UNIVERSITÀ DEGLI STUDI DI FOGGIA

DIPARTIMENTO DI SCIENZE AGRARIE, DEGLI ALIMENTI E DELL'AMBIENTE

DOTTORATO DI RICERCA IN
"BIOTECNOLOGIE DEI PRODOTTI ALIMENTARI" (XXVIII CICLO)

COORDINATORE: PROF. MATTEO ALESSANDRO DEL NOBILE

"Genetic analysis on biosynthesis of compounds affecting bitterness in almond tree (*Amygdalus communis L.= Prunus Dulcis Mill.*)"

Dottoranda: Francesca Ricciardi

Tutor: Prof.ssa Concetta Lotti
Co-tutor: Prof.ssa Laura De Palma

ANNO ACCADEMICO 2014-2015

INDEX

SUMMARY

Wild-type seeds of Prunoidaeae are bitter and toxic as they accumulate amygdalin, a cyanogenic phytoanticipin thought to have an important role in defense against herbivores and pathogens. The sweet taste of the seeds of cultivated almond (*Prunus dulcis* Miller D.A. Webb syn. *Prunus amygdalus* Batsch) originates from a mutation selected during early domestication and maintained in modern cultivars. Previous genetic analyses revealed that the almond seed taste depends on the genotype of the mother plant. Moreover, it is under monogenic control, and the allele associated with sweetness (*Sk*) is dominant over the allele conferring bitterness (*sk*). In previous studies, the *Sk* locus was positioned on almond linkage group (LG) 5. However, the molecular nature of the *Sk* gene is still elusive.

The genome sequence of peach (*Prunus persica* L.), a species evolutionary close to almond, has been recently released and is publicly available at the Genomic Database of Rosaceae (GDR). In this thesis, we exploited peach genomic resources to identify new molecular markers linked to the *Sk* locus in almond. Moreover, we provide a fine-map of the *Sk* genomic region and identified a few candidate genes for kernel taste.

The SSR markers BPPCT037, CPDCT028, EPDCU2584 and UDA045 were previously shown to be closely linked to the *Sk* locus on the almond LG5. BLAST search for peach nucleotide sequences similar to these markers

resulted in the identification of best matches all positioned on scaffold 5, suggesting the identification of a synthenic region between the peach and almond genomes. The two markers UDA045 and CPDCT028, previously reported to flank the *Sk* locus, were found to delimit a region of about 810 Kb in the physical map of peach, likely containg the *Sk* gene ortholog.

Aiming to saturate the *Sk* region, we designed several primer pairs on peach genes included in the 810 Kb interval flanked by UDA045 and CPDCT028. Sequence analysis of PCR products revealed the presence of six SNPs segregating in a large F_1 mapping population of 476 individuals, obtained by crossing the two parental genotypes R1000 and Desmayo Largueta. This information was used to develop CAPS markers linked to the *Sk* locus.

Newly identified CAPS markers and previously reported *Sk*-linked SSR markers were used to fingerprint the F_1 population above mentioned. Marker data were merged with phenotypic scores and used to produce a saturated map of the *Sk* region. Notably, the marker order on almond LG5 resulted to be fully in accordance with the order of ortholog sequences in peach scaffold 5, indicating collinearity between the two genomes. Having provided strong evidence for genome collinearity between almond and peach and the identification of almond markers flanking the *Sk* locus, we concluded that the *Sk* ortholog in peach is likely located in a physical region of about 80 Kb,

delimited by the genes referred to as ppa003882m and ppa005388m according to the nomenclature used in the GDR.

Primers designed on the seven peach genes included in the interval delimited by the genes ppa003882m and ppa005388m were used on almond cDNA obtained by the kernel tegument tissues of sweet and bitter almonds for a real-time qPCR assay, in order to detect differences in expression levels that would have indicated the implication of these genes in amygdalin accumulation. Two of them were found to be more expressed in sweet genotypes, although difference were not always significant in all the time points of kernel ripening stage under test.

Researches performed during the PhD period brought to the identification of a series of molecular markers linked to the *Sk* locus in almond. This is of great interest for almond breeding, as they can aid the selection of sweet genotypes without the need to wait for a long juvenile phase and carry out phenotypic tests. Specific advantages of the CAPS markers identified in this study regard their close linkage with the *Sk* locus, which minimizes errors due to recombination, and their co-dominant nature, which allows to distinguish between sweet homozygous and heterozygous individuals. Moreover, CAPS markers can be obtained by means of simple protocols that require relatively inexpensive laboratory equipment.

This thesis also provides a fine-scale map of the genomic region around the *Sk* gene, which lays a foundation for its isolation via positional cloning and

corroborates previous reports indicating high level of synteny and collinearity between species of the *Prunus* genus.

According to the map provided in this thesis, and assuming collinearity between peach and almond, the *Sk* gene should be one of the ortholog of the seven peach genes flanked by the genes ppa003882m and ppa005388m. Five of them (ppa022201m, ppa025417m, ppa027182m, ppa015634m and ppa005343m, according to the GDR nomenclature) encode MYC transcription factors. The remaining two genes, ppa011942m and ppa023406m, respectively encode the MED10 component of the mediator complex, functioning as transcriptional co-activator in eukaryotes, and a putative galactose oxidase. The almond ortholog of the peach MYC transcription factor ppa022201m is a very good candidate for being the *Sk* gene, as its expression was found to be higher in the tegument tissue of a sweet genotype than in a bitter genotype, as resulting from real-time qPCR.

Overall, we believe that this thesis might be of interest for basic research and applied breeding. Currently, we are testing newly identified CAPS polymorphisms on several almond genotypes, in order to test their possible applicability in practical breeding programs, and sequencing the genomic interval presumably containing the *Sk* locus in sweet and bitter almond genotypes. Moreover, we are carrying out studies addressed to the functional characterization of MYB transcription factor above mentioned which is a candidate for being the *Sk* gene.

INTRODUCTION

INTRODUCTION

1. Almond species [(*Prunus dulcis* (Mill.) D.A.Webb; syn. *Prunu amygdalus* Batsch)]

1.1 Almond origin and history of crop

The almond [*Prunus dulcis* (Mill.) D.A.Webb; syn. *Prunus amygdalus* Batsch], species belonging to the *Prunus* genus, is an ancient tree crop cultivated for its edible kernel and nowadays its growing area is still raising worldwide.

Many studies were performed to establish the origin of almond. Several wild species, as *Prunus fenzliana* (Fritsch) Lipsky, *Prunus bucharica* Korschinsky and *Prunus kuramica* (of the section Euamygdalus), are found from Western China (Tian Shan mountain) to the deserts and mountainous area of Kurdistan, Turkestan, Afghanistan and into Iran and Iraq (Grasselly, 1976b; Kester and Gradziel, 1996). They were described to be closely related to almond, and might be the ancestral species of the modern cultivated almond (Grasselly, 1976a; Kester *et al.,* 1991). Anyway, Ladizinsky in 1999 suggested *P. fenzliana* as the only wild ancestor of almond, contrariwise Grasselly, (1976a, b) and Browick and Zohary (1996) described *P. webbii* (Spach) Vieh, native to the Balkan peninsula, as a species strongly related to almond. Furthermore, in 1979 Watkins suggested the cultivated almond came from selection within species listed originally as *Amygdalus communis* (syn. *Prunus communis* Archangeli) based on studies of two natural populations containing many sweet kernelled individuals rather than the bitter ones typically found in the wild species. After that, *Amygdalus communis* (syn. *P. communis* Archangeli) cultivation was employed in Iran, Transcaucausus, Eastern Turkey, Syria, and thus overlapped with known sites of antique almond cultivation (Denisov, 1988; Kester *et al.,* 1991).

According with these theories, Ladizinsky (1999) realised the natural sweet seed populations looked like the phenotypic range of almonds grown today and considered that the Kobet Dagh and Tian Shan populations were recent remnants of later domesticated or semi-domesticated almond field.

Recently, Gradziel (2009) and Zeinalabedini *et al.,* (2010) proposed that cultivated almond represented a generalized and interchangeable kernel phenotype, probably derived from natural inter-specific cross-hybridization among *P. fenzliana* and a range of related species as *P. bucharica, P. kuramica* and *P. triloba* that occurred naturally (**Fig. 1**).

The further evolution and distribution of almonds both in cultivation and in the associated semi wild state, occurred in three stages: Asiatic, Mediterranean, and Californian, corresponding to the geographical areas where almond is grown and spread (**Fig. 1**) (Grasselly, 1976a; Kester *et al.*, 1991; Kester and Gradziel, 1996; Gómez, 2007; Gradziel, 2009). Particularly the first stage is associated to the beginning of the almond domestication and its subsequent diffusion through Afghanistan, Iran, Turkey and the Southern Republics of the old USSR.

In the Mediterranean stage, it was supposed that almond was brought in Greece prior to 300 BCE, maybe during the Phoenicians and Greeks trades and by moving of their colonies to Sicily and other Mediterranean places (Bacarella *et al.,* 1991). Romans and Berbers also concurred to the almond distribution throughout that area. Moreover, around 500-600 CE, Arabs spread almonds in Southern Spain and Portugal, by the conquering of the North Africa. The continuous cultivation (2000 and more years) in the Mediterranean allowed getting local almond ecotype selections very well suited to different climate and environmental conditions, specific of each regions.

Finally, the Californian stage began as an extension of Mediterranean culture and almond germplasm has been brought from Spain, but later, the first almond varieties from the Languedoc region, in Southern France, were

introduced in California in the XIX, signing a "new époque" of high diffusion and production of Californian almonds in the world.

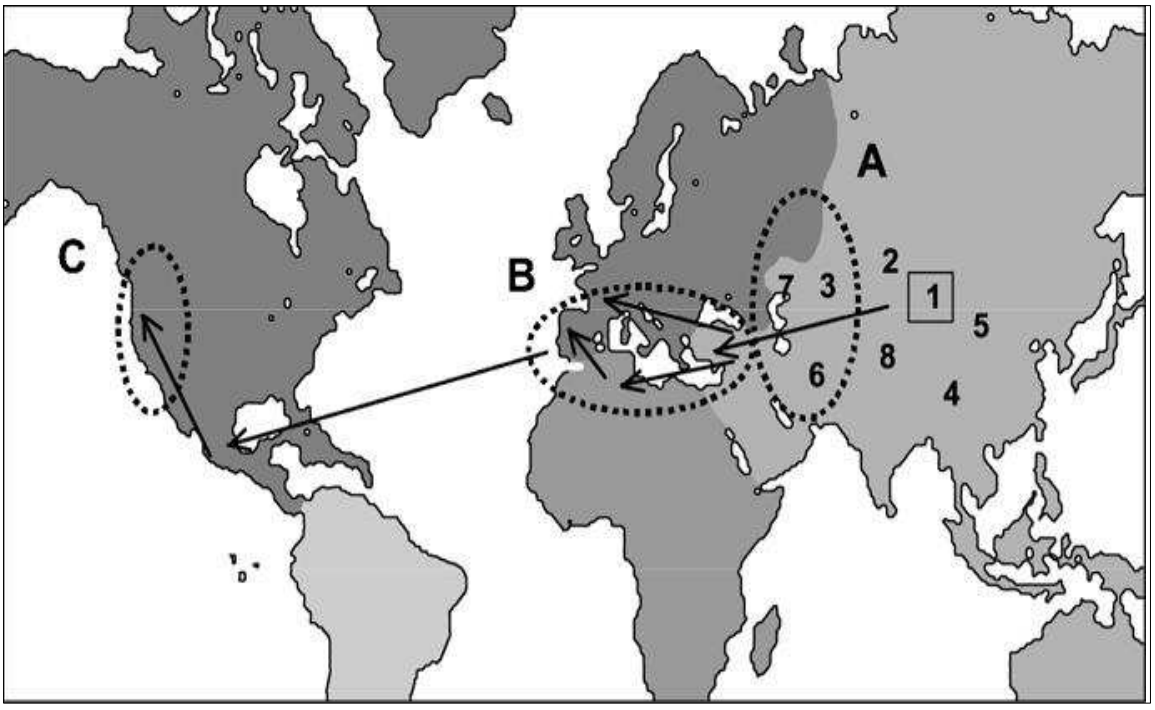


Fig. 1- Map of world showing the origin of almond [*Prunus dulcis* (*1*)] and different relative *Prunus* species [*P. bucharica* (*2*), *P. fenzliana* (*3*), *P. davidiana* (*4*), *P. persica* (*5*), *P. scoparia* (*6*), *P. webbii* (*7*), and *P. argentea* (*8*)]. The dissemination routes for the cultivated almond [→], and the three main areas for diversification and cultivation of almonds [Asiatic (*A*), Mediterranean (*B*), and Californian (*C*)]. Source: Gómez, 2007

1.2 Taxonomy, morphology and floral biology

Almond (*Prunus dulcis* Mill., D.A.Webb) is classified within the genus *Prunus,* in the subgenus *Amygdalus,* and belongs to the *Rosaceae* family, as many stone fruit trees like peach (*Prunus persica*), apple (*Malus domestica*), plum *(Prunus* spp.*),* apricot *(Prunus armeniaca)* and black cherry *(Prunus serotina*) (**Tab. 1**).

The botanical nomenclature of almond was a riddle for many years; Vavilov (1930), Jukovski (1950), Rickter (1953), Browicz (1974) and Seramifov (1971), as reported by Muncharaz Pou (2003), adopted the *Amygdalus communis* denomination (Grasselly, 1984) according with Linnaeus classifications (1953). *Prunus dulcis* was the qualification set by Miller in

1768 and then accepted by Grasselly in 1997, to distinguish almond from other *Prunus* species (Muncharaz Pou, 2003). Finally in 1964, the General Committee of Botanical Nomenclature proposed *Prunus dulcis* (Miller) D.A. Webb, including it in the International Code of Botanical Nomenclature (Muncharaz Pou, 2003), and it considered synonyms *Prunus amygdalus* (Batsch) and *Prunus communis* (Archangeli).

Tab. 1- Taxonomic classification of almond

Kingdom	Plantae
Phylum	Spermatophyta
Subphylum	Angiospermae
Class	Dicotyledoneae
Subclass	Rosidas
Super Order	Rosanae
Order	Rosales
Family	Rosaceae
Subfamily	Prunoidea
Genus	*Prunus*
Subgenus	*Amygdalus*
Specie	*Prunus dulcis*

The almond is a perennial tree with a trunk of about 30 cm in diameter and 4-10 m in height. The leaves are deciduous, alternate, narrow, elongated and pointed; their size and disposition within the stem are distinctive traits among the varieties. A very recent work gave importance to almond leaves as discreet source of phenols, with antibacterial and antioxidant activities, which can be used in food and pharmaceutical industries (Tiwari *et al.,* 2015).

Flowers are peryginous self-incompatible, typical of *Prunus* and they appear before the leaves in early spring (Griffiths and Huxley, 1992; Rushforth,

1999). They show five petals produced singly or in pairs, colored from white to pale pink.

Cultivars have different petal size, shape, color, arrangement and length of anthers. The number of stamens can range from 20 to 40, but usually it is around 30 (Ensminger, 1994) (**Fig. 2**).

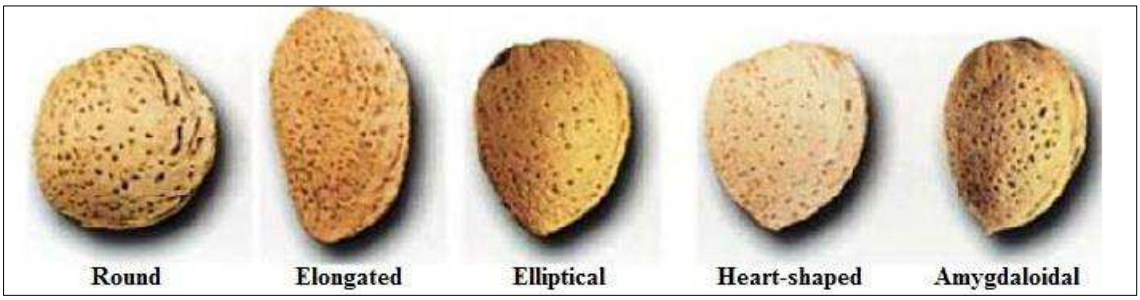


Fig. 2- Examples of almond fruit shape (Muncharaz Pou, 2003)

One of the most important adaptive traits in almond is the time of blooming, because of the vulnerability to spring frosts and rains. It is strongly affected from the breaking dormancy, depending in turn by the amount of chilling in winter, the exposure to warm temperatures in spring before bloom, and the threshold temperatures of the bud growth (Muncharaz Pou, 2003).

Almond reproduction pattern, which occurs in Angiosperm subphylum, has the peculiarity to show double fecundation events, with the purpose to produce the endosperm, a triploid vegetal tissue highly nutritive to feed the embryo during the early stages of development (Kordium, 2008), (**Fig. 3**). Moreover, both cultivated and most related almond species express gametophytic self-incompatibity controlled by a single *locus* (S) with multiple co-dominant alleles series (Socias i Company and Felipe, 1988; Dicenta and García, 1993a); hence, self-fertilization favors cross-pollination, but high genetic variability, within seedling populations, is preserved.

Many "cross-incompatibility groups" and also the corresponding alleles have been identified (Tamura *et al.,* 2000; Boskovic *et al.,* 2003; Channuntapipat *et al.*, 2003; Lòpez *et al.*, 2004) and still they manage the

selection of cultivar combinations used in orchard by marker assisted selection (MAS). In California, for example, these groups have been identified for all major and important almond varieties (Barckley *et al.*, 2006).

Already in the eighties, self-compatibility was one of the main objectives for almond breeding programs in Europe and the USA (Grasselly *et al.*, 1981; Vargas *et al.*, 1984; Socias i Company and Felipe, 1988; Dicenta and García, 1993a; Gradziel and Kester, 1998). Nowadays, different degrees of self-compatibility and self-fruitfulness have been discovered (Gradziel *et al.*, 2002) and self-compatible almond cultivars were also reported, as in Apulia where *P. webbii* has became a source of self-fertility (Godini, 2002).

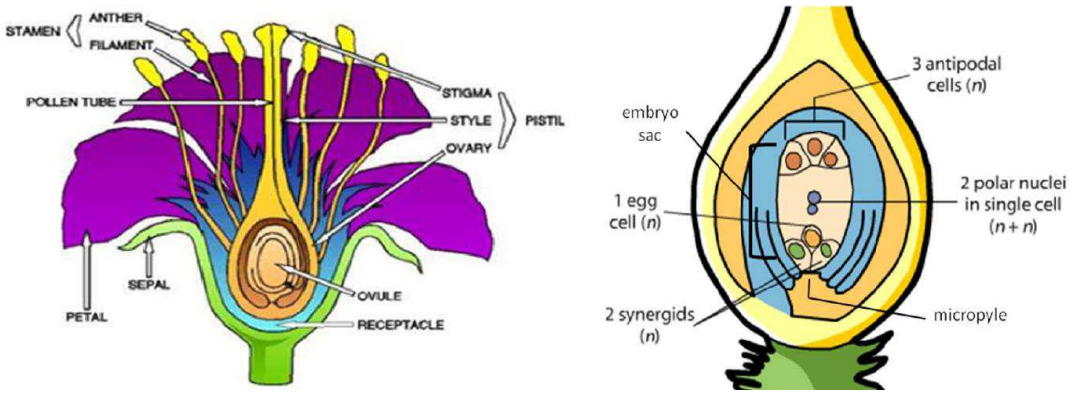


Fig. 3- Parts of almond flower and action on the pistil during double fertilization

1.3 The almond tree crop production and diffusion

Almond tree is one of the most valuable nut crop globally (Chen *et al.,* 2005; Wijeratne *et al.*, 2006) due to its fruits of high commercial value (Cordeiro and Monteiro, 2001; Martins *et al.*, 2003; Moure *et al.*, 2007). However, there is a great diversity among almond cultivars which exhibit different productivity and yields of seed in the fruit (Martinez *et al.,* 1995).

Basing on data collected by Food and Agriculture Organization (FAO) in 2013, the world produced around 2.92 million tons of almonds, despite to 2011 with a production of 2.00 million tons. The United States of America is confirmed to be the largest producer, growing its production from 0.73 million

tones in 2011 to 1.8 million tonnes of almonds in 2013 (**Fig. 4**). At in the past, most of the US almonds are grown in California (Jahanban *et al.*, 2009; Wijeratne *et al*., 2006), in an area of almost 400.000 ha (Sathe *et al*., 2002) of irrigated land, and generate an annual crop value of $6 billion US in 2012 (https://www.cdfa.ca.gov/). Italy produces 0.10 millions tons of almonds with a yield for hectares of 1.3 tons. Other important Countries are Spain, Iran and Marocco. USA, Afghanistan, Iran and Turkey shown the highest values of yield/ hectares.

Country	Production (2010- 2011) in million of tons	Yields (2010-2011) in tons/hectare
USA	1.41 ▼ 0.73	4.85 ▼ 4.50
Spain	0.22 – 0.21	0.40 – 0.40
Iran	0.16 – 0.17	2.97 ▼ 1.91
Italy	0.11 – 0.10	1.26 ▲ 1.39
Morocco	0.10 ▲ 0.13	0.98 ▲ 1.52
Syria	0.073 ▲ 0.13	1.49 ▲ 2.52
Afghanistan	0.056 ▲ 0.061	5.00 ▼ 4.50
Turkey	0.055 ▲ 0.070	3.23 ▲ 3.41
Tunisia	0.052 ▲ 0.061	0.32 – 0.32
Algeria	0.039 – 0.050	1.16 ▲ 1.80
World Total	**2.51 ▼ 2.00**	**1.62 ▼ 1.27**

Fig. 4- Country productions and yield per hectare recorded in two years (2010/2011). Source: FAOSTAT, 2011

Considering the period 2007-11 (FAOSTAT, 2011), the production of almonds in the EU on average remained stationary with 37456 tons, surface (ha) of 69954 and yield per ha of 10755.

In EU almond is widespread especially in Spain, Italy and Portugal, while has little importance in other Countries (**Tab. 2**). In particular, Spain has the largest area under cultivation, estimated at over 436,500 ha. In the last five years, almond growing areas and total production have increased in US, Spain, Tunisia, Morocco and Italy.

Tab. 2- Area and almond production in shell and shelled in EU in 2007-2011

Countries	Area (ha)	Production (Mg)	Yield (Mg/ha)
Bulgaria	1326	287	0,237
Croatia	371	788	1,942
France	1276	1043	0,815
Greece	15095	36570	2,414
Hungary	196	139	0,703
Italy	80115	110195	1,376
Portugal	27820	8071	0,290
Spain	572763	216517	0,377
Macedonia	488	944	1,933
EU Average	63595	34051	1,075

Source: Faostat 2007/2011

1.4 Almond cultivation in Italy

In the 40's and 50's, the trade of almonds represented an essential component of the agricultural economy of Italy, that was the largest almond producer in the world. The main hub of almond production was located in Sicily and Apulia, however, in that years, regions as like as Sardinia, Basilicata, Calabria, Abruzzo, Campania also reached a significant production.

In the sixties and seventies, almond production greatly increased in USA, contributing to the growth of almonds production in the world. During this period, Italian production remained substantially the same. In the next two decades of the 80' and 90', the situation worsened increasingly, causing a crisis in the trade, so deep as to cause concern an actual disappearance of cultivation

of the almond tree in all Italian regions concerned. Although Spain suffered the effects of the expansion of production in California, it has significantly increased its production. Similarly, Australia, in just 10 years since 2000, has reached high levels of production becoming the second largest producer in the world. Therefore, while in Italy a massive reduction of areas dedicated to cultivation of almonds has been noted, world production, compared to the postwar period, increased hugely from just over 87,000 tons of shelled product in the five years from 1947 to 1951 to over 921,000 tons in 2010/2011. The expected increase from 2010 is due in part because of growing demand from emerging countries, such as with China and India (*World Nut & Dried Fruit Congress*, 2011).

The cultivation of almond in Italy is actually focused in two key areas, the islands and the South. Although islands have a major cultivated area, the South has the largest number of farms cultivating almond. As reported in **Tab. 3**, in which Italian cultivation area and almond productions are showed as mean of the 2007-2011 period, the main sites of production are in Sicily (mean: 47,746 ha), which occupies the first place, and Apulia (mean: 26868 ha), with production of 748,3 and 274,9 hq, respectively.

Almond Italian trade, in decline in the last years might return to represent for many Southern regions an important economic resource, in particular increasing the quality of our almonds in the domestic and international markets.

Tab. 3-Total area and almond production in major Italian regions recorded as means of 2007-2011 period

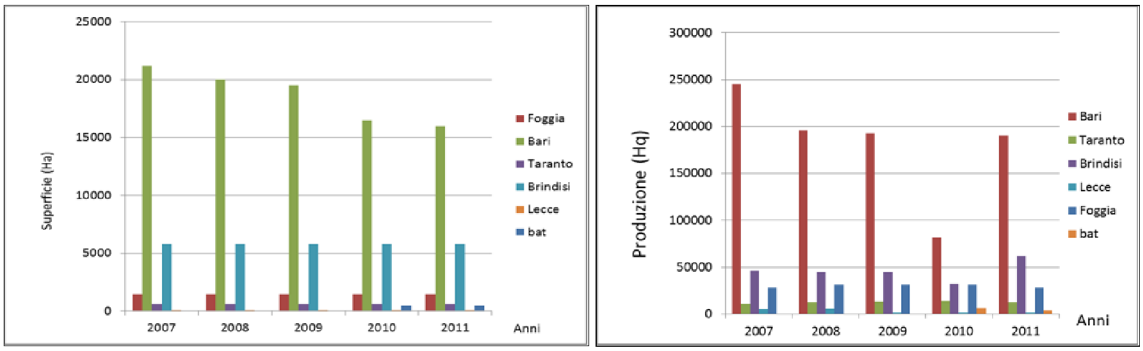
	Almond	
Italian regions	**Total Surface (Ha)**	**Harvested production (Hq)**
Abruzzo	141	366
Apulia	26.868	274.960
Basilicata	46	2.151
Calabria	241	5.266
Sicily	47.746	748.333
Sardinia	3.276	25.234

ISTAT: 2007-2011

1.5 Cultivation of almond in Apulia

During the last years a decrease of almond cultivation in Apulia region was recorded, but actually an opposit trend in relation to a higher demand from the market. The main Apulian area cultivated with almond, with over 68% (16000 ha), falls mostly in the area of Murgia, near Bari, and in Ceglie Messapica, Carovigno and Ostuni, close to Brindisi, with 15.7% (5800 ha). In the other four districts of the region there is an area of 2265.95 ha (15.4%) planted with almond.

Nowadays, in Apulia, the whole almond cultivation holds an area around 24520 Ha (32.4%) of Italy, providing a total production of 300875 quintals (q) of almond (28,65% on the total production of Itlay), with a production respectively of 190.000q for the province of Bari and 62000q for Brindisi (**Fig. 5** and **6**).



Figg. 5-6- Total cultivated area (ha) and production (Hq) of almond in the six provinces of Apulia. Source: ISTAT, 2007-2011

The main farms are located near Bari (11147), followed by Brindisi (4.632), Taranto (971), Foggia (731) and Barletta-Andria-Trani (546) (ISTAT, 2007-2011).

Apulia is characterized by a still high number of varieties and/or ecotypes, which were selected in the past decades and handed down with the practice of grafting.

2. Almond seed: a powerful nutritive and multi-purpose source

2.1 Morphology and fruit development

The fruit of almond, as other *Prunus* species, is a drupe with a single seed and two large cotyledons. Some varieties have often double seeds (**Fig. 7**) and this trait is considered as an important commercial defect due to the lacking of yield uniformity, which complicates the industrial transformation process (Gulcan, 1975).

Morphologically, the fruit is arranged by a outer pubescent skin (exocarp), a middle fleshy and thin hull (mesocarp), that at the maturity becomes dry and split, revealing an inner woody brown shell (endocarp) with function to protect the seed.

Fig. 7- Example of almond double seed

Almond is classified as “nut”, whose seed (kernel or meat) is the edible part and the commercial product (Martìnez-Gomez *et al.*, 2008). Fruit development follows three stages after fertilization, in which the pericarp, kernel and nucellus develop during the first stage, while embryo and endosperm enlarge during the second, and the latter is slowly replaced by growing of the cotyledons, and finally the dry weight of the embryo increases during the third one. The tegument (seedcoat or skin) is a diploid mother brown tissue covering and protecting the kernel by the oxidation and microbial contamination (Werner and Creller, 1997; Dicenta *et al.,* 2007).

The fruit ripening is 7–8 months after flowering (Griffiths and Huxley, 1992) and during the early steps of growth in the spring, linear dimensions of length, width and thickness can be measured to express the kernel size, till early summer. Shape, instead, is a function of differences in length, width, and thickness and, irregularities between the last two parameters may change the visual effect of the seed significantly. Moreover, dehiscence of the hull, its splitting along the fruit suture and dehydration, and peduncle abscission, are all physiological processes involved during fruit ripening and important commercial traits in almond (Felipe, 1977). The entire process of hull and nut maturation and drying may require 2–6 weeks to complete. Usually, maturity is most reliably characterized by the initiation and progress of hull splitting

(King *et al.*, 1970). The dates for the initiation of 5%–10% splitting and it completion is useful criteria in comparison with standard cultivars.

The size, endocarp consistency, shape, pubescence, retention of the pistil remnants, nature of suture lines (Monastra *et.al.*, 1982), such as tegument color, roughness, of almond meets, are strongly distinctive traits among varieties, suggesting specific marketing categories and uses. Shell hardness, is associated with the total amount of lignin deposited to the shell during nut development (**Fig 8**). Shelling proportion (dry weight of kernel/dry weight of in-shell nut) is used to obtain a quantitative measure of shell density and is an example of trait used in commercial activities. For several years, almond fruit quality has been reported only considering physical parameters (Socias i Company, 1998) and consumer preferences made the main differences in a definition of fruit quality (Janick, 2005; Socias i Company *et al.*, 2008).

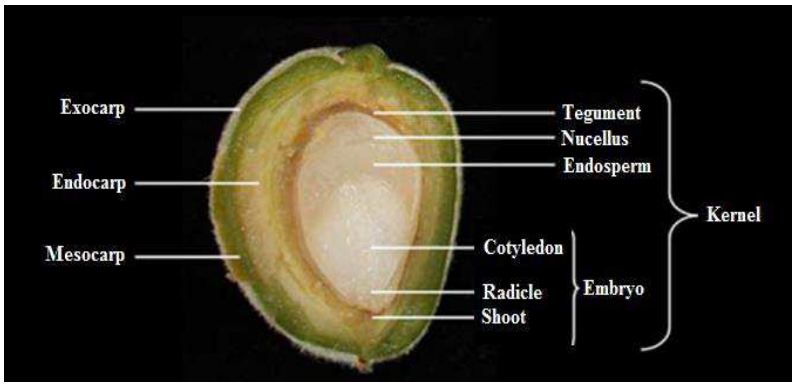


Fig. 8- Parts of almond fruit and seed

Although Torre Grossa *et al.* (1994) and Dicenta *et al.* (2002) carried out several studies about almond traits by using different pollination treatments, no difference was found among the varieties tested and none information about the chemical composition as a function of pollination type was available. However, seed composition in relation to its possible industrial utilization (Alessandroni, 1980) or quality stability (Kester *et al.*, 1993) should be an important features to evaluate the actual effect of self-compatibility. So

recently, Kodad and Socias i Company (2015) carried out an important research about the almond quality related to the type of pollination in self-compatible genotypes. In particular, results shown how the self-pollination could be responsible for positive effect on the fatty acid composition, improving the quality of almond kernel.

2.2 Potential application and properties

2.2.1 Almond kernel and nutritional values

Almond kernels are valuable for their health, sensory and nutritional attributes (Venkatachalam and Shridhar, 2006) being one of the complete sources of energy because of their high protein and lipid content (50%–55%), good flavor, crunchy texture and good visual appeal. Kernels have many important food applications (Rosengarten, 1984; Fraser *et al.*, 2002; Kendall *et al.,* 2003). Indeed, nutritional value of almond fruit (tab. 4) is related to its kernel, other parts such as brown skin, shell and hull which represent the 55% of fruit weight, but they are discarded as agricultural by-products (Martinez *et al.*, 1995; Fadel, 1999); only recently, researchers have begun to study alternative methods to employ them.

The seeds are usually harvested as mature and eaten either raw or roasted, if used as snack foods, or as ingredient in a variety of processed foods, especially in bakery and confectionery industry (Woodroof, 1979; Kester and Gradziel, 1996; Schirra, 1997; Sang *et al.,* 2002b). In particular, natural almonds (with shell) may be sold as whole or processed into various forms without shell, indeed, meat is used either un-blanched (with skin) or blanched in which the tegument is removed by hot water or steam (Sang *et al.,* 2002b; Wijeratne *et al.* 2006). It can also be sliced or diced to be used in pastry, ice cream, breakfast cereals, vegetable mixtures and also ground into paste to be used in bakery products, confectionary and in the production of marzipan. The flavor

and texture of almonds can be intensified or moderated through proper selection of seed combinations belonging to different cultivars, on the basis of variation in origin, moisture content and processing, handling procedures (Kester, 1993), and the content of amygdalin, a cyanogenic glucoside (McCarty *et al.*, 1952; Chandler, 1957; Woodroof, 1979; Frehner *et al.,* 1990) common in several Mediterranean almonds (Dicenta and Garcia, 1993b; Vargas *et al.*, 2001), which gives a singular bitter flavor known as "amaretto".

In Italy there is almost no region that does not have a dessert where almond appears as main ingredient: the *Mandeltorte* of Alto Adige, the *Mandorlato* of Cologna Veneto, the *Torrone* of Cremona, the *Ricciarelli* of Siena, the *Parrozzo* in Abruzzo, the *Diavolacci* in Apulia, the *Frutta Martorana* in Sicily, just to name a few. Moreover in confectionery field, Sicilian almond cultivar, *Pizzuta* e *Fascionello*, commercially known as *Mandorla di Avola*, have always been the soul of another typical sweet of the tradition "*confetti*", produced by Italian companies, among which those of the territory of Sulmona, are renowned but also by French, Belgian, Greek industries (Akparov and Khidirova, 2006; MiPAAF, 2012/2014).

Among the beneficial properties of almond kernel, when incorporated in the diet, there are the possibility to decrease total low density lipoproteins (LDL, bad cholesterol) levels, increasing high density lipoproteins (HDL, good cholesterol) ones (Abbey *et al.,* 1994; Cordeiro and Monteiro, 2001; Hyson *et al.*, 2002; Martins *et al.,* 2003; Moure *et al.,* 2007). Research also shows a connection between regular nut consumption and decreased incidence of coronary heart disease (Dreher *et al.*, 1996), hypertension, congestive heart failure and chronic diseases (Fulgoni *et al.*, 2002; Lovejoy *et al.*, 2002; Pellegrini *et al.,* 2006; Socias i Company *et al.,* 2008).

Apart from its nutritive value, the kernel contains a wide variety of phenolics and flavonoids, showing interesting biological effects such as sedative, anti-inflammatory, anti-hyperlipidemic, anti-tumour and antioxidant

activities (Donovan *et al.*, 1998; Wang *et al.,* 1999a; Sang *et al.*, 2002b; Chen *et al.*, 2005; Milbury *et al.*, 2006; Esfahlan *et al.*, 2010).

Almond seed provides a high quality of extracted mono-unsaturated oil (Abdallah *et al.*, 1998; Kodad *et al.*, 2005) which can constitute over 50% of the kernel dry weight. The oil is composed mostly by stable oleic (70%) and linoleic (15%) fatty acids (García-López *et al.*, 1996) used from ancient times as ointments, emollient and pharmaceuticals.

As shown in **Tab. 4,** the meat is also rich of minerals such as manganese, potassium, calcium, molydenum, phosphorus, iron, magnesium, zinc, copper, selenium and contains good amounts of riboflavin (vitamin B2) and vitamin E (Kester *et al.,* 1991; Kester and Gradziel 1996; Sabate and Haddad, 2001; Chen *et al.*, 2005;). Vitamin E is involved in heart disease and some kinds of cancer prevention and cataract formation (Kodad *et al.,* 2006). Almonds also represent a convenient source of folic acid and fiber (Vezvaei and Jackson, 1996; Schirra, 1997). Futhermore, raw almonds are one of the best plant sources of protein, but free from *gluten* ones and they are one of the popular ingredients in the preparation of gluten-free food formulas, becoming valid healthy alternatives in people with wheat food allergy and celiac disease.

