFAM]
Oeu-miR5140g	AT5G65540.1	1.0	18.219	1418	1435	GUUGCUGAAGAUUUUGGG	CCCAAAAUUUUCAGCAGC	Cleavage	| Symbols: | unknown protein | chr5:26195689-26198322 FORWARD [PFAM]
Oeu-miR5140h	AT5G10278.1	2.0	23.833	598	615	GCUGGGGAGGAGCUGGUG	AACCAGCUCCUCCCUAGU	Cleavage	| Symbols: | other RNA | chr5:3228286-3232723 FORWARD [PFAM]
Oeu-miR5142a	AT1G12700.1	1.5	14.674	550	568	CAAUUGAAUUAUAAGUGAU	GUUACUUAUAAUUCGAUUG	Cleavage	| Symbols: | helicase domain-containing protein / pentatricopeptide (PPR) repeat-containing protein | chr1:4322913-4326197 REVERSE [PFAM]
Oeu-miR5142b	AT3G59520.1	3.0	15.718	563	581	CUAUCAAUUGACAAGUGAU	AUCAUUUGUUGAUUGCUAG	Cleavage	| Symbols: ATRBL13 | ATRBL13 (ARABIDOPSIS RHOMBOID-LIKE PROTEIN 13) | chr3:21991826-21993263 FORWARD [PFAM]
Oeu-miR5202	None								
Oeu-miR5301	AT5G45770.1	3.0	8.452	566	586	UGUGAUUUGGGUUGAAAGAUU	AAUCUUUCCACUCAAAUCUCA	Cleavage	| Symbols: AtRLP55 | AtRLP55 (Receptor Like Protein 55); protein binding | chr5:18563568-18564921 FORWARD [PFAM]

Discussion

Olive (*Olea europaea* L.) is a very old crop plant with a great economic impact in several countries of the Mediterranean basin and, more recently, widespread in many parts of the word, including Australia and USA. The main drawbacks that occur in olive trees and that result in yield loss are female sterility, male sterility and self-incompatibility (SI). The main goal of the thesis was to shed light on SI in olive and, to the best of our knowledge, this work represents the first deep study on the matter. The information about self-incompatibility in olive currently available in literature is scanty and often very confused. In fact, the classification of different cultivars as self-compatible or self-incompatible is quite contradictory and different authors have classified the same cultivar either as self-compatible or self-incompatible. However, it is generally accepted the classification of olive as self-incompatible species, even though some cultivars are not affected by incompatibility barriers. In particular, we focalized our attention on two cultivars, Leccino and Frantoio, which are among the most widespread cultivars in Italy. The cv. Leccino is classified as self-incompatible cultivar, whereas the cv. Frantoio as self-compatible. In order to get a more robust set of data, we further analyzed two self-incompatible Italian cultivars: Dolce Agogia and Moraiolo.

Olive is currently classified as a gametophytic self-incompatible (GSI) species, but the classification is based on morphological traits shared with taxa manifesting GSI (*e.g. Rosaceae*, *Solanaceae* and *Plantaginaceae*), such as wet-type stigma and bi-nucleate pollen. However, neither cytological nor molecular evidences supporting or confirming the occurrence of a GSI system in olive are currently available in literature.

On the basis of the classification of olive as GSI species, several attempts were performed in order to isolate genes belonging to RNase-based GSI, such *S-RNase* and *SLF/SFB*, the female and male determinants, respectively. In spite of several attempts, no sequences

having reference with the searched genes could be isolated. At the same time, we tested by means of aniline blue staining the germination rate and growth behavior of more than 34,000 pollen grains and pollen tubes collected from about 300 flowers from 4 different cultivars, three self-incompatible and one self-compatible. Results after self-pollination clearly showed that no pollen tubes were visualized through pistils of self-incompatible cultivars, whereas several pollen tubes were marked in pistils of the self-compatible cultivar (**Figure 1.27**). By contrast, after cross-pollination, no significant differences were visualized between self-compatible and self-incompatible cultivars. Moreover, molecular evidences, *i.e.* no genes having reference with RNase-based GSI genes isolated, and cytological results, *i.e.* no pollen tubes growing into the pistils of self-incompatible cultivars after self-pollination, allowed us to demonstrate that RNase-based GSI is not the actual SI system adopted by olive trees. It is worth noting that RNase-based GSI is characterized by an incompatible reaction that takes place at different levels of the style.

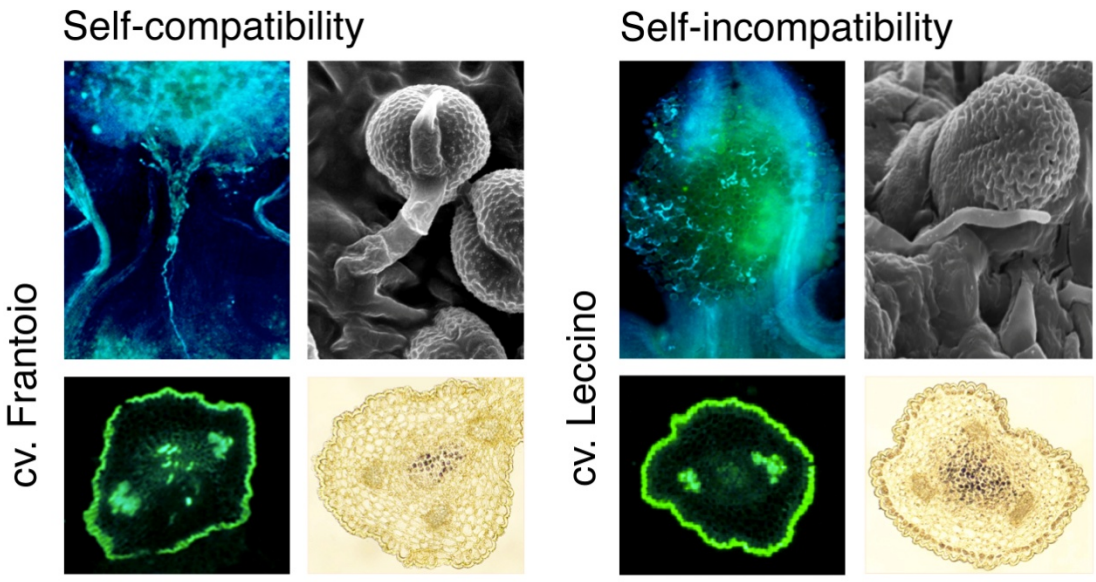


Figure 1.27. Summary results emerged from the cytological investigations that support a SSI system in olive.

Detailed cytological analyses were useful to demonstrate that in olive the rejection of the self-pollen occurs at the stigma level. According to the currently known SI systems, this observation opens two possibilities for SI in olive, which are both characterized by stigmatic self-pollen rejection: self-incompatibility similar to *Papaveraceae*-type GSI or to SSI of *Brassica*.

We isolated a putative ortholog of *PrsS*, female determinant in *Papaveraceae*-type GSI, which exhibits the four conserved cysteine residues as reported in Higashiyama (2010). Several amplification reactions performed by using three different primer combinations enable to recover amplicons from olive genomic DNAs, but not from olive flower cDNAs. This finding proved that the *OePrsS-like* gene is not expressed in flowers and thus the *Papaveraceae*-type GSI cannot be taken into account in olive (**Figure 1.28**).

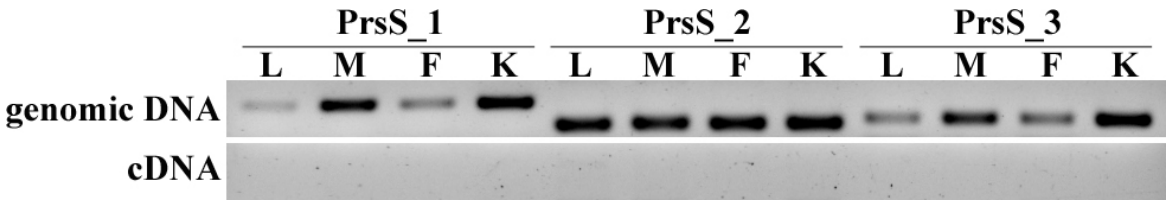


Figure 1.28. Amplification of the putative ortholog of *PrsS* by three different primer combinations (PrsS_1, PrsS_2 and PrsS_3). L: cv. Leccino (SI). M: cv. Moraiolo (SI). F: cv. Frantoio (SC). K: cv. Kalamata (SC).

On the basis of our preliminary cytological and molecular findings, we started to investigate SSI in olive and several experimental results were collected to support this new hypothesis.

The search for genes involved in SSI of *Brassica* allowed us to isolate two genes having reference with *SRK*, the candidate as female determinant, one gene for *SCR*, the male determinant, and two genes for *SLG*, the enhancer of the incompatible response in some *Brassica* lines, which were named *OeSRK*, *OeSRK-like*, *OeSCR-like_12*, *OeSLG* and *OeSLG-like*, respectively.

Gene expression analyses performed in pistils and anthers of both cultivars, Leccino and Frantoio, allowed us to note that *OeSRK* was not expressed in later stages of pistils of cv. Frantoio. Taking into account the crucial role played by *SRK* in the SSI of *Brassica*, the specific or preferential expression of *OeSRK* in pistils of the self-incompatible cv. Leccino along with the lack or low expression of *OeSRK* in pistils of the self-compatible cv. Frantoio, strongly support the hypothesis that a SSI similar to that of *Brassica* may occur also in *Olea*. However, even if the gene was preferentially expressed in pistils, qRT-PCR analyses also showed that the expression of *OeSRK* was not pistil-specific, as instead expected by a candidate gene for the female determinant. In fact, *OeSRK* expression was also found in other tissues such as fruits and, in low amounts, in anthers and roots. It is worth mentioning that *OeSRK* belongs to the receptor protein kinase family, which includes several proteins involved in a large number of pathways. Further analyses aimed at isolating the full-length of *OeSRK* in pistils and anthers revealed that the entire gene was present exclusively in pistils (**Figure 1.29**). These findings open at least two different scenarios. Considering that the average size of amplicons used for qRT-PCR ranges between 100-200 bp and that *OeSRK* belongs to a wide protein family, signal detection of *OeSRK* by means of qRT-PCR in tissues such as anthers, fruits and roots may be due to an amplification of unspecific conserved sequences, likely belonging to the same receptor protein kinase family. The other possibility should be a post-transcriptional regulation of *OeSRK* occurring in all tissues but not in pistils.

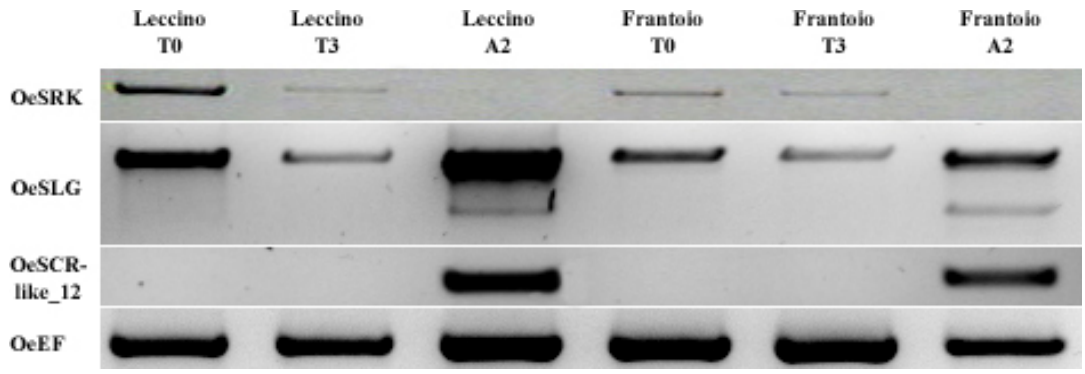


Figure 1.29. Full-length amplification of SI genes, *OeSRK*, *OeSLG* and *OeSCR-like_12*, in pistils and anthers in the self-incompatible cultivar Leccino and in the self-compatible cultivar Frantoio. T0: pistils during the flowering day. T3: pistils three days after flowering. A2: dehiscent anthers. *OeEF*: Elongation Factor gene used as internal control.

After several attempts aimed at isolating an ortholog of *SCR* of *Brassica*, the best candidate was *OeSCR-like_12*. The nucleotide sequence similarity between *SCR* of *Brassica* and *OeSCR-like_12* was low but both of them showed the conserved cysteine-rich pattern and the expected small size (<10 kDa), which are crucial features for those proteins involved as ligand in pollen-pistil interactions. Furthermore, both the gene expression analysis and the isolation of its full-length revealed a strong anther-specificity, as expected by a candidate gene for the male determinant (**Figure 1.29**).

The two genes having reference with *SLG* of *Brassica*, *OeSLG* and *OeSLG-like*, were differing each other for few SNPs and was not possible to discriminate between them by conventional PCR-based amplification techniques. Moreover, qRT-PCR analyses showed a preferential, but not exclusive, expression in pistils rather than in other tissues. Intriguingly, tests aimed at isolating full-lengths in pistils and anthers revealed a single band in pistils and two bands in anthers (**Figure 1.29**). We hypothesize that the lower band obtained in anthers is a splice variant of the gene but further analyses are needed to find out the real nature of that transcript and whether it may be involved in the regulation of *OeSLG* in anthers.

We also tested the localization of SI genes in pistils by means of *in situ* hybridizations. Analyses revealed that *OeSRK* and *OeSLG* were preferentially expressed on the external surface of papillar cells, whereas no expression in the same cells was found for *OeSRK-like* and *OeSCR-like_12*. The localization of the expression of *OeSRK* and *OeSLG* was in agreement with that expected from genes involved in self-incompatibility reaction, since papillar cells are the first cells to interact with landed pollen grains (**Figure 1.30**). Furthermore, since *OeSCR-like_12* was expected to be an anther-specific gene, the lack of its signal in papillar cells was a further confirm of the reliability of the localization of *OeSRK* and *OeSLG* genes.

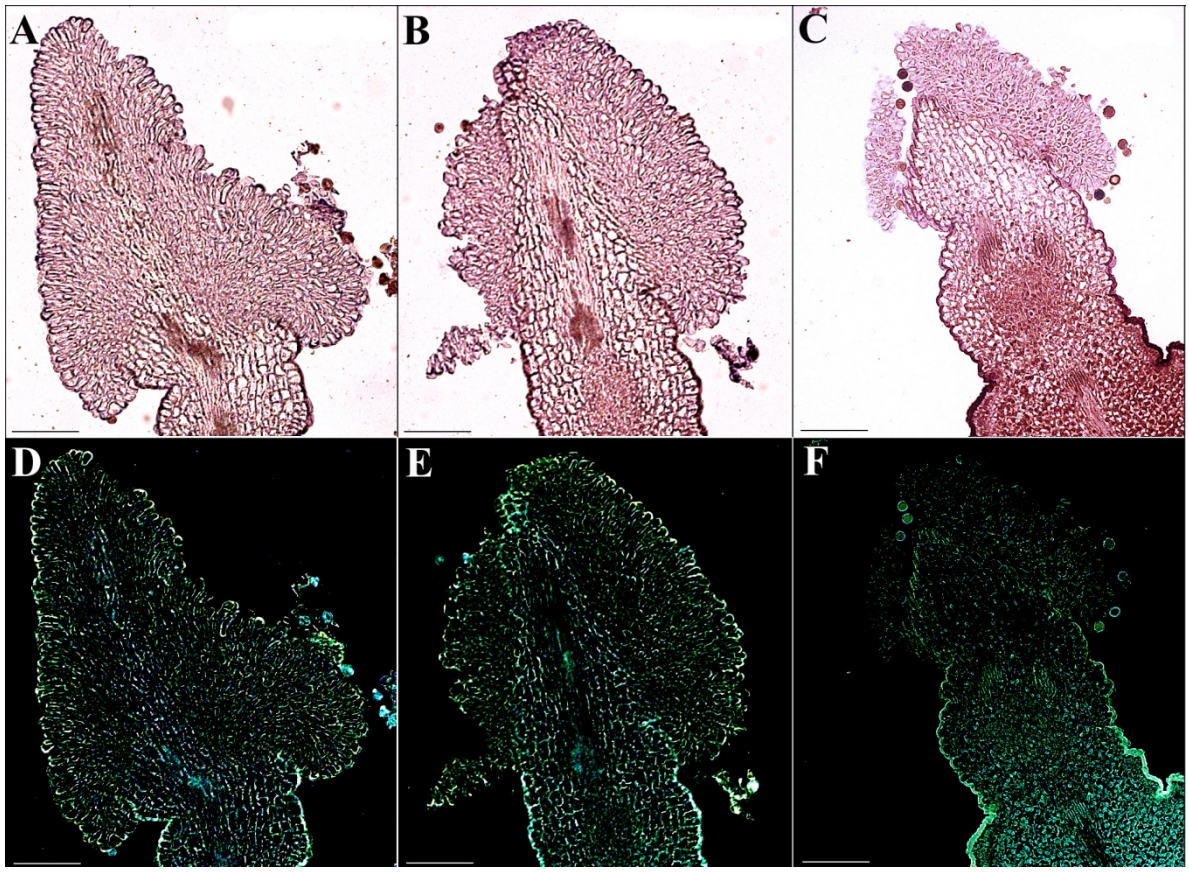


Figure 1.30. *In situ* hybridizations revealed that the expression of *OeSRK* (A, D) and *OeSLG* (B, E) was restricted to the external surface of papillar cells. C, F: *in situ* performed without probe, used as negative control. D-F: the same samples of A-C but with inverted colors in order to under light signals on papillar cells.

Summarizing, we found one ortholog of *SRK* of *Brassica*, *OeSRK*, which was pistil-specific expressed and one good candidate as male determinant, *OeSCR-like_12*, which was characterized by a conserved cysteine rich pattern and it was anther-specific expressed. Furthermore, the ortholog of *SLG* of *Brassica*, *OeSLG*, showed an unexpected lower band only in anthers and its expression pattern in pistils was similar to that of *OeSRK*, being more expressed during the flowering day (T0) and less expressed three days after flowering (T3). Theoretically, *OeSRK* and *OeSLG* should be linked in the same locus (S-locus, along with *OeSCR-like_12*) and the similar expression pattern could be a proof of their association. However, in order to get a further confirm of their association, we have identified SNPs in *OeSRK* and *OeSCR-like_12* useful to test gene association in segregating populations. We have also cloned full-lengths of *OeSRK*, *OeSLG* and *OeSCR-like_12* in expression vectors and we are currently testing their proteins interaction by a Yeast-2-Hybrid screen.

To determine whether small RNAs could play a role in the regulation of SI-related genes, we have also identified 204 putative miRNAs, which were used to predict interaction with our SI candidate genes. 38 miRNAs were predicted to interact with *OeSRK*, 12 with *OeSLG* and 19 with *OeSCR-like*, but further analyses are needed to confirm predicted interactions between isolated miRNAs and SI-related genes.

On the basis of our results, we supposed that olive adopts a sporophytic self-incompatible system similar to that of *Brassica*, contrarily to the gametophytic self-incompatibility method currently reported in literature. We further hypothesize that, although olive could adopt a SSI similar to *Brassica*, a variant might also occur. In fact, olive belongs to the *Oleaceae* family, which is the same family of another species recently characterized as sporophytic self-incompatible: *Phillyrea angustifolia* (Saumitou-Laprade *et al.*, 2010). *Phillyrea angustifolia* is a species characterized by a high degree of androecious flowers and has been characterized as sporophytic self-incompatible species according to the

existence of only two incompatibility groups, G1 and G2. Two incompatibility groups can be explained only by sporophytic self-incompatibility, because in the case of gametophytic self-incompatibility three is the minimum number of incompatibility groups required (Saumitou-Laprade *et al.*, 2010). A recent collaboration with the group of Saumitou-Laprade allowed us to sequencing orthologs of our SI genes in *Phillyrea angustifolia* demonstrating that *Olea europaea* and *Phillyrea angustifolia* share a high conservation of both nucleotide composition and intron distribution. It is worth noting that in olive, especially in some cultivars, a high degree of female sterility and, consequently, a high incidence of male flowers can also occur. These evidences support the hypothesis of a high conservation of reproductive features between the two species and, consequently, we hypothesize that also olive might be characterized by only two groups of incompatibility. Potential differences between olive and *Phillyrea* may be the result of centuries of anthropic selection in olive cultivars despite no selection occurred in *Phillyrea*. To finely define the conservation of SI method between olive and *Phillyrea*, we are currently testing the SI genes association in a segregating population of *Phillyrea* and we will compare the results with those obtained from SI genes segregation test performed in olive.

On the whole, we demonstrated that olive is a self-incompatible species and the rejection of self-pollen occurs at the stigma level. We further cloned candidate genes for female and male determinants showing a tissue-specific expression in pistils and anthers, respectively, and a differential expression pattern of SI genes was documented between self-compatible and self-incompatible cultivars. Therefore, we hypothesize that olive is a sporophytic self-incompatible species and that the pathway leading to the recognition and rejection of self-pollen is similar to the one of *Brassica*.

Reference

Alché J.D., Butowt R., Castro A.J., Rodriguez-Garcìa M.I. (2000). Ubiquitin and ubiquitin-conjugated proteins in the olive (Olea europea L.) pollen. Sex Plant Reprod 12, 285-291.

Androulakis I.I., Loupassaki M.H. (1990). Studies on self-fertility of some olive cultivars in the area of Crete. Acta Horticulturae 286, 159-162.

Ayerza R., Sibbett G.S. (2001). Thermal adaptability of olive (Olea europaea L.) to the Arid Chaco of Argentina. Agr. Ecosyst. Environ. 84, 277-285.

Badr S.A., Hartmann H.T., Martin G.C. (1970). Endogenous gibberellins and inhibitors in relation to flower induction and inflorescence development in the olive. Plant Physiology 46, 674-679.

Bartolini S.G.R. (1995). Self-compatibility in several clones of olive oil cv. 'Leccino'. Adv. Hortic. Sci. 9, 71-74.

Bateman A.J. (1952). Self-incompatibility systems in angiosperms. I. Theory. Heredity 6: 285-310.

Besnard G., Khadari B., Villlemur P., Berville A. (2000). Cytoplasmic male sterility in the olive (Olea europaea L.). Theoretical and applied genetics. 100, 1018-1024.

Besselat B (1992). La prévision du potentiel de récolte a l'aide de la technique du dosage pollinique de l'atmosphére. In: Méthodes aéropalinologiques, phénologie de la floraison et prévision de rendements. Int. Conf. Bordeaux 1992. Proc. (ed. Org. Comm.) 9-12.

Boggs N.A., Dwyer K., Shah P., McCuuoch A., Bechsgaard J., Schierup M., Nasrallah M., Nasrallah J. (2009a). Expression of Distinct Self-Incompatibility Specificities in Arabidopsis thaliana. Genetics 182, 1313-1321.

Boggs N.A., Nasrallah J.B., Nasrallah M.E. (2009b). Independent S-locus mutations caused self-fertility in Arabidopsis thaliana. PLoS Genet 5, e1000426.

Bradley MV, Griggs WH, Hartmenn HT (1961). Studies on self- and cross- pollination of olives under varying temperature conditions. California Agriculture 4-5.

Burd M. (1998). "Excess" flower production and selective fruit abortion: a model of potential benefits. Ecology 79, 2123-2132.

Busch J.W. Schoen D.J. (2008). The evolution of self-incompatibility when mates are limiting. Trends in plant science. 13, 128-136.

Chen X.M. (2009). Small RNAs and Their Roles in Plant Development. Annu. Rev. Cell Dev. Biol. 25, 21-44.

Connor DJ, Fereres E (2005). The physiology of adaptation and yield expression in olive. Horticultural Reviews 31.

Cuevas J, Diaz Hermoso AJ, Galian D, Hueso JJ, Pinillos V, Prieto M, SolaD, Polito VS (2001). Response to cross pollination and choice of pollinisers for the olive cultivars (Olea europaea L.) 'Manzanilla de Sevilla', 'Hojiblanca' and 'Picual'. Olivae 85, 26-32.

Cuevas J, Polito VS (1997). Compatibility relationships in ‘Manzanillo’ olive. Hortic. Sci. 32, 1056-1058.

Cuevas J, Polito VS (2004). The role of staminate flowers in the breeding system of Olea europea (Oleaceae): an andromonoecium, wind pollinated taxon. Ann. Bot. 93, 547-553.

Cuevas J, Rallo L, Rapoport HF (1994). Initial fruit set at high temperature in olive, Olea europaea L. Journal of Horticultural Science 69, 655-672.

Dai X., Zhao P.X. (2011). psRNATarget: A Plant Small RNA Target Analysis Server. Nucleic Acids Research, 2011.

De Graaf BHJ, Derksen JWM, Mariani C (2001). Pollen and pistil in the progamic phase. Sex Plant Reprod 14, 41-55.

de la Rosa R, Rallo L, Rapoport HF (2000). Olive floral bud growth and starch content during winter rest and spring budbreak. HortScience 35, 1223-1227.

Diaz A., Martin A., Rallo P., Barranco D., Rosa R.d.l. (2006). Self-incompatibility of 'Arbequina' and 'Picua' olive assessed by SSR markers. Journal of the American Society for Horticultural Science. 131, 250-255.

Dimassi K., Therios I., Balatsos A. (1999). The blooming period and self-fruitfulness in twelve greek and three foreign olive cultivars. Acta horticulturae.275-278.

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

Fabbri A. Benelli C. (2000). Flower bud induction and differentiation in olive. Journal Of Horticultural Science & Biotechnology 75, 131-141.

Fernàndez-Bolanòs P., Frìas L. (1969). Autofertilidad y austerilidad en el olivo. Agricultura 443:150–151.

Fernandez-Escobar R., Gomez-Valledor G., Rallo L. (1983). Influence of pistil extract and temperature on in vitro pollen germination and pollen tube growth of olive cultivars. Journal of horticultural science. Journal of horticultural science 58, 219-227.

Fernandez-Escobar R., Benlloch M., Navarro C., Martin G. (1992). The time of floral induction in the olive. J. Am. Soc. Hort. Sci. 117, 304-307.

Finnegan E.J., Liang D., Wang M.B. (2011). Self-incompatibility: Smi silences through a novel sRNA pathway. Trends in Plant Science 16, No. 5.

Foote H.C.C., Ride J.P., Franklin-Tong V.E., Walker E.A., Lawrence M.J., Franklin F.C.H. (1994). Cloning and expression of a distinctive class of self-incompatibility (S) gene from papaver rhoeas L. Proceedings of the National Academy of Sciences of the United States of America. 91, 2265-2269.

Fornaciari M., Pieroni L., Ciuchi P., Romano B. (1998). A regression model of the start of pollen season in Olea europaea L. Grana 37:110–113.

Franklin-Tong (1999). Signaling and the modulation of pollen tube growth. Plant Cell 11, 727.

Franklin-Tong N., Franklin F.C.H. (2003). Gametophytic self-incompatibility inhibits pollen tube growth using different mechanisms. Trends in plant science. 8, 598-605.

Goodwillie C. (1997). The genetic control of self-incompatibility in Linanthus parviflorus (Polemoniaceae). Heredity : an international journal of genetics. 79, 424-432.

Ghrisi N., Boulouha B., Benichou M., Hilali S. (1999). Agro-physiological evaluation of the phenomenon of pollen compatibility in olive. Case of the Mediterranean Collection at the Menera Station, Marrakech. Olivae 79: 51-59.

Griggs W.H., Hartmann H.T., Bradley M.V., Iwakiri B.T., Whisler J.E. (1975). Olive pollination in California. California Agricultural Experimental Station, Bulletin 869.

Gu T., Mazzurco M., Sulaman W., Matias D.D., Goring D.R. (1998). Binding of an arm repeat protein to a kinase domain of the S-locus receptor kinase. Proceedings of the National Academy of Sciences of the United States of America (USA).

Guerin J., Sedgley M. (2007). Cross pollination in Olive Cultivars. Australian Government, Rural Industries Research and Development Corporation, 2007.

Hartmann H.T., Opitz K.W. (1966). Olive production in Callifornia. Division of Agricultural Sciences, University of California, California, U.S.A., 64 p.

Higashiyama T. (2010). Peptide Signaling in Pollen-Pistil Interactions. Plant & cell physiology. 51, 177-189.

Hiscock S.J., McInnis S.M. (2003). Pollen recognition and rejection during the sporophytic self-incompatibility response: Brassica and beyond. Trends Plant Sci. 8, 606.

Hoebe P.N., Stift M., Tedder A., Mable B.K. (2009). Multiple losses of self-incompatibility in North-American Arabidopsis lyrata?: phylogeographic context and population genetic consequences. Mol. Ecol. 18, 4924.

Hua Z.H., 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.

Iannotta N., Palopoli A., Perri L. (1996). Effetti di trattamenti con molecole diverse sulla biologia fiorale di Olea europaea L. (cv. Carolea). p. 393–397 (In Italian). In: N. Lambardo and N. Iannotta (eds.). Atti del convegno. L'olivicoltura mediterranaea: stato e propettive della coltura e della ricerca.

Ida K., Norioka S., Yamamoto M., Kumasaka T., Yamashita E., (2001). The 1.55 ˚A resolution structure of Nicotiana alata SF11-RNase associated with gametophytic self-incompatibility. J. Mol. Biol. 314:103–12

Igic B., Kohn J. (2001). Evolutionary relationships among self-incompatibility RNases. Proc. Natl. Acad. Sci. U. S. A. 98, 13167-13171.

Igic B., Lande R., Kohn J.R. (2008). Loss of self-incompatibility and its evolutionary consequences. International journal of plant sciences. 169, 93-104.

Ivanov R., Gaude T. (2009). Endocytosis and endosomal regulation of the S-receptor kinase during the self-incompatibility response in Brassica oleracea. Plant Cell 21, 2107-2117.

Jones-Rhoades M., Bartel D. (2004). Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. Mol. Cell 14, 787-799.

Jordan N.D., Kakeda K., Conner A., Ride J.P., Franklin-Tong V.E., Franklin F.C.H. (1999). S-protein mutants indicate a functional role for SBP in the self-incompatibility reaction of Papaver rhoeas. Plant journal: for cell and molecular biology. 20, 119-125.

Kachroo A., Nasrallah M.E., Nasrallah J.B. (2002). Self-incompatibility in the Brassicaceae: receptor-ligand signaling and cell-to-cell communication. Plant Cell 14 Suppl, S227.

Kakeda, K. (1998). Identification of residues in a hydrophilic loop of the Papaver rhoeas S protein that play a crucial role in recognition of incompatible pollen. Plant Cell 10, 1723–1731

Kakita M., Murase K., Iwano M., Matsumoto T., Watanabe M., Shiba H., Isogai A., Takayama S. (2007). Two distinct forms of M-locus protein kinase localize to the plasma membrane and interact directly with S-locus receptor kinase to transduce self-incompatibility signaling in Brassica rapa. Plant Cell 19, 3961.

Kao T.H. Tsukamoto T. (2004). The molecular and genetic bases of S-RNase-based self-incompatibility. Plant Cell 16, S72-S83.

Kitashiba H. (2011). Functional test of Brassica self-incompatibility modifiers in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. 108, 18173-18178.

Kondo K. (2002). Cultivated tomato has defects in both S-RNase and HT genes required for stylar function of self-incompatibility. Plant journal: for cell and molecular biology.627-636.

Kowyama Y., Tsuchiya T., Kakeda K. (2000). Sporophytic self-incompatibility in Ipomoea trifida, a close relative of sweet potato. Annals of botany. 85, 191-196.

Kramer E.M. (2005). Methods for studying the evolution of plant reproductive structures: Comparative gene expression techniques. Molecular evolution: producing the biochemical data, part b 395, 617-636.

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

Larson B., Barrett S. (2000). A comparative analysis of pollen limitation in flowering plants. Biol. J. Linn. Soc. 69, 503-520.

Lavee S., Rallo L., Rapoport H., Troncoso A. (1996). The floral biology of the olive: Effect of flower number, type and distribution on fruitset. Scientia Horticulturae 66, 149-158.

Lavee S., Taryan J., Levin J., Haskal A. (2002). The significance of cross-pollination for various olive cultivars under irrigated intensive growing conditions. Olivae 91:25–36.

Lee H.S., Huang S., Kao T.H. (1994). S proteins control rejection of incompatible pollen in Petunia inflata. Nature 367, 560-563.

Levin D.A. (1993). S-gene polymorphism in Phlox drummondii. Heredity : an international journal of genetics. 71, 193-198.

Liu J.H., Honda C., Moriguchi T. (2006). Involvement of polyamine in floral and fruit development. Jarq-Japan Agricultural Research Quarterly 40, 51-58.

Lu C., Kulkarni K., Souret F.F., MuthuValliappan R., Tej S.S., Poethig R.S., Henderson I.R., Jacobsen S.E., Wang W., Green P.J., Meyers B.C. (2006). MicroRNAs and other

small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. Genome Research 16: 1276-1288.

Malik N., Bradford J. (2007). Different flower-inducing conditions elicit different responses for free polyamine levels in olive (Olea europaea) leaves. Journal Of The Japanese Society For Horticultural Science 76, 205-209.

Martin G.C. (1990). Olive flower and fruit population dynamics. Acta Hort. 286: 141-153.

Martin G.C. (1994). Botany of the olive. In: Ferguson L, Sibbett SG, Martin GC eds) Olive Production Manual. University of California, Division of Agriculture and Natural Resources, Oakland, CA, California, pp 19-21.

Matsuura T., Sakai H., Unno M., Ida K., Sato M. (2001). Crystal structure at 1.5-˚A resolution of Pyrus pyrifolia pistil ribonuclease responsible for gametophytic self-incompatibility. J. Biol. Chem. 276:45261–69

Matton d.p., Nass N., Clarke A., Newbigin E. (1994). Self-incompatibility - how plants avoid illegitimate offspring. Proc. Natl. Acad. Sci. U. S. A. 91, 1992-1997.

McClure B.A., Haring V., Ebert P.R., Anderson M.A., Simpson R.J. (1989). Style self-incompatibility gene products of Nicotiana alata are ribonucleases. Nature 342:955–57

McClure B. (2004). S-RNase and SLF determine S-haplotype-specific pollen recognition and rejection. Plant Cell 16, 2840.

McCubbin A.G. Kao T. (2000). Molecular recognition and response in pollen and pistil interactions. Annu. Rev. Cell Dev. Biol. 16, 333-364.

Meyers B.C., Axtell M.J., Bartel B., Bartel D.P., Baulcombe D., Bowman J.L., Cao X., Carrington J.C., Chen X., Green P.J., Griffiths-Jones S., Jacobsen S.E., Mallory AC., Martienssen R.A., Poethig R.S., Qi Y., Vaucheret H., Voinnet O., Watanabe Y., Weigel D., Zhu J.K. (2008). Criteria for Annotation of PlantMicroRNAs. The Plant Cell, december 2008.

Minero F.J.G., Candau P., Morales J., Tomas C. (1998). Forecasting olive crop production based on ten consecutive years of monitoring airborne pollen in Andalusia (Southern Spain). Agricult Ecosyst Environ 69: 201-215.

Mishima M., Takayama S., Sasaki K., Jee J., Kojima C., Isogai A., Shirakawa M. (2003). Structure of the male determinant factor for Brassica self-incompatibility. J. Biol. Chem. 278, 36389-36395.

Mookerjee S., Guerin J., Collins G., Ford C., Sedgley M. (2005). Paternity analysis using microsatellite markers to identify pollen donors in an olive grove. Theoretical and applied genetics. 111, 1174-1182.

Murase K., Shiba H., Iwano M., Che F., Watanabe M., Isogai A., Takayama S. (2004). A membrane-anchored protein kinase involved in Brassica self-incompatibility signaling. Science 303, 1516.

Murfett J., Atherton T.L., Mou B., Gasser C.S., McClure B.A. (1994). S-RNase expressed in transgenic Nicotiana causes S-allele-specific pollen rejection. Nature 367, 563-566.

Nasrallah J.B., Kao T.H., Goldberg M.L., Nasrallah M.E. (1985). A cDNA clone encoding an S-locus-specific glycoprotein from Brassica oleracea. Nature 318, 263-267.

Nasrallah J.B., Kao T.H., Chen C.H., Goldberg M.L., Nasrallah M.E. (1987). Amino-acid sequence of glycoproteins encoded by three alleles of the S locus of Brassica oleracea. Nature 326, 617-619.

Newbigin E., Uyenoyama M.K. (2008). RNase-Based Self-Incompatibility: Puzzled by Pollen S. Plant Cell 20, 2286-2292.

Okada K., Tonaka N., Moriya Y., Norioka N., Sawamura Y., Matsumoto T., Nakanishi T., Takasaki-Yasuda T. (2008). Deletion of a 236 kb region around S4-RNase in a stylar-part mutant S4 sm-haplotype of Japanese pear. Plant Mol. Biol. 66, 389-400.

Orlandi F., Romano B., Fornaciari M. (2005). Relations between pollen emission and fruit production in olive (Olea europaea L.) Grana 44:98–103.

Pan D., Qui D., Lin Y., Shi M., Xie H. (1999). Effects of storage temperature on polyamines and activity of protective enzymes of Chinese olive fruit. J. Fojian Agric. Univ. 28: 284–288.

Qiao H., Wang HY, Zhao L., Zhou JL, Huang J., Zhang YS, Xue YB (2004a). The F-box protein AhSLF-S-2 physically interacts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in antirrhinum. Plant Cell 16, 582-595.

Qiao H.,Wang F., Zhao L., Zhou J., Lai Z., Zhang Y., Robbins T.P., Xue Y. (2004b). The F-box protein AhSLF-S2 controls the pollen function of S-RNase-based self-incompatibility. Plant Cell 16, 2307.

Rallo L., Martin G.C., Lavee S. (1981). Relationship between abnormal embryo sac development and fruitfulness in olive. Journal of the American Society for Horticultural Science. 106, 813-817.

Rallo L., Martin G.C. (1991). The role of chilling in releasing olive floral buds from dormancy. Journal of the American Society for Horticultural Science. 116, 1058-1062.

Rallo L. (1998). Frutificacíon y Produccíon. p. 112-144. In: Barranco D, Fernández-Escobar R, and Rallo L (eds.). El cultivo del olivo. Junta de Andalucía y Mundi-Prensa, Madrid.

Rapoport H.F., Rallo L. (1991). Postanthesis flower and fruit abscission in 'Manzanillo' olive. Journal of the American Society for Horticultural Science. 116, 720-723.

Rapoport H.F. (1998). Botanica y morfología. p. 34-60. In: D. Barranco, R. Fernández Escobar, and L. Rallo (eds.). El cultivo del olivo. Junta de Andalucía y Mundi-Prensa, Madrid.

Rea A.C., Liu P., Nasrallah J.B. (2010). A transgenic self-incompatible Arabidopsis thaliana model for evolutionary and mechanistic studies of crucifer self-incompatibility. Journal of experimental botany. 61, 1897-1906.

Reale L., Sgromo C., Bonofiglio T., Orlandi F., Fornaciari M., Ferranti F., Romano B. (2006). Reproductive biology of olive (Olea europaea L.) DOP umbria cultivars. 19, 151-161.

Robertson K., Goldberg E.E., Igić B. (2010). Comparative evidence for the correlated evolution of polyploidy and self-compatibility in solanaceae. Evolution.

Rodriguez-Garcia M.I., M'rani-Alaoui M., Fernàndez M.C. (2003). Behavior of storage lipids develoopment and germiantion of olive (Olea europea L.) pollen. Protoplasma (2003) 221: 237-244.

Rugini E. Mencuccini M. (1985). Increased yield in the olive with putrescine treatment. HortScience 20, 102-103.

Samuel M.A., Chong Y.T., Haasen K.E., Aldea-Brydges M.G., Stone S.L., Goring D.R. (2009). Cellular Pathways Regulating Responses to Compatible and Self-Incompatible Pollen in Brassica and Arabidopsis Stigmas Intersect at Exo70A1, a Putative Component of the Exocyst Complex. Plant Cell 21, 2655-2671.

Saumitou-Laprade P., Vernet P., Vassiliadis C., Hoareau Y., de Magny G., Dommee B., Lepart J. (2010). A Self-Incompatibility System Explains High Male Frequencies in an Androdioecious Plant. Science 327, 1648-1650.

Schopfer C.R., Nasrallah M.E., Nasrallah J.B. (1999). The male determinant of self-incompatibility in Brassica. Science 286, 1697.

Serrano I., Suárez C., Olmedilla A., Rapoport H.F., Rodríguez-García M.I. (2008). Structural organization and cytochemical features of the pistil in Olive (Olea europaea L.) cv. Picual at anthesis. Sexual Plant Reproduction 21, 99-111.

Shiba H., Takayama,S., Iwano M., Shimosato H., Ishimoto K., Watanabe M.,Isogai A. (2002). The dominance of alleles controlling self-incompatibility in Brassica pollen is regulated at the RNA level. Plant And Cell Physiology 43, S34-S34.

Shimizu K.K., Shimizu-Inatsugi R., Tsuchimatsu T.,Purugganan M.D. (2008). Independent origins of self-compatibility in Arabidopsis thaliana. Molecular ecology. 17, 704-714.

Sijacic P., Wang X, Skirpan A.L.,Wang Y.,Dowd P.E., McCubbin A.G, Huang S., Kao T.H. (2004). Identification of the pollen determinant of S-RNase-mediated self-incompatibility. Nature 429, 302.

Silva N.F. (2001). Expression of the S receptor kinase in self-compatible Brassica napus cv. Westar leads to the allele-specific rejection of self-incompatible Brassica napus pollen. Molecular genetics and genomics : MGG. 265, 552-559.

Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E., Nasrallah J.B. (1991). Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of Brassica oleracea. Proceedings of the National Academy of Sciences of the United States of America. 88, 8816-8820.

Stein J.C., Dixit R., Nasrallah M.E., Nasrallah J.B. (1996). SRK, the stigma-specific S locus receptor kinase of Brassica, is targeted to the plasma membrane in transgenic tobacco. Plant Cell 8, 429-445.

Stephenson A.G. (1981). Flower and fruit abortion: proximate causes and ultimate functions. Annual review of ecology and systematics.253-279.

Stephenson A.G., Doughty J., Dixon S., Elleman C., Hiscok S., Dickinson H.G. (1997). The male determinant of self-incompatibility in Brassica oleracea is located in the pollen coating. Plant journal: for cell and molecular biology. 12, 1351-1359.

Stone S.L., Arnoldo M., Goring D. (1999). A breakdown of Brassica self-incompatibility in ARC1 antisense transgenic plants. Science 286, 1729-1731.

Stone S.L., Anderson E.M., Mullen R.T., Goring D.R. (2003). ARC1 is an E3 ubiquitin ligase and promotes the ubiquitination of proteins during the rejection of self-incompatible Brassica pollen. Plant Cell 15, 885.

Stutte G.W., Martin G.C. (1986). Effect of killing the seed on return bloom of olive. Scientia Horticulturae (Netherlands).

Sunkar R., Zhu J.K. (2004). Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell 16, 2001-2019.

