UNIVERSITÀ DEGLI STUDI DI SASSARI

SCUOLA DI DOTTORATO DI RICERCA

Scienze dei Sistemi Agrari e Forestali

e delle Produzioni Alimentari

Indirizzo: Produttività delle Piante Coltivate

Ciclo XXIV

Analysis of almond (*Prunus amygdalus*) biodiversity in Sardinia

D.ssa Maria Pia Rigoldi

Direttore della Scuola prof. ssa Alba Pusino
Referente di Indirizzo prof. ssa Rosella Motzo
Docente Guida prof. Andrea Porceddu

Anno accademico 2011- 2012

INDEX

CHAPTER 2

Study on genetic diversity and phenotypic traits of Sardinian almond accessions

CHAPTER 3

Study on self-incompatibility alleles of Sardinian accessions

3.1 Introduction

CHAPTER 4

Main fatty acids and α-tocopherol concentrations in oil of Sardinian almond accessions

ABSTRACT

Almond (*Prunus amygdalus* Batsch) is the most important nut crop worldwide. Based on 2010 FAO data, the USA were the first almond producing country, followed by Spain, Iran and Italy.

In Sardinia the cultivations are mainly constituted by local cultivars, used for the production of typical sweets. Several of these cvs. are under the risk of genetic erosion.

Focusing on the assessment of local almond materials, the purpose of this thesis was to characterize a 20-years old collection, containing genotypes gathered from different areas of Sardinia.

Chapter 1 studied, by means of eleven SSR markers, the genetic diversity between our collection and a Central-South Italian collection.

Chapter 2 analyzed more deeply the genetic structure of Sardinian collection through the same SSR set. Moreover, the analysis of a four-year phenotypical data set allowed to investigate the existance of phenotypic clusters and the association between molecular and phenotypic data.

Chapter 3 was focused on the molecular characterization of the self-incompatibility (SI) genotypes of the sweet almond accessions of the collection. The comprehension of crossing compatibilities for the local accessions and the identification of possible pollinators is one essential prerequisite to increase the productive levels.

Chapter 4 analyzed the local germplasm for the quality of almond oil in the perspective of specific utilization/purposes (e.g. sweet productions, cosmetic/pharmaceutical uses, nutritional purposes).

CHAPTER 1

GENETIC DIVERSITY BETWEEN SARDINIAN AND CENTRAL-SOUTH ITALIAN ALMOND COLLECTIONS

1.1 Introduction

Almond (*Prunus amygdalus* Batsch, syn. *Prunus dulcis* Miller, syn. *Amygdalus communis* L.*)* is a species native of south-west Asia (Ladizinsky, 1999), that belongs to the Rosaceae family, Amygdaloideae subfamily (Igic and Kohn, 2001). According to phylogenetic studies based on chloroplast DNA analysis, almond and peach (*Prunus persica*) appear to be the most closely related species among cultivated *Prunus* and are classified into the subgenus *Amygdalus* (Badenes and Parfitt, 1995). Actually almond and peach are thought to have evolved from the same primitive stock: peach seems to have evolved eastward into China at lower elevations in regions of higher humidity, whereas almond spread along the deserts and lower mountain slopes to the West, developing many subspecies along the way (Rieger, 2006). From the early bitter-seeded wild genotypes, that still grow in western China, Kazakhstan, Uzbekistan, Afghanistan (Kester and Ross, 1996), almond inherited a remarkable tolerance to drought and poor soils.

In 3,000 BC, domesticated almond (sweet-seeded) was in use in Mediterranean civilizations. With regard to Italy, almond was first brought to Sicily by the Greeks, in the 5th century B.C. (Avanzato, 2006).

Almond is the most important nut crop worldwide. According to the last available FAO data (FAOSTAT), the USA were the first almond producing country in 2010, with 1,413,800 tons of shelled almonds (mainly from California), followed by Spain with 221,000, Iran with 158,000 and Italy with 108,160 tons.

Almond cultivation in Italy is located in the south of the country, particularly in Sicily (46,255 hectares and 71,154 tons), and Apulia (24,510 hectares and 30,087 tons) (data 2011 from ISTAT, Italian National Institution for Statistics). In these two regions productions derive mainly from a few originally local varieties that have largely spread at national level,

but also from several, more restricted, local varieties The species is also present in other Italian regions (Sardinia with 3,419 hectares, and 2,667 tons, followed by Calabria, Abruzzo, and Basilicata), where the cultivations are settled at small-scales, mainly as traditional local varieties which date to ancient times.

Most of traditional varieties are self-incompatible and early-blooming, both facts causing them to have a little production, due to low fertility and frost damage to the flowers. That caused, in the last decades, a regression of traditional almond cultivation in Italy both in terms of lands and production. Such a decline is also due to the application of agronomic techniques not suitable to the needs of a modern almond cultivation.

However, in recent years, the interest in recovering ancient genotypes arise for many species, almond included, in order to protect biodiversity and tradition.

Local accessions could be adequately exploited only if they are previously identified and assessed. Phenotypic traits usually used for cultivar identification could have the limit of being affected by environmental fluctuations and may only be visible in adult materials.

In recent years, molecular markers have been used to study genetic diversity and cultivar identification of peach and almond (Dangl *et al*, 2009; Shiran *et al*., 2007; Sorkheh *et al.*, 2007; Amirbakhtiar *et al.*, 2006; Aranzana *et al.*, 2003; Dirlewanger *et al.*, 2002; Sosinski *et al.*, 2000; Testolin *et al.*, 2000; Cipriani *et al.*, 1999). The first linkage map for almond was constructed by Viruel *et al*. (1995) for the F1 progeny between "Ferragnes" x "Tuono", (F x T map), with RFLPs and isozymes; it detected the expected eight linkage groups (Martinez-Gomez *et al*., 2007), that coincide with its chromosom number (almond has a 300-Mbp, diploid genome with 2n=16, Baird *et al.*, 1994; Corredor *et al.*, 2004). Thereafter a saturated linkage map of *Prunus* was obtained in a F2 progeny of almond x peach (cv. "Texas=Mission" x cv. "Earlygold" respectively) (Joobeur *et al.*, 1998): again, it contained with RFLPs and isozymes (many in common with the F x T map), that mapped in the eight linkage groups found. The T x E map, considered the *Prunus* reference map, has been progressively improved (Aranzana *et al*. 2003,) with the addition of more markers, such as additional RFLPs and simple-sequence repeats (SSRs or microsatellites).

The possibility of studying the genetic diversity among different cultivars and populations will benefit almond germplasm management and breeding programs, by helping to take decisions on parental genotypes for crosses.

Sardinia owns a considerable number of local almond varieties, deriving from the interaction, in an insular environment, between autocthonal and imported individuals. Several of them are gathered in a very precious collection field, the only one owned by a public Regional institution (AGRIS Sardegna). Some of those accessions are under the risk of genetic erosion and are listed in the Rural Development Plan of Sardinian Government (PSR 2007-2013)

The present study investigated, through SSRs molecular markers, the genetic diversity between the Sardinian collection and another germplasm collection from Central-Southern Italy, using some foreign commercial cultivars as reference. Its purpose is to contribute to almond germplasm management, by identifying and maximizing the conserved almond diversity of our Region; that relates also with the promotion of the fruit as a distinctive product for the local agrifood industry that produces almond-based typical sweets (Chessa, 2006).

1.2 Materials and methods

1.2.1 Plant material and DNA extraction

Ninety-six almond accessions were analyzed in this study (tab. 1.1).

Sardinian collection (located in Uta, Southern Sardinia, 39°10'N latitude and 10 metres above sea level) included 43 accessions (40 sweet and 3 bitter almonds) and 4 commercial varieties.

The DNA of 6 more commercial varieties was supplied by G.S. Dangl from the Foundation Plant Services, University of California, Davis.

Central-South (CS) Italian collection was constituted by 43 cultivars, some of which are, by now, extended at a national level in Italy; their DNA was provided by P. Resta from Department of Biology and Agroforestry and Environmental Chemistry - University of Bari (Italy).

Total genomic DNA of Sardinian collection was extracted from young leaves after grinding in liquid nitrogen and using the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich).

Tab .1.1 Plant material included in this study

cultivar	collection site/putative origin	cultivar	collection site/putative origin	cultivar	collection site/putative origin
ANTIOCO PALA	Sardinia	REBECCU 2 (bitter)	Sardinia	NOCELLA	C.S. Italy
ANTONI PIRAS	Sardinia	REBECCU 3 (bitter)	Sardinia	OCCHIOROSSO DI TRANI	C.S. Italy
ARRUBIA	Sardinia	RIU LOI	Sardinia	PAPPAMUCCO	C.S. Italy
BASIBI	Sardinia	SCHINA DE PORCU	Sardinia	PEPPARUDDO	C.S. Italy
BIANCA	Sardinia	STAMPASACCUSU	Sardinia	PIANGENTE	C.S. Italy
BOCCHINO	Sardinia	SUNDA G.	Sardinia	PIGNATIDDE	C.S. Italy
CIATTA INGLESE	Sardinia	SUNDA N.	Sardinia	PISCALZE	C.S. Italy
CIATTA MALISSA	Sardinia	TROITO A	Sardinia/Unknown	PIZZUTA D'AVOLA	C.S. Italy
CORROCHINA	Sardinia	TROITO B	Sardinia/Unknown	PUTIGNANO	C.S. Italy
COSSU	Sardinia	VARGIU	Sardinia	RACHELE TENERA	C.S. Italy
DE EFISI SINZOBA	Sardinia	VAVANI PERRA	Sardinia	RANA	C.S. Italy
DE MRASCIAI	Sardinia	ALBANESE	Puglia	RANA GENTILE	C.S. Italy
EFISI SINZOBA	Sardinia	ANTONIO DE VITO	C.S. Italy	REALE	C.S. Italy
EMILIO 91	Sardinia	BANCHIERE	C.S. Italy	RIVIEZZO	C.S. Italy
FARCI	Sardinia	BARLETTANA	C.S. Italy	ROSSA	C.S. Italy
FIORI	Sardinia	CAPUTO	C.S. Italy	SANTERAMO	C.S. Italy
FOLLA 'E PRESSIU	Sardinia	CATALINI	C.S. Italy	SANTORO	C.S. Italy
FRANCISCU	Sardinia	CATUCCIA	C.S. Italy	TENENTE	C.S. Italy
GHIRONI	Sardinia	CATUCEDDA	C.S. Italy	TUONO	C.S. Italy
IBBA	Sardinia	CENTOPEZZE	C.S. Italy	VISCARDA	C.S. Italy
IS STUMBUS	Sardinia	CIAVEA	C.S. Italy	ZIA COMARA	C.S. Italy
LUTZEDDU	Sardinia	COSIMO DI BARI	C.S. Italy	ZIN ZIN	C.S. Italy
MALISSA TUNDA	Sardinia	CRISTOMORTO	C.S. Italy	JORDANOLO	Sardinia/USA
NIEDDA I	Sardinia	D'ALOIA	C.S. Italy	NE PLUS ULTRA	Sardinia/USA
NIEDDA II	Sardinia	FERRANTE	C.S. Italy	NONPAREIL	Sardinia/USA
NUXEDDA	Sardinia	FILIPPO CEO	C.S. Italy	PICANTILI	Sardinia/Ukraine
OLLA	Sardinia	FRAGIULIO	C.S. Italy	ALDRICH	USA
ORRI	Sardinia	FRANCISCUDDA	C.S. Italy	MISSION	USA
PITICHEDDA	Sardinia	GALGANO	C.S. Italy	RUBY	USA
PROVVISTA	Sardinia	IRENE LANZOLLA	C.S. Italy	SONORA	USA
FARRAU	Sardinia	MINCONE	C.S. Italy	SWEETHEART	USA
REBECCU 1 (bitter)	Sardinia	MONTRONE	C.S. Italy	WINTER	USA

1.2.2 SSR marker genotyping

From a set of 31 single-locus peach SSR, previously assayed on some of the cultivars above (data not shown), 11 were selected, based on their position in the *Prunus* genetic map and their polymorphism (tab. 1.2 and fig. 1.1).