The research by Sathe *et al.* (2001) was interesting because no significantly higher allergy risk has been found in a range of cultivated almond, as well as breeding lines derived from interspecies hybrids, differently from other kind of nuts. So, the beverage industry takes advantage by almond milk production because of its flavor and being an alternative to cow's milk to cope with intolerance and allergies or yet for vegan diet.

Tab. 4- Main nutritional values of almond for edible portion and mineral, vitamin and lipid content

Nutrient	**Unit**	**Value per 100g**
Water	g	4,51
Energy	Kcal	590
Protein	g	21,40
Total lipid (fat)	g	52,52
Carbohydrates, by difference	g	18,67
Fiber, total dietary	g	9,9
Total sugars	g	4,63
Minerals		
Calcium, Ca	mg	236
Iron, Fe	mg	3,28
Magnesium, Mg	mg	268
Phosphorus, P	mg	481
Potassiun, K	mg	659
Sodium, Na	mg	19
Zinc, Zn	mg	2,97
Vitamins		
Thiamin	mg	0,191
Riboflavin	mg	0,711
Niacin	mg	3,500
Vitamin B-6	mg	0,115
Folate, DFE	µg	49
Vitamin A, IU	IU	7
Vitamin E (alpha-tocopherol)	Mg	23,75
Lipids		
Fatty acids, total saturated	g	3,953
Fatty acids, total monounsaturated	g	33,425
Fatty acids, total polyunsaturated	g	12,368
Fatty acids, total trans	g	0.019

Source: USDA National Nutrient Database for Standard Reference Release 27, 2015

2.2.2 Meat brown seedcoat

For many years, almond kernel was considered as the main source of almond production, but recently bibliographic evidences have been reported in literature about the utility and employment of other almond products in several different trades (Takeoka *et al.*, 2000; Sang *et al.*, 2002b; Takeoka and Dao, 2003; Pinelo *et al.*, 2004; Rabinowitz, 2004; Jahanban *et al.,* 2009).

The skin of ripe almonds constitutes about 4% of the fruit (Chen *et al.*, 2005) and when seeds are pealed because aimed to confectionary industry or bakery, large amounts of agricultural by-products remain unused (Frison and Sporns, 2002). An alternative and cheap employ was reported by Harrison and Were (2007), suggesting to feed animals or as a fuel in processing plants, after burning. Moreover, the seedcoat contain a wide variety of available phenolic acids and flavonoids which give a support to antioxidant activity (Frison and Sporns, 2002; Heim *et al.*, 2002; Chen *et al.*, 2005; Wijeratne *et al.*, 2006; Monagas *et al.*, 2007). Many other compounds have been found in the tegument, as triterpenoids, betulinic, oleanoic and ursolic acid and they were described by Takeoka *et al.* (2000) to have anti-inflammatory (Singh *et al.,* 1994), anti-HIV (Kashiwada *et al.*, 1998), and anti-cancer activities (Pisha *et al.*, 1995).

2.2.4 Almond shell

Almond shell is the name given to the highly lignified material (30–38% of the dry weight), forming the thick endocarp of the almond (Martinez *et al.*, 1995). When it is processed to obtain the edible seed, big ligneous fragments are separated, and being not useful, hence, they are normally incinerated or dumped without control (Urrestarazu *et al.*, 2005).

Many researchers have reported the high xylan content of almond shells, that could become a suitable substrate for the production of xylose (Pou-Ilinas *et al.*, 1990), furfuraldehyde (Quesada *et al.*, 2002) or for fractionation

into cellulose, pentosans and lignin (Martinez *et al.*, 1995), producing oxyaromatics, an useful compound for health, cosmetics and food production (Quesada *et al.*, 2002). An antioxidant activity has been observed also in almond shells (Pinelo *et al.*, 2004; Jahanban *et al.,* 2009) as in the whole almond fruit.

Other discoveries about almond shell, as potential applications, have been developed and reported in literature (Bansode *et al.,* 2003; Daifullah and Girgis, 2003; Estevinho *et al.,* 2006), showing the heavy metal low-cost absorbent power of endocarp against the pollution of surface and ground waters (Bulut and Tez, 2003); or again, the absorption of harmful dyes, produced by the textile industry dangerous for the environment, was studied by using shell to reduce costs and obtaining the maximum efficiency (Ardejani *et al.*, 2007). Furthermore, almond shell used (100% pure) as growing media can be more ecological-friendly and less expensive substitute than the traditional rockwool for soilless vegetable production (Urrestarazu *et al.*, 2005).

2.2.5 Green shell mesocarp (hull)

In the past, almond hulls, a by-product of the almond industry, were removed from almonds after harvesting and used as supplemental livestock feed. Recently there is interest in using dry almond hulls as a natural source for sweetener concentrate and dietary fibre (Takeoka and Dao, 2003) or as food, food additives, pharmaceuticals and roughage or cat litter (Rabinowitz, 1991; 2002; 2004).

The mesocarp in senescence remains stable retaining its high sugars, flavonoids, lignin, fibre, constructing of cellulose, hemicellulose, triterpenoids, pectins, tannin-like complex, polyphenols and ash (Sang *et al.,* 2002b; Takeoka *et al.,* 2000).

Comparing almond to its botanical relatives, it has an unusual high concentration of flavonoids in the hull, a reason could be the protective role of

them because of the long exposure time of the mesocarp to several biotic and abiotic stresses as the intense heat, ultraviolet radiation and pest infestation. Recent researchs have been performed with the aim to investigate the characterization and identification of almond hull phenolic compounds, using them as natural antioxidants and antiradicals in foods and oxidative damage (Takeoka *et al.*, 2000; Takeoka and Dao, 2003; Sang *et al.*, 2002a;b;c; Rabinowitz, 2004; Jahanban *et al.*, 2009).

3. Cyanogenic glucosides and bitterness in almond

3.1 Cyanogenic glucosides and the key role of β-glucosidases

This thesis has been performed studying the role of the *Sk* gene involved in the control of bitterness in almond. The trait is determined by the presence and content of amygdalin which is a cyanogenic glucoside, this is the reason for which in the next pages important information is reported about cyanogenic compounds.

More than 3.000 plant species from all major vascular plant taxa are cyanogenic, belonging to over 130 families of flowering plants and representing 11% of the total number of plant species tested, surely a overestimated amount of cultivated plants (Gleadow and Møller, 2014). Indeed this value is higher the 5% found in screens of natural population systems (Gleadow *et al.,* 2008, Miller *et al.,* 2006) and it could be due to humans unconscious selection of cyanogenic plants for cultivation, perhaps because of their increased resistance to herbivores or because the need to process the plant material before consumption makes them less attractive to trespassers (Jones, 1998, McKey *et al.,* 2010).

Cyanogenic glucosides (CGs) are bioactive plant products found in the oldest of terrestrial plants (the ferns) and in gymnosperms and angiosperms (Zagrobelny and Møller, 2011; Bak *et al*., 2006). They are also called

secondary metabolites or specialized plant products, classified as phytoanticipins and may be defined chemically as glycosides of α–hydroxynitriles (cyanohydrins) and stabilized by glucosylation. They are synthetized from aromatic and aliphatic amino acids (tyrosine, phenylalanine, valine, isoleucine, and leucine) (Bjarnholt *et al*., 2008) and a few nonproteinogenic amino acids [2S,1'R and 2S,1'S epimers of 2-(2'-cyclopentenyl)-glycine as well as (2S,1'S,2'S)-2-(2'-hydroxy- 3'-cyclopentenyl) glycine], (Møller and Conn, 1979; Sibbesen *et al*., 1995; Bak *et al*., 1998; Jones *et al*., 1999; Bjarnholt *et al*., 2008; Clausen *et al.,*2002; Conn, 1980; Møller, 2010; Selmar, 2010).

The most common cyanides (CNs) are: the monoglucosides dhurrin [(*S*)-4-hydroxymandelonitrile-β-D-glucopyranoside], linamarin (2 hydroxyisobutyronitrile-β-D-glucopyranoside), lotaustralin [(*R*)-2-hydroxy-2-methyl-butyronitrile-β-D-glucopyranoside] which derive respectively from tyrosine, phenylalanine, valine, and isoleucine and prunasin [(*R*)-mandelonitrile-β-Dglucopyranoside]. Prunasin is an epimer derived by phenylalanine and amygdalin is a prunasin-derived diglucoside. The sugar residue are glucose, xylose and apiose, the latter of which is acylated by a cinnamic acid residue (**Fig. 9**).

Besides, the number of CNs found in nature is further expanded by the structural diversity of the sugar moiety. In the CNs currently known, the cyanohydrin is always stabilized by a β- glucosidic bond to D-glucose. In cyanogenic diglycosides, the second sugar moiety may also be a D-glucose residue, with the most common examples being amygdalin (prunasin-6'-glucoside) and linustatin (linamarin-6'-glucoside), in both of which the second glucose residue is attached by a β-1,6 linkage.

The main function of CGs in plants is to perform an immediate chemical defense when a plant tissue is broken by herbivores and pathogens. In addition, previous studies have shown that CGs play also an important role in the

primary metabolism of plants as transporters of nitrogen and glucose (Selmar *et al*., 1988). In some plant species, they are even implicated in the control of germination (Flematti *et al.,* 2013; Richmond and Ghisalberti, 1994) and bud burst via cyanohydrin and hydrogen cyanide (HCN) formation (Barros *et al.,* 2012), or in transport of forms of carbon and nitrogen (Selmar *et al.,* 1988), and endogenous turnover processes which may release the nitrogen from CNs in the form of ammonia (Jenrich *et al.,* 2007; Neilson *et al.,* 2013; Piotrwski, 2008). More recently, it has been suggested that CNs have a role in modulation of oxidative stress (Kadow *et al.,* 2012; Møller, 2010; Neilson *et al.,* 2013).

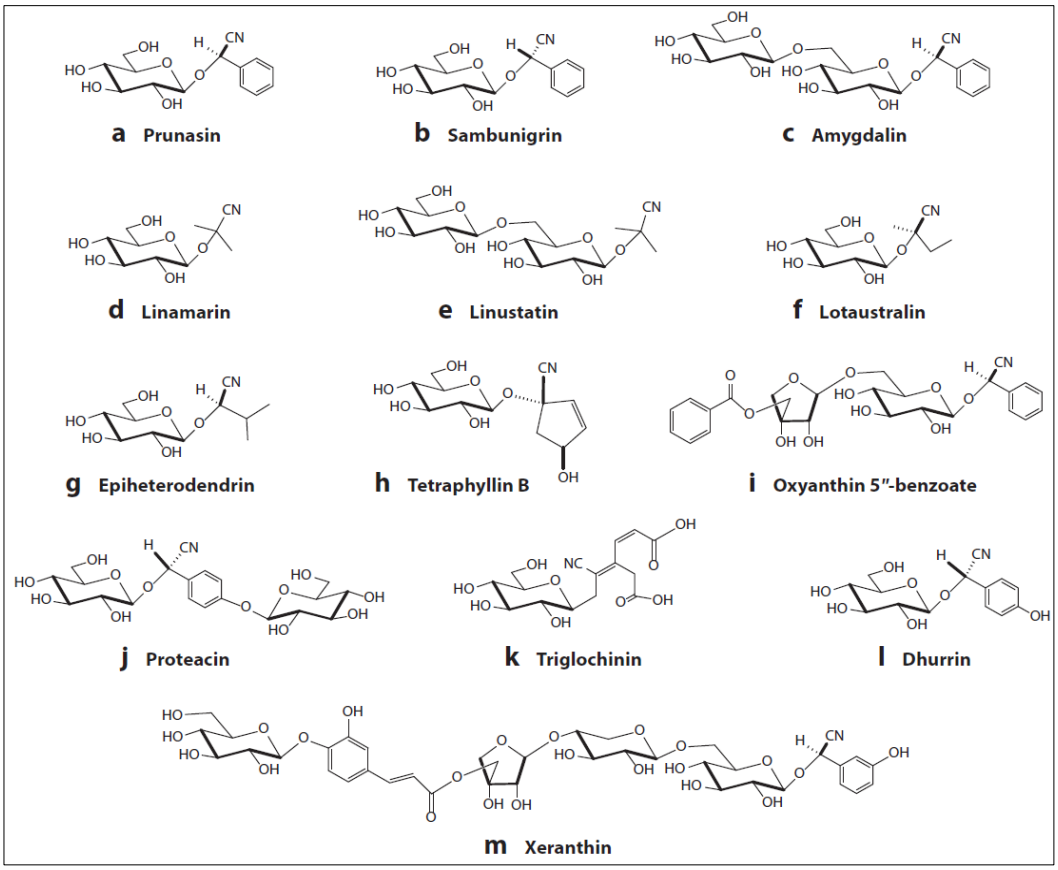


Fig. 9- The structural diversity of CGs. Source: Gleadow and Møller, (2014)

Concerning plant defense, upon crushing, the cellular structure of plant tissue containing CGs, the enzyme β-glucosidases, catalyzes the hydrolysis of the β-glycosidic bond in b-glycosides composed of carbohydrate moieties, or

of a carbohydrate moiety linked to an aryl or alkyl aglycone (Morant *et al*., 2008).

β-glucosidases are among the oldest classes of enzymes known, belonging to the Family-1 b-glycosidases, (Morant *et al*., 2008). They are also involved in plant metabolism, like in cell degradation, in endosperm growth, during the germination (Leah *et al*., 1995; Suda and Giorgini, 2003), activation of phytohormones (Kristoffersen *et al*., 2000; Kiran *et al*., 2006; Lee *et al*., 2013) or in lignifications (Dharmawardhana *et al*., 1995; Escamilla-Trevino *et al*., 2006).

Furthermore, β-glycosidases play important roles in the food and beverage industries, such as tea, wine and fruit juice production, despite of their relatively high stability, (Mizutani *et al*., 2002; Fia *et al*., 2005; Maicas and Mateo, 2005).

In particular, β-glucosidases allow the dissociation of the labile cyanohydrin that is formed by broken tissues, releasing toxic HCN, glucose and a kethone or an aldeyde in a process known as cyanogenesis. CGs are synthesized from specific amino acids in a series of reactions catalyzed by two multifunctional, membrane-bound cytochrome P450s (P450aa and P450ox) and a soluble UDP-glucosyl-transferase, with an oxime and an α-hydroxynitrile (cyanohydrin) as key intermediates. Cyanogenesis occurs when the β-glucosdic linkage is hydrolyzed by a specific β-glycosidase to form an unstable α-hydroxynitrile that dissociates into HCN and a ketone either spontaneously at high pH or catalyzed by an α-hydroxynitrilase (Morant *et al.,* 2008; Conn, 1980; Poulton, 1990; Møller and Seigler, 1999; Morant *et al*., 2003; Bak *et al*., 2006) (**Fig. 10**).

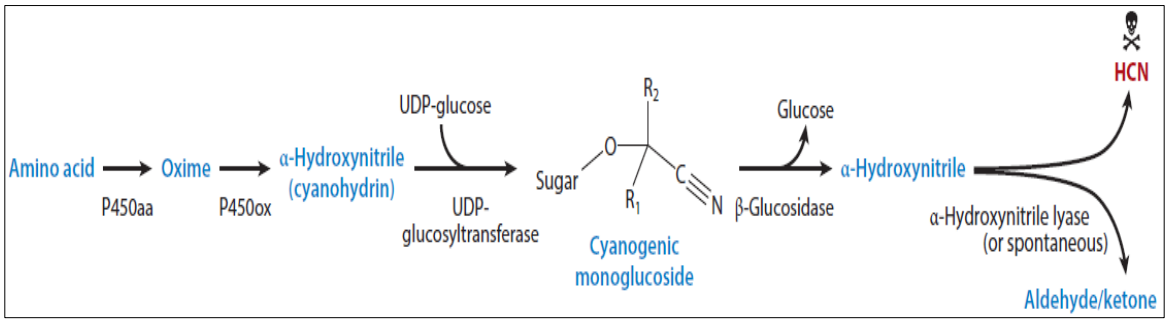


Fig. 10- Biosynthesis and bioactivation of CGs. Source: Gleadow and Møller, (2014)

All plants produce tiny amounts of HCN as an additional product in the biosynthesis of ethylene, but some plant species can release large amounts from the endogenously stored CNs, and most of them are employed widely in human diet, representing a danger for humans and animals. Indeed, HCN is really toxic because of the inhibition of the activity of metalloenzymes, principally cytochrome *c* oxidase, the final enzyme involved in the respiratory electron transport chain (Nelson, 2006; Leavesley *et al.,* 2008). Acute or chronic exposure to HCN can lead to intoxication, mild to severe illness and, in extreme cases, even death in humans (Loyd and Gray, 1970) and animals (Finnie *et al.,* 2011).

Furthermore, hydrolysis of CNs results in the concomitant release of carbonyl compounds that may further increase the toxic and repellant effects of HCN (Peterson *et al.,* 1987). So, CNs are effective deterrents to herbivores (Ballhor *et al.,* 2008a and b; Gleadow and Woodrow, 2002a and b; Tattersall *et al.,* 2001; Zagrobelny *et al.,* 2004) and this is most likely the main evolutionary driver in their occurrence across the plant kingdom (Agrawal *et al.,* 2012; Bak *et al.,* 2006; McKey *et al.,* 2010). Nevertheless, several specialist herbivores not only tolerate CNs, but sequester them for defense purposes together with other compounds against predators (Nishida, 2002; Peterson, 1986; Schappert and Shore, 1999a and b; Zagrobelny *et al.,* 2007a and b).

Although the CGs cyanogenesis is dangerous for living being, CGs compartmentalization is essential to prevent auto-toxicity of plants and to ensure that the HCN released is for a targeted response, e.g., to herbivore attack (Morant *et al.,* 2008). The most common kind of compartmentalization is at the level of the organelle. CGs, at least in leaves, are typically confined to the vacuole of the plant tissues (Gruhnert *et al*., 1994; Conn, 1991; Kadow *et al.,* 2012; Saunders and Conn, 1978), as illustrated in **Fig. 11**.

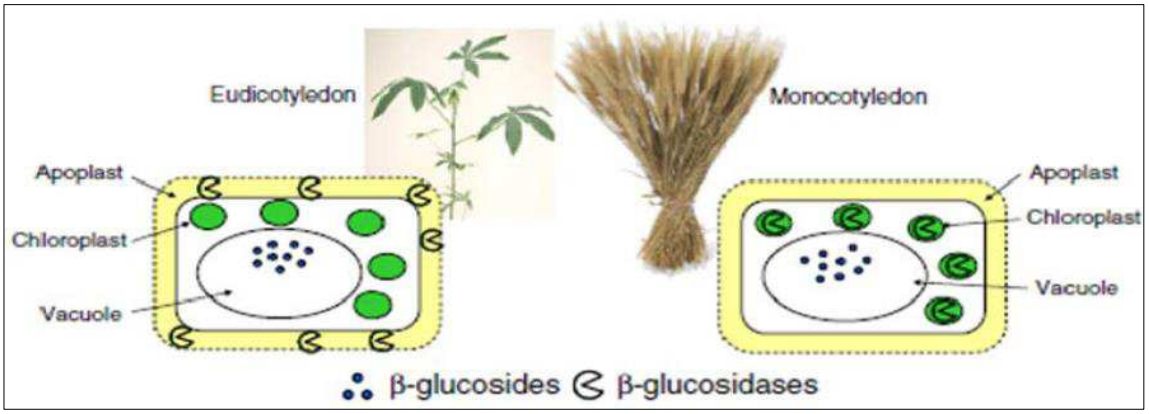


Fig. 11- Compartimentalization of β-glucosidases and their substrates in eudicotyledenous and monocotyledonous plants (Morant *et al.,* 2008)

Contrariwise, the localization of β-glycosidases involved in hydrolysis of CNs is more variable (apoplastic space, bound to the cell wall, cytoplasm, small vesicles, or chloroplasts) depending by species (Elias *et al*., 1997; Frehner and Conn, 1987; Kakes, 1985; Kogjima *et al*., 1979; Sanchéz- Perez *et al*., 2012; Thayer and Conn, 1981) and if plants are monocotyledonous or eudicotyledenous, where are stored respectively in plastids, and apoplastically or intracellularly in protein bodies (Morant *et al*., 2008).

Separation of CNs and their degradative enzymes may also occur at the tissue level as like as in sorghum, in which dhurrin is found almost exclusively in the epidermal cell layer of the leaf blade, whereas dhurrinase and α-hydroxynitrilase are located predominantly in the mesophyll cells (Kogjima *et al*., 1979; Saunders and Conn, 1978; Thayer and Conn, 1981). Some plants

concentrate the CNs in seed like *H. brasiliensis* (Selmar *et al*., 1988), *Prunus serotina* (Swain *et al.*, 1992a and b), and *Canthium schimperianum* (Schwarzt *et al*., 1996); in others, the cyanogenic component is the fruit itself as in *Passiflora edulis* (Chassagne *et al*., 1996). Several species have only trace amounts of CNs as in the dry seeds sorghum (Loyd and Gray, 1970) and clover (Hughes, 1991; Vickery *et al*., 1987), whereas others have completely a-cyanogenic seeds, e.g. *E. cladocalyx* (Gleadow and Woodrow, 2000 a and b).

Prunasin and amygdalin are instead located in the parenchyma of the cotyledons in the seeds of rosaceous stone fruits, whereas the β-glucosidase and α-hydroxynitrilase in the procambium (Poulton and Li, 1994; Sanchèz - Perez *et al*., 2012; Swain *et al*., 1992 a and b).

Actually, there are many genetic studies about β-glycosidases and those responsible of hydrolysis of CNs were isolated, identifying also the genes in several species. Sorghum contains two β-glucosidases, the dhurrinases Dhr1 and Dhr2, which hydrolyze dhurrin with high specificity and organ-specific expression (Cicek and Esen, 1998). White clover contains a single linamarase encoded by the *Li locus* (Olsen *et al*., 2007). *L. japonicus* contains the two paralogous β-glucosidases BGD2 and BGD4.

3.2 Byosinthesis and turnover

3.2.1 Byosinthesis

The biosynthetic pathway for CNs was first elucidated in *Sorghum bicolor* (McFarlane *et al*., 1975; Møller and Conn, 1979; Shimada and Conn, 1977), where the CGs dhurrin is synthesized from L-tyrosine in a series of steps catalyzed by a combination of: two enzymes CYP79A1 and CYP71E1, the multifunctional cytochrome P450 (Bak *et al.,*1998; Kahn *et al.,* 1997; Kahn *et al.*, 1999; Koch *et al,*, 1995; Sibbesen *et al.,* 1995) a family 1 UDP-glucosyltransferase (UGT85B1) (Jones *et al.*, 1999) and a P450 redox partner

NADPH-dependent cytochrome P450 oxidoreductase (POR) (Halkier and Møller, 1991) (**Fig. 12**). The enzymes, catalyzing dhurrin formation in sorghum (CYP79A1, CYP71E1, UGT85B1 and POR), support the formation of a metabolon to promote rapid channeling of the toxic and labile intermediates into dhurrin formation and preventing undesired metabolic crosstalk (Jensen *et al.,* 2011; Jørgensen *et al.,* 2005; Kristensen *et al.,* 2005; Nielsen *et al.,* 2008).

The types of intermediates identified in the CGs pathway in sorghum have also been found in other species, such as flax (*Linum usitatissimum*) (Cutler and Conn, 1981), white clover (*T. repens*) (Hughes, 1991), cassava (*Manihot esculenta*) (Andersen *et al.,* 2000), *L. japonicus* (Forslund *et al.,* 2004), almond (*Prunus amygdalus*) (Sanchèz-Perez *et al.,* 2008) and *Triglochin maritime* (Nielsen and Møller, 1999). In addition to sorghum, the genes encoding the entire pathway for CGs synthesis have now been identified in cassava (Andersen *et al.,* 2000; Jørgensen *et al.*, 2011; Kannangara *et al.,*, 2011) and *L. japonicus* (Takos *et al.,*, 2010). Genes encoding CYP79s and UGTs involved in CGs synthesis have also been identified in white clover (Oldsen *et al.,*, 2008) and almond (Franks *et al.,*, 2008), respectively. Thus, the key intermediates involved in biosynthesis of CNs appear to be the same across the plant kingdom.

The CYP79s are considered the signature enzyme of the CGs pathway and all tested current members of this P450 family catalyze the conversion of amino acids to the corresponding oximes. Both CYP79 and the second P450 in the pathway, catalyzing the conversion of oxime to cyanohydrin, belong to the CYP71 clan of P450s (Bak *et al.*, 2011). The latter P450 and the UGT that catalyzes the final step in CGs synthesis have a much broader substrate specificity compared with that of CYP79 (Hansen *et al.*, 2003; Jones *et al.*, 1999; Kahn *et al.*, 1999).

3.2.2 Turnover

Plants containing large amounts of CGs have high amounts of nitrogen and sugar which could represent substantial percentages of the total organic matter present (Neilson *et al.*, 2013). The ability of plants to reuse the carbon and nitrogen deposited in CNs (e.g., to balance resource demands in primary metabolism) offsets some of the energy and resource costs associated with their synthesis and storage (Neilson *et al.*, 2013). After the classical catabolic pathway by means of β-glycosidase action, part of HCN may be sorted for primary metabolism in a reaction catalyzed by β-cyanoalanine synthase that involves stoichiometric consumption of cysteine (Blumenthal *et al.*, 1968; Piotrowski and Volmer, 2006). Subsequent action of nitrilase 4 (NIT4) family enzymes results in asparagine and aspartate formation (Jenrich *et al.*, 2007).

Recently, a different pathway, independent by crushing of the tissues, was discovered in sorghum. The intermediate *p*-Hydroxyphenylacetonitrile compound belonging to this alternative way, is converted into *p*-hydroxyphenylacetic acid by the action of the nitrilase heteromer NIT4A/NIT4B2s with a concomitant release of ammonia (Jenrich *et al.*, 2007) (**Fig. 12**).

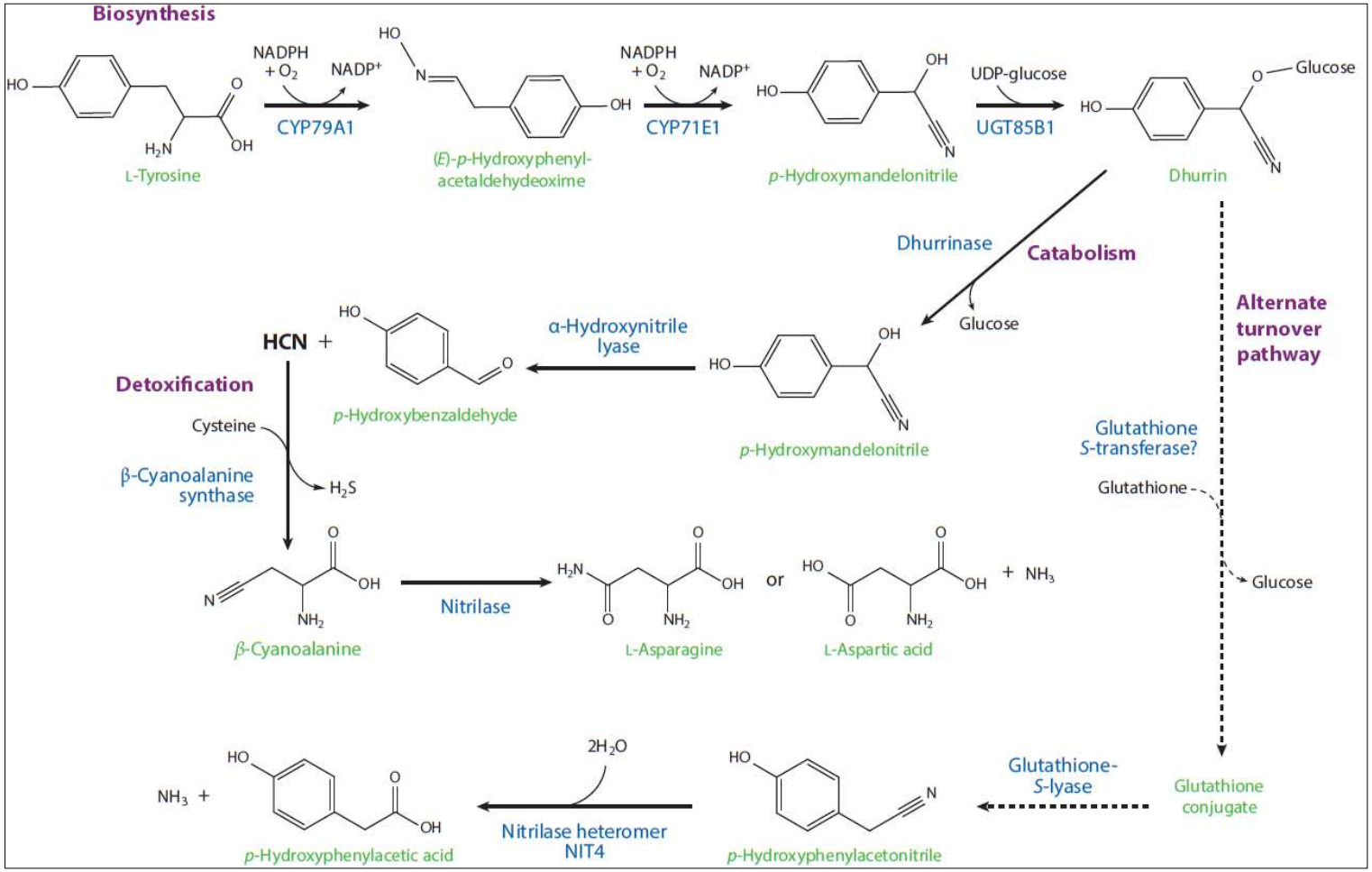


Fig. 12- Synthesis and turnover of dhurrin in *Sorgum bicolor*. Source: Gleadow and Møller, (2014)

3.3 Bitterness in almond: the cyanogenic di-glycoside amygdalin, biosynthesis and catabolism

The peculiar bitter taste in certain almond kernels is due to the accumulation of a secondary metabolite and cyanogenic diglucoside amygdalin. This compound is widespread in seeds of other members of the *Rosaceae*, like apple (*Malus* spp.), peach (*Prunus persica*), apricot (*Prunus armeniaca*), black cherry (*Prunus serotina*) and plum (*Prunus* spp.), (McCarty *et al*., 1952; Chandler, 1957; Woodroof, 1979;; Conn, 1980; Frehner *et al*., 1990; Swain *et al.*, 1992a and b; Poulton and Li, 1994; Møller and Seigler, 1991; Arrázola, 2002; Dicenta *et al*., 2002; Mizutani *et al*., 2002Franks *et al.,* 2008 a and b; Sánchez-Pérez *et al*., 2008; Wirthensohn *et al*., 2008).

Amygdalin was the first cyanogenic diglucoside to be isolated. Indeed, the first known detection of HCN liberated from damaged plant tissue was made in 1802 by the pharmacist Bohm in Berlin upon distillation of bitter almonds. In 1830, Robiquet and Boutron-Chalard discovered the chemical structure of the HCN-liberating compound in bitter almonds (Lechtenberg and Nahrstedt, 1999). Because the compound was isolated from *Prunus amygdalus* (synonym *Prunus dulcis*) it was named amygdalin.

Amygdalin is a diglucoside of R-mandelonitrile, composed by two glucoses with a β orientation and joined by a 1,6 bond (**Fig. 13**). Besides, the cyanogenic monoglucosyde prunasin is related to amygdalin having the same structure, but with only one glucose and being its precursor in the biosynthetic pathway (Frehner *et al*., 1990) (**Fig. 13**).

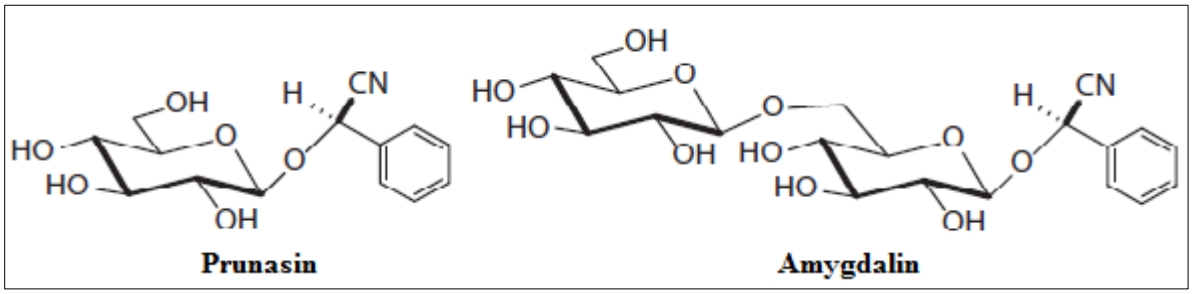


Fig. 13- Prunasin and amygdalin chemical structures. Source: Gleadow and Møller, (2014)

Since 1999, Lechtenberg and Nahrstedt shown that wild almonds contain the cyanogenic glucoside amygdalin in their seeds and fruit pulp. Although prunasin and amygdalin are present in the vegetative (roots, leaves, kernels of sweet, bitter and slightly bitter varietes) and floral parts of almond (Dicenta *et al*. 2002; London-Shafir *et al*. 2003), in bitter types the kernel is the principal site of amygdalin accumulation. On the other hand, Sánchez-Pérez *et al*., (2009; 2010) detected prunasin in leaf, petiole, endosperm and cotyledon in sweet and bitter almond varieties; amygdalin instead, was found in the fruit tissues of the bitter genotype with a maximum quantity during the ripening, in particular in nucellus and endosperm.

Assuming, as reported by Dicenta *et al*., (2000), that a phenotype of the kernel is determined by the genotype of the tree, the control of amygdalin accumulation in seeds, could occur in tissues of the plant which show the same genotype of the tree. Therefore, kernel seed coat is a good candidate site for the control of amygdalin accumulation, as it derived from ovule in teguments (Haughn and Chaudhury, 2005).

Metabolic and genetic studies of sweet and bitter almonds show that a seed coat UDP-glucosyltransferase (GT1) is associated with bitterness (Franks *et al*. 2008). Furthermore, distinct cellular localisations of the enzymes involved in the degradation pathway, possibly involving a seed coat prunasin hydrolase, have also been suggested to be related to bitterness in almond (Sánchez-Pérez *et al*. 2008). More specifically, some authors have suggested that the difference between sweet and bitter almonds could derive from a different turnover of prunasin and this hypothesis was supported by several tests showing a higher activity of beta-glucosidase detected in the seed coat of sweet almonds (Sanchez-Perez *et al*., 2008; 2012).

3.3.1 Biosynthesis

The biosynthetic pathway of amydalin is described in **Fig. 14**: two multifunctional cytochrome P450 monooxigenases, CYP79 and CYP71, located on the membrane, convert phenylalanine to mandelonitrile. This latter is converted into prunasin after

the addition of glucosidic residue by a glucosil-transferase (GT1) and UDPG-mandelonitrile (5'-uridin difosfoglucose) glucosyltransferase. Prunasin is then converted in amygdalin by adding a second glucose molecule by a GT2 and UDPG-prunasin glucosyltransferase.

3.3.2 Catabolism

The genes and proteins involved in amygdalin degradation have been also identified (**Fig. 14**). Upon crushing, in bitter almond kernels, amygdalin is degraded in benzaldehyde, HCN, with releasing of glucose. Three enzymes are involved in this first step: the beta-glucosidases amygdalin hydrolase (AH), prunasin hydrolase (PH) and mandelonitrile lyase 1 (MDL1), (Evreinoff, 1952; Mentzer and Favre-Bonvin, 1961; Frehner *et al*., 1990; Swain and Poulton, 1994b; Kester and Gradziel, 1996; Dicenta *et al*., 2002; Sánchez-Pérez *et al*., 2008; Wirthensohn *et al*., 2008).

The conversion catalyzed by MDL1 leading benzaldeyde and HCN may also proceed in non-enzymatic via at neutral or alkaline pH. Indeed, HCN (being an inhibitor of cell respiration) could be detoxified by the action of the enzyme β-cyano-ala synthase, which converts HCN and Cys into β -cyano-ala (Floss *et al*., 1965). By the action of nitrilases, β -cyano-ala is converted into aspartic acid and asparagine (Swain and Poulton, 1994 a and b; Piotrowski *et al*., 2001; Piotrowski and Volmer, 2006; Jenrich *et al*., 2007; Kriechbaumer *et al*., 2007).