Sutherland S. (1987). Why hermaphroditic plants produce many more flowers than fruits: experimental tests with Agave mckelveyana. Evolution. 41, 750-759.

Suzuki G., Tanaka S., Yamamoto M., Tomita R.N., Kowyama Y., Mukai Y. (2004). Visualization of the S-locus region in Ipomoea trifida: toward positional cloning of self-incompatibility genes. Chromosome research. 12, 475-481.

Takasaki T., Hatakeyama K., Suzuki G., Watanabe M., Isogal A., Hinata K. (2000). The S receptor kinase determines self-incompatibility in Brassica stigma. Nature 403, 913-916.

Takayama S., Isogai A., Tsukamoto C., Ueda Y., Hinata K., Okazaki K., Suzuki A. (1987). Sequences of S-glycoproteins, products of the Brassica campestris self-incompatibility locus. Nature 326, 102-105.

Takayama S., Shiba H., Iwano M., Shimosato H., Che F.S., Kai N., Watanabe M., Suzuki G., Hinata K., Isogai A. (2000). The pollen determinant of self-incompatibility in Brassica campestris. Proceedings of the National Academy of Sciences of the United States of America. 97, 1920-1925.

Takayama S., Shimosato H., Shiba H., Funato M., Che F.S., Watanabe M., Iwano M., Isogai A. (2001). Direct ligand-receptor complex interaction controls Brassica self-incompatibility. Nature 413, 534-538.

Takayama S., Isogai A. (2005). Self-incompatibility in plants. Annual Review Of Plant Biology 56, 467-489.

Tantikanjana T., Nasrallah M.E., Nasrallah J.B. (2010). Complex networks of self-incompatibility signaling in the Brassicaceae. Curr. Opin. Plant Biol. 13, 520-526.

Tarutani Y., Shiba H., Iwano M., Kakizaki T., Suzuki G., Watanabe M., Isogai A., Takayama S. (2010). Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility. Nature 466, 983.

Tomita R.N., Fukami K., Takayama S., Kowyama Y. (2004). Genetic mapping of AFLP/AMF-derived DNA markers in the vicinity of the self-incompatibility locus in Ipomoea trifida. Sexual plant reproduction. 16, 265-272.

Tsuchimatsu T., Suwabe K., Shimizu-Inatsugi R., Isokawa S., Pavlidis P., Städler T., Suzuki G., Takayama S., Watanabe M., Shimizu K.K. (2010). Evolution of self-compatibility in Arabidopsis by a mutation in the male specificity gene. Nature 464, 1342.

Tsukamoto T., Ando T., Watanabe H., Marchesi E., Kao T.H. (2005). Duplication of the S-locus F-box gene is associated with breakdown of pollen function in an S-haplotype identified in a natural population of self-incompatible Petunia axillaris. Plant molecular biology. 57, 141-153.

Ushijima K., Sassa H., Dandekar A., Gradziel T., Tao R., 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.

Vanoosthuyse V., Miege C., Dumas C., Cock J.M. (2001). Two large Arabidopsis thaliana gene families are homologous to the Brassica gene superfamily that encodes pollen coat proteins and the male component of the self-incompatibility response. Plant molecular biology. 46, 17-34.

Vassiliadis C., Lepart J., Pierre Saumitou-Laprade, Vernet P. (2000). Self-Incompatibility and Male Fertilization Success in Phillyrea angustifolia (Oleaceae). International Journal of Plant Sciences 161, 393-402.

Wang Y., Tsukamoto T., Yi K., Wang X., Huang S., McCubbin A.G., Kao T. (2004). Chromosome walking in the Petunia inflata self-incompatibility (S-) locus and gene identification in an 881-kb contig containing S2-RNase. Plant Mol. Biol. 54, 727.

Watanabe M., Ito A., Takada Y., Ninomiya C., Kakizaki T., Takahata Y., Hatakeyama K., Hinata K., Suzuki G., Takasaki T., Satta Y., Shiba H., Takayama S., Isogai A. (2000). Highly divergent sequences of the pollen self-incompatibility (S) gene in class-I S haplotypes of Brassica campestris (syn. rapa) L. FEBS Lett. 473, 139.

Wheeler M.J., de Graaf B., Hadjiosif N., Perry R., Poulter N., Osman K., Vatovec S., Harper A., Franklin F., Franklin-Tong V. (2009). Identification of the pollen self-incompatibility determinant in Papaver rhoeas. Nature 459, 992-U118.

Whitehouse H.L.K. (1950). Multiple-allelomorph incompatibility of pollen and style in the evolution of angiosperms. Annals of Botany 14: 198–216.

Willi Y. (2009). Evolution towards self-compatibility when mates are limited. Journal of evolutionary biology. Journal of evolutionary biology 22, 1967-1973.

Yang Q., Zhang D., Li Q., Cheng Z., Xue Y. (2007). Heterochromatic and genetic features are consistent with recombination suppression of the self-incompatibility locus in Antirrhinum. Plant journal. 51, 140-151.

Zhang Y., Zhao Z., Xue Y. (2009). Roles of Proteolysis in Plant Self-Incompatibility. Annual review of plant biology.21-42.

Zuker M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31 (13), 3406-3415.

Appendix 1.1. Gene sequences.

>OeSLG
ATGGAGAAATCGATTAAAGATATATCCCTTCTTCTTCTCTTAACATCCATCTTGTTCATCCTAGAGATTTCAC
CTGCAATTGATACCATTAGCACAACTCAGAGCCTCAAAGATGGAGATACCCTGGTTTCATCAGGGGGAACCTT
CGAACTGGGATTTTTCAGCCCAGGTGACTCCAAGAATCGGTACGTGGGAATTTGGTATAAGAAGGTGCCTAGC
ATAACAGCAGTTTGGGTCCTCAACAGAGAAATTCCGCTAAATAGTAGATCAGGTATACTAAAGTTCAATGAGC
TAGGCCACTTGGTTCTTGTGAATGACACTAACAACCTCTTGTGGTCTTCAAATACATCAAGAATTGCTCGAAC
TCCAATTTTGCAATTGCTGGACTCGGGAAATCTTGTTCTTCGAGAAGCAAATGATGATAATCTGGAGAATTTC
CTTTGGCAGAGTTTCGATTATTTAAGTGACACTTACCTACCAGGCATGAACTTCGGTTGGAATGCTGCAACAG
GTGTACAGAACTATTTGTCATCATGGACGAGCAATGAGGATCCGGCTCCAGGAGATCTTACATTTTACCTGGA
TCCAACTGGATATCCACAAGTCTTCATCAAAAGAGGCACAGGTGCCATATACAGGATGGGACCTTGGAATGGT
CTTCGCTTTAGTGGAACACCATATGTGAGTCCTACATTCAGACATGGAATATTCAAGAATAAGAACACGACGT
ACTATAGAGAAGACTCCAATGACAAATCAGTTATTTCAAGGGTTACCCTGAATCAGAGTGGTGTTGTACAGCG
CTGGGTATGGGTTGATCGAACCAGAGGTTGGGTCCTCTACTTGACTGTACCAAAAGATGATTGTGACACTTAT
AGTGACTGTGGCGCTTATGGGACTTGCTATATTGGGAATTCTCCGGCTTGTGGATGTCTGAGTAAATTTCAAC
CAAAAGATCCAGAAGGATGGAATAAGGGAGATTGGTCAAATGGGTGCATTAGAAGGACTCCCTTAAATTGTCA
AGAAGGTGATGTATTTTTGAAGTATTCTAGCGTTAAATTACCAGACGCACAATATTCCACGTATAATGAGAGT
ATGACACTTGACGAAAGCGAAGTGAAGTGTTTACAGAATTGTTCTTGTATGGCATATTCACAATTGGATATCA
GCCGAGGGAGTGGATGCCTGTTTTGGTTTCGAGAATTGATTGACATCAGAGATATGTCTTCAGACGGACAAGA
TATTTACATTAGAATGGCTTCTTCTTAG

>OeSRK
ATGGAAGGCTTAACAAGAATTGCTTTTCTCTGTAATTTAATTTTCTCTGTTCTTTTCACTTGTGAAGCAGTAG
ACACCATAACTGTAAACACTTCTCTTACAGATGGAAACTACGTCGTTTCATCAGGGGGAACATTTGAAATGGG
ATTCTTCAGCCCAGATAATTCTATCAACAGGTACATAGGAATATGGTACAAACAAATTTCAGTGAAAACAATT
GTATGGGTGGCCAATAGAGAAACCCCTTTTAATGATACTTCAGGAGCTTTGAAGTTGACAGAGAATGGGAATT
TGGTACTTGTCAATGGTACTGATGGAGTTGTTATGTGGTCATCCAATTCCACTGGTGGTACTATGAGTATGAG
TCCGGTGGCACAACTACTCGACTCGGGAAATCTTGTCATACGAGGTTCGAGGGACGGAAATTACACTTGGCAG
AGTTTTGATCATCCAACAGACACGGCCTTGCCTGGGCTGAAAATGGGGAAGAATTTAGTGACGGGCGTTGACC
GGATTCTATATTCGAGGAAGAGTAACAACGATCCTTCGCGGGGAGACTATATGTATCTTATGGATACTCATGG
ATATCCACAACACATGATGATGACTGGTTCGACTGTTCGTTTTCGATCAGGACCATGGAATGGTTTGGCGTTT
AGTGGGTCTCCAGGCCTAAAAACGAATCCGATCTACACATTTCAATTCGTGTTCAATCAAGAGGAAGTATACT
ATAGCTTTGACCTCGTTAACCCTCATGTTTACTCAAGGCTTGTATTGGATCCGGATGGAGTTCTTCGACGATT
TTCGTGGAATAATAGAACACAGGTTTGGACTAATTTGGTTAGTGCACCAGCTGATAATTGTGACATTTATGGA
CAGTGTAATGGATATGGTAAATGTACAATTGGTGAGTCTCCGATTTGTAGTTGTTTGGATAAATTTAAGCCTA
AAAACCCGAAAGATTGGCTCTCGGCCGTTTGGTCAGATGGTTGTGTTCGTAGAACGCCGTTGAATTGTAATAG
CGATGGTTTTGTAAAGTATTCTAGAGTTAAATTACCTGATACAAGAAAATCTTGGTATAATTTGAGCATGTCA
CTTAAGGAATGCAGGCAAATGTGTAAGAATAATTGTTCTTGTATGGCTTATTCCAATATAGATATAAGAGGGA
AAGGAAGTGGTTGCTTTCTTTGGTTTGAAGATCTGATGGACATTAGATATTACGATGGAAACGATGGTCAAGA
TATCTATATTAGGATGGCATCCTCTGAGTTAGGCTCAAGTGGGTTGAGAAAGAAGATACTTAGGGCATGTTTG
GCATCACTTGGGGCAGTTCTTATATTGTGCCTCATACTCATCTCATTTACTTGGAAGAAGAAGAGGGATAGAG
AAAAACAGCAGCAGGTGCAGCAACAGCTGACTCGAGAAGGATCAATAGGAAGCAGTTCAAGACAATTTTACAC
AGCAGAAAATGACAATGGAGATCTTGACCTACCATTGTTTGATGTAACAACAATATTAGAAGCTACCAATTAC
TTTTCACCTGGCAACAAGATTGGAGAAGGTGGATTTGGACCAGTGTACAAGGGTGTACTCAGAAAGGGGAAAG
AAATTGCTGTGAAGAGGCTATCAAAGTATTCAATACAAGGAGATGACGAGTTCAAAAATGAAGTAATTTTAAT
TGCAAAGTTACAACATAGGAATCTTGTGAATCTTATTGGATGTTGCATTCATGAAGAAGAGAAAATTCTTATC
TATGAATTCATGCCTAACAATAGCTTGGACAGTTACATATTTGATAAGGATCGAGGTAGATTGTTGGATTGGG
AAAAGCGGTTTCAAATTATCAATGGAATTGCTCGTGGATTACTTTATCTTCACCAAGACTCCCGATTGAGAAT
CATTCATAGAGATCTAAAGGCTGGAAACATTTTGTTGGATGCTGATATGAACCCGAAAATATCAGACTTCGGC
ATGGCCAGAAGTTTTGGAGGGAATGAAATTGAAGCAAACACACGAAGGGTTGTCGGTACATATGGCTATATGT
CACCCGAATATGTGGTTGATGGTCACTTCTCAGTCAAATCTGACATATTCAGCTTTGGTGTCCTCATTCTGGA
AATCATAAGCGGACAAAAGAACAGAGGATTTTTTCATCAAGATCACCATCACAATTTGCTTGGACATGCATGG
ATTCTACACAATGAAGGACGATCTCTAGAATTAATCGACTCACATTTAGCTCAATCGTGCTACTTGTCTGAAG
TACTTCGATCAATGCATGTCGCTCTATTATGTGTACAGAGAAATCCTGAAGATCGGCCAAATATGTCGAATGT
TGTCCTAATGTTGGCAAGTGCAGGGGCATTGCCTAAGCCTAAAGAGCCCGGTTTTTTCACAGAAAGAAATTCA
TTTCTTGGATTTGAAACTTCATCAAGTAAGCCTACGGTAAGTTCTGCCAATGAATTGAGCTTCACGGAAATGG
AAGGTAGATAA

>OeSCR-like_12
ATGGCAGGTCGTCCGTACAATACAGGCATTTTGTTGGTTACATCGCTACTTGTATTGATTACATTATCCAACG
TTGTTGAGGGATATAATAGACTTCATCCTCGAGATTGCAATCCAAGATGCACGTACCGTTGCTCAGCAACATC
CCATAAGAAGCCATGCATGTTCTTCTGTCAAAAGTGCTGTGCGAAGTGCTTGTGTGTGCCACCAGGCACTTAT
GGCAATAAACAAACATGCCCTTGTTACAATAATTGGAAGACTAAGGAAGGCGGACCTAAATGCCCCATAAATT
AA

>OeSRK-like
AGCAGGAGCAGGAGCAGGAGCAGGATCATCAGGAAATGATTCCAATAAGACACAAAAAATTATCATGGTTGCT
GGCATTACGGTTGGCATTGGCGTCGTTCTGGCTGGATTCTGCATTTTCTTCATCTGGAAGAGGAAGAAATCCC
GTACTGAGTTAGGGAATAGGACGGAGTTTAGAGGTTCAAGGGAAAGAAGTCAAGATTTTCTGATGAATGCACC
GTCTATTCTCAGCAAAAGAGACCATTCAGGTGAAAGTACAGTGGACGAAGTGGATTTACCGTTATATGATTTC
AGTACCATAGTAATGGCTACTGATAAATTTTCCGATGCTAATAAGCTTGGACAAGGTGGTTTTGGTTGTGTTT
ATAAGGGTATATTGGTGGAAGGCCAGGAAATAGCAGTGAAGAGGCTTTCGAAAAATTCTGGCCAAGGAATAGA
GGAATTTAAAAATGAGATGAAGTTGATTGCAAGACTTCAGCACAGAAATCTGGTTCGTTTGCTTGGCTGCTGC
ATTGATATGGAGGAGAAGATGCTGGTATACGAGTACATGGAAAACAAAAGCCTGGATTCGATTTTATTCAAGA
AAGAGAAAAGTTCATTGCTCGACTGGAACAGGAGGTTCAACATCATATGTGGGATTGCACGAGGGCTTCTCTA
TCTTCATCAAGATTCAAGATTTAGAATTATCCATAGAGATCTCAAAGCGAGCAACATTCTTTTGGATAAAGAG
ATGAACCCTAAAATATCAGATTTTGGCATGGCAAGAATTTTTGGAGGAGATCAGACCGAAGCGAATACAAAAA
GAGTAGTCGGAACATATGGTTACATGTCACCTGAATATGCAATGGATGGCCTATTCTCTATGAAATCTGATGT
ATTCAGTTTCGGAGTTCTAGTAGTAGAAATAGTAACTGGCAAGAAGAACAGAGGATTTTATCATACAAATAAC
CAACTTAATCTTCTTGCGCATGCATGGAGGCTATGGAGAGAAGGAAGAGGTATGGAACTAATGGATTCAGCAG
CTGGTGAATCATATTCCCCAAGTGAGGTATTGAGGTGCATACAAGTCGGCCTCTTGTGCGTGCAGGAACAAGC
CGAAGATAGACCTAACATGGACACTGTAGTCCTGATGTTAAGCAGTGATACTGCATCAATGTCCCAACCTAAG
AACCCGGGATTTTGCATGGGAAGAAGACCTGCTGATACAGAATCATCATCAAGTAAACAGGATGAATCCTGCA
CTGTGAATCAAGTTACTGTTACCATTCTAGATGGTCGATAG

>OeSCR-like_defensin_1
GTACGCGGGGGACCCCATTCCATTTTCACTTTCTACATCTAAGATGGGACGCTATGCAGCTGTGTTCTTCATG
CTCATGCTTGTCATGCTCACCGGTCCTTTAATGGTGGCAGAGGCAAGGACTTGTGAATCCCAAAGCCACCGGT
TCAAGGGATCTTGTGTTAGTAAAAGCAACTGCGCAGCGGTTTGCCAGACTGAGGGATTTCCCGATGGCTATTG
CAGGGGATTTCGCCGCCGCTGCTTCTGCTCCAAACATTGTTAACTCACCGTCTTGACGGTTGTTTTCTGGCAG
TCTCTGGTGGTGTCGGTCGACGGTCATCAATTTTCCTGAGAAAATTATGTCCTTTAATTAACGTGCTTTTCCT
GCTCATGGTGTTTGGTGTATTAATAATATTTAGTGTGTAATAATCTGTGAGATGTTTCTAATAACACGTTATT
TTGTACTG

>OeSCR-like_defensin_2
GTACGCGGGGGACCCCATTCCATTTTCACTTTCTACATCTAAGATGGGACGCTATGCAGCTGTGTTCTTCATG
CTCATGCTTGTCATGCTCACCGGTCCTTTAATGGTGGCAGAGGCAAGGACTTGTGAATCCCAAAGCCACCGGT
TCAAGGGATCTTGTGTTAGTAAAAGCAACTGTGCAGCGGTTTGCCAGACTGAGGGATTTCCCGACGGCTATTG
CAGGGGCTTTCGCCGCCGCTGCTTCTGCTCCAAACATTGTTAAATCACCGTCTTGACGGTTGTTTTCTGGCAG
TCTCTGGTGGTGTCGGTCGACGGTCATCAATTTTCCTGAGAAAATTATGTCCTTTAATTAACGTGCTTTTCCT
GCTCATGGTGTTTGGTGTATTAATAATATTTAGTGTGTAATAATCTGTGAGATGTTTCTAATAACACGTTATT
TTGTACTGA

>OeSCR-like_defensin_3
GTACGCGGGGGACCCCATTCCATTTTCACTTTCTACATCTAAGATGGGACGCTATGCAGCTGTGTTCTTCATG
CTCATGCTTGTCATGCTCACCGGTCCTTTAATGGTGGCAGATGCAAGGACTTGTGAATCCCAAAGCCACCGGT
TCAAGGGATCTTGCGTTAGTAAAAGCAACTGCGCGGCGGTTTGCCAGACTGAGGGATTTCCCGACGGCTATTG
CAGGGGCTTTCGCCGCCGCTGCTTCTGCTCCAAACATTGTTAAATCACCGTGTTGACGGTTGTTTTCTGGCAG
TCTCTGGTGGTGTCGGTCGACGGTCATCAATTTTCCTGAGAAAATTATGTCCTTTAATTAATGTGCTTTTTCT
GCTCACGGTCTTTGGTGTATGAATAATATTTAGTGTGTAATAATCTGTGAGATGTTTCTAATAACACGTTAAG
TGATTTTGT

>Oe-SCRlike_8
GTGGTGGTGGCCATCTGCAGGTCAGCATCGGATAGTATTACATTTCAATGCCTTTGTAGCTGCCTCAGCAAAA
CACTGAGCAGCCTCTGCCTCAACTTCTGCCTCCTCTGCTTCCCTTGCCGCCACCTCAGCTTCTGCAATTGCAG
CTTCTGCCTCTGCGACTGCTTGTGCAGCAGCTGCAGCTGCCTCTTGTGCAAGTACAATAGTTACTTACAATAC

>Oe-SCRlike_9
GTGGTGGTGGCCATCTGCAGGTCAGCATCGGATAGTATTACATTTCAATGCCTTTGTAGCTGCCTCAGCAAAA
CACTGAGCAGCCTCTGCCTCAGCTTCTGCCTCCTCTGCTTCCCTTGCCGCCACCTCAGCTTCTGCAATTGCAG
CTTCTGCCTCTGCGACTGCTTGTGCAGCAGCTGCAGCTGC

>Oe-SCRlike_10
AGTGGTATCAACGCAGAGTACGCGGGGGAGCTGAATCCTTTCAAATACCCCTCCAGTAAACAATATTATTATA
GGAGCATCCTCCACAAGCGAACAGATTTTTATACTTTCTGCTTTCAGTTTTGTGGTTTTAATGTGACGGTTGT
TTTTACATGTGGATCTAACCAGTTAATAAATGTATTGCATTGCATATGCATCTTTTTATTAGAGAAAAAGGTG
GGCATTAGAGGGCGAAAAAACAACTGCATATGCCTCCGGTTTCTGTGCTCCAGAATTTTG

>Oe-SCRlike_11
AAACGTGTCGTTGTAGTAGGTGGTGGACCAGGAGGTTACGTGGCTGCCATCAAGGCCGCTCAGCTTGGGCTCA
AGACTACCTGTATTGAGAAGCGTGGTGCCCTCGGTGGAACTTGCCTCAATGTCGGATGTATACCTTCCAAGGT
ATTGTTCCGGTTTTGTTTTAGATATACTGAATTGAATTTCGGTTGTTGTGTTCTTGTTCTGTTTTATTTAGAT
TGTAAAAATTGTGCTTGTTTTGTATTATTTTATTGTTGTATTTGTAAGTATTTG

>Oe-SCRlike_12
ATGGCAGGTCGTCCGTACAATACAGGCATTTTGTTGGTTACATCGCTACTTGTATTGATTACATTATCCAACG
TTGTTGAGGGATATAATAGACTTCATCCTCGAGATTGCAATCCAAGATGCACGTACCGTTGCTCAGCAACATC
CCATAAGAAGCCATGCATGTTCTTCTGTCAAAAGTGCTGTGCGAAGTGCTTGTGTGTGCCACCAGGCACTTAT
GGCAATAAACAAACATGCCCTTGTTACAATAATTGGAAGACTAAGGAAGGCGGACCTAAATGCCCCATAAATT
AA

>Oe-SCRlike_14
GAGTGCAGCAAGCGGCATCCATATGTGGCAGCAAAACCTTGTGCTACTTCTGTATCTTTTTGCAGGAAATCTT
TAGTGTACGTAGAAATACGATGTTTGTGTAAGAAGATTACTGGTTCTTGGAGCCACATAGCACTTTTTTGTGC
ATGTTTTGCACTGCTGAGGGATTTCAATGAGAACCTTTTTTTCGCTAGT

>Oe-SCRlike_15
CTGCATTCGACTTACTGCTGGACCCTTGCATGCTCTTGCTGCTGCTACACGGCCTGCACGAGCTGCTCCTGTT
TCTGGGGTTCCTGGTGTTGCTGCATCTTTTCTTCTACTTTGAAGCAGAGTGGGTACCTGCCGTCTCAGGGTCA
TCTGCCAAGCAATTGTAGCACAAATACAAGTCAAGCTTCCTCTGGTTCTGCAGGTAATCCTGGTGGACCTACC
CCGTCTGGTCCTGTTGGGAGTCACAACCACCCTACTGCTTCAATGTTGGCAGCAGCCGCTGCAGCTGCTGCTG
CTGGAAGAGGTGGTTCTGGCATTGTTCCAAGTTCTCTGTTGCCTATTGACGTTTCTGTAGTACTTCGCGGCCC
ATACTGGATTCGGATAATATACCGCAAAAACTTTGCTGTAGACATGCGGTGTTATGCAGGAGATCAAGTCTGG
TTG

>OeSCR-like_17
GCCGCTTTTCTTCACAAATTCATTCACTTGATCCCTCACTCTCTTTCTATTTCAATTTGCTTCAAGGCCTCGT
ATACTTTCCAAGCAATGGCAAAGAAACTAACCTTCTTGTTCTGTCTAATATTCTTCATGGTTGCTATTCTAGC
CCACAATCAGGCTTCTAGCACCTTGCTACCAGCTTCTCAGCTTGAACCACAGGGAAACGGCCAAATGTATGGC
ACTACTCAGGGGAGCCTTCAGCCTCAAGAATGTGCACCTAGATGCACTGGTAGATGCTCAAAAACTTCATTCA
AGAAGCCATGCATGTTTTTCTGTCAAAAATGCTGTGCAAAGTGTTTGTGTGTGCCTCCTGGCACATATGGCAA
CAAGCAAGTATGCCCTTGCTACAATAACTGGAAAACCAAGAGAGGTGGACCAAAGTGCCCT

>OeSCR-like_18
CAGAGAGAGACCTATTATATGTATCACTCTATTCTTATATTCACCTCCAAAATACTTGCATTCAACTCACTAG
AGTTTAAAGACTTCAGAAATATGAAGAAGTTCTTCTTTACAGCTCTTCTTCTTGTAGCCCTTCTCCTCAGTTC
TTCACACATCCGCACCACAATGGCTTCAAGTTTTTGCGATAGTAAATGTGCAGTGAGGTGTTCAAAGGCAGGA
TTGACGGACAGGTGCTTGCGTTACTGCGGAATTTGCTGCGAGGAGTGCCATTGTGTGCCTTCAGGGACCTATG
GGCACAAGGATGAATGTCCCTGTTATAGGGACAAGATTAACTCTAAGGGCAAACCAAGTGCCCTTAGAAAGAT
CACAAGATTCTGCCATGGCTCT

>OeSCR-like_19
ATCAACGCAGAGTACGCGGGGACTCACAACCACACGACAAAGGACTTTCCTCCTCCGTCTAATTCTTGCAACA
TCAGTATCAAATCCAAACAACAGAGGAGAATGAAGACTTTTTTACTTCACTTTTTGTCTCATTTTTGTCTCTT
CTGCTCAGCAACTCATTTACTGAACCCACCCGTGGCCGCATACTATTACTGTGGACAGAAATGCAAGGAAAGG
TGCGCGAAGGCGGGCGTGGTGGACCGTTGCCTCAAGTACTGTGGAATATGCTGTCAAATATGCAAGTGCGTTC
CTTCTGGGACTTATGGGAACAAGCACCAGTGCCCTTGCTACAGGGACATGAGAAACTCCAAGGGCAAGCCCAA
ATGCCCT

Appendix 1.2. Olive miRNAs and their pre-miRNAs sequences.

miRNA name	miRNA sequence	pri-miRNA sequence
Oeu-miR156_a	UUGACAGAAGAGAGAGAGCAU	AGAGUACUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUCACUAGCAAAACCCACAACAUGUCAUUGCAUUUAUUUAAAAUUCAAAUUUAGGCUAUAUUCACUUAUUAGUAAAUUUUGAAGCUUAAUCCAACGCACUCUUUAUUGUUUGAUCGACAAACUACAUCAGAAGAAAUCCAGAAAUUGACAGAAGAGAGAGAGCAUUUUAU
Oeu-miR156_b	CAGAAGGUAGCGAGCAGCUC	CAGAAGGUAGCGAGCAGCUCGCCCAGCGGAUUUGGGGUAUUCUGCAAAAGAAAUCUUUAAAGCGAAGGAUUAUCCGAGAGGUGAAGCUGUGCAGCUUUCUUUA
Oeu-miR156_c	CAGAAGGUAGCGAGCAGCUC	CAGAAGGUAGCGAGCAGCUCGCCCAGCGGAUUUGGGGUAUUCUGCAAAAGAAAAUCUUUAAAGCAAAGGAUUAUCCGAGAGGCGAAGCUGUGCAGCUUUCUUAC
Oeu-miR158	UGCAAAAUGUAGACAGAGUA	GUUAACUUGUCUAAAUUAUGCACCAUGCAGUGGAAUUGAUCCACUCAAUUGUCUUAAUCAUUGCAUCCAUUUACGAACUUAUUCAUGCAAAAUGUAGACAGAGUAACAA
Oeu-miR159_a	AUUGGAUUGAAGGGAGCUCUC	GCUGAAGAGCUCCUUUCAAUCCAAUGUGGAGGGCCAAGGAGUGGUUAGAGCUGCCAUCUCAUGCAUUUGGUUAUGCUUAAUGUGGAGAUUUUGAGGGUUAAGUUUAACAAGUGCAUGAUAUGAGAGCAACUCUUCCGUUUGCUUCGCCUGUUAAUUGGAUUGAAGGGAGCUCUCAGUGA
Oeu-miR159_b	UUGUAUUGAAGAGAGCAUCA	UUUGUAUUGAAGAGAGCAUCACUAGACAUCAUAGCAUAAGAUAAAAAGUGUGAAGAAUUGAAAAAUAACAUUAGUGAUCUUUAUCGAGAGCAGACU
Oeu-miR159_c	CUUGAGGUUAUGGGUGCUCCU	UCCUUGAGGUUAUGGGUGCUCCUCUAACAGGGACUGAUUUGAGAGCAUAUGUAAUGUCAGGAGG
Oeu-miR161	UAAAUACACUGAUAGUUACUA	CAGGCUAAAUACACUGAUAGUUACUAAAAUUUUAAGCGAAAAUUUCACGUUGGUUAAGUUUGAUAACCAGCUUACAACAAAAUGAAUUUAGCAAUAUACAUGAAAUUUUGCUAAAGAUUUAGUGGGCAUCAGUGUUUUUAACCUGUU
Oeu-miR162_a	UCGAGAAAUCUGUGCCUCCAG	UCGAGAAAUCUGUGCCUCCAGGUUAUUUGAAAUUUAUUAACAAUUUGGAGAACGGAAUCUUGAAU
Oeu-miR162_b	UCAAUAGACCUCUGCCUCCUG	ACGCGGGGCAGAUGUGUUGUUGUGGAAGAAUUGGUUUGGGAGUGCUGCUGUGCUUAUCGGUUCCACCGCCCUUUGGUUACUAUCUGAAAGAGCCGGCUAUAAUCUCUUAUCAUUUAUUUCAAAUGCCUUAUUUUUGCUUGUAGUGAUCUUGUUCUUCUGGGCCAAAUCUGCAUUACUUCUCAAUAGACCUCUGCCUCCUGUUC
Oeu-miR167_a	UGAAUGUGCCAGCAAGAUAUU	GCUGAAUGUGCCAGCAAGAUAUUAUGCCCCACAAUGUAUGUUUCUUUUGAGGAAUUACCCCAAGAGCACCUUGUUCUGUCUGAACAGCGCCCCGGGGCAAUGAUACCUUGCUCAUAUCCAGCCA
Oeu-miR167_b	UGAAUGUGCCAGCAAGAUAUU	GCUGAAUGUGCCAGCAAGAUAUUAUGCCCCACAAUGUAUGUUCUUUUGAGGAAUUACCCCUAAGUAGCACCUUGUCUGUCUGAACAGCGCCCCGGGGCAAUGAUACCUUGCUCAUAUCCAGCCA
Oeu-miR167_c	UGAAUGUGCCAGCAAGAUAUU	GCUGAAUGUGCCAGCAAGAUAUUAUGCCCCACAAUGUAUGUUUCUUUUGAGGAAUUACCCCAAGGGCACCUUGUUCUGUCUGAACAGCGCCCCGGGGCAAUGAUACCUUGCUCAUAUCCAGCCA
Oeu-miR169	AGCCAAGAAUGACUUGCCGA	UAGCCAAGAAUGACUUGCCGAAUCAAUGGCAAGCCCAAAUUUAGUCGGCAAACCAUUCCUGGCUGCG
Oeu-miR171_a	UCAUUGAGACGAGGCCUUGUC	AGGACGGUGUGAUUGUCCAAGAAGAAAUAGCAGUGUUGAUUCAUUGAGACGAGGCCUUGUCCUUC
Oeu-miR171_b	UCAUUGAGACGAGGCCUUGUC	AGGACGAUGUGAUUGUCCAAGAAGAAAUAGCAGUGUUGAUUCAUUGAGACGAGGCCUUGUCCUUC
Oeu-miR171_c	AUGAGCCGAAUCAAUAUCACU	GGUGUGGUAUUGGUUUCGGCUCAUAAUCAUGCAGCUUCUGCAAAAAUGAGCCGAAUCAAUAUCACUCUUG
Oeu-miR171_d	UGAGCCGCGCCAAUAUCACUU	AGAAAGCGAUGUUGGUGAGGUUCAAUCCGAAGACGGAUUUACAAUGGUACAGUAUUUAUGGUAAACGACGAUCUCAGAUUGAGCCGCGCCAAUAUCACUUUCUAU
Oeu-miR172_a	UCAAUCCUGAUGCAGCUGCAG	CUCUCAAUCCUGAUGCAGCUGCAGUGUUGAGGGAUAACCCAAAGAUGUUUGAAUCCAAUGUGCGAAGGGCGAUGGCUGGUGGCUAUGUGGGACAGAGUUUCUUCAAUCGCUGCAUAUAGCGUUUGGAGACUGUGUCAGUCAGUUUCUCAACUCAGUCGACAUUGUAGAUGCUUUAAGUUGAGAGAU
Oeu-miR172_b	UCAAUCCUGAUGCAGCUGCAG	CUCUCAAUCCUGAUGCAGCUGCAGUGUUGAGGGAUAACCCAAAGAUGUUUGAAUCCAAUGUGCGAAGGGCGAUGGCUGGUGGCUAUGUGGGACAGAGUUUCUUCAAUCGCUGCAUAUAGCGUUUGGAGACAGUGUCAGUCAGUUUCUCAACUCAGUCGACAUUGUAGAUGCUUUAAGUUGAGAGAU
Oeu-miR394	UGGUCAUUCUGUCCAGUUCC	AAUUUGGUCAUUCUGUCCAGUUCCGUUAGUAAGAUGAUGUGGCAGUCUAGUUGGAAAAUAACCGGAUUGCCAUGUCACCUUUUUUACUAAUUAAAAAUGGGAUAUGGAUGGAGAAUCGAAUUAU
Oeu-miR395_a	CUGAAGUGUCUGGAGGAAAGC	CACUGAAGUGUCUGGAGGAAAGCAAUACAUACAGCAGGGCAUAUCUAAUGGUUUACCAGAUGUUGAUUCGGUUUUCUAUUCAACUAGGAAGACCGGUUCGGUGCCAUUUGAUAUAAGAUUUGAUGAACAAGCACCAAAAGUUGCCUUGCCCCGAGAUACUAUGGUGCC
Oeu-miR395_b	CUGAAGUGUCUGGAGGAAAGC	CACUGAAGUGUCUGGAGGAAAGCAAUACAUACAGCAGGGCAUAUCUAAUGGUUUACCAGAUGUUGAUUCGGUUUUCUAUUCGACUAGGAAGACCGGUUCGGUGCCAUUUGAUAUAAGAUUUGAUGAACAAGCACCAAAAGUUGCCUUGCCCCGAGAUACUAUGGUGCC
Oeu-miR395_c	UGAAGUGUUUGGGGGAACUC	UAGAGUUCCCCUGACCGCUUCAUGGGGAUCCUUGGCUACUGGAUAAAAUGCAGCCAUUUAGCCUUCUGAAGUGUUUGGGGGAACUCUACA
Oeu-miR395_d	CUGAAGUGUUUGGGGGAACUC	GAGUUCCCUUCAAUCACUUCAUUGCAAUUGUUCUUUUUGAAGAACAAUACCUAGUACUGAAGUGUUUGGGGGAACUCGU
Oeu-miR395_e	CUGAAGUGUUUGGGGGAGCUC	GAGUUCCCUUCAAUCACUUCAUUGCAAUUGUUCUUUUUGAAGAACAAUACCUAGUACUGAAGUGUUUGGGGGAGCUCGU
Oeu-miR396	UCUCACAGCUUUCUUGGCCUU	UUCUCACAGCUUUCUUGGCCUUCCAUGGGGAGUGCAUCACCAACUGGACAAGGUAGGAAUGUUGUGGGAAAU
Oeu-miR402	UUGGCCUAUUCUGCCUCUGCUU	GUACGGGAUUGGCCUAUUCUGCCUCUGCUUUCACAGCACUUCCUCUCAAAGCCAUAAACGGAGGACAGCUUCAUCCACCAGAAUUCAAGAUGCCGUCGCACAAGACAUUCAUGAUCAAGAAGAGCUUGCGAAGAAGCAGAGGCAGAAUAGGCCAAUCCCGUAUGG

miRNA name	miRNA sequence	pri-miRNA sequence
Oeu-miR414_a	UCUUCAUCGUCAUCAUCGUCA	UCUUCAUCGUCAUCAUCGUCAUCAUCCUCUUCUUGGUCAUUAGUUGCUGAUGUCAUCAUAUUCAUGAAACUCUGUUUCUUCUGAGAUGAUGAUGGCUGAGAGGACA
Oeu-miR414_b	UCAUCUUCUUCGUCAUUGUCA	UGAAGAUGAUGAAGAUUAUGGGAAUGUGAUUUGGUCACAACAGUCAUUUGGAUUUGGAUUUGAGACCAAGAAAAUGAAGAAAGGUUUGUUGAGAUCGGCAGGAAGUUCAUCUUCUUCGUCAUUGUCAUC
Oeu-miR414_c	UCAUUUUUAUCAUCCUCUUCA	UCUUCAUUUUUAUCAUCCUCUUCAUUUUCUUCUUCCUCAUUCUGUUCUUCUGCAGAAUCUUGAUUUUGAGGAGGAGGUGGAGGUGGAGGUG
Oeu-miR414_d	UCAUAUUCUUCAUCAUCAGCA	UAUGUAGCGAGGAAGAAGAAUGUGAAAGUCUCCUUCUUCAACUUCUCCUCAUAUUCUUCAUCAUCAGCAGCAUAAU
Oeu-miR414_e	UCAUCUUCAUCAGCAGCUGCA	UCAUCUUCAUCAGCAGCUGCAGCACAAGAAUAACAGCCAUGGAUACAGAAAUCUAUCCGCAACAUGGGGUCUUGUGAUUGUAGCAGCUGGUUCGGAUGGAA
Oeu-miR437	AAAAAGAAAGAAGUUUAACUU	AAAGUAAAGGGAAAAAGAAAGAAGUUUAACUUUAAUUAAGAGGAGCAGACCACAUUUGAAUUUUUCUUUUUUAUUUUUUCAUCUUUAUUUUAG
Oeu-miR444_a	AGCUGGUGCUGUUUCAAGCUU	AUCGAGUGAAGCAGUGCCUGGUUUGAUUUGUCGUAGGGAAACGAUAGUCAGCUGGUGCUGUUUCAAGCUUGAUCU
Oeu-miR444_b	UGCAGAUGAUUUGUCAAGCUU	UGCAGAUGAUUUGUCAAGCUUCAUGUUGAUUUCUUUGAGAAUUUGAUGUCAUAGUUGUCAGUUGUUGCUUGCUUUCAACAAAAGUACUACCACUCUUACUUACAUGCAUAAAAUUAGAAUUUUAUGUGAUUCUUGCAAACGCAUCUGCAUU
Oeu-miR472_a	CAUUUCCUAUUCCUCCCAUACC	UUCGGUUGGGAGCACUAGGAAAUGAAUUUUUUCAUUUACUGAAAUUCUUUCAUUUCCUAUUCCUCCCAUACCGAAGG
Oeu-miR472_b	CAUUUCCUAUUCCUCCCAUACC	UUCGGUUGGGAGCACUAGGAAAUGAAUUUUUUCAUUUACUGAAAUUUCUUUCAUUUCCUAUUCCUCCCAUACCGAAGG
Oeu-miR472_c	AUUUCCUAUACCUCCCAUACC	UUUCGGUUGGGAGGCAUUAGGAAUUGAAUUCCUUUCAUUUCCAGAAAUUCCUUUCAUUUCCUAUACCUCCCAUACCGAAGGU
Oeu-miR473	ACUCAUUCUCACGGCUUCCA	UGGAAGUUUUGAAGUGAGUAAUUCCCUUAAUCGUCGUACAGCUUUAACUUUGUUUGCCAAAAUGUCGAGCAAUGGUAGAAGCUCUUCAGUUUGUAGUGGGAAUUCUGGAGAAAGCCAUAGAACAGGCCUCAAUCCUUUCUUGUACUCAUUCUCACGGCUUCCAUC
Oeu-miR477_a	UUCUUCUUUAGAGGCUUCCA	CUUCUUCUUUAGAGGCUUCCAAAUUUAUGAAGCAAAGGUUUCUUCUUCUUCUCGACUUGGCUGCUUUGAAGGUGAGGUU
Oeu-miR477_b	AGAUGCCCUUGUGGAAGACGA	AGAUGCCCUUGUGGAAGACGAGCUGAUGGCAUCCAGCUCUUCUGCUCCCAUGGUAUCUGA
Oeu-miR477_c	UUCUCUUUCAAAGGCUUGAA	UUCUCUUUCAAAGGCUUGAAAGGUGAGAGAUAGACUGAGGACUCCACCCUUCACAUUCAGAUCAGGAUCUUCCGGCCAUGGAUAUUCGUGCUUUUUCACUUUCCCAACCUUUGCAGAGGAUA
Oeu-miR482_a	UCUUUCCGAGUCCUCCCAUUCC	UUGGAAUGGAGUUCUGCGGAAGGAAGGAAUUUCCUCCUGUUUCUUUCCGAGUCCUCCCAUUCCAACA
Oeu-miR482_b	UCUUUCCGAGUCCUCCCAUUCC	UUGGAAUGGAGUUCUGCGGAAGGAAGGAAUCUCCUCCUGUUUCUUUCCGAGUCCUCCCAUUCCAACA
Oeu-miR482_c	CAUUUCCUAUUCCUCCCAUACC	UUCGGUUGGGAGCACUAGGAAAUGAAUUUUUUCAUUUACUGAAAUUCUUUCAUUUCCUAUUCCUCCCAUACCGAAGG
Oeu-miR482_d	CAUUUCCUAUUCCUCCCAUACC	UUCGGUUGGGAGCACUAGGAAAUGAAUUUUUUCAUUUACUGAAAUUUCUUUCAUUUCCUAUUCCUCCCAUACCGAAGG
Oeu-miR482_e	CAUUUCCUAUACCUCCCAUACC	UUUCGGUUGGGAGGCAUUAGGAAUUGAAUUCCUUUCAUUUCCAGAAAUUCCUUUCAUUUCCUAUACCUCCCAUACCGAAGGU
Oeu-miR482_f	UCUUUCUUGUACCUCCCUUC	GAUGUGGGUAUGUGCAGGGAAGGUGAGAUCAGUGAAGUAUAAUCUUUCUUGUACCUCCCUUCGUCCA
Oeu-miR482_g	UCUUCGUUGUUUCUCCCUUU	GUGAGAGAGAGAGAGCAAGAGAAGAGAACAUGAUCGCACUUGUCAUUCAACUUGUCUCUUCGUUGUUUCUCCCUUUUUGCUU
Oeu-miR482_h	UCUUCCUUUUUUUUCUCAUU	UCUUCCUUUUUUUUCUCAUUGAAUAGAUUUGUUUCUAUUGGUUUGGCUCGGAUUGGUAUCAUAAAGAAAGGGGAAUGGCUCGGCUAUCCCACUUAGCCAAGCCAGAAAAAUAGAUUAGAUAUAAAUGAGUAAAAAAAAAAGAAA
Oeu-miR482_i	UCUUCAUUCUUCCUCCCUUG	UUUUCUUCAUUCUUCCUCCCUUGUCUUCAUUAUUUAGAAAAAGGGACGAAGAAUAAAGGAAAGG
Oeu-miR482_j	UCUUCUUCGUUCUUCUCAUU	UCUUCUUCGUUCUUCUCAUUUCCAUCGCACUCAACUCGCACGUUUCAUCUCAGUCCGAGGAACCCAGCUUGAAACUUCCGACUGAUGCAUCACUUGAAUGGCCAAUGGGGAUGUCGUUUUACGAGAGCGAUUUCGGAGAAUUUAAUGAAGAAGGACAAGAAGACG
Oeu-miR529_a	CUAUUCCCUUUCUCUUCUUC	GAAGAAGAAGAAUGAAAGUAGAGUACAAUCACUGAUAUAUCAUCCAUUGCUAUUCCCUUUCUCUUCUUCUUCCA
Oeu-miR529_b	CUAUUCCCUUUCUCUUCUUC	GAAGAAGAAGAAUGAAAGUAGAGUACAAUUACUGAUAUAUCAUCUAUUGCUAUUCCCUUUCUCUUCUUCUUCCA
Oeu-miR530_a	UACAUUAGCACCUGAGUCUC	GAGCACUCAAGUCUCAAUGUAACUUUGCUUAAUAGAGAUUCUUUUAGUUCUUCACAGUUGUCACAUGCCUCAGAAGCUAUUUGGACAACAUCAUGGGAAAUUCUUGAAGCCUCUGUCUUCAGAUGUUGGAUCUUUGCAUUGGUUUCUUGUUCAAGGCGCUUUACAUUAGCACCUGAGUCUCCA
Oeu-miR530_b	UGCAUUUGGUGCUGCAACUU	UGCAUUUGGUGCUGCAACUUCAGCUUAUCAGAUUGAAGGUGGAUGGAAUGCAGGAACUAAAGGCAUU
Oeu-miR781	UUAGAUUUUUCUGGUUAAAUA	GUUUAGAUUUUUCUGGUUAAAUAGCAAGCUUGUAGAGACAGGGGCAUGCGAUGUUCCCUCAUCUUACCAUAAGUUUAGAUUCAGCACCGUUCCUAUAUACAACUGUGACCAGGAGGGUCUAAAUGC
Oeu-miR783_a	AUGCUUUGUGGGCUCAUUUUC	AUGCUUUGUGGGCUCAUUUUCAGGUAGGACAAUGCAAAACUGGAAAAGGCAUCACGGGAGCAUAC
Oeu-miR783_b	AUGGUUGGCAGCUUCAUUUUC	AUGGUUGGCAGCUUCAUUUUCAUGACCAUUUUUCAGAACAUCAUCAUUCUUCGGAGAAAAUUCACCCUCUUCAUGAGAUGAGCUGCCAUCUGUCU
Oeu-miR813_a	GGUUUUUGGCAUGGGUUUGACC	GGUCAAAACCACGGCCAAAAACCAACAAUGCCACUUGACCAUAACCAAAAAACAAUGAUGCCACGUGGAGCCGGCUCCAAAACAGCCACCCCUCCAAGUAGGGGGCGCCAUGUUGGUUUUUGGCAUGGGUUUGACCCU