Tab. 1.2 SSRs used in molecular analysis

SSR code	Linkage group	primers (5'→3') sequence		Tm (°C)	Ta (°C)	Repeated motif	Bibliographic reference
UPD 96-003	4	TTGCTCAAAAGTGTCGTTGC	F	66	57	$(CT)_{11}(CA)_{28}$	Testolin *et al.* (2000)
		ACACGTAGTGCAACACTGGC	R	*70*			
UPD 96-005	1	GTAACGCTCGCTACCACAAA	F	68	57	$(AC)_{16}TG(CT)_2CA(CT)_{11}$	Testolin *et al.* (2000)
		CACCCAGCTCATACACCTCA	R	*70*			
UPD 96-013	2	ATTCTTCACTACACGTGCACG	F	69	57	$(AG)_{22}(TG)_8TT(TG)_{10}$	Testolin *et al.* (2000)
		CCCCAGACATACTGTGGCTT	R	*70*			
UPD 98-024	4	CCTTGATGCATAATCAAACAGC	F	68	57	$(GT)_{19}TC(TG)_7$	Testolin *et al.* (2000)
		GGACACACTGGCATGTGAAG	R	*70*			
UPD 98-409	8	GCTGATGGGTTTTATGGTTTTC	F	68	57	$(AG)_{19}$	Testolin *et al.* (2000)
		CGGACTCTTATCCTCTATCAACA	R	*70*			
BPPCT- 004	2	CTGAGTGATCCATTTGCAGG	F	60	58	(CT)22	Dirlewanger *et al.* (2002)
		AGGGCATCTAGACCTCATTGTT	R	*64*			
BPPCT- 007	3	TCATTGCTCGTCATCAGC	F	65	58	$(AG)_{22}$ $(CG)_2$ $(AG)_4$	Dirlewanger *et al.* (2002)
		CAGATTTCTGAAGTTAGCGGTA	R	*68*			
BPPCT- 039	3	ATTACGTACCCTAAAGCTTCTGC	F	66	58	$(GA)_{20}$	Dirlewanger *et al.* (2002)
		GATGTCATGAAGATTGGAGAGG	R	*64*			
BPPCT- 025	6	TCCTGCGTAGAAGAAGGTAGC	F	71	58	$(GA)_{29}$	Dirlewanger *et al.* (2002)
		CGACATAAAGTCCAAATGGC	R	*73*			
BPPCT- 028	1	TCAAGTTAGCTGAGGATCGC	F	68	58	$(TC)_{15}$	Dirlewanger *et al.* (2002)
		GAGCTTGCCTATGAGAAGACC	R	*71*			
CPPCT-33	7	TCAGCAAACTAGAAACAAACC	F	*65*	50	$(CT)_{16}$	Aranzana *et* al (2002)
		TTGCAATCTGGTTGATGTT	R	*61*			

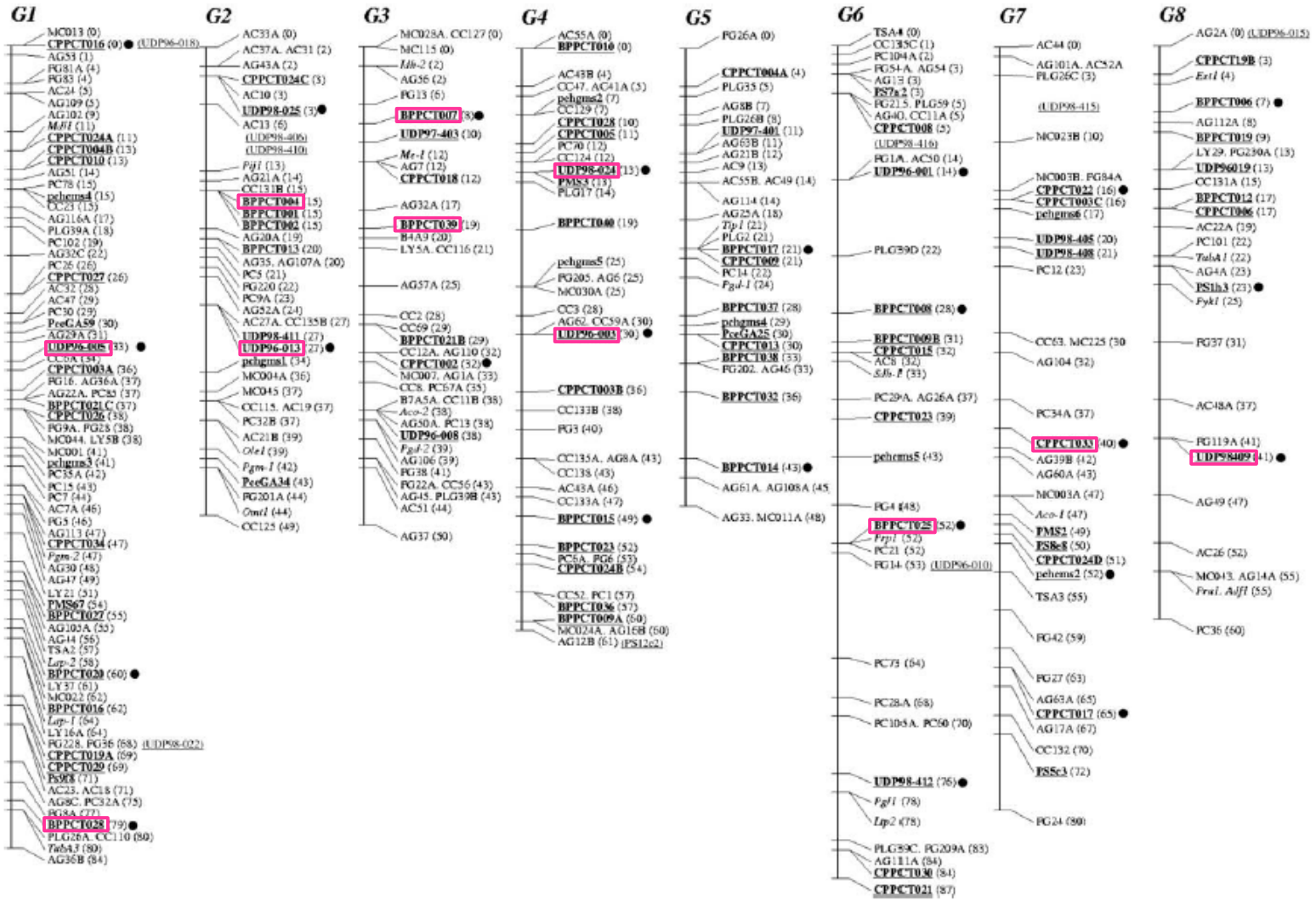

Fig 1.1. *Molecular marker map of Prunus genus (from Aranzana et al., 2003)*

The PCR amplifications were carried out as described by Cipriani et *al.* (1999) for UDP and CPPCT microsatellites, and by Dirlewanger et *al.* (2002) for BPPCT, using 6-FAM dye forward primers. PCR products were separated by electrophoresis in 1,5% agarose gels (Sigma Aldrich) and stained with ethidium bromide to check the PCR amplification and determine approximately the size of the amplified fragments. Amplified PCR products were analyzed by Automated Sequencer Capillary Elettrophoresis (ASCE) using ABI® PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). The software 310 GeneScan® 3.1.2 (Applied Biosystem) was used to evaluate the standard size (GeneScan™ - 500 Liz™ Size Standard).

1.2.3 Data analyses

Descriptive statistics assessing the information content of microsatellite loci were calculated by the following softwares: TFPGA (Miller, 1997), to calculate observed and expected heterozygosity (H_o and H_e), FST and significance p for HW equilibrium test; FSTAT version 2.9.3.2. (Goudet, 2002) for allelic richness (Rs).

Genetic relationships among genotypes were analyzed by a model based hypothesis (*STRUCTURE 2.2*, Pritchard *et al*., 2007) which identified the most likely number of genetic groups or clusters (K) that minimized Hardy Weinberg and linkage disequilibria. The admixed model was used, which allows for analyzing admixture and correlated allele frequencies. For each K, 20 runs (100,000 burn-in generations and 200,000 Markov chain generations) were carried out.

The most likely K was determined following the procedure of Evanno *et al*. (2005).

1.3 Results and discussions

1.3.1 SSR polymorphism

All primer pairs produced polymorphic SSRs consistent with their inheritance as single locus in all almond genotypes (tab.1.3 and 1.4).

Tab 1.3 Range of SSR lengths

SSR code	Fragment length range (bp) for almond accessions addressed in this study
UPD 96-003	85-135
UPD 96-005	106-169
UPD 96-013	185-241
UPD 98-024	72-122
UPD 98-409	116-167
BPPCT 004	182-227
BPPCT 007	124-159
BPPCT 039	118-184
BPPCT 025	180-204
BPPCT 028	166-190
CPPCT 33	142-170

Tab. 1.4 Allele sizes (in base pairs) of SSRs in the almond accessions addressed in this study

Accession	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
ANTIOCO PALA	88	88	128	139	188	235	80	87	136	153	204	207	149	149	143	143	184	192	172	186	158	160
ANTONI PIRAS	88	104	121	141	188	188	72	74	146	146	182	207	142	149	140	140	182	198	172	176	154	162
ARRUBIA	96	124	113	126	200	208	72	87	151	151	216	227	132	149	143	143	182	182	184	184	160	162
BASIBI	88	104	128	146	218	241	74	80	146	153	184	207	124	144	143	150	182	186	180	184	146	162
BIANCA	88	124	113	139	188	200	85	87	140	153	207	207	147	147	143	170	192	192	180	184	156	158
BOCCHINO	114	114	115	115	188	218	87	89	134	146	182	198	132	142	138	152	186	188	172	172	146	162
CIATTA INGLESE	88	100	113	143	200	218	92	105	151	151	195	195	132	158	147	156	188	196	172	186	146	158
CIATTA MALISSA	88	100	141	141	188	218	87	105	140	140	197	197	133	133	134	149	192	196	172	172	142	152
CORROCHINA	100	122	113	113	206	227	72	89	134	151	182	182	144	156	145	158	186	194	182	182	158	162
COSSU	100	104	143	143	232	239	74	83	134	134	182	182	149	158	152	152	182	182	186	186	146	152
DE EFISI SINZOBA	100	114	113	130	200	241	74	74	116	122	182	182	133	133	120	149	182	198	186	186	146	156
DE MRASCIAI	100	100	113	143	206	206	72	80	138	138	185	190	149	156	143	156	182	182	172	172	154	158
EFISI SINZOBA	88	104	113	128	200	218	83	89	140	140	196	207	144	158	138	160	192	200	182	182	152	160
EMILIO 91	114	114	115	115	188	218	87	89	134	147	182	198	132	142	140	152	186	188	172	172	146	162
FARCI	100	122	115	115	218	232	72	74	151	151	182	194	133	133	137	137	192	198	178	178	142	154
FARRAU	133	135	143	143	197	197	119	122	124	124	196	198	141	141	143	143	196	204	170	170	142	154
FIORI	88	104	115	141	235	239	80	87	148	148	182	207	125	158	149	158	182	198	176	182	146	162
FOLLA 'E PRESSIU	114	114	113	113	188	218	87	89	134	146	182	196	132	142	140	152	186	188	172	172	146	162
FRANCISCU	100	104	113	113	188	188	80	80	136	144	182	196	133	133	138	138	182	192	170	182	152	158

continues...

Accession	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
GHIRONI	100	124	141	141	188	218	83	87	148	148	196	196	133	133	137	149	190	194	178	178	146	146
IBBA	100	114	141	141	218	235	74	87	142	144	185	207	144	158	137	156	182	182	176	186	146	160
IS STUMBUS	114	122	128	128	200	218	74	80	134	140	185	185	132	142	140	150	182	194	174	174	158	162
LUTZEDDU	88	104	115	143	218	235	74	83	142	142	196	204	132	142	136	136	192	198	172	182	158	158
MALISSA TUNDA	100	100	121	126	188	206	80	87	144	144	182	194	124	140	143	143	186	188	170	182	142	152
NIEDDA I	122	122	115	141	188	188	78	105	134	138	185	207	124	134	141	141	188	192	176	180	158	160
NIEDDA II	100	114	128	141	200	241	80	92	151	157	190	202	133	133	141	150	182	186	176	180	152	154
NUXEDDA	118	118	119	141	188	206	72	72	153	153	182	204	132	142	122	150	190	198	172	180	142	146
OLLA	88	104	113	128	200	218	83	89	138	140	196	207	144	159	141	160	190	198	184	184	150	158
ORRI	100	104	113	128	200	206	74	74	140	144	182	202	142	142	120	140	182	192	184	184	158	158
PITICHEDDA	110	122	121	128	185	185	74	85	134	138	182	194	150	159	134	150	186	188	172	184	146	152
PROVVISTA	100	104	113	143	188	188	76	83	138	140	185	206	132	144	120	149	182	182	176	176	160	160
REBECCU 1	85	100	106	119	188	188	80	85	144	144	216	225	130	145	126	162	188	198	172	182	146	160
REBECCU 2	85	100	106	119	188	188	80	85	144	144	216	225	130	145	129	150	192	192	172	182	146	160
REBECCU 3	100	100	115	128	188	206	72	85	144	144	182	182	132	145	141	162	192	192	172	172	m.d.	m.d.
RIU LOI	114	118	115	128	188	218	80	89	147	147	182	182	132	142	141	141	196	200	172	172	146	154
SCHINA DE PORCU	88	100	113	143	200	218	92	105	151	151	194	194	132	158	147	158	188	196	172	186	158	158
STAMPASACCUSU	104	104	106	126	188	210	74	83	151	151	208	216	124	132	133	143	184	198	182	182	154	162
SUNDA G.	114	122	113	146	200	218	74	87	140	140	182	182	142	142	141	156	186	198	178	190	146	158

continues...