Since the ancient Egypt, traitorous priests in Memphis and Thebes were poisoned to death with pits of peaches (Davis, 1991). Even though, benzaldehyde and HCN are olfactorily similar and responsible for the bitter taste (the "cherry" or "amaretto" flavor) of kernels, HCN is highly toxic (Merck Co. Inc., 2001). Benzaldehyde is also known in the chemical and flavoring industries as "oil of bitter almond" because of its preponderance in bitter rather than swe*et* almonds. Bitter almonds contain 3 – 5% of amygdalin, taste very bitter, and develop a characteristic cyanide aroma with moisture (Rumsey, 1927). Sweet almonds have a slightly nutty fragrance and taste, while semi-bitter almonds have a "marzipan-like" taste. However, a chemical basis to

distinguish this phenotype, and the semi-bitter one, from the bitter flavor has not been determined.

Recently Wirthensohn *et al*., (2008) demonstrated that amygdalin is not only the one to give kernel flavor, confirming the presence of benzyl alcohol and identifying the compound 2,3- butanediol by using chemical sensor analysis.

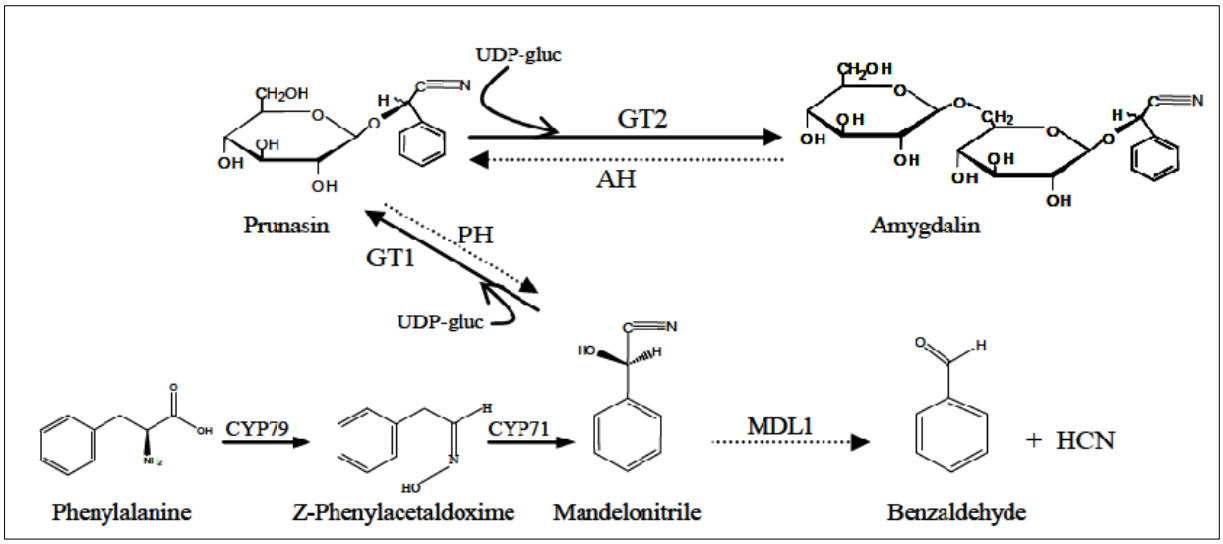


Fig. 14- The metabolic pathways for biosynthesis and catabolism of the cyanogenic glucosides prunasin and amygdalin in almonds. Biosynthetic enzymes (black lines): CYP79 and CYP7; GT1, UDPG-mandelonitrile glucosyltransferase; GT2, UDPG-prunasin glucosyltransferase; Catabolic enzymes (dashed lines): AH, amygdalin hydrolase; PH, prunasin hydrolase; MDL1, mandelonitrile lyase. Source: Sánchez-Pérez *et al*., (2008)

4. The Sk gene

Many important agronomic traits in almond seem to be controlled by major genes, **Fig. 15** (Dirlewanger *et al*., 2004b).

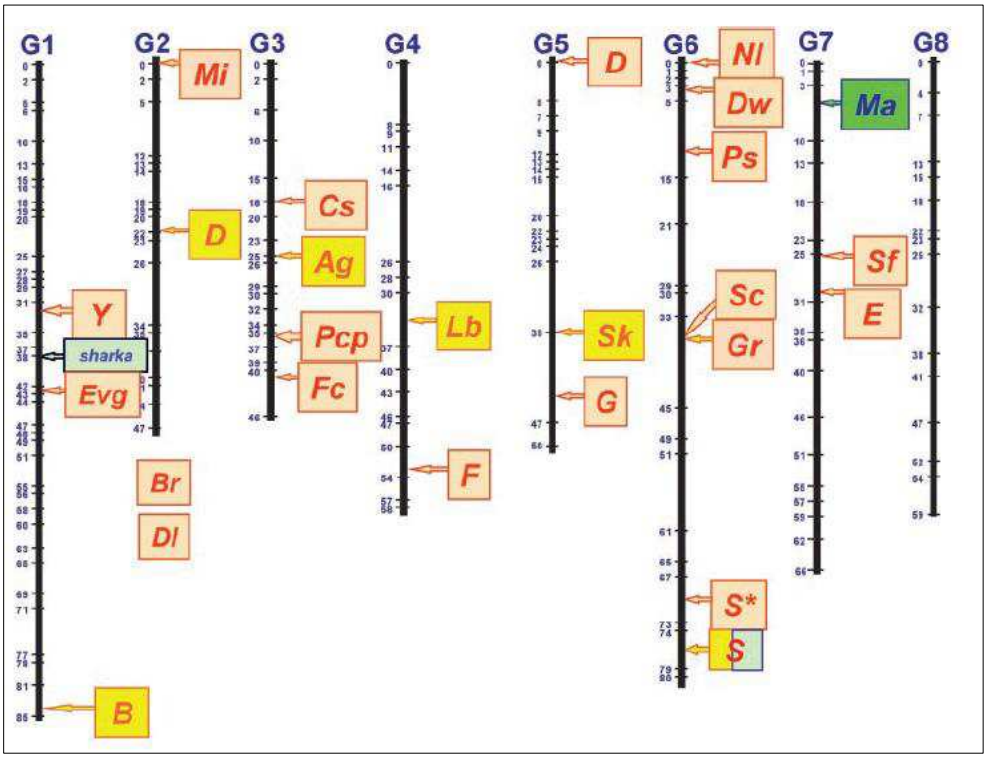


Fig. 15- Approximate position of 28 major genes ([1]) mapped in different populations of apricot (Blue background), peach (orange), almond or almond x peach (yellow), and Myrobalan plum (green) on the framework of *The Prunus Reference Map* (Joobeur *et al*., 1998 and Aranzana *et al*., 2003). Source: Dirlewanger *et al*., 2004b

The kernel taste is an important feature studied by almond growers and breeders, as the selection of sweet cultivars is a prioritary goal. In literature is reported that bitterness is a chemical defense of wild almond, protecting the kernels from herbivores, insects and pathogens feeding (Conn, 1969; Nahrstedt, 1985; Jones, 1988; Morant *et al*., 2003; Zagrobelny *et al*., 2004, 2007a, 2007b). In addition to this role,

[1] Gene abbreviations: Y, peach flesh color; sharka, plum pox virus resistance; B, flower color in almond x peach; Mi, nematode resistance for peach; D, almond shell hardness; Br, broomy plant habit; DI, double flower; Cs, fesh color around the stone; Ag, anther color; Pcp, polycarpel; Fc, flower color; Lb, blooming date; F, flesh adherence to stone; D, non-acid fruit in peach, Sk, bitter kernel; G, fruit skin pubescence; NI, leaf shape; Dw, dwarf plant; Ps, male sterility; Sc, fruit skin color; Gr, leaf color; S*, fruit shape; S, self-incompatibility (almond and apricot); Ma, nematode resistance from Myrobalan plum; Sf, resistance to powdery mildew; E, leaf gland shape;. Genes DI and Br are located on an unknown position of G2.

accumulation of cyanogenic glucosides in certain angiosperm seeds may provide a storage deposit of reduced nitrogen and sugar for the developing seedlings (Lieberei *et al*., 1985; Selmar *et al*., 1988, 1990; Swain *et al*., 1992 a and b).

In 1974 Spiegel-Roy and Kochba suggested a complex kind of inheritance with three genes involved. It is a monogenic trait with simple inheritance. Despite what has been said so far. However, today it is generally accepted the sweet kernel taste allele is dominant over the bitter one (recessive) as the result of several studies (Heppner, 1923, 1926; Kester *et al*. 1977; El Gharbi, 1981; Grassely and Crossa-Raynaud, 1983;Vargas *et al*., 1984; 2001; Dicenta and García, 1993a; Dicenta *et al*., 2007). Dicenta *et al*. (2001), by studying a high number of families, classified several sweet cultivars as heterozygous or homozygous and found that most almond cultivars are heterozygous.

The *Sk* gene is located on the linkage group five (G5) of almond genome as reported in two almond genetic linkage maps (Joobeur *et al*., 1998; Bliss *et al*., 2002; Sánchez-Pérez *et al*., 2010).

During the domestication, humans have selected sweet almond trees for cultivation (Ladizinsky, 1999). So, sweet kernels is originated in almond by a mutation and it led to the domestication. As a consequence of grower selection, a decrease on the frequency of alleles responsible for the bitter taste in this species has occurred (Browicz and Zohary, 1996). Then, most European (and New World) cultivated almonds are heterozygous for bitterness owing to its discouragement of foliar feeding, so that many open pollinated seed derived local landraces typically segregate for bitter kernels (Grasselly, 1972).

Although in recent studies molecular markers linked to the *Sk locus* have been found (Sánchez-Pérez *et al*., 2010), its exact localization on linkage group 5 (G5) and its biochemical function is elusive. The **Fig. 16** reported the genetic map of G5 obtained using a F_1 population derived by crossing the two parental lines R1000 (R) and Desmayo Largueta (D) both heterozygous for the *Sk locus*.

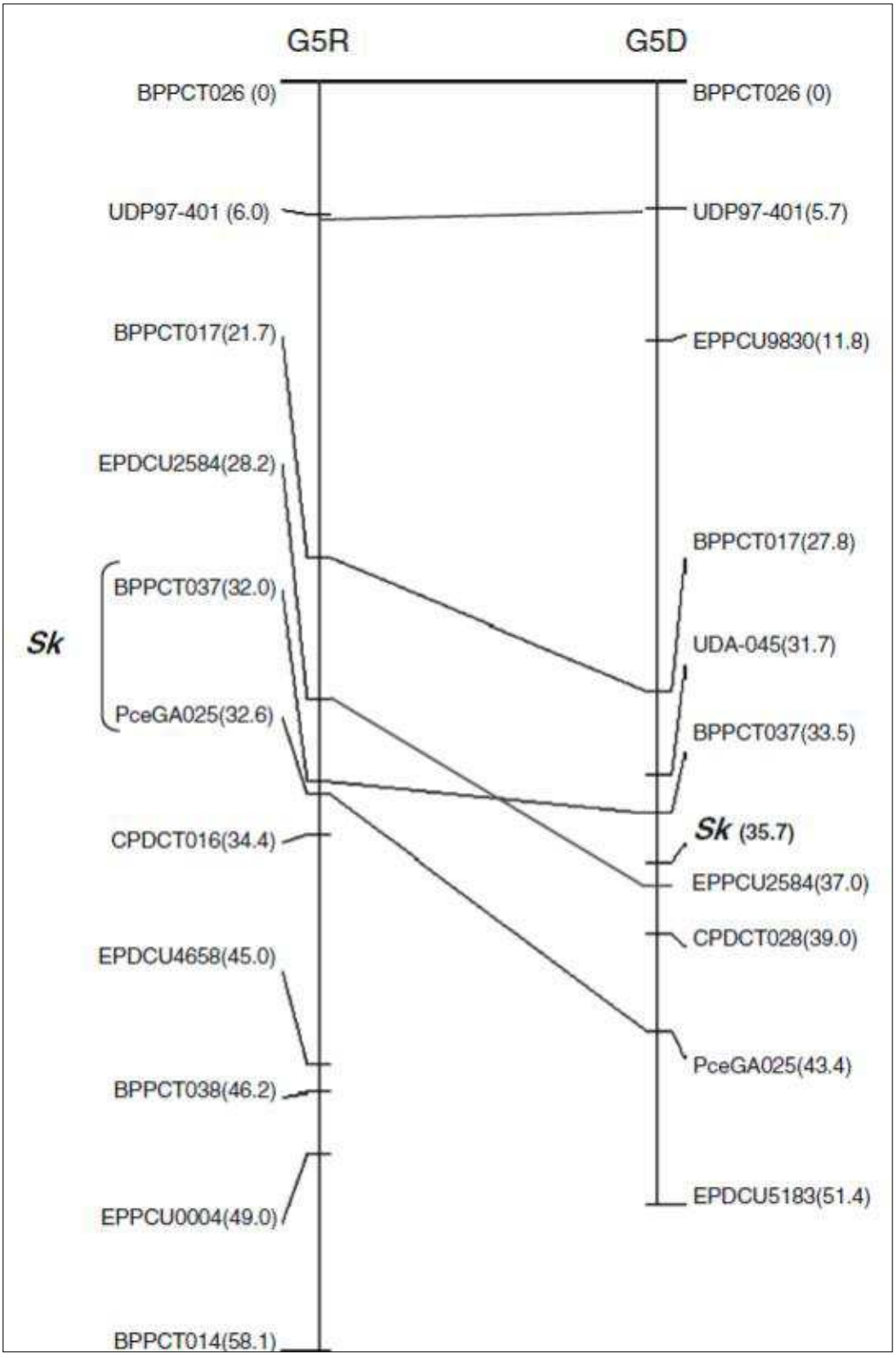


Fig. 16- Genetic map of almond group 5 around the kernel taste locus *Sk/sk* (G5R and G5D) of the progeny RxD. In parentheses genetic distances in centimorgans (cM) are reported. Source: Sánchez-Pérez *et al*., (2010)

Nowadays, recent inheritance studies have provided evidence that more than one gene is involved in bitterness in other *Prunus* species, such as apricot, where five unlinked genes involved in two distinct biochemical pathways (three in the biosynthesis of CGs or their transport and two involved in the breakdown of CGs) have been associated with bitterness (Negri *et al*., 2008). In contrast, genetic studies in peach have proposed a mono-factorial inheritance, with bitterness dominant to sweetness in kernels (Werner and Creller, 1997).

In the case of almond, the determination of sweetness or bitterness depends on the genetic background of the sporophyte. Indeed, the maternal genotype is responsible for the taste of the seed, therefore all fruits of an almond tree will have either sweet or bitter kernels (Kester and Asay 1975), showing the influence of both parents in the following generation (Dicenta *et al*., 2000). Anyway, the same authors reported a different detection of slightly bitter seeds in different years, suggesting that this trait could be independent from the sweet or bitter character and other genes could be involved in the determination of it, as like as in other *Prunus* species, i.e. apricot (Dicenta *et al*., 2000; Negri et al., 2008).

An approach to map candidate genes involved in the metabolic and catabolic pathways of amygdalin and encoding for three glucosyltransferases, one prunasin hydrolase and one amygdalin hydrolase, has been carried out by Sánchez-Pérez *et al*., (2010). However, no correlation between prunasin hydrolase and bitterness has been found because the above-mentioned genes are located on other linkage groups. The aim to identify candidate genes that could be involved in amygdalin accumulation, a further comparative genomic analysis has been performed in Prunoideae (Koepke *et al*., 2013).

5. Genetic improvement and molecular approach

5.1 Genetic improvement goals

The main aim of plant breeding is to develop successful cultivars which must provide an improvement with respect to the cultivar to replace. So, the new cultivar

will be at least as good as the previous one, adapted to the environment, with the same or better quality, disease/pest resistant and appreciated by market.

The almond genome is relatively small, of approximately 294 million base pairs (Mbp), (Baird *et al.*, 1994; Arumuganathan and Earle, 1991), with sixteen ($2n = 2x = 16$) small, but distinguishable chromosomes (Corredor *et al.*, 2004; Martınez-Gomez *et al.*, 2005; Yousefzadeh *et al.*, 2010).

Almond cultivars have superior fitness among tree crops, showing adaptation to a wide range of environmental conditions (Gradziel, 2012). Authors as Socias i Company and Felipe (1988), Dicenta and Garcia (1993a), Ortega and Dicenta (2003) found a relation between the higher genetic heterozygosity, typical of almond cultivars, and the versatility in adaptation, related to different environments.

Several centuries of selection occurred to obtain materials with a broad environmental adaptability, identifying rare, elite selections, where the maximum potential of additive, dominance, epistatic, genomic and epigenetic interactions were combined. Clonal propagation allowed to store these genotypes for further and future studies on the genetic variation of almond. Recently, i.e. the extensive genetic variability found in the Sardinian and Apulian almond germplasm has indicated that these materials represent an important source of genes for the improvement of the crop (Rigoldi *et al.*, 2015).

Nowadays, a direct accessibility to a rich germplasm of almond is possible (**Tab. 5**) because of the absence of extensive crossing barriers among the different *Prunus* species in the initial hybridization and the subsequent backcrosses (Browicz and Zohary, 1996; Gradziel *et al.*, 2001; Martínez-Gómez *et al.*, 2003b) encouraging the almond breeding and the effort to perform inter-specific hybrids and backcrosses with the aim to transfer useful traits (**Tab. 6**), (Gradziel *et al.*, 2001). Many authors also discuss about the crossing among these related almond species as rootstock, mainly under non-irrigated native conditions (Grasselly, 1975; Denisov, 1988; Kester and Hansen, 1966; Felipe, 1975) or to improve the quality of kernels or in respect of self-

compatibility (Kester and Gradziel, 1996; Gradziel and Kester, 1998; Gradziel *et al.*, 2001).

A genetic improvement program begins with a series of crossing inside the same species (intra-specific cross) or between related species (inter-specific crosses) in order to exploit natural existing variability, developing new almond cultivars. This process is useful for the species with a small genetic background, especially when a new useful traits has been found in the wild related ones (Foulongne *et al.,* 2003). After crossing, the next step is the selection of the best genotypes to carry on and grow, but this is a very long process. Indeed, the biggest critical point of the genetic improvement of the woody plants is their long juvenile phase (Martínez-Gómez *et al.*, 2003b, c).

Tab. 5- Main almond germplasm banks in the World and its consistency

Country	Research center	City	Accessions
Europe			
France	INRA-Avignon	Avignon	180
Italy	University of Udine	Udine	65
Spain	CEBAS-CSIC	Murcia	70
	CITA	Zaragoza	80
	USDA-UC-Davis	Reus	83
Asia			
Iran	National Plant Gene Bank	Karaj	67
Syria	Centre for studies of Arid Zones	Sednaya	130
Turkey	Ege University	Izmir	51
America			
USA	University of Davis	Davis	165
Africa			
Morocco	INRA-Rabat	Rabat	120
Oceania			
Australia	University of Adelaide	Adelaide	45

Source: FAO, (2009)

Tab. 6- Example of *Prunus* species of value for genetic improvement of almond

Section	Species	Use in peach and almond breeding[2]
Amygdalus Spach.	*P. persica* (L.) Batsch.	Self-compatibility and pest and desease resistance in amond
	P. davidiana (Carr.) Franch.	Disease resistance in peach and self-compatibility in almond
	P. mira Koehne.	Disease resistance in peach and self-compatibility in almond
	P. dulcis (Mill.) D.A. Webb	Pest and disease resistance in peach
	P. argentea Lam.	Self-compatibility and frost resistance in almond
	P. bucharica Korschinsky	Self-compatibility, growth habit and frost resistance in almond
	P. kuramica Korschinsky	Self-compatibility and disease resitance in almond
	P. webii (Spach) Vieh.	Self-compatibility and growth habit in almond
Chameamygdalus Spach.	*P. petunikowii* Lits.	Pest and disease resistance in almond
	P. tangutica Batal.	Pest and disease resistance in almond
Spartioides Spach.	*P. scoparia* Batal.	Self-compatibility and drought resitance in almond
Lepoptus Spach.	*P.pedeuncolata* Pall.	Pest and disease resistance in almond

Being almond a predominantly self-incompatible species, self-compatibility is one of the main goals for almond breeding programs in Europe and the USA (Grasselly *et al.*, 1981; Vargas *et al.*, 1984; Socias i Company and Felipe, 1988; Dicenta and García, 1993a; Gradziel and Kester, 1998). In addition, it is important to underline that several self-compatible almond cultivars were reported in Puglia,

Many quantitative traits have been studied in almond because of their agronomic importance, giving a good response to the selection (Kester *et al.* 1973; Kester and Asay, 1975; Grasselly and Crossa-Raynaud, 1980; Dicenta *et al.*, 1993; Socias i Company, 1998). For example, late flowering allowed to avoid the spring frosts in colder areas (Kester, 1965; Vargas 1984; Dicenta *et al.,* 1993; Socias i Company *et al.,* 1999). Additionally, flowering density and productivity were studied by Kester

[2] References: Hesse, 1975; Grassely, 1976; Denisov, 1988; Kester *et al*., 1991; Kester and Gradziel, 1996; Scorza and Sherman, 1996; Gradziel *et al*., 2001a, 2002.

and Asay (1975), Grasselly and Crossa-Raynaud (1980), Vargas (1984) and Dicenta *et al.,* (1993), and the time of ripening by Kester and Asay (1975) and Dicenta (1993b) to reach a better and consistent production; or yet, shell hardness, double kernels and kernel weight (Kester *et al.*, 1977; Vargas *et al.*, 1984; Dicenta *et al.*, 1993).

Although quantitative features are important for almond breeding, the agronomic traits, controlled by major genes also affects the almond business worldwide. Nowadays, almond quality has became an important objective for breeding (Socias i Company *et al.,* 2008), especially focusing on the resistance to biotic and a-biotic stresses, aimed to limit costs during the whole cycle of growing and ripening (Socias I Company, 1998).

A special attention is given also to the ripening dates and bitterness of the almond seed, which is one of the most important traits to improve, conforming it to the market request (Kester *et al.,* 1977; Vargas *et al.,* 1984; Dicenta *et al.,* 1993b; Janick, 2005). In particular, the genetic improvement for kernel quality must consider not only the chemical composition conferring specific organoleptic properties, but also physical traits affecting use as a different kind of shell, which is hard in most of Mediterranean countries and soft in California and regions with a similar growing system.

The chemical composition of almond kernels also represents an evolving goal for breeding because of the beneficial aspects of almond on human health. Although these aspects have been recently discovered, they are in the spotlight not only among almond breeders, but also among growers, processors and consumers.

5.2 Analysis of the variability and early selection

From long time, the identification and characterization of almond cultivars, as for other tree crops, focused on morphological markers. However, some of them are not always available for analysis, as they were affected by environmental conditions and may only be visible in adult materials, requiring a long time for their analysis. Studies

on large-sized populations, with long juvenile time (Martınez-Gomez *et al.*, 2003a; Scorza, 2001), on recessive gene expression, or aimed to study resistance to biotic or abiotic stress (Luby and Shaw, 2001), are really difficult to be carried on in tree crops. These obstacles gave the opportunities to develop new technologies much more revolutionary than those used in traditional breeding, being careful to transfer the wished traits, and avoiding to obtain negative ones (Quilot *et al.*, 2004).

Nowadays, molecular markers are routinely used in plant breeding aiming to perform the selection of economically important traits. A molecular marker can be defined as a genomic locus of variable dimension (generally between 50 and 3000 base pairs), detectable with specific probes or primers, whose simple presence is sufficient to mark in a unequivocal way a chromosomal stretch.

A simple mendelian inheritance characterize this kind of markers, so no epistatic interaction or pleiotropic effects are involved. Moreover, they should be widespread and easily available in the genome and not affected by environment and development stage of the plant (Frey,1981).

Apart from these features, markers often find limited application due to the frequent recombination during the first meiotic division which occurs and breaks linkage with genes of interest, or because they result monomorphic among the genotypes under selection. Indeed, the practical value of a marker, is closely related with its genetic distance from genes or quantitative trait *loci* (QTLs) determining the phenotype, as well as from their capacity to retain their diagnostic value across breeding populations.

So, an ideal potential marker should be also highly polymorphic at the single *locus*; with a high Polymorphism Information Content (PIC), defined as the product of the probability that the parent is heterozygous and the probability that the offspring is informative (Botstein *et al*., 1980); co-dominant, with the possibility to identify more *loci*; and being transferable intra and inter-species.

The main peculiarities of molecular markers are to recognize the parents and their segregating progeny, to give light on genetic variability, making a good choice to

perform, at the same time, crosses and by using cheap and fast techniques as gel electrophoresis on agarose or polyacrylamide.

Molecular markers give a support also in mapping and tagging of genes of agronomic interest and for disease resistance, genome mapping (genetic mapping, comparative mapping, physical and association mapping), marker assisted selection (MAS) and marker assisted backcrossing (MAB) during breeding programs, gender identification, studying the population structure and taxonomic and phylogenetic relationships. In particular, MAS is a powerful tool that, by direct analysis of the plant genotype with a marker tightly linked to the desired trait, allows to identify desidered individuals before these would be showed in the field. This supports the early assisted selection, lowering times, costs and spaces usually required in the traditional breeding (Wünsch and Hormaza, 2002; Martínez-Gómez *et al.*, 2003b; Sánchez- Pérez *et al.*, 2004a).

Finally, quantitative traits, as like as yield, disease resistance, stress tolerance, seed and fruit quality, etc., are really important for breeding programs and molecular markers are also used in studies concerning quantitative trait *loci* (QTLs), aiding the identification of candidate genes or the development of candidate genes markers associated with the trait under investigation (Neeraja *et al*., 2007; Romero *et al*., 2009).

5.3 Genetic mapping

There are two distinctive types of "maps" used for genome mapping: physical maps and genetic maps. Physical maps are based on the assessment of physical distances between genes along a chromosome, they are expressed in base pairs (bp). They could be at high or low resolution, obtained by means of several molecular biology techniques, as like as Radiation Hybrid Mapping (RH), Fluorescent in situ hybridization (FISH), DNA sequencing.

The first physical map for peach has been generated using high-information content fingerprinting (HICF), opening the door to the complete sequencing of other Prunus species such as almond (Zhebentyayeva *et al*., 2008).

In contrast, genetic maps, also known as linkage or meiotic maps, indicate the relative position and genetic distance between markers along chromosomes (Paterson, 1996; Collard *et al*., 2005), estimating the recombination frequency among the individuals of a segregant population. Indeed, parental chromosomes undergo recombination events (crossing over) during meiosis, changing DNA segments, obtaining a ri-arrangement of genes. Therefore, new recombinant chromosomes and allelic variants will be inherited in next generations of offspring.

The crossing over, producing recombination between genes, is directly related to their distance. Then, the relative distance between *loci* is measured in centiMorgan (cM) or map units (m.u.). In particular, 1 cM is equivalent to 1% of recombinant gametes obtained during meiosis. However, there is no direct linear relationship between units of genetic distances in centimorgan (cM) and physical distances in kylobase pairs (kb). The number of base-pairs corresponding to 1 cM could vary in relation to different species; for example, it is about 1 million in humans (Lodish *et al.,* 2004); in rice, on average, it is equal to 258.5 kb; in wheat it ranges from 118 to 22,000 kb (Gill *et al*., 1996a, b). Therefore, genetically close markers may actually be far apart in terms of base pairs (or vice versa) due to differences in the frequency of recombination along the length of a chromosome.

Although physical map could be a more "accurate" representation of the genome, genetic maps are essential to understand recombination rates, giving information about the organization of inheritance of traits (Myers *et al*., 2005).

The construction of a detailed genetic map requires four steps as described below:

1. The availability of a mapping population generated by two genetically divergent parents that show clear genetic differences for one or more traits of interest. The large size and the type of a mapping population are two of the most important features, which could exert an influence on the accuracy of genetic map (Ferreira *et*

al., 2006). The main populations used to develop genetic mapping are: progenies from the second filial generation (F_2), backcross (BC), recombinant inbred lines (RILs), double haploids (DHs) and near isogenic lines (NILs) (Burr *et al*.,1988; He *et al*., 2001; Doerge, 2002).

2. The development of molecular markers to genotype the mapping population. Several classes of molecular markers have been developed during the years using two different molecular biology techniques: the Southern Blot Hybridization (SBH) and the Polymerase Chain Reaction (PCR) (**Fig. 17**). The former allowed to obtain two classes of molecular markers: Restriction Fragment Length Polymorphisms (RFLP) and Variable Number of Tandem Repeat (VNTR). However, due to several disavantages related to laboriousity, time consuming, costs, etc., they were replaced by PCR markers, such as microsatellites or Simple Sequence Repeat (SSR), Expressed Sequence Tags (EST), Cleaved Amplified Polymorphic Sequence (CAPS), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphisms (Aflps), Inter Simple Sequence Repeat (ISSR), Diversity Arrays Technology (DArT), and Single Nucleotide Polymorphism (SNP).

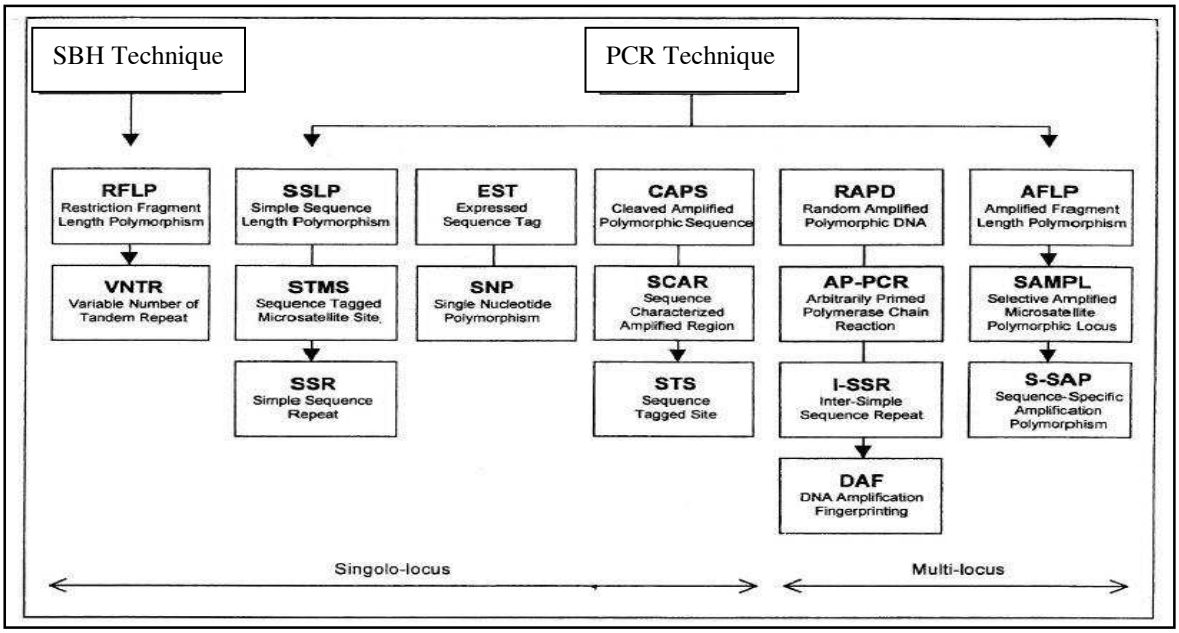


Fig. 17- Methods of analyses to find molecular markers and their grouping. Source: Barcaccia *et al*. 2000

In the present thesis, the fine mapping of the *Sk locus* has been performed using SSR and CAPS markers. Therefore, below more specific information for these two classes of molecular markers is reported.

5.3.1 Simple Sequence Repeat (SSR)

SSR markers, also called microsatellites, allow to detect differences among DNA sequence repeats which are simple, short (2-5 base pairs, and dispersed in the whole genome. They are easily reproducible and often single-locus markers.

SSRs show high polymorphism, co-dominant inheritance, they are abundant (Tauz, 1989; Morgante and Olivieri, 1993; Bell and Ecker, 1994) and are widely usd to describe genetic diversity in a wide range of plant species (Gupta *et al.*, 1996; Powell *et al.*, 1996). SSRs were used to establish genetic diversity and genetic relationships among peach and almond cultivars and related *Prunus* species (Serrano *et al.*, 2002; Martínez-Gómez *et al.*, 2003b).

Moreover, in *Prunus*, primer pairs flanking SSRs have been cloned and sequenced in peach (Cipriani *et al.*, 1999; Downey and Iezzoni, 2000; Sosinski *et al.*, 2000; Testolin *et al.*, 2000; Aranzana *et al.*, 2002, 2003a; Dirlewanger *et al.*, 2002; Georgi *et al.*, 2002; Wang *et al.*, 2002; Yamamoto *et al.*, 2002), apricot (Decroocq *et al.*, 2003), cherry (Downey and Iezzoni, 2000; Cantini *et al.*, 2001; Schueler *et al.*, 2003). All of them were used for molecular characterization and identification of cultivars in different species by means of different methods.

In addition, SSRs were employed also in genetic mapping for peach (Sosinski, 2000; Dettori *et al.*, 2001; Etienne *et al.*, 2002; Aranzana *et al.*, 2003b), almond (Joobeur *et al.*, 2000; Bliss *et al.*, 2002), and apricot (Hurtado *et al.*, 2002b; Vilanova *et al.*, 2003). Moreover, Testolin *et al.* (2004) found the first set of almond SSRs, then successfully used for almond cultivars characterization by Martínez-Gómez *et al.* (2003a) and Testolin *et al.* (2004); and for related *Prunus* species by Martínez-Gómez *et al.* (2003c).

5.3.2 Cleaved Amplified Polymorphic Sequence (CAPS)

CAPS markers allow to discriminate the polymorphism by using a PCR fragment digested with a restriction enzyme (Konieczny and Ausubel, 1993). They are developed on Single Nucleotide Polymorphisms (SNP) which are common kind of polymorphism in plant genomes and genes (Wicks *et al.,* 2001; Jehan and Lakhanpaul, 2006; Sehgal and Raina, 2008). Studies show that 30–40% of SNPs alter restriction endonuclease recognition sites, which is the basis to develop CAPS markers.

The conversion of SNPs in CAPS markers is facilitated by several freely available softwares, as like as SNP2CAPS (Thiel *et al* 2003), CapsID (Taylor and Provart, 2006) and CAPS Finder 2.0 (Neff *et al*., 2002), which held choosing appropriate and inexpensive restriction enzymes based on information on SNP polymorphism.

CAPSs are specific, reproducible, co-dominant and cheap (Varshney, 2010), making it accessible for plant breeding.

Furthermore, combining SNPs and SSR, several almond genotypes have been characterize, and genetic maps were also elaborated (Tavassolian Varshney, 2010).

3. A software to perform linkage analyses and to obtain a genetic map. Several computer packages are available for this kind of studies, such as JoinMap (Stam, 1993a and b), MAPMAKER/EXP (Lander *et al*., 1987), GMENDEL (Echt *et al*., 1992), LINKAGE (Suiter *et al*., 1983) and Map Manager QTX (Manly *et al*., 2001).

The basic principles in map construction consist in the use of statistical programs that calculate pair-wise recombination frequencies among markers, establish linkage groups, estimate map distances and determine map order.

Plant breeding can benefit from the availability of genetic maps for several reasons:

- Allow localization of qualitative genes or quantitative trait *loci* (QTL) (Mohan *et al.*, 1997; Doerge, 2002; Yim *et al*., 2002).

- Facilitate the introgression of desirable genes or QTLs through marker-assisted selection.

- Allow the comparison between maps of different species, aiming to evaluate similarity between genes, orders and function in the expression of a phenotype (Ahn and Tanksley, 1993; Paterson *et al*., 2000).

- Provide a framework for anchoring with physical maps based on chromosome translocations, DNA sequence or other direct measures (Yim *et al*., 2002).

- Constitute the first step towards positional or map based cloning of genes responsible for economically important traits (Mohan et al., 1997; Vuysteke *et al*., 1999).

However, a genetic linkage map should show requirements as like as simplicity, robustness, transferability, speed and cost effectiveness, to provide the benefits above cited (Lorieux *et al*., 2000).

5.4 Construction of almond genetic maps

The first linkage map analysis in almond was performed using isozymes (Arús *et al.,* 1994a; Vezvaei *et al.,*1995). In 1980s, the development of RFLPs provided a virtually source of high quality markers located all over the genome, allowing to build the first map for almond by Viruel *et al.,* (1995). It was obtained from the F_1 progeny by the crossing of "Ferragnès" and "Tuono" (FxT) and based on 120 RFLPs and 7 isoenzymes, confirming the eight expected linkage groups (corresponding to the eight chromosome peculiar of genus *Prunus*) and spanning approximately 400 cM.