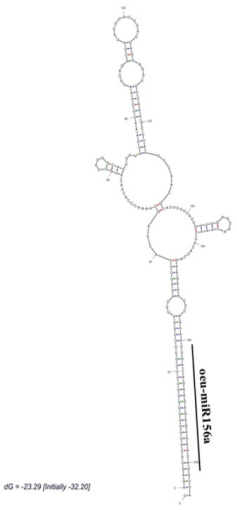
miRNA name	miRNA sequence	pri-miRNA sequence
Oeu-miR813_b	GGUUUUUGGCAUGGGUUUGACC	GGUCAAAACCACGGCCAAAAACCAACAAUGCCACUUUGACCAUAACCAAAAAACAAUGAUGCCACGUGGAGCCGGCUCCAAAACAGCCACCCCUCCAAUAGGGGGCGCCAUGUUGGUUUUUGGCAUGGGUUUGACCCU
Oeu-miR813_c	GGUUUUUGGCAUGGGUUUGACC	GGUCAAAACCACGGCCAAAAACCACAAUGCCACUUUGACCAUACCAAAAACAAUGAUGCCACGUGGAGCCGGCUCCAAAACAGCCACCCCUCCAAUAGGGGGCGCCAUGUGGUUUUUGGCAUGGGUUUGACCCU
Oeu-miR816	GUCAGAUAUUAUGCUACAAC	GGUUGUUGGCACAAUAUCGAGAAAUAGCUGCCCUUGCUCUAUUUGGGUCAGAUAUUAUGCUACAACUCA
Oeu-miR825	UUCUCAAGAACUUUUAUGAAC	UCAUUUAUAGUCGUUCUUCAGAGAAAUCCUGGAUUGAAUGCUGCCUUUAUCCUUCUCCAAGGUAGCUAAUUCCUUCUCAAGAACUUUUAUGAACUGAUC
Oeu-miR834	UGGUAGCAUUAGCGUUGGUGA	CGUUGUUGGUGAUGACGUCCCCGGUGGGGUGGUGGUAGCAUUAGCGUUGGUGAAAAUCGCAGCAUCAGAAGAAUUAGCAGGUAAGUGUAGAAGCUUGGGGCAUUCACUGACGCUAAUGCUACCACCACCCCACCGGGGACGUCAUCACCAACAACGCC
Oeu-miR837_a	CCUUUUUUCUUUUUUUUUUUA	CCUUUUUUCUUUUUUUUUUUACUCAUUUUAUAUCUAAUCUAUUUUUUCUGGCUUGGCUAAGUGGGAUAGCCGAGCCAUUCCCCUUUCUUUAUGAUACCAAUCCGAGCCAAACCAAUAGAAACAAAUCUAUUCAAUGAGAAAAAAAAGGAGAGAGAGGGAU
Oeu-miR837_b	UUCAUUUUCUUGUUCUUUUCG	UUUCAUUUUCUUGUUCUUUUCGAAAACCAACAACGUUGCCUCCUCAACCCAACGUCAAACAGAGCAAAAUGCAAAGAAUCGAAAGAAGAAGAGAAGAAAAA
Oeu-miR837_c	UAUUUUUACUUGUUUUAUUCA	UAUUUUUACUUGUUUUAUUCACCGAUCGUACCAAAGUAAACAACUAAGCAAGAUUGAUGAGUAAACAGCUACAGAGUACG
Oeu-miR838_a	UUUUCUUCUUCUUCUUCUCCU	GGGGAGAAUUGAAGAAGAAGAAACAAGAAUUCUUGAUUAUUUUCUUCUUCUUCUUCCUUCU
Oeu-miR838_b	UUUUCUUCUUCUUCUUGCCAA	GUUUUCUUCUUCUUCUUGCCAACAGCGGCUGCUGCUUUGACCAAAGCUAGAUGAUCAAACUGUGCAGGCGAGAAGGCAAGAGGACUU
Oeu-miR847	UCAUUCUUCUUCUUCUUCAUG	GCAUGGAAGAAGAAGAGAAAGGGAAUAGCAAUAGAUGAUAUAUCAGUAAUUGUACUCUACUUUCAUUCUUCUUCUUCUUCAUGUGA
Oeu-miR854	GAGGAGGAUGUGGAGGUGGAG	CCUCUUCCAACCCCGAUUCCUCUUUCGAUUAUACUACCACUGCGAAUGCUAAUGAAGACGGUAAUAAUAGUAACUCGAAAGGUGGAGGAGGAUGUGGAGGUGGAGGAGGUA
Oeu-miR857	UUUUUUUUGUUGAAGUUGUGU	AUUAUUUUUUUUGUUGAAGUUGUGUAAUAUUAUUAAAUUUAAAAUGUGCGGUUACAGUUUCAGCUGAAAAGGAUGGUAA
Oeu-miR863	UUGAGAGCAAGAAGAAGUAAU	AGUUAACUCUUCUGCCUCAAGUUGAGAAUUGACAGCAAAAAAGACAAGACCUUGAGAGCAAGAAGAAGUAAUUGA
Oeu-miR948	CCAGGCUGUGUUGGAGCUGG	GGUACACUUCCGUACGCACCUGCUGGUUGAGGAAUGGACUGAUCAUAGCCAGGCUGUGUUGGAGCUGGCUGA
Oeu-miR950	UCAGCUCCUCGGUGGUUGUG	UCAGCUCCUCGGUGGUUGUGCCCCUCUUGUUAAACAGAGGCACAGCACCAUAGACGAGCUGACU
Oeu-miR952	AAGUGAGAAUGAAAAUGGUG	GUAUAUGAUUUCGAUUUUACUACUAAUGAAAAUUGCCAGAAAAGCAAGAUUUACUUCGUUGCUUGGUCACCAGAUACCUCUAAAGUGAGAAUGAAAAUGGUGUAUGC
Oeu-miR1023	CCACUCUCUCCGCAUCCAUUUC	AGAAGAGAUGCAGAGGAGUGGAAGUUGAUGCAUCUCCCUCUCUUCCACUCUCUCCGCAUCCAUUUCUAC
Oeu-miR1027	UUUUUAUCUUCUUUUCCUAUA	UGCUUUUUAUCUUCUUUUCCUAUAUUUUGGAUUUGAAAAAAUUAUUAUUUUCCAUAUCAGAUAGAACUAAGGAUAAAAGAAGUGGAGGGUACA
Oeu-miR1038	GAGGUGCCUUUGCCCCAAAG	CUUUGGAGGGAGGGUAUUUCGUCGGGGGAUUUGAUUUGUUCGAACGGGGUAUUCAAUAUGAGGUGCCUUUGCCCCAAAGUA
Oeu-miR1039	UCUUUGGGUCAUUGUGUGCUG	AUUCUUUGGGUCAUUGUGUGCUGAAAGCAGAAUUUCCAAAAGGUAGAAAGGAGCUGGCUGUUUCUCUAUUUCAGACGAUUGUGUUGAUGCUAUUUAACAAUGCACAAAACCUUAGCUUACAAGAUAUUAAGGAAUCU
Oeu-miR1040	UCAACGCAGAUGGCCAUGUUC	UAUCAACGCAGAUGGCCAUGUUCUAUUCGGAGGAAUAAAAUAAAAAUUGUCUUUCGGAGAGAUGGCUGAGUGGUUGAUAGC
Oeu-miR1042	GCUGGGCUGGAACGAGAUAGU	GCGCUGGGCUGGAACGAGAUAGUAGCUGUUCUACGCAACCCCGUGUACUUCAUCUUCCUCAUCCUUUUGGGCUCGGCGCCU
Oeu-miR1044_a	UAGUAGAGCAUAUAUGUAUU	CCUAGUAGAGCAUAUAUGUAUUAUUUUGAUGAAUUAUAUUUAAGAUUUAUCUUGCAUAUAAUUAUGUAUAUCUCUAUAUGGUC
Oeu-miR1044_b	UAGUAGAGCAUAUAUGUAUU	CCUAGUAGAGCAUAUAUGUAUUAUUUUGAUGAAUUAUAUUUAAGAUUUAUCUUGCAUAUAAUUAUGUAUAUCUCUAUAUGGUC
Oeu-miR1052_a	UUGCUUUCAUUAAUUGUGAUA	UGAUUGCUUUCAUUAAUUGUGAUAGAAAAAUAAUUUAAACAUAAAUACGUCUUUGUUUUGAUGAUCACAAUGUUAAUGAAAGGAUUAUU
Oeu-miR1052_b	UUGCUUUCAUUAAUUGUGAUA	UGAUUGCUUUCAUUAAUUGUGAUAGAAAAAUAAUUUAAACUAUAAAUACGUCUUUGUUUUGAUGAUCACAAUGUUAAUGAAAGGAUUAUU
Oeu-miR1056	UGGAUCUUUGCAUCUAAUCUC	GAGAGGAUUGGAAUGUAAGGUUUCAUAUCAUUCGUGAACUUCGAGCUGUUGUUGAAUCCAUAGAAAGUCUGAGAGGUGCAACUGUGGAACCAUUUGGAUCUUUGCAUCUAAUCUCUUCAC
Oeu-miR1063	CGUAUUGGAGUACUGCCUUUU	UAAAAUGCUUGUGCCCGAUAGCUGACGUGUUAUGUUUAUAUCGUGUGUGUGAGCGUAUUGGAGUACUGCCUUUUAAG
Oeu-miR1067	GAAUACUGAAGUUUGAUAUCA	UGAUUAUGGGACUUUGUGUUUUUGGAUAUUUUUUGAAUCGAAGGUGCAUUUUGGAAGGGAUUGGAGUUGUUUUGUGAAGAAUACUGAAGCUUUGAUUUCGUUCGCAGAUUGCGCAUUUUAAAGCGGAUUGGAGUUGUGUUAAGAAGAAUACUGAAGUUUGAUAUCAUU
Oeu-miR1097	UAGCCAUAUUUCUUGCUGGAA	UAGCCAUAUUUCUUGCUGGAACCUCGCCUCGACAGGUUUCGGGCCAGAACUGCGGCUGUG
Oeu-miR1122	UACAUACAUACGUAUAUAUA	AUACAUACAUACGUAUAUAUACAUAGAUACACACAUAUAUAUAUGUGUGUGUGUGUGUCUAUAUACAUGUAUAUGUGUGUGUGUAUAU
Oeu-miR1313	UACCAAUGAAAUUUCUGUUAG	AAAAUUUUACUGAAGAAUUUUCCGUUGGUAAAGUGUGAAUCAUUUACCAAUGAAAUUUCUGUUAGUAAAAAUUUUAA
Oeu-miR1320	UGUUAAAUUGAUUUGUUUCAA	GUCAAACAAAAGUGAUCUAACACNCCUUUUUUUAAAAAGUUAUCACACUUUCAUUAGGUCAUAACGUUCGUUUGAAGAGUGUUAAAUUGAUUUGUUUCAACAU
Oeu-miR1428	AAUACACAGCCCCUAUUUGGUG	GUAAUACACAGCCCCUAUUUGGUGGGAAUUAAAAUCAGAUUUUCAGUAAAAUGUGGUUGAAAAGUUGUGGAACCCAUGAAAAAUGUGGUUAAAAUUAAUUGUUAAAUUUAACUUAGAUUAAAAUCUGAAUUUUAUUCAGUACCAAGCAGGGGCAUUGUUUAUAA

miRNA name	miRNA sequence	pri-miRNA sequence
Oeu-miR1435_a	UUUCUGAAGUUCAACUUGUU	AACAGCUUCGACUUGAGAAAUUCCAGUGGGAACAUUGAACCCAUGAGGAAAACUAGGGUUCAUCAUUGGACCAUUUAGAUUAUCAUUUCUGAAGUUCAACUUGUUGG
Oeu-miR1435_b	UUUCUUGAGGUAAAAUUUUC	UUUCUUGAGGUAAAAUUUUCUAUUAGUGAAAAAAGUUUCUAUUUGCGUUUGGCGCAUUGGACUAAAUUGUAUUGAGAAUUUAGUAUAAAAUAAUAAGAUGGUAUAGGUAGAUGAUAUUAUAGUAGGAAAAA
Oeu-miR1444	UGCACUUUUGGUUAAUGUUC	UGCACUUUUGGUUAAUGUUCUUCGUUUCGCCAUUAAAGGAGCUCCAUGAAGAAAGCCUGAGCUUGUAUGGAGAAUGUUGACCAGAGUGUGUAUC
Oeu-miR1450_a	UUCAGUGGCUCGGUGAAGUUGC	GUAAUUUCACCGAGACGAUUGAGCUUCAAAUUGGGCUGAAAAAUUAUGAUCCCCAGAAGGAUAAGCGAUUCAGUGGCUCGGUGAAGUUGCCU
Oeu-miR1450_b	UUCAGUGGCUCGGUGAAGUUGC	GUAAUUUCACUGAGACGAUUGAGCUUCAAAUUGGGCUGAAAAAUUAUGAUCCCCAGAAGGAUAAGCGAUUCAGUGGCUCGGUGAAGUUGCCU
Oeu-miR1450_c	UUCAGUGGCUCGGUGAAGUUGC	GUAAUUUACACUGAGACGAUUGAGCUUCAAAUUGGGCUGAAAAAUUAUGAUCCCCAGAAGGAUAAGCGAUUCAGUGGCUCGGUGAAGUUGCCU
Oeu-miR1450_d	UUCAGUGGCUCGGUGAAGUUGC	GUAAUUUCACCGAGACGAUUGAGCUUCAAAUUGGGCUGAAAAAUUAUGAUCCCCAGAAGGAUAAGCGAUUCAGUGGCUCGGUGAAGUUGCCU
Oeu-miR1450_e	UUCAGUGGCUCGGUGAAGUUGC	GUAAUUUACACUGAGACGAUUGAGCUUCAAAUUGGGACUGAAAAAUUAUGAUCCCAGAAGAUAAGCGAUUCAGUGGCUCGGUGAAGUUGCCU
Oeu-miR1450_f	UUCAGUGGCUCGGUGAAGUUGC	GUAAUUUCACUGAGACGAUUGAGUCUUUCUAAAUUGGGCUGAAAAAUUAUGAUCCCCAGAAGGAUAAGCGAUUCAGUGGCUCGGUGAAGUUGCCU
Oeu-miR1508_a	GAGAAAGGGGAGGAGCAGUCG	UUUUCACUGUCUUCACUCUUUCUCGCAAUCGGGAGAUCGAAUUUGUCCUUUUUGCGUAUAGAUACGUGAUUCACGCACAUUUUUACCAUGUAUAGAGAAAGGGGAGGAGCAGUCGGGAAGG
Oeu-miR1508_b	GAGAAAGGGGAGGAGCAGUCG	UUUUCUACUGUCUUCACUCUUUCUCGCAAUCGGGAGAUCGAAUUUGUCCUUUUGCGUAUAGAUACGUGAUUCACGCACAUUUUUACCAUGUAUAGAGAAAGGGGAGGAGCAGUCGGGAAGGG
Oeu-miR1510	UUUUGUUUUGUCAAUUCCACC	UUUUUUUUGUUUUGUCAAUUCCACCAUACUAUUGCAUAAAAAAACUUGUCAUAUUACAAUAUUUGGUUGGGGUGGGAUAGAAACAAUGCAGGAGAAGCC
Oeu-miR1511	AACCAGUCUCUGUUGCCAUU	CAACCAGUCUCUGUUGCCAUUGAUGCUAGUGGCUAUGACUUUCAACUGUACUCUAGCGGGAUCUUCUCUGGGUAUUGUGGGAAAAAUCUCAAUCAUGGCGUAGCGGCUGUUGGC
Oeu-miR1518	UGUGUUGAAAAGUUGAUAUGA	UCAUGUUGUGACAUUUUGCUGUUUGUUGUAGCUUAAAAAUUAGUAGAUUUUGAUAAAAGUUAGAAGAAAUGUGUUGAAAAGUUGAUAUGAAU
Oeu-miR1521	CAGUUAAUGGAGAAGGUUGA	UCAGCUUUUCUAUUCGCUGAGGCAAACUCAUUGCUGUACAUAGGGAGGCUUGCUACCGUUUCUGCAGUUAAUGGAGAAGGUUGAAG
Oeu-miR1523	AUGGGAGAAUUCUGAGCUGA	AAUAAUGGGAGAAUUCUGAGCUGAAAACUUUUCUUUUAUGUGUAUUAACAACUCUGCAUAAUCUGCUAAAGCUUGUGCUGAGUUGAAAUAUCCACGAGCAUUCUCAUCAUUCAUGGCUUCUUCCAUUGUUCC
Oeu-miR1527	UAAUUCAGCCUUAGAAAAGC	AGUAAUUCAGCCUUAGAAAAGCAAAGGCAUAUCUGCCGUAGGAGCUGCUUGGUUUUGCUGGGGAGGGAUGACUGU
Oeu-miR1533_a	AGGAUAAAAAUUAUAACGA	UAAUUGUAUAGUUAUUAUACCAAUAGCAUCUGUGGACACAAGUUGGCCUGCAAAAGGAUAAAAAUUAUAACGAUUACA
Oeu-miR1533_b	AGAAAAAGAAAAAUAAUGA	UUUUCUUUUGUUUCUAUUUUCUCUGUUUUCUUAAUUUAUGUGGGUAUUUAGUAGAAAAAAAAGAAAAAGAAAAAUAAUGAAUAAAGGGAAAAA
Oeu-miR1533_c	AAAACAAAAAGAAUAGUGA	CAAAAACAAAAAGAAUAGUGAUCACUCACAUUUACCAAACAACGCAAGAAAUUUACCUGCUUCCUUGUGAGGCUGAAUGCUAUCAAACAUUGACUUCUACUGGGCUUUUGGUUUUUGAA
Oeu-miR1533_d	AUGAUAUAAAUUAUAUUGA	AAAUGAUAUAAAUUAUAUUGAGAAAAUUAUCGUAAAAAAUAAUAAACAUUAUUAAAAUAAUUAUUUUUAAAAAUACAAGUUAUACGCAUUUUUUAACGCACGCAUUUGAUCAUGUACGUUUAAAAUUUUCGAUAUUAAUUUAGAUCAUUUGA
Oeu-miR1533_e	AUAAUUGAAAAAAUAAGGA	UGAUAAUUGAAAAAAUAAGGAAAUUAAAAAAUUUGAGAUUUAAUUUUGCAUUUGGGGAUUCUGUGUUUUUUUUAGUUUAGUCAGU
Oeu-miR1533_f	AGAAUAAAAAAUAUAAUUA	AGAAUAAAAAAUAUAAUUAUUACGUGCUCCACACAAAAAAUAUAUUUUCAAAUAUGGCUACUUAAUGUGAUCAAUAAUAAUCAUAUUUUUUAUUUUAA
Oeu-miR1533_g	AUAAUAAAGAAAAUAAUGG	AUAAUAAAGAAAAUAAUGGUUAUGAAGCCGGUAGAACUGGAAAUUUUCUGAUCACAGUUUUGCAAGUUGUAGUUUUAAAGCAGUCCUAGACGAGCUGACAGCGGCAUGGUUUAUCUCAGAAUGUUUAAUUUCUAUGUUGUAA
Oeu-miR1533_h	AUAAUGAAAAUAAAAAAAA	ACUUUUUUAUUUUUUAUUUUUUAUUUUUUUAAGCUAGCAAUAGAAUGGUUGCAGAUGAAAGUAAUAAUGAAAAUAAAAAAAAGUUA
Oeu-miR1535_a	UUUGUUUGUGGUGGUGUGG	GGCUAUGCCUUGCAAGGAUAAGGUCUGCCCCCAGGGAUUCAGGUGCGCGGCCAUCCCUCAUCAGGGUGUUCUCUUUGUUUGUGGUGGUGUGGCCUC
Oeu-miR1535_b	UUUGUUUGUGGUGGUGUAG	GGCUAUGCCUUGCAAGGAUAAGGUCUGCCCCCAGGGAUUCAGGUGCGCGGCCAUCCCUCAUCAGGGUGUUCUCUUUGUUUGUGGUGGUGUAGCCUC
Oeu-miR1535_c	AUUGAUUGUGGGGAUGACU	ACUUCAUUGAUUGUGGGGAUGACUGGGAAUUAUACGCACACAACAGCCUUGGCGGCAAACUGGUAUGAUUGAUUACUUACAGAGUUCUUGUUGAAGUAUUGCUUCUCAUAUGUUCCUAUAAGCAAGUAGUUU
Oeu-miR1535_d	CCUGUUUGCUGUGAUCUCU	UACCUGUUUGCUGUGAUCUCUGGUUGAGAAGCCUGUGGAGGAGGAGGAGGAGGAGGGGGCGGCGGAGGCGGGUAGU
Oeu-miR1852	AGAUGGAUUCAGGAAGCAUGU	UUAGAUGGAUUCAGGAAGCAUGUGACGAGCUUGGCGGCUCUAGACUUCUUGGUCUGUUUGAAG
Oeu-miR1862_a	UGAAGGUUGGUGUAUUUUGG	GAUUUAAAGUCAUCCAAUCUUCUCGUCGACAGGAACUGGAAUGUGAAGGUUGGUGUAUUUUGGUCUC
Oeu-miR1862_b	UUGAGGUUGUUUAAUUUUGA	CUUUUUCUAAAAAUGAAAACGAAUUAAACCCUAUAAAUUGUAUAUUUGUAUGACUAUGUGAGAAAAUCUGUGGAGGGGAGGGAUUGAGGUUGUUUAAUUUUGAGGAGAAGAA
Oeu-miR1888	UUGUUUAAGAUUUGUGAAGAU	GACUUGUUUAAGAUUUGUGAAGAUUUGGAAAAUAUUGAUGGUCUUCACAUGAUUUUCAAAUUAGUUAG
Oeu-miR2083	ACAUUGACGAGUGAAACCAGA	CCCACAUUGACGAGUGAAACCAGACUAAAGGGAGGAAGUGGAUUGUCAUAUAUAUGAGAUUGCUACACUUGGUUACCACUCACUAAUGUGGGUU

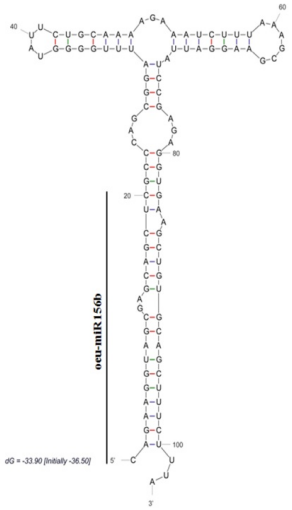
miRNA name	miRNA sequence	pri-miRNA sequence
Oeu-miR2093_a	UUGCUUUAAUUUGAAGAAGA	GAAUUUGCUUUAAUUUGAAGAAGAUUUUGAUAUUCUUGAUUUAUCAUGGACUUGCUUUAAUUUAAAGAAGAUUUUA
Oeu-miR2093_b	GUGCAUUUCUUGGAAGGGCA	GAGUGCAUUUCUUGGAAGGGCACCACUAGUGAAUCUUCCUUGAGGGUUUUGAUUGGAGGGUGGUUGGCCAUGGGAAUCACUUUG
Oeu-miR2592	GAUUCUUUGAGUCAUGUUGUU	GAUUCUUUGAGUCAUGUUGUUCAGUCUGGUGACCACGAUCUGGGACUGUCAUCAAAUGGUAUAUUUGAGAUCAGUUGCUCAAAGGAUAUACAAAAUUCAGGGAAUUAU
Oeu-miR2595_a	UUUAUUUUCCUUUUUAUGUCU	ACAAGAGAAUGAACCAGAAGAGGACGAUGGACCAGACGAGCCCCUGAAUUUUUUAUUUUAUUUUCCUUUUUAUGUCUUAUUCUUUCGUUG
Oeu-miR2595_b	UACAUUUUCUUGUAUAUCUCU	UCUCAUACAUUUUCUUGUAUAUCUCUACAAGUCCCAUGCCUAUAAAUAGGAUAGAGAGGCUGAUACGUGAAUGGGUGAGCAAGAGAAUGUGGGGAAC
Oeu-miR2612	UGUUGGUGUCAACUCGUACGG	AUGUUGGUGUCAACUCGUACGGGCUCAAGAACAGUCUAAUUGGAAUUGAUUCUCCUCUAACUGGAGCACCAUCCAUUAGCUCAUAUUUCGCGAGUGUCUCUGUCUCAACAUGGGUGUUCACACCUGAACCAGUUGAUUCUCGACGUCU
Oeu-miR2628	CAAGAAAAAAAUGAAGAGGAA	UUUUUUGUUUUUUUUUUUUUUUCUUUAAACAAAAAUCAAUCUUUAACGGCAUGAUAGCUAUCCAAGAAAAAAAUGAAGAGGAAACAGAGAAUU
Oeu-miR2630	UAGUUUUGUUCCUUGCUAAUU	CGCAGAGUACGCGGGGGAGUAAGAUUCAGUAAGAACAAUUUAGUUUCUGGGAAAAUUUUAGCUGUAUCUAUUUGAUAGUUUUGUUCCUUGCUAAUUGCGAU
Oeu-miR2633	UGAAAUUUUGCUAAAGAUUUA	UUACUAAAAUUUUAAGCGAAAAUUUCACGUUGGUUAAGUUUGAUAACCAGCUUACAACAAAAUGAAUUUAGCAAUAUACAUGAAAUUUUGCUAAAGAUUUAGUGGGC
Oeu-miR2641	GUUCGAACCUUUACAUGUAU	CAUAUAUGUAAUAGUUUGAUUCAGGGGCGGUGGGGGACUCUUAAGGUCGGAGUUCGAACCUUUACAUGUAUGUA
Oeu-miR2651	UUUGAGAGGUAUGGGUGCAUU	GAUGACUUAUGCUUCUUUGGUUCUAUUUUGUACGGUAUAUUACGUGGAGUGUCUGAUUUGAGAGGUAUGGGUGCAUUAG
Oeu-miR2654	AUUCAGAGAGAAAGUAUCCG	AUUAUUCAGAGAGAAAGUAUCCGUUUCAUGGUGAAAAUCAGACCAAGUUAGUCUGAGGACGAAGUGUGUAUUUUUUUUUAAAUAAUAC
Oeu-miR2657	UGGUUUUUCUUCUAUUUUGUUG	CGACAGAAUUCAGAAGAAAGUUUAAUGAAGAUGAAGGUGAAAUAUUGGCAUCAACCGUUUCCUGAUUAGAUCUGGAGCUACUGAAAUCAUUGGUUUUUCUUCUAUUUUGUUGUA
Oeu-miR2868	UUUGUGUUUUGUAGUAGAAC	GGUGUUUGGUGAAGAAUAUAAAGAGCUGAUGAAGCUGAAUUUUUUUUGUGUUUUGUAGUAGAACCGG
Oeu-miR2873	GUGGACUUGUGGUUUGAUAUG	CAGUGGUAUCAACGCAGAGUACGCGCCAGAGAACAGUUGAGCACUUAGUAGCAGUGGACUUGUGGUUUGAUAUGCUGGG
Oeu-miR2911_a	CGGCCGGGGACGGACUGGGA	UCCCAGUCCCGAACCCGUCGGCUGUCGGUGGACUGCUCGAGCUGCUCCCGCGGCGAGAGCGGGUCGCCGCGUGCCGGCCGGGGACGGACUGGGAAC
Oeu-miR2911_b	GGCCGGCGAACCGGCUGGGA	UGGCCGGCGAACCGGCUGGGAGGAUAGUGAGUGGGAAGAGCAGAAAGAAUCUUUGACGUCAUUUCCCGGCCAAUCGCCGCCGGU
Oeu-miR2913	CAGGUCGGGGAUUGAAAGUUG	GCAGGUCGGGGAUUGAAAGUUGCUGGACGAGGACGCUUGCUUCGAGCACGUUGUGGACCCCGGUGACUUGGACUGAGAGGCAUAGUCUUCGUCCAGCAACUUUCAAUCCCCGUACUCUGCGU
Oeu-miR2919	GCGGGGGGGGAGGGAAGGA	GCGGGGGGGGAGGGAAGGAGAAAAGAAUUUUCAACGGUCCUCUACUCCUCUUUUCCUUUUCCUUUCCCUUGCAU
Oeu-miR2923	AGACAAAAAAAUAAAAUAAAAAA	UUUUUUUUUUUUUUUUUUUUUUUUUAGUUGGAAGCACUCAAACAGAUCGAUUUACAUUUAAGCAAGUUACAAAAUCCUUACAGUGCCACCAACAAGACAAAAAAUAAAAUAAAAAAAAAUA
Oeu-miR2925_a	UGGCGGCCGGGGUCUUACU	GCGGGGAUUUUGGGGUGGCGGCCGGGGUCUUACUCCCAUCGUCAUCGCCUGCACCUAAAUCAUCUUCCCCAAAAGCUGCACCAUCGCCGAAGGAAAGUAAGACCCCGGCCGCCACCCCAAAAUCCUCGCCA
Oeu-miR2925_b	GGGCGGCCGGGAGCAUCGU	GGGGCGGCCGGGAGCAUCGUUGAUCGGUUGGUUGAGGCUUCUAGUCCCCGUGAUAUGCUUUCAAUUCUCUGUGAGGCAUUAGGUUCUCGAGAUGAGCUGUUCGCUGUUCCUC
Oeu-miR2925_c	GGGCUGCCUCCGGCUUCGU	AGGGCUGCCUCCGGCUUCGUGCUGCUUGCGAGAAGUUGAAAAAGGUUCUCAGCGCAAAUCCUGAGGCACCUUUG
Oeu-miR2925_d	UGGUGGCCGCCGCCUUAGU	GUGAAAUCCGGCGGUGGCGCCGGUCUUCAUUUUCCUCAUGCUAGUGGUGGCCGCCGCCUUAGUUUUAUCC
Oeu-miR2925_e	UGGUGGCCGCCGCCUUAGU	GUGAAAUCCGGCGGUGGCGCCGGUCUCAUUUUCCCCAUGCUAGUGGUGGCCGCCGCCUUAGUUUUAUCC
Oeu-miR2928_a	AAGAUGAGGACAUUUUGGUC	GAAGAUGAGGACAUUUUGGUCCUUUUCUUAUUACUGUGCUUAAGGUUGUGGCUAUGGCGAACGUGGCUUCGAUCUGUAGAAGAGAUCGAAGUGUUCUCUCUUCCG
Oeu-miR2928_b	AAGAUGAGGACAUUUUGGUC	GAAGAUGAGGACAUUUUGGUCCUUUUCUUAUUACUGUGCUUAAGGUUGUGGCUAUGGCGAACGUGGCUUCGAUCUGUAGAAGAGAUCGAAGUGUUCUUUCUCUUCCG
Oeu-miR2928_c	GAGAUGACGAGAUUUUAUUG	UGCUGAGAUGACGAGAUUUUAUUGGCUGAUUUAAUAUUUUUUUUUGUCCAAAAUGGUGUGAAUUCUUGUCAACUCAGCAUU
Oeu-miR2937_a	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUUCAUUGUGGUUUACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2937_b	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGACAAUGCUAUGGGGCAAGUAUGUUUUUCAUUGUGGUUUACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2937_c	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCUAAGUAUGUUUUUCAUUGUGGUUUUACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2937_d	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUUCAUUUGUGGUUCACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2937_e	AUGAGAGAUGUUGAAGGGGGC	AUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUUCAUUGUGGUUUACCUUCAUAUUUUUCGUGG
Oeu-miR2937_f	AUGAGAGGUGUUGAAGGGGGC	AUUAUGAGAGGUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUUCAUUGUGGUUUACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2937_g	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUAGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUUCAUUGUGGUUUACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2937_h	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUCAUUGUGGUUUACCUUCAUUAUUUUUCGUGGUCU
Oeu-miR2937_i	AUGAGAGAUGUUGAAGGGGGC	AUUAUGAGAGAUGUUGAAGGGGGCUGGUUGCUCCGUUAUAUGCAUGCUAAUGGGGCAAGUAUGUUUUCAUUGUGGUUUACCUUCAUAUUUUUCGUGGUCU
Oeu-miR2950_a	AUCCAUCUCUUGCAUAAUGGA	AUCCAUGAUGGAGGAGAUUGUUCAAGAAGCUCUGGAUUCUGUGCUCCCUUGCUCAAGUGAAGCUGCACUCCUUGAAUCCAUCUCUUGCAUAAUGGAUCC

miRNA name	miRNA sequence	pri-miRNA sequence
Oeu-miR2950_b	GGGUGUGCUGGGGGUUGAAUG	CUGGGGUGUGCUGGGGGUUGAAUGAGCCAAAAAUCUCAACACCCAAGUUGCAAGAAAUUUGUCGAAAUGACAAGUUCAGUUCCUUCCUAUAUACUACCAGAA
Oeu-miR3434_a	AAUUGGUUGAUUGUAUCAUU	GAACUGGAAUGAAAUACUUAACCGAAUGGAAAUAAAUUAAAGUACAAUAAAUUACAAUUGGUUGAUUGUAUCAUUAACCAUUUCU
Oeu-miR3434_b	UCUUGGCUUUUUCUAUGAUU	UCUUGGCUUUUUCUAUGAUUAUUGCUGUUUAUGCUCUCAUUUUUCUCCCUGUUGAGUUUUUCAUUAUCAGAUUUUUUAUUCGCAUCAGGAAAAAACGAUGAAUUUGAGGACAAAGAUGAGGAUGGAGAGGAUUGAGGAAGAGGUGAGGAUG
Oeu-miR3438	UCGAAAGCUUCUUCUCUGACG	AUGUUGAGGACGAGCUUCUGUACUUUAAUCUCUCUCAUUGGGUUCGAAAGCUUCUUCUCUGACGCC
Oeu-miR3447	UUUGAGUUGGUUUGUUACAAA	UUUGUGGCAAUCCAAUUAGGAAUAGUUCAAAAUCAUACAUAUAGAACAGUUGGCUUAAAGGAAUGCCUUGUGUUGUCUUUAACUUUUGAGUUGGUUUGUUACAAAAU
Oeu-miR3476	UGAUGUGGGUUUGUUGGAUGA	GUGUGAUGUGGGUUUGUUGGAUGAAUUGGGAAUGCUGGUCUGUUUGCCAAAAGGGUUCCUUCCAUUGCUCAGUUUGUUCAGAAAUUUAUAUGAACGCGU
Oeu-miR3511	UAUUACUAUGGCCUAUGGUGA	GUAUUACUAUGGCCUAUGGUGAGGAUGGGACAAGAGUGGGGAUUUGUUAAAAUUUUCCAUUUUUAACCAGUAGUGUCAGUGGUACAA
Oeu-miR3513	UGAAUUUCUGAGUUUGCAAUG	UUUGAAUUUCUGAGUUUGCAAUGUGAAAAUUGAAAAUUGACAAGAUUGAUCGAACCGAAACUCCUCGAAGAAGAGAGAAUUGGGGAAAUUGCAAAGUCUCUAGAAUUAGAAA
Oeu-miR3706_a	UUUCGGAGAGAUGGCUGAGU	CUAUGGCCAUGUUCUAUUCGGAGGAAUAAAAUAAAAAUUGUCUUUCGGAGAGAUGGCUGAGUGGUU
Oeu-miR3706_b	UUUCGGAGAGAUGGCUGAGU	CCUAUGGCCAUGUUCUAUUCGGAGGAAUAAAAUAAAAUUGUCUUUCGGAGAGAUGGCUGAGUAGGUU
Oeu-miR3706_c	UUGUGGAAAAAUGGAGAAGA	UUGUGGAAAAAUGGAGAAGAAGAUGCUGAUAGUAGUUUUAGCGACGCUUCUGGUCAUAUUUUCAGCAAAC
Oeu-miR3712	UUUGAUCAAGAUGAGAGUACGA	UUCUACUCGUGCUACUCCUUGUUUACCAUGUUGGGUGCGUUUGAUCAAGAUGAGAGUACGAAGAAAU
Oeu-miR3950	UUUUUUGGUAACAUUAUUUCU	UUUUUUUGGUAACAUUAUUUCUGUAUUUCGAUCAAGUACAACAAAUACAAAUAAUAACAUUAUAAGAACUCGAAAUCGUCUGACAAGAAAGAUGGAUGCUAAAACUGUCUAAUGACCUAAUGCUAAAUUGAAAUACAGAAAUAAUAUUAUACAACAAAACA
Oeu-miR3953	UUCAGUUCCGGAAGCCGUAGA	UUUCAGUUCCGGAAGCCGUAGAUAACUCUCAUAAGCCGAGCAUGCCGAGUCAGACGCCGACUAGAUAACUCUGUAACACGAGCUUGCGGCAUCGGGAUUGAAAUC
Oeu-miR4221	UUUUUUUCUGUUUAAUUCGUGC	UUUUUUUCUGUUUAAUUCGUGCAGUGGAAAACUUGUUACUUGAUAUCACUUUUUAUCCAAUAUGUUUAAUAAAGCUAGACAAAGUAGCAUGACUUAAGUAAAAAAAAAA
Oeu-miR4350_a	UCAAGGGAUUUUGUGUCAUUGG	AUCAAUAACACGAAGUCUUUUGACCCCGAGUACCUUAUAAUCCGGACUCACCACCUUGCCCACAUGGCAGCCUCCAUGGUAGUGCCAAAUUGUAAUUACCGUGUCUUUACAAAACUGCUCAAGGGAUUUUGUGUCAUUGGUAU
Oeu-miR4350_b	UCAAGGGAUUUUGUGUCGUUGG	AUCAAUAACACGAAGUCUUUUGACCCCGAGUACCUUAUAAUCCGGACUCACCACCUUGCCCACAUGGCAGCCUCCAUGGUAGUGCCAAAUUGUAAUUACCGUGUCUUUACAAAACUGCUCAAGGGAUUUUGUGUCGUUGGUAU
Oeu-miR5021_a	AGAGAAGAGGAGGAAGAAGA	UUUUCUUCUUCAAUCUUUGUCUUGGUUUUCUGGGAAUUGAAUUGAAGUGAAAGAGAGAAGAGGAGGAAGAAGACA
Oeu-miR5021_b	CUAGAAAAAUAAGAAGAAAA	UUUUUUUUUUUUUUUUUCUAUUCAACCGCAAAUUUUUAUUUCAAAGUGAUUUGUCCAUUGGAAUUACUAGAAAAAUAAGAAGAAAAAU
Oeu-miR5021_c	AGAGAAGACGAAGAAGAAAG	AUUCUUUUUUCUUGCCGGGUUCUUUGGUUCUAGUCUCUUCUCUCAGCAGGGAGUGUCGAGUACUUCUUUAAGGACUUGGCCUCGGGGACUGCAGGAAGAGAAGACGAAGAAGAAAGAAUUU
Oeu-miR5021_d	UGAGAGCAAGAAGAAGUAAU	AGUUAACUCUUCUGCCUCAAGUUGAGAAUUGACAGCAAAAAAGACAAGACCUUGAGAGCAAGAAGAAGUAAUUGA
Oeu-miR5021_e	UGACACAAAGAGGAAGAAAA	UUUUCUAUCUUUUUUCUUUUAUUUUUCUUUUAAGAUCCUUAGGGCUCUGAUAUUACUGAAGGAAUUGUGACACAAAGAGGAAGAAAAUA
Oeu-miR5021_f	GAAGAAGAAGAAGAAGAAAU	GUUUCUUUUGCUGCUCUUCAACACCCUUCUUUUCUUCAACAGUGGCUGAUGACUCGACGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAAUGC
Oeu-miR5138	GAAAAUCAUCAGGCGCUC	UGAGCCCGAUGAUUCUCCCCGGGUAGGAAAAGAAGAAGCUCUUAUCAGAAAAUCAUCAGGCGCUCAUA
Oeu-miR5140_a	GCUCUUGAAGCUAUGGUG	CUUUCUUUAUAUCUUCGAGUCUUUAUGUAGAGAAGGAAGAGUCAUUGAGUUACUAGACGCUCUUGAAGCUAUGGUGAAAGAU
Oeu-miR5140_b	GCUGGUGAAAUUUUGUUU	GGAUCAAAAUAUCCUCAGUCAAAAGAUCAUCUUGAGAAAAGUUGUAAAUUUCUUCAACCUCAAACUUCCCUUUAUUGAAAGAGAAAUCAAAUAAGUGAGGAUCUCUGACAACCUCUGGAGAUAUUUUCUUGCUGGUGAAAUUUUGUUUCCCU
Oeu-miR5140_c	GCUCGUAAUGAUUUGGUA	CAAGGCCAGAUGAUUGAAGAGGUCCCAAAUUGCAGAGGCUGGCCGUGAAGAAGGUUAAAGGACUUUCAGCAGAAGAACUUGCCGCUCGUAAUGAUUUGGUACUUGCA
Oeu-miR5140_d	GCGGGUGAAAACCUGGUG	GCGGGUGAAAACCUGGUGACGACCCAAUGAUUUUUUCGCAUGUUGUGCUGGUCUUGCAAUCCGUUCCGGAAAUCAUUGCAUUUGUGCACCAGGUUUUCACCUCGCUU
Oeu-miR5140_e	GCUGCUGGAGAUUCAGUG	GCUGCUGGAGAUUCAGUGAUUUUCAUCUGGAAUGAAAAUAAUCAGUUGUUAUUGGGAAUUCGGCGUGCUA
Oeu-miR5140_f	GUUGUUGAAGAUAUAGUG	CAUUGAGUUCUUCACAAUGAAUUUCAUACAUUUUGUGUAUAUAAAAUUACAUAUAUAUACAUAGUUACGUUGUUGAAGAUAUAGUGGA
Oeu-miR5140_g	GUUGCUGAAGAUUUUGGG	GCUUUGCUGCUGAAGUCUCGCAAUACCAAGCGCAGGGGGAAAGUAAAGAGUCUGCAUUUGUUGUGAGUUAUCAGGUUGCUGAAGAUUUUGGGCAAAGCCU
Oeu-miR5140_h	GCUGGGGAGGAGCUGGUG	CUGCACAAGUUUUUCUAUGUUGGCUUUUUAUUCAAGCAUUUGGCUGGGGAGGAGCUGGUGCGGCU
Oeu-miR5142_a	CAAUUGAAUUAUAAGUGAU	GCAAUUGAAUUAUAAGUGAUUCAUUCUAUUUGCUGGAUAGUGAAAAUCAUAUAAUUGGAUUGUCU
Oeu-miR5142_b	CUAUCAAUUGACAAGUGAU	UUCUAUCAAUUGACAAGUGAUUCCGUAAAUUUGCGCCUCUUCUUCUUCUUUGACUUAAUUUUGGGAAUUAUGGGUGUUGGUGGAGGG
Oeu-miR5202	UUACGUGAUCGAAAUCGUCAA	UCUUACGUGAUCGAAAUCGUCAAUAUAGAAGUUAAUAUAAAAUCGCACUGUUAAAAGUGGUUUGAUGGUUCAUUCCUCAAGCUUUAAUAAUGUGAAAUUCGCAAGUAAAUUUUAACGUUUAUGUUGAUUAGUUUUGAUCGCAUGGGACC
Oeu-miR5301	UGUGAUUUGGGUUGAAAGAUU	GGGAUGUGAUUUGGGUUGAAAGAUUUGCAAAUGUUUGAAGGUUCUUUUUCAAAUUUCUGAAAGCUGUUCAUAGACCGAUUGUGAGAGGAUUUUUCGGGCCAGAUUAUUGCUUUC

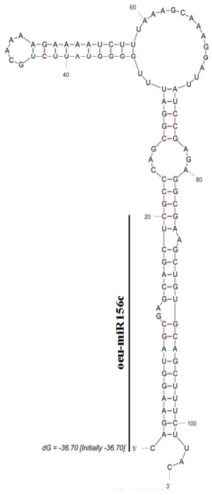
Appendix 1.3. Stem loop structures of 204 olive pre-miRNAs.