Accession	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
SUNDA N.	100	122	113	143	213	232	72	74	140	155	182	182	124	149	138	138	192	198	172	184	152	162
TROITO A	100	114	123	130	221	221	85	85	138	138	182	182	144	153	158	180	180	192	172	184	160	160
TROITO B	108	114	130	141	218	218	83	92	138	138	200	200	124	149	143	149	m.d.	m.d.	172	180	146	156
VARGIU	100	104	106	141	200	218	83	85	134	134	182	185	142	142	122	150	182	182	172	182	154	154
VAVANI PERRA	88	104	113	128	200	218	83	89	140	140	196	206	145	158	138	160	192	198	184	184	152	156
JORDANOLO	110	116	109	143	195	200	74	89	140	140	194	204	132	149	129	129	184	194	176	180	146	152
NE PLUS ULTRA	100	114	141	141	185	200	74	92	140	140	182	184	132	132	129	147	186	186	180	186	146	146
NO PAREIL	100	110	141	141	200	200	89	92	140	140	182	194	132	150	129	147	186	186	180	180	146	160
PICANTILI	100	110	115	146	188	200	87	89	138	138	182	194	124	149	143	143	184	194	182	182	146	160
ALDRICH	110	114	130	141	200	218	74	92	134	140	182	196	132	150	122	129	186	186	m.d.	m.d.	146	160
MISSION	108	114	130	146	218	218	74	85	134	140	196	216	132	150	122	145	190	190	172	180	146	158
RUBY	108	110	141	141	200	213	85	89	116	140	194	196	150	150	122	141	190	190	180	180	146	160
SONORA	100	100	146	146	200	200	89	92	140	153	182	194	150	150	147	147	186	186	180	180	146	160
SWEETHEART	110	114	141	141	206	218	80	89	140	140	194	194	132	150	129	147	186	186	180	180	148	154
WINTER	116	116	146	146	200	200	74	74	140	140	182	200	132	150	129	136	186	186	176	180	146	160
ALBANESE	96	108	141	157	185	213	80	103	116	138	196	196	132	149	133	150	188	188	176	176	156	166
ANTONIO DE VITO	96	114	126	141	197	206	85	103	138	140	190	190	124	156	141	158	188	192	168	174	156	166
BANCHIERE	96	100	124	128	197	197	76	85	138	140	190	190	124	132	134	134	188	194	168	174	144	146
BARLETTANA	100	100	141	141	197	237	103	103	138	138	190	206	124	156	152	160	188	196	m.d.	m.d.	156	156
CAPUTO	100	100	124	141	197	227	103	103	138	153	182	190	149	156	150	184	186	196	168	174	168	168

continues...

Accession	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
CATALINI	85	124	117	117	206	218	76	76	116	116	190	195	131	131	141	149	188	192	168	176	152	166
CATUCCIA	100	108	124	143	197	221	87	103	138	138	182	182	141	154	141	158	186	186	178	178	147	158
CATUCEDDA	100	108	124	143	197	221	87	103	138	138	182	182	141	154	141	158	186	186	178	178	147	158
CENTO PEZZE	96	100	115	115	197	237	85	85	116	153	182	190	156	156	134	143	188	188	166	176	168	168
CIAVEA	100	100	113	141	221	241	92	117	138	138	190	195	149	149	133	149	186	192	174	178	156	166
COSIMO DI BARI	96	108	124	128	197	206	76	87	138	138	190	190	132	156	136	136	188	192	168	174	147	168
CRISTOMORTO	96	100	124	139	197	197	89	103	138	138	190	206	156	156	126	150	192	196	170	174	146	156
D'ALOIA	104	114	115	115	197	241	87	87	138	138	184	194	156	156	134	141	188	190	168	174	146	160
FERRANTE	96	96	124	128	197	197	87	87	138	138	197	197	124	124	141	141	192	196	174	174	144	156
FILIPPO CEO	96	100	113	139	197	237	76	105	136	136	190	190	124	156	137	160	188	188	176	176	156	168
FRAGIULIO	m.d.	m.d.	139	167	188	241	76	87	140	140	184	200	147	150	141	147	188	194	178	178	156	168
FRANCISCUDDA	100	114	128	139	197	206	89	89	138	138	190	198	124	156	141	150	m.d.	m.d.	176	176	m.d.	m.d.
GALGANO	100	100	139	139	218	227	85	89	138	138	190	216	156	156	126	150	186	196	176	176	144	156
IRENE LANZOLLA	96	100	139	139	197	208	89	105	138	138	194	206	132	156	150	150	188	196	170	174	156	168
MINCONE	100	124	119	124	197	208	76	89	138	138	190	206	141	156	152	152	192	196	170	170	162	168
MONTRONE	96	100	121	121	188	197	87	87	138	138	190	200	124	149	141	184	188	192	178	178	148	170
NOCELLA	100	100	113	167	197	241	87	105	138	153	182	190	139	154	141	141	188	194	180	180	156	162
OCCHIOROSSO DI TRANI	100	114	169	169	188	197	76	85	138	138	190	206	128	156	134	134	184	188	168	172	144	146
PAPPAMUCCO	96	100	124	128	197	197	76	85	138	138	190	190	124	139	134	134	188	192	168	172	144	146

continues...

Accession	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
PEPPARUDDO	96	100	124	128	197	227	87	89	138	167	195	206	124	139	141	150	188	188	178	178	156	170
PIANGENTE	100	100	m.d.	m.d.	188	227	85	89	138	138	190	198	149	156	126	149	186	186	172	174	m.d.	m.d.
PIGNATIDDE	96	108	124	167	197	208	76	87	153	153	184	190	124	124	126	141	188	190	170	180	146	170
PISCALZE	96	96	135	139	227	227	87	89	138	144	190	198	139	159	141	141	182	188	170	174	m.d.	m.d.
PIZZUTA D'AVOLA	96	100	113	146	218	218	76	76	140	140	184	200	124	132	137	156	182	192	172	172	m.d.	m.d.
PUTIGNANO	100	100	115	115	227	237	87	89	142	142	190	190	149	153	133	149	186	186	174	174	148	156
RACHELE TENERA	100	122	141	141	188	188	87	87	138	142	198	216	124	124	141	141	188	188	168	184	m.d.	m.d.
RANA	100	108	141	169	208	237	76	85	155	155	184	216	154	154	126	149	m.d.	m.d.	168	174	m.d.	m.d.
RANA GENTILE	100	108	141	169	208	237	76	85	155	155	184	216	131	131	126	149	186	190	168	174	146	154
REALE	96	122	124	130	188	197	76	103	153	153	190	198	139	156	134	134	190	192	174	176	146	170
RIVIEZZO	96	100	143	143	197	197	105	105	155	155	185	198	124	124	126	141	188	188	182	182	156	170
ROSSA	100	100	143	143	197	227	89	105	138	138	185	190	141	156	141	152	186	196	184	184	142	156
SANTERAMO	100	100	117	141	221	221	105	105	155	155	184	195	132	156	133	152	182	186	170	170	146	146
SANTORO	100	100	121	124	188	227	76	87	138	153	184	216	154	154	126	141	180	192	170	186	144	156
TENENTE	96	100	124	141	197	227	105	105	138	138	190	198	124	150	118	141	183	192	184	184	146	156
TUONO	100	100	121	141	227	227	89	89	138	138	198	198	124	150	141	141	188	192	182	182	146	156
VISCARDA	100	100	141	141	197	237	105	105	140	140	190	216	124	156	m.d.	m.d.	188	196	170	180	156	156
ZIA COMARA	96	100	126	141	197	241	76	105	138	138	182	190	139	147	150	160	183	198	168	170	146	146
ZIN ZIN	96	100	115	143	221	241	92	117	138	138	190	198	149	156	136	149	186	192	176	176	156	170

1.3.2 Descriptive statistics

Sardinian accessions showed the highest polymorphism and the highest number of private alleles across microsatellite loci (tab. 1.5 of descriptive statistics). The number of alleles in Sardinian accessions ranged from 9 of CPPCT 033 to 23 of BPPCT 039, with a mean of 14.5.

Mean H_o (Nei, 1978) ranged from 0.42 in UDP98-409 to 0.88 in UDP96-024.

Nine out of eleven loci significantly deviated ($p<0.01$) from Hardy-Weinberg equilibrium (values in italics).

As to the CS Italian varieties, the allele number ranged from 8 of UDP96-003 to 17 for UDP 96-005, with a mean of 11.9.

Mean H_o across microsatellite loci was 0.66, ranging from 0.23 of UDP 98-409 to 0.84 of CPPCT 033.

Six out of eleven loci significantly deviated from Hardy-Weinberg equilibrium (values in italics).

Within the commercial cultivars, the allele number per locus was 5.6 on average, ranging from 4 for BPPCT 007 to 7 for BPPCT 039 and UDP 96-013. .

Mean H_o across microsatellite loci ranged from 0.20 in BPPCT 025 to 0.90 in UDP98 024, BPPCT 004 and CPPCT 033 .

Only two loci significantly deviated from Hardy-Weinberg equilibrium (values in italics).

An estimation of allele number more correct than the absolute allele number was made through the allelic richness, that allows to measure it independently from the sample size (hence allowing to compare allele numbers among two 43-sized sample and a 10-sized sample: see table 1.5)

The FST index was statistically significant, with a value of 0.099 on average (ranging from 0.051 of UDP 96-005 to 0.175 of UDP 96-409), with a 99% interval confidence from bootstrapping over loci (10000 replications) between 0.075 and 0.129 (data not shown).

Tab. 1.5 Descriptive statistics for Sardinian, CS-Italian and commercial almond accessions

SSR locus	**N. OF ALLELES**						**ALL. RICHNESS (Rs)**			H_o	H_e	**HW eq. p**	H_o	H_e	**HW eq. p**	H_o	H_e	**HW eq. p**	**FST**
	Tot	(Priv)	Tot	(Priv)	Tot	(Priv)													
	Sardinian		**CS Italian**		**Comm.**		**Sardinian**	**CS Italian**	**Comm.**	**Sardinian**			**CS Italian**			**Comm.**			
UDP 96003	13	(4)	8		5	(1)	7.2	4.7	5	0.77	0.84	*0.0009*	0.67	0.64	0.35	0.80	0.82	0.66	0.108
UDP 96005	13	(2)	17	(6)	6	(1)	8.4	9.7	5.7	0.72	0.88	*0.0005*	0.69	0.91	*<0.0001*	0.40	0.73	*0.007*	0.051
UDP 96013	15	(5)	11	(1)	7	(1)	7.8	7.5	6.5	0.74	0.84	*<0.0001*	0.77	0.82	0.14	0.60	0.68	0.78	0.135
UDP 98024	13	(5)	9	(2)	6		8.5	6.7	5.8	0.88	0.89	0.26	0.65	0.85	*0.0007*	0.90	0.81	0.90	0.065
UDP 98409	16	(8)	9	(1)	5		9.3	5.5	4.8	0.42	0.91	*<0.0001*	0.23	0.63	*<0.0001*	0.40	0.51	0.24	0.175
BPPCT 004	18	(5)	12		7		9.2	7.8	6.6	0.65	0.86	*0.0004*	0.74	0.82	0.0137	0.90	0.79	0.85	0.104
BPPCT 007	18	(9)	13	(4)	4		9.7	8	3.9	0.72	0.91	*<0.0001*	0.70	0.85	*0.0032*	0.70	0.66	0.24	0.107
BPPCT 039	23	(6)	16	(2)	7		11.7	9	6.7	0.72	0.94	*<0.0001*	0.71	0.88	*0.0005*	0.70	0.82	0.13	0.058
BPPCT 025	12	(2)	11	(1)	4		7.7	6.6	4	0.79	0.86	*0.0018*	0.76	0.81	0.0738	0.20	0.61	*0.0003*	0.098
BPPCT 028	10	(1)	11	(2)	5		7.3	7.9	5	0.51	0.84	*<0.0001*	0.52	0.88	*<0.0001*	0.44	0.56	0.27	0.112
CPPCT 033	9	(1)	14	(5)	6		7.2	8.1	5.6	0.83	0.86	0.64	0.84	0.85	0.18	0.90	0.68	0.07	0.085
Mean	14.5		11.9		5.6		8.5	7.4	5.4	0.71	0.88		0.66	0.81		0.63	0.70		0.099
SD	4.03		2.88		1.12		1.37	1.45	0.98	0.14	0.03		0.16	0.16		0.24	0.11		

1.3.3 Model-based cluster identification

The model based analysis identified, as the most important level of structure (Evanno *et al*. 2005), two genetic clusters (k=2, fig.1.2). Bars represent individuals, and colours correspond to proportional membership (Q) of that individual to specific clusters. In order to assign a genotype to one of the K clusters, based on the probability expressed by its Q (membership coefficient), a Q of 0.5 was considered. The red cluster includes all the commercial accession, plus the Sardinian varieties with the exception of cv. Farrau; the blue cluster comprises all the CS-Italian varieties, except for Pizzuta d'Avola, closer to the commercial plus Sardinian group.