Afterwards, another map, longer than the previous one (800 cM), was obtained by Foolad *et al.* (1995), using 111 RFLPs and 6 isozymes. The map was obtained by using an F_2 population of the inter-specific cross between the almond cultivar "Padre" and a peach selection "54P455" (P×5).

Another map was obtained in 1998 by Joobeur, who developed the *Prunus* reference map, a saturated linkage map created by using a F_2 population derived by the crossing between the almond cultivar Texas (T, syn. "Mission") and the peach Earlygold (E) cultivar (Arús *et al.,*, 1994b). TxE map contained eight linkage groups

and 246 markers: 11 isoenzymes and 235 RFLP, covering a total distance of 491 cM (Joobeur *et al.,*, 1998).

The development of markers obtained with simpler methods, such as RAPDs and SSRs, fostered studies addressed to the development of almond linkage map such as an improvement of F×T map (Joobeur *et al.,* 2000), with 174 markers, and an improved map P×5 by Bliss *et al.,* (2002), with 161 markers, including 6 morphological genes and 8 resistance-gene analog sequences.

Aranzana *et al.,*, (2003) improved further the (T×E) map by adding some good quality RFLPs and 96 SSRs, allowing to close some gaps and increasing map density.

The version of (TxE) map by Dirlewanger *et al.,* (2004a) includes 562 markers (361 RFLPs, 185 SSRs, 11 isozymes and 5 STSs), covering a total distance of 519 cM, with a map density of 0,92 marker/cM and largest gap of 7 cM (**Fig. 18**).

The inter-specific crosses performed between almond and peach or among other closely related fruit trees like apricot (Hurtado *et al*., 2002) and cherry (Wang *et al*., 2000), allowed to discover other main traits characterizing *Prunus*. The similar order of molecular markers observed in different *Prunu*s maps, when compared to the *Prunus* reference one, suggested, as previously reported, high level of synteny within the genus (Aranzana *et al.,*, 2003; Dirlewanger *et al.*, 2004a, b; Lambert *et al.,*, 2004). This homology among the genomes of *Prunus* species is in agreement with the low level of breeding barriers to inter-specific gene introgression and supports the opportunity for successful gene transfer between closely related species (Gradziel *et al.,* 2001; Martínez-Gómez *et al.*, 2003b). In addition, SSR markers developed by several other research groups have also been shown to be useful in characterization and genetic diversity studies across species in the genus *Prunus* (Cipriani *et al*., 1999; Sosinski *et al*., 2000; Dirlewanger *et al*., 2002; Wang *et al*., 2002).

Anyway, the *Prunus* reference map has been compared with the Arabidopsis genomic sequences, finding 23 syntenic blocks between them, which covered 23% of the *Prunus* map distance and 16% of the Arabidopsis genome (Dominguez *et al.,*, 2003). Micro-synteny studies have also found a fractional conservation between these

two distant taxons (Georgi *et al.,*, 2003) and indicate that the sequence of Arabidopsis can be employed to a limited extent for gene or marker search in *Prunus.*

The development of a new map by Sànchez-Pérez *et al*. (2007a and b; 2010) was an interesting discovery to place some major genes and QTLs in association group of almond (**Fig. 19**). Indeed, six tree traits (self-compatibility, blooming date, blooming density, productivity, leafing date and ripening time) and five pomological traits (kernel taste, in-shell weight, shell hardness, kernel weight and double kernel) have been studied in an F_1 almond progeny of 167 seedlings by the cross between the French cultivar "R1000" and the Spanish one "Desmayo Largueta". In addition, a set of 135 SSR markers developed from peach, cherry and almond were used for the molecular characterization of the progeny.

A genetic linkage map (RxD) was constructed with 56 of these SSRs. Co-segregation analysis allowed the identification of the map positions of two major genes to be confirmed for kernel taste (*Sk*) in linkage group five (G5) and for self-incompatibility (S) in G6. Moreover, QTLs mapped include two genes for leafing date (Lf-Q1 and Lf-Q2) in G1and G4, one for shell hardness (D-Q) in G2, one for double kernel (Dk-Q) and productivity (P-Q) in G4, one for blooming date (Lb-Q) in G4, two for kernel weight (Kw-Q1 and Kw-Q2) in G1 and G4 and two for in-shell weight (Shw-Q1 and Shw-Q2) in G1 and G2. Finally, four of SSR loci studied, (BPPCT011, UDP96-013, UDP96-003 and PceGA025) were linked to the important bio-agronomical traits of leafing date, shell hardness, blooming date and kernel taste respectively.

More recently, Sanchez-Pèrez *et al*., (2009) identified a major QTL for flowering time in G4 (Lb), together with other minor QTLs in G1, G6, and G7. In agreement with these results, a major QTL for chilling requirements was also located in G4 together with other minor QTLs in G1, G3, G6, and G7. Finally five minor QTLs were located controlling heat requirement in G1, G3, G4, G6 and G7.

All these results obtained and showed in this paragraph point out as the availability of high-density linkage maps may allow recent successes in establishing the approximate map position of major genes of horticultural value in almond (**Tab. 7**).

Tab. 7- Class of markers associated to main bio-agronomical traits in almond

Trait	Symbol	Linkage group	Marker	References
Flower color	B	G1	RFLP	Jàuregui *et al*. (2001)
Shell hardness	D	G2	RFLP	Arùs *et al*. (1999)
	D	G2	SSR	Sànchez-Pérez *et al*. (2007)
Nematode resistance	Mi	G2	RFLP	Bliss *et al*. (2002)
Anther color	Ag	G3	RFLP	Joobeur *et al*. (1998)
Blooming time	Lb	G4	RAPD	Ballester *et al*. (2001)
	Lb	G4	SSR	Sànchez-Pérez *et al*. (2007)
	Lb	G1, G4, G6, G7	SSR	Sànchez-Pérez *et al*. (2009)
Chill requirements	CR	G1, G3, G4, G7	SSR	Sànchez-Pérez *et al*. (2009)
Heat requirements	HR	G1, G3, G4, G7	SSR	Sànchez-Pérez *et al*. (2009)
Blooming density	Bd	G4	SSR	Sànchez-Pérez *et al*. (2007)
Leafing tim	Lf	G1, G4	SSR	Sànchez-Pérez *et al*. (2007)
Reaping time	Rd	G4, G5	SSR	Sànchez-Pérez *et al*. (2007)
In-shell weight	Shw	G1, G2	SSR	Sànchez-Pérez *et al*. (2007)
Kernel weight	Kw	G1, G4	SSR	Sànchez-Pérez *et al*. (2007)
Kernel taste	Sk	G5	RFLP	Bliss *et al*. (2002)
	Sk	G5	RFLP	Joobeur *et al*. (1998)
	Sk	G5	SSR	Sànchez-Pérez *et al*. (2007)
	Sk	G5	SSR	Sànchez-Pérez *et al*. (2010)
Double kernel	Dk	G4	SSR	Sànchez-Pérez *et al*. (2007)
Selfcompatibility	S	G6	RFLP	Ballester *et al*. (1998)
	S	G6	RFLP	Arùs *et al*. (1999)
	S	G6	SSR	Sanchez-Pérez *et al*. (2007)
	S	G6	RFLP	Bliss *et al*. (2002)

Source: Gradziel and Gomez, 2012

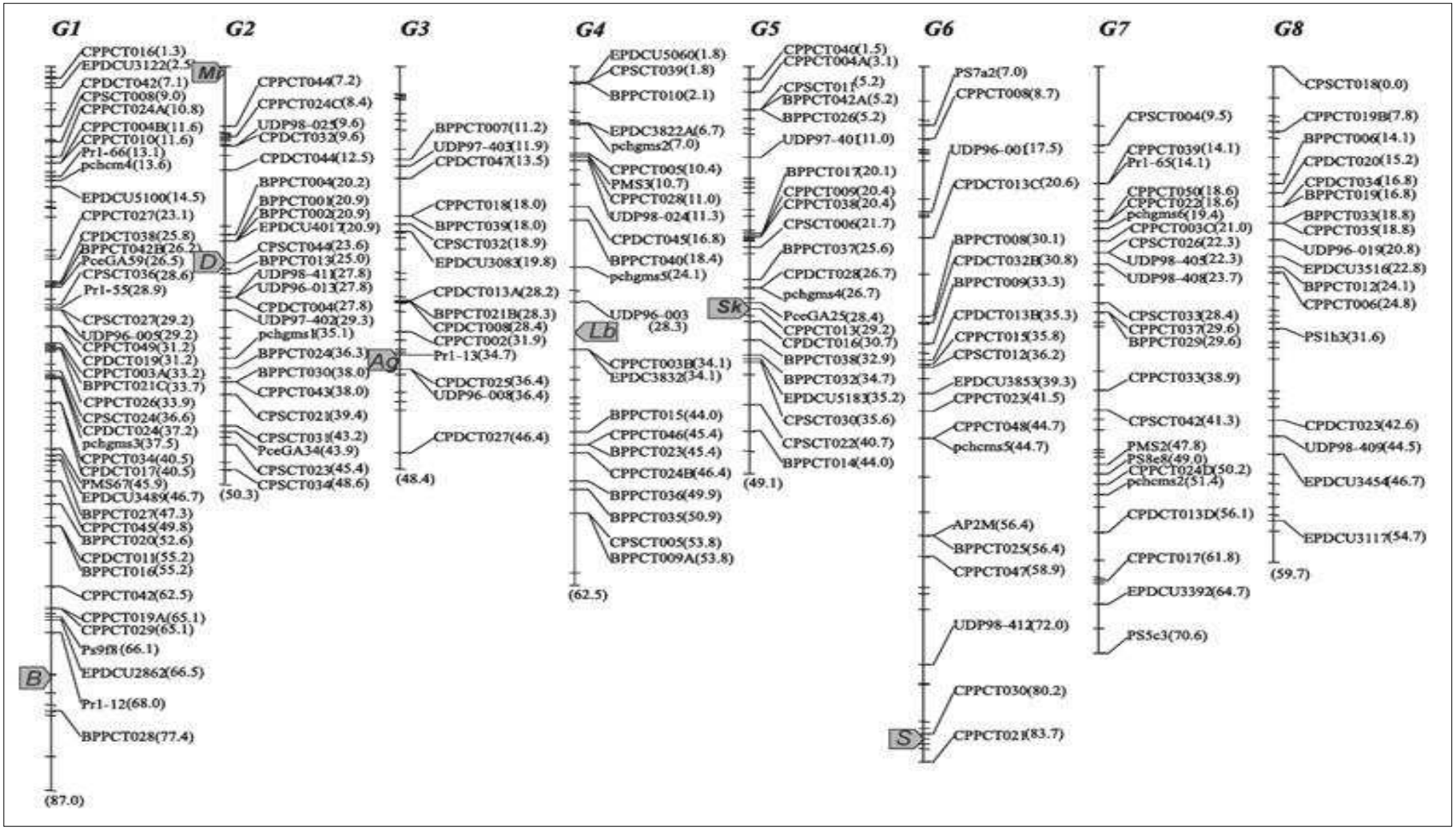


Fig. 18- Map of the "Texas" (almond) x "Earlygold" (peach) F_2 population obtained only with the SSR markers isolated by Dirlewanger *et al*., (2004a) and with the approximate location of gene involved in the control of flower color (*B*), nematode resistance (*Mi*), shell hardness (*D*), anther color (*Ag*), blooming time (*Lb*), kernel taste (*Sk*) and self-incompatibility (*S*)

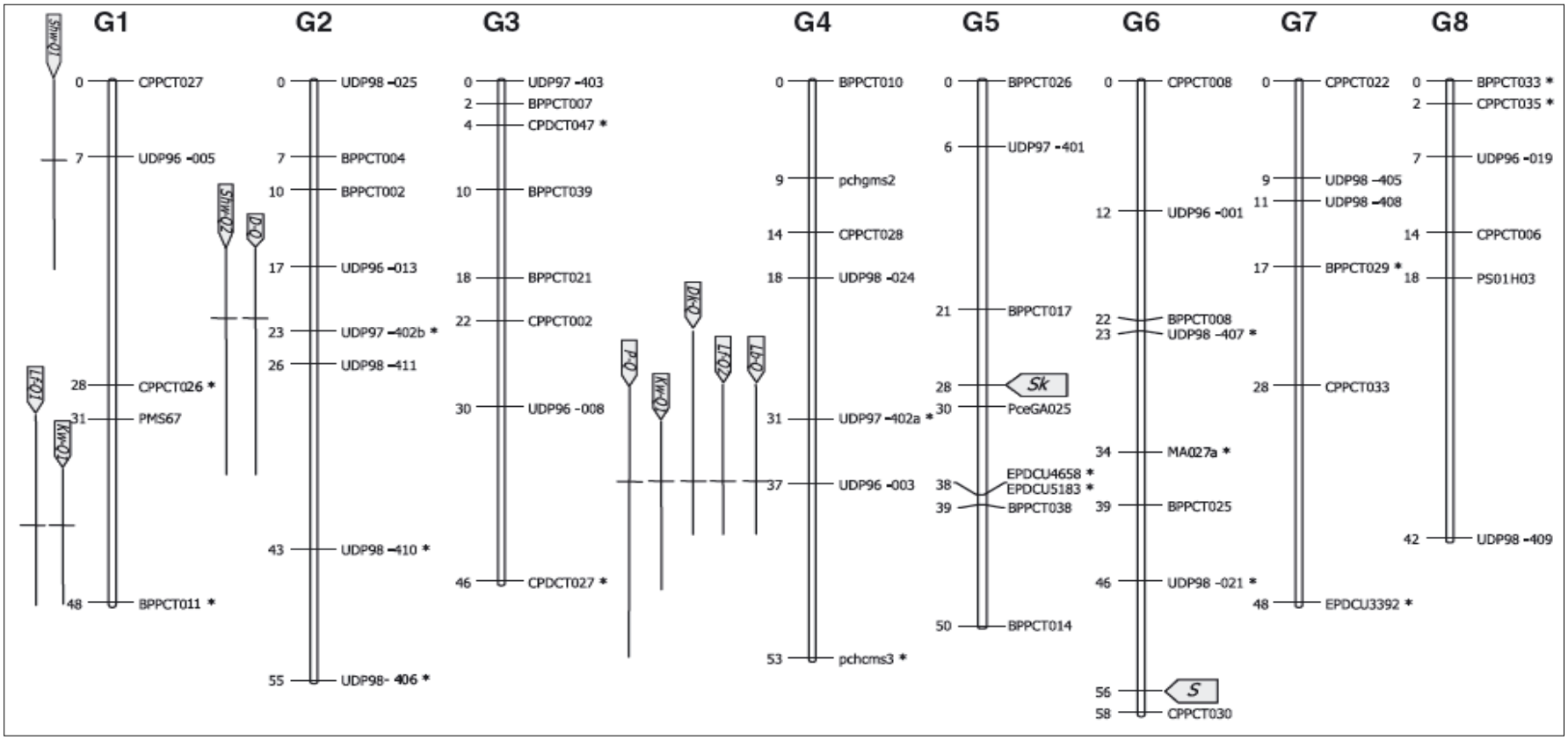


Fig. 19- Molecular linkage map of RxD F_1 progeny. The map has been obtained with 56 SSRs, locating approximately two major almond genes: *Sk* (kernel taste) and *S* (self- incompatibility, and 10 QTLs, including leafing date (Lf-Q1 and Lf-Q2), shell hardness (D-Q), double kernel (Dk-Q) and productivity (P-Q), blooming date (Lb-Q), kernel weight (Kw-Q1 and Kw-Q2), in-shell weight (Shw-Q1 and Shw-Q2). *indicates SSR mapped for the first time

5.5 The peach physical map

The first physical map has been generated for peach (Zhebentyayeva *et al.*, 2008) and it is composed by 2,138 contigs, containing 15,655 Bacterial Artificial Chromosome (BAC) clones. This physical map also integrates 2,633 markers including peach ESTs, cDNAs and RFLPs and is anchored to the *Prunus* (TxE) reference map using 152 core genetic probes (**Fig. 18**).

The International Peach Genome Initiative has released in 2010 by Sosinski the complete peach genome sequence (http://www.rosaceae.org/peach/genome), essential to support almond and other *Prunus* species researches. The peach v1.0 genome sequence consisted of eight pseudo-molecules (the first eight scaffolds) representing the eight chromosomes of peach and they are numbered according to their corresponding linkage groups. Nowadays, a second uploaded version 2.0 (The International Peach Genome Initiative, 2013) of peach genome is available (http://www.rosaceae.org/gb/gbrowse/prunus_persica_v2.1/) being of great interest for future molecular studies in *Prunus*, particularly for closely related species such as almond. This new version aims at improving several issues such as the chromosome-scale assembly and the annotation of the repeated and gene sequences.

MATERIALS AND METHODS

6. MATERIALS AND METHODS

6.1 Plant material

Plant materials assayed in this work belong to the "Almond Breeding Program" of "Centro de Edafòlogia y Biologìa Aplicada del Segura Consejo Superior de Investigaciones Cientìficas (CEBASIC-CSIC), Campus Universitario de Espinardo (Murcia)", Spain.

The materials were:

- The two genotypes "R1000" and "Desmayo Largueta", heterozygous at the *Sk* locus (sweet, *Sksk*). "R1000" is a late-blooming, mid- ripening, self-compatible selection of Institut National de la Recherche Agronomique (INRA) - Station de Recherches Fruitières Méditerranéennes Domaine Saint-Paul, Avignon, France, obtained from a cross between the North American cultivar "Tardy Nonpareil" and the Italian self-compatible cultivar "Tuono". On the other hand, "Desmayo Largueta" is a traditional Spanish cultivar, early blooming, late ripening and self-incompatible.

- A F_1 population progeny of 476 individuals was obtained by the cross "R1000"x"Desmayo Largueta" (RxD) performed in different years; the first in 1996 at INRA- Avignon (France) by Mr. Henri Duval;. the second and the third in 2009 and 2010 at CEBASIC-CSIC in Murcia (Spain). The F_1 population is maintained in the experimental orchard of CEBAS-CSIC in Murcia (Spain).

- The two almond cultivars "Lauranne" (sweet, *SkSk*) and "S3067"(bitter, *sksk*).

- "Lauranne" is a late blooming self-compatible cultivar originated in 1978 by the cross "Ferragnès" × "Tuono" from the breeding program

INRA between Spain and France (Socias i Company and Felipe, 1987; Grasselly, 1991; Vargas and Romero, 1994; Dicenta *et al*., 1999).

-'S3067' is a self-compatible bitter clone obtained in CEBAS in Murcia (Spain).

6.2 Sk fine-mapping

5.2.1 DNA extraction and quantification

Genomic DNA from "R1000" and "Desmayo Largueta" and their F_1 offspring was isolated according to a cetyltrimethylammonium bromide (CTAB) protocol described by Doyle and Doyle (1990): for each sample, 2-3 cm^2 of young leaves were frozen with liquid nitrogen and, by mortar and pestles, a fine powder was obtained and re-suspended with 400 µl of *extraction buffer* pH 7.5 (Sorbitol 0.35 M, Tris 100 mM, EDTA 5 mM), 500 µl of *nuclei lysis buffer* (Tris-Hcl 200 mM, EDTA 50 mM, NaCl 2 M, CTAB 2%) and 50 µl of *sarkosyl buffer* (10g di N-lauroyl sarcosine in 100 ml of water) in 2.0 ml Eppendorf tubes. After vortexing, samples were incubated at 65°C for 60 min and then, 800 µl of chloroform:isoamyl alcohol (24: 1) was added; a centrifugation for 4 min, speed 8000 rpm at room temperature, was performed to obtain DNA separation from other cellular components. The supernatant containing DNA was then transferred to a clean tube, and the DNA was precipitated by adding 800 µl of frozen isopropanol. A subsequent centrifugation for 1 min (8000 rpm) and a washing step of 20 min in 250 µl 76% ethanol with 10 mM of sodium acetate allowed to obtain a whitish pellet that subsequently has been air dried and dissolved in 100µl of sterile water. Finally, samples were incubated with 2µl of RNase (µg/ml) for 60 min at 37°C, in order to remove RNA residues.

The concentration and purity of isolated DNAs were estimated through the spectrophotometer Nanodrop 2000 UV-Vis (ThermoScientific), which

evaluates DNA absorbance at 260 nm and the optical density (OD) absorption ratios OD_{260nm}/OD_{280nm} and OD_{260nm}/OD_{230nm} . Optimal ranges of the two ratios are 1.8-2.0 and 2.0-2.2, respectively, which means low contamination from proteins, carbohydrates and polyphenols.

DNA integrity of each sample was verified by electrophoresis, loading 2 µl of DNA, 2 µl of the DNA dye (GelRed, Biotium) and 6 µl of water in 1% agarose gel in TBE 1X (Tris-HCl 1M pH 8.0; boric acid 1M; EDTA 0.5 M, pH 8.0) for 30 min at 100V, and visualizing the signal by means of at the UV transilluminator BioDoc-It Imaging system (UVP).

6.2.2 Identification of CAPS markers

For the identification of CAPS markers linked to the *Sk* locus, we used the peach genome version 1.0 (International_Peach_Genome_Initiative, 2013) available at the public database Genomic Database for Rosaceae (GDR) (www.rosaceae.org) (Jung *et al.,* 2008).

In detail, DNA sequences of almond SSR markers previously reported to flank the *Sk* locus (Sanchez-Pèrez *et al.,* 2010) were used to perform a Basic Local Alignment Search Tool (BLAST) analysis against the peach genome available at GDR.

The BLAST search allowed to identify a peach genomic region putatively containing the *Sk* ortholog and delimited by the two markers UDA045 and CPDCT028. Therefore, peach genes included in this region were selected based on their putative function, inferred using the InterPro (http://www.ebi.ac.uk/interpro/) and Gene Ontology (GO) (http://geneontology.org/) web resources. Then, a series of primer pairs were designed on peach genes, which are listed in **Tab. 8.** Primer pairs were designed on exon regions by means of the "Primer 3 plus" software (http://www.bioinformatics.nl/cgibin/primer3plus/primer3plus.cgi/) (Rozen and Skaletsky, 2000).

Tab. 8- Features of peach genes and primers used to develop cleaved amplified polymorphic sequence (CAPS) markers in this study

GDR accession code	Position on physical map	Primer pairs sequences (5'- 3')
ppa003882m	12562194	FW: CATAACGTCGCCAAGGAGAT
		REV: CATCCTTGCCAAAATCCACT
ppa018792m	13454839	FW: ACGTTGTCTCGTTCGTGGTT
		REV: AGGTGCTGCAAAGACACTGA
ppa006282m	12547702	FW: GTTTCGCTCGATTGGGTCTC
		REV: ATCATTTCCCGCCTGAATGC
ppa005388m	12644406	FW: GCTTCAAGGCAAGATTGGAG
		REV: ATTCCACAATTCGGTGGTTC
ppa001838m	12434063	FW: GGTTGTTCTGGGAGATGGAA
		REV: ACTTGACCGCAACCAAAATC

Primer pairs were tested on R, D and 6-8 individuals of the F_1 population by Polymerase Chain reaction (PCR), using the Phusion High-Fidelity DNA Polymerase (New England Biolabs). Sometimes, the optimal conditions for some primer pairs have been defined using a gradient thermocycler, the SimpliAmp Thermal Cycler, performing a gradient PCR (Lopez and Prezioso, 2001). The PCR protocol is described below:

Reagent	Initial concentration	Quantity per PCR reaction (µl)
Phusion Reaction Buffer (HF)	5X	5
$MgCl_2$	50mM	0,75
Nucleotides mix (dNTPs)	10mM	0,5
Forward	10µM	1,25
Reverse primer	10µM	1,25
Phusion DNA polymerase	2U/µl	0,25
DNA template	50 ng/µl	1
Bi-distillated sterile water	-	13,3
Final volume:		**25**

PCR conditions were:

Temperature (°C)	Time		
98	30"		
98	10"		
65	30"	x cycles	35
72	30'		
72	7’		
12	Hold		

Each PCR product was then checked by electrophoresis on 1.5-3% agarose gel in 1X TBE. Agarose concentration, as well as the running time and voltage, were in function of the amplicon lenght.

PCR products were purified by using the “Nucleo Spin Extract II” kit (Macherey Nagel, Germany) or the “QIAquick Gel Extraction kit” (Quiagen) in order to obtain high specific amplicons to sequence. Finally, purified PCR products were sequenced by the companies Macrogen Europe (The Netherlands) or Eurofins Genomics (Germany).

The electropherograms thus obtained were analyzed either visually or using the software “CLC Sequence Viewer 7” (http://clcbio.com), aiming to detect SNPs between the two parents segregating in F_1 individuals. Once the SNPs were identified, restriction enzymes were searched for which the SNP polymorphism was associated to the gain or loss of a restriction site, using the software “CAPS Designer” (genomics.net/tools/caps_designer/caps_input.pl).

Finally, PCR products were digested with the appropriate restriction enzyme according to the protocol of Neff *et al.,* (1998).

The digestion mix used was the following:

Reagent	Quantity for each sample
Restriction enzyme	5U
Buffer	1 µl
Water	5.4 µl
PCR product	3.5 µl
Final volume:	10 µl

All samples were incubated in a water bath with temperature and time indicated by the manufacturer (New England Biolabs) (**Tab. 9**). Afterwards, digested fragments were visualized by agarose or MetaPhor agarose gel electrophoresis, using gel concentrations dependent on the expected fragment length polymorphism.

Tab. 9- Restriction Enzymes used to get CAPS markers

Restriction enzyme	Temperature reaction (°C)
*Alu*I	37
*Msp*I/*Hpa*II	37
*Hpy*188I	37
*Hpy*188III	37
*Hpy*CH4V	37
*Bsa*WI	60
*Taq*I	65

Finally, CAPS markers founded were used to genotype the large RxD segregating F_1 population above described in this Chapter.

6.2.3 Microsatellite marker assays

Microsatellite markers linked to the *Sk* locus were selected based on the linkage map reported by Sànchez-Pèrez *et al.,* 2010, whose primer pairs, also reported by Sànchez-Pèrez *et al.,*2007*,* are listed in the following **Tab. 10**.

Tab. 10- SSR markers tested in this study and reported on the linkage map by Sànchez-Pèrez *et al.,* 2007

SSR marker *locus* name	**Primer Sequence 5'- 3'**	**Expected fragment size (bp)**
UDA-045	FW: CCATCCCAGGCCTTAGTACA	R: 168/168
	REV:GGAGGATGCTATTGGGGTC	D: 168/152
EPDCU2584	FW: TTCAGCTCATCTAGTTTCAT	R: 132/130
	CACGGTTCGAACAACATCTG	D: 132/130
CPDCT028	FW: TGAACGTTGCACTCCTTCAC	R: 190/190
	REV:ACCACCACCATAACCACCA	D: 186/166
BPPCT037	FW: CATGGAAGAGGATCAAGT	R: 131/132
	REV:CTTGAAGGTAGTGCCAAAG	D: 119/148

The SSR markers UDA045 and CPDCT028 were scored in the F_1 population performing a 20-µl PCR reaction with the DreamTaq DNA Polymerase (Thermo Scientific) as follows:

Reagent	**Concentration**	**Quantity per PCR reaction (µl)**
Reaction Buffer	10X	2
Nucleotide mix (dNTPs)	10mM	0.4
Forward primer	10µM	1
Reverse primer	10µM	1
Taq polymerase	5U/µl	0.08
DNA template	50 ng/µl	1
Bi-distillated sterile water	-	13.9
Final volume:		**20**

PCR conditions are described as follows:

Temperature (°C)	Time		
94	3 '		
94	30"	x	40
56	1’	cycles	
72	1'		
72	5'		
12	Hold		

Finally, PCR products were visualized on 2% agarose gel (Lonza) in 1X TBE.

The SSR markers: EPDCU2584 and BPPCT037 were scored on the RxD F_1 population by capillary electrophoresis. The software Primer3plus was used to design a modified forward primer containing a M13 tail of 18 bases (5’TGTAAAACGACGGCCAGT3’) at the 5’end (Schuelke, 2000). Moreover, a universal M13 primer complementary to the tail and labeled with the Fam or Hex fluorescent dyes (Sigma Genosys), was also added to the PCR reaction. Amplifications were carried out using the Dream*Taq* DNA Polymerase (Thermo Scientific) as follow:

Reagent	Concentration	Quantity perPCR reaction (µl)
Reaction Buffer	10 X	1.25
Nucleotides mix (dNTPs)	10mM	0.25
Forward M13	10µM	0.4
Reverse primer	10µM	2
Universal primer	10µM	1
Taq polymerase	5U/µl	0.05
DNA template	50 ng/µl	1
Bi-distillated sterile water	-	5.9
Final volume:		**12.5**

with the following PCR conditions:

Temperature (°C)	Time	
94	3 '	
94	30"	
56	1’	x 40 cycles
72	1'	
72	5'	
12	Hold	

All PCRs were performed either in the SimpliAmp or the GeneAmp PCR System 9700 thermocyclers (Applied Biosystem).

PCR product were checked by electrophoresis on 2-3% agarose gel in 1X TBE. Finally, samples to analyze by capillary electrophoresis were prepared in *MicroAmp™ Optical 96-Well Reaction Plates* (Applied Biosystems) by adding: 1.2 µl of PCR product, 15 µl HiDi formamide (Applied Biosystems) and 0.3µl GeneScan™-500 ROX™ Size Standard; after a denaturation at 94°C for 5’, they were run on a 36 cm electrophoretic capillary (ABI-3500 Genetic Analyzer, Applied Biosystems, : see also **Figg**. **20 and 21**), following the parameters below reported:

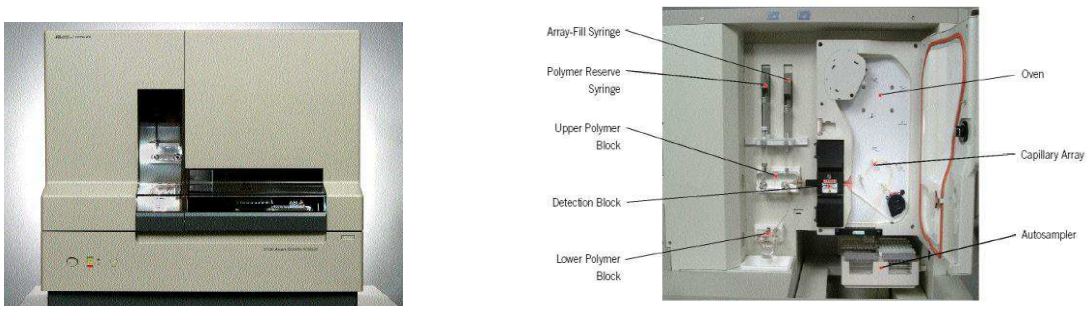


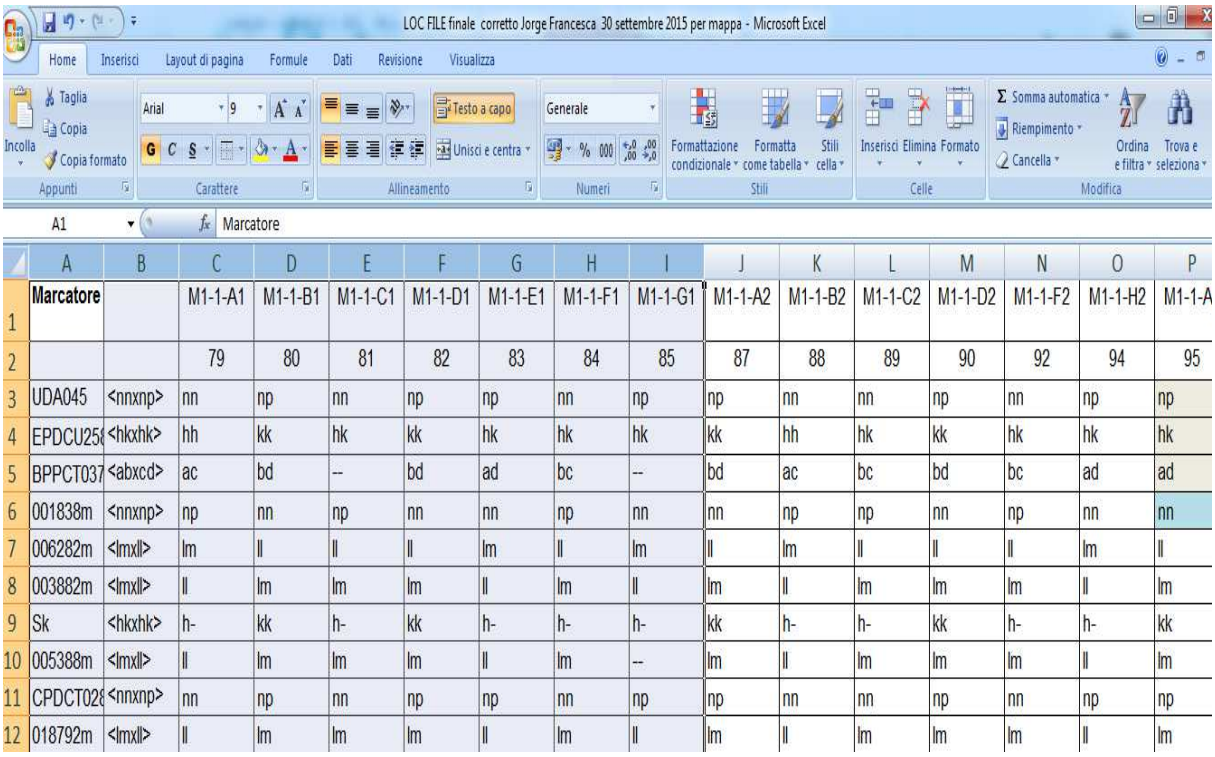
Fig. 20 - 21- Automated sequencer ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystem)

Project Name	3100-Avant-Project
Dye Set	D
Run Module	GeneScan36-POP4 Default Module
Analysis Module	GS 500 Analysis gsp
Oven Temperature	60°C
Laser Power	15,0 mW
EP Voltage	150 KV
EP Current	30 mA
Laser Current	4,0 A

Finally, data were automatically transferred from the sequencer to the *Workstation* and analyzed with the software GeneMapper v. 5.0. This is able to transform received data in electropherograms. For each individual, alleles at each SSR locus were called on the basis of their estimated molecular weight and the quality of the fluorescence peak.

6.3 Sk fine mapping

Phenotyping was carried out on the F_1 segregating population, by tasting at least two seeds for each of the 476 individuals. Phenotypic segregation data were merged with marker segregation data in an excel file suitable to be imported in the *JoinMap* v. 4.1 software (Van Ooijen, 2011) and create a so-called LOC file (**Fig. 22**). The logarithm of the odds (LOD) score threshold value of 3 was set to test for significant linkage between markers. The order and distance among the markers were determined using the Maximum Likelihood mapping algorithm for cross-pollinated populations (CP), which result from crossing heterozygous individuals. Default mapping parameters were assumed with spatial sampling thresholds of 0.100, 0.050, 0.030, 0.020 and 0.0010, the number of Monte Carlo EM cycles of 4 and increasing the number of optimization rounds to 5.



LOC FILE finale corretto Jorge Francesca 30 settembre 2015 per mappa - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Marcatore		M1-1-A1	M1-1-B1	M1-1-C1	M1-1-D1	M1-1-E1	M1-1-F1	M1-1-G1	M1-1-A2	M1-1-B2	M1-1-C2	M1-1-D2	M1-1-F2	M1-1-H2	M1-1-A
2			79	80	81	82	83	84	85	87	88	89	90	92	94	95
3	UDA045	<nnxnp>	nn	np	nn	np	np	nn	np	np	nn	nn	np	nn	np	np
4	EPDCU25	<hkxhk>	hh	kk	hk	kk	hk	hk	hk	kk	hh	hk	kk	hk	hk	hk
5	BPPCT037	<abxcd>	ac	bd	--	bd	ad	bc	--	bd	ac	bc	bd	bc	ad	ad
6	001838m	<nnxnp>	np	nn	np	nn	nn	np	nn	nn	np	np	nn	np	nn	nn
7	006282m	<lmxll>	lm	ll	ll	ll	lm	ll	lm	ll	lm	ll	ll	ll	lm	ll
8	003882m	<lmxll>	ll	lm	lm	lm	ll	lm	ll	lm	ll	lm	lm	lm	ll	lm
9	Sk	<hkxhk>	h-	kk	h-	kk	h-	h-	h-	kk	h-	h-	kk	h-	h-	kk
10	005388m	<lmxll>	ll	lm	lm	lm	ll	lm	--	lm	ll	lm	lm	lm	ll	lm
11	CPDCT028	<nnxnp>	nn	np	nn	np	np	nn	np	np	nn	nn	np	nn	np	np
12	018792m	<lmxll>	ll	lm	lm	lm	ll	lm	ll	lm	ll	lm	lm	lm	ll	lm

Fig. 22- Example of a excel file with the segregation data obtained in this work, suitable to be imported in JoinMap.