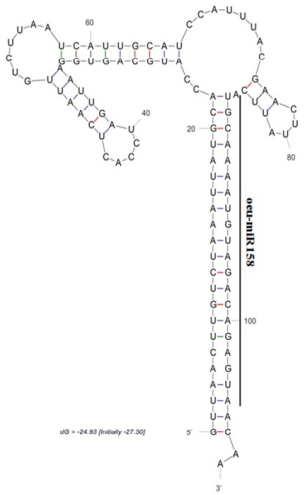
oeu-miR156a



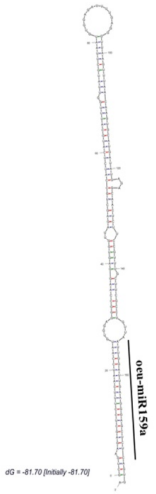
oeu-miR156b



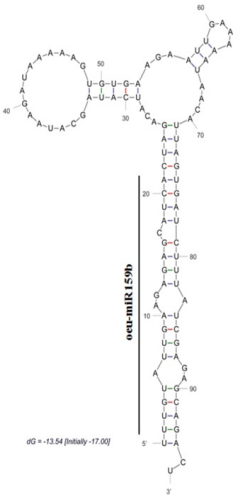
oeu-miR156c



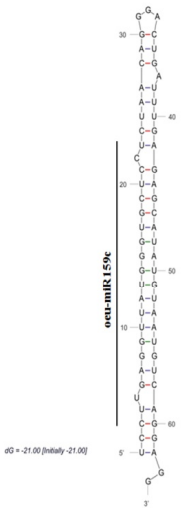
oeu-miR158



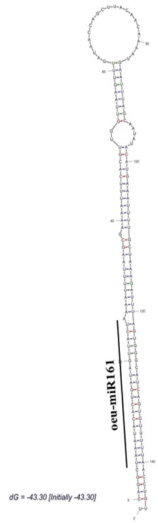
oeu-miR159a



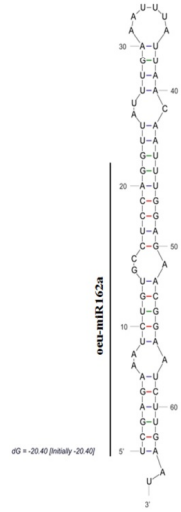
oeu-miR159b



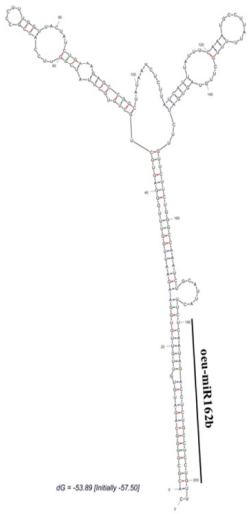
oeu-miR159c



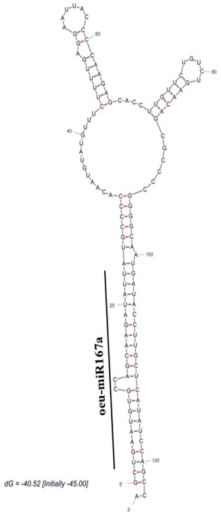
oeu-miR161



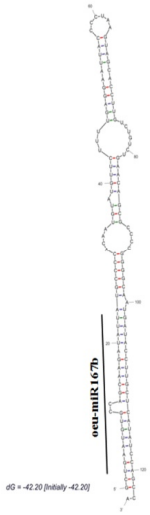
oeu-miR162a



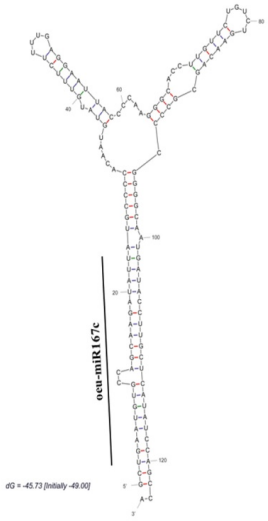
oeu-miR162b

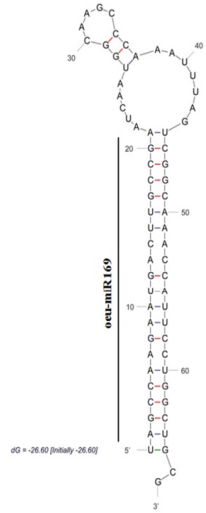


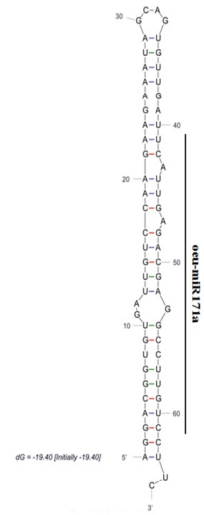
oeu-miR167a

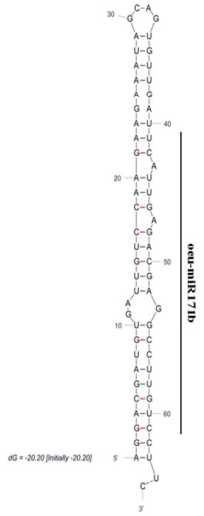


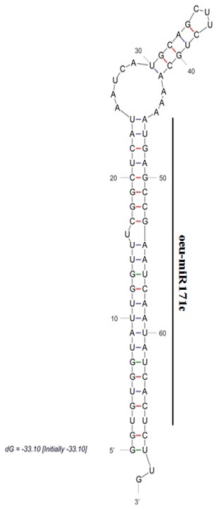
oeu-miR167b

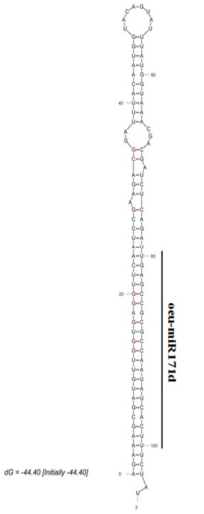

oeu-miR167c
dG = -45.73 [Initially -49.00]

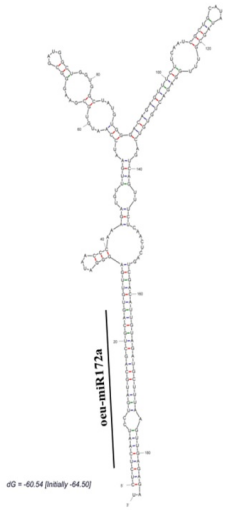

oeu-miR169
dG = -26.60 [Initially -26.60]


oeu-miR171a
dG = -19.40 [Initially -19.40]

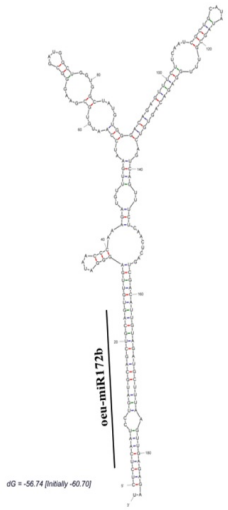

oeu-miR171b
dG = -20.20 [Initially -20.20]


oeu-miR171c
dG = -33.10 [Initially -33.10]

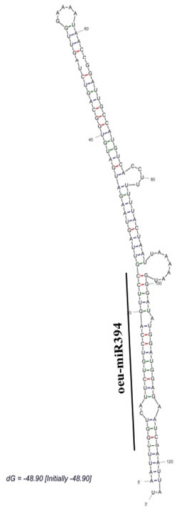

oeu-miR171d
dG = -44.40 [Initially -44.40]



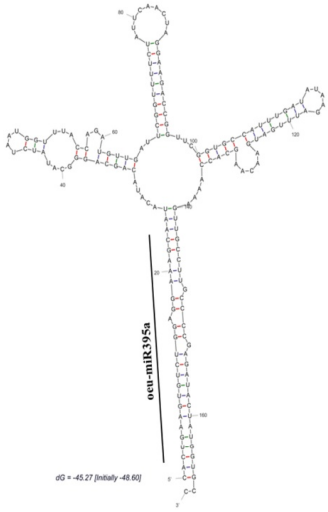
oeu-miR172a



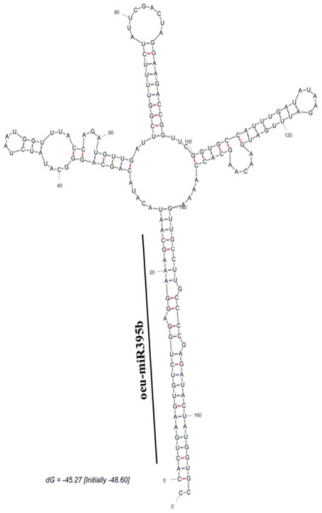
oeu-miR172b



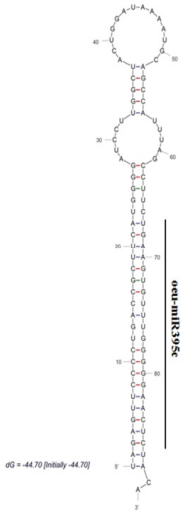
oeu-miR394



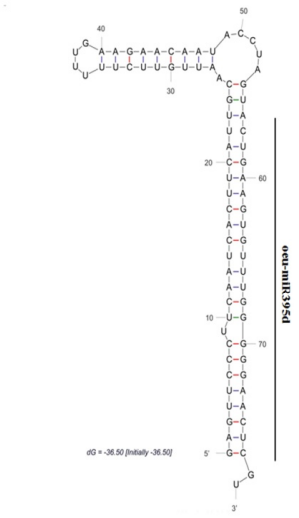
oeu-miR395a



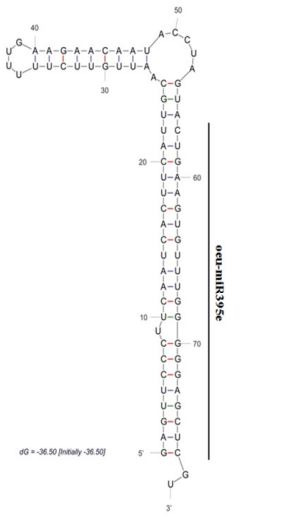
oeu-miR395b



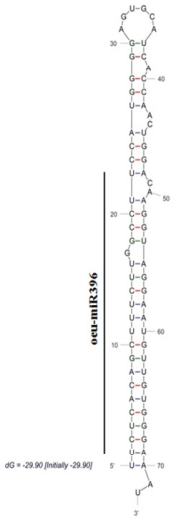
oeu-miR395c



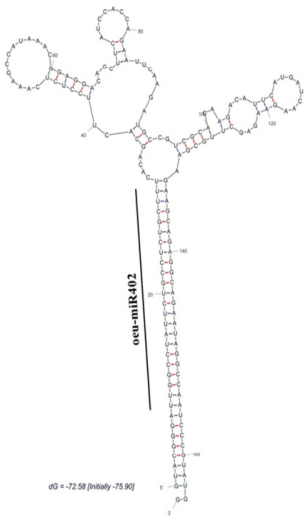
oeu-miR395d



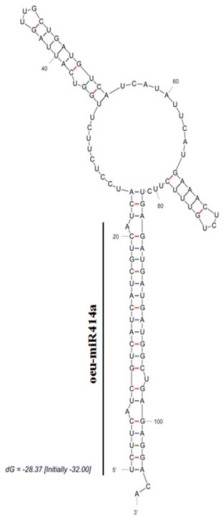
oeu-miR395e



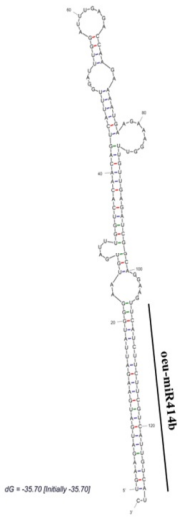
oeu-miR396



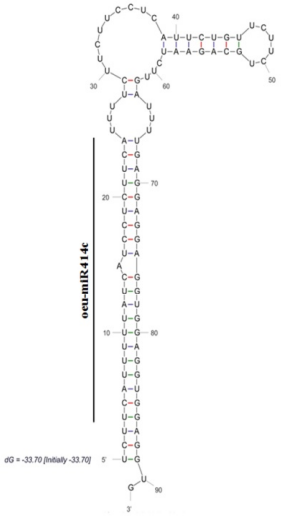
oeu-miR402



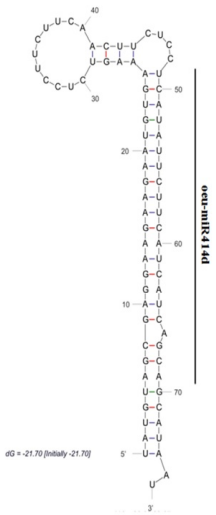
oeu-miR414a



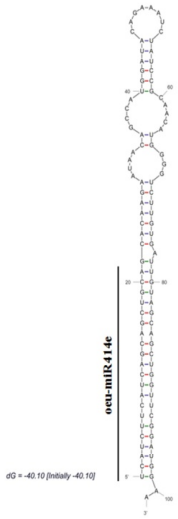
oeu-miR414b



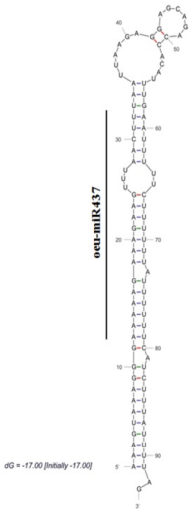
oeu-miR414c



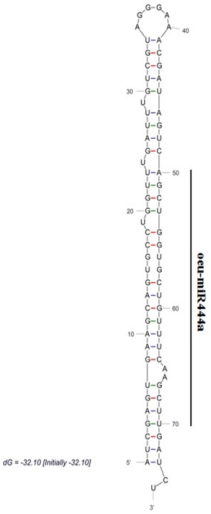
oeu-miR414d



oeu-miR414e

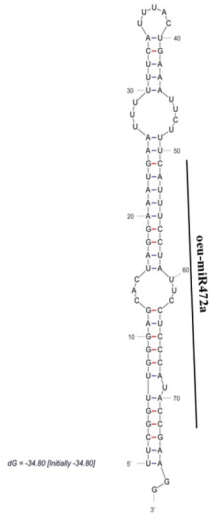


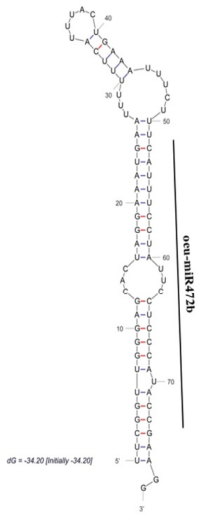
oeu-miR437

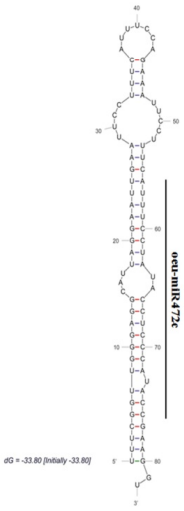


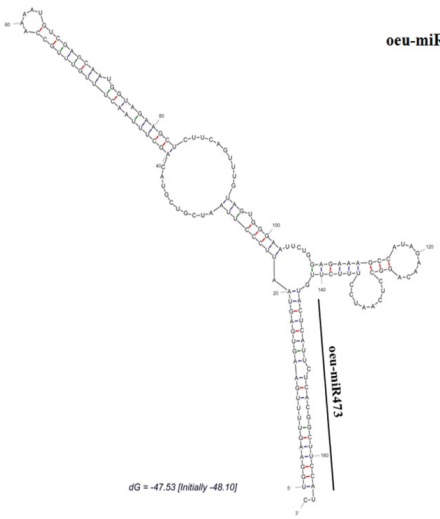
oeu-miR444a

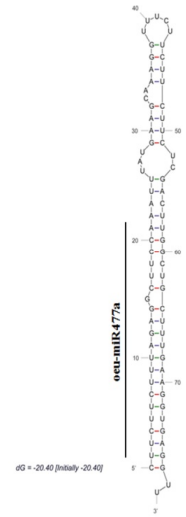
oeu-miR444b

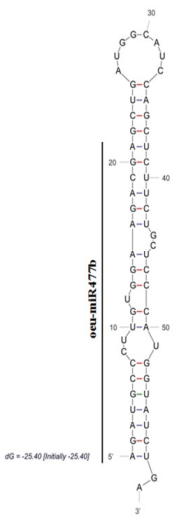

oeu-miR472a
dG = -34.80 [Initially -34.80]

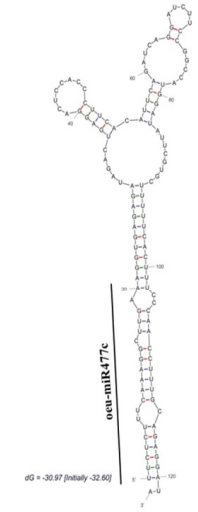

oeu-miR472b
dG = -34.20 [Initially -34.20]


oeu-miR472c
dG = -33.80 [Initially -33.80]

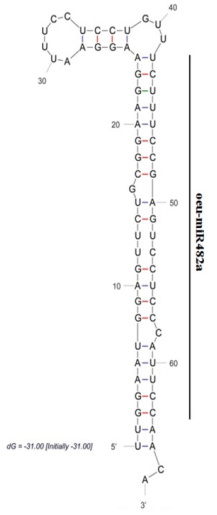

oeu-miR473
dG = -47.53 [Initially -48.10]


oeu-miR477a
dG = -20.40 [Initially -20.40]

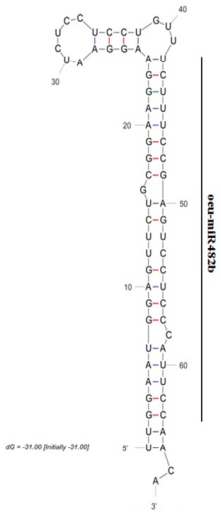

oeu-miR477b
dG = -25.40 [Initially -25.40]



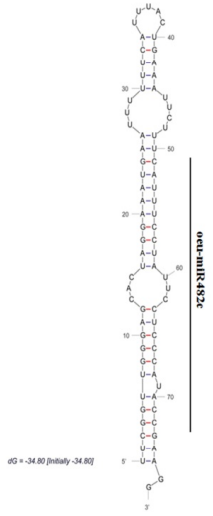
oeu-miR477c



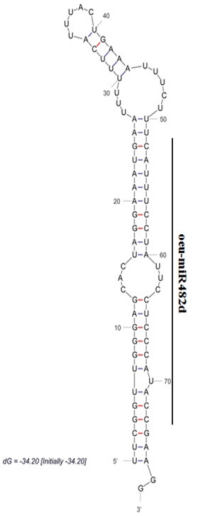
oeu-miR482a



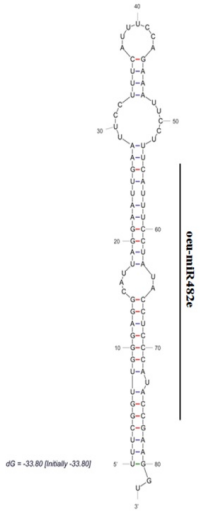
oeu-miR482b



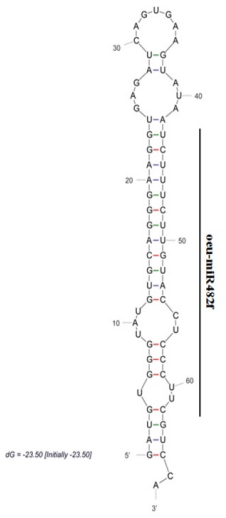
oeu-miR482c

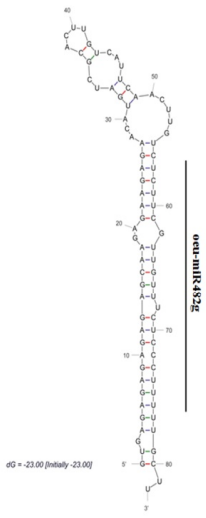


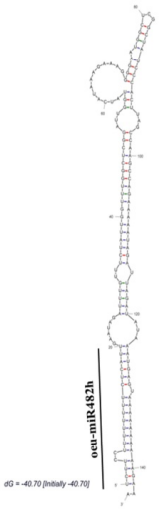
oeu-miR482d

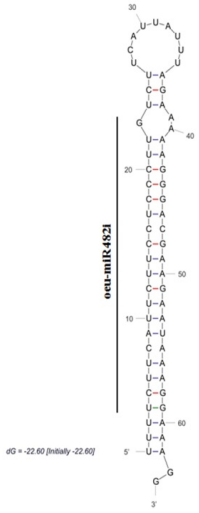


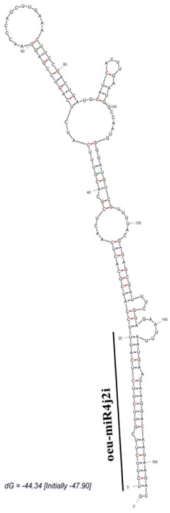
oeu-miR482e

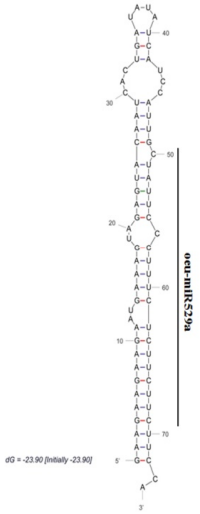

oeu-miR482f
dG = -23.50 [Initially -23.50]

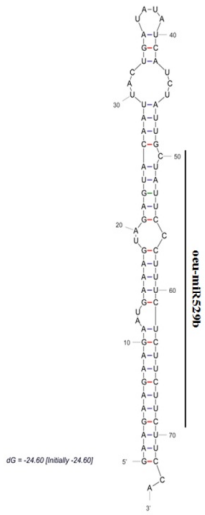

oeu-miR482g
dG = -23.00 [Initially -23.00]


oeu-miR482h
dG = -40.70 [Initially -40.70]

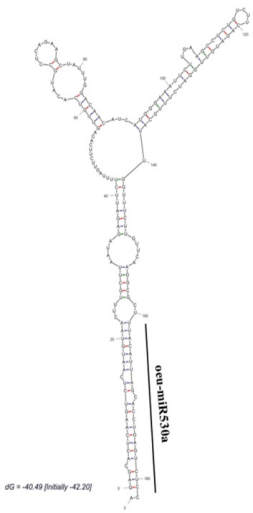

oeu-miR482i
dG = -22.60 [Initially -22.60]


oeu-miR482j
dG = -44.34 [Initially -47.90]

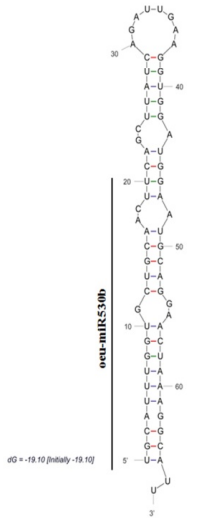

oeu-miR529a
dG = -23.90 [Initially -23.90]