Considering the following level of genetic structure (k=3, fig. 1.3), Sardinian group, CS Italian group and commercial cultivars were clearly distinguished, with some exceptions: Troito B and Pitichedda (Sardinia) resulted admixed genotypes (none of their Q coefficient reached 0.5), slightly closer to the commercial group; whereas Pizzuta d'Avola (CS Italian) and Picantili (commercial) resulted definitely closer to Sardinian group.

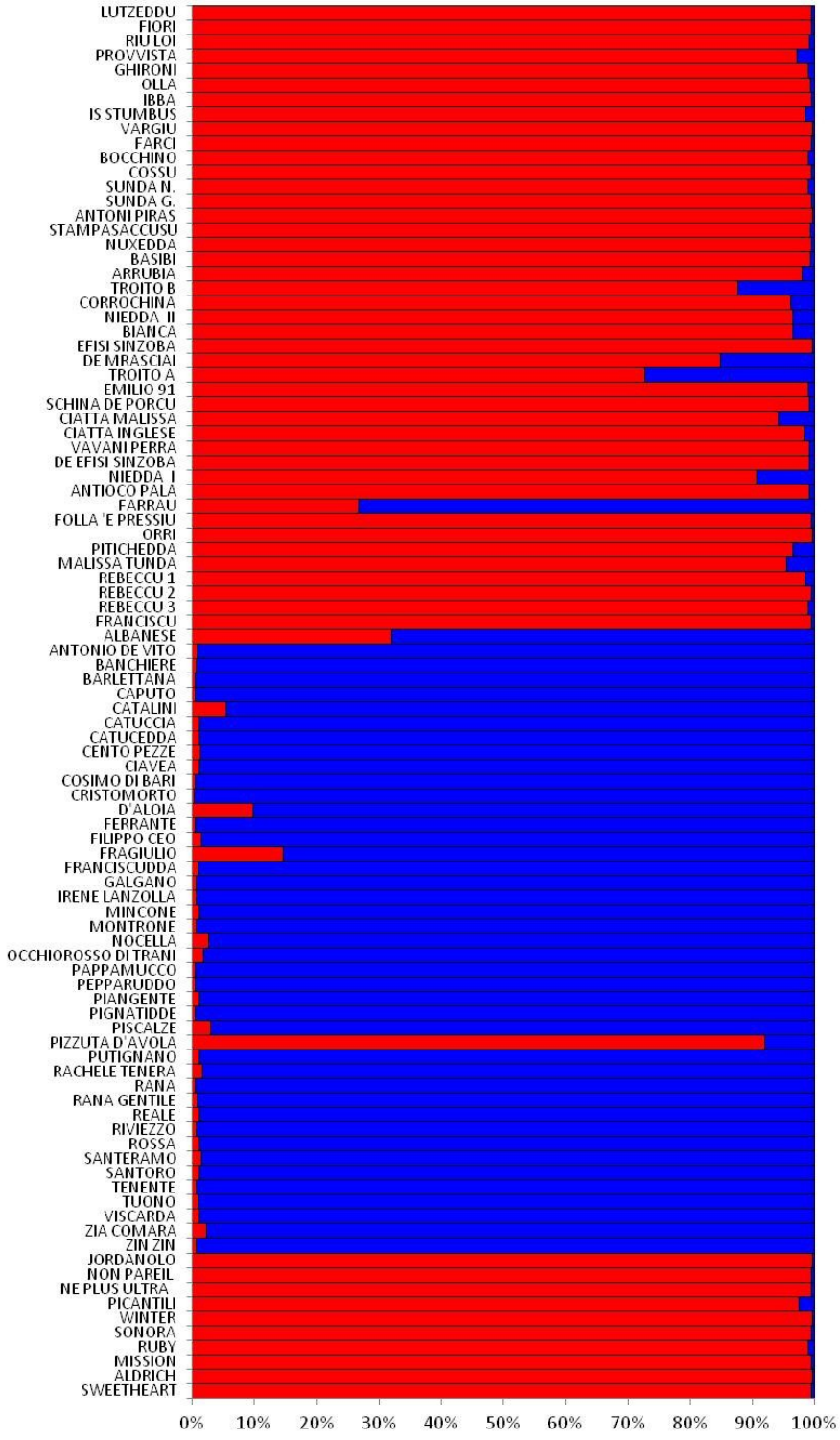

Fig. 1.2 *Individuals grouped basing on their "Q" (membership coefficient). For each individual, Q coefficient for cluster* ▬ *or* ▬ *is shown*

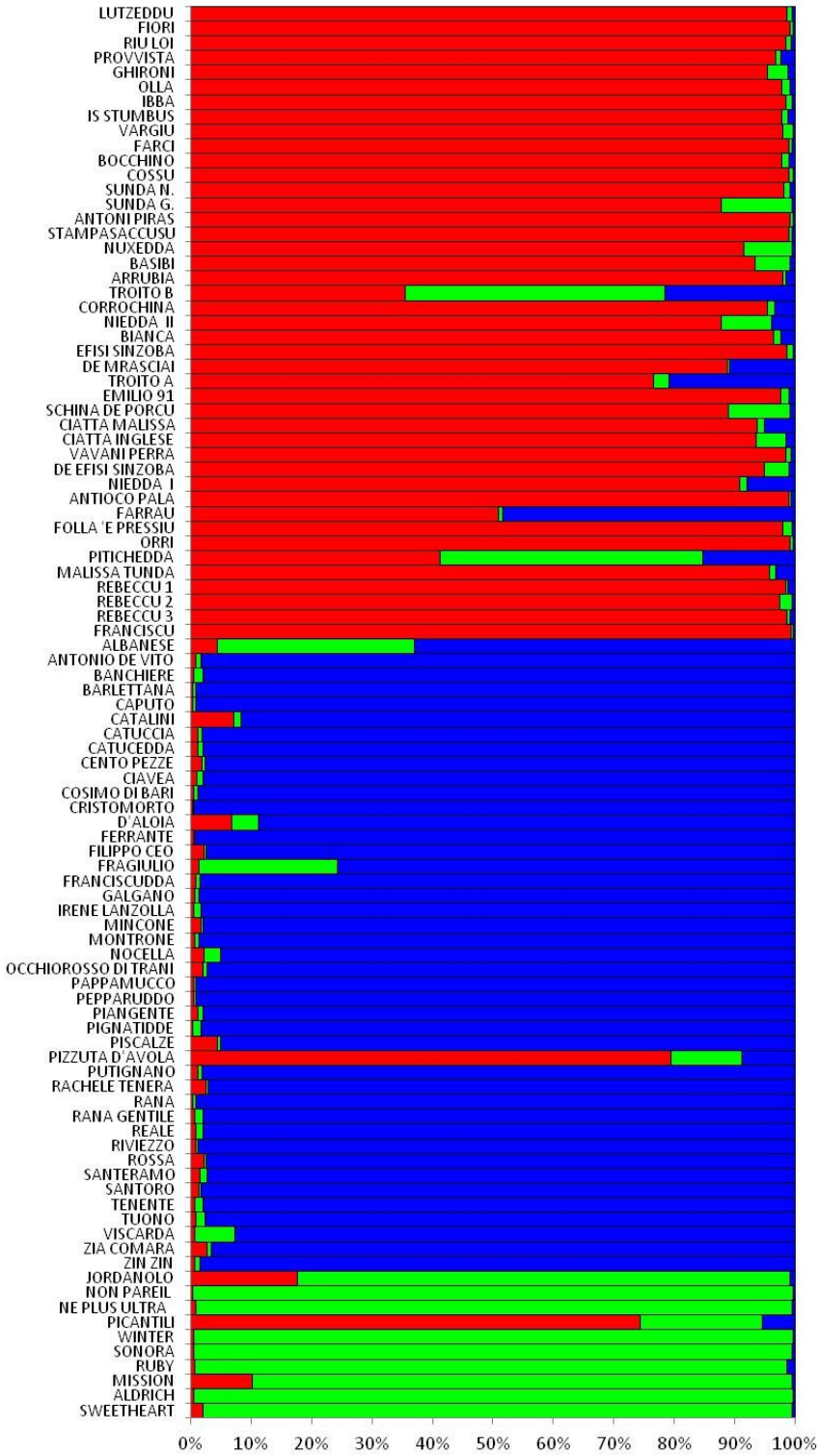

Fig.1.3 *Individuals grouped basing on their "Q" (membership coefficient). For each individual, Q coefficient for cluster , or is shown*

1.3.4 Conclusions

The measures of genetic variability of almond germplasm from Central-South Italy and Sardinia were based on SSR polymorphysm and their comparisons with analogous data from other characterized collections indicated that the Italian collections are an important "reservoir" of genetic diversity. Sanchez-Perez *et al.*, 2006, have found a mean number of SSR alleles of 10.7 and an average level of heterezigosity of 0.86 in a core collection of 21 almond genotypes from mediterranean regions (Spain, Italy, France and Portugal), Ukraine and USA. A mean number of SSR alleles of 8.76 and a level of heterozigosity of 0.67 were reported for an Iranian local collection of almond 57 genotypes (Fathi *et al.*, 2008).

It is worth mentioning that the choice of SSR markers was guided by their map position in chromosomes and thus the values of genetic diversity reported should be a "*bona fide*" representation of the mean diversity at genome level.

All these observations are in sharp contrast with the reduced level of variability of commercial varieties. According to the Almond Board of California 2010 report, the 80 % of almond world production is provided by the USA and within it, 39% is constituted by cv. "Nonpareil" and 40% by only four more cultivars ("Monterey", "Carmel", "Butte" and "Padre"; the first three have been previously reported to be chance seedling selections probably originating from 'Nonpareil' x 'Mission' crosses (Asai *et al.*, 1996; Brooks *et al.*, 1997).

The level of genetic variation of commercial cultivars detected in this study confirmed the scenario of a narrow genetic base for the commercial genotypes: actually the local commercial varieties showed a mean number of alleles of 5.6 and an expected heterozigosity of 0.63.

The analysis of population structure revealed that the Sardinia germplasm presents distinctive features from the continental germplasm. Interestingly we found that at a lower level of germplasm partition (k=2), Sardininan genotypes clustered together with commercial varieties. At higher germplasm partitions (k=3), Sardinian genotypes were separated from the continental germplasm and from the commercial varieties.

Such a finding leaves room for two hypotheses. One possibility is that the Sardinian germplasm belongs to a gene pool comprising also the progenitors of commercial varieties.

The other is that the commercial gene pools has become in contact with the local gene pool in Sardinia whereas a similar event did not occured in the penisula.

It is important to mention at this point that almond cultivation in Sardinia is restricted to domestic orchards which are generally of small size and conducted with traditional techniques. In addition, literature indicates that in Sardinia most of the traditional almond plantations started from seeds (Agabbio *et al.*, 1984). Therefore the possibility of introgressions from the commercial germoplasm to the local gene pools is "technically" possible. Interestingly, in analogy with our finding, Fathi *et al.* (2008) reported that several almond genotypes (Yalda, Shokofe and Shand) from Shiraz were genetically similar to commercial varieties from Spain and USA. The authors hypothesized that migration between commercial varieties and local genotypes could be the cause of the observed genetic similarity.

Altogether these data suggest that even in regions where almond cultivation is carried out following traditional practices introgression of genes from the commercial varieties is possible. Such information should be taken into due account for the establishment of new germplasm collections.

1.4 Bibliography

Agabbio M., Frau A. M. and Chessa I. (1984). Remarks on a five year survey based on ninety-two almond selections of the Sardinian patrimony variety. Option Mediterraneennes. Serie Etudes II: 39-50.

Almond Board of California, available on web at http://www.almondboard.com/AboutTheAlmondBoard/Documents/2010%20Almanac%20FINAL.pdf

Amirbakhtiar N., Shiran B., Moradi H. and Sayed-Tabatabaei B.E. (2006). Molecular Characterization of Almond Cultivars Using Microsatellite Markers Proc. IVth IS on Pistachios and Almonds. Acta Hort. 726, ISHS 2006.

Aranzana M.J, Garcia-Mas J., Carboa J. and Arus P. (2002). Development and variability of micro satellite 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 Arus P. (2003) A set of simple-sequence repeat (SSR) markers covering the Prunus genome. Theor Appl Genet 106: 819–825.

Asai W.K., Micke W.C., Kester D.E. and Rough D. (1996.) The evaluation and selection of current varieties. In: Almond Production Manual. Publication 3364, University of California, Division of Agriculture and Natural Resources (ANR): 51–59.

Avanzato D. and Vassallo I. (2006). In "Following Almond Footprints *(Amygdalus communis* L.*)"*. Scripta Horticulturae n.4, pg 80 (a publication of ISHS).

Badenes M.L. and Parfitt D.E. (1995). Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. Theor. Appl. Genet. 90: 1035–41.

Baird W.V., Estager A.S., Wells J.K. (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.

Brooks R.M. and Olmo H.P. (1997) The Brooks and Olmo register of fruit and nut varieties, 3rd edn. ASHS Press, Alexandria.