6.4 Expression analysis of candidate Sk genes

6.4.1 Total RNA extraction, quantification and cDNA synthesis

Total RNA from the almond varieties "Lauranne" (sweet, *SkSk*) and "S3067"(bitter, *sksk*) was extracted in the tegument tissues at different time points (01/04/2015, 08/04/2014, 15/04/2015, 22/04/2015, 30/04/2015, 05/05/2015, 13/05/2015, 28/05/2015, 12/06/2015, 17/06/2015)using the Ultra Clean Plant RNA Isolation Kit (MO BIO).

The concentration of isolated total RNA for each sample was estimated by means of the spectrophotometer Nanodrop 2000 UV-Vis (ThermoScientific), which evaluates RNA absorbance at 260 nm and the optical density (OD) absorption ratios OD_{260nm}/OD_{280nm} and OD_{260nm}/OD_{230nm} . Optimal ranges of the two ratios are 1.8-2.05 and 2.0-2.60, respectively, which means low contamination from proteins, carbohydrates and polyphenols. RNA integrity

was evaluated by electrophoresis on 1% Agarose gel in TBE 1X, staining 2 µl of RNA with2 µl of GelRed (Biotium) and 6µl of water. Intact rRNA subunits of 18S and 28S were observed on the gel.

Afterwards, the corresponding cDNAs were synthesized using the "Superscript III reverse Transcriptase" kit (Life technologies), with oligo(dT) primers. Genomic DNA was eliminated by treating each sample with RNase-free DNase I (TaKaRa, Japan), according to the instructions manual.

6.4.2 Gene Expression analyses: quantitative Real-Time PCR (qRT-PCR)

Primer pairs of the seven candidate genes were designed by means of the software Primer 3 plus on intron/eson junctions (Table giu). This was to avoid the possibility that genomic DNA was also amplified during RT-qPCR reactions. Primer pairs were designed to have a melting temperature between 60–62°C, a length to 20–26 bases and a content about 50% of GC. The length of the expected amplicons were optimized to 150-250 bp to ensure optimal polymerization efficiency and minimize the impact of RNA integrity on relative quantification of gene expression.

Before performing the RT-qPCR all the primer pairs have been tested *via* standard PCR (whose protocol has been described above). Each PCR product has been finally checked through 1.5% agarose gel electrophoresis in TBE 1X.

A real-time qPCR was carried out in order to perform gene expression analyses only on three of the seven candidate genes (ppa022201m, ppa0023406m, ppa011942m).

The peach housekeeping gene *TEF2* was chosen as an internal control to normalize mRNA levels between different samples and for an exact comparison of gene expression level, according with Tong *et al*., (2009). All details about the genes assayed for the expression levels and corresponding primer pairs are listed in **Tab. 11**.

After the previous optimizations, the RT-qPCRs have been carried on using the SYBR Select Master Mix for CFX, Applied Biosystem (Life technologies) and the CFX96 Touch-Real Time PCR Detection System termocycler (Bio-Rad) following the protocol shown below:

Reagent	Quantity (µl)
SYBR Select Master Mix	7
Forward (10µM)	1
Reverse primer (10µM)	1
DNA template (10 ng/µl)	1
Bi-distillated sterile water	17,8
Final volume:	**12.5**

RT-qPCR conditions were set as:

Temperature (°C)	Time	
95	2 '	x 1 cycle
95	15"	x 40
60	1	cycles

The experiment was planned in order to analyze three biological replicates for different time points. Each PCR reaction was carried out in duplicate and the mean values were used for RT-qPCR analysis. A melting curve was generated for each sample at the end of each run to ensure the purity of the amplified products.

The gene expression level compared with respect to the reference gene has been determined using the $2^{-\Delta(\Delta CT)}$ method (Livak and Schmittgen, 2001).

The Student's test (t-test) was performed in order to infer significant genotypic differences at each time point.

Tab. 11- Features of peach genes and primer pairs designed to perform Real Time qPCRs on seven candidate genes included in a region of about 80 Kb of peach map

GDR accession code	Position on physical map	Forward Primer Sequence 5'- 3'	Product size (bp)
ppa011942m	12.576.856	CCAACGCCTAAATGCTCTTGTT	210
		AGATGTTTGCGTAAACTCTTGA	
ppa023406m	12.587.426	CAGTTGAGCCAACTAATCTACC	221
		ATTCAACAGAATGAGCATAGC	
		A	
ppa022201m	12.597.330	CCAAACAAGAAGGTGACACAA	218
		AGAGCGGAGGGCATAAAATC	
ppa025417m	12.603.688	AAGGCAGCCAGGGTGATACA	220
		AACAGAGCGGAGGACGTAGA	
ppa027182m	12.612.821	CCAGACAAGGTGGATCATCA	218
		TTTCGACACATTCGGAACAA	
ppa015634m	12.625.785	CAGGAGGGCAGCTGTTGTA	241
		AGCGGAGGACATAGAAACGA	
ppa005343m	12.636.946	CACAGAGCAGTCGTGGTACA	213
		GCTTCTCTCTCCGCTGTCTC	
		GGTGTGACGATGAAGAGTGAT	
TEFII		G	129
		TGAAGGAGAGGGAAGGTGAAA	
		G	

AIMS

Aims of the work

This thesis is the result of a collaboration among the Department of Agricultural, Food and Environmental Sciences (S.A.F.E) of the University of Foggia, the Department of Soil, Plant and Food Sciences (Di.S.S.P.A.) of the University of Bari "Aldo Moro", the Department of Plant and Environmental Sciences of the University of Copenhagen and the Department of Plant Breeding of the University of Murcia.

The thesis addressed different aims. The first was the identification of new almond molecular markers tightly linked to the *Sk* locus, which controls the biosynthetic pathway of the cyanogenic di-glucoside amygdalin and its accumulation, thus determining sweetness or bitterness of kernels.To pursue this aim, SNP polymorphisms in almond were detected on the basis of peach gene sequences occurring in a genomic region putatively synthenic to the one containing the *Sk* locus.

A second fundamental goal of the thesis was to provide a fine map of the *Sk* region, aiming to narrow down the interval in which the *Sk* gene is located, thus supplying the basis for its isolation *via* positional cloning.

Finally, research activities were addressed to the identification of a candidate *Sk* gene. To this aim, almond genes located in the *Sk* region were analyzed with respect to their putative function and expression levels during the development of the tegument of sweet and bitter almond genotypes.

RESULTS

7. RESULTS

7.1 Identification of a peach genomic region synthenic to the one containing the Sk ortholog in almond

As reported in the introduction of this thesis, the genomes of peach and almond are highly similar, there is evidence for genome co-linearity within the *Prunus* genus and the genomic sequence of peach is freely available at the GDR database.

Aiming to detect a peach genomic region putatively containing the ortholog of the *Sk* gene, we performed a BLAST analysis in the peach genome, using as query sequences of linkage group V almond markers previously shown to flank the *Sk* locus, i.e. UDA045 and CPDCT028 (**Figg. 16 and 23,** Sanchez-Perez *et al.*, 2010). Notably, this resulted in the identification of identical sequences on peach chromosome 5, thus suggesting the detection of a synthenic region containing the *Sk* ortholog. The markers UDA045 and CPDCT028 delimit a physical region of 813.897 Kb (**Fig. 24**).

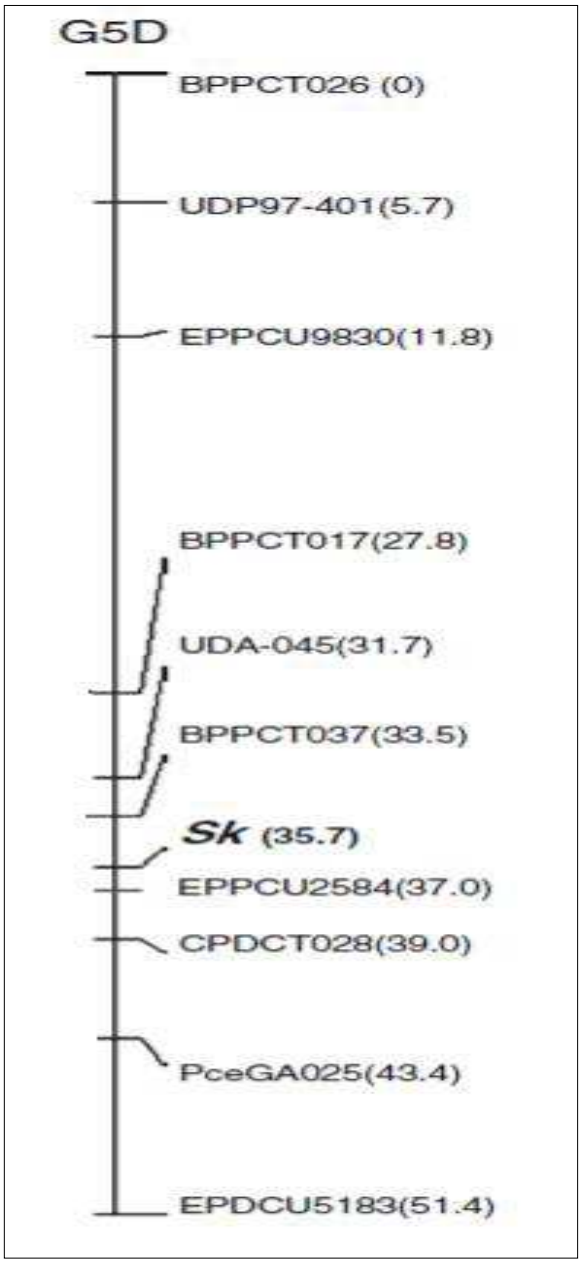


Fig. 23- Linkage map of the *Sk* locus region identified on almond linkage group V (G5D), (Sánchez-Pérez *et al.*, 2010). Numbers in brackets indicate genetic distances expressed in centiMorgans (cM)

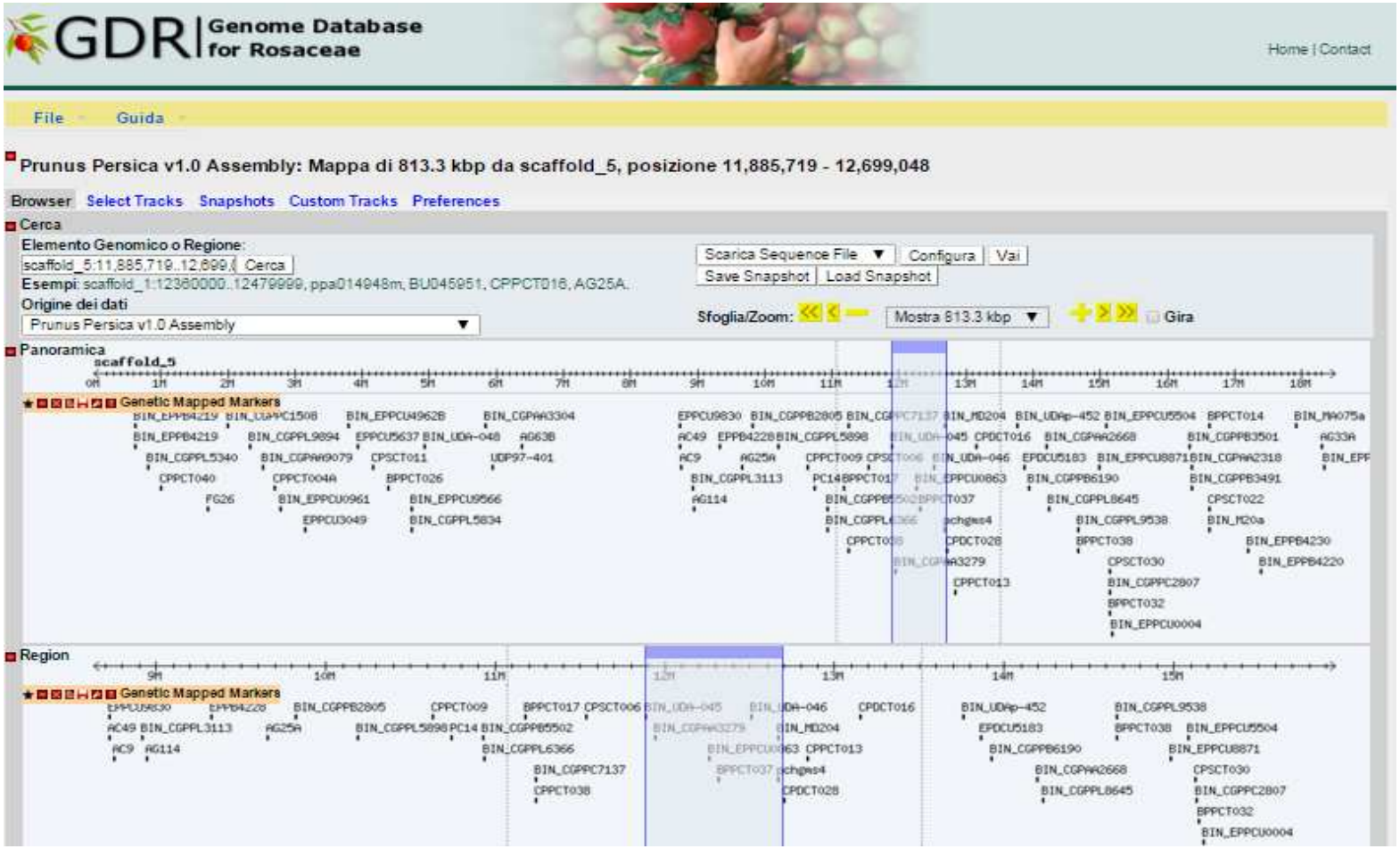


Fig. 24 Detail of a physical region on peach chromosome 5 visualized by means of the GenomeBrowse tool available at the Genomic Database for Rosaceae (GDR). The box delimits the 813.897 Kb interval flanked by the ortholog sequences of the *Sk*-linked SSR markers UDA045 and CPDCT028

7.2 Identification of CAPS markers

The identification of a peach physical region putatively synthenic to the one harbouring the *Sk* locus in almond paved the way to the development of new molecular markers linked with the *Sk locus*, useful to saturate a previous map of the *Sk* region (**Figg. 16 and 23,** Sanchez-Perez *et al,.* 2010,) and facilitate gene isolation and marker assisted selection (MAS) in almond breeding.

In more detail, several primer pairs were designed on peach genes included in the region between the markers UDA045 and CPDCT028. In addition, a primer pair was designed on the gene ppa018792m, which is located outside the UDA045/CPDCT028 interval, but encodes a sweet sugar transporter that could have explained differences in the localization of the glucoside prunasine in sweet and bitter almond genotypes, described in Sanchez-Perez *et al*., (2008), (**Tab. 12)**. Primer pairs were then tested in PCR reactions using, as templates, the gDNAs of the two almond parental lines Desmayo Largueta (D) and R1000 (R) and six/eight individuals of a large RxD F_1 segregant population of 476 individuals. Sequencing of amplification products and inspection of electropherograms allowed to detect six SNPs polymorphisms in five different genes, which were segregating in F_1 (**Tab. 13**) (**Fig. 25**).

Table 12- Genome Database for Rosaceae (GDR) identification code (ID) and physical location of peach genes associated with CAPS markers. Information on putative gene functions based on InterPro and Gene Ontology (GO) databases is also provided

GDR ID	Peach physical map position	InterPro Function	GO assignment Function
ppa001838m	12434063	Helix-loop-helix DNA-binding domain Transcription factor MYC/MYB N-terminal	Not available
ppa006282m	12547702	Uncharacterised protein family UPF0017, hydrolase-like, conserved site	Molecular function: carboxylesterase activity
ppa003882m	12562194	Cytochrome P450 (family)	- Biological process: a) redox process - Molecular function: a) electron carrier activity b) heme binding c) iron ion binding
ppa005388m	12644406	Alpha/beta hydrolase fold-1	Not available
ppa018792m	13454809	RAG1-activating protein-1-related	- Cellular component: a) integral to membrane

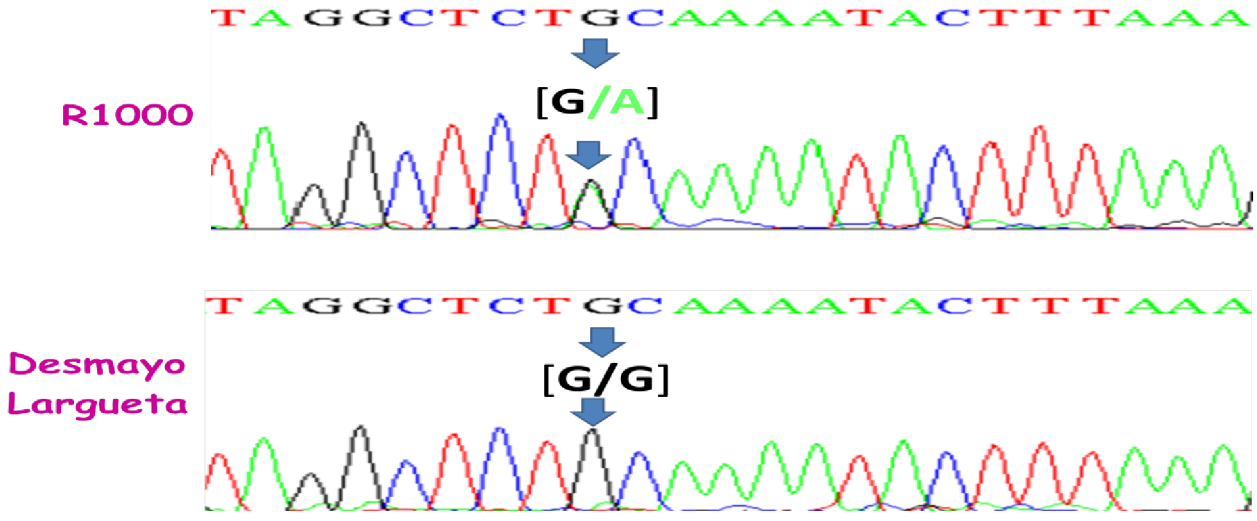


Fig. 25 - Example of electropherograms obtained by sequencing of the DNA of parents R and D using the primer pairs ppa005388m. SNP polymorphisms used to develop CAPS markers are highlighted by arrows

By means of the software "CAPS designer", SNP polymorphisms were used to search for changes of restriction sites, useful to the development of CAPS molecular markers. In total, five CAPS markers were developed, as only one of the two polymorphism identified in the gene ppa005388m was used. In more detail:

1. A T→ G polymorphism in position 77 with respect to the peach gene ppa001838m (R genotype: T/T; D genotype: G/T) was revealed by the restriction enzyme *Hpa*II, resulting in three fragments in R (approximate sizes 550, 400, 300 and 150) and two fragments in D (approximate sizes 400, 300 and 150), expected to segregate according to a 1:1 ratio (**Fig. 26**).
2. A G → A polymorphism in position 425 with respect to the peach gene ppa006282m (R genotype: G/A; D genotype: G/G) was revealed by the restriction enzyme *HpyCH4V*, resulting in three fragments in R (approximate sizes 400, 320 and 280) and four fragments in D (approximate sizes 500, 400, 320 and 280), expected to segregate according a 1:1 ratio (**Fig. 27**).
3. A G→ A polymorphism in position 459 with respect to the peach gene ppa003882m (R genotype: G/A; D genotype: G/G) was revealed by the restriction enzyme *BsaW*I, resulting in four fragments in R (approximate sizes 600, 550, 320 and 280) and three fragments in D (approximate sizes 550, 320 and 280), expected to segregate according a 1:1 ratio (**Fig. 28**)
4. A C→ T polymorphism in position 528 with respect to the peach gene ppa005388m (R genotype: C/T; D genotype: T/T) was revealed by the restriction enzyme *Taq*I, *re*sulting in three fragments in R (approximate sizes 750, 520 and 400) and two fragments in D (approximate sizes 520 and 400), expected to segregate according a 1:1 ratio (**Fig. 29**).
5. A T → C polymorphism in position 317 with respect to the peach gene ppa018792m (R genotype: T/C; D genotype: T/T) was revealed by the restriction enzyme *Alu*I, *re*sulting in three fragments in R (approximate sizes

450, 380 and 280) and two fragments in D (approximate sizes 380 and 280), expected to segregate according a 1:1 ratio (**Fig. 30**).

Table. 13- Features of the CAPS markers identified in this study. Each marker was designed on the sequence of a peach gene, whose accession code in the Genomic Database for Rosaceae (GDR) and position on the peach physical map are reported. Information on primer sequences forward (Fw) and reverse (Rev), SNP polymorphisms, restriction enzymes used and size of the digestion products is also provided

GDR ID	SNP	Primer sequences 5'-3'	Enzyme	Digestion product size (bp)
ppa001838m	$T^{77}\rightarrow G$	Fw: GGTTGTTCTGGGAGATGGAA Rev: ACTTGACCGCAACCAAAATC	*MspI/H* *pa*II	D: 115, 690 R: 805
ppa006282m	$G^{425} \rightarrow A$	Fw: GTTTCGCTCGATTGGGTCTC Rev: ATCATTTCCCGCCTGAATGC	*HpyCH* *4*V	R:14, 251, 329, 85 D:14, 251, 160, 169, 85
ppa003882m	$G^{459} \rightarrow A$	Fw: CATAACGTCGCCAAGGAGAT Rev: CATCCTTGCCAAAATCCACT	*Bsa*WI	R:551, 515, 240, 18, 36 D: 515, 240, 180, 36
ppa005388m	$C^{528}\rightarrow T$	Fw: GCTTCAAGGCAAGATTGGAG Rev: ATTCCACAATTCGGTGGTTC	*Taq*I	R: 994 D: 528, 466
ppa018792m	$T^{317} \rightarrow C$	Fw: ACGTTGTCTCGTTCGTGGTT Rev: AGGTGCTGCAAAGACACTGA	*Alu*I	R: 64, 252, 288 D: 540, 64

Fig. 26- Electrophoretic pattern obtained using the CAPS marker ppa001838m/*Hpa*II on the two parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F_1 individuals

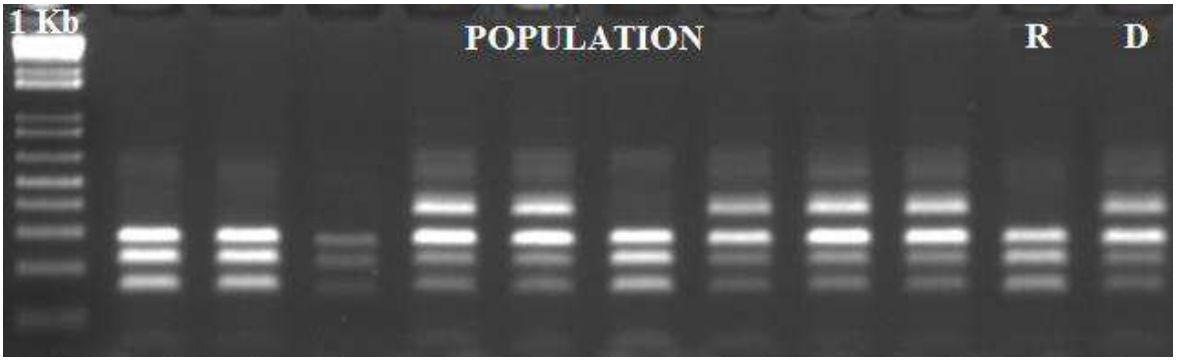


Fig. 27- Electrophoretic pattern obtained using the CAPS marker ppa0062820m/*HpyCH4*V on the two parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F1 individuals

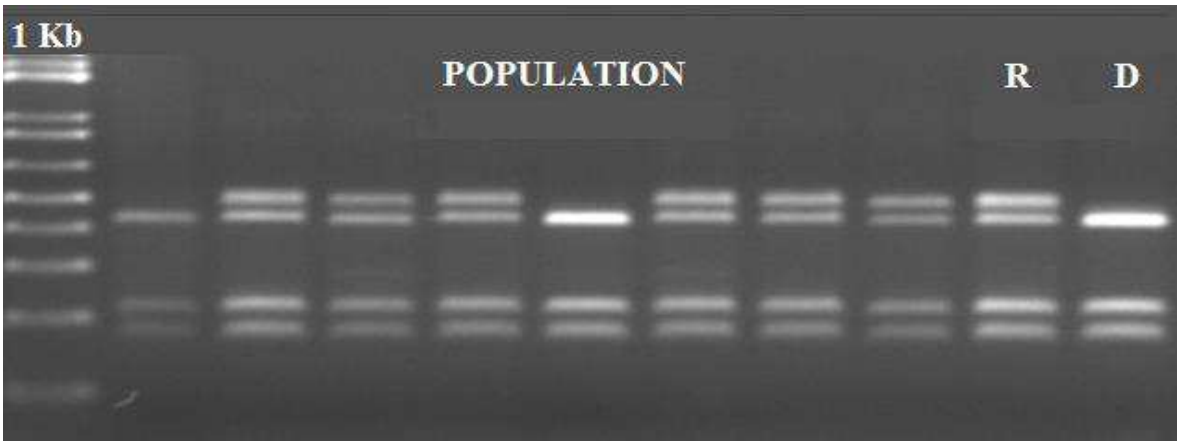


Fig. 28- Electrophoretic pattern obtained using the CAPS marker ppa003882m/*BsaWI* on the two parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F_1 individuals

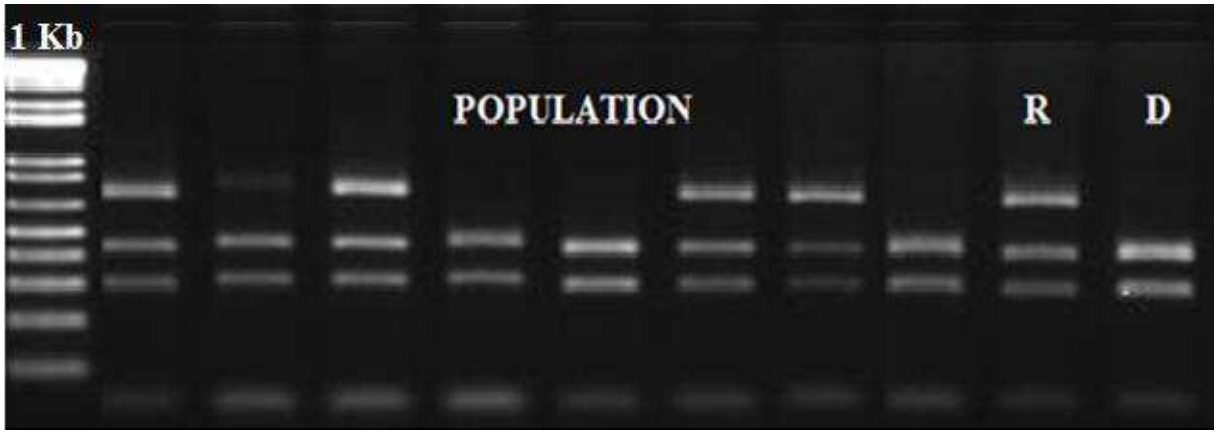


Fig. 29- Electrophoretic pattern obtained using the CAPS marker ppa005388m/*Taq*I on the two parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F_1 individuals

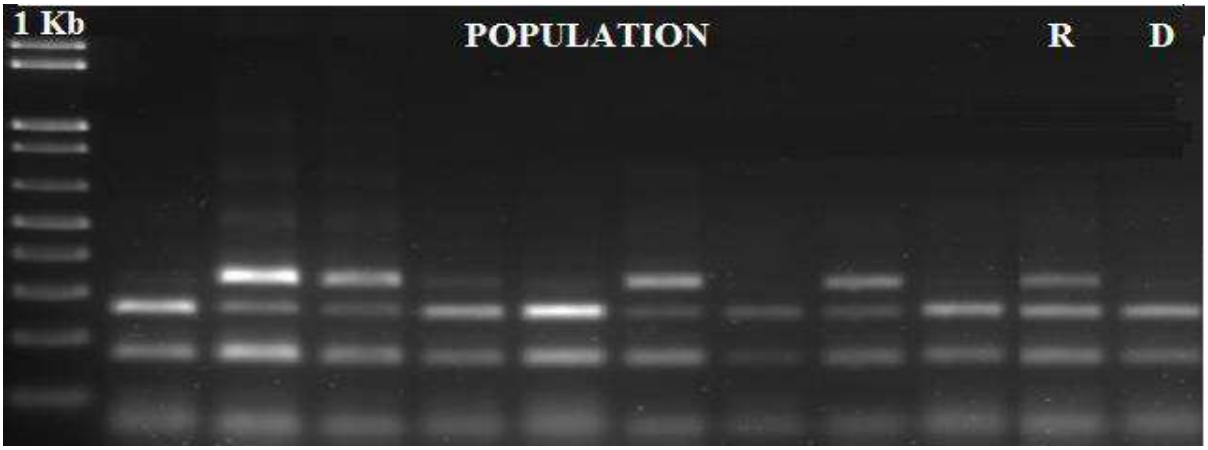


Fig. 30- Electrophoretic pattern obtained using the CAPS marker ppa018792m/*Alu*I on the two parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F_1 individuals

7.3 Sk fine mapping and assessment of almond/peach genomic collinearity

In order to provide genotypic data for *Sk* fine-mapping, the five CAPS markers above mentioned were used to genotype a large F_1 population (RxD) of 476 individuals. Furthermore, four *Sk*-linked SSR markers previously described by Sanchez-Perez *et al.* (2010), i.e. UDA045, CPDCT028, BPPCT037 and EPDCU2584 (**Figg 31, 32, 33 and 34**), were also used to genotype the same F_1 population. Differently from Sanchez-Perez *et al,.* (2010), the two markers BPCCT037 and EPDCU2584 were scored by means of capillary electrophoresis, rather than by acrylamide gels, as the former methodology allows the accurate detection of fragment length polymorphisms of a single base pair and thus minimizes risks of scoring errors.

Features of the SSR markers reported in this study are provided in **Tab. 14** and **Figg**. **31, 32, 33 and 34.**

Table 14- Features of the SSR markers used in this study. For each SSR markers the size of the amplicons obtained in the two genotypes of almond "R1000" and "Desmayo Largueta" is also reported

SSR marker	Extimated amplicon sizes
UDA045	R: 180/180 D: 180/160
CPDCT028	R: 200/200 D: 200/180
BPPCT037	R: 150/151 D: 137/166
EPDCU2584	R: 152/150 D: 152/150

It should be noticed that the size of the alleles obtained for the markers BPPCT037 and EPDCU2584 is about 20 bp longer than the one reported by Sánchez-Pérez *et al.,* (2010) (**Figg. 31** and **32),** as capillary electrophoresis requires the addition of a M13

universal tail to the forward primer, to ensure the staining with a fluorescent dye (see Materials and Methods).