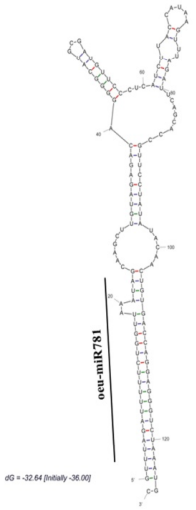
oeu-miR529b



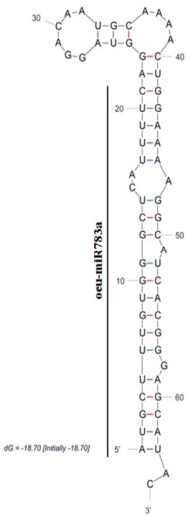
oeu-miR530a



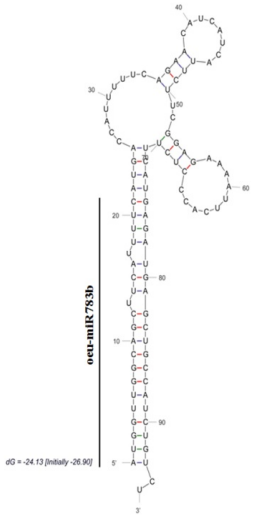
oeu-miR530b



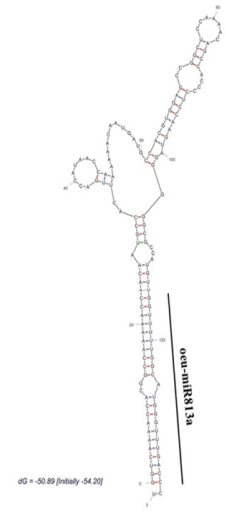
oeu-miR781



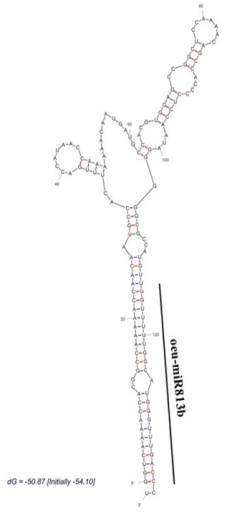
oeu-miR783a



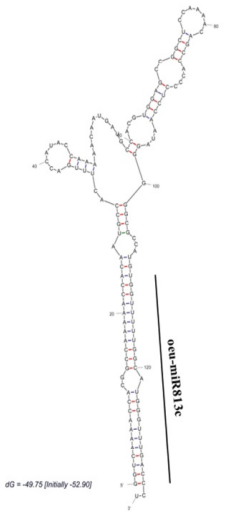
oeu-miR783b



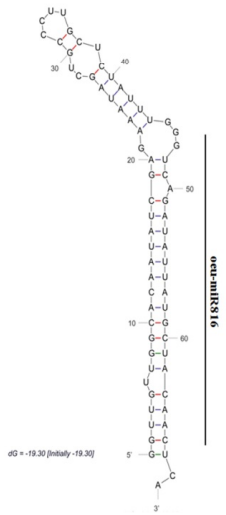
oeu-miR813a



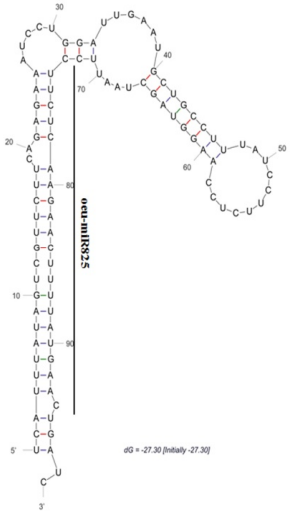
oeu-miR813b



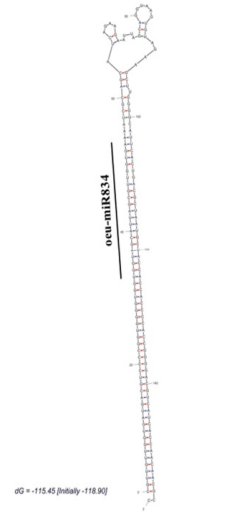
oeu-miR813c



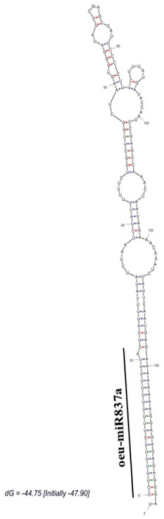
oeu-miR816



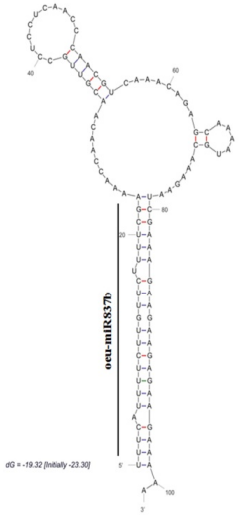
oeu-miR825



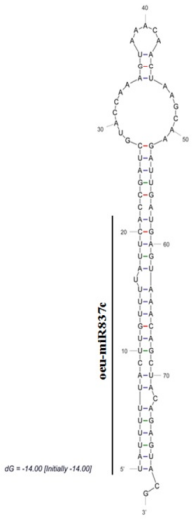
oeu-miR834



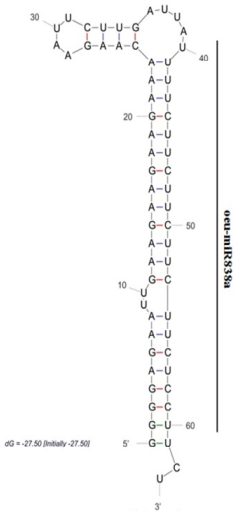
oeu-miR837a



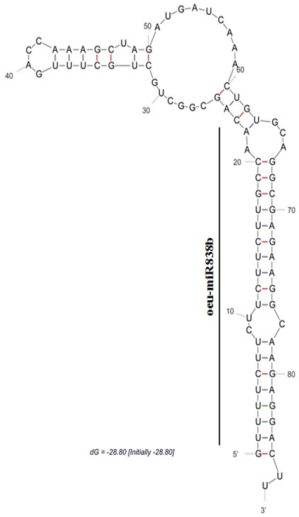
oeu-miR837b



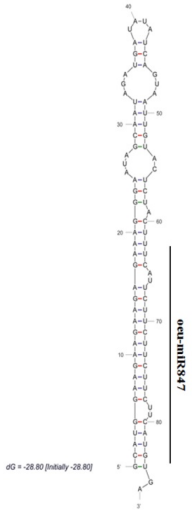
oeu-miR837c



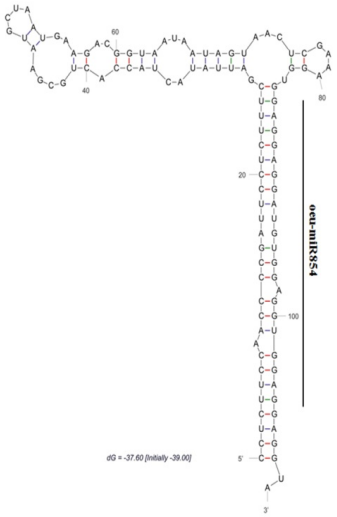
oeu-miR838a



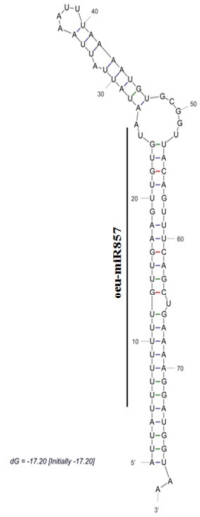
oeu-miR838b



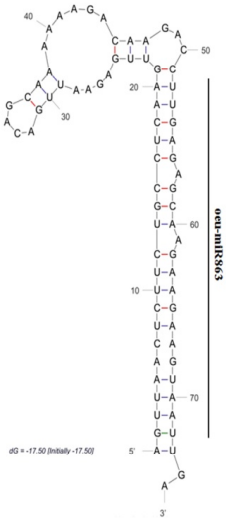
oeu-miR847



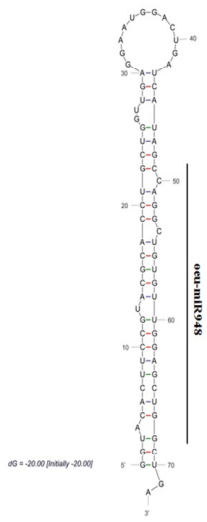
oeu-miR854



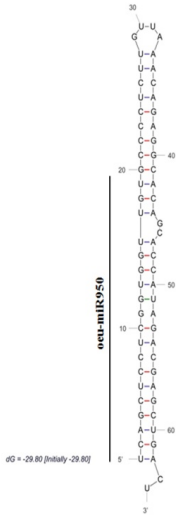
oeu-miR857



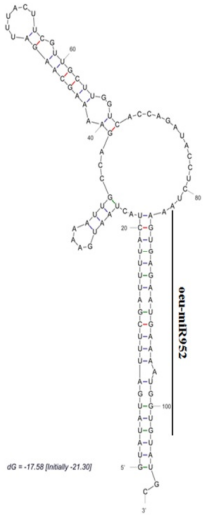
oeu-miR863



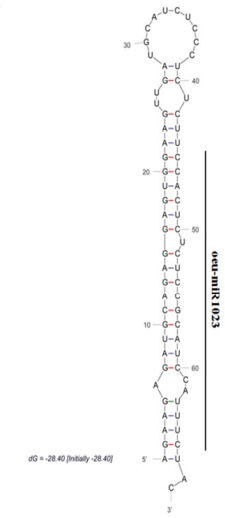
oeu-miR948



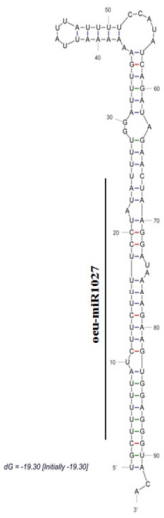
oeu-miR950



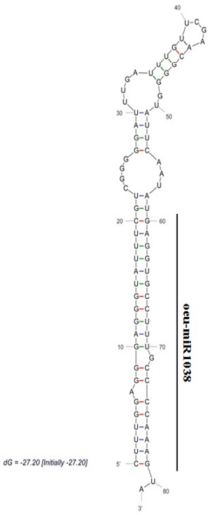
oeu-miR952



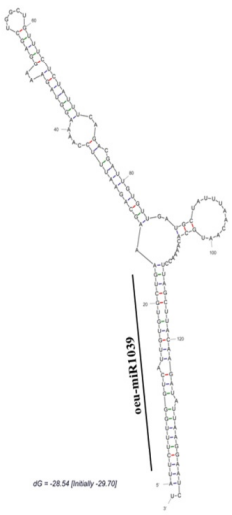
oeu-miR1023



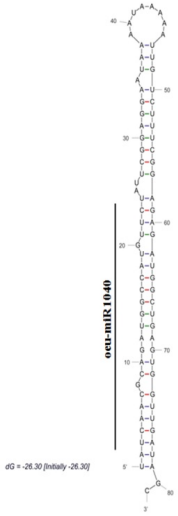
oeu-miR1027



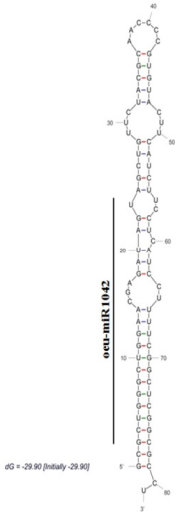
oeu-miR1038



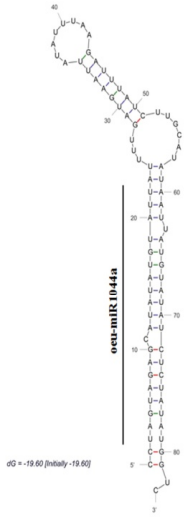
oeu-miR1039



oeu-miR1040

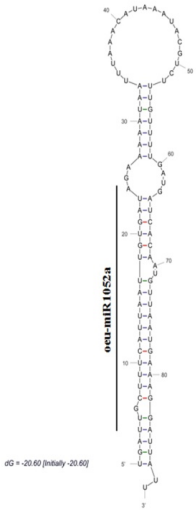


oeu-miR1042

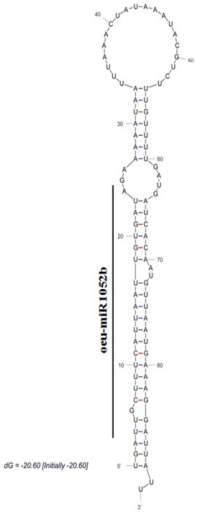


oeu-miR1044a

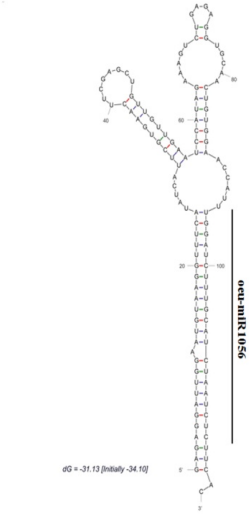
oeu-miR1044b



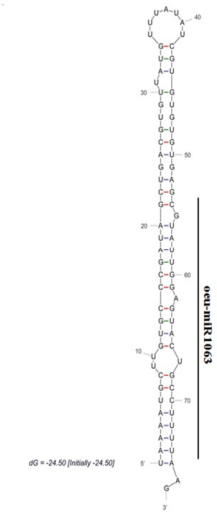
oeu-miR1052a



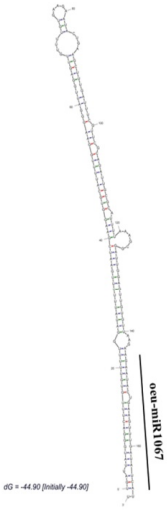
oeu-miR1052b



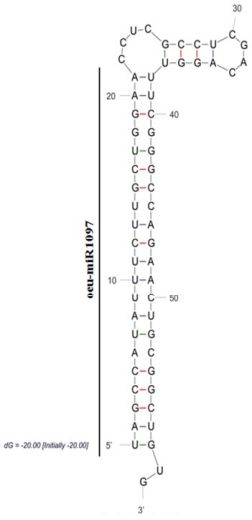
oeu-miR1056



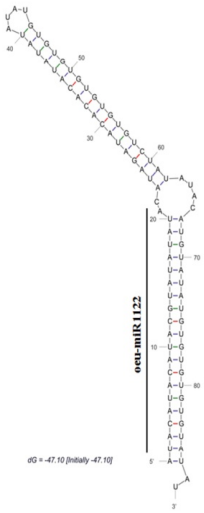
oeu-miR1063



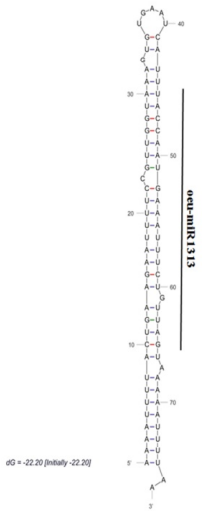
oeu-miR1067



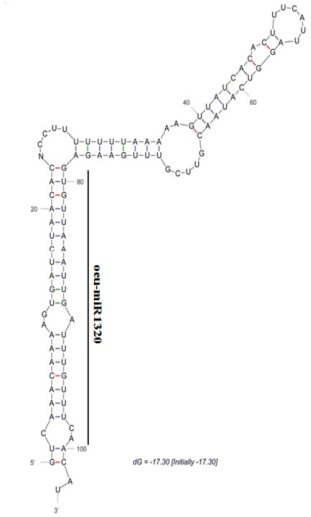
oeu-miR1097



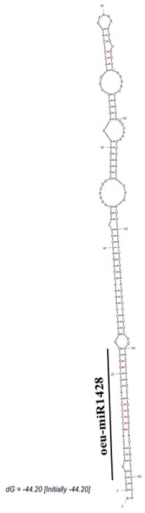
oeu-miR1122



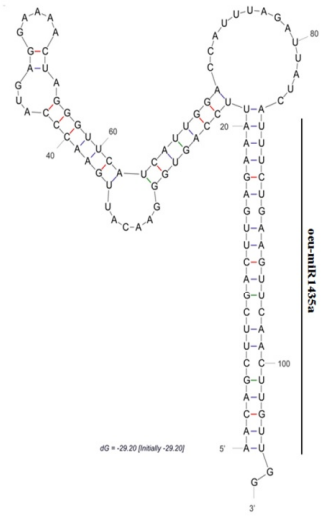
oeu-miR1313



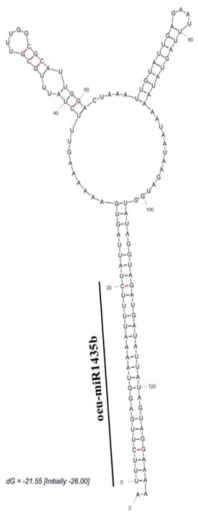
oeu-miR1320



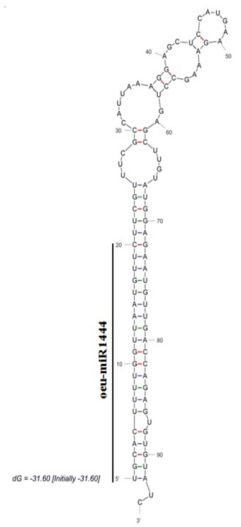
oeu-miR1428



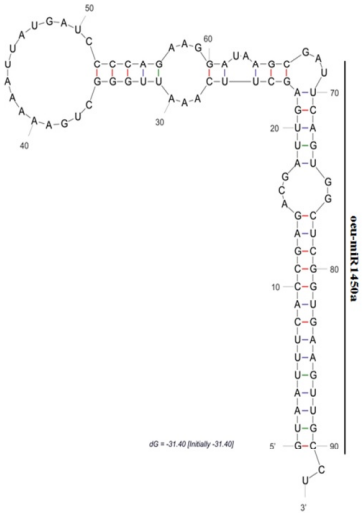
oeu-miR1435a



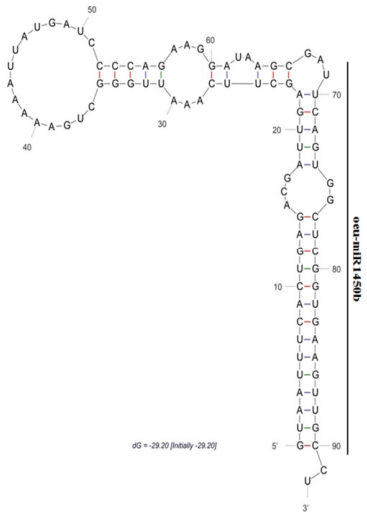
oeu-miR1435b



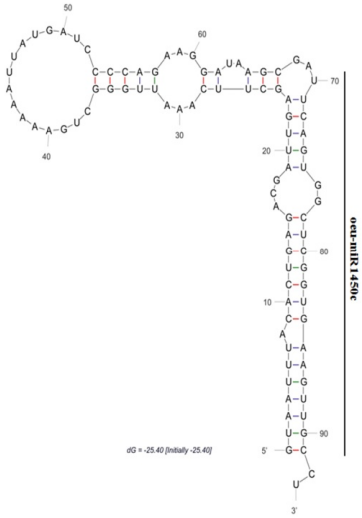
oeu-miR1444



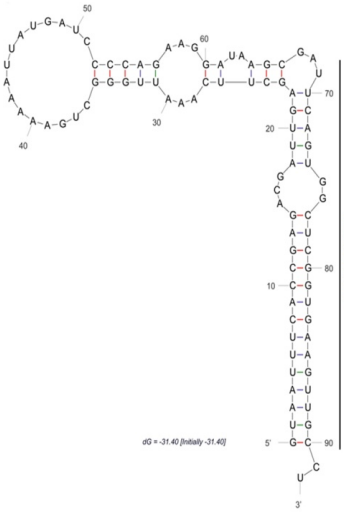
oeu-miR1450a



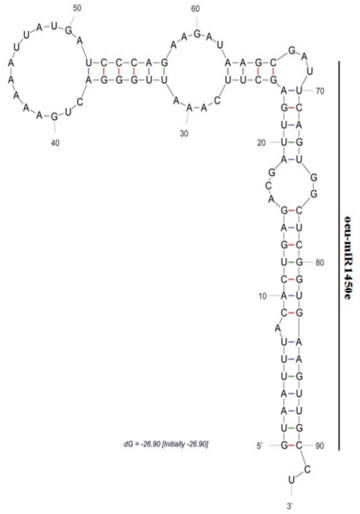
oeu-miR1450b



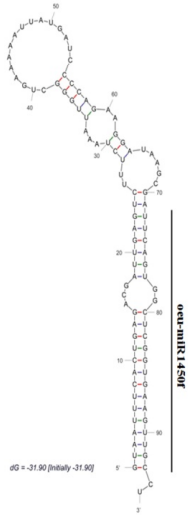
oeu-miR1450c



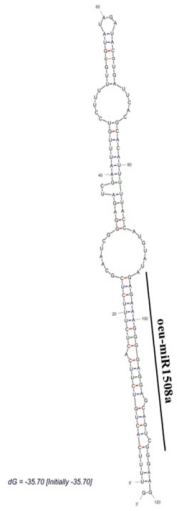
oeu-miR1450d



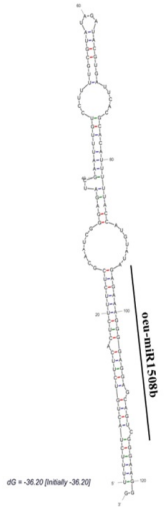
oeu-miR1450e



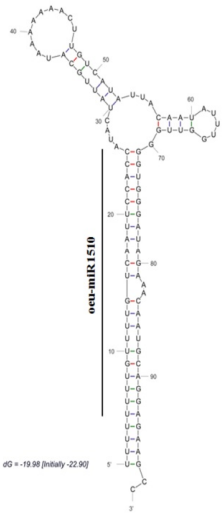
oeu-miR1450f



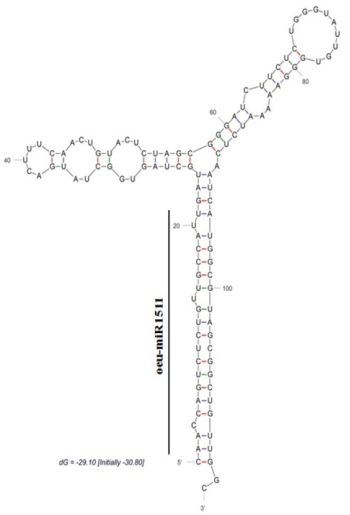
oeu-miR1508a



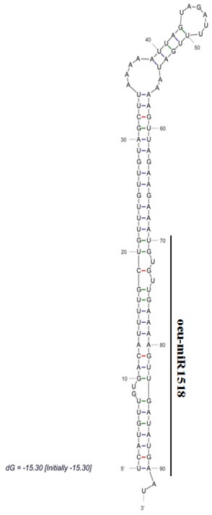
oeu-miR1508b



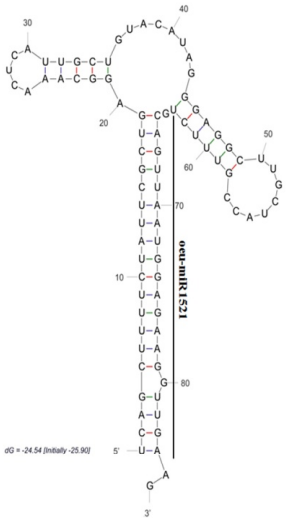
oeu-miR1510



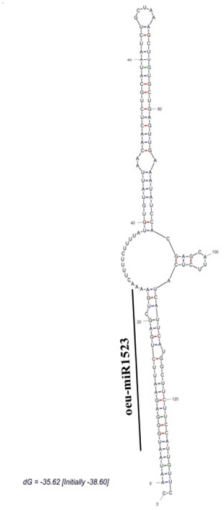
oeu-miR1511



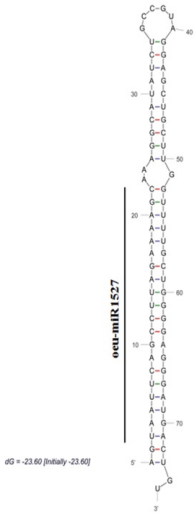
oeu-miR1518



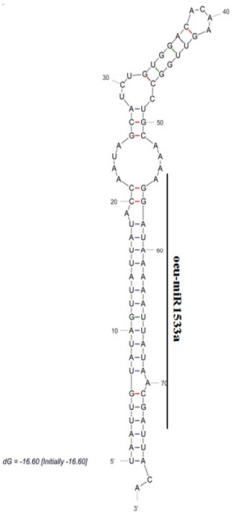
oeu-miR1521



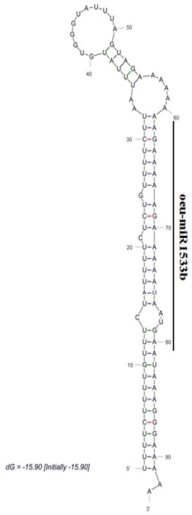
oeu-miR1523



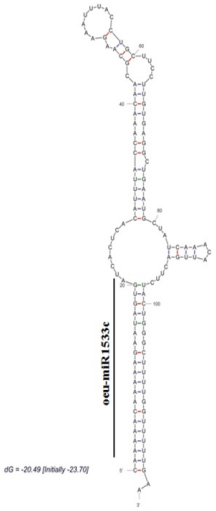
oeu-miR1527



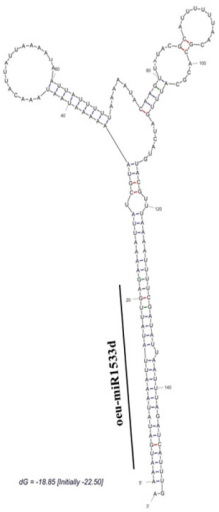
oeu-miR1533a



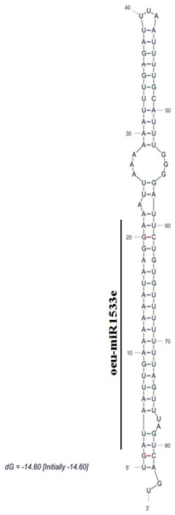
oeu-miR1533b



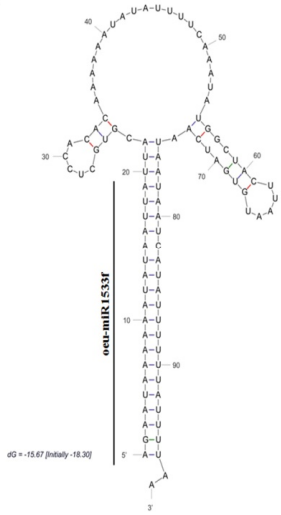
oeu-miR1533c



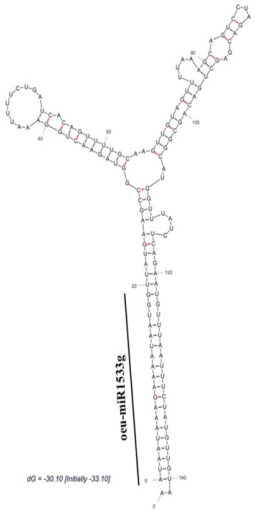
oeu-miR1533d



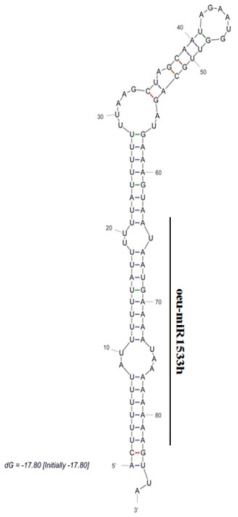
oeu-miR1533e



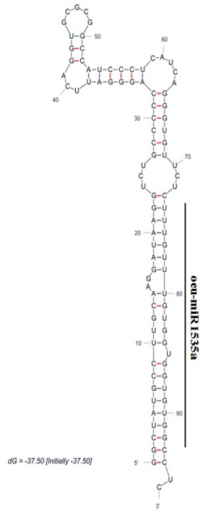
oeu-miR1533f



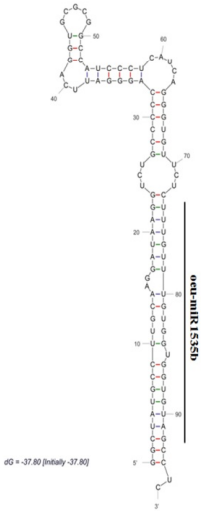
oeu-miR1533g



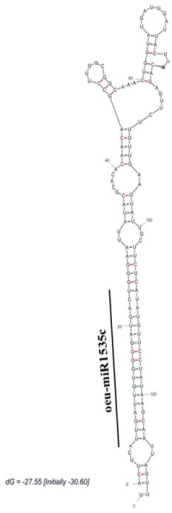
oeu-miR1533h



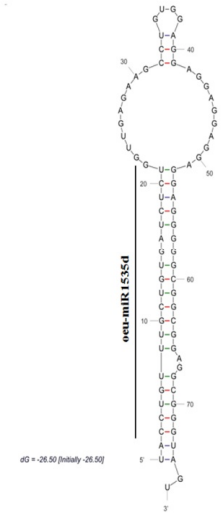
oeu-miR1535a



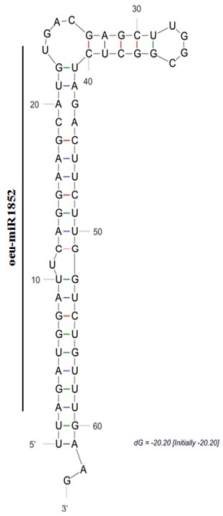
oeu-miR1535b



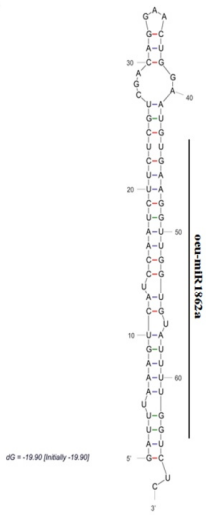
oeu-miR1535c



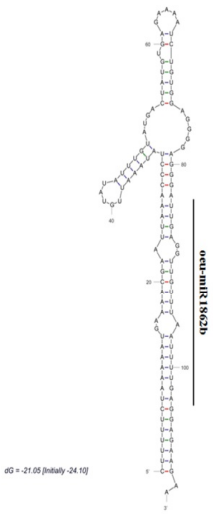
oeu-miR1535d



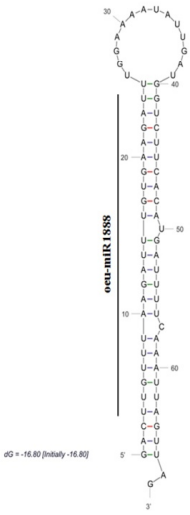
oeu-miR1852



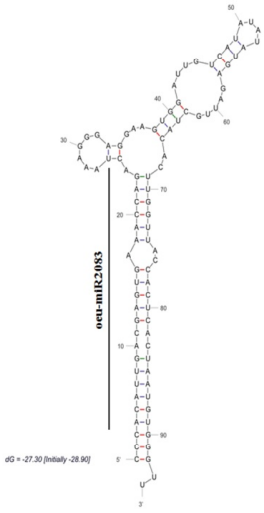
oeu-miR1862a



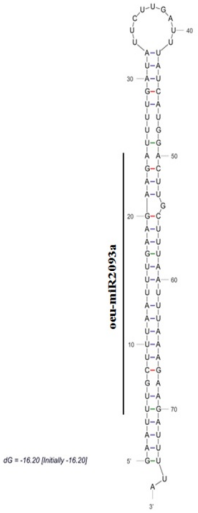
oeu-miR1862b



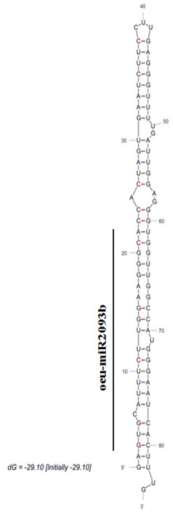
oeu-miR1888



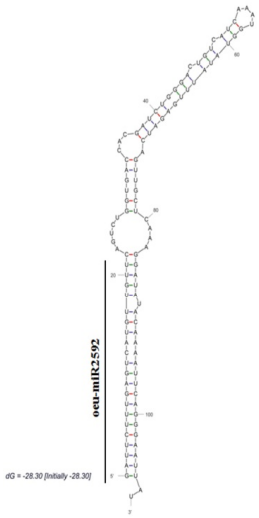
oeu-miR2083



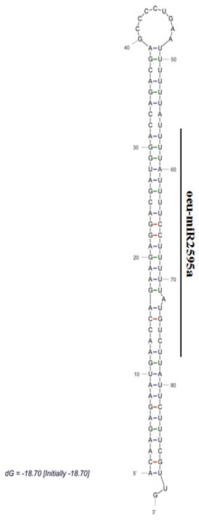
oeu-miR2093a



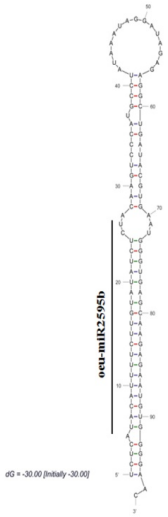
oeu-miR2093b



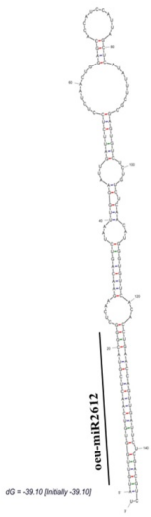
oeu-miR2592



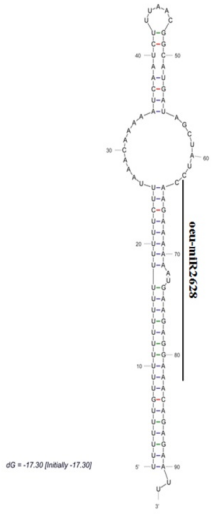
oeu-miR2595a



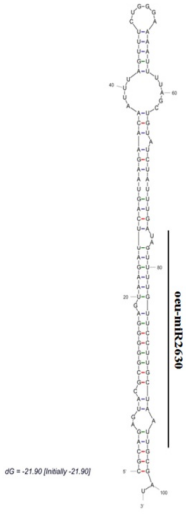
oeu-miR2595b



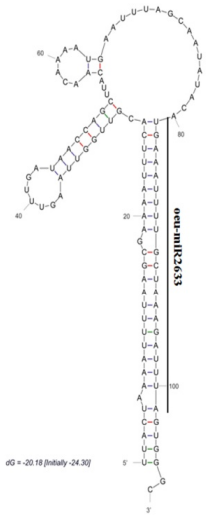
oeu-miR2612



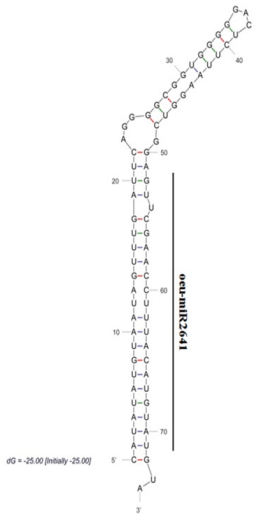
oeu-miR2628



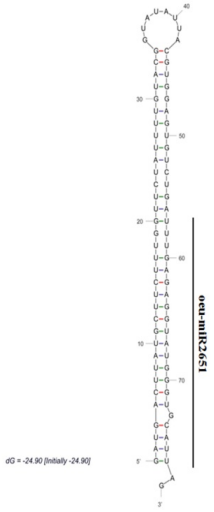
oeu-miR2630



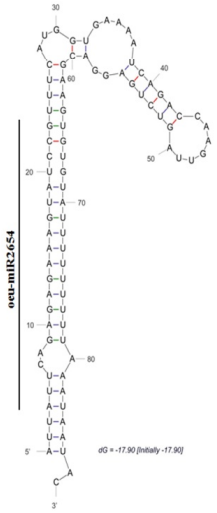
oeu-miR2633



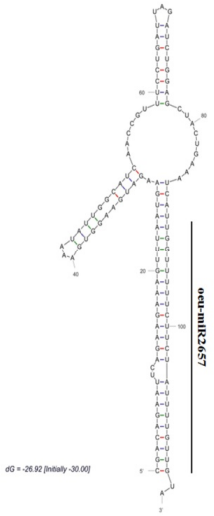
oeu-miR2641



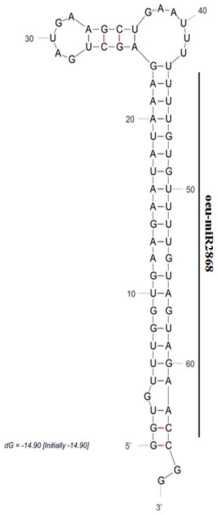
oeu-miR2651



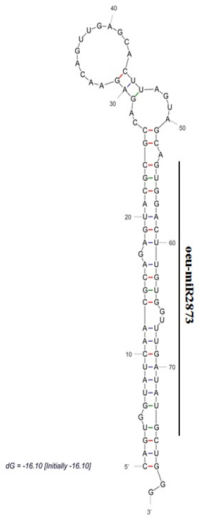
oeu-miR2654



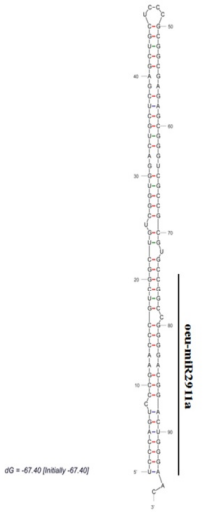
oeu-miR2657



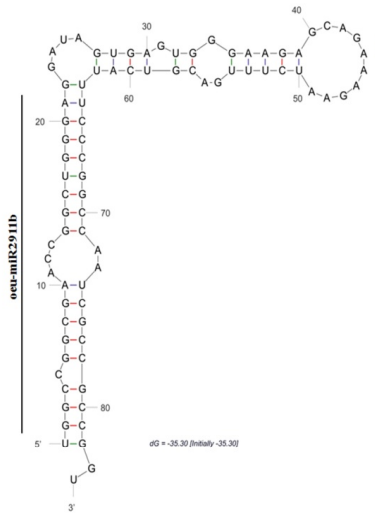
oeu-miR2868



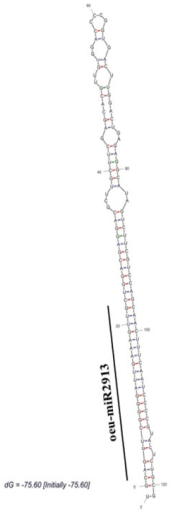
oeu-miR2873



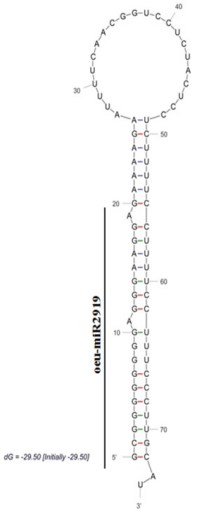
oeu-miR2911a



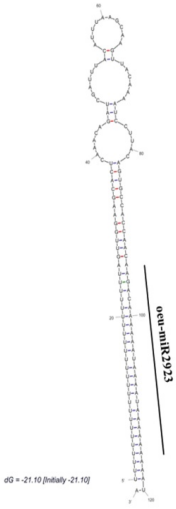
oeu-miR2911b



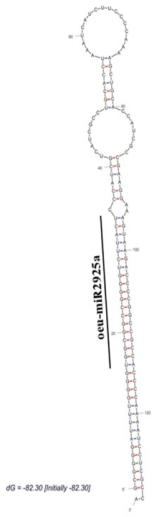
oeu-miR2913



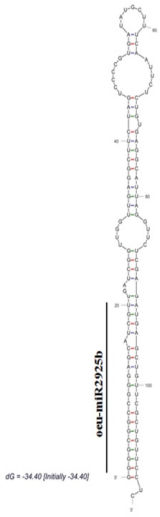
oeu-miR2919



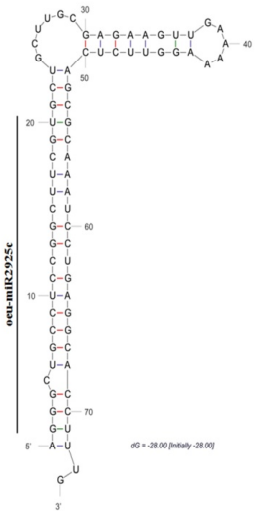
oeu-miR2923



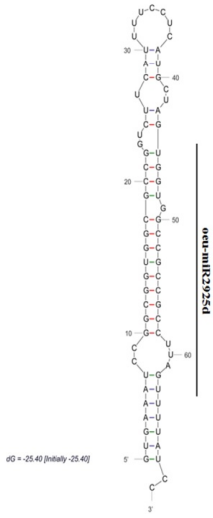
oeu-miR2925a



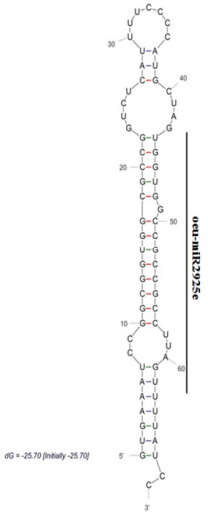
oeu-miR2925b



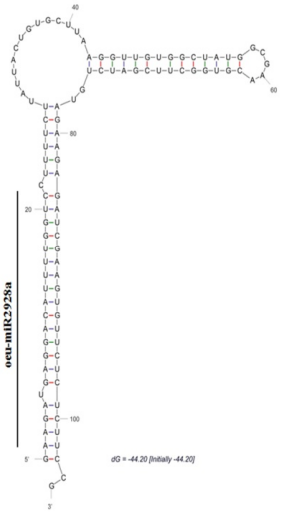
oeu-miR2925c



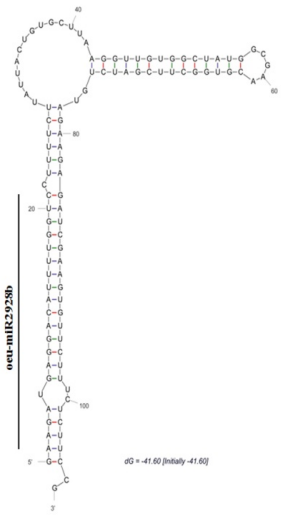
oeu-miR2925d



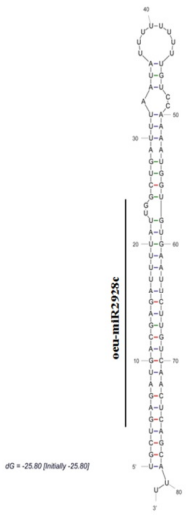
oeu-miR2925e



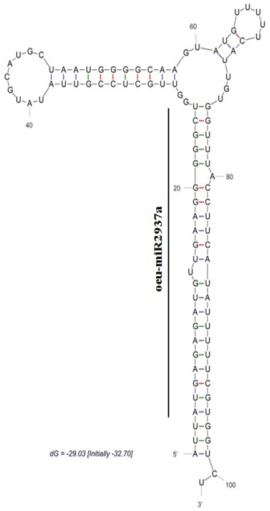
oeu-miR2928a



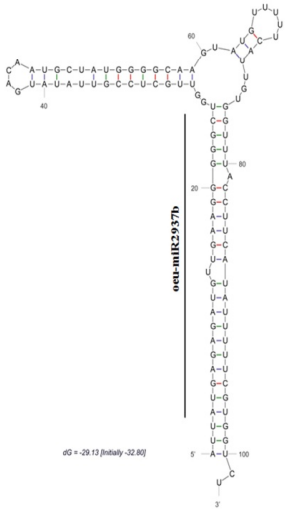
oeu-miR2928b



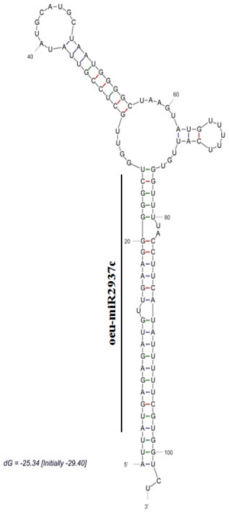
oeu-miR2928c



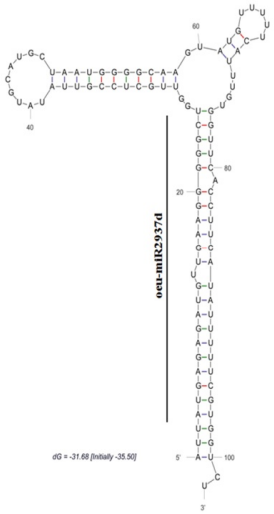
oeu-miR2937a



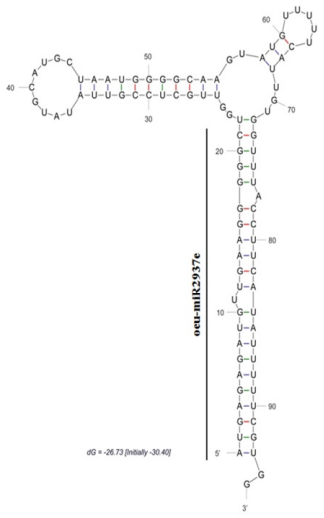
oeu-miR2937b



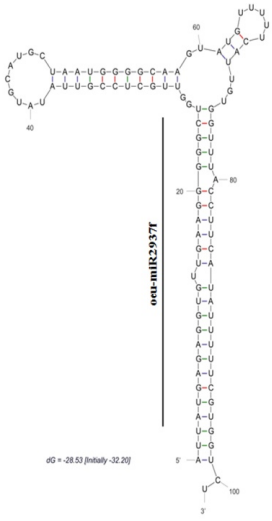
oeu-miR2937c



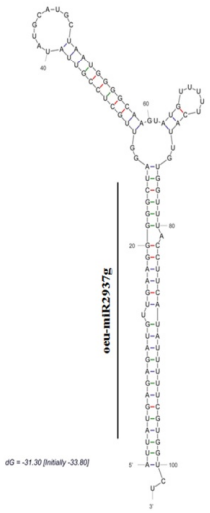
oeu-miR2937d



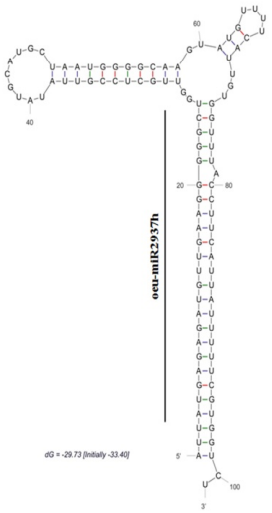
oeu-miR2937e



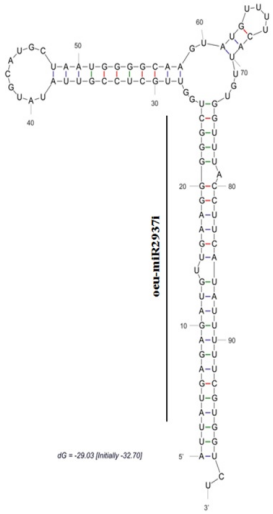
oeu-miR2937f



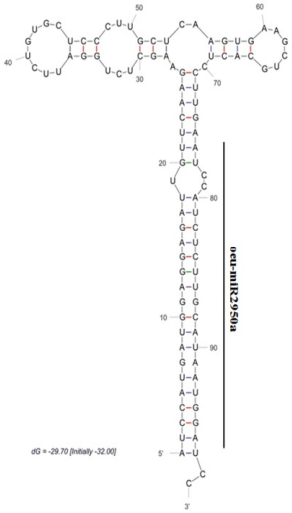
oeu-miR2937g



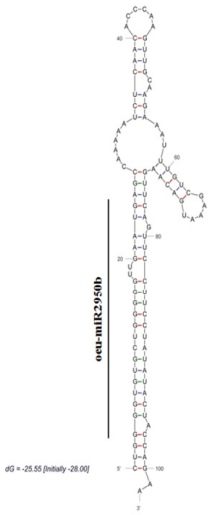
oeu-miR2937h



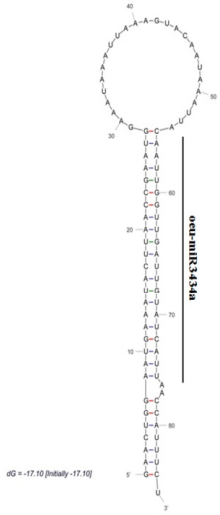
oeu-miR2937i



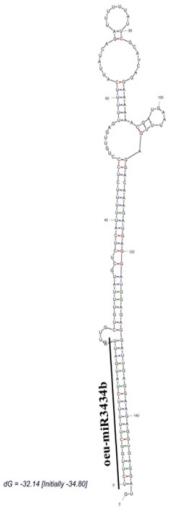
oeu-miR2950a



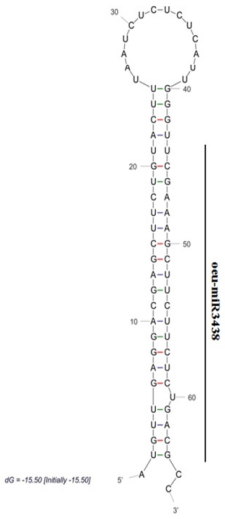
oeu-miR2950b



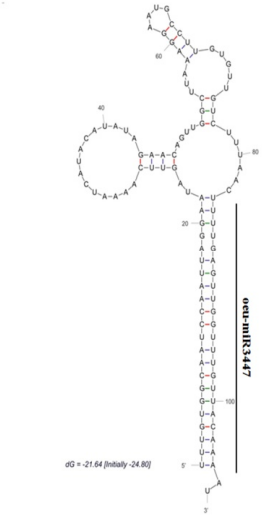
oeu-miR3434a



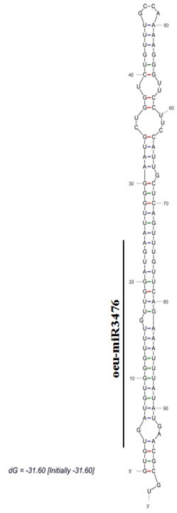
oeu-miR3434b



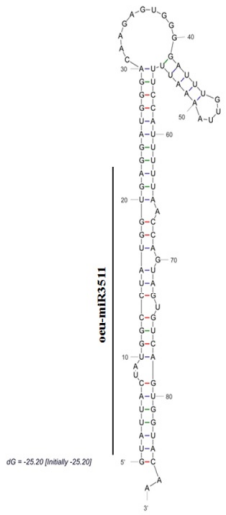
oeu-miR3438



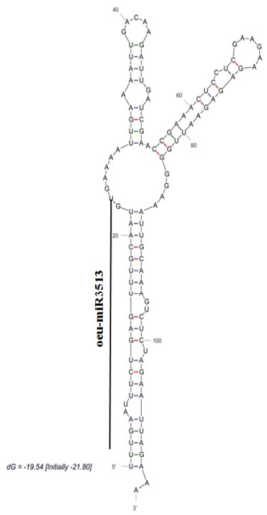
oeu-miR3447



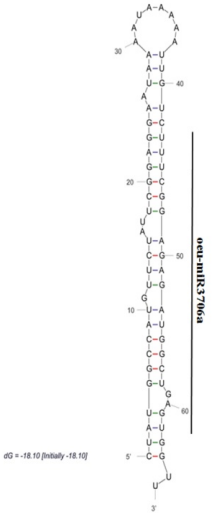
oeu-miR3476



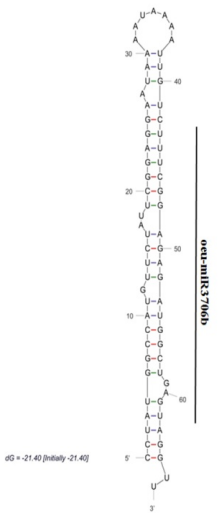
oeu-miR3511



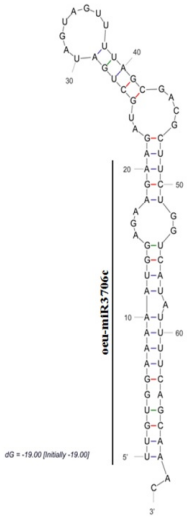
oeu-miR3513



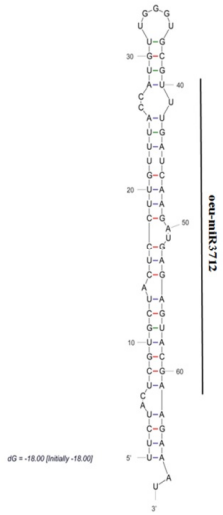
oeu-miR3706a



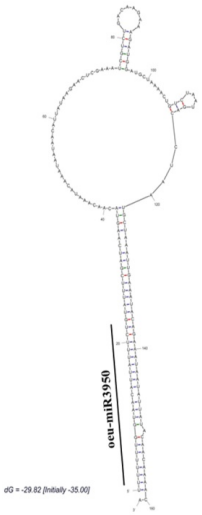
oeu-miR3706b



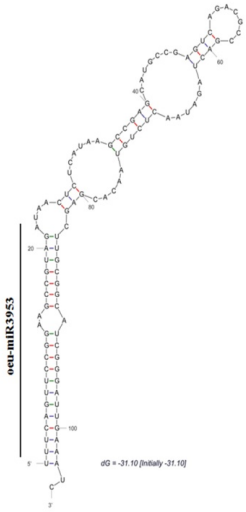
oeu-miR3706c



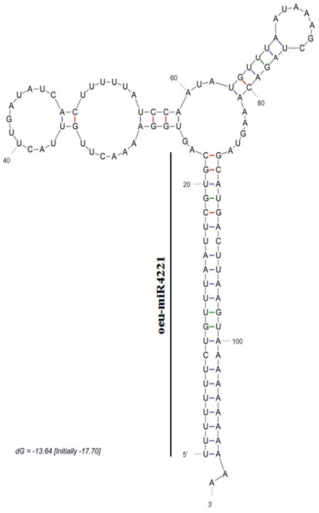
oeu-miR3712



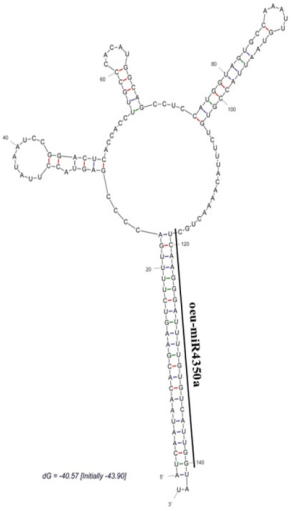
oeu-miR3950



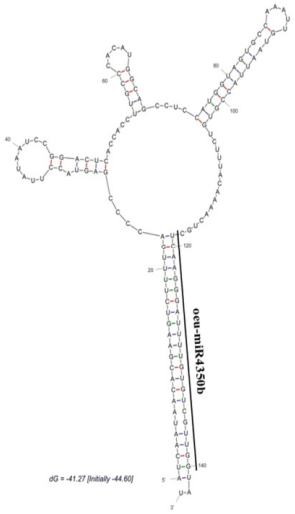
oeu-miR3953



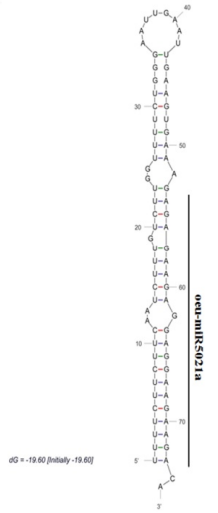
oeu-miR4221



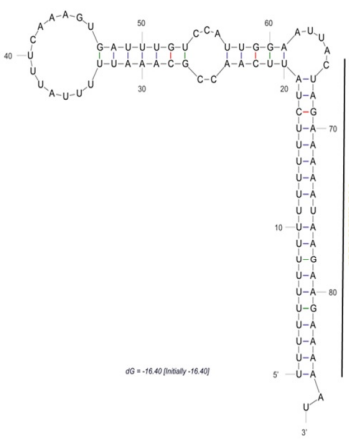
oeu-miR4350a

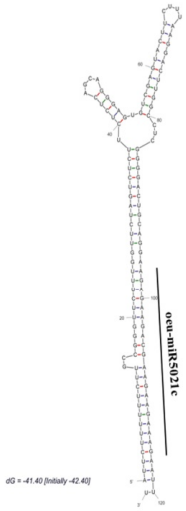


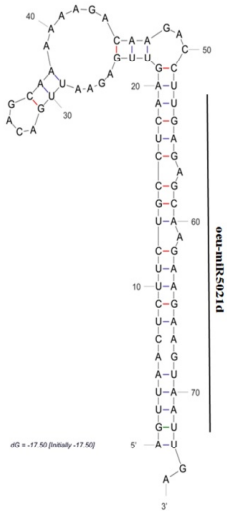
oeu-miR4350b

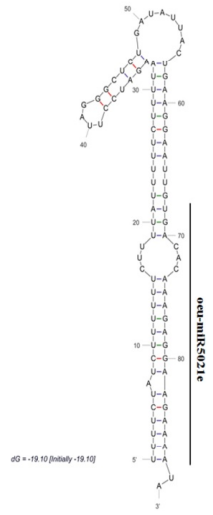


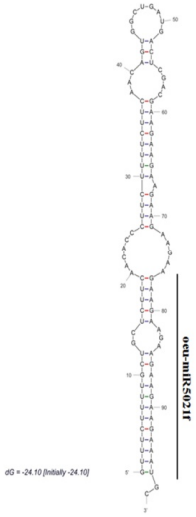
oeu-miR5021a

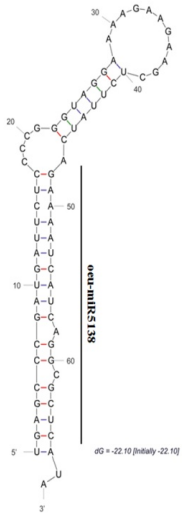

oeu-miR5021b
oeu-miR5021b
dG = -16.40 [Initially -16.40]

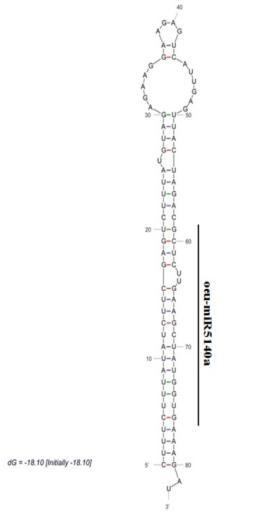

oeu-miR5021c
oeu-miR5021c
dG = -41.40 [Initially -42.40]


oeu-miR5021d
oeu-miR5021d
dG = -17.50 [Initially -17.50]

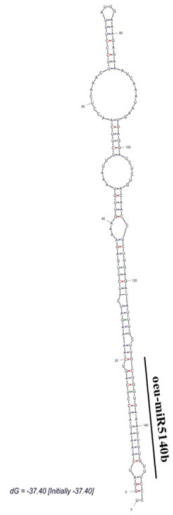

oeu-miR5021e
oeu-miR5021e
dG = -19.10 [Initially -19.10]


oeu-miR5021f
oeu-miR5021f
dG = -24.10 [Initially -24.10]

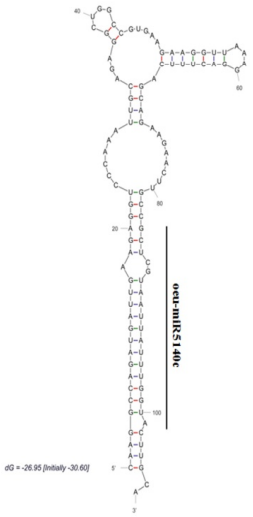

oeu-miR5138
oeu-miR5138
dG = -22.10 [Initially -22.10]



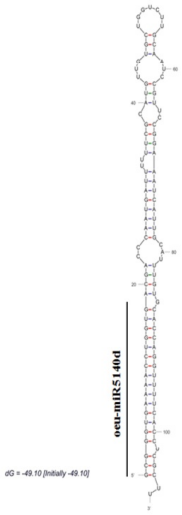
oeu-miR5140a



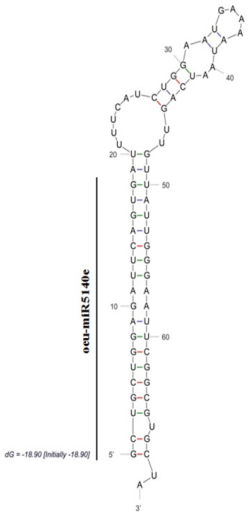
oeu-miR5140b



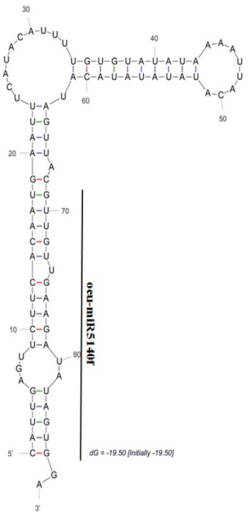
oeu-miR5140c



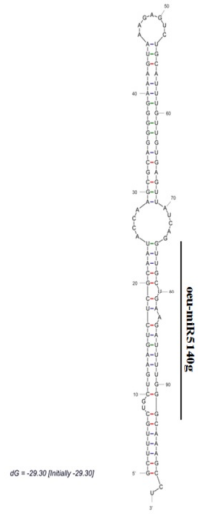
oeu-miR5140d



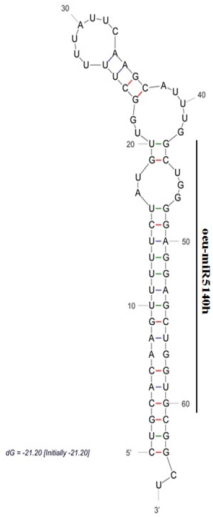
oeu-miR5140e



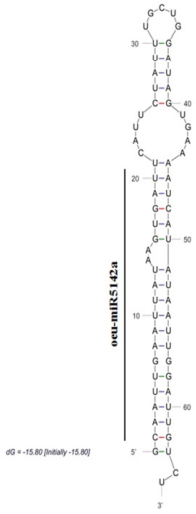
oeu-miR5140f



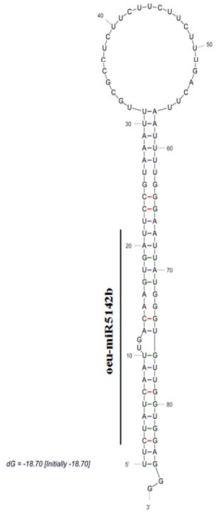
oeu-miR5140g



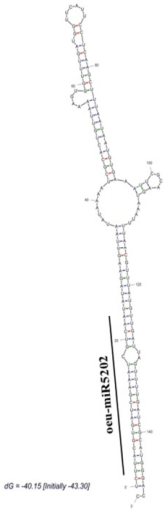
oeu-miR5140h



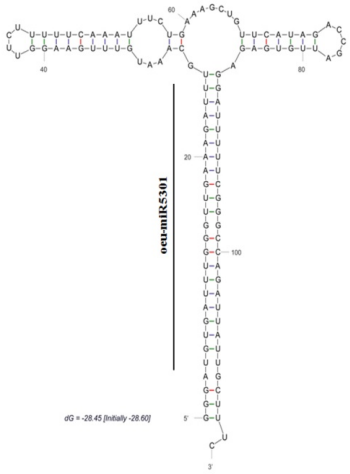
oeu-miR5142a



oeu-miR5142b



oeu-miR5202



oeu-miR5301

CHAPTER 2

MALE-STERILITY IN CHICORY

Male-sterility in chicory

Introduction

Male and female sterility is quite common in flowering plants and it can originate by different mechanisms. The extreme case of male and female sterility is represented by unisexual flowers that arise through the abortion of one of the two sexual organs, such as pistils or anthers, and that can be brought within the same plant, in hermaphroditic species. Unisexual flowers have as direct consequence an obligate cross-fertilization between different plants. Besides the extreme case of unisexual flowers, sterility can also derive from abnormalities of the normal macrosporogenesis and microsporogenesis, as well as gametogenesis, ending with the production of non-functional female or male gametes, respectively. In some cases, even if the male meiosis occurs regularly and the haploid spores are perfectly viable, the lack of microspore release from tetrads or pollen dispersal from anthers leads to a complete male-sterility. In other cases, pollen grains are produced and normally dispersed from anthers but, once landed on compatible stigmas, they are not able to germinate. Therefore, several causes can influence the normal production of viable gametes, from the early differentiation stages of sexual organs to the last steps of gametophyte developmental pathways.