Chessa I. (2006). In "Following Almond Footprints *(Amygdalus communis* L.*)"*. Scripta Horticulturae n. 4, pg 94 (a publication of ISHS).

Cipriani G., Lot G., Huang W.G., Marrazzo M.T., Peterlunger E. and Testolin R. (1999). AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. Theor. Appl. Genet. 99: 65-72.

Corredor E., Roman M., Garcia E., Perera E., Arus P. and Naranjo T. (2004). Physical mapping of rDNA genes enables to establish the kariotype of almond. Ann Appl. Biol. 144: 219-222.

Dangl G.S., Yang J., Golino D.A. and Gradziel T. (2009). A practical method for almond cultivar identification and parental analysis using simple sequence repeat markers. Euphytica 168: 41–48.

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

Evanno G., Regnaut S. and Goudet J. (2005). How to use the software Strucutre: a simulation study. Molecular Ecology 14: 2611-2620.

FAOSTAT 2010 Available from http://faostat.fao.org/site/339/default.aspx.

Fathi A., Ghareyazi B., Haghnazari A., Ghaffari M.R., Pirseyedi S.M., Kadkhodaei S., Naghavi M.R. and Mardi M. (2008). Assessment of the genetic diversity of almond

(*Prunus dulcis*) using microsatellite markers and morphological traits. Iranian Journal of Biotechnology, Vol. 6 (2): 98-106.

Goudet J. (2002). FSTAT vers 2.9.3.2 Available from http:// www.unil.ch/izea/softwares/fstat.html. Updated from Goudet J. (1995). FSTAT (vers. 1.2): a computer program to calculate F-statistics. J. Hered. 86: 485-486.

Igic B. and Kohn J.R. (2001). Evolutionary relationships among self-incompatibility RNases. Proc. Natl. Acad. Sci. 98: 13167-13171.

ISTAT 2011. Data on production of main crops. Available from http://dati.istat.it/Index.aspx?DataSetCode=DCSP_COLTIVAZ&Lang=it#.

JMP ver. 7 . Copyright © 2007 by SAS Institute Inc., Cary, NC, USA.

Joobeur T., Viruel M.A., De Vicente M. C., Jauregui B., Ballester J., Dettori M.T., Verde I., Truco M. J., Messeguer R., Batlle I., Quarta R., Dirlewanger E. and Arus P. (1998). Construction of a saturated linkage map for Prunus using an almond X peach F2 progeny. Theor Appl Genet. 97: 1034-1041.

Ladizinsky G. (1999). On the origin of almond. Genetic Resources and Crop Evolution 46: 143-147.

Kester D.E. and Ross N.W. (1996) in "Almond production manual". Publication 3364, University of California, Division of Agriculture and Natural Resources (ANR) pg 1-2.

Martinez-Gomez P., Sanchez-Perez R., Dicenta F., Howad W., Arus P. and Gradziel T.M. (2007). "Almond" In: Genome mapping and molecular breeding in plants, vol 4 – Fruits and nuts. C. Kole (Ed.) Springer-Verlag Berlin Heidelberg .

Miller M.P. (1997). Tools for population genetic analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author. Copright (C) 1997 Mark P. Miller.

Department of Biological Sciences, Nothern Arizona University, Box 5640, Flagstaff, AZ 86011-5640. E-mail: Mark.Miller@usu.edu.

Nei M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Pritchard J.K., Wen X. and Falush D. (2007). Documentation for Structure software, Version 2.2. In: http://pritch.bsd.uchicago.edu/software.

PSR 2007-2013 available on web at http://www.regione.sardegna.it/documenti/14_43_20071121202527.pdf , pg 9-10.

Rieger M. (2006). "Almond". In: Introduction to fruit crops. Food Products Press, an Imprint of the Haworth Press, Inc., NewYork - London - Oxford 2006, pg 37-46.

Sanchez-Perez R., Ballester J., Dicenta F., Arus P. and Martınez-Gomez P. (2006). Comparison of SSR polymorphisms using automated capillary sequencers, and polyacrylamide and agarose gel electrophoresis: Implications for the assessment of genetic diversity and relatedness in almond. Scientia Horticulturae 108: 310–316.

Shiran B., Amirbakhtiar N., Kiani S.,. Mohammadi Sh., Sayed-Tabatabaei B.E. and Moradi H. (2007). Molecular characterization and genetic relationship among almond cultivars assessed by RAPD and SSR markers. Scientia Horticulturae 111: 280–292.

Sorkheh K., Shiran B., Gradziel T.M., Epperson B.K., Martinez-Gomez P. and Asadi E. (2007). Amplified fragment length polymorphism as a tool for molecular characterization of almond germplasm: genetic diversity among cultivated genotypes and related wild speciesof almond, and its relationships with agronomic traits. Euphytica 156: 327–344.

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

Testolin R., Marrazzo T., Cipriani G. and Quarta R. (2000). Microsatellite DNA in peach and its use in fingerprinting and testing the genetic origin of cultivars. Genome, 43 (3): 512-520.

Viruel M.A., Messeguer R., De Vicente M.C., Garcia-Mas J., Puigdomenech P., Vargas F., and Arfis P. (1995). A linkage map with RFLP and isozyme markers for almond. Theor. Appl. Genet. 91: 964-971.

CHAPTER 2

STUDY ON GENETIC DIVERSITY AND PHENOTYPIC TRAITS OF SARDINIAN ALMOND ACCESSIONS

2.1 Introduction

Almond (*Prunus amygdalus* Batsch, syn. *Prunus dulcis* Miller, syn. *Amygdalus communis L.)* is the most important nut crop worldwide. The species is native of south west Asia, extending from the desertic area of west China to the Mediterranean (Ladizinsky, 1999). Sweet kernel selections from these predominantly bitter-seeded wild plants (that still exist) brought to almond domestication at least 3200 years ago in the Middle East (Kester and Ross, 1996; Ladizinsky, 1999). From its wild ancestors almond acquired a remarkable tolerance to drought, salinity and poor soils.

With regard to Italy, the plant was first brought to Sicily by the Greeks, in the 5th century B.C. (Avanzato, 2006). Main almond cultivating Italian areas are located in Southern Italy, particularly in Apulia and Sicily. The species is also cultivated in other Italian regions (Sardinia, Calabria, Abruzzo, Basilicata), whose small-scale productions are constituted mainly by authoctonal, ancient varieties. Those varieties are not competitive due to their little production, caused by self-incompatibility and early-bloom, and by the application of traditional agronomic techniques. However the recent interest in protecting biodiversity, tradition and some qualitative aspects of local productions refuelled the interest for recovering ancient cultivars for many species, almond included.

The identification and the assessment of local genotypes are crucial for their adequate exploitation. Phenotypic traits, such as kernel size, kernel yield, and blooming time, are usually used for cultivar identification in almond (De Giorgio and Polignano, 1999). However, morphological characteristics are influenced by environmental fluctuations and cultivar identification using phenotypic traits is difficult because trees are planted before

distinguishing traits develop (Dangl *et al.*, 2009). In recent years, molecular markers have been used to study genetic diversity and cultivar identification of almond (Dangl *et al.*, 2009; Shiran *et al.*, 2007; Sorkheh *et al.*, 2007; Amirbakhtiar *et al.*, 2006; Xie *et al*., 2006, Aranzana *et al.*, 2003).

Microsatellite or simple-sequence repeat (SSR) markers seem to be the appropriate marker system for cultivar identification, given their high polymorphism, codominant inheritance and the simplicity of the methods required for their development (Aranzana *et al.*, 2003).

Sardinian almond varieties derive from the interaction, in a confined environment, between autocthonal and imported individuals; they are in considerable number, probably preserved thanks to the insularity of the region. Many of them are gathered in a collection field owned by the Regional institution AGRIS Sardegna. Some of these accessions are under the risk of genetic erosion and are listed in the Rural Development Plan 2007-2013 of Sardinian Government (PSR 2007-2013).

That Sardinian collection had previously been compared, through a set of eleven SSRs markers, with a collection from Central-South Italy and some foreign commercial cultivars (see first chapter): the Sardinian collection had grouped, together with the commercial varieties, in a genetic cluster distinct from the CS Italian accessions.

The present study focused on the assessment of such Sardinian almond cultivars, both from the genetic and phenotypical point of view. First of all it investigated, through the same set of eleven SSRs markers above mentioned, the genetic diversity between Sardinian germplasm and the reference commercial cultivars; then it investigate the existance of association between molecular clusters and phenotypical groups. The main aim was to contribute to the identification of the most suitable local varieties, as they can be still important to sustainable agriculture (that leans on biodiversity safeguard, local products for agrifood industry) and for landscape preservation.

2.2 Materials and methods

2.2.1 Plant material and DNA extraction

Fifty-three almond accessions were analyzed in this study. They have been already mentioned in section 1.2.1 “Material and methods-Plant material” of chapter 1: forty-three Sardinian accessions and four commercial varieties of the collection field, plus six more commercial varieties whose DNA was supplied directly from the U.S.A. (see table 2.1).

Tab. 2.1 Plant material included in this study

cultivar	collection site/putative origin	cultivar	collection site/putative origin
ANTIOCO PALA	Sardinia	OLLA	Sardinia
ANTONI PIRAS	Sardinia	ORRI	Sardinia
ARRUBIA	Sardinia	PITICHEDDA	Sardinia
BASIBI	Sardinia	PROVVISTA	Sardinia
BIANCA	Sardinia	REBECCU 1 (bitter)	Sardinia
BOCCHINO	Sardinia	REBECCU 2 (bitter)	Sardinia
CIATTA INGLESE	Sardinia	REBECCU 3 (bitter)	Sardinia
CIATTA MALISSA	Sardinia	RIU LOI	Sardinia
CORROCHINA	Sardinia	SCHINA DE PORCU	Sardinia
COSSU	Sardinia	STAMPASACCUSU	Sardinia
DE EFISI SINZOBA	Sardinia	SUNDA G.	Sardinia
DE MRASCIAI	Sardinia	SUNDA N.	Sardinia
EFISI SINZOBA	Sardinia	TROITO A	Sardinia/Unknown
EMILIO 91	Sardinia	TROITO B	Sardinia/Unknown
FARCI	Sardinia	VARGIU	Sardinia
FARRAU	Sardinia	VAVANI PERRA	Sardinia
FIORI	Sardinia	JORDANOLO	Sardinia/USA
FOLLA 'E PRESSIU	Sardinia	NE PLUS ULTRA	Sardinia/USA
FRANCISCU	Sardinia	NONPAREIL	Sardinia/USA
GHIRONI	Sardinia	PICANTILI	Sardinia/Ukraine
IBBA	Sardinia	ALDRICH	USA
IS STUMBUS	Sardinia	MISSION	USA
LUTZEDDU	Sardinia	RUBY	USA
MALISSA TUNDA	Sardinia	SONORA	USA
NIEDDA I	Sardinia	SWEETHEART	USA
NIEDDA II	Sardinia	WINTER	USA
NUXEDDA	Sardinia		

Total genomic DNA of Sardinian collection was extracted from young leaves after crushing in liquid nitrogen and using the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich).

2.2.2 SSR marker genotyping

The same 11 SSRs as first chapter were used, that had been selected based on their polymorphism and their position in the *Prunus* genetic map (see table 2.2 and fig.2.1).