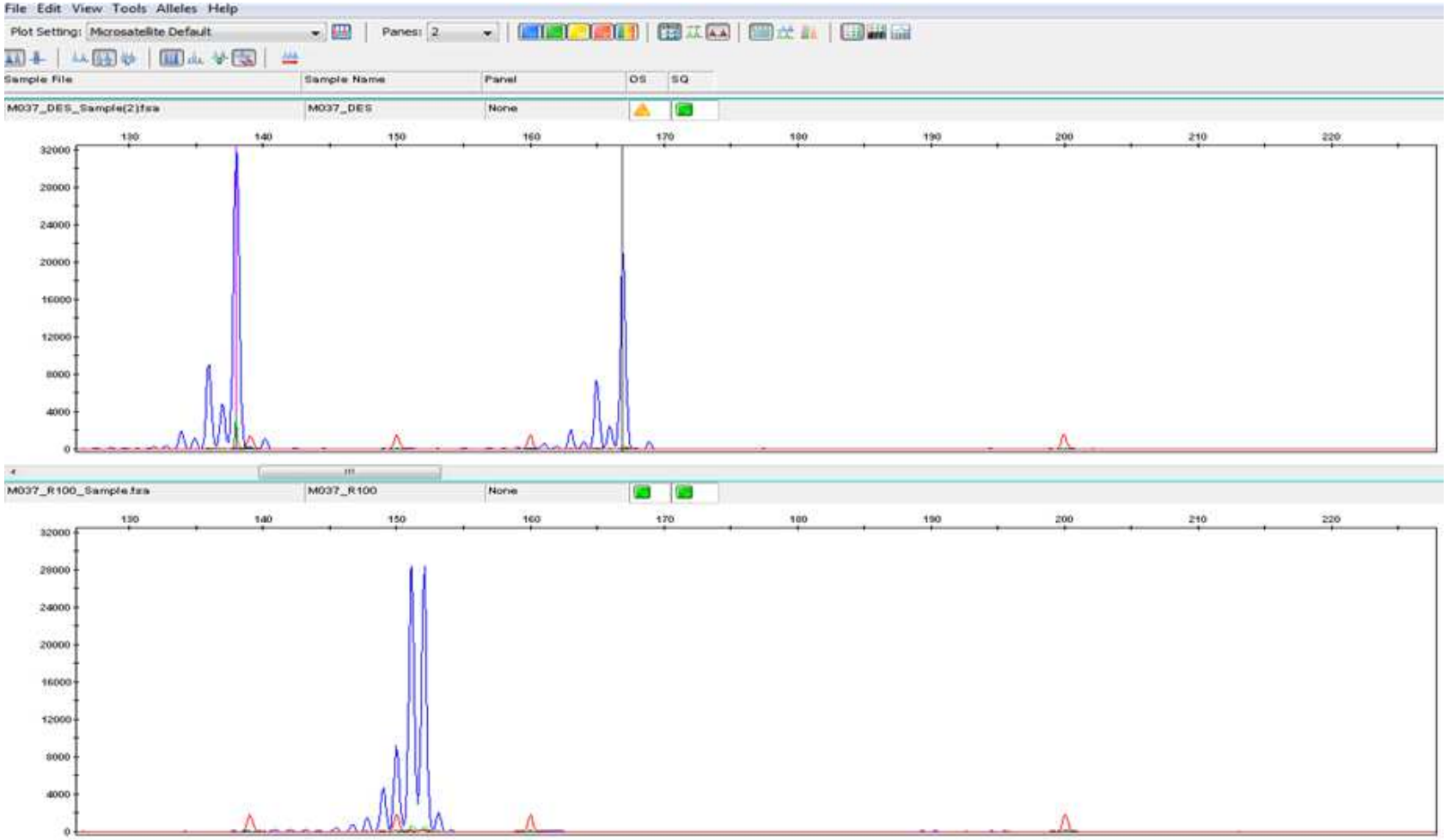


Fig. 31 Example of capillary electrophoresis of the SSR marker BPPCT037, performed by means of the ABI 3550 Genetic Analyzer (Applied Biosystems). In the upper window, peaks corresponding to 137 and 166 bp in the genotype Desmayo Largueta are shown; the lower window shows the two peaks of 150 and 151 bp detected in the genotype R1000. Red peaks refer to the molecular weight of the size standard

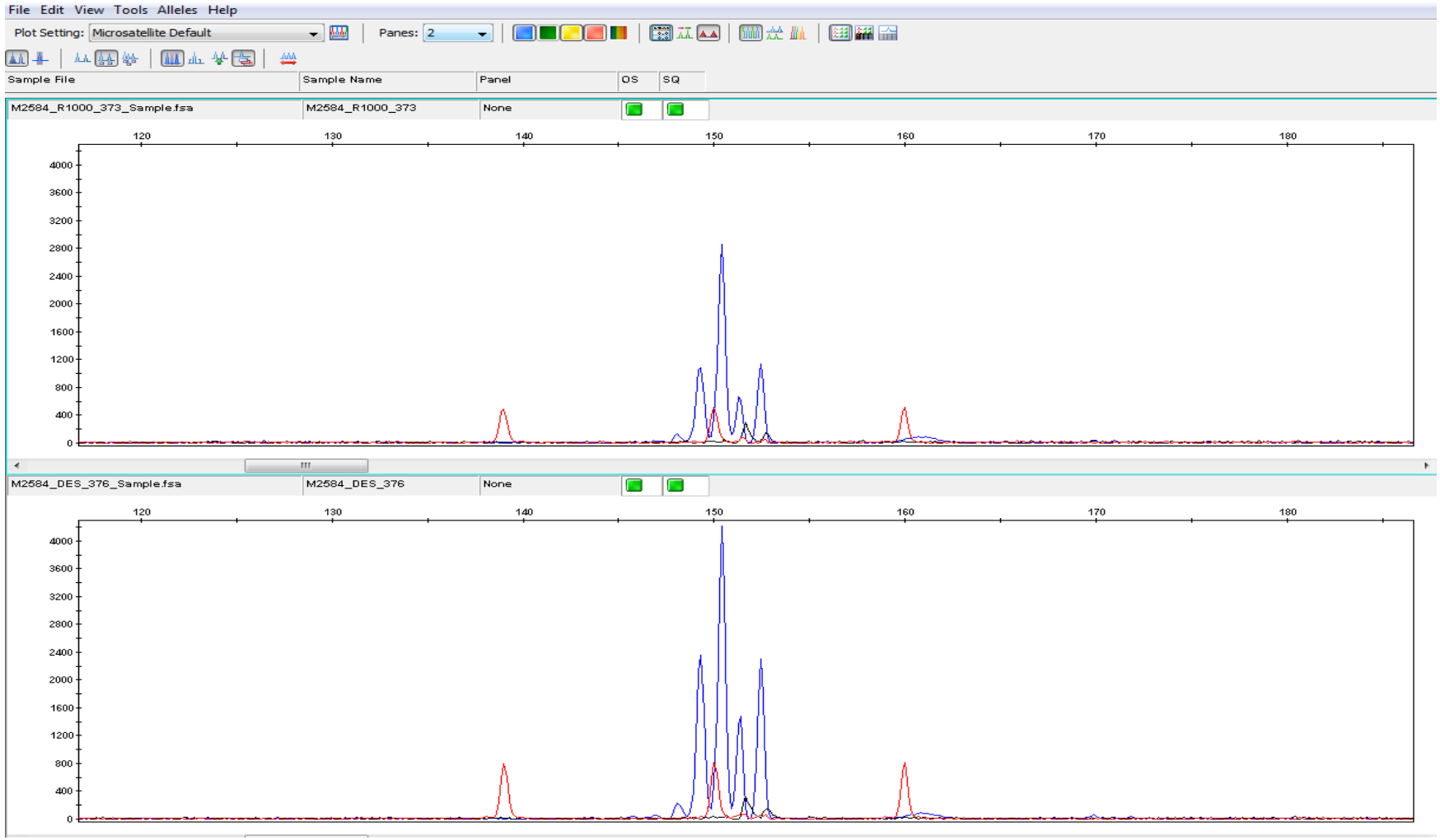


Fig. 32 Example of capillary electrophoresis of the SSR marker EPDCU2584, performed by means of ABI 3550 Genetic Analyzer (Applied Biosystems). In the upper window, peaks corresponding to 150 and 151 bp in the genotype R1000 are shown; the lower window shows the two peaks of 150 and 151 bp detected in the genotype Desmayo Largueta. Red peaks refer to the molecular weight of the size standard

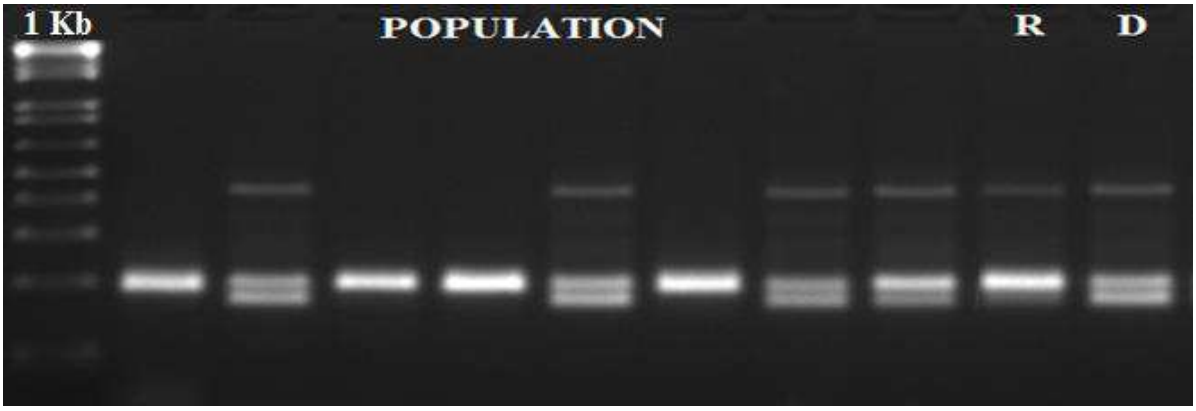


Fig. 33- Electrophoretic pattern obtained using the SSR marker CPDCT028 on the two almond parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F_1 individuals. On the left the DNA ladder 1Kb plus

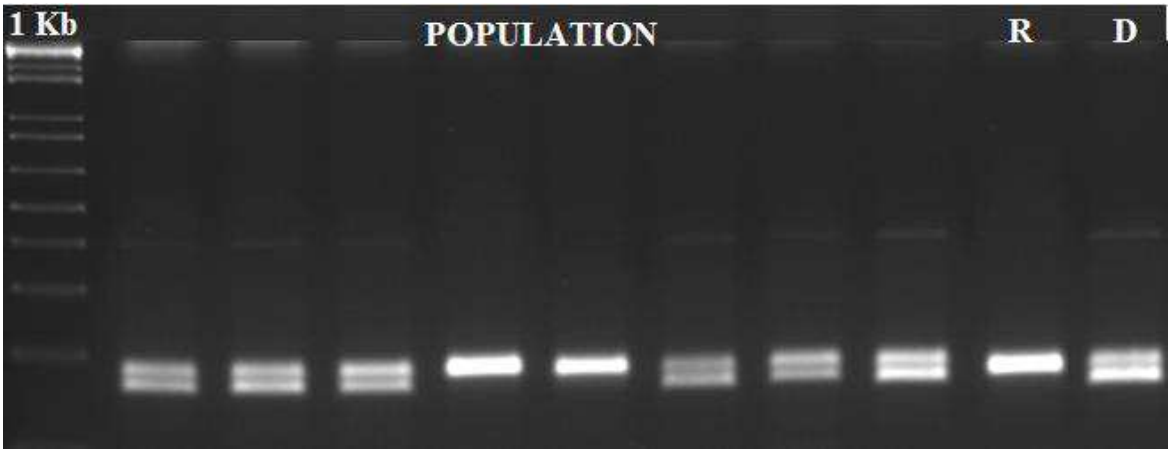


Fig. 34- Electrophoretic pattern obtained using the SSR marker UDA045 on the two parental cultivars "R1000" (R) and "Desmayo Largueta" (D) of almond and 8 F_1 individuals. On the left the DNA ladder 1Kb plus

In order to provide phenotypic data for the *Sk* map, the kernel taste of the (RxD) F_1 population was also assessed by a panel test. This was found to segregate according to a 3:1 (sweet:bitter) ratio, as expected from the two parental genotypes.

The *Sk* fine map was developed using JoinMap 4.1 software, using a mapping algorithm specific for F_1 segregant populations originating from the cross of two heterozygous parental lines. The resulting linkage map is reported in **Fig 35**. Similarly to the previous map reported by Sanchez-Perez *et al.* (2010) (**Fig. 16 and 23**), the two markers UDA045 and CPDCT028 flank the *Sk* locus, and delimit a genomic interval of 5.1 cM. The markers ppa003882m/*BsaW*I,

ppa006282m/*HpyCH4*V, ppa005388m/*Taq*I and ppa001838m/*Hpa*II are positioned within this interval, as well as the SSR markers BPPCT037 and EPDCU2584. The closest markers flanking *Sk* are: on the one side ppa003882m/*BsaW*I and ppa006282m/*HpyCH4V*, which are co-segregating and are distant 0.7 cM; on the other side ppa005388m/*Taq*I, which is distant 0.1 cM. As expected the marker ppa018792m/*Alu*I lies outside the UDA045/CPDCT028 interval, at a distance of 1.7 cM from the *Sk* locus.

Importantly, the almond genetic map developed in this study is fully in accordance with a particular physical region of peach chromosome 5, as reported in **Fig. 36.** Indeed, the order of the markers in the two maps is the same (with the exception of co-segregating markers in the almond genetic map). This supports the notion that the *Prunus* genomes are essentially collinear (Dirlewanger *et al.,* 2004a). Moreover, it suggests that the *Sk* ortholog in peach is located in a physical region of about 82 Kb, delimited by the genes ppa003882m and ppa005388m.

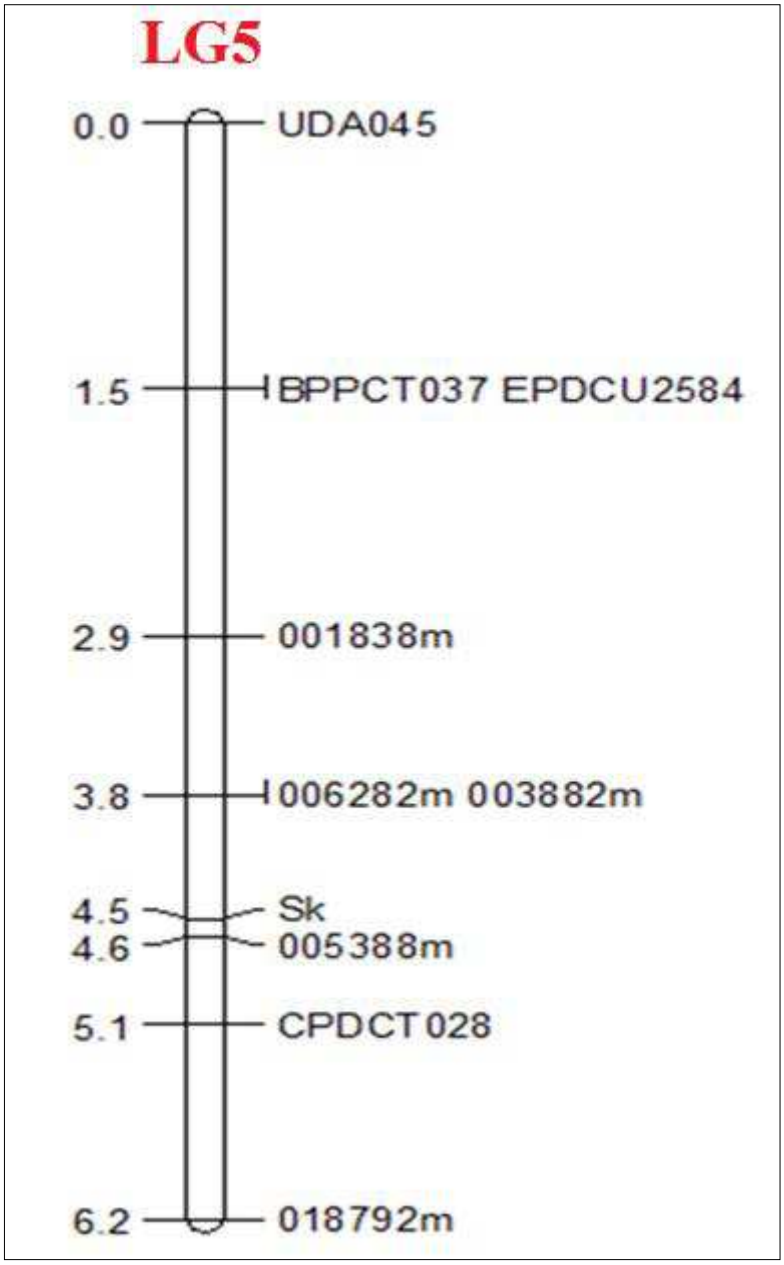


Fig. 35- Fine map of the *Sk* locus genomic region in almond. Genetic distances are expressed in cM

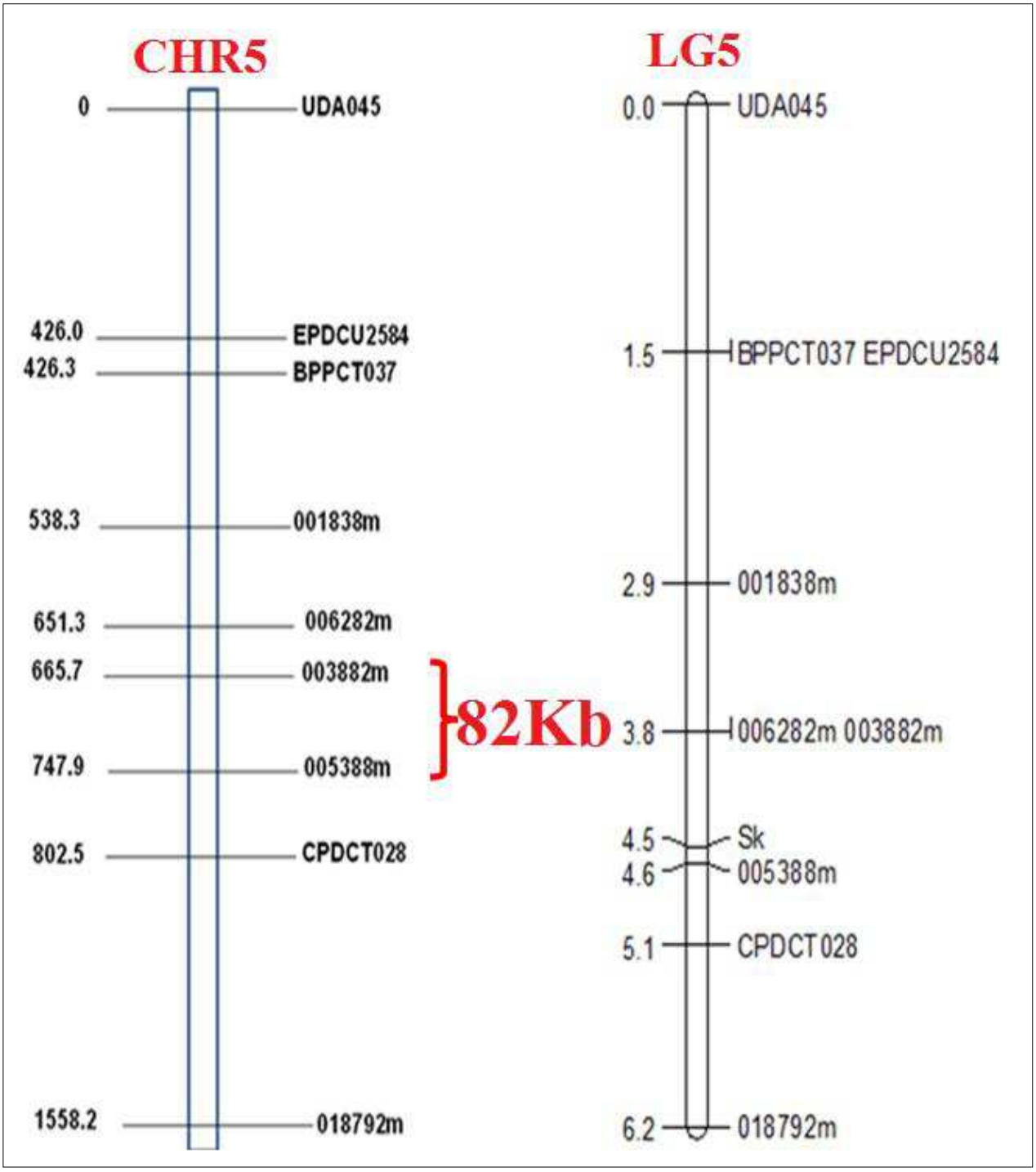


Fig. 36- Collinearity between the physical map of peach chromosome (chr) 5 (left) and the genetic map of almond linkage group (LG) 5 (right) in the *Sk* region. Distances between markers are expressed in Kb in the case of the peach chr5 and in cM in the case of the LG5. The interval of 82 Kb, flanked by the two markers ppa003882m/*BsaW*I and ppa005388m/*Taq*I, and presumably containing the *Sk* ortholog in peach, is highlighted

7.4.1 Identification of candidate Sk ortholog genes in peach

As mentioned in the previous paragraph, based on the results of *Sk* fine mapping in almond and available genomic information on peach, we characterized a peach genomic region, delimited by the two genes ppa003882m and ppa005388m, that contain candidate genes for being the *Sk* ortholog. In total, seven genes are located in this region (**Fig. 37, Tab. 15**).

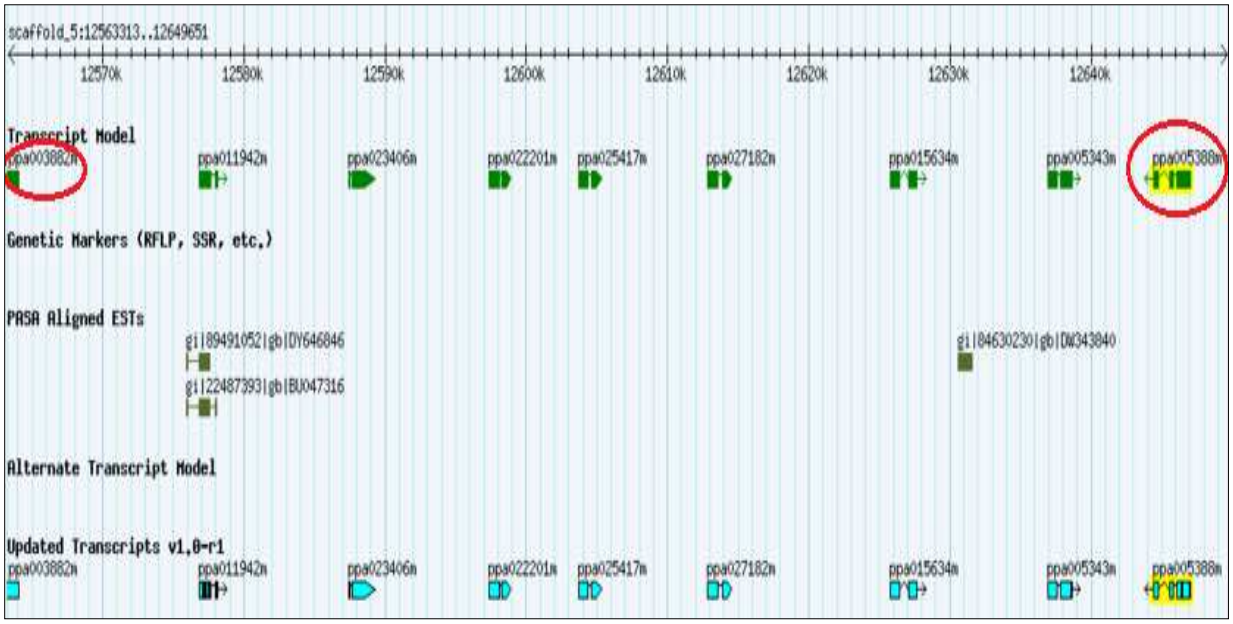


Fig. 37- Detail of the 82 Kb physical region on peach chromosome 5 flanked by the two genes ppa003882m and ppa005388m (circled), visualized by means of the GenomeBrowse tool available at the Genomic Database for Rosaceae (GDR) and presumably containing the *Sk* ortholog.

According to the bioinformatic predictions available at the InterPro and GeneOntology (GO) databases, five of the seven genes included in the ppa003882m/ppa005388m (ppa022201m, ppa025417m**,** ppa027182m**,** ppa015634m and ppa005343m) encode transcription factors of the v-myc avian myelocytomatosis viral oncogene homolog (MYC) family. The gene ppa023406m encodes a putative galactose oxidase, or glyoxal oxidase. Finally, the gene ppa011942m encodes the protein subunit Med10 of the mediator complex.

Table 15- *In silico* functional analysis of seven peach genes included in the ppa003882m/ppa005388m interval, as inferred by the InterPro and GeneOntology (GO) databases

GDR ID	peach physical map position	InterPro function	GO assignment function
ppa023406m	12.587.426	Glyoxal oxidase, N-terminal lactose oxidase/kelch, beta-propeller Immunoglobulin-like fold Immunoglobulin E-set Domain of unknown function DUF1929 Galactose oxidase, beta-propeller	Not available
ppa022201m	12.597.330	Helix-loop-helix DNA-binding Transcription factor MYC/MYB	Not available
ppa025417m	12.603.688	Helix-loop-helix DNA-binding Transcription factor MYC/MYB	Not available
ppa027182m	12.612.821	Helix-loop-helix DNA-binding Transcription factor MYC/MYB	Not available
ppa015634m	12.625.785	Helix-loop-helix DNA-binding Transcription factor MYC/MYB	Not available
ppa005343m	12.636.946	Helix-loop-helix DNA-binding Transcription factor MYC/MYB	Not available
ppa011942m	12.576.856	Mediator complex, subunit Med10	- biological_process: a) regulation of transcription RNA polymerase II promoter - cellular_component: a)mediator complex - molecular_function: a) RNA polymerase II transcription cofactor activity

7.4.2 Expression analysis of candidate genes to be the Sk ortholog in almond

Aiming to provide evidence for the involvement of one of the seven genes above mentioned in amygdalin accumulation/degradation, we monitored their expression at five different stages of kernel development, in sweet and bitter almond genotypes (Lauranne and S3067, respectively).

The tegument tissue was taken into account as it was previously shown that prunasin, the amygdalin precursor, accumulates in the tegument of bitter almonds, whereas it is degraded in the tegument of sweet almond.

The expression of the genes ppa022201m, ppa023406m and ppa011942m was verified in the tegument, as primer pairs designed on these genes allowed the amplification of a single product of the expected size when used for a RT-qPCR assay. Conversely, no amplification was obtained when using primer pairs designed on the genes ppa025417m**,** ppa027182m**,** ppa015634m and ppa005343m.

Relative expression levels of the genes ppa022201m and ppa023406m are reported in **Figg**. **38 and 39.** In more detail, the gene ppa0022201m was found to be more expressed in Lauranne than in S3067 at all the time points under investigation, although differences between were significant only for the first time point (**Fig. 38**).

Thus, also the gene ppa0023406m, resulted to be more expressed in Lauranne tegument in all the time points under investigation (**Fig. 39**). However, again differences did not result to be significant.

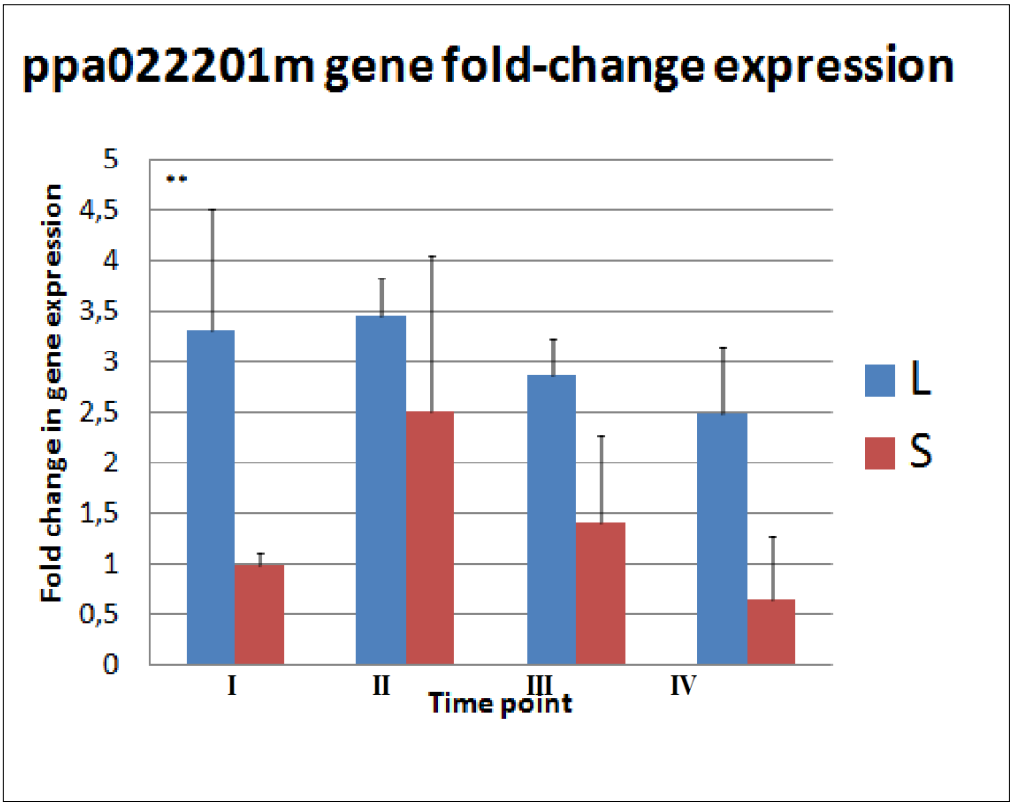


Fig. 38 - Relative fold change in the expression of the gene ppa022201m in the tegument of the sweet genotype Lauranne and the bitter genotype S3067. Bars indicate standard error of four biological replicates. ** indicate significant differences ($p<0.01$)

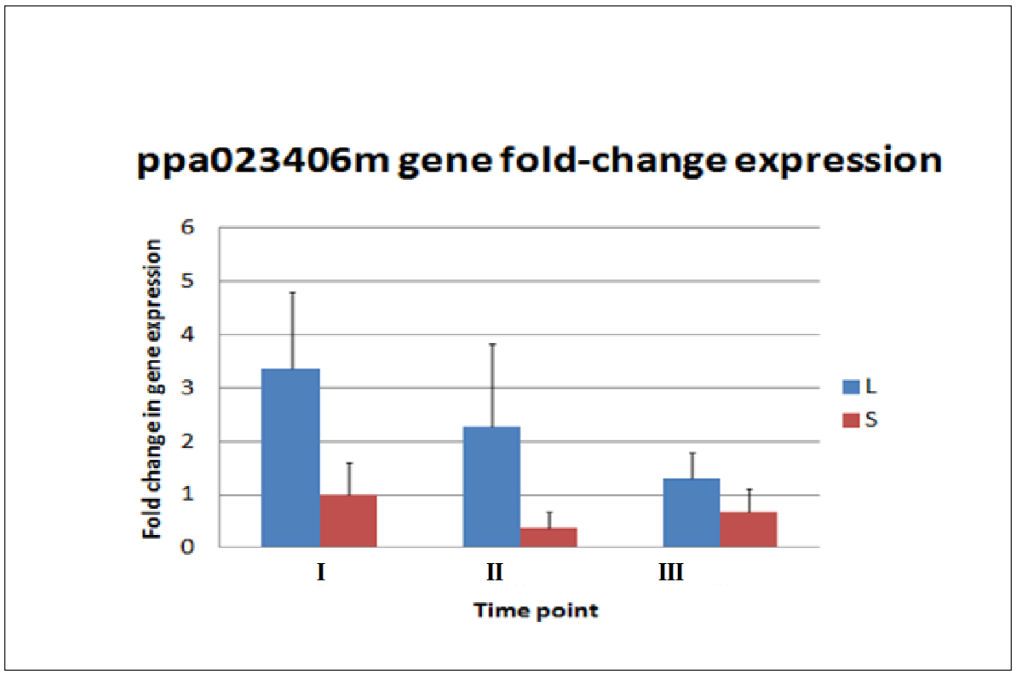


Fig. 39- Relative fold change in the expression of the gene ppa022406m in the tegument of the sweet genotype Lauranne and the bitter genotype S3067. Bars indicate standard error of four biological replicates.

DISCUSSIONS

8. DISCUSSION

In this thesis, we exploited peach genomic data, sequence similarity between *Prunus* genomes and available scientific literature to develop new molecular markers linked to the *Sk* locus, determining the kernel taste in almond, and provide a fine map of the *Sk* region. Furthermore, we performed molecular and genetic analyses to carry out inference on possible *Sk* candidate genes.

In detail, a bioinformatic search allowed to map sequences of almond markers previously shown to flank the *Sk* locus (Sanchez-Perez *et al.,* 2010) on the peach physical map. Peach genes located within this interval were used as a basis to develop five new molecular markers linked to the *Sk* locus. These are expected to be important for marker assisted selection (MAS) in future almond breeding programs. MAS for the sweet kernel taste is particularly desirable in almond, as:

1) the selection of sweet genotypes is a basic requirement of any breeding program;

2) it might avoid drawbacks associated with traditional phenotypic selection, which is expensive and time-consuming because of the long almond juvenile period and the large plant size.

Moreover, all the new markers developed in this work belong to the class of CAPS, which, being co-dominant, allow the discrimination of heterozygous individuals. Another advantage of CAPS markers is that they do not require expensive equipment and thus can be obtained in small laboratories (Pavan *et al*., 2014).

As mentioned above, another important objective of this thesis was to provide a fine map of the *Sk* locus. To this aim, we used the newly developed CAPS and four previously reported *Sk*-linked SSR markers (Sanchez-Perez *et al.* 2010) to genotype a large F_1 segregant population of 476 individuals, including the 167 individuals used by Sanchez-Perez *et al.* (2010) to obtain the first *Sk* genetic map.

Differently from the approach followed by Sanchez-Perez *et al.* (2010), the two *Sk*-linked SSR markers EPDCU2584 and BPPCT037 were scored by capillary electrophoresis, rather than acrylamide gel electrophoresis, which allows to discriminate fragments differing for a single base pair. This ultimately allowed to detect a few scoring errors in the genotyping and the thus mapping performed by Sanchez-Perez *et al.* (2010). Accuracy in genotyping and larger population size together can explain differences in order and distance between loci of

the map reported in this thesis with respect to the two *Sk* linkage maps reported by Sanchez-Perez *et al.* (2010) (**Figg. 16 and 35**)

As shown in **Fig. 35**, our *Sk* fine-map, delimited by the two SSR markers UDA045 and CPDCT028, spans a genetic interval of 5.1 cM. The closest markers flanking the *Sk* locus are, at one side ppa003882m/*Bsaw*I and ppa006282m/*HpyCH4V* and, on the other side, ppa005388m/*Taq*I, delimiting an interval of 0.8 cM.

Notably, the order of markers in our genetic map is fully in accordance with the order of corresponding sequences in the peach physical map. This finding substantiates previous reports suggesting high level of synteny and collinearity between *Prunus* genomes (Dirlewanger *et al.,* 2004a, Arus *et al.,* 2005 and Chen *et al.,* 2008). Collinearity between peach and almond also indicates that the peach ortholog of the *Sk* gene is included in the physical interval between the genes ppa003882m and ppa005388m. This spans about 82 Kb and contains seven candidate genes **(Figg. 36 and 37).**

Interestingly, none of them encodes for a prunasin hydrolase or an enzyme with β-glucosidase activity. This is somehow surprising, as previous studies of Sanchez-Perez *et al.* (2012) hypothesized that difference between sweet and bitter genotypes might be due to a different prunasin hydrolase activity, converting prunasin in mandelonitrile with the release of glucose. Indeed, these authors found that, in sweet genotypes, the inner epidermis in the tegument facing the nucellus is rich in cytoplasmic and vacuolar localized β-glucosidase activity, differently from bitter genotypes.

Five of the seven genes present in the peach ppa003882m/ppa005388m genomic interval (ppa022201m, ppa025417m**,** ppa027182m**,** ppa015634m and ppa005343m) encode for MYC/MYB N-terminal transcription factors. Several studies indicate that transcription factors are important in regulating the biosynthesis of plant secondary me*Tablele*olites, such as amygdalin. For example, in Rosaceae, a R2R3 MYB transcription factor regulates the anthocyanin biosynthetic pathway (Wang *et al.,* 2010). Therefore, the *Sk* otholog could be similarly involved in the biosynthesis of prunasin and allelic variation might be responsible for different prunasin/amygdalin accumulation.

The seven genes included in the ppa003882m/ppa005388m genomic interval were further studied at the functional level, by monitoring their expression in the tegument tissue of sweet and bitter genotypes. Interestingly, no amplification product was obtained for four genes (ppa025417m**,** ppa027182m**,** ppa015634m and ppa005343m) suggesting that probably none of them is expressed in the tegument and thus is involved in amygdalin biosynthesis/degradation.

On the contrary, the two genes ppa022201m and ppa022406m resulted to be more expressed in the sweet genotype at several time points under investigation, although differences were not always significant. Higher expression of the MYB transcription factor ppa022201m further substantiates the hypothesis above discussed suggesting that the Sk gene could be a transcriptional regulator of amygdalin biosynthesis.

Further studies on gene expression are ongoing and are will be necessary to gain insights on how the *Sk* gene influences amygdalin accumulation and thus the typical bitter taste of almond seeds.

CONCLUSIONS

9. CONCLUSIONS AND FUTURE PERSPECTIVES

Results obtained during the three years of this PhD in "Biotechnology of Food Products" at the University of Foggia, which are reported in this thesis, are expected to be of great importance for basic genetic studies on peach and almond. Moreover, they will be useful for almond plant breeding using biotechnological approaches, in particular molecular marker technology. Clearly, this thesis strengthened a national and international cooperation between Institutions (Universities and Research Centres), which will continue in the future with the final aim to isolate the *Sk* gene, a goal which was not possible to achieve despite three years of intense efforts.

As reported in the paragraph 5 of this thesis ("Aims of the work"), main goals of this thesis were the identification of new almond molecular markers linked to the *Sk* locus, determining the kernel taste in almond, and to provide a fine map of the *Sk* genomic region.

Based on sequence similarity between the almond and peach genomes, we developed a series of CAPS markers polymorphic between the two parental genotypes of a large segregant population. This is expected to be of importance to perform selection in future marker-assisted breeding programs, as the sweet kernel taste is one of the primary traits of selection.

Mapping of these markers revealed they are linked to the *Sk* locus and, moreover, showed that the genomes of peach and almond are collinear in the *Sk* region. This result is extremely useful for future studies addressed to gene isolation. Indeed, this indicates that the ortholog of the *Sk* locus in peach likely lies in the interval included between the markers ppa003882m/*Bsaw*I and ppa005388m/*Taq*I, delimiting a narrow physical region of 82 kb and containing only seven genes.

Preliminary functional studies, carried out by means of real-time qPCR, also demonstrated that two genes in this region are expressed at higher level in sweet genotypes. In particular, the gene ppa022201m, encoding a MYC

transcription factor, is a strong candidate for being the *Sk* ortholog in almond, as it was previously shown that these kind of transcriptional regulators often control the biosynthesis of secondary me*Tablele*olites.

As said, aiming at gene isolation, further studies on the topic of this thesis will be continued and will strengthen the collaboration between the Department of Agriculture, Food and Environmental Sciences (S.A.F.E) of the University of Foggia, the Department of Soil, Plant and Food Sciences (Di.S.S.P.A.) of the University of Bari "Aldo Moro", the Department of Plant and Environmental Sciences of the University of Copenhagen and the Department of Plant Breeding, CEBAS-CSIC of the University of Murcia.

One of the first target of this collaboration and efforts will be a *de novo* sequencing of the *Sk* genomic region in sweet and bitter almond genotypes, in order to detect gene polymorphisms putatively associated with the kernel phenotype.

REFERENCES

REFERENCES

Abbey M., Noakes M., Belling G.B. and P.J. Nestel, 1994. Partial replacement of saturated fatty acids with almonds or walnuts lowers total plasma cholesterol and low-density-lipoprotein cholesterol. Am. J. Clinic. Nutrit., 59, pp.:995–999.

Abdallah A., Ahumada M.H. and T.M. Gradziel, 1998. Oil content and fatty acid composition of almond kernels from different genotypes and California production regions. J. Am. Soc. Hort. Sci., 123:1029–1033.

Agrawal A.A., Hastings A.P., Johnson M.T.J., Maron J.L. and J.P. Salminen, 2012. Insect herbivores drive realtime ecological and evolutionary change in plant populations. Science, 338:113–16.

Ahn S. and S.D. Tanksley, 1993. Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci., USA 90: 7980-7984.

Ahn Y.O., Saino H., Mizutani M., Shimizu B. and K. Sakata, 2007. Vicianin hydrolase is a novel cyanogenic β-glycosidase specific to β-vicianoside (6-*O*-α-L-arabinopyranosyl-β-D-glucopyranoside) in seeds of *Vicia angustifolia*. Plant Cell Physiol., 48:938–39.

Akparov Z. and E. Khidirova, 2006. Sulle orme del mandorlo. Scriptae Hort., 4:24-28.