Hereafter will be deeply treated the male sterility trait in plants under both theoretical and practical standpoints. Male sterility (MS) is currently classified into two main categories: cytoplasmic male-sterility (CMS) and genic male-sterility (GMS) controlled by mitochondrial and nuclear genes, respectively. Male sterility systems and the comprehension of their genetic determinants are crucial to shed light on mechanisms at the basis of the microsporogogenesis and microgametogenesis, the pathways involved in normal anther development, and the interactions between nuclear and mitochondrial genes

and proteins. Furthermore, a deep comprehension of male sterility systems is extremely important in plant breeding applications, being male sterility currently used as one of the most effective methods to produce F1 hybrids in crop plants. F1 hybrids are usually constituted by crossing of two highly homozygous parental lines obtaining so a highly heterozygous progeny, which is usually characterized not only by high uniformity of phenotypic traits, but also by strong heterosis in terms of productivity. The best way to breed F1 hybrid varieties is based on a female line, seed parent, unable to produce viable male gametes hence avoiding self-fertilization. For those species in which varieties are represented by F1 hybrids, such as maize, the theoretical condition can be obtained by mechanical emasculation of the plants belonging to inbred lines selected as seed producers. However, male sterility is a less expensive and more effective solution and, above all, can be applied to a large number of species. The possibility to exploit male sterility for the constitution of F1 hybrids in crop species is one of the reasons why breeders have always been attracted by male sterility.

Male sterile phenotypes include a wide range of reproductive abnormalities, such as i) stamens converted to petals or to carpels; ii) degeneration of stamens or specific stamen components, such as the anther of the tapetal cells; iii) pollen that fails the complete development; iv) pollen that completes its development but fails to function (Chase, 2007). Since abnormalities leading to a male-sterile phenotype are different, it may be deduced that different genes are involved in male sterility. In our specific case, male sterility in chicory mutants is due to microspores that fail their development at an early stage and the plants that carry the mutant gene are not able to form pollen grains.

Cytoplasmic male sterility (CMS)

CMS is under extra nuclear genetic control, since it does not show Mendelian inheritance, and it is under the regulation of cytoplasmic factors, usually carried by mitochondrial genome, and as a consequence it is maternally inherited (Vinod 2005). CMS occur in many plant species and it is usually due to chimeric mitochondrial open reading frames (ORFs) encoding for proteins that appear to interfere with mitochondrial function and pollen development. CMS is usually characterized by two types of cytoplasm: the normal (N) and the sterile (S) cytoplasm. Several mitochondrial genes are known to play a role in male sterility and they have been characterized as ORFs which are expressed because they are fused directly to plant mitochondrial promoter sequences or co-transcribed with upstream mitochondrial genes (Chase, 2007). Despite mutations in mitochondrial genes often lead to severe problems of plant development and, in some cases, lethality of plants, mitochondrial genes involved in CMS usually do not affect normal respiratory pathways of the organelle and male sterile plants show a phenotype similar to fertile plants. For these reasons, some of the mitochondrial ORFs determining CMS are considered to be gain-of-function mutations (Chase, 2007). However, some cases of loss-of-function mutations leading to CMS have been characterized. For instance, wild beet (*Beta* spp.) showing a male sterile cytoplasm is widely distributed along Western European coast. Within this male sterile cytoplasm, two mutations of mitochondrial genes were found, the NAD9 subunit has a C-terminal extension and the COX2 subunit has a truncated C-terminal end, which were proven to be involved in activity reduction of the cytochrome *c* oxidase, even though an alternative pathway compensate for the cytochrome *c* oxidase deficiency allowing the life cycle of the plant (Ducos *et al.*, 2001). Similarly, the activation of internal and external alternative NAD(P)H dehydrogenases allowed the survive of those *Nicotiana sylvestris* plants which were characterized by lack of crucial mitochondrial genes (Sabar *et al.*, 2000).

Specific nuclear genes, called restorer of fertility (*Rf*), can suppress or counteract the expression of most CMS-determining genes regaining the fertility to those plants carrying the S cytoplasm. *Rf* genes differ among species and S-type cytoplasms, and they act both in dominant homozygous (*Rf*/*Rf*) and heterozygous (*Rf*/*rf*) conditions. Fertility-restoration systems can be sporophytic, functioning in the diploid plant, or gametophytic, functioning in the haploid pollen. In sporophytic restoration all pollen produced by the plants characterized by *Rf*/- is functional, whereas in gametophytic restoration only pollen carrying the restorer allele is functional (Chase, 2007). CMS-determining mRNAs can be affected by restorer genes through different pathways ending with a decrease of transcripts and proteins or with a fail of translation. Restorer genes usually encode members of the pentatricopeptide-repeat (PPR) protein family, which are proteins that might be untargeted or targeted to the plastidial or mitochondrial organelle. The *Arabidopsis* nuclear genome encodes about 450 PPR proteins, of which 19% are predicted targeted to plastids and 54% targeted to mitochondria (Lurin *et al.*, 2004). Moreover, PPR proteins are supposed to work as site-specific RNA-binding adaptor proteins that mediate interactions between RNA substrates and the enzymes acting on them (Lurin *et al.*, 2004; Chase, 2007).

CMS-determining and *Rf* genes were characterized for the first time in maize (*Zea mays*). Maize has a number of male sterile cytoplasms, such as Texas (T), USDA (S) and Charrua (C). These sterile cytoplasms are found to differ from fertile or normal (N) cytoplasms in mitochondrial translation products, mitochondrial RFLPs and mitochondrial RNAs (Vinod, 2005). For each of those S cytoplasms, a number of restorer genes were isolated. CMS-S has the nuclear *Rf3* gene as restorer of fertility and its restoration is gametophytic. Thus, in S cytoplasm a pollen grain carrying the *Rf3* allele is starch filled and functional, while a pollen grain carrying the *rf3* allele collapses and fails the complete development. Recently, a further gametophytic nuclear *Rf* gene was discovered in CMS-S phenotype and it was called *Rf9* (Gabay-Laughnan *et al.*, 2009). *Rf9* is a less effective restorer of fertility than

Rf3 and its expression is influenced by both inbred nuclear background and temperature (Gabay-Laughnan *et al.*, 2009). The sporophytic restoration system of CMS-T requires at least two nuclear restorer genes, *Rf1* and *Rf2*, which have to be carried together in order to restore the fertility (Schnable and Wise, 1998; Chase, 2007) Vinod, 2005). Interestingly, the *Rf1* restorer of T-cytoplams is quite rare among maize lines, whereas the *Rf2* restorer is present in almost all maize lines, even though most of these lines have never been exposed to the T-cytoplasm, suggesting a conservation of this gene during the evolution (Schnable and Wise, 1998). Two additional restorer genes, *Rf8* and *Rf**, were isolated in CMS-T, though their fertility restoration is distinct from that of *Rf1* and *Rf2* (Vinod, 2005). In CMS-C a restorer gene, *Rf4*, was found in chromosome 8 but the same author affirmed that since the region of chromosome 8 near to *Rf4* is duplicated in chromosome 3, an additional restorer gene for CMS-C might eventually be found (Sisco, 1991).

In natural populations CMS is partially maintained through a selective advantage to females showing CMS, because in some cases gynodioecy flowers set more seeds than hermaphrodite ones. It has also been proposed that the expression of CMS-determining genes in vegetative tissues might induce the expression of nuclear genes involved in stress response, improving the prospects for survival and reproduction of those plants carrying these mitochondrial genes (Budar *et al.*, 2003). Furthermore, since mitochondria are usually inherited by the maternal parent, mitochondrial mutations that compromise plant viability or female fertility would be eliminated quickly from the gene pool, whereas mitochondrial mutations that affect only male fertility can be passed to the subsequent generations through the maternal germ line (Chase, 2007).

Nuclear male sterility

Genic male sterility (GMS) is under control of nuclear genes, which can act both in dominant and recessive homozygous states. Contrarily to CMS, GMS is biparentally inherited and therefore it shows a Mendelian segregation according to the number of genes involved. Currently GMS is divided into three main categories according to its genetic control: i) GMS controlled by double recessive genes; ii) GMS controlled by interacting recessive genes; iii) GMS inherited as a dominant gene (Huang *et al.*, 2007). Theoretically, in order to produce F1 hybrids, CMS is more suitable since it is inherited by maternal parents and, by crossing a maternal parent carrying S cytoplasm and *rf*/*rf* restorer genes with a paternal parent showing N cytoplasm and *rf*/*rf* restorer genes, all the progeny plants will be sterile and the maintaining of male sterility trait is therefore easy. On the other hand, GMS can be maintained only by crossing a male sterile parent as seed producer with a male fertile parent as pollen donor, with the progeny that may segregate for the trait. In particular, the best segregation ratio is 1:1, male sterile vs. male fertile plants, that can be observed by crossing a homozygous recessive male sterile seed producer with an heterozygous male fertile pollen donor. However, GMS has more advantages than CMS, such as stability and complete expressivity of the male sterility trait, many sources of cytoplasm, and easy transfer of male sterility genes (Huang *et al.*, 2007).

Gene mutation, interspecific and intraspecific hybridization, radiation, and genetic engineering can lead to the nuclear male sterility. Natural and transgenic GMS is also classified into three categories: i) sporogenous, ii) structural, and iii) functional (Kaul *et al.*, 1988). Sporogenous GMS occurs in stamen that appears normal but, due to abnormalities occurring between early microsporogenesis and late microgametogenesis, plants produce non-functional microspores or non-viable pollen grains. Structural GMS instead is characterized by a pollen-less anther or abnormal pollen (Johns *et al.*, 1981). Functional

GMS is characterized by the production of viable pollen that is not released from the anther (Horner and Palmer, 1995).

Several GMS cases, differing each other for genes involved, were so far documented. Problems leading to GMS can affect tapetal cells, pollen mother cells (PMC) and microspores. In anthers, tapetal cells provide nourishment to PMCs and to developing microspores and it represents the source of a great amount of genes involved in pollen formation. Abnormalities affecting the anther tapetum are crucial for the normal pollen development. For instance, it has been demonstrated that in one male sterile line of *Brassica napus*, where the pollen development is affected during tetrad stage, that tapetal cells swelled with expanded vacuoles at the early tetrad stages (Wan *et al.*, 2010). Authors supposed that a lack of callase, which is an enzyme required to degrade the callose of tetrads, is the cause of the sterility (Wan *et al.*, 2010). Furthermore, after the cross between *Brassica napus* and *Capsella bursa-pastoris*, a male sterile mutant was isolated and it showed a segregation of male fertile and male sterile of 3 : 1 and 1 : 1 in F2 and BC1 respectively, meaning that the trait was controlled by a single nuclear recessive gene (Chen *et al.*, 2009). Histological analyses revealed that the meiotic chromosome pairing and segregation were normal, but the tapetum was multiple layers and it was hypertrophic compressing so the tetrads that degenerated (Chen *et al.*, 2009). Abnormalities of tapetum development were also found in rice. A systematic cytological comparison of anthers in GMS mutants and fertile plants of rice showed that the pollen abortion first occurred before the PMC stage and continued during the entire process of pollen development ending with the pollen degradation (Shi *et al.*, 2009). It has been demonstrated that the abortive process was closely associated with the abnormal behavior of the tapetum, which initiated its degeneration at the PMC stage and proceeded slowly missing the final pollen development (Shi *et al.*, 2009).

Several genes have been demonstrated to be the cause of GMS and they can act as single locus or after the interaction among genes at different loci. Usually genes involved in GMS produce sterile phenotypes in recessive condition. For instance, in one spontaneous male sterile line of rice, pollen development proceeds normally until vacuolization stage. Male sterile plants fail to vacuolate and no viable pollen is produced. Genetic analyses have demonstrated that a single nuclear recessive gene (*vr1*), mapped to the chromosome 4, was the responsible of sterility (Chu *et al.*, 2011). In one male sterile line of *Brassica napus*, the sterility was controlled by a recessive epistatic genic two-type line system. The sterility was demonstrated to be under control of two duplicate genes, *Bnms3* and *Bnms4* and one recessive epistatic inhibitor gene, *Bnrf*. Homozygosity at the locus *Bnrf* (*Bnrfrf*) inhibits the expression of the two recessive male sterility genes in homozygous *Bnms3ms3ms4ms4* plants and produces a male fertile phenotype (Huang *et al.*, 2007). In another spontaneous male sterile mutants of *Brassica napus*, the sterility was controlled by two duplicate recessive genes, *Bnms1* and *Bnms2*, which were supposed to be the cause of sterility since in F2 there was a segregation of 15:1 of male fertile against male sterile (Yi *et al.*, 2006).

Even though GMS is often controlled by one or more recessive genes, dominant GMS can also occur. One male sterile line of *Brassica napus* is conditioned by epistatic interaction of two genes: the dominant gene *Ms* and suppressor gene *Rf* (Li *et al.*, 1985). In this case, the expression of *Ms*- results in male sterile phenotype but the co-expression of the suppressor *Rf*- regain the fertility (Lu *et al.*, 2004).

In addition to the genetic mechanisms, normal stamen and pollen development can also be affected by internal substances, called plant growth substances (PGSs), such as cytokinins, auxin, abscisic acid and ethylene, which are thought to directly or indirectly influence male sterility (Sawhney and Shukla, 1994).

Use of male sterility in F1 hybrid seed production

High level of heterozygousity often positively affects plant phenotypes in terms of adaptativeness and fitness of several species, and these effects are pooled under the term of "heterotic vigor". This heterotic vigor is directly associated to a better status of plants along with a higher productivity seeds. The best method to maximize the level of heterozygosity is to cross two pure or inbred lines characterized by high level of divergent homozygosity. Highly homozygous lines can be obtained by repeated cycles of selfing in those species self-compatible or by recurrent back-crossing with the same parent in those species characterized by a self-incompatibility or male sterility. Theoretically, after each selfing or back-crossing generation the level of homozygosity increases by 50%, thus assuming to start from a 100% heterozygous genotype, after the first selfing or back-crossing the percentage of heterozygosity decreases to 50%, after the second round this percentage decreases to 25%, then to 12.5%, and so on. At the end, the controlled cross between two highly divergent homozygous lines may yield F1 hybrid seeds having nearly 100% heterozygosity.

The high potential of F1 hybrids and its application to crop plants have always been very attractive for breeders. A female plant unable to set viable male gametes is the best way to produce F1 hybrids, since a female homozygous lines can be easily used as seed producer to be crossed with a high homozygous hermaphrodite or male plant. Emasculation, action aimed to manually remove male organs, can be applied to those species characterized by female and male flowers spatially separated within the same plant, like maize having the male organs located in the upper part of the plant. For all those species characterized by hermaphrodite flowers, emasculation is only applicable for research purposes but it is not eligible as technique to breed commercial F1 hybrids. Thus, male sterile plants may represent the only way exploitable for large-scale production of F1 hybrid seeds in hermaphrodite plants, which are the most common among crops.

Male sterility can independently arise many times for several reasons and, once a mutant showing this trait has been identified in a given crop plant, its maintaining is crucial. Several breeding schemes to preserve the male sterility trait are variable according to cytoplasmic or nuclear genetic control (**Figures 2.1, 2.2 and 2.3**). At the end, crosses between antagonist highly homozygous female and male parental lines generate F1 hybrids, which are highly heterozygous and characterized by great phenotypic uniformity. F1 hybrid seeds can give rise to sterile or fertile plants, according to the male sterility system, *i.e.* CMS or GMS, and according to the breeding scheme adopted.

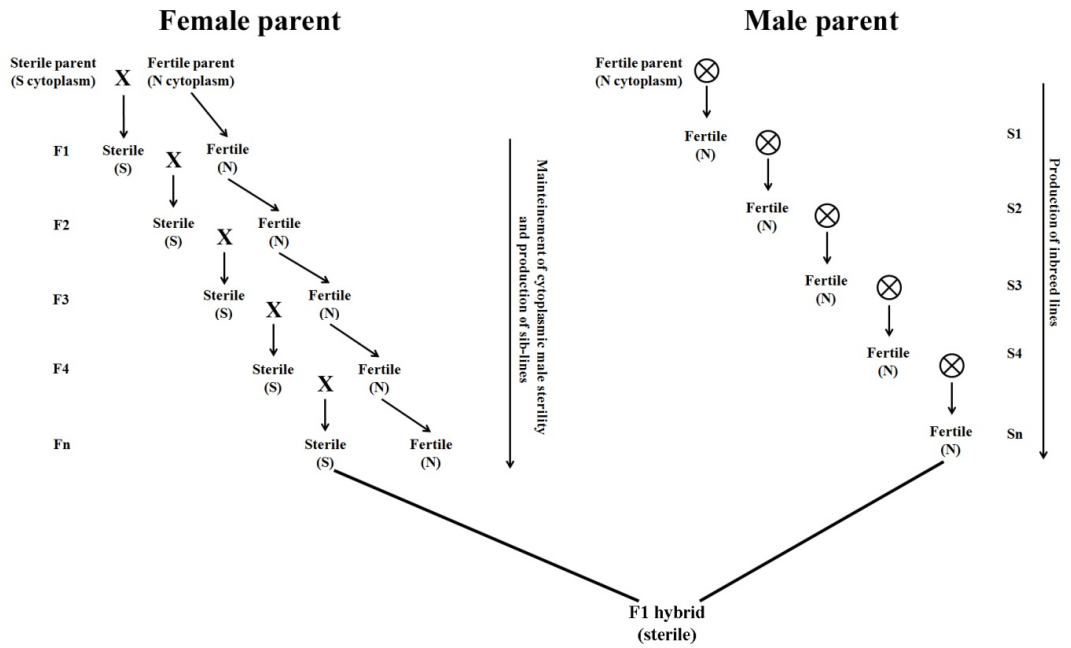


Figure 2.1. Breeding schemes for maintaining the sterility in CMS without the presence of restorer genes (*Rf*) and for obtaining highly homozygous lines to produce F1 hybrid seeds. In this example F1 hybrids are sterile due to the cross between the sterile cytoplasm (S) from the seed producer (female parent) and the normal cytoplasm (N) from the pollen donor (male parent).

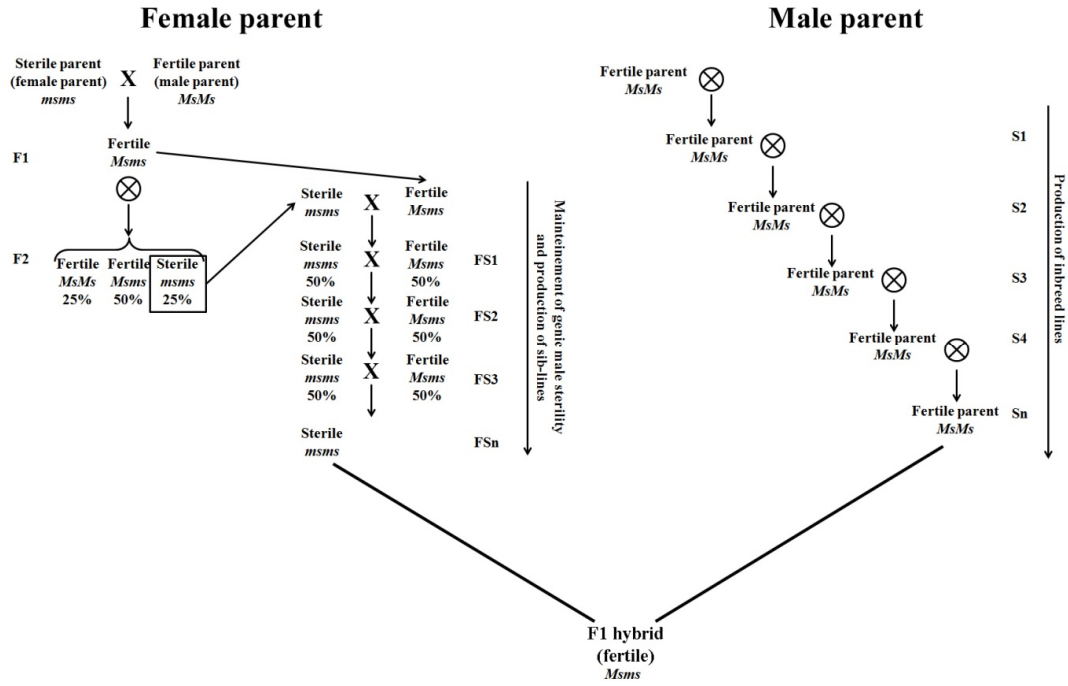


Figure 2.2. Breeding schemes for maintaining the sterility in GMS and obtaining highly homozygous lines to produce F1 hybrid seeds. In this example F1 hybrids are fertile (*Msms*) due to the cross between a homozygous recessive (*msms*) seed producer (female parent) and a homozygous dominant (*MsMs*) pollen donor (male parent).

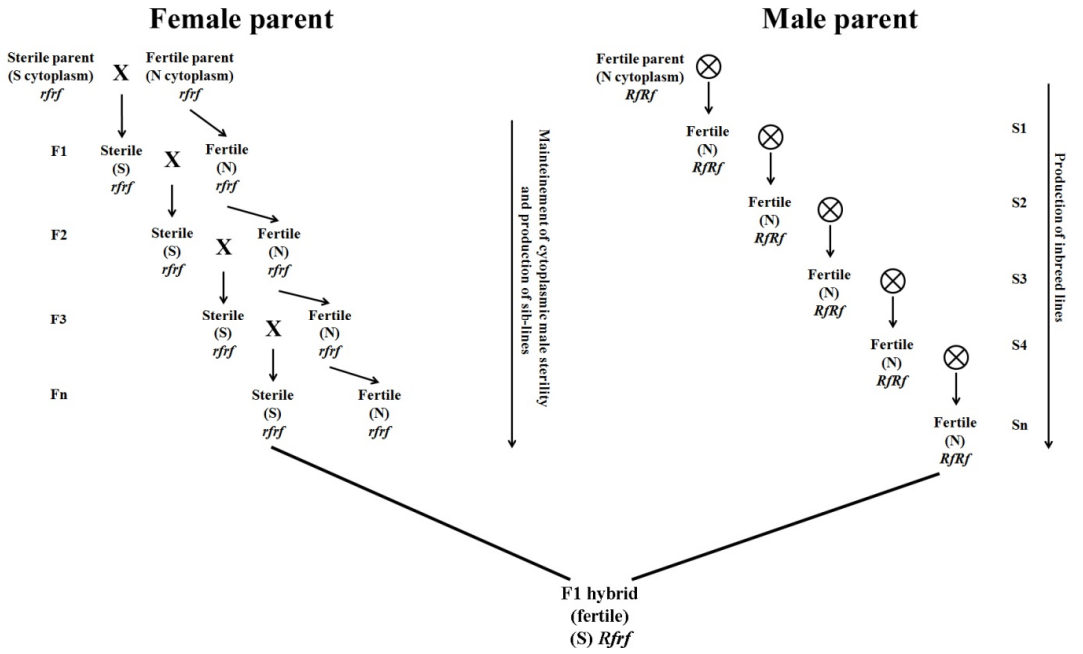


Figure 2.3. Breeding schemes for maintaining the sterility in CMS with restorer genes (*Rf*) and obtaining highly homozygous lines to produce F1 hybrid seeds. In this example F1 hybrids are fertile due to the cross between seed producer (female parent) carrying a sterile cytoplasm (S) and a homozygous recessive restorer gene (*rfrf*) and pollen donor (male parent) carrying a normal cytoplasm (N) and homozygous dominant (*RfRf*) restore gene. Pollen donor can also exhibit a sterile cytoplasm (S) and a homozygous dominant (*RfRf*) restorer gene.

Male sterility in chicory

Chicory (*Cichorium intybus* L.) is a diploid plant species (2*n*=18), belonging to the *Asteraceae* family, subfamily *Cichoriodeae*, tribe *Lactuceae* or *Cichorieae* (Bremer, 1994; Panero and Funk, 2002). This species is naturally allogamous, due to an efficient sporophytic self-incompatibility system (Eenink, 1981)(Varotto *et al.*, 1995)(Barcaccia *et al.*, 2003). In addition, outcrossing is promoted by a floral morpho-phenology (*i.e.*, proterandry, having the anthers mature before the pistils) unfavourable to selfing in the absence of pollen donors (Pécaut, 1962; Desprez *et al.*, 1994) and by a favourable competition of allo-pollen grains and tubes (*i.e.*, pollen genetically diverse from that produced by the seed parents, usually called auto-pollen) (Desprez and Bannerot, 1980; (Eenink, 1982). Long appreciated as a medical plant by ancient Greeks and the Romans (Wittop Koning and Leroux, 1972; Lucchin *et al.*, 2008), red chicory is nowadays one of most important cultivated vegetable crops, being used mainly as component for fresh salads or more rarely cooked according to local traditions and alimentary habits (reviewed by Lucchin *et al.*, 2008). At present, this species is grown all over continental Europe, in South Western Asia, and on limited areas in Northern America, South Africa, and Australia. It is worth mentioning that in Mediterranean Basin countries during the last centuries red chicory gradually underwent a process of naturalization and although this species cannot be considered as autochthonous, it became part of the natural and agricultural European flora.

Two main groups can be recognized within *C. intybus* subsp. *intybus* to which all the cultivated types of chicory belong: the first, which refers to the var. *foliosum*, traditionally includes all the cultivar groups whose commercial products are the leaves, while the second regards the var. *sativum* and comprises all the types whose commercial product, either destined to industrial transformation or direct human consumption, is the root. Among the cultivar groups, “Radicchio” is the Italian common name that has been adopted by all the most internationally used languages to indicate a very differentiated group of chicories,

with red or variegated leaves, traditionally cultivated in North Eastern Italy (Lucchin *et al.*, 2008). There is no documented history about the origin of coloured chicory in Italy. All the red types of Radicchio now being cultivated seem to derive from red-leaved individuals firstly introduced in XV century. According to Bianchedi (1961) the cultivation of red chicory goes back to the first half of XVI century. For sure, the original type has to be identified with the "Rosso di Treviso" which has been for long the only cultivated Radicchio in the Venetian territories. Originally selected around 1930, nowadays "Rosso di Chioggia" is by far the most widely grown among the various types of Radicchio and the one which presents the highest within-type differentiation as far as the availability of cultivars able to guarantee an almost complete year round production. As a matter of fact, it has shown a great adaptability to very different environmental situations all around the world, becoming the most grown type of Radicchio outside the Italian country and the most known at international level (Lucchin *et al.*, 2008).

The original populations of *C. intybus*, as far as their genetic structure is concerned, could be considered as natural since, independently of their historic background, the production of Radicchio has for long time relied on populations, maintained by farmers for their own use, on which very little selection might have been applied according to personal criteria ((Barcaccia *et al.*, 2003) Lucchin *et al.*, 2008). All these populations, obtained by mass selection and maintained through the intercrossing of selected parents, have to be considered highly heterozygous and genetically heterogeneous whose behavior and level of adaptation to different environments and/or cultural conditions depend on the frequency of favorable genes or gene combinations. As the interest for the edible product grew, farmer's selection criteria became more and more attentive to the consumer's request and most of them elaborated their own ideotype. This brought about a great deal of genetic and morphological differentiation that has been entirely preserved until organized breeding programs have been established, firstly by public institutions and, in more recent times, by

private companies. In Radicchio, although with some differences among the various types, the major part of the crop is still based on farmer's populations which are yearly selected and maintained and whose seed is usually reutilized on farm but may also be sold through private and not officially registered transactions. These populations are very well distinguishable among types, but they are often recognizable within type as well, on the basis of morphological and physiological characters and agronomic performances, although they present also an acceptable phenotypic uniformity among individuals. Regarding their genetic variation, as estimated by genetic analysis performed through the application of appropriately chosen molecular markers, it is a common observation that the major part of the genetic variation takes place within populations, while a minor portion is attributable to among population polymorphisms (Barcaccia *et al.*, 2003). Since this is the plant material which, in recent years, has been representing, and still represents, the starting point for the constitution of new commercial varieties, it seems reasonable to state that, if preserved from extinction, it is an invaluable source of genetic diversity on which chicory breeding may rely for long in the future.

Traditionally, varieties of chicory were developed by mass selection in order to obtain uniform populations characterized by valuable production and acceptable commercial head size and shape. Newly released varieties are mainly synthetics produced by intercrossing a number of phenotypically superior plants, selected on the basis of morpho-phenological and commercial traits. More rarely, plants are also evaluated genotypically by means of progeny tests. Synthetics have a rather large genetic base and are represented by a heterogeneous mixture of highly heterozygous genotypes sharing a common gene pool. In recent years, methods for the constitution of F1 hybrids have been developed by private breeders and seed firms. Details on the procedure for the constitution of such hybrids are not available in the current literature and it may be presumed that each company has developed its own protocol, mainly in accordance to the genetic material it has at disposal

and to the possibility of applying a more or less efficient control on the F1 hybrid seed production phase. As a matter of fact, the strong self-incompatibility system, which hinders obtaining highly homozygous parents, and the absence of a male-sterility factor within the species or in sexually compatible species, made it generally difficult to propose an efficient F1 seed production scheme and, most of all, to consider these newly commercial varieties as true F1 hybrids.

As it happens for most allogamous species, in *C. intybus* detectable heterosis effects are present and hybridization between genotypes selected on the basis of their specific combining ability gives vigorous and uniform progenies. Consequently, the constitution of F1 hybrid varieties is profitable in a practical breeding scheme and it is also feasible on a large commercial scale by the selection of self-compatible genotypes, for the production of inbred lines, and the identification of genotypes showing male-sterility, to be used as see parents for the hybridization with pollen donors. It is therefore expected that F1 hybrid varieties will be bred and adopted with increasing frequency for Radicchio. This is particularly true for the cultivated types that take a great advantage from the uniformity of the marketed products, as this is often the key for the customer's appreciation.

Male sterility is defined as the failure of plants to produce functional anthers or pollen grains. It is more prevalent than female sterility, likely because the male sporogenesis and gametophytes (*i.e.* pollen tubes) are less protected from environment than the female sporogenesis and gametophytes (*i.e.* embryo sacs). Male sterile plants have propagation potentials in nature because they can still set seeds, being female fertility unaffected by most of the mutations responsible for male sterility. Male sterility is known to occur spontaneously via mutations in nuclear and/or cytoplasmic genes involved in the development of anthers and pollen grains. In the model plant Arabidopsis, about 3,500 genes may be specifically expressed in anthers and many of them are required for the production of pollen grains (Goldberg *et al.*, 1993; Ma, 2005). Studying male sterility

mechanisms is immensely important not only for the understanding of reproductive barriers that act in flowering plants, but also because of their potential applications for the breeding of F1 hybrid varieties. If it is true that hybrid varieties in most of the crop species have been dramatically successful so far, it also seems likely that hybrid varieties will continue to expand in the future.

In the last century, male sterile mutants have allowed the exploitation of heterosis (*i.e.* hybrid vigour, meant as improved characteristics in terms of size, resistance, growth rate, fertility and crop yield of a hybrid offspring over those of its parents) through the constitution of F1 hybrid varieties in many agricultural and horticultural crops. Two kinds of male sterility can be observed in plants: nuclear and cytoplasmic male sterility. The former type of genetic male sterility is based solely on recessive mutations that affect different functions in nuclear genes (*msms*), while cytoplasmic male sterility (CMS) is maternally inherited and mainly due to mutations in the expression of mitochondrial genes. Moreover, in genotypes showing CMS, male fertility can be eventually restored by nuclear-encoded fertility restorer (*Rf*) genes. In several species, nuclear and/or cytoplasmic male-sterility has been used to produce female parental lines and exploited for the production of hybrid seeds through controlled pollination with male parental lines showing specific combining ability. In absence of an efficient genetic male-sterility system, anthers must be removed mechanically from flowers of seed parents.

This chapter deals with the discovery and genetic analysis of male sterile mutants of red chicory. These mutants, which to the best of our knowledge are the first spontaneous male sterile mutants ever discovered and described in the genus *Cichorium*, were characterized in great details for the developmental pathway of micro-sporogenesis and gametogenesis, and the inheritance pattern of the gene underlying the male-sterility trait. Moreover, the fine mapping of the mutant locus was attempted by using microsatellite markers. Experimental

results on the male sterile mutants are presented and their use for the breeding of new F1 hybrid varieties is critically discussed.

Materials and Methods

Plant materials

Four distinct but genetically related male sterile mutants of red chicory, belonging to the biotype "Rosso di Chioggia", were recently discovered on the basis of morphological observations of anthers within experimental synthetic varieties bred by T&T® Produce, Sant'Anna di Chioggia, province of Venice, Italy (**Figure 2.4**).

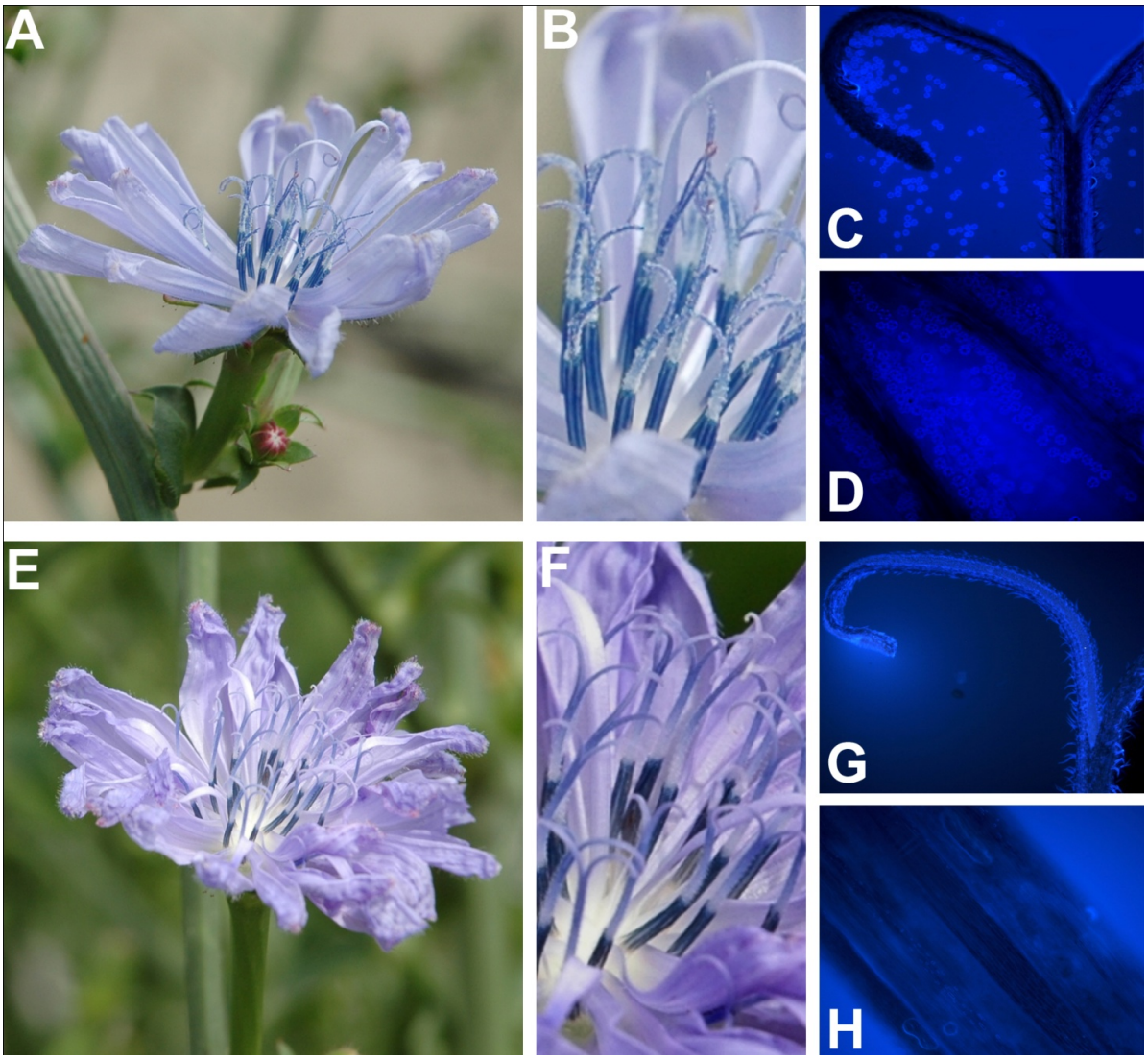


Figure 2.4. Phenotype of wild-type plants (A-D) and male-sterile mutants (E-H) of red chicory (*Cichorium intybus* L.). These mutants belong to experimental synthetic varieties of the biotype "Rosso di Chioggia" bred by T&T® Produce, Sant'Anna di Chioggia, province of Venice, Italy. Details of macroscopic (A, B) and microscopic (C, D) features are given for the wild-type anthers in parallel with mutant anthers (E, F and G, H, respectively).