Tab. 2.2 SSRs used in molecular analysis

SSR code	Linkage group	primers (5'→3') sequence		Tm (°C)	Ta (°C)	Repeated motif	Bibliographic reference
UPD 96-003	4	TTGCTCAAAAGTGTCGTTGC	F	66	57	$(CT)_{11}(CA)_{28}$	Testolin *et al.* (2000)
		ACACGTAGTGCAACACTGGC	R	*70*			
UPD 96-005	1	GTAACGCTCGCTACCACAAA	F	68	57	$(AC)_{16}TG(CT)_2CA(CT)_{11}$	Testolin *et al.* (2000)
		CACCCAGCTCATACACCTCA	R	*70*			
UPD 96-013	2	ATTCTTCACTACACGTGCACG	F	69	57	$(AG)_{22}(TG)_8TT(TG)_{10}$	Testolin *et al.* (2000)
		CCCCAGACATACTGTGGCTT	R	*70*			
UPD 98-024	4	CCTTGATGCATAATCAAACAGC	F	68	57	$(GT)_{19}TC(TG)_7$	Testolin *et al.* (2000)
		GGACACACTGGCATGTGAAG	R	*70*			
UPD 98-409	8	GCTGATGGGTTTTATGGTTTTC	F	68	57	$(AG)_{19}$	Testolin *et al.* (2000)
		CGGACTCTTATCCTCTATCAACA	R	*70*			
BPPCT- 004	2	CTGAGTGATCCATTTGCAGG	F	60	58	(CT)22	Dirlewanger *et al.* (2002)
		AGGGCATCTAGACCTCATTGTT	R	*64*			
BPPCT- 007	3	TCATTGCTCGTCATCAGC	F	65	58	$(AG)_{22}$ $(CG)_2$ $(AG)_4$	Dirlewanger *et al.* (2002)
		CAGATTTCTGAAGTTAGCGGTA	R	*68*			
BPPCT- 039	3	ATTACGTACCCTAAAGCTTCTGC	F	66	58	$(GA)_{20}$	Dirlewanger *et al.* (2002)
		GATGTCATGAAGATTGGAGAGG	R	*64*			
BPPCT- 025	6	TCCTGCGTAGAAGAAGGTAGC	F	71	58	$(GA)_{29}$	Dirlewanger *et al.* (2002)
		CGACATAAAGTCCAAATGGC	R	*73*			
BPPCT- 028	1	TCAAGTTAGCTGAGGATCGC	F	68	58	$(TC)_{15}$	Dirlewanger *et al.* (2002)
		GAGCTTGCCTATGAGAAGACC	R	*71*			
CPPCT-33	7	TCAGCAAACTAGAAACAAACC	F	*65*	50	$(CT)_{16}$	Aranzana *et* al (2002)
		TTGCAATCTGGTTGATGTT	R	*61*			

(a) Linkage group according to Aranzana *et al.* (2003)

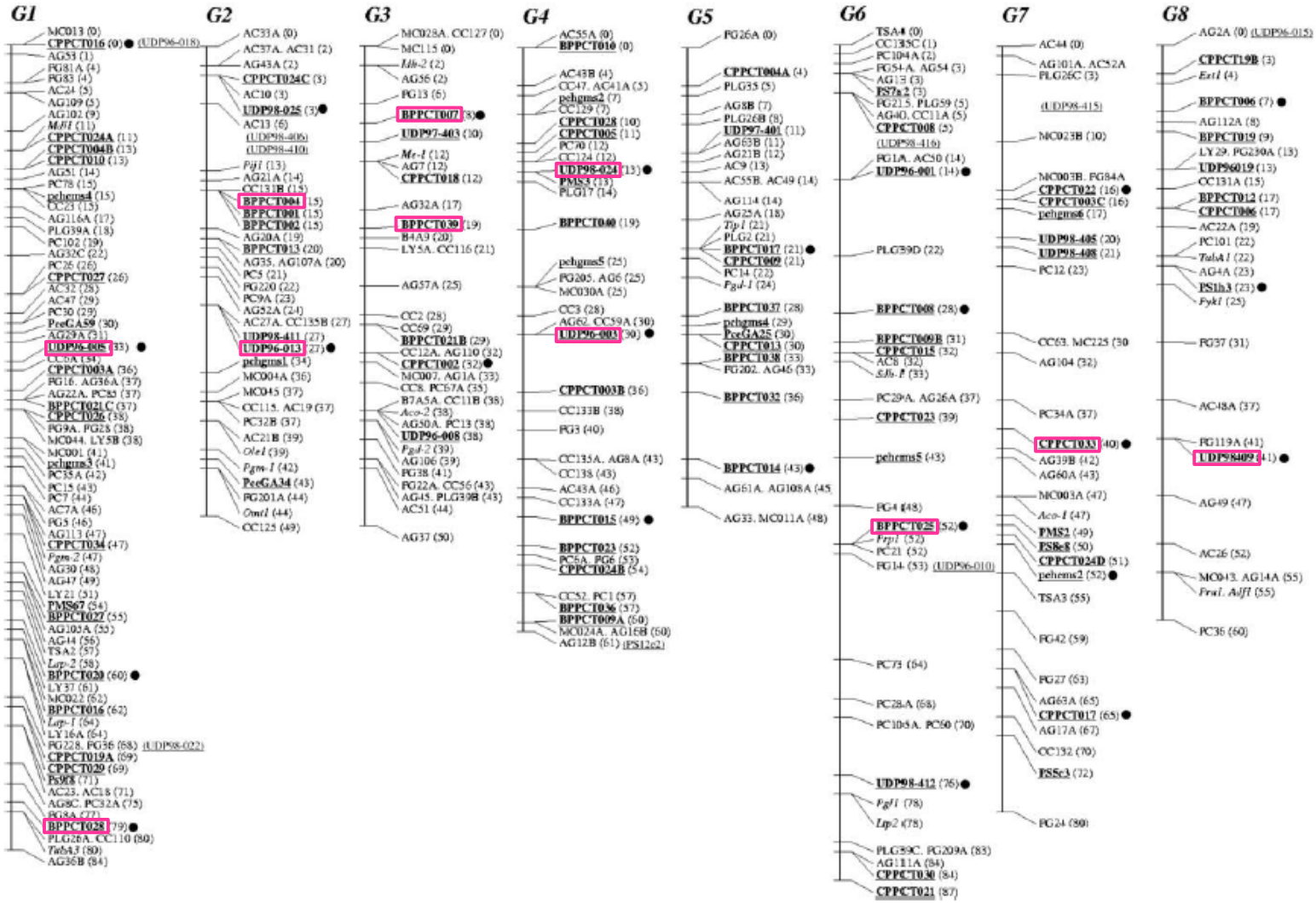

Fig 2.1. *Molecular marker map of Prunus genus (from Aranzana et al., 2003)*

The PCR amplifications were carried out as describe by Cipriani *et al.* (1999) for UDP and CPPCT microsatellites, and by Dirlewanger *et al.* (2002) for BPPCT, using 6-FAM dye forward primers. PCR products were separated by electrophoresis in 1,5% agarose gels (Sigma Aldrich) and stained with ethidium bromide to check the PCR amplification and determine approximately the size of the amplified fragments. Amplified PCR products were analyzed by Automated Sequencer Capillary Elettrophoresis (ASCE) using ABI® PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). The software 310 GeneScan® 3.1.2 (Applied Biosystem) was used to evaluate the standard size (GeneScan™ - 500 Liz™ Size Standard).

2.2.3 Phenotypic traits evaluation

All the sweet almond accession of the collection field were evaluated for the following phenotypical traits: three phenological traits (bloom starting, peak and end, expressed as days from Jan 1st); four productive characteristics, such as nut weight (g), kernel weight (g), kernel %, nuts with kernel failed (%); five morphological fruit traits: double kernel (%), nut length and width (cm); kernel length and width (cm), and the two ratios between the respective dimensions to define fruit shape (fruit width/length) and kernel shape (kernel width/length).

Data were collected during four years. The bitter almonds "Rebeccu 1", "Rebeccu 2", "Rebeccu 3" were not considered for phenotypic analysis as their not-edible fruits are non interesting from the productive point of view.

2.2.4 Data analyses

Descriptive statistics assessing the information content of microsatellite loci were calculated by the following softwares: TFPGA (Miller, 1997), to calculate observed and expected heterozygosity (H_o and H_e), FST and significance (p) for HW equilibrium test; FSTAT version 2.9.3.2. (Goudet, 2002), for allelic richness (Rs).

Genetic relationships among genotypes were analyzed by a model based hypothesis (*STRUCTURE 2.2*, Pritchard *et al*., 2007) and the genetic distance among genetic clusters was analyzed by *TFPGA*.

STRUCTURE 2.2 identified the most likely number of genetic groups or clusters (K) that minimized Hardy Weinberg and linkage disequilibria. The admixed model was used, which allows for analyzing admixture and correlated allele frequencies. For each K, 20 runs (100,000 burn-in generations and 200,000 Markov chain generations) were carried out. The most likely K was determined following the procedure of Evanno *et al*. (2005).

TFPGA software was used to calculate the distance among genetic clusters according to Nei (1978). A dendrogram was obtained using the Unweigthed-Pair-Group-Mathematical Averages method (UPGMA, Sneath and Sokal, 1973). The levels of support for the nodes were evaluated using the bootstrap analysis (adopting 10000 replicates for the bootstrap resampling technique).

Analyses on phenotypic traits and the association between them and genetic clusters were carried out by using *JMP ver.7* software (SAS Institute Inc. 2007) as follows:

I) phenotypic traits were analized by a one-way ANOVA, either with genotype, or *Structure*-derived genetic cluster as independent variable; Tukey test was used in order to detect significant differences among means;

II) to investigate the existance of phenotypic clusters within the collection, a multivariate cluster analysis (according to Ward method) was performed, followed by a MANOVA analysis that identified the most probable number of phenotypic clusters (groups of accessions more similar for overall phenotypic performances);

III) Contingency Analysis was carried out in order to investigate the association between genetic clusters and phenotypic traits.

2.3 Results and discussion

2.3.1 SSR polymorphism

All primer pairs produced polymorphic SSRs consistent with their inheritance as single locus in all almond genotypes (tab 2.3 and 2.4).

Tab 2.3 Range of SSRs lengths

SSR code	Fragment length range (bp) for almond accessions addressed in this study
UPD 96-003	85-135
UPD 96-005	106-146
UPD 96-013	185-241
UPD 98-024	72-122
UPD 98-409	116-157
BPPCT 004	182-227
BPPCT 007	124-159
BPPCT 039	120-180
BPPCT 025	180-204
BPPCT 028	170-190
CPPCT 33	142-162

Tab. 2.4 Allele sizes (in base pairs) of SSRs in the almond accessions addressed in this study

Accession	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
ANTIOCO PALA	88	88	128	139	188	235	80	87	136	153	204	207	149	149	143	143	184	192	172	186	158	160
ANTONI PIRAS	88	104	121	141	188	188	72	74	146	146	182	207	142	149	140	140	182	198	172	176	154	162
ARRUBIA	96	124	113	126	200	208	72	87	151	151	216	227	132	149	143	143	182	182	184	184	160	162
BASIBI	88	104	128	146	218	241	74	80	146	153	184	207	124	144	143	150	182	186	180	184	146	162
BIANCA	88	124	113	139	188	200	85	87	140	153	207	207	147	147	143	170	192	192	180	184	156	158
BOCCHINO	114	114	115	115	188	218	87	89	134	146	182	198	132	142	138	152	186	188	172	172	146	162
CIATTA INGLESA	88	100	113	143	200	218	92	105	151	151	195	195	132	158	147	156	188	196	172	186	146	158
CIATTA MALISSA	88	100	141	141	188	218	87	105	140	140	197	197	133	133	134	149	192	196	172	172	142	152
CORROGHINA	100	122	113	113	206	227	72	89	134	151	182	182	144	156	145	158	186	194	182	182	158	162
COSSU	100	104	143	143	232	239	74	83	134	134	182	182	149	158	152	152	182	182	186	186	146	152
DE EFISI SINZOBA	100	114	113	130	200	241	74	74	116	122	182	182	133	133	120	149	182	198	186	186	146	156
DE MARSCIAI	100	100	113	143	206	206	72	80	138	138	185	190	149	156	143	156	182	182	172	172	154	158
EFISI SINZOBA	88	104	113	128	200	218	83	89	140	140	196	207	144	158	138	160	192	200	182	182	152	160
EMILIO 91	114	114	115	115	188	218	87	89	134	147	182	198	132	142	140	152	186	188	172	172	146	162
FARCI	100	122	115	115	218	232	72	74	151	151	182	194	133	133	137	137	192	198	178	178	142	154
FARRAU	133	135	143	143	197	197	119	122	124	124	196	198	141	141	143	143	196	204	170	170	142	154
FIORI	88	104	115	141	235	239	80	87	148	148	182	207	125	158	149	158	182	198	176	182	146	162

continues...

Genotype	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
FOLLA 'E PRESSIU	114	114	113	113	188	218	87	89	134	146	182	196	132	142	140	152	186	188	172	172	146	162
FRANCISCU	100	104	113	113	188	188	80	80	136	144	182	196	133	133	138	138	182	192	170	182	152	158
GHIRONI	100	124	141	141	188	218	83	87	148	148	196	196	133	133	137	149	190	194	178	178	146	146
IBBA	100	114	141	141	218	235	74	87	142	144	185	207	144	158	137	156	182	182	176	186	146	160
IS STUMBUS	114	122	128	128	200	218	74	80	134	140	185	185	132	142	140	150	182	194	174	174	158	162
LUTZEDDU	88	104	115	143	218	235	74	83	142	142	196	204	132	142	136	136	192	198	172	182	158	158
MALISSA TUNDA	100	100	121	126	188	206	80	87	144	144	182	194	124	140	143	143	186	188	170	182	142	152
NIEDDA I	122	122	115	141	188	188	78	105	134	138	185	207	124	134	141	141	188	192	176	180	158	160
NIEDDA II	100	114	128	141	200	241	80	92	151	157	190	202	133	133	141	150	182	186	176	180	152	154
NUXEDDA	118	118	119	141	188	206	72	72	153	153	182	204	132	142	122	150	190	198	172	180	142	146
OLLA	88	104	113	128	200	218	83	89	138	140	196	207	144	159	141	160	190	198	184	184	150	158
ORRI	100	104	113	128	200	206	74	74	140	144	182	202	142	142	120	140	182	192	184	184	158	158
PITTICCHEDDA	110	122	121	128	185	185	74	85	134	138	182	194	150	159	134	150	186	188	172	184	146	152
PROVVISTA	100	104	113	143	188	188	76	83	138	140	185	206	132	144	120	149	182	182	176	176	160	160
REBECCU 1	85	100	106	119	188	188	80	85	144	144	216	225	130	145	126	162	188	198	172	182	146	160
REBECCU 2	85	100	106	119	188	188	80	85	144	144	216	225	130	145	129	150	192	192	172	182	146	160
REBECCU 3	100	100	115	128	188	206	72	85	144	144	182	182	132	145	141	162	192	192	172	172	m.d	m.d
RIU LOI	114	118	115	128	188	218	80	89	147	147	182	182	132	142	141	141	196	200	172	172	146	154

continues...