Alessandroni A., 1980. Le mandorle. Panificazione Pasticceria, 8: 67–71.

Andersen M.D. Busk P.K., Svendsen I. and B.L. Møller, 2000. Cytochromes P450 from cassava (*Manihot esculenta* Crantz) catalyzing the first steps in the biosynthesis of the cyanogenic glucosides linamarin and lotaustralin: cloning, functional expression in *Pichia pastoris* and substrate specificity of the isolated recombinant enzymes. J. Biol. Chem., 275:1966–75.

Aranzana M.J., García-Mas J., CarbóJ. and P. Arús, 2002. Development and variability analysis of microsatellite markers in peach. Plant Breed., 121:87–92.

Aranzana M.J., Pineda A., Cosson P., Dirlewanger E., Ascasibar J., Cipriani G., Ryder C.D.,Testolin R., Abbott A., King G.J., Iezzoni A.F. and P. Arús, 2003. A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theor. Appl. Genet., 106:819–825.

Ardejani F.D., Badii K., Limaee N.Y., Shafaei S.Z. and A.R. Mirhabibi, 2007. Adsorption of Direct Red 80 dye from aqueous solution onto almond shells: effect of pH, initial concentration and shell type. Journal of Hazardous Materials, 151:730–737.

Arràzola G., 2002. Anàlisis de glucòsidos cianogénicos en variedades de almendro: implicaciones en la mejora genética. PhD thesis. Universidad de Alicante, Alicante, Spain.

Arumuganathan K. and E.D. Earle, 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep., 9:208–218.

Arús P., Messeguer R., Viruel M.A., Tobutt K., Dirlewanger E., Santi F., Quarta R. and E. Ritter, 1994a. The European *Prunus* mapping project. Euphytica, 77:97–100.

Arús P., Olarte C., Romero M. and F.J., Vargas, 1994b. Linkage analysis of tenisozymegenes in F_1 segregating progenies of almond. J. Am. Soc. Hort. Sci., 119:339–344.

Arús P., Ballester J., Jáuregui B., Joobeur T., Truco M.J.and M.C. de Vicente, 1999. The European *Prunus* mapping project: update of marker development in almond. Acta Hort., 484:331–336.

Arús P., Yamamoto T., Dirlewanger E. and A.G. Abbott, 2005. Synteny in the Rosaceae. Plant Breed. Rev., 27:175–211.

Bacarella A., Chironi G. and G. Barbera, 1991. Aspetti tecnici, economici e di mercato del mandorlo in Sicilia. Quarderni di Ricerca di Sperimentazione, (Palermo, Sicily), 40: 1–191.

Baird W.V., Estager A.S. and J.K. Wells, 1994. Estimating nuclear DNA content in peach and related diploid species using lasser flow-cytometry and DNA hybridization. J. Am. Soc. Hort. Sci., 119:1312–1316.

Bak S., Kahn R.A., Nielsen H.N., Møller B.L. and B.A. Halkier, 1998. Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor* (L.) Moench by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside dhurrin. Plant Mol. Biol., 36:393–405.

Bak S., Paquette S.M., Morant M., Rasmussen A.V., Saito S., Bjarnholt N., Zagrobelny M., Jørgensen, K., Hamann T. and S. Osmani, 2006. Cyanogenic glycosides: a case study for evolution and application of cytochromes p450. Phytochem. Rev., 5:309–329.

Bak S., Beisson F., Bishop G., Hamberger B., Höfer R., Paquette S. and D. Werck-Reichhart, 2011. Cytochromes P450. Arabidopsis Book 9:e0144.

Ballester J., Bošković R., Batlle I., Arús P., Vargas F. and M.C.de Vicente, 1998. Location of the self-incompatibility gene on the almond linkage map. Plant Breed., 117:69–72.

Ballester J., Socias i Company R., Arús P. and M.C.de Vicente, 2001. Genetic mapping of a major gene delaying blooming time in almond. Plant Breed., 120:268–270.

Ballhorn D.J., Kautz S., Lion U. and M. Heil, 2008a. Trade-offs between direct and indirect defences of lima bean (*Phaseolus lunatus*). J. Ecol., 96:971–80.

Ballhorn D.J., Schiwy S., Jensen M. and M. Heil, 2008b. Quantitative variability of direct chemical defense in primary and secondary leaves of lima bean (*Phaseolus lunatus*) and consequences for a natural herbivore. J. Chem. Ecol., 34:1298–301.

Bansode R.R., Losso J.N., Marshall W.E., Rao R.M. and R.J. Portier, 2003. Adsorption of volatile organic compounds by pecan shell and almond shellbased granular activated carbons. Biores. Technol., 90:175–184.

Barcaccia G., Lucchin M. and P. Parrini, 2000. Analisi del genoma mediante marcatori molecolari: i fondamenti metodologici. Sementi elette, 5:5-15.

Barckley K., Uratsu S.L, Gradziel T.M. and A.M. Dandekar, 2006. Multidimensional analysis of S-alleles from cross-incompatible groups of California almond cultivars. J. Am. Soc. Hort. Sci., 131:632–636.

Barros P.M., Gonçalves N., Saibo N.J.M. and M.M. Oliveira, 2012. Cold acclimation and floral development in almond bud break: insights into the regulatory pathways. J. Exp. Bot., 63:4585–96.

Bell C.J. and J.R. Ecker, 1994. Assignment of 30 microsatellite loci to the linkage map of Arabidopsis. Genomics, 19:137-144.

Bjarnholt N., Laegdsmand M., Hansen H.C., Jacobsen OH. and B.L. Møller, 2008. Leaching of cyanogenic glucosides and cyanide from white clover green manure. Chemosph., 72:897–904.

Bliss F.A., Arulsekar S., Foolad M.R., Becerra A.M., Gillen A., Warburton M.L., Dandekar A.M., Kocsisne G.M. and K.K. Mydin, 2002. An expanded genetic linkage map of *Prunus* based on and interspecific cross between almond and peach. Genome, 45, 520–529.

Blumenthal S.G., Hendrickson H.R., Abrol Y.P. and E.E. Conn, 1968. Cyanide metabolism in higher plants. III, The biosynthesis of β-cyanoalanine. J. Biol. Chem., 243:5302–7.

Bošković R., Tobutt KR., Batlle I., Duval H., Martínez-Gómez P. and T.M. Gradziel, 2003. Stylar ribonucleases in almond: correlation with and prediction of self-incompatibility genotypes. Plant Breed., 122:70–76.

Botstein D, White R.L, Skolnick M. and R.W. Davis, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet., 32:314–331.

Browicz K. and D. Zohary, 1996. The genus *Amygdalus* L. (*Rosaceae*): species relationships, distribution and evolution under domestication. Genet. Resour. Crop Evol., 43:229–247.

Bulut Y. and Z. Tez, 2003. Removal of heavy metal ions by modified sawdust ofmwalnut. Fresenius Environmental Bulletin, 12:1499–1504.

Burr B., Burr F.A., Thompson K.H., Albertson M.C. and C.W. Stuber, 1988. Gene mapping with recombinant inbreds in maize. Genetics, 118: 519-526.

Cantini C., Iezzoni A.F., Lamboy W.F., Boritzki M. and D. Struss, 2001. DNA fingerprinting of tetraploid cherry germplasm using SSR. J. Am. Soc. Hort. Sci., 126:205–209.

Chandler W.H., 1957. “Deciduous orchards”. Lea and Febiger, Philadelphia

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

Chassagne D., Crouzet J.C., Bayonove C.L. and R.L. Baumes, 1996. Identification and quantification of passion fruit cyanogenic glycosides. J. Agric. Food Chem., 44:3817–20.

Chen C.Y., Milbury P.E., Lapsley K. and J.B. Blumberg, 2005. Flavonoids from almond skins are bioavailable and act synergistically with vitamins C and E to enhance hamster and human LDL resistance to oxidation. J. Nutr., 135:1366–1373.

Chen L., Zhang S., Illa E., Song L.,Wu S., Howad W., Arús P.,Weg E., Chen K. and Z. Gao, 2008. Genomic characterization of putative allergen genes in peach/almond and their synteny with apple. BMC Genomics, 9:543.

Cicek M. and A. Esen, 1998. Structure and expression of a dhurrinase (β-glucosidase) from *Sorghum*. Plant Physiol., 116:1469–78.

Cipriani G., Lot G., Huang H.G., Marrazzo M.T., Peterlunger E. and R. Testolin, 1999. AC/GT and AG/CT microsatellite repeats in peach (*Prunus persica* (L) Basch): Isolation, characterization and cross-species amplification in *Prunus*. Theor. Appl. Genet., 99:65–72.

Clausen V., Frydenvang K., Koopmann R., Jørgensen L.B., Abbiw D.K., Ekpe P. and J.W. Jaroszewski, 2002. Plant analysis by butterflies: occurrence of cyclopentenylglycines in Passifloraceae, Flacourtiaceae, and Turneraceae and discovery of the novel nonproteinogenic amino acid 2-(3_-cyclopentenyl)glycine in Rinorea. J. Nat. Prod., 65:542–47.

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B. and E.C.K. Pang, 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica, 142:169-196.

Conn E.E., 1969. Cyanogenic glycosides. J. Agric. Food Chem., 17(3):519-526.

Conn E.E., 1980. Cyanogenic compounds. Annu. Rev. Plant Physiol., 31:433–51.

Conn E.E., 1991. The metabolism of a natural product: lessons learned from cyanogenic glycosides. Planta Med., 57:1–9.

Cordeiro V. and A. Monteiro, 2001. Almond growing in Trasos-Montes region (Portugal). Acta Hortic., 591:161-165.

Corredor E., Román M., García E., Perera E., Arús P. and T. Naranjo, 2004. Physical mapping of rDNAgenes enables to establish the karyotype of almond. Ann. Appl. Biol., 144:219–222.

Cutler A.J. and E.E. Conn, 1981. The biosynthesis of cyanogenic glucosides in *Linum usitatissimum* (linen flax) in vitro. Arch. Biochem. Biophys., 212:468–474.

Daifullah A.A.M. and B.S. Girgis, 2003. Impact of surface characteristics of activated carbon on adsorption of BTEX. Colloids and Surface A – Physicochem. Engineer. Aspects, 214:181–193.

Davis R.H., 1991. Cyanogens. In: “Toxic Substances in Crop Plants”. (J.P.F. D’Mello, C.M. Duffus, J.H. Duffus, Eds), Royal Society of Chemistry, Cambridge, UK, 202–225.

Decroocq V., Favé M.G., Hagen L., Bordenave L. and S. Decroocq, 2003. Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor. Appl. Genet., 106:912–922.

Denisov V.P., 1988. Almond genetic resources in the USSR and their use in production and breeding. Acta Hort., 224:299–306.

Dettori M.T., Quarta R. and I. Verde, 2001. A peach linkage map integrating RFLPs, SSRs, RAPDs, and morphological markers. Genome, 44:783–790.

Dharmawardhana D.P., Ellis B.E. and J.E. Carlson, 1995. A beta-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. Plant Physiol., 107:331-339.

Dicenta F. and J.E. García, 1993a. Inheritance of self-compatibility in almond. Heredity, 70:313–317.

Dicenta F. and J.E. García, 1993b. Inheritance of the kernel flavor in almond. Heredity, 70:308–312.

Dicenta F., García J.E. and E. Carbonell, 1993. Heritability of flowering, productivity and maturity in almond. J. Hort. Sci., 68:113–120.

Dicenta F., Berenguer T. and J. Egea, 1999. 'Antoñeta' y 'Marta': Dos nuevas variedades de almendro autocompatibles y de floración tardía. In: Fruticultura Profesional, Especial Frutos Secos II, 104:48- 53.

Dicenta, F., Martìnez-Gòmez, P., Ortega E. and H. Duval, 2000. Cultivar pollinizer does not affect almond flavour. Hort. Sci. 35(6):1153–1154.

Dicenta F., Ortega E., Cánovas J.A. and J. Egea, 2001. Self-pollination versus cross-pollination of six self-compatible almond cultivars: Pollen tube growth and fruit set. Cah. Options Méditerr., 56:369–372.

Dicenta F., Martìnez-Gòmez P., Granè N., Martìn M.L., Leòn A., Cànovas J.A. and V. Berenguer, 2002. Relationship between cyanogenic compounds in kernels, leaves, and roots of sweet and bitter kernelled almonds. J. Agric. Food Chem., 50:2149–2152.

Dicenta F., Ortega E. and P. Martinez-Gomez, 2007. Use of recessive homozygous genotypes to assess the genetic control of kernel bitterness in almond. Euphytica, 153:221–225.

Dirlewanger E., Crosson A., Tavaud P., Aranzana M.J., Poizat C., Zanetto A., Arús P. and L. Laigret 2002. Development of microsatellite markers in peach and their use in genetic diversity analysis in peach and sweet cherry. Theor. Appl. Genet., 105:127–138.

Dirlewanger E., Cosson P., Howad W., Capdeville, G., Bosselut N., Claverie M., Voisin R., Poizat C., Lafargue B. and O. Baron, 2004a. Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid location of root-knot nematode resistance genes. Theor. Appl. Genet. 109 (4): 827–838.

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

Doerge R., 2002. Mapping and analysis of quantitative trait loci in experimental populations. Nature Rev., 3:43-52.

Dominguez I., Graziano E., Gebhardt C., Barakat A., Berry S., Arús P., Delseny M. and S. Barnes, 2003. Plant genome archaeology: evidence for conserved ancestral chromosome segments in dicotyledonous plant species. Plant Biotechnol. J., 1:91–99.

Donovan J.L., Meyer A.S.and A.L. Waterhouse, 1998. Phenolic composition andantioxidant activity of prunes and prune juice (*Prunus domestica*). Journal of Agriculture and Food Chemistry, 46:1247–1252.

Downey L.D. and A.F. Iezzoni, 2000. Polymorphic DNA markers in black cherry are identified using sequences from sweet cherry, peach, and sour cherry. J. Am. Soc. Hort. Sci., 125:76–80.

Doyle J.J. and J.L. Doyle, 1990. Isolation of plant DNA from fresh tissue. Focus, 12:13-15.

Dreher M.L., Maher C.V. and P. Kearney, 1996. The traditional and emerging role of nuts in healthful diets. Nutrition Reviews, 54:241–245.

El Gharbi A., 1981. Résultats préliminaires des croisements intervariétaux d'amandier réalisés en Tunisie. Opt. Med. Serie Etudes, 8 1/1, 23-35.

Elias M, Nambisan B and P.R. Sudhakaran, 1997. Characterization of linamarase of latex and its localization in petioles in cassava. Arch. Biochem. Biophys., 341:222–28.

Ensminger A., 1994. Foods & Nutrition. Encyclopedia, 2nd Edition. United States of America: CRC Press, pp: 1167.

Escamilla-Trevino L.L., Chen W., Card M.L., Shih M.C., Cheng C.L. and J.E. Poulton, 2006. Arabidopsis thaliana beta-glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. Phytochem., 67:1651–1660.

Esfahlan A.J., Jamei R. and R.J. Esfahlan, 2010. The importance of almond (Prunus amygdalus L.) and its by-products. Food Chem., 120:349–360.

Estevinho B.N., Ratola N., Alves A. and L. Santos, 2006. Pentachlorophenol removal from aqueous matrices by sorption with almond shell residues. J. Hazard. Mat., 137:1175–1181.

Etienne C., Rothan C., Moing A., Plomion C., Bodenes C., Svanella-Dumas L., Cosson P., Pronier V., Monet R. and E. Dirlewanger, 2002. Candidate genes and QTLs for sugar and organic content in peach [*Prunus persica* (L.) Batsch]. Theor. Appl. Genet., 105:145–159.

Evreinoff V.A., 1952. Quelques observations biologiques sur l'Amandier. In: "Revue internationale de botanique appliquée et d'agriculture tropicale", 32e année, bull. n° 359-360, Septembre-Octobre, 1952,. pp.: 442-459.

Fadel J.G., 1999. Quantitative analyses of selected plant by-product feedstuffs, a global perspective. Animal Feed Science and Technology, 79:255–268.

FAOSAT DATABASE. Food and Agriculture Organization of the United Nations, 129, 1252–1264.

Felipe A.J., 1975. F_1 hybrids of peach and almond trees as a model for both species. Agricultura, 44:661–663.

Felipe A.J., 1977. Almendro. Estados fenol´ogicos. Inf. Tecn. Econ. Agr., 27:8–9.

Ferreira A., da Silva M.F., da Costa e Silva L. and C.D. Cruz, 2006. Estimating the effects of population size and type on the accuracy of genetic maps. Genet. Mol. Biol., 29:182-192.

Fia G., Giovani G. and I. Rosi, 2005. Study of beta-glucosidase production by wine related yeasts during alcoholic fermentation. A new rapid fluorimetric method to determine enzymatic activity. J. Appl. Microbiol., 99: 509-517.

Finnie J.W., Windsor P.A. and A.E. Kessell, 2011. Neurological diseases of ruminant livestock in Australia. II. Toxic disorders and nutritional deficiencies. Aust. Vet. J., 89:247–53.

Flematti G.R., Waters M.T., Scaffidi A., Merritt D.J., Ghisalberti E.L., Dixon K.W., Smith S.M.. 2013. Karrikin and cyanohydrin smoke signals provide clues to new endogenous plant signaling compounds. Mol Plant., 6(1):29-37.

Floss H.G., Hadwiger L. and E.E. Conn, 1965. Enzymatic formation of b-cyanoalanine from cyanide. Nature, 208: 207–1208.

Foolad M.R., Arulsekar S., Becerra V. and F.A. Bliss, 1995. A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theor. Appl. Genet., 91:262–269.

Forslund K., Morant M., Jørgensen B., Olsen C.E., Asamizu E., Sato S., Tabata S. and S. Bak, 2004. Biosynthesis of the nitrile glucosides rhodiocyanoside A and D and the cyanogenic glucosides lotaustralin and linamarin in *Lotus japonicus*. Plant Physiol., 135:71–84.

Foulongne M., Pascal T., Arùs P.and J. Kervella, 2003. The potential Prunus davidiana for introgression into peach [*Prunus persica* (L.) Batsch] assessed by comparative mapping. Theor. Appl. Genet., 107:227-238.

Franks T.K., Yadollahi A., Wirthensohn M.G., Guerin J.R., Kaiser B.N., Sedgley M., and C.M. Ford, 2008. A seed coat cyanohydrin glucosyltransferase is associated with bitterness in almond (*Prunus dulcis*) kernels. Funct. Plant Biol., 35:236–46.

Fraser G.E., Bennett H.W., Jaceldo K.B. and J. Sabate, 2002. Effect on body weight of a free 76 kilojoule (320 calorie) daily supplement of almonds for six months. J. Am. Coll. Nutr., 21:275–283.

Frehner M., Scalet M. and E.E. Conn, 1990. Pattern of the cyanide-potential in developing fruits. Plant Physiol., 94:28–34.

Frehner M., Luscher A., Hebeisen T., Zanetti S., Schubiger F. and M. Scalet, 1997. Effects of elevated partial pressure of carbon dioxide and season of the year on forage quality and cyanide concentration of *Trifolium repens* L. from a FACE experiment. Acta Oecologica, 18:297–304.

Frey K.J., 1981. Cababilities and limitations of conventional plant breeding. In: "Genetic engineering for crop improvement". Rakye and Lyman (Eds), pp:15-63.

Frison S. and P. Sporns, 2002. Variation in the flavonol glycoside composition of almond seedcoats as determined by MALDI-TOF mass spectrometry. Journal of Agriculture and Food Chemistry, 50:6818–6822.

Fulgoni V.L., Abbey M., Davis P., Jenkins D., Lovejoy J., Most M., Sabate J. and G. Spiller, 2002. Almonds lower blood cholesterol and LDL-cholesterol but not HDL-cholesterol in human subjects: results of a meta-analysis. FASEB J., 16:A981–A982.

García-López C., Grané-Teruel N., Berenguer-Navarro V., García- García J.E. and M.L. Martín-Carratalá, 1996. Major fatty acid composition of 19 almond cultivars of different origins. A chemometric approach. J. Agri. Food Chem., 44:1751–1755.

Georgi L.L.,Wang Y., Yvergiaux D., Ormsbee T., Iñigo M., Reighard G. and A.G. Abbott, 2002. Construction of a BAC library and its application to the identification of simple sequence repeats in peach [*Prunus persica* (L) Batsch]. Theor. Appl. Genet., 105:1151–1158.

Georgi L.L.,Wang L., Reighard G.L., Mao L.,Wing R.A. and A.G. Abbott, 2003. Comparison of peach and *Arabidopsis* genomic sequences: fragmentary conservation of gene neigborhoods. Genome, 46:268–276.

Gill K.S., Gill B.S., Endo T.R. and E.V. Boyko, 1996a. Identification of highdensity mapping of gene-rich regions in chromosome group 5 of wheat. Genetics, 143:1001-1012.

Gill K.S., Gill B.S., Endo T.R. and T. Taylor, 1996b. Identification and highdensity mapping of gene-rich regions in chromosome group 1 of wheat. Genetics, 144:1883-1891.

Gleadow R.M. and I.E. Woodrow, 2000a. Polymorphism in cyanogenic glycoside content and cyanogenic β-glucosidase activity in natural populations of *Eucalyptus cladocalyx*. *Aust.* J. Plant Physiol., 27:693–99.

Gleadow R.M. and I.E. Woodrow, 2000b. Temporal and spatial variation in cyanogenic glycosides in *Eucalyptus cladocalyx*. Tree Physiol., 20:591–98.

Gleadow R.M. and I.E. Woodrow, 2002a. Constraints on effectiveness of cyanogenic glycosides in herbivore defense. J. Chem. Ecol., 28:1301–13.

Gleadow R.M. and I.E. Woodrow, 2002b. Defense chemistry of cyanogenic *Eucalyptus cladocalyx* seedlings is affected by water supply. Tree Physiol., 22:939–45.

Gleadow R.M., Haburjak J., Dunn J.E., Conn M.E. and E.E. Conn, 2008. Frequency and distribution of cyanogenic glycosides in *Eucalyptus* L'H´erit. Phytochem., 69:1870–74.

Gleadow R.M. and B.L. Møller, 2014. Cyanogenic Glycosides: Synthesis, Physiology and Phenotypic Plasticity. Annu. Rev. Plant Biol., 65:155–185.

Godini A., 2002. Almond fruitfulness and role of self-fertility. Acta Hort., 591:191–203.

Gomez P.M., Sánchez-Peréz R., Dicenta F., Howad W., Arùs P. and T.M. Gradziel, 2007. 11 almond. In: Genom mapping and molecular breeding in plants. Vol 4, Fruits and nuts, C. Kole Ed..

Gradziel T.M. and D.E. Kester, 1998. Breeding for self-fertility in California almond cultivars. Acta Hort., 470:109–117.

Gradziel T.M., Martínez-Gómez P., Dicenta F. and D.E. Kester, 2001. The utilization of related almond species for almond variety improvement. J. Am. Pomol. Soc., 55:100–109.

Gradziel T.M., Martínez-Gómez P. and A.M. Dandekar, 2001. The use of S-allele specific PCR analysis to improve breeding efficiency for self-fertility in almond. HortSci., 36:440.

Gradziel T.M., Martınez-Gomez P., Dandekar A., Uratsu S. and E. Ortega, 2002. Multiple genetic factors control self-fertility in almond. Acta Hort., 591:221–227.

Gradziel T.M., 2009. Almond (*Prunus dulcis*) breeding. In: "Breeding of plantation tree crops", S.M. Jain and M. Priyadarshan (Eds.), pp.: 1–31, Springer Science, New York.

Gradziel T.M., 2012. Traditional genetics and breeding. In: "Genetics, genomics and breeding of stone fruits", A.G. Abbott and C. Kole. (Eds.), Science Publications, Plymouth, UK.

Gradziel T.M. and P. Martínez-Gómez, 2012. Almond breeding. Plant Breeding Reviews, First Edition. Jules Janick (Ed.), 37:199-251.

Grasselly C., 1975. L'arboretum amygdalus. II Colloque du groupe de recherche et d'étude meditarranéen pour l'amandier. pp.: 1–2.

Grasselly C., 1976a. Origine et évolution de l'amandier cultivé. Options Méditerr., 32:45–50.

Grasselly C., 1976b. Les espèces sauvages d'amandier. Options Méditerr., 32:28–44.

Grasselly C., Crossa Raynaud P., Olivier G., Gall H., 1981. Transmission du caractère d'autocompatibilité chez l'amandier (*Amygdalus communis*). Options Méditerr., 1:71–75.

Grassely C.H. and R. Crossa-Raynaud, 1983. Mejora genética. In: "El Almendro", Grasselly C.H. and P. Crossa-Raynaud (Eds)", Ediciones Mundi-Prensa, Madrid, pp.:165—207.

Grasselly C.H., Crossa-Raynaud P., 1984. El almendro. Ed. Mundi Prensa, Madrid, p.465.

Grasselly CH., 1991. Avijor 'Lauranne'. In: "L'Arboriculture Fruitière", 436(2), p. 75.

Griffiths, M.D.and A.J. Huxley, 1992. The New Royal Horticultural Society dictionary of gardening. London: MacMillan Press.

Gruhnert C, Biehl B. and D. Selmar, 1994. Compartmentation of cyanogenic glucosides and their degrading enzymes. Planta, 195:36–42.

Gulcan R., 1975. Cytological studies on young seedlings of double almond seeds. II Colloque du Groupe de Recherche et d'ètude Mediterranéen pour l'Amandier, Montpellier, France.

Gupta P.K., Balyan H.S., Sharma P.C. and B. Ramesh, 1996. Microsatellites in plants. A new class of molecular markers. Curr. Sciences, 70(1): 45-54.

Hansen K.S., Kristensen C., Tattersall D.B., Jones P.R., Bak S. and B.L. Møller, 2003. The in vitro substrate regiospecificity of recombinant UGT85B1, the cyanohydrin glucosyltransferase from Sorghum bicolor. Phytochem., 64:143–51.

Harrison K. and L.M. Were, 2007. Effect of gamma irradiation on total phenolic content yield and antioxidant capacity of almond skin extracts. Food Chem., 102:932–937.

Haughn G. and A. Chaudhury, 2005. Genetic analysis of seed coat development in Arabidopsis. Trends Plant Sci., 10:472–477.

He P., Li J.Z., Zheng X.W., Shen L.S., Lu C.F., Chen Y. and L.H. Zhu, 2001. Comparison of molecular linkage maps and agronomic trait loci between DH and RIL populations derived from the same rice cross. Crop Sci., 41:1240-1246.

Heim K.E., Tagliaferro A.R. and D.J. Bobilya, 2002. Flavonoid antioxidants: Chemistry, metabolism and structure–activity relationships. J. Biochem., 13:572–584.

Heppner J., 1923. The factor for bitterness in the sweet almond. Genetics, 8:390–392.

Heppner J., 1926. Further evidence on the factor for bitterness in the sweet almond. Genetics, 11: 605–606.

Hughes M.A., 1991. The cyanogenic polymorphism in *Trifolium repens* L. (white clover). Heredity, 66:105–15.

Hurtado M.A., Romero C., Vilanova S., Abbot A.G., Llacer G.and M.L. Badenes, 2002. Genetic linkage map of two apricot cultivars (Prunus armeniaca L.) and mapping of PPV (sharka) resistance. Theor. Appl. Genet., 106:819–825.

Hyson D., Schneeman B.O. and P.A. Davis, 2002. Almonds and almond oil have similar effects on plasma lipids and LDL oxidation in healthy men and women. J. Nutrit., 132:703–707.

Jahanban A. S., Mahmoodzadeh A., Hasanzadeh A., Heidari R. and R. Jamei, 2009. Antioxidants and antiradicals in almond hull and shell (*Amygdalus communis* L.) as a function of genotype. Food Chem., 115:529–533.

Janick J., 2005. Breeding intractable traits in fruit crops: Dream the impossible dream. HortSci., 40:1944.

Jáuregui B., de Vicente M.C., Messeguer R., Felipe A., Bonnet A., Salesses G. and P. Arús, 2001. A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theor. Appl. Genet., 102:1169–1176.

Jehan T. and S. Lakhanpaul, 2006. Single nucleotide polymorphism (SNP) methods and applications in plant genetics: a review. Ind. J. Biotechnol., 5:435-459.

Jenrich R., Trompetter I., Bak S.S., Olsen C.E., Møller B.L., and Piotrowski, 2007. Evolution of heteromic nitrilase complexes in Poaceae with new functions in nitrile metabolism. Proc. Natl. Acad. Sci. USA, 104:18848–53.

Jensen K., Osmani S.A., Hamann Naur T.P. and B.L. Møller, 2011. Homology modeling of the three membrane proteins of the dhurrin metabolon: catalytic sites, membrane surface association and protein-protein interactions. Phytochem., 72:2113–23.

Jones D.A. 1998. Why are so many food plants cyanogenic? Phytochem., 47:155–62.

Jones P.R., Møller B.L. and P.B.Hoj, 1999. The UDP-glucose:*p*-hydroxymandelonitrile-*O*-glucosyltransferase that catalyzes the last step in synthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor*: isolation, cloning, heterologous expression, and substrate specificity. J. Biol. Chem., 274:35483–91.

Joobeur T., Viruel M.A., De Vicente M.C., Jàuregui B., Ballester J., Dettori M.T., Verde I., Truco M.J., Messeguer R. and I. Battle, 1998. Construction of a saturated linkage map for *Prunus* using an almond × peach F_2 progeny. Theor. Appl. Genet., 97:1034–1041.

Joobeur T., Periam N., deVicente M.C., King G.J. and P. Arús, 2000. Development of a second generation linkage map for almond using RAPD and SSR markers. Genome, 43:649–655.

Jørgensen K., Rasmussen A.V., Morant M., Nielsen A.H., Bjarnholt N., Zagrobelny M., Bak S. and B.L. Møller, 2005. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. Curr. Opin. Plant Biol., 8:280–91.

Jørgensen K., Morant A.V., Morant M., Jensen N.B. and C.E. Olsen, 2011. Biosynthesis of the cyanogenic glucosides linamarin and lotaustralin in cassava: isolation, biochemical characterization, and expression pattern of CYP71E7, the oxime-metabolizing cytochrome P450 enzyme. Plant Physiol., 155:282–92.

Jung S., Staton M., Lee T., Blenda A., Svancara R., Abbott A. and D. Main, 2008. Genomic, genetic, and breeding resources for Rosaceae crop improvement. Update 1 Oct. 2012.

Kadow D., Voß K., Selmar D. and R. Lieberei, 2012. The cyanogenic syndrome in rubber tree *Hevea brasiliensis*: tissue-damage-dependent activation of linamarase and hydroxynitrile lyase accelerates hydrogen cyanide release. Ann. Bot., 109:1253–62.

Kahn R.A., Bak S., Svendsen I., Halkier B.A. and Møller, 1997. Isolation and reconstitution of cytochrome P450ox and in vitro reconstitution of the entire biosynthetic pathway of the cyanogenic glucoside dhurrin from sorghum. Plant Physiol., 115:1661–1670.

Kahn R.A., Fahrendorf T., Halkier B.A. and Møller, 1999. Substrate specificity of the cytochrome P450 enzymes CYP79A1 and CYP71E1 involved in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* (L.) Moench. Arch. Biochem. Biophys, 363(1), 9–18.

Kakes P., 1985. Linamerase and other β-glucosidases are present in the cell walls of *Trifolium repens* L. leaves. Planta, 166:156–60.

Kannangara R., Motawia M.S., Hansen N.K.K., Paquette S.M., Olsen C.E., Møller B.L. and K. Jørgenses, 2011. Characterization and expression profile of two UDP-glucosyltransferases, UGT85K4 and UGT85K5, catalyzing the last step in cyanogenic glucoside biosynthesis in cassava. Plant J., 68:287–301.

Kashiwada Y., Wang H.K., Nagao T., Kitanaka S., Yasuda I., Fujioka T., Yamagishi T., Cosentino L.M., Kozuka M., Okabe H., Ikeshiro Y., Hu C.Q., Yeh E. and K.H. Lee, 1998. Anti-AIDS agents. 30. Anti-HIV activity of oleanolic acid, pomolic acid, and structurally related triterpenoids. J. Nat. Produc., 61:1090–1095.

Kendall C.W., Jenkins D.J., Marchie A., Ren Y., Ellis P.R. and K.G. Lapsley, 2003. Energy availability from almonds: implications for weight loss and cardiovascular health. A randomized controlled dose-response trial. FASEB J., 17:A339.

Kester D.E., 1965. Inheritance of time of bloom in certain progenies of almond. Proc. Am. Soc. Hort. Sci., 87:214–221.

Kester D.E. and C.J. Hansen, 1966. Rootstock potentialities of F_1 hybrids between peach (*Prunus persica* L.) and almond (*Prunus amygdalus* Batsch.). J. Am. Soc. Hort. Sci., 89:100–109.

Kester D.E., Rady P. and R.N. Asay, 1973. Correlation among chilling requirements for germination, blooming and leafing in almond. Genetics, 74:5135.

Kester D.E. and R.N. Asay, 1975. Almonds. In: "Advances in fruit Breeding". Janick J., Moore J.N. (eds), Purdue Univ. Press, West Lafayette, Indiana, USA, pp.:387–419.

Kester D.E., Hansche P.E., Beres W. and R.N. Asay, 1977. Variance components and heritability of nut and kernel traits in almond. J. Am. Soc. Hort. Sci., 102:264–266.

Kester D.E., Gradziel T.M. and C. Grasselly, 1991. Almonds (*Prunus*). In: "Genetic resources of temperate fruit and nut crops", Moore J.N., Ballington H.J. (eds), Int. Soc. Hort. Sci., The Netherlands, pp.: 701–758.

Kester D.E., Cunningham S. and A.A. Kader, 1993. Almonds. In: Encyclopedia of food science, food technology and nutrition, pp.: 121–126, Academic Press, London.

Kester D.E. and Gradziel, 1996. Almonds (*Prunus*) In: "Fruit Breeding", Moore J.N., Janick J. (eds), Wiley & Sons, New York, USA, pp.:1–97.

King A.D., Miller J.R. M.J. and L.C. Eldridge, 1970. Almond harvesting, processing and microbial flora. Appl. Microbiol., 20:208–214.

Kiran N.S., Polanska L., Fohlerova R., Mazura P., Válková M., Šmeral M., Zouhar J., Malbeck J., Dobrev P.I. , Macháčková I. and B. Brzobohatý, 2006. Ectopic over-expression of the maize {beta}-glucosidase Zm-p60.1 perturbs cytokinin homeostasis in transgenic tobacco. J. Exp. Bot., 57:985–996.

Kodad O., Gracia-Gomez M.S. and R. Socias i Company, 2005. Fatty acid composition as evaluation criterion for kernel quality in almond breeding. Acta Hort., 663: 301–304.

Kodad O., Socias i Company R., Prats M.S. and M.C. Lopez Ortiz, 2006. Variability in tocopherol concentrations in almond oil and its use as a selection criterion in almond breeding. J. Hort. Sci. Biotechnol., 81:501–507.

Kodad O. and R. Socias i Company, 2015. Fruit Quality in Almond as Related to the Type of Pollination in Self-compatible Genotypes. Int. J. Curr. Microbiol. App.Sci., 4 (5):728-736.

Koepke T., Schaeffer S., Harper A., Dicenta F., Edwards M., Henry R.J., Moller B.L., Meisell L., Oraguzie M., Silva H., Sanchez- Perez R. and A. Dhingra, 2013. Comparative genomics analysis in Prunoideae to identify biologically relevant polymorphisms. Plant Biotechnol. J., 11:883–93.

Kojima M., Poulton J.E., Thayer S.S. and E.E. Conn, 1979. Tissue distributions of dhurrin and of enzymes involved in its metabolism in leaves of *Sorghum bicolor*. Plant Physiol., 63:1022–28.

Konieczny A. and F.M. Ausubel, 1993. Procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based. The Plant J., 4(2):403-410.

Kordium E.L., 2008. Double fertilization in flowering plants: 1898-2008. Tsitol Genet., 42(3):12-26.

Kriechbaumer V., Park W.J., Piotrowski M., Meeley R.B., Gierl A. and E. Glawischnig, 2007. Maize nitrilases have a dual role in auxin homeostasis and b-cyanoalanine hydrolysis. J. Exp. Bot., 58: 4225–4233.