In particular, one mutant each was found among the mother plants of synthetics coded as CH03/85 and CH04/99, while two mutants were found among the mother plants of CH05/01. The male-sterile mutants analyzed in this study were named L11ms, IG9ms, CS1ms and CS2ms. The three populations from which they originate have been bred through genetic selection based on progeny tests performed using mother plants chosen for uniformity and superiority of their morphological and agronomic traits. In particular, the synthetic CH05/01 is a direct derivative of CH04/99, being they bred for a different earliness (90 and 110 days, respectively), whereas the synthetic CH03/85 was bred independently for an even shorter cycle (80 days). Nevertheless, these experimental varieties are differentiated each other not only for their earliness, but also for esthetical and organoleptic traits of the commercial head (for details on varietal traits of Radicchio, see www.vegetableseeds.it). Both productive and qualitative traits have been main objectives of the red chicory breeding programs fulfilled in recent years by T&T® Produce. General and common goals in breeding new varieties mainly concerned: i) single plant size, weight and yield; ii) resistance to biotic (*e.g.*, fungal diseases and insects) and abiotic stresses; iii) adaptation to a specific climatic or agronomic environment; iv) uniformity of crop maturity; v) good market acceptance regarding extrinsic (*i.e.*, leaf color, head shape, uniformity and consistency) and intrinsic (*i.e.*, taste and texture) traits. It is worth mentioning that the three T&T experimental synthetics, from which all male sterile mutants come from, share a common ancestral genetic background because they derive from experimental populations produced around 1970-80s by recurrent massal selection from an old Venetian ecotype of red chicory (dated back to 1940-50s). In fact, in red chicory this type of selection was shown to be effective for increasing the frequency of favourable genes and genotypes underlying important Mendelian (*i.e.* monogenic) and quantitative (*i.e.* polygenic) traits, especially if under the control of QTLs showing high heritability (Barcaccia and Parrini, 2001). The best synthetic varieties among those showing narrow genetic base recently bred

by T&T® Produce and T&T® Vegetable Seeds through marker-assisted breeding programs were positively evaluated by the technical commissions to be registered in the European Catalogue of Plant Varieties.

Cytological analysis of male sporogenesis and gametogenesis

The presence of pollen within anthers was assayed by whole mount staining with DAPI (4',6-diamidino-2-phenylindole), a fluorescent stain that binds strongly to A-T rich regions in DNA. Anther heads isolated from five flowers for each of the male-sterile mutants and the wild-types were squashed on a microscope slide and treated with 10 µl of staining solution (DAPI 5 µg/ml). After an incubation of 10 min, a detailed observation of stained anthers was done by a Leica DM4000B imagine microscope using the appropriate filter combination for DAPI fluorescent detection. Pictures were taken by the Leica DC300F camera and digital images at 10X or 20X magnification were screened in great details for the presence vs. absence of pollen grains using Adobe Photoshop® CS4 (Adobe Inc., U.S.A., www.adobe.com/it/products/photoshop).

An alternative staining technique was used to investigate the pattern of micro-sporogenesis and the development of pollen grains in each male-sterile mutant in comparison with wild-type. Flowers at four different developmental stages, spanning from young buds to full anthesis, were collected from mutants and wild-type plants, fixed in Carnoy's solution (ethyl alcohol-acetic acid 3:1) and stored at +4°C for 24-48 hours. After this pre-treatment, flowers were transferred in 70% ethyl alcohol at +4°C until their use for cytological analysis. Anthers were dissected from individual flowers, opened on microscope slides using a pair of teasing needles with the aid of a stereomicroscope. Specimens containing pollen mother cells, tetrads, microspores and pollen grains were squashed using a drop of 4% aceto-carmine and mounted in lacto-phenol with acid fuchsin.

For the preparation of meiocyte chromosomes, anther specimens of mutants and wild-types were treated with citrate buffer (10 mM citric acid, 10 mM sodium citrate, pH 4.5) for 3 min and incubated in a six times diluted pectolytic enzyme mixture containing 1% pectolyase Y23, 1% cellulase RS and 1% cytohelicase (Sigma Aldrich, www.sigmaaldrich.com) in 10 mM citrate buffer at 37°C for about 1-2 hours, according to the anther stage. Anther preparations were squashed on microscope slides using a drop of purified and deionized water (Milli-Q Integral Water Purification System, www.millipore.com) and then transferred on a hot plate at 45°C. Cells were spread on microscope slides using a teasing needle by adding one drop of 45% acetic acid, then maintained at 45°C for 2 min and washed with Carnoy's solution. Each slide was dried on the hot plate at 45°C and specimens were stained with DAPI.

Cytological observations of male meiosis and gametogenesis as well as karyological analysis of meiocyte chromosomes were made under natural and fluorescent light using a photomicroscope (Zeiss Axiophot photomicroscope, www.zeiss.com) equipped with epifluorescence illumination and single-band filters for DAPI. Photograph films were scanned at 1,200 dpi for digital image processing with Adobe Photoshop® CS4 (Adobe Inc., U.S.A.).

Genetic analysis of mutants and inheritance of male sterility

Each of the male sterile mutants was crossed as seed parent with a wild type pollinator belonging to the same variety. Several F1 plants from each hybrid population were then selfed and crossed in pair-wise combinations in order to obtain segregating F2 progenies. Moreover, F1 plants were also backcrossed as pollen donors with either male sterile mutants belonging to F2 progenies or wild type plants of S1 progenies stemmed from selfing in order to obtain segregating BC1 progenies (**Figure 2.5**). The experimental populations segregating for the male sterility/fertility trait were composed of about 100

plants for each of the four mutants. These populations were used for genetic analyses in order to establish the inheritance pattern of the mutation (*e.g.*, dominant/recessive nuclear vs. cytoplasmic) and to finely map the male sterility gene using microsatellite markers.

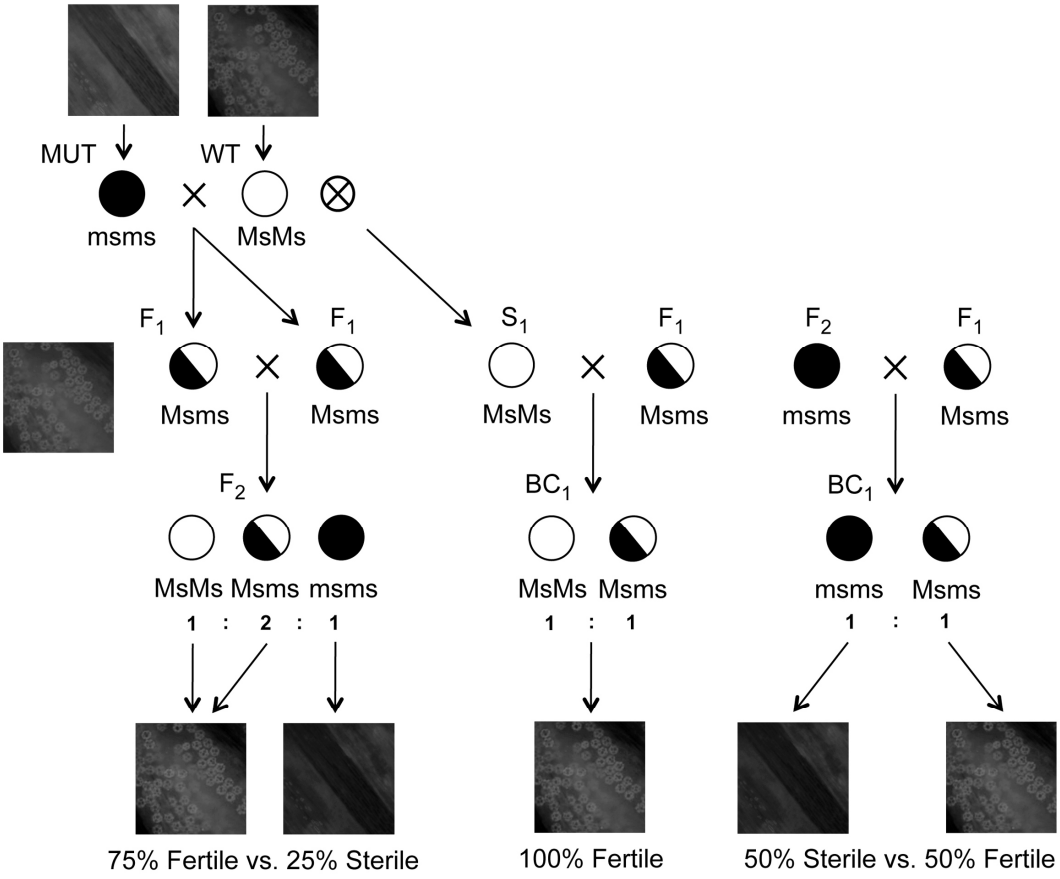


Figure 2.5. Genetic analysis of male-sterile mutants based on segregation patterns observed in F2 and BC1 progenies. Each of the male sterile mutants was crossed as seed parent with a wild type pollinator belonging to the same variety. Several F1 plants from each hybrid population were then selfed and crossed in pair-wise combinations in order to obtain segregating F2 progenies. Moreover, F1 plants were also backcrossed as pollen donors with either male sterile mutants belonging to F2 progenies or wild type plants of S1 progenies stemmed from selfing in order to obtain segregating BC1 progenies. These experimental populations were used to establish the inheritance pattern of the mutation and to map the male sterility gene using microsatellite markers.

Molecular mapping of the gene for male sterility

A total of 118 F2 progeny plants and 92 BC1 progeny plants segregating, respectively, 3:1 and 1:1 for the male fertility vs. sterility trait were used for mapping the *ms* locus using AFLP and SSR markers. The F2 plants derived from the progenies of mutants CS1ms and CS2ms, while the BC1 progenies included plants of mutants L11ms. Moreover, 100 plants of the segregating progenies of mutant IG9ms were also analyzed in order to validate molecular markers tightly co-segregating with male sterility.

Total genomic DNA was isolated from 100 mg of fresh leaf tissue using the DNeasy® Plant mini-kit (QIAGEN, www.qiagen.com) following the recommendations of the manufacturer. The DNA pellets were washed twice with 70% ethanol, dried and resuspended in 100 µl of TE 0.1X buffer (Tris-HCl 100 mM, EDTA 0.1 mM pH 8). The quality of DNA samples was assessed by electrophoresis on 0.8% (p/v) agarose gels, and its concentration was determined by optical density reading (DU650 spectrophotometer, Beckman) at 260 nm (1 O.D. = 50 µg/ml). The purity was calculated by the O.D.260/O.D.280 ratio and by O.D.210-O.D.310 pattern (as described by Barcaccia *et al.*, 2003).

A subset of 48 progeny plants with a contrasting microgametogenesis pattern, (*i.e.*, 24 male sterile plants and 24 male sterile plants) were selected and used for performing a bulked segregant analysis, BSA (Michelmore *et al.*, 1991) in the attempt to identify molecular markers linked to the male-sterility trait. Genomic DNA bulks of 12 plants each from two progeny sets were prepared by combining equal amounts of DNA from male fertile and male sterile plants. All bulked DNA samples were investigated by AFLP markers using the parental lines as controls.

Genomic AFLP fingerprinting was performed using the protocol of (Vos *et al.*, 1995) with modifications described by (Barcaccia *et al.*, 2003). AFLP analysis was based on the detection of *Eco*RI-*Mse*I genomic restriction fragments by PCR amplification with 9

different primer combinations having three selective nucleotides (E+CAC, E+CCA, E+CTG and M+ATC, M+AGG and M+AAG), chosen during preliminary tests according to their ability to find homologous binding sites in red chicory templates. Briefly, after restriction of 500 ng of genomic DNA with *Eco*RI and *Mse*I endonucleases, pre-amplification reactions were performed in a final volume of 20 µl with *Eco*RI and *Mse*I primers carrying one selective nucleotide. Then, 20 cycles were carried out at 94°C for 30 s, 56°C for 60 s and 72°C for 60 s in a thermal cycler GeneAmp® System 9700 (Applied Biosystems). The *Eco*RI primer was labelled by phosphorylating the 5' end with [γ-^{33}P]ATP and T4 kinase, incubating the reaction at 37°C for 1 h, as described in the manufacturer's instructions. The pre-amplified DNA was diluted 1 : 1 in Tris-EDTA buffer and was used as template for hot-PCRs with a *Mse*I primer carrying three selective nucleotides in combination with a *Eco*RI radiolabelled primer, carrying two selective nucleotides at the 3' end. Selective amplification was carried out under the cycling conditions set up by (Barcaccia *et al.*, 2000), which begins with one cycle at 94°C for 30 s, 65°C for 30 s, and 72°C for 60 s. The annealing temperature was then reduced each cycle by 0.7°C according to a touch-down profile of 13 cycles to reach the optimal annealing temperature of 56°C. Twenty-three cycles were run to complete the final amplification at 94°C for 30 s, 56°C for 30 s and 72°C for 60 s.

After amplification, PCR reactions were stopped with equal volume of loading buffer (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol) and denatured at 94°C for 5 min. The labelled, restricted and selectively amplified DNA fragments were separated by electrophoresis on 5% denaturing polyacrylamide gels with 8 M urea at 80 W constant power using a standard DNA sequencing unit Sequi-Gen GT-system (BIO-RAD, www.bio-rad.com). Gels were dried at 80°C for 1 h and then visualized by autoradiogram after overnight exposure on an X-ray film at -80°C using intensifying screens. The AFLP fragment analysis was performed using the 1D® Image analysis

software (Kodak Digital Science). Overall data were recorded as a binary matrix by assigning the molecular weight to each quantitatively polymorphic marker identified by comparing DNA fingerprints with known DNA ladders.

Microsatellite (SSR) loci analysis was carried out following an already tested PCR protocol (Ambrosi *et al.*, 2010) with some changes to adapt it to red chicory templates. The detection was performed with the use of the 5′ M13-tailed primer method (Hayden *et al.*, 2008). DNA fragments were visualized by capillary electrophoresis after amplification reactions performed with the universal M13 primer (the sequence of the tail is the following: 5'-TTGTAAAACGACGGCCAGT-3') labeled with a HEX, FAM or TAMRA fluorophore (by Life Technologies, www.invitrogen.com). PCR experiments were conducted in a 20 µl total volume, including 10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl2, 200 mM of each dNTP, 3 pmol of primer forward, 8 pmol of primer reverse, 6 pmol M13-labeled primer, 1 U *Taq* DNA polymerase (GE Healthcare) and 25 ng of genomic DNA as template. All individual DNA samples were then investigated with 9 SSR markers belonging to as many mapped loci, one for each of the nine linkage groups (*i.e.*, basic chromosomes) of the genetic map recently constructed by (Cadalen *et al.*, 2010). Amplification reactions were performed in a 9700 Thermal Cycler (Applied Biosystems): the temperature profile consisted of an initial denaturation step of 5 min at 95°C followed by 40 cycles of 30 sec at 95°C, 30 sec at annealing temperature of 55–58°C, and 30 second at 72°C, followed in turn by 7 min at 72°C and then held at 4°C. DNA fragment analysis was carried out using a fully automated capillary electrophoresis system (Applied Biosystems 3130) and SSR patterns were visualized and scored in replicated analysis using the software GeneScan® v. 2.1 e Genotyper® v. 2.0 (Applied Biosystems).

As a preliminary screening based on SSR markers, 12 male sterile and 12 male fertile genomic DNA plants were randomly selected from segregating populations of each mutant, for a total of 96 plants. For the marker alleles showing to significantly co-segregate with the

male-sterility/fertility genotypes, the analysis was extended to all the 300 plants available on the whole for this study. The observed segregation ratio of SSR markers was tested by chi-square analyses for goodness-of-fit to the expected 3:1 or 1:1 segregation ratios, as well as for independent assortment in the male-sterile vs. wild-type progenies by a 2 × 2 contingency test. Segregation data for the markers were analyzed with JOINMAP® v. 2.0 (Stam and Van Ooijen, 1995) using the cross pollination (CP) population type option (*i.e.*, segregating populations resulting from a cross between two heterogeneous parents that were heterozygous and/or homozygous at the loci being tested). The association between microsatellite markers and male sterility was assessed by recording the target *ms* locus as a putative monogenic marker fully co-segregating with the trait being mapped. For the genotype code option, presence of marker allele and wild-type phenotype were assigned to *aa* = homozygous dominant or *ab* = *heterozygous*, and absence of marker allele and male-sterile phenotype to *bb* = homozygous recessive. For the identification of the linkage group carrying the *ms* locus with the selected SSR markers, the grouping module was applied by setting a minimum LOD = 3 and a maximum recombination frequency, r = 30% (Barcaccia *et al.*, 2000). The genetic distance between each pair-wise comparison of SSR marker locus and *ms* locus, expressed in centimorgans (cM), was calculated from the recombination frequency corrected by using the Kosambi mapping function (Kosambi, 1944).

Characterization of a MADS-box as candidate gene for male sterility

According to the genetic linkage map recently constructed for *C. intybus* (Cadalen *et al.*, 2010), our SSR marker co-segregating with male sterility was mapped in a linkage group with other SSR and EST markers. Among the latter, one MADS-box gene named *Ci*MPF2-like was selected as candidate. It is worth noting that orthologs of this gene have been demonstrated to be involved in male-sterility in other species, including rice. Specific primers for the *CiMPF2*-like gene were designed according to the partial CDS deposited in

NCBI database (Accession number: AF101420). Replicated cDNAs from flowers of both sterile and fertile plants were used to amplify the gene by PCR and RACE experiments were then performed to obtain the full-lengths. All the amplicons obtained were subcloned and sequenced. Genomic DNA samples of both male-sterile and male-fertile plants were also used to obtain nucleotide sequences of introns for this gene in order to discover informative SNPs.

Results

Male sterility is controlled by a nuclear gene that acts as recessive

Three distinct inheritance models for the genetic basis of male-sterility could be postulated for the male sterile mutations of red chicory: cytoplasmic, related to a mitochondrial gene, and nuclear, which can be associated to either a dominant or a recessive gene. In case of cytoplasmic origin, F1 progenies had to be composed exclusively of male sterile plants (with cytoplasm of *S* type), whereas in case of nuclear origin, two were the expected results: all F1 progeny plants (with a heterozygous genotype *Msms*) had to manifest male-sterility, for a trait controlled by a dominant Mendelian factor (*i.e.*, *Ms*), or male-fertility, for a trait controlled by a recessive Mendelian factor (*i.e.*, *ms*).

All crosses between male-sterile mutants and wild-type pollinators resulted in 100% male-fertile F1 progenies, whereas F2 and BC1 progenies showed to segregate for this trait and to be composed of both male-fertile and male-sterile plants, with proportions equal to 3 : 1 and 1 : 1, respectively. These findings suggested that the male-sterile mutants used as seed parents are homozygous recessive at the locus responsible for male-sterility (*i.e.*, *msms*) and that the wild-type pollinators are homozygous for the dominant allele accounting for male-fertility (*i.e.*, *MsMs*). Segregation ratios observed in the F2 and BC1 progenies developed for each of the four male-sterile mutants along with chi-square values are reported in **Table 2.1**.

Table 2.1. Results from F2 and BC1 progeny tests. Segregation ratios were in agreement with those expected for a trait under control of a single nuclear gene: male-fertile vs. male-sterile plants 3 : 1 in F2 and 1 : 1 in BC1, respectively. Chi-square values were not significant.

			Expected ratios		Observed ratios		
Mutants	**Progeny type**	**Progeny size**	**male-fertile plants**	**male-sterile plants**	**male-fertile plants**	**male-sterile plants**	**Chi-square values**
CS1ms	F2	107	80	27	82	25	0,153
CS2ms	F2	92	69	23	71	21	0,232
IG9ms	F2	100	75	25	78	22	0,480
L11ms	F2	84	63	21	66	18	0,571
Overall	F2	383	287	96	297	86	1,324
CS1ms	BC1	94	47	47	49	45	0,170
CS2ms	BC1	102	51	51	54	48	0,353
IG9ms	BC1	88	44	44	41	47	0,409
L11ms	BC1	96	48	48	43	53	1,042
Overall	BC1	380	190	190	187	193	0,095

Overall data clearly support a nuclear origin and a monogenic control of recessive type for the male-sterility trait in each of the red chicory mutants. Taking together all segregating progeny sets of the F2 and BC1 populations, which included 383 and 380 plants respectively, chi-squares values were non-significant, being as low as 1.324 and 0.095 (Table 1). It is worth mentioning that all flowering plants could be easily scored as male-fertile or male-sterile by a rapid observation of squashed anthers and pollen grains stained with aceto-carmine under a stereomicroscope. No doubtful cases of classification were ever experienced.

Male gametogenesis is arrested at the stage of uninucleate microspores

Regular meiosis was normally found in wild-type plants. After meiosis, each microspore of the tetrads was shown to develop into a binucleated pollen grain through a mitotic division that originated a vegetative and a generative nucleus. Moreover, at anthesis when the pollen grains were mature, they germinated and emitted the pollen tubes (*i.e.*, the microgametophyte), in which the generative nucleus underwent another mitotic division, giving rise to two distinct sperm nuclei.

In the male-sterile mutants, the cytological analysis showed that microsporogenesis proceeds normally up to the development of microspore tetrads. Then the miscropores arrested their development at the uninucleate stage, as documented in **Figure 2.6** using a

parallel with unrelated wild-type plants. In particular, cytological observations revealed that microspores degenerate before their release from the tetrads showing a collapse of the exine. At the end of male meiosis, most of the microspores were found arranged in tetrads while some others were released, becoming shapeless even though the cytoplasm stained well with aceto-carmine. At the beginning of gametogenesis, non-viable shrunken microspores were clearly visible within anthers (details given on **Figure 2.6**).

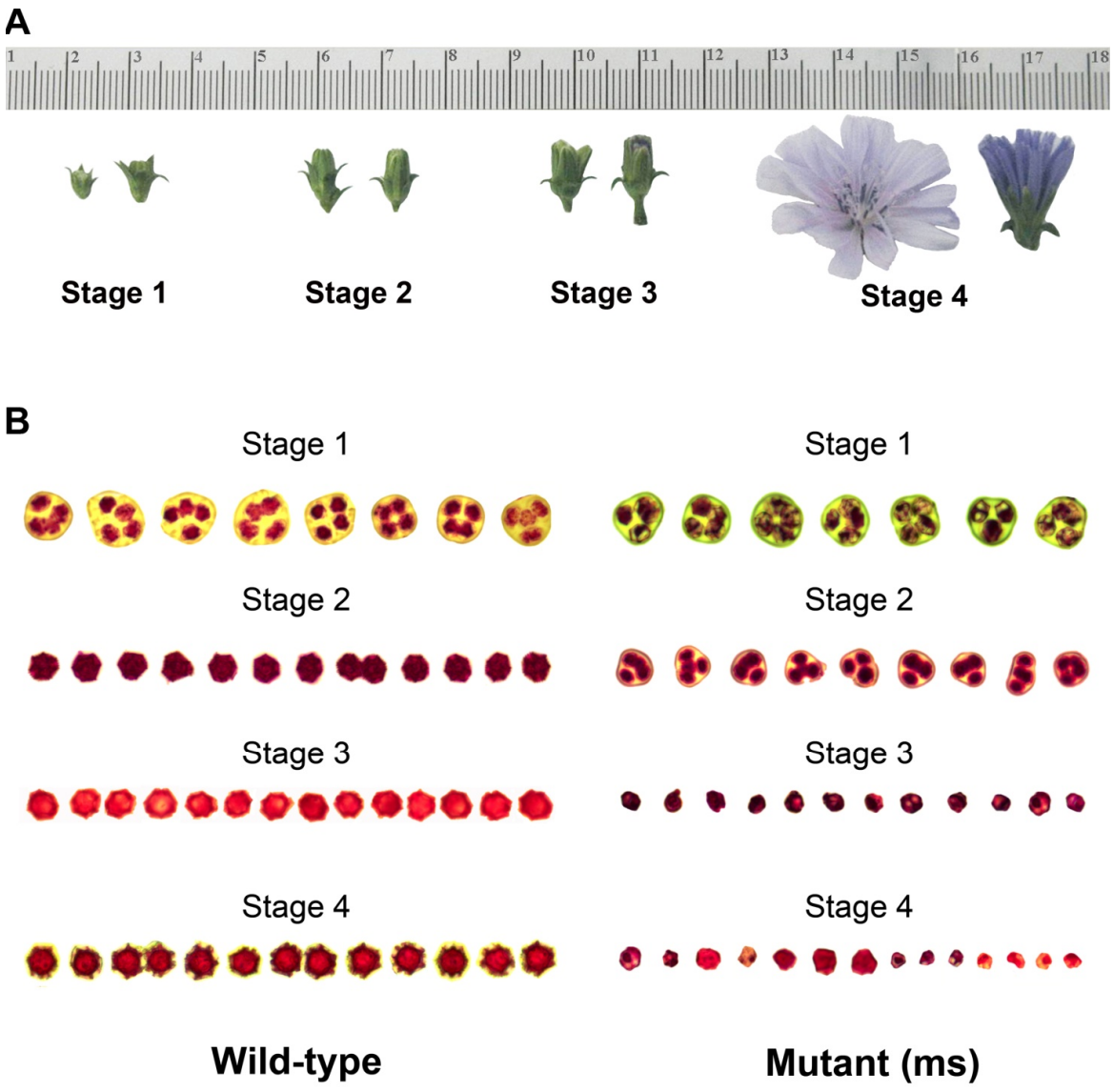


Figure 2.6. Flower developmental stages in red chicory (panel A) and patterns of male gametogenesis in the male-sterile mutants in parallel with wild-type plants (panel B). In the mutants, the microspores of each tetrad arrest their development at the uninucleate stage, degenerating before their release from the tetrads. At full flowering, most of the microspores of dehiscent anthers were found shapeless, shrunken and much smaller than wild-type ones.

It is worth mentioning that microspore tetrads were comparable for their size and shape between mutants and wild-types, whereas mutant microspores at the uninucleate stage proved to be shrunken and much smaller than wild-type ones, as shown on **Figure 2.7**. The most important evidence is that pollen grains were never detected in mature anthers of all four male sterile mutants. This cytological finding was also supported by DAPI staining of squashed anthers (see Figure 1, panels C-D and G-H).

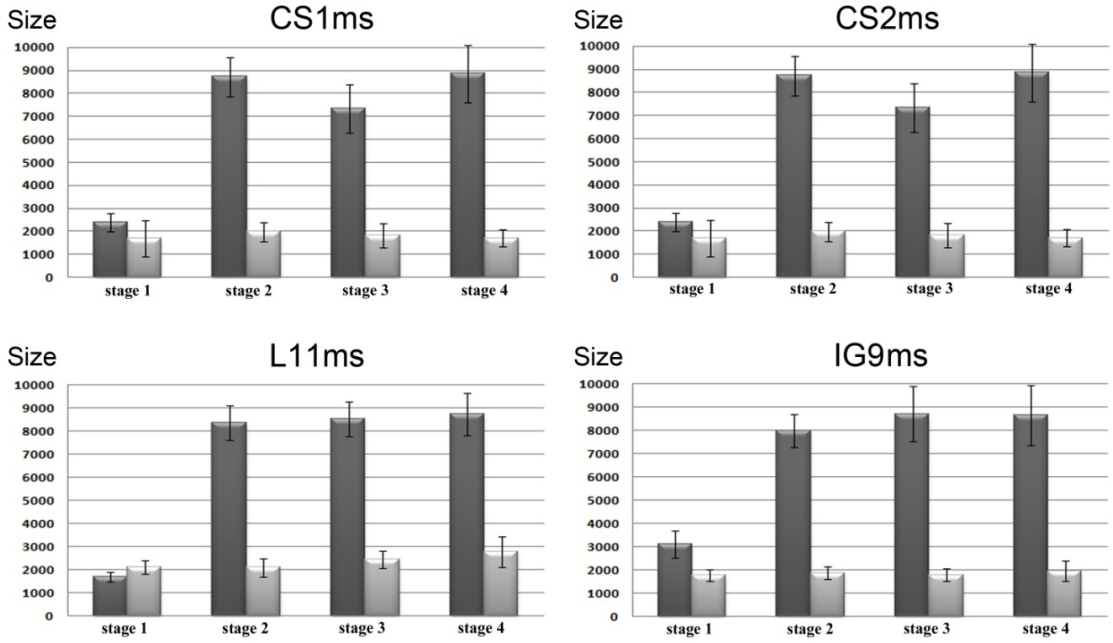


Figure 2.7. Microspore size in the male-sterile mutants compared to wild-type plants, expressed as mean value (histograms) with standard error (bars). At the stage of tetrad, the microspores were comparable for their size and shape between mutants and wild-types, whereas mutant microspores at the uninucleate stage proved to be about three times smaller than wild-type ones. Dark grey: WT plants. Light grey: mutant plants.

Furthermore, the cytological analysis of microsporogenesis and gametogenesis was performed also in the plants belonging to F2 and BC1 progenies. At the cellular level, male meiosis was shown to proceed regularly until the stage of microspore tetrads in both male-sterile mutants and male-fertile plants. Gametogenesis followed a regular pathway in male-fertile plants, giving rise to mature pollen grains, whereas microspores collapsed within

each tetrad in the male-sterile plants, without any further developing process (**Figure 2.8**). In fact, at the end of gametogenesis, a similar phenotype of non-viable shrunken microspores was observed for male-sterile mutants belonging to each of the segregating progenies (**Figure 2.8**). This finding demonstrated that the gene responsible for male-sterility is inherited in the offspring from each mutant by recovering an unaltered maternal genotype, which was always found associated to an unchanged phenotype for male-sterility.

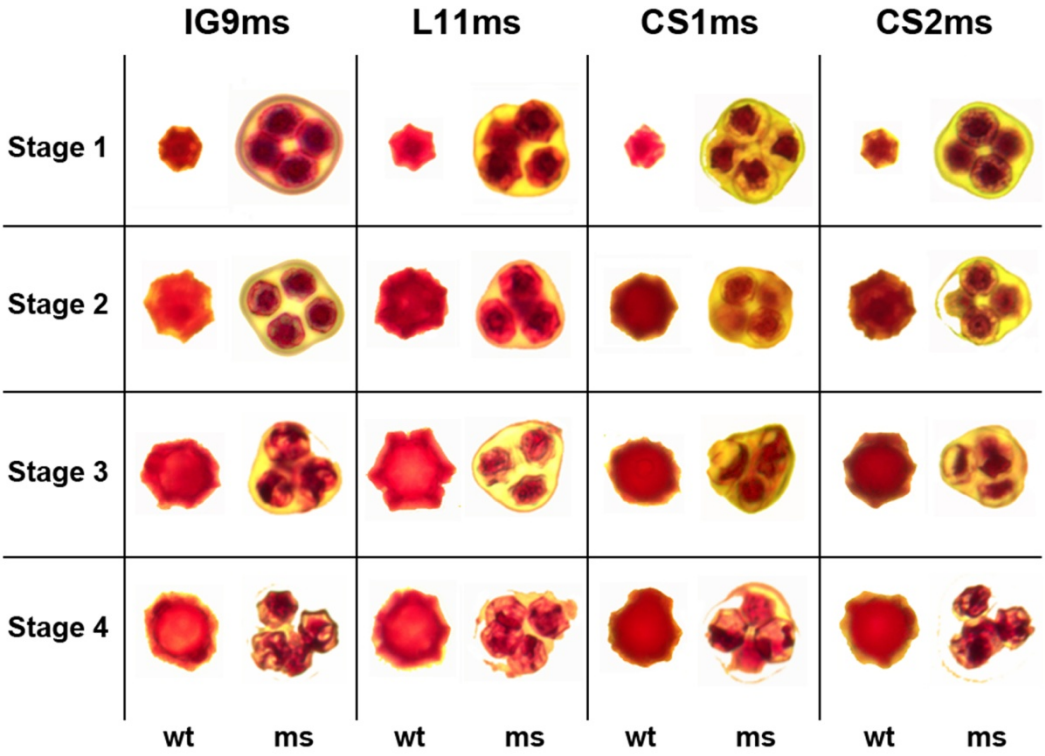


Figure 2.8. Parallel between male gametogenesis in wild-type plants and male-sterile mutants belonging to segregating progenies at flower stages 1-4. Gametogenesis followed a regular pathway in male-fertile plants, giving rise to mature pollen grains, whereas microspores collapsed within each tetrad in the male-sterile plants, without any further developing process. This finding demonstrated that the gene responsible for male-sterility is inherited in the offspring from each mutant by recovering an unaltered maternal genotype, which is always associated to an unchanged phenotype for male-sterility.

The chromosome behavior of male-sterile mutants was also investigated during meiosis: the male meiocyte chromosomes were further analyzed by means of DAPI staining in both wild-type and mutant flowers. Different forms of meiotic abnormalities were found in the male-sterile mutants compared to wild-types, especially at prophase I. In fact, during pachytene, the stage when chiasmata take place and crossing-over occurs between non-sister chromatids of homologous chromosomes, abnormal pairings and chromosomal loops were observed in several sites. Moreover, chromatin bridges were also observed in ana-telophase II. **Figure 2.9** shows some examples of normal chromosome pairing in wild-types (panel A) and miss-pairing of certain chromosome pairs in male-sterile mutants (panels B-D).

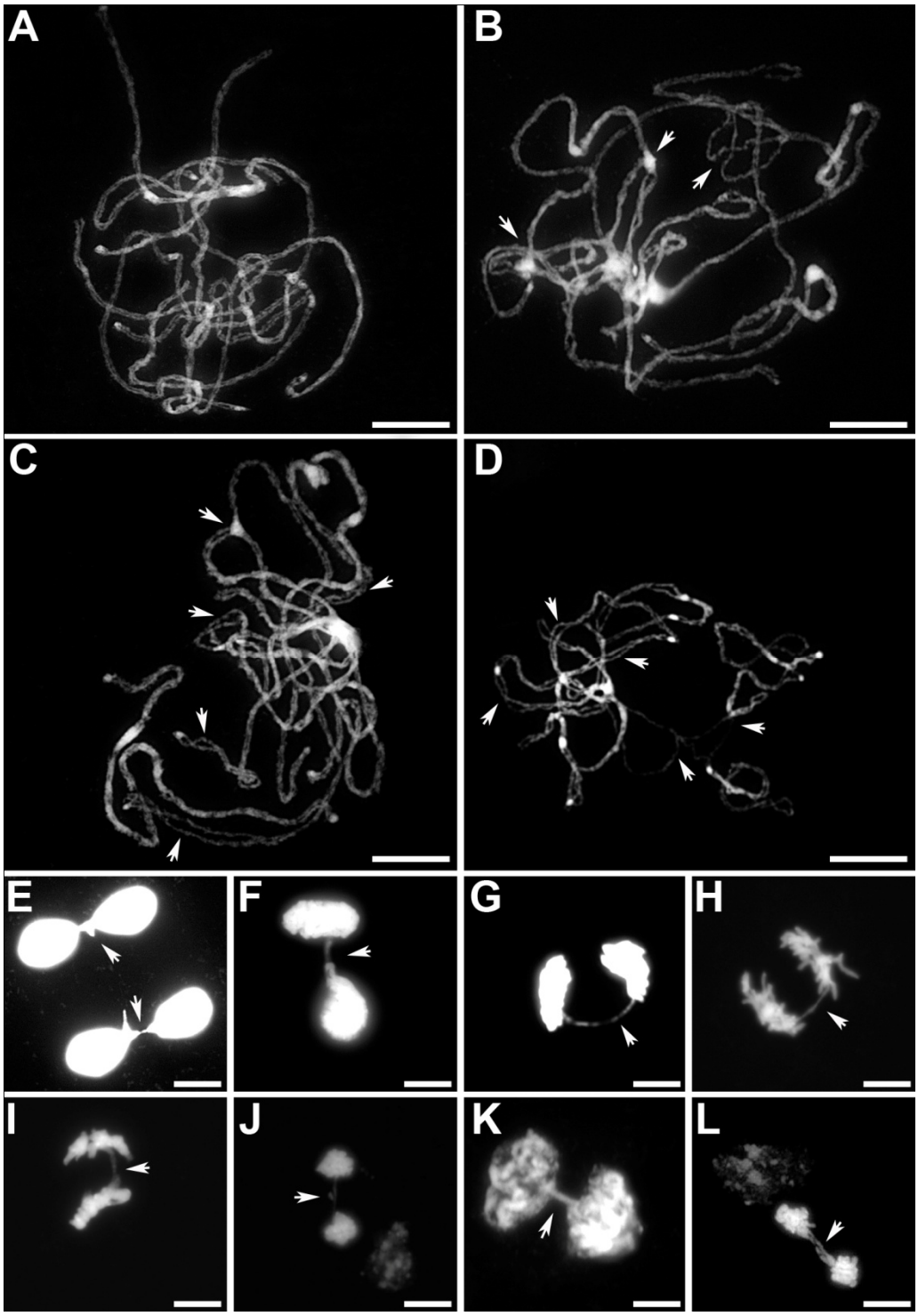


Figure 2.9. Results of cytogenetic analyses of male-sterile mutants: different types of meiotic abnormalities were found in the male-sterile mutants compared to wild-types, especially at prophase I, along with chromatin bridges observed in ana-telophase II. Some examples of normal chromosome pairing in wild-types (panel A) and miss-pairing of certain chromosome pairs in male-sterile mutants (panels B-D). The main aberrant feature in the mutants was recovered at pachytene stage when the homologue chromosomes reached their full pairing: homologues were not completely pairing each other and aberrant structures characterized by one or more loops, due to partial or aspecific pairing between homologous chromosomes, were often observed (see white arrows in panels B-D). Moreover, several cases of chromatin bridges, *i.e.* bridges made of chromatin occurring between newly forming cells, were found in the male-sterile mutants (see white arrows in panels E-L). Bars= 10 µm.

The main aberrant feature was recovered at pachytene stage when the homologous chromosomes reached their full pairing. It was evident that in the mutant, the homologues were not completely pairing each other and aberrant structures characterized by one or more loops, due to partial or aspecific pairing between homologous chromosomes, were often observed (see white arrows in panels B-D of **Figure 2.9**). Moreover, several cases of chromatin bridges, *i.e.* bridges made of chromatin occurring between newly forming cells, were found in the male-sterile mutants (**Figure 2.9**, panels E-L).

Male sterility is genetically linked to a microsatellite marker

In order to map the *ms* locus, a subset of F2 progenies was initially screened to find out molecular marker alleles co-segregating with the male-sterility/fertility trait. Then the selected markers were validated using BC1 progenies on the basis of chi-square values against independent assortment patterns. This strategy allowed us to detect a molecular marker qualitatively polymorphic between DNA bulks of male fertile and male sterile progeny plants and to precisely calculate the genetic distance between the male-sterility trait and the co-segregating marker.

The Mendelian factor responsible for male-sterility was found tightly linked with the molecular marker coded as E02M09/230. When the datasets for both the trait and the marker were analyzed together, there was a significant deviation in the segregation data from the expected 1 : 1 : 1 : 1 ratio. The genetic determinant for male-sterility was found tightly associated with the diagnostic marker, as their alleles were preferentially inherited together (Fisher's 2×2 contingency test: $\chi 2=75.3$ with $P<0.0001$). However, recombination events were apparently possible in the chromosome block carrying the male-sterility gene. In fact, this gene was associated with the marker E02M09/230 in a chromosome window likely characterized by active crossing-over sites and densely saturated by expressed sequence tags. The mean recombination frequency between the male-sterility trait and the

microsatellite marker was equal to 5,8%, corresponding to about 6 cM after correction with the Kosambi's mapping function. This means that the size of the chromosome window covering the *ms* locus may be around 3,000 Kb (assuming 500 Kb/cM).

Amplicons corresponding to the marker E02M09/230 identified and characterized in the chicory progenies were recovered from the agarose gels, subcloned into plasmid vectors and sequenced in order to obtain information on the whole genomic sequence. PCR reactions were performed for both strands using three genomic DNA templates belonging to male sterile and male fertile plants of each segregating population. The sequence recovered from the molecular marker genetically linked with the male sterility trait was the following:

TTGGAGGTGTGAGTGATTCTCGGAGAGTT(TC)nCAGAGATCATTGCTTGTGTAA TTCTCGCTGATTTCAGTTCATTGTCGTCTCTCTTTGCTGTTTCGT. The molecular marker of interest proved to include a microsatellite showing a perfect dinucleotide repetition of the motif (TC/GA)n, with n ranging from 27 to 33. As a consequence, a novel SSR assay for the detection of this marker, which includes the basic dinucleotide repeat TC/GA and whose size ranges around 160-170 bp, was implemented by the design of a specific and stringent forward primer (5'-CTTGGAGGTGTGAGTGATTCT-3') and reverse primer (5'-TACGAAACAGCAAAGAGAGAC-3').

Figure 2.10 shows the SSR genotypes detected in the progeny plants of segregating populations: male-sterile plants were homozygous *AA* or *BB*, with marker alleles of 160 and 162 bp, respectively, whereas male fertile-plants could be either heterozygous *AD* or *BC* and homozygous *DD* or *CC*, with marker alleles of 170 and 168 bp, respectively. We concluded that the SSR assay developed in this study can be profitably adopted as a tool of marker-assisted breeding and exploited for an early screening of the T&T® male-sterile plants within segregating progenies stemmed from back-crosses with a genotyping error around 2.9%.