Genotype	UDP96 003		UDP96 005		UDP96 013		UDP98 024		UDP98 409		BPPCT 004		BPPCT 007		BPPCT 039		BPPCT 025		BPPCT 028		CPPCT 33	
SCHINA DE PORCU	88	100	113	143	200	218	92	105	151	151	194	194	132	158	147	158	188	196	172	186	158	158
STAMPASACCUS	104	104	106	126	188	210	74	83	151	151	208	216	124	132	133	143	184	198	182	182	154	162
SUNDA G.	114	122	113	146	200	218	74	87	140	140	182	182	142	142	141	156	186	198	178	190	146	158
SUNDA N.	100	122	113	143	213	232	72	74	140	155	182	182	124	149	138	138	192	198	172	184	152	162
TROITO A	100	114	123	130	221	221	85	85	138	138	182	182	144	153	158	180	180	192	172	184	160	160
TROITO B	108	114	130	141	218	218	83	92	138	138	200	200	124	149	143	149	m.d	m.d	172	180	146	156
VARGIU	100	104	106	141	200	218	83	85	134	134	182	185	142	142	122	150	182	182	172	182	154	154
VAVANI PERRA	88	104	113	128	200	218	83	89	140	140	196	206	145	158	138	160	192	198	184	184	152	156
JORDANOLO	110	116	109	143	195	200	74	89	140	140	194	204	132	149	129	129	184	194	176	180	146	152
NE PLUS ULTRA	100	114	141	141	185	200	74	92	140	140	182	184	132	132	129	147	186	186	180	186	146	146
NONPAREIL	100	110	141	141	200	200	89	92	140	140	182	194	132	150	129	147	186	186	180	180	146	160
PICANTILI	100	110	115	146	188	200	87	89	138	138	182	194	124	149	143	143	184	194	182	182	146	160
ALDRICH	110	114	130	141	200	218	74	92	134	140	182	196	132	150	122	129	186	186	m.d	m.d	146	160
MISSION	108	114	130	146	218	218	74	85	134	140	196	216	132	150	122	145	190	190	172	180	146	158
RUBY	108	110	141	141	200	213	85	89	116	140	194	196	150	150	122	141	190	190	180	180	146	160
SONORA	100	100	146	146	200	200	89	92	140	153	182	194	150	150	147	147	186	186	180	180	146	160
SWEETHEART	110	114	141	141	206	218	80	89	140	140	194	194	132	150	129	147	186	186	180	180	148	154
WINTER	116	116	146	146	200	200	74	74	140	140	182	200	132	150	129	136	186	186	176	180	146	160

2.3.2 Descriptive statistics

Descriptive statistics are reported in tab 2.5. It is important to underline the finding that both in terms of allelic number and H_o (observed heterozigosity) the Sardinian genotypes present higher level of genetic diversity than the commercial varieties. The allele number of Sardinian accessions ranged from 9 of CPPCT 033 to 23 of BPPCT 039, with a mean of 14.5.

Mean H_o (Nei, 1978) ranged from 0.42 in UDP98-409 to 0.88 in UDP96-024.

Nine out of eleven loci significantly deviated ($p<0.01$) from Hardy-Weinberg equilibrium (values in italics).

Within the commercial cultivars, the allele number per locus was 5.6 on average, ranging from 4 for BPPCT 007 to 7 for BPPCT 039 and UDP 96-013.

Mean H_o across microsatellite loci ranged from 0.20 in BPPCT 025 to 0.90 in UDP98 024, BPPCT 004 and CPPCT 033 .

Only two loci significantly deviated from Hardy-Weinberg equilibrium (values in italics).

An estimation of allele number more correct than the absolute allele number was made through the allelic richness, that allows to measure it independently from the sample size (hence allowing to compare allele numbers among a 43-sized sample and a 10-sized sample: see table 2.5)

The FST index was statistically significant with a value of 0.115 on average (ranging from 0.037 of UDP96-013 to 0.209 of BPPCT 007), with a 99% interval confidence from bootstrapping over loci (10000 replications) between 0.073 and 0.16 (data not shown).

Tab. 2.5 Descriptive statistics for Sardinian and commercial almond accessions

SSR locus	N. OF ALLELES				ALL. RICHNESS (Rs)		H_o	H_e	HW equil p	H_o	H_e	HW equil p	FST
	Tot	(Priv)	Tot	(Priv)									
	Sardinian		Commercial		Sardinian	Commercial	Sardinian			Commercial			
UDP 96003	13	(9)	5	(1)	7.2	5	0.77	0.84	*0.0018*	0.80	0.82	0.67	0.038
UDP 96005	13	(8)	6	(1)	8.4	5.7	0.72	0.88	*0.0008*	0.40	0.73	*0.008*	0.106
UDP 96013	15	(9)	7	(1)	7.8	6.5	0.74	0.84	*<0.0001*	0.60	0.68	0.77	0.188
UDP 98024	13	(7)	6		8.5	5.8	0.88	0.89	0.25	0.90	0.81	0.90	0.118
UDP 98409	16	(11)	5		9.3	4.8	0.42	0.91	*<0.0001*	0.40	0.51	0.23	0.037
BPPCT 004	18	(11)	7		9.2	6.6	0.65	0.86	*0.0003*	0.90	0.79	0.85	0.085
BPPCT 007	18	(14)	4		9.7	3.9	0.72	0.91	*<0.0001*	0.70	0.66	0.23	0.209
BPPCT 039	23	(16)	7		11.7	6.7	0.72	0.94	*<0.0001*	0.70	0.82	0.14	0.079
BPPCT 025	12	(8)	4		7.7	4	0.79	0.86	*0.0018*	0.20	0.61	0.0003	0.189
BPPCT 028	10	(5)	5		7.3	5	0.51	0.84	*<0.0001*	0.44	0.56	0.28	0.118
CPPCT 033	9	(4)	6	(1)	7.2	5.6	0.83	0.86	0.64	0.90	0.68	*0.06*	0.081
Mean	14.5		5.6		8.5	5.4	0.71	0.88		0.63	0.70		0.115
SD	4.03		1.12		1.37	0.98	0.14	0.03		0.24	0.11		

2.3.3 Model-based cluster identification

The model-based analyses with STRUCTURE were carried out at three hierarchical levels (figs. 2.2.1, 2.2.2, 2.2.3). In each analysis, bars represent individuals, and colours correspond to proportional membership (Q) of that individual to specific clusters. In order to assign a genotype to one of the K clusters, basing on the probability expressed by its membership coefficient, the mayor Q was considered (Q > 50%). In this way six clusters were identified.

At first, the software highlighted two principals clusters (k=2): cluster 1 and cluster 2 (Fig. 2.2.1), that included respectively 44 accessions (all the Sardinian plus the commercial cultivar Picantili), and 9 genotypes (all the other commercial cultivars).

Cluster analysis within cluster 1 (fig. 2.2.2) distinguished, again, two clusters (k=2): cluster 1A includes thirty-five Sardinian accessions, plus cv. Picantili; cluster 1B includes the remaining eight Sardinian accessions. Cluster 2 did not split further.

Within cluster 1A (fig. 2.2.3), *STRUCTURE* identified four clusters (k=4), named 1A1, 1A2, 1A3, 1A4, including respectively nine, seven, seventeen and three accession.

The main distinction among the almond accessions addressed in this study, detected through SSR markers, was between all the Sardinian genotypes and almost all the commercial varieties (nine, out of ten, of them). On the other hand, the distance-based analysis (TFPGA dendrogram, fig. 2.3) indicates that the most distant cluster is 1A4, that is constituted by the bitter almond accessions, followed by the american cultivars (cluster 2).

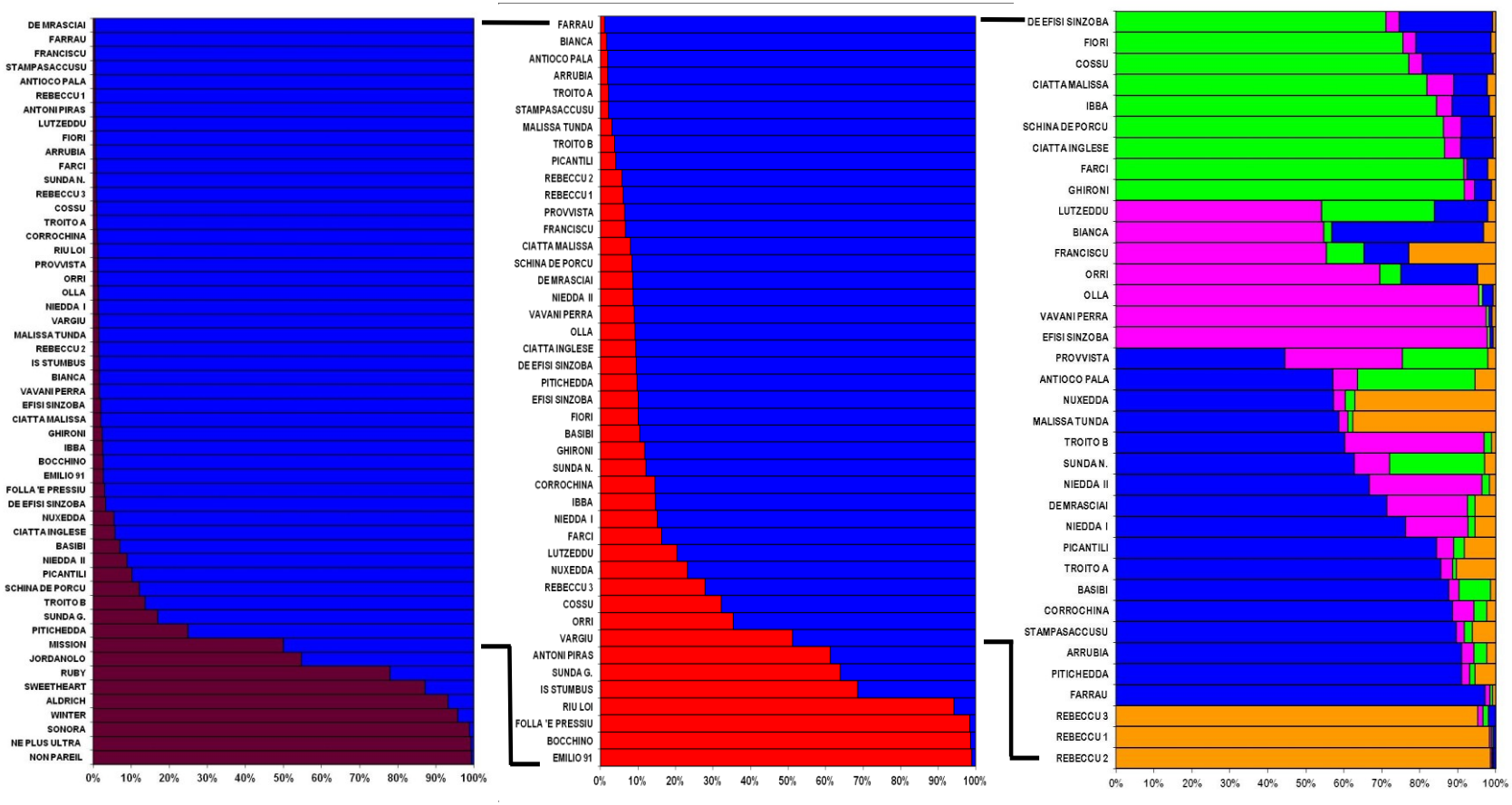

Fig. 2.2.1 For each individual, Q coefficient for the first (1) or second (2) cluster is shown (1 st subdivision)

Fig. 2.2.2 For each individual, Q coefficient for cluster 1A or 1B is shown (2 nd subdivision)

Fig. 2.2.3 For each individual, Q coefficient for cluster 1A1, 1A2, 1A3, or 1A4 is shown (3 rd subdivision)

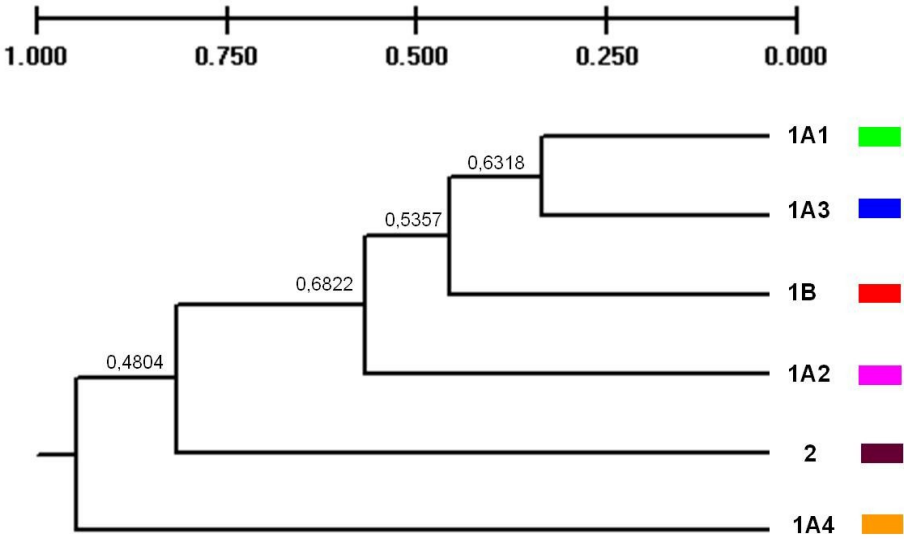

Fig. 2.3. Dendrogram (UPGMA) of genetic distances among clusters (Nei, 1978).