Kristensen C., Morant M., Olsen C.E., Ekstrom C.T., Galbraith D.W., Moller B.L. and S. Bak, 2005. Metabolic engineering of dhurrin in transgenic *Arabidopsis* plants with marginal inadvertent effects on the metabolome and transcriptome. Proc. Natl. Acad. Sci. USA, 102:1779–84.

Kristoffersen P., Brzobohaty B., Hohfeld I., Bako L., Melkonian M. and K. Palme, 2000. Developmental regulation of the maize Zm-g60. 1 gene encoding a beta-glucosidase located to plastids. Planta, 210:407-415.

Ladizinsky G., 1999. On the origin of almond. Genet. Resour. Crop Evol., 46:143–147.

Lambert P., Hagen L.S., Arùs P. and J.M. Audergon, 2004. Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas × peach Earlygold reference map for *Prunus*. Theor. Appl. Genet., 108:1120–1130.

Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S.E. and L.A. Newberg, 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, 1 (2):174-81

Leah R., Kigel J., Svendsen I. and J. Mundy, 1995. Biochemical and molecular characterization of a barely seed beta-glucosidase. J. Biol. Chem., 270:15789-15797.

Leavesley H.B., Li L., Prabhakaran K., Borowitz J.L. and GE. Ison, 2008. Interaction of cyanide and nitric oxide with cytochrome c oxidase: implications for acute cyanide toxicity. Toxicol. Sci., 101:101–11.

Lechtenberg M. and A. Nahrstedt, 1999. Cyanogenic glycosides. In: "Naturally Occurring Glycosides", R. Ikan (ed.), Wiley, Chichester, pp.: 147-91.

Lee J., Zhang G., Wood E., Castillo C.R. and A.E. Mitchell, 2013. Quantification of amygdalin in nonbitter, semibitter, and bitter almonds (*Prunus dulcis*) by UHPLC-(ESI)QqQ MS/MS. J. Agric. Food Chem., 61:7754–59.

Lieberei R., Selmar D. and B. Biehl, 1985. Metabolisation of cyanogenic glucosides in *Hevea brasiliensis*. Plant Syst. Evol., 150:49–63.

Livak K.J. and T.D. Schmittgen, 2001. Analysis of Relative Gene Expression Data Using Real- Time Quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. METHODS, 25:402–408,

Lodish H., Berk A., Matsudaira P., Kaiser C.A., Krieger M., Scott M.P., Zipursky L., and J. Darnell, 2004. Molecular Cell Biology, Fifth Edition. San Francisco: W. H. Freeman, pp.: 396.

London-Shafir I., Shafir S. and D. Eisikowitch, 2003. Amygdalin in almond nectar and pollen facts and possible roles. Plant Syst. Evol., 238: 87–95.

Lopez J. and P. Prezioso, 2001. A better way to optimize: two-step gradient PCR. Eppendorf Bionews Application Notes, 16:3-4.

López M., Mnejja M., Rovira M., Colins G., Vargas F.J. and I. Arús P., Batlle, 2004. Self-incompatibility genotypes in almond re-evaluated by PCR, stylar ribonucleases, sequencing analysis and controlled pollinations. Theor. Appl. Genet., 109:954–964.

Lorieux M., Ndjiondjop M.N., Ghesquière A., 2000. A first interspecific Oryza sativa x Oryza glaberrima microsatellite-based genetic linkare map. Theor. Appl. Genet., 100: 593–601.

Lovejoy J.C., Most M.M., Lefevre M., Greenway F.L. and J.C. Rood, 2002. Effect of diets enriched in almonds on insulin action and serum lipids in adults with normal glucose tolerance or type 2 diabetes. Am. J. Clin. Nutr., 76:1000–1006.

Loyd R.C., Gray E., 1970. Amount and distribution of hydrocyanic acid potential during the life cycle of plants of three sorghum cultivars. Agron. J., 62:394–97.

Luby J.J. and D.V. Shaw, 2001. Does marker-assisted selection make dollars and sense in a fruit breeding program? HortSci., 36:872–879.

Maicas S. and J.J. Mateo, 2005. Hydrolysis of terpenyl glycosides in grape juice and other fruit juices: a review. Appl. Microbiol. Biotechnol., 67:322–335.

Manly K.F., Cudmore Robert H. Jr. and J.M. Meer, 2001. Map Manager QTX, cross-platform software for genetic mapping. Mamm. Genome, 12: 930–932

Martinez J.M.,Granado J.M., Montane D., Salvado J. and X. Farriol, 1995. Fractionation of residual lignocellulosics by dilute-acid prehydrolysis and alkaline extraction: application to almond shells. Biores. Technol., 52:59–67.

Martínez-Gómez P. and T.M. Gradziel, 2003. Sexual polyembryony in almond. Sex Plant Reprod., 16:135–139.

Martínez-Gómez P., Arulsekar S., Potter D. and T.M. Gradziel, 2003a. An extended interspecific gene pool available to peach and almond breeding as characterized using simple sequence repeat (SSR) markers. Euphytica, 131:313–322.

Martínez-Gómez P., Sozzi G.O., Sánchez-Pérez R., Rubio M. and T.M. Gradziel, 2003b. New approaches to Prunus tree crop breeding. J. Food. Agri. Environ., 1:52–63.

Martínez-Gómez P., Arulsekar S., Potter D. T.M. Gradziel, 2003c. Relationships among peach and almond and related species as detected by SSR markers. J. Am. Soc. Hort. Sci., 128:667–671.

Martìnez-Gòmez, P., Sánchez-Pérez R., F. Dicenta and T.M. Gradziel, 2005. Improved technique for counting chromosomes in almond. Sci. Hort., 105:139–143.

Martínez-Gómez P., Sánchez-Pérez R. and F. Dicenta, 2008. Fruit Development in Almond for Fresh Consumption. Journal of the American Pomological Society, 62 (2), pp.:82-86.

Martins M., Tenreiro R. and M.M. Oliveira, 2003. Genetic reladness of Portuguese almond cultivars assessed by RAPD and ISSR markers. Plant Cell Rep., 22:71–78.

McCarty C.D., Leslie J.W. and H.B. Frost., 1952. Bitterness of kernels of almond x peach hybrids and their parents. Amer. Soc. Hort. Sci., 59 (1):254-258.

McFarlane I.J., Lees E.M. and E.E. Conn, 1975. The in vitro biosynthesis of dhurrin, the cyanogenic glycoside of *Sorghum bicolor*. J. Biol. Chem., 250:4708–13.

McKey D., Cavagnaro T.R., Cliff J. and R.M. Gleadow, 2010. Chemical ecology in coupled human and natural systems: people, manioc, multitrophic interactions and global change. Chemoecol., 20:109–33.

entzer C. and J. Favre-Bonvin, 1961. Sur la biogenése du glucoside cyanogé- neétique des feuilles de laurier-cerise (Prunus lauro-cerasus). C. R. Acad. Sci. (Paris), 253: 1072–1074.

Merck Company Inc., 2001. The Merck Index. An Encyclopedia of Chemicals, Drugs and Biologicals. 13th Edition. Merck Research Laboratories, Whitehouse Station, NJ, USA, pp:.1818.

Milbury P.E., Chen C.Y., Dolnikowski G.G. and J.B. Blumberg, 2006. Determination of flavonoids and phenolics and their distribution in almonds. J. Agric. Food Chem., 54:5027–5033.

Miller R.E., Jensen R. and I.E. Woodrow, 2006a. Frequency of cyanogenesis in tropical rainforests of Far North Queensland, Australia. Ann. Bot., 97:1017–44.

Ministero delle politiche agricolealimentari e forestali. Piano del settore mandorle, noci, pistacchi e carrube, 2012/1014.

Mizutani M., Nakanishi H., Ema J., Ma S.J., Noguchi, E., Inohara-Ochiai, M., Fukuchi-Mizutani M., Nakao M. and K. Sakata, 2002. Cloning of betaprimeverosidase from tea leaves, a key enzyme in tea aroma formation. Plant Physiol., 130:2164–2176.

Mohan M., Nair S., Bhagwat A., Krishna T.G. and M. Yano, 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol. Breed., 3:87-103

Møller B.L. and E.E. Conn, 1979. The biosynthesis of cyanogenic glucosides in higher plants: Nhydroxytyrosine as an intermediate in the biosynthesis of dhurrin by *Sorghum bicolor* (L.) Moench. J. Biol. Chem., 254:8575–83

Møller B.L. and D.S. Seigler, 1999. Biosynthesis of cyanogenic glucosides and related compounds, in: Singh B.K. (Ed.), Plant amino acids. Marcel Dekker, New York, pp:563–609.

Møller B.L., 2010. Functional diversifications of cyanogenic glucosides. Curr. Opin. Plant Biol., 13:337–46.

Monagas M., Garrido I., Lebron-Aguilar R., Bartolome B. and C. Gomez-Cordoves, 2007. Almond [(*Prunus dulcis* (Mill.) D.A. Webb] skins as a potential source of bioactive polyphenols. J. Agric. Food Chem., 55:8498–8507.

Monastra F.A., Crisafulli F., Marchese G., Ondradu R., Pavia and L. Rivalta, 1982. Monografia di cultivar di mandorlo. Istituto Sperimentale Frutticoltura, Roma.

Morant M., Bak S., Møller B.L. and D. Werck-Reichhart, 2003. Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Curr. Opin. Biotechnol., 14:151–162.

Morant A.V., Jørgensen K., Jørgensen C., Paquette S.M., Sanchez-Perez R., Moller B.L.and S. Bak, 2008. β-glucosidases as detonators of plant chemical defense. Phytochem., 69:1795–813.

Morgante M. and A.M. Olivieri, 1993. PCR-amplified microsatellites as markers in plant genetics. Plant J., 3 (1):175-82.

Moure A., Pazos M., Medina I., Dominguez H. and J.C. Parajo, 2007. Antioxidant activity of extracts produced by solvent extraction of almond shells acid hydrolysates. Food Chem., 101:193-201.

Muncharaz Pou M., 2003. El Almendro. Manual Técnico. 1st Edition. Barcelona: Eds Mundi-Prensa, pp. 413

Myers S., Bottolo L., Freeman C., McVean G. and P. Donnelly, 2005. A fine-scale map of recombination rates and hotspots across the human genome. Science, 310:321–324.

Nahrstedt A., 1985. Cyanogenic compounds as protecting agents for organisms. Plant Syst. Evol., 150, (1-2), 35-47.

Neeraja C.N., Maghirang-Rodriguez R., Pamplona A., Heuer S., Collard B.C., Septiningsih E.M., Vergara G., Sanchez D., Xu K., Ismail A.M. and D.J. Mackill, 2007. A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. Theor. Appl. Genet., 115:767–776.

Neff M.M., Neff J.D., Chory J. and A.E. Pepper, 1998: dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. Plant J., 14:387–392.

Neff M.M., Turk E. and M. Kalishman, 2002. Web-based primer design for single nucleotide polymorphism analysis. Trends Genet., 18:613-615.

Negri P., Bassi D., Magnanini E., Rizzo M. and F. Bartolozzi, 2008. Bitterness inheritance in apricot (*P. armeniaca* L.) seeds. Tree Genet. Genomes, 4:767–776.

Neilson E.H., Goodger J.Q.D., Woodrow I.E. and B.L. Møller, 2013. Plant chemical defence: at what cost? Trends Plant Sci., 18:250–58.

Nelson L., 2006. Acute cyanide toxicity: mechanisms and manifestations. J. Emerg. Nurs., 32:S8–S11.

Nielsen J.S., Møller B.L., 1999. Biosynthesis of cyanogenic glucosides in *Triglochin maritima* and the involvement of cytochrome P450 enzymes. Arch. Biochem. Biophys., 368:121–30.

Nielsen K.A., Tattersall D.B., Jones P.R. and B.L. Møller, 2008. Metabolon formation in dhurrin biosynthesis. Phytochemistry, 69:88–98.

Nishida R., 2002. Sequestration of defensive substances from plants by Lepidoptera. Annu. Rev. Entomol., 47:57–92.

Olsen K.M., Sutherland B.L. and L.L. Small, 2007. Molecular evolution of the *Li/li* chemical defence polymorphism in white clover (*Trifolium repens* L.). Mol. Ecol., 16:4180–93.

Olsen K.M., Hsu S.C. and L.L. Small, 2008. Evidence on themolecular basis of the *Ac/ac* adaptive cyanogenesis polymorphism in white clover (*Trifolium repens* L.). Genetics, 179:517–26.

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

Paterson A.H., 1996. Making genetic maps. In: "Genome mapping in plants", Paterson A.H. (ed.), San Diego, California: Academic Press, Austin, Texas, pp.:23–39.

Paterson A.H., Bowers J.E., Burow M.D., Draye X., Elsik C.G., Jiang C.X., Katsar C.S., Lan T.H., Lin Y.R., Ming R. and R.J. Wright, 2000. Comparative genomics of plant chromosomes. Plant Cell, 12:1523-1539.

Pavan S., Schiavulli A., Lotti C., Ricciardi L., 2014. CAPS technology as a tool for the development of genic and functional markers: study in peas. In: "Cleaved Amplified Polymorphic Sequences (CAPS) Markers in Plant Biology", 83-90. Nova Publisher, ISBN: 978-1-63117-553-4.

Pellegrini N., Serafini, M., Salvatore S., Del Rio D., Bianchi M. and F. Brighenti, 2006. Total antioxidant capacity of spices, dried fruits, nuts, pulses, cereals and sweets consumed in Italy assessed by three different in vitro assays. Molec. Nutrit. Food Res., 50:1030–1038.

Peterson S.C., 1986. Breakdown products of cyanogenesis repellency and toxicity to predatory ants. Naturwissenschaften, 73:627–28.

Peterson S.C., Johnson N.D. and J.L. LeGuyader, 1987. Defensive regurgitation of allelochemicals derived from host cyanogenesis by eastern tent caterpillars. Ecology, 68:1268–72.

Pinelo M., Rubilar M., Sineiro J. and M.J. Nunez, 2004. Extraction of antioxidant phenolics from almond hulls (*Prunus amygdalus*) and pine sawdust (*Pinus pinaster*). Food Chem., 85:267–273.

Piotrowski M., Schonfelder S. and E.W. Weiler, 2001. The Arabidopsis thaliana isogene *NIT4* and its orthologs in tobacco encode beta-cyano-L-alanine hydratase/nitrilase. J Biol. Chem., 276:2616–2621.

Piotrowski M. and J. Volmer, 2006. Cyanide metabolism in higher plants: cyanoalanine hydratase is a *NIT4* homolog. Plant Mol. Biol., 61:111–22.

Piotrowski M., 2008. Primary or secondary? Versatile nitrilases in plant metabolism. Phytochem., 69:2655–67.

Pisha E., Chai H., Lee I.K., Chagwedera T.E., Farnsworth N.R., Cordell, G.A., Beecher C.W.W., Fong H.H.S., Douglas Kinghorn A., Brown D.M., Wani M.C., Wall M.E., Hieken T.J., Gupta T.K. D. and J.M. Pezzut, 1995. Discovery of betulinic acid as a selective inhibitor of human melanoma that functions by induction of apoptosis. Nature Medicine, 1:1046–1051.

Pou-Ilinas J., Canellas J., Driguez H., Excoffier G. and M.R. Vignon, 1990. Steam pre-treatment of almond shells for xylose production. Carbohydrate Res., 207:126–130.

Poulton, J.E., 1990. Cyanogenesis in plants. Plant Physiol., 94:401–405.

Poulton J.E. and C.P. Li, 1994. Tissue level compartmentation of (R)-amygdalin and amygdalin hydrolase prevents large-scale cyanogenesis in undamaged *Prunus* seeds. Plant Physiol., 104:29–35.

Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S. and A. Rafalski, 1996. Comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed., 2:225–238.

Quesada J., Teffo-Bertaud F., Croue J.P. and M. Rubio, 2002. Ozone oxidation and structural features of an almond shell lignin remaining after furfural manufacture. Holzforschung, 56:32–38.

Quilot B., Wu B.H., Cervella J., Génard M., Foulongne M. and K. Moreau, 2004. QTL analysis of quality traits in an advanced backcross between Prunus persica cultivars and the wild relative species P. davidiana. Theor. Appl. Genet., 109:884–897.

Rabinowitz I.N., 1991. Increasing the inositol yield from almond hulls. US Patent, US 5064762.

Rabinowitz I.N., 2002. Method of producing aqueous extracts from almond hulls. US Patent, US 0132031.

Rabinowitz I. N., 2004. Dietary fiber, process for preparing it, and augmented dietary fiber from almond hulls. US Patent, US 0018255.

Richmond G.S. and E.L. Ghisalberti, 1994. Seed dormancy and germination mechanisms in Eremophila (Myoporaceae). Aust. J. Bot., 42:705–15.

Rigoldi M.P., Rapposelli E., De Giorgio D., Resta P. and A. Porceddu, 2015. Genetic diversity in two Italian almond collections. Electronic J. Biotechnol., 18 (1):1-6.

Romero G., Adeva C. and Z. Battad, 2009. Genetic fingerprinting: advancing the frontiers of crop biology research. Philipp. Sci. Lett., 2:8–13.

Rosengarten F.Jr., 1984. The book of edible nuts. Walker and Co., New York, 384.

Rozen S. and H. Skaletsky, 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol., 132:365–386.

Rumsey H.J., 1927. Australian Nuts and Nut Growing in Australia. Part I. The Australian Nut. Herbert J. Rumsey Ltd., Sydney, Australia. 120

Rushforth K., 1999. Collins wildlife trust guide trees: a photographic guide to the trees of Britain and Europe. London: Harper Collins.ISBN 0-00-220013-9.

Sabate J. and E. Haddad, 2001. Almond-rich diets simultaneously improve plasma lipoproteins and alpha-tocopherol levels in men and women. Ann. Nutr. Metab., 45:596.

Sánchez-Pérez R., Dicenta F., Gradziel T.M., Arús P. and P. Martínez- Gómez, 2004a. Application of molecular markers in almond breeding programmes. Nucis-Newsl., 12:9–12.

Sánchez-Pérez R., Dicenta F. and P Martínez-Gómez, 2004b. Identification of S-alleles in almond using multiplex-PCR. Euphytica, 138:263–269.

Sànchez-Pérez R., Ortega E., Duval H., Martìnez-Gòmez P. and F. Dicenta, 2007a. Inheritance and relationships of important agronomic traits in almond. Euphytica, 155:381–391.

Sánchez-Pérez R., Howad W., Dicenta F., Arùs P. and P. Martìnez-Gòmez, 2007b. Mapping major genes and quantitative trait loci controlling agronomic traits in almond. Plant Breed., 126:310– 318.

Sánchez-Pérez R., Jørgensen K., Olsen C.E., Dicenta F. and B.L. Møller, 2008. Bitterness in almonds. Plant Physiol., 146:1040–52.

Sánchez-Pérez R., Jørgensen K., Motawia M.S., Dicenta F. and B.L. Møller, 2009. Tissue and cellular localization of individual β-glycosidases using a substrate-specific sugar reducing assay. Plant J., 60:894–906.

Sánchez-Pérez R., Howad W., Garcia-Mas J., Arús P., Martínez-Gómez P. and F. Dicenta, 2010. Molecular markers for kernel bitterness in almond. Tree Genet. Gen., 6 (2):237–245.

Sánchez-Pérez R., Belmonte F.S., Borch J., Dicenta F., Møller B.L. and K. Jørgensen, 2012. Prunasin hydrolases during fruit development in sweet and bitter almonds. Plant Physiol., 158:1916–32.

Sang S., Cheng, X., Fu H.Y., Shieh D.E., Bai N., Lapsley K., Stark R.E., Rosen R.T. and C.T. Ho, 2002a. New type sesquiterpene lactone from almond hulls (*Prunus amygdalus* Batsch). Tetrahedron Letters, 43:2547–2549.

Sang, S., Lapsley, K., Jeong, W.S., Lachence, P.A., Ho, C.T. and R.T. Rosen, 2002b. Antioxidative phenolic compounds isolated from almond skins (*Prunus amygdalus* Batsch). J.Agric. Food Chem., 50:2459–2463.

Sang, S., Lapsley, K., Rosen, R.T., and C.T Ho, 2002c. New prenylated benzoic acid and other constituents from almond hulls (*Prunus amygdalus* Batsch). J.Agric. Food Chem., 50:607–609.

Sathe S.K., Teuber S.S., Gradziel T.M. and K.H. Rouxm, 2001. Electrophoretic and immunological analyses of almond genotypes and hybrids. J. Agri. Food Chem. 49:2043–2052.

Sathe S.K., Wolf W.J., Roux K.H., Teuber S.S., Venkatacham M. and K.W.C. Sze-Tao, 2002. Biochemical characterization of amandin, the major storage protein in almond (Prunus dulcis L.). J. Agri. Food Chem., 50:4333–4341.

Saunders J.A. and E.E.Conn, 1978. Presence of the cyanogenic glucoside dhurrin in isolated vacuoles from sorghum. Plant Physiol., 61:154–57.

Schappert P.J. and J.S. Shore, 1999a. Cyanogenesis, herbivory and plant defense in *Turnera ulmifolia* on Jamaica. Ecoscience, 6:511–20.

Schappert P.J. and J.S. Shore, 1999b. Effects of cyanogenesis polymorphism in *Turnera ulmifolia* on *Euptoieta hegesia* and potential anolis predators. J. Chem. Ecol., 25:1455–79.

Schirra M., 1997. Post-harvest technology and utilization of almonds. Hort. Rev., 20:267–292.

Singh, Z., and S. Sansavini, 1998. Genetic transformation and fruit crop improvement. Plant Breed. Rev., 16:87–134.

Schueler S., Tusch A., Schuster M. and B. Ziegenhagen, 2003. Characterization of microsatellites in wild and sweet cherry (*Prunus avium* L.) –markers for individual identification and reproductive processes. Genome, 46:95–102.

Schuelke M., 2000. An economic method for the fluorescent labelling of PCR fragments. Nature Biotechn. 18:233-234.

Schwarz B.,Wray V. and P. Proksch, 1996. A cyanogenic glycoside from *Canthium schimperianum*. Phytochem., 42:633–36.

Scorza R., 2001. Progress in tree fruit improvement through molecular genetics. Hort. Sci., 36:855–857.

Sehgal D. and S.N. Raina, 2008. DNA markers and germplasm resource diagnostics: new perspectives in crop improve-ment and conservation strategies. In: Arya ID, Arya S (eds) Utilization of biotechnology in plant sciences. Microsoft Printech (I) Pvt. Ltd, Dehradun, 39–54.

Selmar D., Lieberei R. and B. Biehl, 1988. Mobilization and utilization of cyanogenic glycosides: the linustatin pathway. Plant Physiol., 86:711–16.

Selmar D., Grocholewski S. and D.S. Seigler, 1990. Cyanogenic lipids: utilization during seedling development of Ungnadia speciosa. Plant Physiol., 93:631–636.

Selmar D., 1993. Transport of cyanogenic glucosides: linustatin uptake by Hevea cotyledons. Planta, 191:191–99.

Selmar D., 2010. Biosynthesis of cyanogenic glycosides, glucosinolates and non-protein amino acids. Annu. Plant Rev., 40:92–181.

Serrano B.B., Gomez-Aparisi J. and J. Hormaza, 2002. Molecular fingerprinting of *Prunus* rootstocks using SSRs. J. Hort. Sci. Biol., 77(3):368-372.

Shimada M. and E.E. Conn, 1977. The enzymatic conversion of *p*-hydroxyphenylacetaldoxime to *p*-hydroxymandelonitrile. Arch. Biochem. Biophys., 180:199–207.

Sibbesen O., Koch B., Halkier B.A. and B.L. Møller, 1995. Cytochrome P-450TYR is a multifunctional hemethiolate enzyme catalyzing the conversion of L-tyrosine to *p*-hydroxyphenylacealdehyde oxime in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* (L.) Moench. J. Biol. Chem., 270:3506–11.

Socias i Company R. and A.J. Felipe, 1987. La mejora genética del almendro. Frutic. Prof., 11:64-66.

Socias i Company R. and A.J. Felipe, 1988. Self-compatibility in almond: transmission and recent advances. Acta Hort., 224:307–317.

Socias i Company R., 1998. Fruit tree genetics at a turning point: the almond example. Theor. Appl. Genet., 96:588–601.

Socias i Company R., Felipe AJ., Gomez and J. Aparisi, 1999. A major gene for flowering time in almond. Plant Breed., 118:443–448.

Socias i Company R., Kodad O., Alonso J.M. and T.M. Gradziel, 2008. Almond quality: a breeding perspective. Hort. Rev., 34:197–238.

Sosinski B., Gannavarapu M., Hager L.E., Beck L.E., King G.J., Ryder C.D., Rajapakse S., Baird W.V., Ballard R.E. and A.G. Abbott, 2000. Characterization of microsatellite markers in peach (*Prunus persica* (L) Basch). Theor. Appl. Genet., 101:421–428.

Spiegel-Roy P. and J. Kochba, 1974. The inheritance of bitter and double kernel characters in the almond. Z. Pflanz., 71:319–329.

Stam P., 1993a. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J., 3:739-744.

Stam P., 1993b. JoinMap Version 1.4: A computer program to generate genetic linkage maps. CPRO-DLO, Wageningen. Centre for Plant Breeding and Reproduction Research CPRO-DLO, Wageningen, The Netherlands.

Suda, C.N.K. and J.F. Giorgini, 2003. Multiple forms of endo-1,4-b-glucanases in the endosperm of Euphorbia heterophylla L. J. Exp. Bot., 54, 2045–2052.

Suiter K.A., Wendel J.F. and J.S. Case, 1983. Linkage-I: A pascal computer program for the detection and analysis of genetic linkage. J. Hered., 74:203-204.

Swain E., Li C.P. and J.E. Poulton, 1992a. Development of the potential for cyanogenesis in maturing black cherry (*Prunus serotina* Ehrh.) fruits. Plant Physiol., 98:1423–1428.

Swain E., Li C.P. and J.E. Poulton, 1992b. Tissue and subcellular localization of enzymes catabolizing (*R*)-amygdalin in mature *Prunus serotina* seeds. Plant Physiol., 100:291–300.

Swain E. and J.E. Poulton, 1994a. Immunocytochemical localization of prunasin hydrolase and mandelonitrile lyase in stems and leaves of *Prunus serotina*. Plant Physiol., 106:1285–1291.

Swain E. and J.E. Poulton, 1994b. Utilization of amygdalin during seedling development of *Prunus serotina*. Plant Physiol., 106: 437–445.

Takeoka G., Dao L., Teranishi R., Wong R., Flessa S., Harden L. and R. Edwards, 2000. Identification of three triterpenoids in almond hulls. J. Agric. Food Chem., 48:3437–3439.

Takeoka G.R. and L.T. Dao, 2003. Antioxidant constituents of almond [*Prunus dulcis* (Mill.) D.A. Webb] hulls. J. Agric. Food Chem., 51:496–501.

Takos A.M., Lai D., Mikkelsen L., Maher A.H., Shelton D, *et al*., 2010. Genetic screening identifies cyanogenesis-deficient mutants of *Lotus japonicus* and reveals enzymatic specificity in hydroxynitrile glucoside metabolism. Plant Cell, 22:1605–19.

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

Tattersall D.B., Bak S., Jones P.R., Olsen C.E., Nielsen J.K., Hansen M.L., Høj P.B. and B.L. Møller, 2001. Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science, 293:1826–28

Tautz D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res., 17(16):6463–6471.

Tavassolian I., Rabiei G., Gregory D., Mnejja M., Wirthensohn M.G., Hunt P., Gibson J.P., Ford C.M., Sedgley M. and S.B. Wu, 2010. Construction of an almond linkage map in an Australian population Nonpareil × Lauranne. BMC Genomics, 11:551.

Taylor J. and N.J. Provart, 2006. CapsID: a web-based tool for developing parsimonious sets of CAPS molecular markers for genotyping. BMC Genetics, 10:7-27.

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

Testolin R., Messina R., Lain O., Marrazo T., Huang G. and G. Cipriani, 2004. Microsatellites isolated in almond from an ACrepeat enriched library. Mol. Ecol. Notes, 4:459–461.

Thayer S.S. and E.E. Conn, 1981. Subcellular localisation of dhurrin β-glucosidase and hydroxynitrile lyase in the mesophyll cells of sorghum leaf blades. Plant Physiol., 67:617–22.

Thiel T., Michalek W., Varshney R.K. and A. Graner, 2003. Exploiting EST databases for the development of cDNA derived microsatellite markers in barley (*Hordeum vulgare* L.). Theor. Appl. Genet., 106:411–422.

Tiwari O., Dubey N. and S. Kaushik, 2015. Antibacterial and antioxidant efficacy analysis of leaves extracts of *Prunus amygdalus* (Badam) in different solvents. Int. J. Curr. Microbiol. App. Sci., 4(5): 728-736.

Tong Z., Gao Z., Wang F., Zhou J. and Z. Zhang, 2009. Selection of reliable reference genes for gene expression studies in peach using real-time PCR. BMC Mol. Biol., 10:71.

Torre Grossa J.P., Vaissiere B.E., Rodet G., Botella L. and M. Cousin, 1994. Besoins en pollinisation de la vari´et´e d'amandier autocompatible 'Lauranne'. Acta Hort., 373:145–152.

Urrestarazu M., Martinez G.A. and M.C. Salas, 2005. Almond shell waste: possible local rockwool substitute in soilless crop culture. Sci. Hortic., 103:453–460.

USDA National Nutrient Database for Standard Reference Release 27 Software v.2.2.6., 2015. The National Agricultural Library. Basic Report 12062, Nuts, almonds, blanched a 09:08 EDT.

Van Ooijen J.W., 2011. JoinMap 4.1, Software for the Calculation of Genetic Linkage Maps, Plant Res. Internat., Wageningen, the Netherlands.

Vargas F.J., Romero M.A., Rovira M. and J. Girona, 1984. Amélioration de l'amandier par croisement de variétés. Résultats préliminaires à Tarragone (Espagne). Options Méditerr., 1:101–122.

Vargas F.J. and M.A. Romero, 1994. 'Masbovera', 'Glorieta', and 'Francolí', three new almond varieties from IRTA. Acta Hortic., 373:75-82.

Vargas F.J., Romero M.A. and I. Batlle, 2001. Kernel taste inheritance in almond. Options Méditerr., 56:129–134.

Varshney R.K., 2010. Gene-based marker systems in plants: high-throughput approaches for discovery and genotyping. In: Jain SM, Brar DS (eds) Molecular techniques in crop improvement, Springer, The Netherlands, 119-142.

Venkatachalam M. and K.S. Shridhar, 2006. Chemical composition of selected edible nut seeds. J. Agric. Food Chem., 54:4705–4714.

Vezvaei A., Hancock T.W., Giles L.C., Clarke G.R. and J.F.Jackson, 1995. Inheritance and linkage of isozyme loci in almond. Theor. Appl. Genet., 91:432–438.

Vezvaei A. and J.F. Jackson, 1996. Almond nut analysis. In: "Modern methods of plant analysis". H.F. Linskens and J.F. Jackson (eds.), Fruit analysis. Springer- Verlag, Berlin. Vol. 18. pp.:135-148

Vickery P.J., Wheeler J.L. and C. Mulcahy, 1987. Factors affecting the hydrogen cyanide potential of white clover (*Trifolium repens* L.). Aust. J. Agric. Res., 38:1053–59.

Vilanova S., Romero C., Abbot A.G., Llacer G. and M.L. Badenes, 2003. An apricot (*Prunus armeniaca* L.) F_2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. Theor. Appl. Genet., 107:239–247.

Viruel M.A., Messeguer R., de Vicente M.C., García-Mas J., Puigdomènech P., Vargas F. and P. Arús, 1995. A linkage map with RFLP and isozyme markers for almond. Theor. Appl. Genet., 91:964–971.

Vuysteke M., Mank R., Antonise R., Bastiaans E., Senior M.L. and C.W. Stuber, 1999. Two high density AFLP linkage maps of Zea mays L.:analysis of distribution of AFLP markers. Theor. Appl. Genet., 99:921–935.

Wang H., Nair M.G., Strasburg G.M., Booren A.M. and J.I. Gray, 1999a. Novel antioxidant compounds from tart cherries (*Prunus cerasus*). J. Natur. Prod., 62:86–88.

Wang R.L., Stec A., Hey J., Lukens L. and J. Doebley, 1999b. The limits of selection during maize domestication. Nature, 398 (6724):236-9.

Wang D., Karle R. and A.F. Iezzoni, 2000. QTL analysis of flower and fruit traits in sour cherry. Theor. Appl. Genet., 100:535–544.

Wang Y., Georgi L.L., Zhebentyayeva N., Reighard G.L., Scorza R. and A.G. Abbott, 2002. High-throughput targeted SSR marker development in peach. Genome, 45:319–328.

Wang K.L., Bolitho K., Grafton K., Kortstee A., Karunairetnam S., McGhie T.K, Espley R.V., Hellens R.P. and A.C. Allan, 2010. An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biology, 10:50.

Watkins R., 1979. Cherry, plum, peach, apricot and almond. In: "Evolution of crop plants". *Prunus* spp., N. W. Simmonds (ed.), Longman, London, pp:242–247.

Werner D.J. and M.A. Creller, 1997. Genetic studies in peach: Inheritance of Sweet Kernel and Male Sterility. J. Amer. Soc. Hort. Sci., 122 (2):215-217.

Wicks S.R., Yeh R.T., Gish W.R., Waterston R.H. and R.H. Plasterk, 2001. Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat. Genet., 28:160-164.

Wijeratne S.S.K., Abou-zaid M.M. and F. Shahidi, 2006. Antioxidant polyphenols in almond and its coproducts. J. Agric. Food Chem., 54:312-318.

Wirthensohn M.G., Chin W. L., Franks T.K., Baldock G., Ford C.M. and M. Sedgley, 2008. Characterising the flavour phenotypes o f almond (*Prunus dulcis* Mill.) kernels. J. Hortic. Sci. & Biotechn., 83(4):462–468.

Woodrof J.G., 1979. Tree nuts, production and processing products. Vol. III, 2nd ed. AVI Publications, Westport, CT.

Wünsch A. and J.I. Hormaza, 2002. Cultivar identification and genetic fingerprinting of fruit tree species using DNA markers. Euphytica, 125:56–67.

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

Yim Y.S., Davis G.L., Duru N.A., Musket T.A., Linton E.W., Messing JW, McMullen M. D., Soderlund C.A., Polacco M.L., Gardiner J.M., and E.H. Coe, 2002. Characterization of three maize bacterial artificial chromosome libraries toward anchoring of the physical map to the genetic map using high-density bacterial artificial chromosome filter hybriddization. Plant Physiol., 130:1686–1696.

Yousefzadeh A., Houshmand S., Madani B. and P. Martìnez-Gòmez, 2010. Karyotypic studies in Iranian wild almond species. Caryologia, 63:117–123.

Zagrobelny M., Bak S., Rasmussen A.V., Jorgensen B., Naumann C.M. and B.L. Møller, 2004. Cyanogenic glucosides and plant–insect interactions. Phytochem., 65:293–306.

Zagrobelny M., Bak S., Ekstrøm C.T., Olsen C.E. and B.L. Møller, 2007a. The cyanogenic glucoside composition of *Zygaena filipendulae* (Lepidoptera: Zygaenidae) as effected by

feeding on wild-type and transgenic Lotus populations with variable cyanogenic glucoside profiles. Insect Biochem. Mol. Biol., 37:10–18.

Zagrobelny M., Bak S., Olsen C.E. and B.L. Møller, 2007b. Intimate roles for cyanogenic glucosides in the life cycle of *Zygaena filipendulae* (Lepidoptera, Zygaenidae). Insect Biochem. Mol. Biol., 37:1189–1197.

Zagrobelny M. and B.L. Møller, 2011. Cyanogenic glucosides in the biological warfare between plants and insects: the Burnet moth-birdsfoot trefoil model system. Phytochem., 72:1585–92.

Zeinalabedini M., Khayam-Nekoui M., Grigorian V., Gradziel T.M. and P. Martinez-Gomez, 2010. The origin and dissemination of the cultivated almond as determined by nuclear and chloroplast SSR marker analysis. Sci. Hort., 125:593–601.

Zhebentyayeva, T.N., Swire-Clark G., Georgi L.L., Garay L., Jung S., Forrest A., Blackmon B., Horn R., Howad W., Arus P., Main D., Sosinski B., Baird W.V., Reighard G.L. and A.G. Abbott, 2008. A framework physical map for peach, a model Rosaceae species. Tree Gen. Gen., 4:745–756.