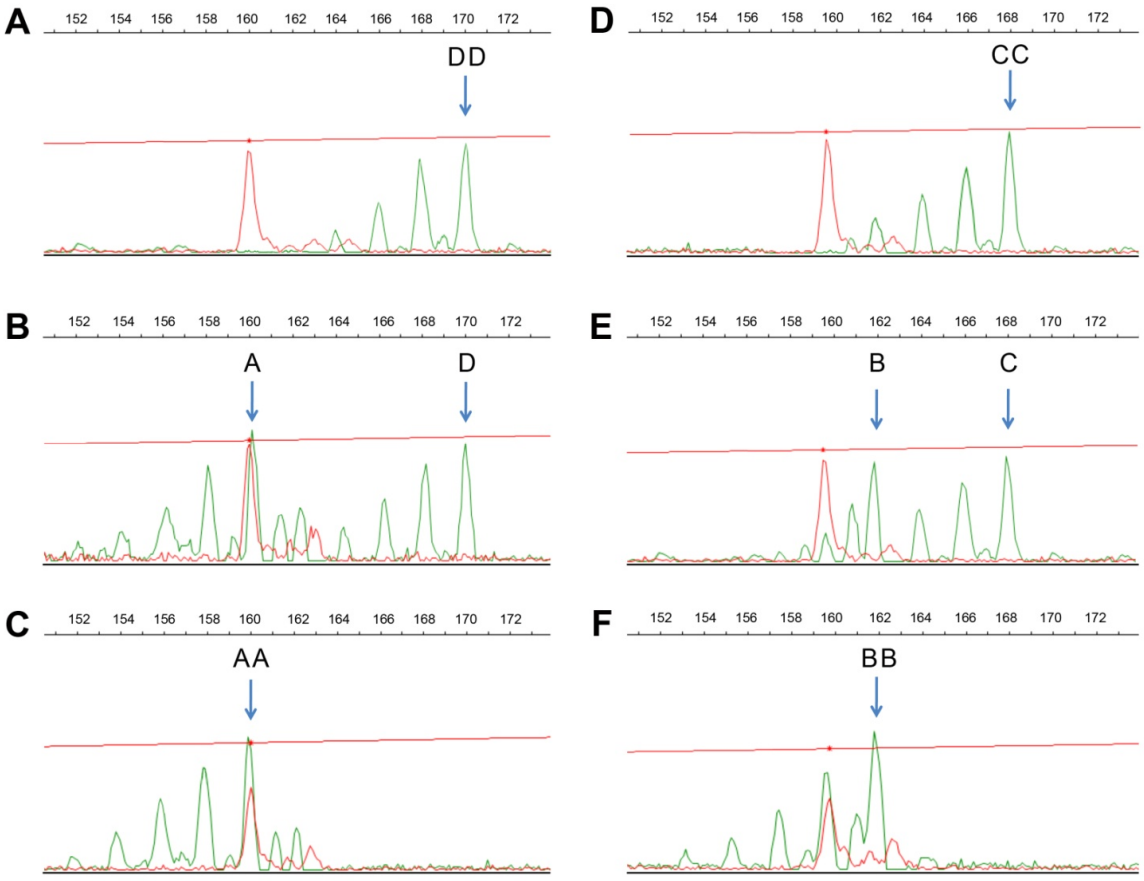


Figure 2.10. Molecular SSR marker diagnostic assay for discriminating male-sterile from male-fertile plant genotypes. The male-sterile plants are homozygous AA or BB, with marker alleles of 160 and 162 bp, respectively, whereas male fertile-plants can be either heterozygous AD or BC and homozygous DD or CC, with marker alleles of 170 and 168 bp, respectively. This SSR assay is useful as a tool of marker-assisted breeding for an early screening of the male-sterile plants within segregating progenies, with a genotyping error around 2.9%. The nuclear DNA sequence of the marker associated to male-sterility in red chicory was deposited as CiSSR_TC/GA31 locus in the NCBI databases under the GenBank accession number JF748831.

The nuclear DNA sequence of the marker associated to male-sterility in red chicory was deposited as CiSSR_TC/GA31 locus in the NCBI databases under the GenBank accession number JF748831 (BankIt1443036, see also www.ncbi.nlm.nih.gov/Genbank).

Characterization of a MADS-box gene mapped in the linkage group carrying the *ms* gene

On the basis of the genetic map available for chicory (Cadalen *et al.*, 2010), the CiSSR_TC/GA_31 marker co-segregating with male-sterility was located in a linkage group next to a sequence tagged-site marker deriving from a MADS-box gene, which was termed *Ci*MPF2-like. A partial CDS sequence of the gene was deposited in NCBI database (Accession number AF101420). PCR amplifications by specific primers and RACE experiments enabled to obtain full-lengths of the gene from both male sterile and male fertile genotypes. Interestingly, two different forms were isolated differing each other at the 3' region: one form was shorter than the other (**Figure 11**). Furthermore, the long sequence form, called *Ci*MPF2-like_long of the MADS-box gene was presumably duplicated, because at least three alleles/members were found in the genome. Thus, the classification as STS marker (Cadalen *et al.*, 2010) seems not verified due to the detection of three copies (two replicated long forms, *Ci*MPF2-like_long, and one short form, *Ci*MPF2-like_short).

The short form of the gene, *i.e*. *Ci*MPF2-like_short, was amplified from genomic DNAs of both male sterile and male fertile plants, but it was discarded as candidate gene for male sterility because it was not expressed in floral tissues (data not shown). Vice versa, the long form of the gene, *i.e*. *Ci*MPF2-like_long, proved to be strongly expressed in floral tissues and it was selected as candidate gene for further experiments. This latter *Ci*MPF2-like gene was amplified and sequenced from either male fertile and male sterile plants belonging to F2 and BC1 progenies, but their CDS nucleotide sequences did not show any polymorphism. It is worth mentioning that sequencing of the *Ci*MPF2-like_long genomic member was a difficult task because of the high number and large size of introns (**Figure 2.11**). Furthermore, the lack of similar gene sequences belonging to the same taxonomic entity or botanic family of chicory in the nucleotide NCBI databases, did not allow to reliably predicting the gene structure in terms of position and succession of exons/introns.

Nevertheless, the site of three introns of length equal to 426 bp, 284 bp and 88 bp and the presence of five additional introns were predicted. Above all, the detailed analysis of the sequence of the three introns did not reveal any SNP between mutants and wild-type plants. The long form of the MADS-box gene strongly expressed in flowers proved to share the whole coding sequence in plants characterized by either male-fertility or male-sterility.

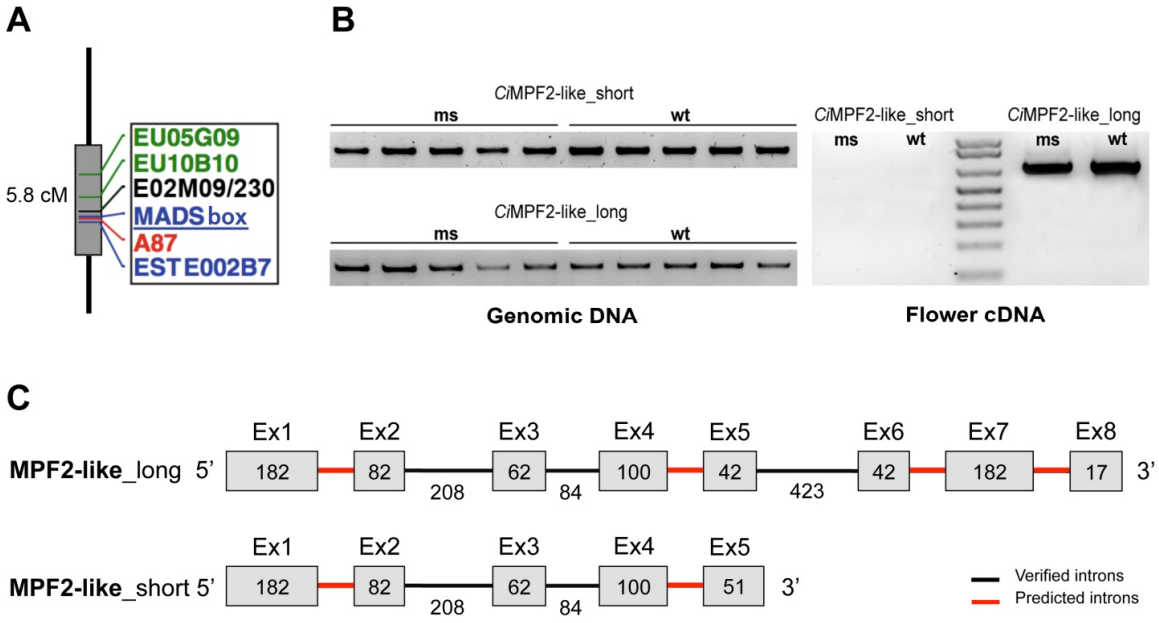


Figure 2.11. Schematic representation of the structure of the *Ci*MPF2-like genes. *Ci*MPF2-like_long is characterized by 8 exons and 7 predicted introns, whereas *Ci*MPF2-like_short is characterized by 5 exons and 4 predicted introns. Numbers inside the square indicates the length of exons/introns in base pairs (bp). The light and dark grey indicates the exons and introns, respectively. PI refers to the predicted introns of which the nucleotide sequence is still unknown.

The longer introns belonging to MPF2-long were cloned in specific vectors and their sequencing is in progress. Current experiments aim at discovering SNPs between gene copies isolated from male sterile and male fertile plants to be exploited to test the co-segregation of this genetic factor with the sterility trait. A genetic analysis like this presented will allow us to conclusively accept or discard the MADS-box gene as good candidate for male sterility in chicory.

Discussion

Discovery of genetic male sterility in chicory

To the best of our knowledge, this is the first time that naturally occurring male-sterile mutants are discovered and genetically characterized in *C. intybus*. The objectives of this study were to describe the genetic basis (inheritance patterns and linkage with molecular markers), and the cytological and karyological features (male meiosis and gametogenesis) of the first male-sterile mutants ever discovered in the genus *Cichorium*.

We have demonstrated that male-sterility in our red chicory "radicchio" mutants is controlled by a single nuclear gene (*ms*) that acts at the recessive status. In fact, all crosses between male-sterile mutants and wild-type pollinators resulted in 100% male-fertile F1 progenies (*Msms*), whereas F2 and BC1 progenies segregated for this trait being composed of both male-fertile and male-sterile plants, with ratios equal to 3 (25% *MsMs* and 50% *Msms*): 1 (25% *msms*) and 1 (50% *Msms*) : 1 (50% *msms*), respectively.

We have also documented that the male gametogenesis is arrested at the stage of microspore tetrads. In all male sterile mutants, the cytological analysis showed that microsporogenesis seems to proceed regularly up to the development of tetrads, the microspores then arrest their developmental program. At the beginning of microgametogenesis, non-viable shrunken microspores were clearly visible within anthers. Interestingly, meiotic abnormalities were found in the male-sterile mutants, especially at prophase I. In fact, abnormal pairings and chromosomal loops were observed during pachytene. It is well known that the central function of synapsis is the recognition of homologues by pairing, an essential step for a successful meiosis. Irregular synapsis for some of the homologous chromosomes may alter the further development of microspores. This feature would stop the process of male gametogenesis.

Genetic factors affecting meiotic chromosome pairing in plants are of special interest to geneticists and especially breeders. The nuclear male-sterile mutations can affect microsporogenesis or microgametogenesis, hampering the formation of pollen grains. A crucial step for male fertility is the conjugation of chromosomes during the first meiotic prophase, a phenomenon termed synapsis. When synapsis occurs irregularly in the male meiocytes, meiosis may lead to non-functional microspores. Several mutants characterized by the lack of chromosome pairing during the first meiotic prophase (*i.e.*, asynapsis) have been found in plant species, as well as mutants in which chromosomes initially pair in early meiotic prophase but fail to remain paired at later meiotic stages (*i.e.*, desynapsis) (Chaudhury, 1993; Horner and Palmer, 1995).

Our findings suggest that the exchange of DNA segments over regions of homology is strongly prevented in the male-sterile mutants and that the lack of regular synapsis for some of the homologous chromosomes may alter the further development of microspores. In addition, the occurrence of chromatin bridges between newly forming cells is usually an indicator of abnormalities related to cellular division. All together these features provide karyological evidences that support chromosome features and factors negatively influencing the process of male gametogenesis, resulting in a phenotype that can be described as “anthers with no pollen grains”.

It is worth noting that the sterility of gametes occurs only in male organs. The quantity of seeds set by the mutant flowers was not significantly different from that of wild-type plants (data not shown), demonstrating that the female organs of mutant flowers are completely fertile. As a consequence, our observations suggest that the mutant phenotype is attributable to a gene expressed in an anther-specific manner.

We successfully implemented a PCR-based assay that can be profitably adopted for an early screening of the male-sterile plants within segregating progenies, with a genotyping error lower than 3%. The gene responsible for male-sterility was found genetically linked to

a molecular marker (*i.e.*, E02M09/230), about 6 cM apart from the *ms* locus. The molecular marker linked to male-sterility was sequenced and its analysis allowed us to disclose a perfect dinucleotide microsatellite of the repetitive motif (TC/GA)n, with n ranging from 27 to 33. As a consequence, an SSR assay for the detection of this marker, whose size ranges around 160-170 bp, was implemented by the design of a specific pair of primers. The male sterile plants were homozygous for the smaller marker alleles, whereas male fertile plants could be either heterozygous or homozygous for marker alleles of larger size. Our PCR-based assay will find application not only for the marker-assisted selection of male-sterile seed parents but also for the genetic identification and legal protection of these valuable mutant genotypes of red chicory.

According to the genetic map reported in (Cadalen *et al.*, 2010), the selected SSR marker located at about 6 cM apart from the *ms* locus allowed us to identified, within the same linkage group, a MADS-box gene, *Ci*MPF2-like, which is involved in male-sterility in other species. However, CDS and introns nucleotide sequences from male sterile and male fertile plants did not show difference and this is in contrast with a putative candidate for the sterility trait. Furthermore, assuming that *Ci*MPF2-like gene had to be the responsible for the sterility, our genetic distance of about 6 cM revealed between the selected SSR marker and *Ci*MPF2-like gene was in contrast with that measured by (Cadalen *et al.*, 2010) which consist of about 0,6 cM. These two evidences do not encourage the hypothesis of *Ci*MPF2-like as good candidate of male-sterility in our mutants of red chicory.

Since the history dawn of plant breeding after the rediscovery of Mendel's laws, the exploitation of heterosis is an effective approach to increase crop yields. F1 hybrid varieties in major crops such as cereals and vegetables can show more than a 100% yield advantage over the best conventional ones under the same cultivation conditions (Melchinger and Gumber, 1998; Lamkey and Edwards, 1999; see also review by Barcaccia *et al.*, 2006). Difficulties in breeding elite male sterile lines and producing commercial hybrid seeds

hamper the development of F1 hybrid varieties. An important role in chicory breeding could be played by male-sterility in hybrid seed production: this is particularly true in "Radicchio" since self-incompatibility of parental lines was found inadequate for reliable production of F1 hybrids.

Genetically engineered male sterility in chicory

In chicory, the presence of a naturally occurring CMS system has not been reported, whereas strategies to genetically engineering male sterility were used in Magdeburg, Witloof and Chioggia genotypes (reviewed by Lucchin *et al.*, 2008).

In a first approach, transgenic male sterile lines of chicory were produced by expressing the ribonuclease gene RNase from *Bacillus amyloliquefaciens* ("BARNASE") under the control of a tapetum-specific promoter originally isolated from tobacco (TA-29), according to (Mariani *et al.*, 1990). Restorer lines for these male-sterile lines were obtained by expressing the gene coding for "BARSTAR", the intracellular inhibitor of "BARNASE" under control of the same promoter (Denis *et al.*, 1993; REYNAERTS *et al.*, 1993). The development of inbred lines and male-sterile lines provided a reliable system for genetically engineering pollination in plants and allowed a new hybrid seed production system, which has been registered as SeedLink^{TM}.

Somatic hybridization by means of protoplast symmetric fusion between chicory and the CMS line of sunflower "PET-1" was also attempted in order to promote the regeneration of interspecific hybrid plants. This kind of CMS in sunflower was identified in an interspecific cross between *Helianthus petiolaris* and *Helianthus annuus*, and it was associated with the expression of the mitochondrial gene ORF522, encoding a 15-kD polypeptide. The ORF522 gene was originated by a recombination event at the 3' of *atp1* gene and its protein is detectable in flowers of CMS but not of restored lines (Horn *et al.*, 1991; Moneger *et al.*, 1994). The hybrid plants obtained after somatic symmetric fusion were cytoplasmic

hybrids, cybrids, and showed mtDNA rearrangements, indicating that symmetric fusion had the tendency to maintain the chicory mitochondrial genome. Three different kinds of sterility were observed: i) plant with anthers lacking dehiscence without, or with non-viable, pollen; ii) complete absence of the anthers; and iii) absence of both anthers and styles or the presence of reduced styles. One of these male-sterile plants was used for the production of F1 hybrids whose yields were equal to or higher than those of traditional varieties. In a subsequent work, three different CMS chicory cybrids were backcrossed to Witloof chicory in order to transfer the male sterile cytoplasm from an industrial chicory to a Witloof genetic background. The transcript analysis revealed that the ORF522 is weakly expressed or not expressed at all in the cybrids. This finding led (Dubreucq *et al.*, 1999) to conclude that ORF522 cannot be associated to the CMS observed in the chicory cybrids and to suggest that they presented a novel form of CMS, different from that of sunflower. Protoplast asymmetric fusion was used to produce male sterile somatic hybrids between a "Rosso di Chioggia" genotype and a "PET-1" sunflower CMS line. At anthesis the regenerated cybrids had fewer and non-viable pollen grains but they could set seeds when free-pollination occurred (Varotto *et al.*, 2001). Overall results collected so far using interspecific protoplast fusion experiments suggest that male-sterile cybrid plants can be actually regenerated in chicory. Nevertheless, it appears that mitochondrial genome re-arrangements lead to the creation of novel CMS chicory types instead of transferring the desired trait from CMS sunflower lines. The methods of transgenesis useful for making cytoplasmic male sterile chicory plants comprising the ORF522 of *Helianthus annuus* was patented in 2004 (see United States Patent documents and files available at www.patentgenius.com/patent/6803497.html). As a matter of fact, the development of inbred lines and male-sterile lines based on this biotechnological approach failed to provide any reliable hybrid seed production system in chicory.

Breeding F1 hybrid varieties of chicory

Currently, the main types of radicchio cultivated in Veneto are: “Rosso di Chioggia”, early “Rosso di Treviso”, late “Rosso di Treviso”, “Variegato di Castelfranco” and “Rosso di Verona”. The first of these is the most widespread, while the others represent locally valuable high-quality crops. Although a clearcut morphological differentiation among the five types does exist, their genetic identification is becoming increasingly important. Materials grown are usually represented by local populations known to possess a high variation and adaptation to the natural and anthropological environment where they have originated and are still cultivated. These populations are maintained by farmers through phenotypical selection according to their own criteria and sometimes by exploiting to controlled hybridizations among different types in order to obtain recombinant genotypes showing superior agronomic and commercial traits.

The high value of the fixation index (G_{ST}=0.315) found by (Barcaccia *et al.*, 2003) is consistent with a DNA polymorphism rate more pronounced among entries within type than between types of radicchio. This result suggested that, in each radicchio type, populations produced by breeders through controlled intercrossing or repeated polycrossing conserved their well-separated gene pools over the years, as confirmed by the rather low estimates of gene flow. On the basis of the reproductive system of radicchio, such a finding may be explained taking into account two factors: first, in this species, inbreeding is hampered by a sporophytic incompatibility mechanism that limits both selfing and intercrossing between plants with an identical phenotype at the multi-allelic S-locus, thus allowing a certain amount of heterozygosity to be maintained even in inbred populations; second, the selection criteria of mother plants to be used for seed production applied each year by each farmer probably allowed contamination between types to be limited, preserving the phenotypic identity of each type. Moreover, seed multiplication carried out

over the years by each farmer according to their own criteria produced a clear genetic differentiation between types within the species (Barcaccia *et al.*, 2003b).

The breeding programs at present under way by local breeders and regional seed institutions aim (1) to isolate, within the best local selections, individuals amenable for use as parents for the constitution of synthetic varieties and, although not easily feasible (2) to select inbred lines suitable for the production of commercial F1 hybrids. These breeding procedures could be greatly helped by the use of molecular markers that allow the discarding of molecular off-types, to better exploit parental genetic polymorphisms for synthetics and to identify the most genetically distant inbreds as parental lines for hybrids (Barcaccia *et al.*, 2003).

A few genetic studies using molecular markers have been carried out on *Cichorium* spp. mainly to characterize commercial varieties and experimental materials (Bellamy *et al.*, 1996; Barcaccia *et al.*, 2003; Van Stallen *et al.*, 2003) to evaluate the genetic homogeneity and purity, respectively, of inbreds and hybrids (Bellamy *et al.*, 1996), and to investigate phylogenetic relationships between cultivars and cultivar groups of *C. intybus* and other species, both cultivated and wild, belonging to the same genus (Kiers *et al.*, 2000; Barcaccia *et al.*, 2003; Van Stallen *et al.*, 2003). Amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) markers were also used to construct the first genetic map of *C. intybus* (De Simone *et al.*, 1997). More recently, a new genetic map was constructed for chicory using simple sequence repeat (SSR or microsatellite) markers by (Cadalen *et al.*, 2010). This consensus genetic map, which includes 9 homologous linkage groups, was obtained after the integration and ordination of molecular marker data of one witloof chicory and two industrial chicory progenies.

SSR markers are considered as ideal co-dominant markers for genome fingerprinting and gene mapping in crop plants because of their neutral status, abundant sites, high

polymorphism information contents and mostly monogenic inheritance patterns (Morgante and Olivieri, 1993).

Traditionally, varieties of chicory have been developed by mass selection in order to obtain uniform populations characterized by valuable production and acceptable commercial head size and shape. Newly released varieties are mainly synthetics produced by intercrossing a number of phenotypically superior plants, selected on the basis of morpho-phenological and commercial traits. More rarely, plants are also evaluated genotypically by means of progeny tests. Synthetics have a rather large genetic base and are represented by a heterogeneous mixture of highly heterozygous genotypes sharing a common gene pool. In recent years, methods for the constitution of F1 hybrids have been developed by private breeders and seed companies. Details on the procedure for the constitution of such hybrids are not available in the current literature and it may be presumed that each company has developed its own protocol, mainly in accordance to the genetic material it has at disposal and to the possibility of applying a more or less efficient control on the F1 hybrid seed production phase. As a matter of fact, the strong self-incompatibility system, which hinders obtaining highly homozygous parents, and the absence of a male sterility factor within the species or in sexually compatible species, made it difficult if not impossible to propose an F1 seed production scheme and, most of all, to consider these newly commercial varieties as true F1 hybrids.

It is worth emphasizing that molecular markers in *C. intybus* have been exploited for selecting the mother plants of synthetics as well as for determining the distinctiveness, uniformity and stability, *i.e*. DUS testing, of newly bred varieties (Barcaccia and Parrini, 2001). In *C. intybus*, molecular marker should also find utility for assessing the genetic homogeneity and homozygosity of inbred lines produced by repeated selfing, measuring the genetic diversity among inbred lines in order to plan crosses and maximize heterosis in the

experimental F1 hybrids, and evaluating the genetic purity and heterozygosity of seed stocks of commercial F1 hybrids.

In conclusion, the discovery of male-sterility in this species will open new frontiers for breeding new varieties of radicchio, in particular, and of chicory, in general, especially if this trait can be successfully transferred to elite lines and precociously identified by molecular diagnostic assays suitable to perform marker-assisted selection programs.

References

Ambrosi D.G., Galla G., Purelli M., Barbi T., Fabbri A., Lucretti S., Sharbel T.F. and Barcaccia G. (2010). DNA markers and FCSS analyses shed light on the genetic diversity and reproductive strategy of *Jatropha curcas* L. Diversity, 2: 810-836.

Barcaccia G., Albertini E., Rosellini D., Tavoletti S., Veronesi F. (2000). Inheritance and mapping of 2n-egg production in diploid alfalfa. Genome 43, 528-537.

Barcaccia G., Parrini P. (2001). Radicchio breeding and seed production (in Italian). In: Radicchio rosso di Verona. Edited by Veneto Agricoltura (Legnaro, Padova, Italy). pp. 45-56.

Barcaccia G., Pallottini L., Soattin M., Lazzarin R., Parrini P., Lucchin M. (2003). Genomic DNA fingerprints as a tool for identifying cultivated types of radicchio (Cichorium intybus L.) from Veneto, Italy. Plant breeding = Zeitschrift für Pflanzenzüchtung. 122, 178-183.

Barcaccia G., Varotto S., Soattin M., Lucchin M. Parrini P. (2003). Genetic and molecular studies of sporophytic self-incompatibility in *Cichorium intybus* L. Proc. of the Eucarpia meeting on Leafy Vegetables Genetics and Breeding, March 19-21, 2003, Noordwijkerhout, The Netherland, p. 154.

Barcaccia G., Lorenzetti S., Falcinelli M. (2006). Heterosis in plants: from the hypothesis of Jones to the genomics era (in Italian). Dal Seme, 11: 22-48.

Bellamy A., Vedel F., Bannerot H. (1996). Varietal identification in Cichorium intypbus L. and determination of genetic purity of F1 hybrid seed samples, based on RAPD markers. Plant breeding = Zeitschrift für Pflanzenzüchtung. 115, 128-132.

Bianchedi A. 1961. I radicchi di Treviso. L'Italia Agricola. 1: 37-51.

Bremer K. (1994) Asteraceae: cladistics and classification. Timber Press, Portland, USA.

Budar F., Touzet P., De Paepe R. (2003). The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. Genetica 117, 3-16.

Cadalen T.,M., Blassiau C., Clabaut A., Scheer I., Hilbert J., Hendriks T., Quillet M. (2010). Development of SSR markers and construction of a consensus genetic map for chicory (Cichorium intybus L.). Mol. Breed. 25, 699-722.

Chase C.D. (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet. 23, 81.

Chaudhury A.M. (1993). Nuclear genes controlling male fertility. Plant Cell 5, 1277-1283.

Chen H., Ge X., Du X., Zhao Z., Li Z. (2009). Genetic and histological characterization of a novel recessive genic male sterile line of Brassica napus derived from a cross with Capsella bursa-pastoris. Euphytica 167, 31-37.

Chu M.G., Li S., Wang S., Zheng A., Deng Q., Ding L., Zhang J., Zhang M., He M., Liu H., Zhu J., Wang L., Li P. (2011). Fine mapping of a male sterility gene, vr1, on chromosome 4 in rice. Mol. Breed. 28, 181-187.

De Simone M., Morgante M., Lucchin M., Parrini P., Marocco A. (1997). A first linkage map of Cichorium intybus L. using a one-way pseudo-testcross and PCR-derived markers. Mol. Breed. 3, 415-425.

Desprez B., Delesalle L., Dhellemmes C. Desprez M. (1994). Genetics and breeding of industrial chicory. C R Acad. Agric. Fr., 80: 47-62.

Desprez F.F., Bannerot H. (1980). A study of pollen tube growth in witloof chicory. In: Maxon Smith JW, Langton A (eds.) Proc. of the Eucarpia meeting on leafy vegetables, Littlehampton. pp. 47-52.

Denis M., Delourme R., Gourret J.P., Mariani C., Renard M. (1993). Expression of engineered nuclear male sterility in Brassica napus. Genetics, morphology, cytology, and sensitivity to temperature. Plant physiology. 101, 1295-1304.

Dubreucq A., Berthe B., Asset J.F., Boulidard L., Budar F., Vasseur J., Rambaud C. (1999). Analyses of mitochondrial DNA structure and expression in three cytoplasmic male-sterile chicories originating from somatic hybridisation between fertile chicory and CMS sunflower protoplasts. Theoretical and applied genetics. 99, 1094-1105.

Ducos E., Touzet P., Boutry M. (2001). The male sterile G cytoplasm of wild beet displays modified mitochondrial respiratory complexes.Plant Journal 26, 171-180.

Eenink A.H. (1981). Compatibility and incompatibility in witloof-chicory (Cichorium intybus L.). 2. The incompatibility system. Euphytica 30, 77-85.

Eenink A.H. (1982). Compatibility and incompatibility in witloof-chicory (Cichorium intybus L.). 3. Gametic competition after mixed pollinations and double pollinations. Euphytica 31, 773-786.

Gabay-Laughnan S., Kuzmin E.V., Monroe J., Roark L., Newton K.J. (2009). Characterization of a novel thermosensitive restorer of fertility for cytoplasmic male sterility in maize. Genetics 182, 91-103.

Goldberg R.B., Beals T.P., Sanders P.M. (1993). Anther development: basic principles and practical applications. Plant Cell 5, 1217-1229.

Hayden M.J., Nguyen T.M., Waterman A., McMichael G.L., Chalmers K.J. (2008). Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. Molecular breeding. 21, 271-281.

Horn R., Kohler R.H., Zetsche K. (1991). A mitochondrial 16 kDa protein is associated with cytoplasmic male sterility in sunflower. Plant molecular biology : an international journal on fundamental research and genetic engineering. Plant Mol. Biol. 17, 29-36.

Horner H.T., Palmer R.G. (1995). Mechanisms of genic male sterility. Crop science. 35, 1527-1535.

Huang Z., Chen Y., Yi B., Xiao L., Ma C., Tu J., Fu T. (2007). Fine mapping of the recessive genic male sterility gene (Bnms3) in Brassica napus L. Theor. Appl. Genet. 115, 113-118.

Johns C.W., Delanny X., Palmer R. (1981). Structural sterility controlled by nuclear mutations in angiosperms. The Nucleus 24:97–105).

Kaul M.L.H. (1988). Male sterility in higher plants. Monographs on theoretical and applied genetics 10. Springer, New York, p 10).

Kiers A.M., Mes T.H.M., Meijden R.v.d., Bachmann K. (2000). A search for diagnostic AFLP markers in Cichorium species with emphasis on endive and chicory cultivar groups. Genome 43, 470-476.

Kosambi D.D. (1944). The estimation of map distances from recombination values. Annals Eugenics, 12: 172-175.

Lamkey, K.R., Edwards, J.W. (1999). Quantitative genetics of heterosis. In: The genetics and exploitation of heterosis in crops. (J.G. Coors and S. Pandey, ed.). American Society of Agronomy, Crop Science Society of America, Madison, WI, USA. pp. 31-48.

Li, S. L., Qian Y. X., Wu Z. H., (1985). Genetic study on genic male sterility and its utilization in Brassica napus L. Acta Agriculturae Shanghai 1, 1—12.)

Lu G.Y., Yang G.S., Fu T.D. (2004). Molecular mapping of a dominant genic male sterility gene Ms in rapeseed (Brassica napus). Plant breeding = Zeitschrift für Pflanzenzuchtung. 123, 262-265.

Lucchin M., Barcaccia G., Lazzarin R. Parrini P. (2003). Is phenotypic selection suitable to obtain molecularly differentiated populations of radicchio (*Cichorium intybus* L.)? Atti del Congresso della Società Orticola Italiana, Gruppo di Lavoro Miglioramento Genetico: Lo stato dell'arte nel miglioramento genetico delle principali specie ortoflorofrutticole d'interesse mediterraneo (in Italiano). 25-26 Giugno 2002, Valenzano (Bari), Italy. A cura di L. Ricciardi. pp. 387-396.

Lucchin M., Varotto S., Barcaccia G., Parrini P. (2008). Chicory and Endive. In: Handbook of Plant Breeding, Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae. Edited by Jaime Prohens-Tomás and Fernando Nuez. Springer Science, New York, USA. pp. 1-46.

Lurin C., Andres C., Aubourg S., Bellaoui M., Bitton F., Bruyere C., Caboche M., Debast C., Gualberto J., Hoffmann B. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16, 2089-2103.

Ma H. (2005). Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. Annu Rev Plant Biol 56, 393.

Mariani C., Beuckeleer M.d., Truettner J., Leemans J., Goldberg R.B. (1990). Induction of male sterility in plants by a chimaeric ribonuclease gene. Nature 347, 737-741.

Melchinger, A.E., Gumber, R.K. (1998). Overview of heterosis and heterotic groups in agronomic crops. In: Concepts and breeding of heterosis in crop plants (K.R. Lampkey and J.E. Staub, ed.). Crop Science Society of America, CSSA, Madison, WI, USA. Special Publ. n. 25. pp. 99-118.

Michelmore R.W., Paran I., Kesseli R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences of the United States of America. 88, 9828-9832.

Moneger F., Smart C.J., Leaver C.J. (1994). Nuclear restoration of cytoplasmic male sterility in sunflower is associated with the tissue-specific regulation of a novel mitochondrial gene. EMBO journal. 13, 8-17.

Morgante M., Olivieri A.M. (1993). PCR-amplified microsatellites as markers in plant genetics. Plant J. 3, 175.

Panero J.L., Funk V.A. (2002). Toward a phylogenetic subfamilial classification for the Compositae. Proc. Biol. Soc. Wash., 115: 909-922.

Pécaut P. (1962). Etude sur le système de reproduction de l'endive (*Cichorium intybus* L.). Ann Amélior Plantes, 12: 265-296.

Reynaerts A., Vandewiele H., Desutter G., Janssens J. (1993). Engineered genes for fertility-control and their application in hybrid seed production. Theoretical and applied genetics55, 125-139.

Sabar M., Paepe R.d., Kouchkovsky Y.d. (2000). Complex I impairment, respiratory compensations, and photosynthetic decrease in nuclear and mitochondrial male sterile mutants of Nicotiana sylvestris. Plant physiology. 124, 1239-1249.

Sawhney V.P., Shukla A. (1994). Male sterility in flowering plants: are plant growth substances involved? American journal of botany. 81, 1640-1647.

Schnable P.S., Wise R. (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci. 3, 175-180.

Shi Y., Zhao S., Yao J. (2009). Premature Tapetum Degeneration: a Major Cause of Abortive Pollen Development in Photoperiod Sensitive Genic Male Sterility in Rice. Journal of integrative plant biology. 51, 774-781.

Sisco P.H. (1991). Duplications complicate genetic mapping of Rf4, a restorer gene for cms-C cytoplasmic male sterility in corn. Crop science. 31, 1263-1266.

Stam P., Van Ooijen J.W. (1995). JOINMAP™ version 2.0: Software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen, The Netherlands.

Van Ooijen J.W., Voorrips R.E. (2001). JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen (The Netherlands).

Van Stallen N., Vandenbussche B., Verdoodt V., De Proft M. (2003). Construction of a genetic linkage map for witloof (Cichorium intybus L. var. foliosum Hegi). Plant Breeding 122, 521-525.

Varotto S., Pizzoli L., Lucchin M., Parrini P. (1995). The incompatibility system in Italian red chicory (Cichorium intybus L.). Plant breeding = Zeitschrift für Pflanzenzüchtung. 114, 535-538.

Varotto S., Nenz E., Nucchin M., Parrini P. (2001). Production of asymmetric somatic hybrid plants between Cichorium intybus L. and Helianthus annuus L. Theoretical and applied genetics. 102, 950-956.

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

Vinod K.K. (2005). Cytoplasmic genetic male sterility in plants - A molecular perspective. In: proceedings of the training programme on "Advances and Accomplishment in Heterosis Breeding", Tamil Nadu Agricultural University, Coimbatore, India. Downloaded from http://kkvinod.webs.com pp. 147-162.

Wan L.L., Xia X., Hong D., Li J., Yang G. (2010). Abnormal Vacuolization of the Tapetum During the Tetrad Stage is Associated with Male Sterility in the Recessive Genic Male Sterile Brassica napus L. Line 9012A. Journal of Plant Biology 53, 121-133.

Wittop Koning D.A., Leroux A. (1972). La chicoré e dans l'histoire de la médecine et dans la céramique pharmaceutique. Rev. Hist. Pharm. (Suppl.) 215: pp. 28.

Yi B., Chen Y., Lei S., Tu J., Fu T. (2006). Fine mapping of the recessive genic male-sterile gene (Bnms1) in Brassica napus L. Theoretical and applied genetics. 113, 643-650.

Appendix 2.1. CDS and genomic nucleotide sequences of MPF2-like_long and MPF2-like_short genes.

>*CiMPF2-like_long* complete CDS sequences

ATGGCGAGAGAGAAGATAAAGATAAGAAAGATCGATAATATAACGGCAAGGCAGGTGACTTTTTCCAAAAGAA
GAAGAGGGCTTCTGAAGAAAGCTGAAGAACTCGCGGTACTTTGTGACGCCGACGTCGCTCTGGTCATCTTCTC
CGCCACAGGAAAACTTTTCGAGTATGCTAGTTCCAGTATGCAAGAGTTACTCGGAAAATATAAACTCCACTCG
ACTAATAACGTCAATAAAGTTGACGAACCTTCTCTCGATCTACAGCTCGTGGAAAGCCAAGAGTCACGGATGA
GCCAAGAGGTACTAGAAAAGGATCGGGAACTGAGCCAATTACGTGGAGAGGACCTTCAAGGCTTAACTCTAGA
GGAACTTCAGCGTTTGGAGAGCTTGCTCGAAGGAGGGCTCAACCGTGTGCTCCAGACAAAGGATGAACGGATT
GCAAACGAGATAGCAAGTCTTCAACAAAAGGGTTTACAGTTGATGGAAGAGAACAAGCTCTTGAAACAACAAA
TGATGACTTTGACTACGCTAGGCAAGAGGCCACGGACGACGGCTGCCGAGTTGGACAATATTGTAATTAACCC
TGAAGATCAAGGGCAGTCATCGGACTCTGTTGCTACCAACGTATACAGTTGTAACAGTGGACCTCCTCCAGAG
GATGATTGGTCTGATACTTCTCTTAAACTAGCGTTGCCCTTTAACTGA

>*CiMPF2-like_short* complete CDS sequences

ATGGCGAGAGAGAAGATAAAGATAAGAAAGATCGATAATATAACGGCAAGGCAGGTGACTTTTTCCAAAAGAA
GAAGAGGGCTTCTGAAGAAAGCTGAAGAACTCGCGGTACTTTGTGACGCCGACGTCGCTCTGGTCATCTTCTC
CGCCACAGGAAAACTTTTCGAGTATGCTAGTTCCAGTATGCAAGAGTTACTCGGAAAATATAAACTCCACTCG
ACTAATAACGTCAATAAAGTTGACGAACCTTCTCTCGATCTACAGCTCGTGGAAAGCCAGGAGTCACGGATGA
GCCAAGAGGTACTAGAAAAGGATCGGGAACTGAGCCAATTACGTGGAGAGGACCTTCAAGGCTTAACTCTAGA
GGAACTTCAGCGTTTGGAGAGCTTGCTCGAAGGAGGGCTCAACCGTGTGCTCCAGACAAAGGTTTCTTTTAGA
TTCTTTCATGTTCTTGTTCTAGCTTTGGATTTAATATAA

>*CiMPF2-like_long* partial genomic sequence

ATGGCGAGAGAGAAGATAAAGATAAGAAAGATCGATAATATAACGGCAAGGCAGGTGACTTTTTCCAAAAGAA
GAAGAGGGCTTCTGAAGAAAGCTGAAGAACTCGCGGTACTTTGTGACGCCGACGTCGCTCTGGTCATCTTCTC
CGCCACAGGAAAACTTTTCGAGTATGCTAGTTCCAGTATGCAAGAGTTACTCGGAAAATATAAACTCCACTCG
ACTAATAACGTCAATAAAGTTGACGAACCTTCTCTCGATCTACAGGTTTAGTCCCTTGATCTTTGACCAACTT
ATTCGTATTACATTATTTCGTTATGCATGGTTGGTTTAGATGCTTTTACAATGGATATGTCGGAGGAAGTCAT
TAATTTTGAAAATAAAACCAAGTTAGTATTGATATGCTTTTGCTGCTAATTCCCCCATTGCAAATAAACATGA
TTTTCCGATGGTAAAACATGGTATCCTTATGTAGCTCGTGGAAAGCCAAGAGTCACGGATGAGCCAAGAGGTA
CTAGAAAAGGATCGGGAACTGAGGTACTTTATTTTTATTAATTTAAATTTCATAATTCACAATAATGAAACTT
CATGGCTAATCTATGATTTCGTCACGAAATACAGCCAATTACGTGGAGAGGACCTTCAAGGCTTAACTCTAGA
GGAACTTCAGCGTTTGGAGAGCTTGCTCGAAGGAGGGCTCAACCGTGTGCTCCAGACAAAGGATGAACGGATT
GCAAACGAGATAGCAAGTCTTCAACAAAAGGTTATAAGTGTTTTTATACTCAACCTTCTCAAACTAGAGTCCG
TTTTTGACTTTCAAGAGTTAAATGAGTAGGTTTTAATTTCAAAATCATATCAAATTACACATGTTTTAAAAAG
AAAAAAAAGTGAGTATACTTCAGTTATATATGTTGATGTAAAAGTGTTTCTGAAAAATACTAAAAATATTCAT
ATAGAACTTTCAATTCTGCATTAATAGTGACTTTCGAAAGTCAAACTTCAAACTCAAATGCTTTAGGAAACAA
CATCACATTACCGCATTCAACTATTTAATTTGTTCTCAAAATCAATCAAATTTATACTAAGCTTGACTCTATG
ATGTCAAACGTTTGACTCTATAATTTATTAGAAATTGACTTTTATACATTTTCCTGGTATTAAAAGAATTTCT
ATTTATATTATGCAGGGTTTACAGTTGATGGAAGAGAACAAGCTCTTGAAACAACAAATGATGACTTTGACTA
CGCTAGGCAAGAGGCCACGGACGACGGCTGCCGAGTTGGACAATATTGTAATTAACCCTGAAGATCAAGGGCA
GTCATCGGACTCTGTTGCTACCAACGTATACAGTTGTAACAGTGGACCTCCTCCAGAGGATGATTGGTCTGAT
ACTTCTCTTAAACTAGCGTTGCCCTTTAACTGA

Ringraziamenti

Un particolare ringraziamento al Prof. Gianni Barcaccia, supervisore della mia tesi, il quale mi ha consigliato e revisionato i lavori in tutti questi anni. Un ringraziamento anche al Prof. Angelo Ramina, per i preziosi consigli dati.

Un ringraziamento al CNR di Perugia. In particolare vorrei ringraziare la Dott.ssa Baldoni Luciana, per i materiali vegetali messi a disposizione, un particolare ringraziamento alla Dott.ssa Fiammetta Alagna, per il preziosissimo aiuto teorico e pratico, e alla Dott.ssa Eugenia Caceres, per il supporto citologico.

Vorrei ringraziare anche tutti i gruppi di ricerca con i quali abbiamo collaborato in questi anni: il gruppo del Dott. Perrotta (ENEA C.R. Trisaia), il gruppo del Prof. Muleo (Università della Tuscia), il gruppo del Prof. Valle (Università di Padova).

Inoltre vorrei ringraziare il Prof. Hans De Jong (Wageningen University), la Prof. Elena Kramer (Harvard University) e i signori Tiozzo Silvano e Ubaldo (T&T®).

Infine, ma non meno importanti, vorrei anche ringraziare tutti i colleghi ed ex-colleghi che mi hanno accompagnato in questi anni.

Finanziamenti

Le ricerche sull'incompatibilità in olivo sono state condotte nell'ambito del Progetto Strategico MiPAAF “OLEA - Genomica e Miglioramento genetico dell’olivo”, D.M. 27011/7643/10. Nell'ambito di tale progetto di ricerca sono state tenute numerose comunicazioni scientifiche a congressi nazionali ed internazionali, e sono già stati prodotti diversi lavori scientifici, alcuni dei quali già disponibili come pubblicazioni mentre altri sono al momento in fase di preparazione. Le ricerche in radicchio sono state condotte nell'ambito del Progetto "UBALDO" finanziato dalla T&T® di Sant'Anna di Chioggia (Venezia) SS Romea 77/bc. Tale progetto ha già portato alla costituzione di due varietà innovative, T&T4070-F1 e T&T4020-F1, attualmente in fase di registrazione presso l’UPOV (International Union for the Protection of New Varieties of Plants, http://www.upov.int) e nel Catalogo Europeo delle Varietà Vegetali (CPVO, Community Plant Variety Office, http://www.cpvo.europa.eu). Al momento la protezione dei mutanti maschio-sterili è assicurata a livello internazionale da un brevetto pendente (Patent Cooperation Treaty, domanda No. PCT/EP2011/058765), in cui sono riportate le loro caratteristiche cito-genetiche e molecolari distintive che rendono i mutanti stessi facilmente identificabili a livello fenotipico e tracciabili anche a livello genotipico.