2.3.4 Analyses on phenotypic traits data

2.3.4a ANOVA

Phenotypic traits were analyzed by a one-way ANOVA, either with accession, or genetic cluster, as independent variable. All of them resulted significantly influenced (Tab. 2.6) by the accession, either when Sardinian and commercial were considered altogether or by examining only Sardinian accessions. Tukey test was used to separate means significantly different (tab. 2.7).

The one-way ANOVA with genetic cluster as independent variable computed only five genetic clusters, excluding 1A4, as phenotypic analysis was not carried out on the bitter almond group. Moreover, cluster 2, that in SSR analysis included ten commercial varieties, in the present analysis comprised only three commercial cultivars out of the four grown in the collection field, as Picantili belongs to 1A3 genetic cluster. From that analysis, seven phenotypic traits (all related to fruit characteristics) were significantly influenced by the genetic cluster (Tab. 2.8) only when Sardinian and commercial were considered altogether; Tukey test (Tab. 2.9) shows that significant differences among means occurred between clone 2 (commercial genotypes) and all the other clusters.

Tab. 2.6 One-way ANOVA, x = accession

Sardinian + Commercial					*Sardinian*			
Phenotypical character [(a)]	**DF**	**SS**	**F**	**P**	**DF**	**SS**	**F**	**P**
bloom starting (days from Jan 1st)	43	10219.08	6.94	<0.0001	39	9000.60	7.02	<0.0001
bloom peak (days from Jan 1st)	43	8657.06	7.26	<0.0001	39	7210.99	7.05	<0.0001
bloom end (days from Jan 1st)	43	4438.29	2.23	<0.001	39	3861.99	1.99	<0.01
nut weight (gr)	43	319.50	8.84	<0.0001	39	226.96	6.54	<0.0001
kernel weight (gr)	43	8.84	6.29	<0.0001	39	7.91	6.27	<0.0001
kernel %	43	3.31	100.04	<0.0001	39	1.41	48.44	<0.0001
double kernels (%)	43	2.19	6.20	<0.0001	39	2.10	6.33	<0.0001
nuts with kernel failed (%)	43	0.03	2.42	<0.0001	39	0.01	1.79	<0.01
nut length (cm)	43	19.72	6.60	<0.0001	39	13.40	4.88	<0.0001
nut width (cm)	43	11.78	4.00	<0.0001	39	10.34	3.70	<0.0001
Nut width/length	43	1.34	6.12	<0.0001	39	0.86	4.06	<0.0001
kernel length (cm)	43	11.15	13.05	<0.0001	39	7.65	9.63	<0.0001
kernel width (cm)	43	3.37	11.39	<0.0001	39	3.18	11.98	<0.0001
kernel width/length	43	0.91	24.21	<0.0001	39	0.70	19.70	<0.0001

[(a)] parameter expressed in % were transformed in arcsen for statistical elaborations

Tab. 2.7 Tukey significant differences among means of one-way ANOVA, x= accession [(a)]

ACCESSION	Bloom starting (days from Jan 1st)	Bloom peak (days from Jan 1st)	Bloom end (days from Jan 1st)	nut weight (gr)	Kernel weight (gr)	kernel % or shelling %	double kernels %	Nuts with kernel failed %	Nut length (cm)	Nut width (cm)	Nut shape (nut wdt/ lgt)	Kernel length (cm)	Kernel width (cm)	kernel shape (kernel wdt/lgt)
Antioco Pala	43.0 A-I	48.5 A-K	62.3 A	5.41 A-I	1.51 A-H	28.0 F-K	17.8 B-F	0.75 B	3.06 C-J	2.45 A-F	0.81 A-E	2.46 C-J	1.52 A-H	0.62 D-I
Antoni Piras	44.3 A-I	52.5 A-J	59.8 A	5.65 A-H	1.31 A-J	23.4 H-L	4.5 D-F	1.50 B	3.19 C-I	2.30 B-F	0.73 C-H	2.25 E-O	1.42 B-J	0.63 C-I
Arrubia	51.8 A-E	59.0 A-D	72.0 A	6.17 A-F	1.52 A-H	25.1 H-L	0.7 F	0.75 B	3.20 C-I	2.58 A-E	0.81 A-E	2.55 B-F	1.61 A-D	0.63 C-I
Basibi	38.0 C-J	44.8 D-K	63.5 A	6.86 A-C	1.40 A-I	20.9 K L	0.5 F	1.00 B	3.15 C-J	2.59 A-E	0.83 A-E	2.39 C-L	1.66 A	0.70 A-D
Bianca	49.5 A-H	54.3 A-I	70.0 A	6.98 AB	1.82 A	26.1 G-L	40.4 AB	0.25 B	3.27 B-I	2.57 A-E	0.79 A-G	2.38 C-M	1.58 A-E	0.67 A-H
Bocchino	32.5 I J	38.0 J K	55.0 A	6.33 A-F	1.62 A-G	25.7 G-L	3.2 E F	1.75 B	3.10 C-J	2.44 A-F	0.79 A-G	2.41 C-K	1.60 A-D	0.66 B-H
Ciatta inglese	41.8 A-I	48.0 A-K	64.0 A	5.25 A-I	1.59 A-G	31.0 E-H	5.3 D-F	1.50 B	3.19 C-I	2.53 A-F	0.80 A-E	2.41 C-K	1.58 A-F	0.66 B-I
Ciatta malissa	40.8 A-I	47.8 A-K	61.8 A	6.77 A-C	1.50 A-H	22.2 I-L	9.3 C-F	1.00 B	3.01 C-J	2.79 A-C	0.93 AB	2.22 E-O	1.66 AB	0.75 A
Corrochina	36.8 D-J	45.5 D-K	62.8 A	4.93 A-I	1.43 A-I	29.0 F-J	0.7 F	1.00 B	3.41 A-E	2.60 A-E	0.77 A-G	2.71 B-D	1.41 C-K	0.52 LM
Cossu	50.8 A-G	58.8 A-D	72.8 A	5.93 A-F	1.67 A-E	28.4 F-K	6.8 C-F	1.75 B	3.23 B-I	2.63 A-D	0.81 A-E	2.36 C-N	1.68 A	0.71 A-C
De Efisi Sinzoba	40.5 A-I	48.0 A-K	65.3 A	2.28 J K	1.39 A-I	60.7 BC	14.3 C-F	0.25 B	3.11 C-J	1.81 F	0.58 G-H	2.45 C-K	1.21 J-L	0.49 MN
De Mrasciai	51.5 A-F	57.8 A-E	67.5 A	6.00 A-F	1.25 B-J	21.3 J-L	0.0 F	0 B	2.84 D-J	2.72 A-C	0.96 A	2.10 I-P	1.48 A-I	0.71 A-C
Efisi Sinzoba	32.0 I J	40.3 H-K	63.5 A	4.62 A-J	1.18 D-J	25.6 G-L	9.0 C-F	0.50 B	2.73 E-J	2.16 B-F	0.79 A-F	2.06 K-P	1.35 E-L	0.66 B-H
Emilio 91	34.3 G-J	39.5 I-K	54.8 A	6.48 A-E	1.71 AB	26.6 G-L	4.8 D-F	0.50 B	3.26 B-I	2.57 A-E	0.79 A-F	2.44 C-K	1.63 A-C	0.67 A-H
Farci	40.5 A-I	47.5 A-K	62.3 A	4.31 C-K	0.98 I J	22.7 I-L	0.5 F	1.50 B	3.12 C-J	2.25 B-F	0.73 C-H	2.22 E-O	1.28 H-L	0.58 I-L
Farrau	39.3 B-J	52.0 A-J	60.3 A	6.30 A-F	1.47 A-I	23.7 H-L	0.0 F	0.25 B	3.33 B-I	2.48 A-F	0.74 B-H	2.53 B-G	1.53 A-G	0.60 F-L
Fiori	35.5 E-J	43.5 E-K	60.5 A	3.27 G-K	1.70 A-C	52.1 D	3.8 E F	0.25 B	3.22 C-I	2.29 B-F	0.71 C-H	2.50 B-H	1.51 A-H	0.60 F-L
Folla 'e pressiu	35.0 F-J	41.8 F-K	59.5 A	6.52 A-D	1.69 A-D	25.7 G-L	8.8 C-F	1.25 B	3.25 B-I	2.58 A-E	0.79 A-F	2.37 C-N	1.62 A-C	0.68 A-F
Franciscu	41.0 A-I	47.3 B-K	63.8 A	3.75 F-K	1.01 H-J	26.8 G-L	9.8 C-F	0.25 B	2.82 D-J	1.95 D-F	0.69 D-H	2.14 G-O	1.15 L	0.54 J-L
Ghironi	44.3 A-I	51.0 A-J	63.8 A	6.97 AB	1.60 A-G	22.9 I-L	29.6 A-D	0.50 B	3.34 B-H	2.73 A-C	0.82 A-E	2.49 B-I	1.68 A	0.68 A-G
Ibba	30.8 I J	40.3 H-K	58.5 A	3.87 F-K	1.41 A-I	37.1 E	31.8 A-C	1.25 B	2.93 D-J	2.14 B-F	0.76 A-G	2.32 D-O	1.25 I-L	0.54 J-M
Is Stumbus	39.3 B-J	45.3 D-K	63.3 A	5.33 A-I	1.49 A-I	28.0 F-K	0.5 F	0 B	3.51 A-D	2.41 A-F	0.69 D-H	2.75 BC	1.46 A-I	0.53 K-M
Lutzeddu	33.3 H-J	40.8 G-K	57.0 A	7.04 A	1.70 A-D	24.2 H-L	16.8 B-F	1.50 B	3.51 A-D	2.53 A-F	0.72 C-H	2.49 B-I	1.50 A-H	0.60 F-L

continues….

ACCESSION	Bloom starting (days from Jan 1st)	Bloom peak (days from Jan 1st)	Bloom end (days from Jan 1st)	nut weight (gr)	Kernel weight (gr)	kernel % or shelling %	double kernels %	Nuts with kernel failed %	Nut length (cm)	Nut width (cm)	Nut shape (nut wdt/ lgt)	Kernel length (cm)	Kernel width (cm)	kernel shape (kernel wdt/lgt)
Malissa tunda	49.8 A-H	55.0 A-H	65.8 A	6.55 A-D	1.36 A-I	21.2 K L	2.5 E F	0.75 B	3.70 A-C	2.52 A-F	0.68 D-H	2.62 B-E	1.55 A-F	0.59 H-L
Niedda I	41.8 A-I	47.3 B-K	69.0 A	5.94 A-F	1.31 A-J	22.2 I-L	1.5 E F	1.00 B	3.09 C-J	2.32 B-F	0.75 B-H	2.30 E-O	1.47 A-I	0.64 C-I
Niedda II	41.8 A-I	48.5 A-K	59.0 A	6.60 A-D	1.59 A-G	24.3 H-L	17.4 B-F	0.75 B	3.38 B-G	3.11 A	0.91 A-C	2.45 C-K	1.62 A-C	0.67 A-H
Nuxedda	41.5 A-I	47.0 B-K	65.3 A	4.15 D-K	0.81 J	19.8 L	0.0 F	0.50 B	2.43 J	2.19 B-F	0.90 A-C	1.73 P	1.29 G-L	0.75 A
Olla	30.8 I J	40.5 G-K	62.8 A	4.45 A-J	1.13 G-J	25.5 G-L	10.5 C-F	0.25 B	2.64 F-J	2.15 B-F	0.82 A-E	1.97 N-P	1.34 F-L	0.68 A-F
Orri	39.5 B-I	45.0 D-K	58.8 A	4.51 A-J	1.30 B-J	29.4 F-I	26.5 A-E	4.50 AB	2.60 H-J	2.24 B-F	0.86 A-D	2.00 L-P	1.40 C-K	0.70 A-D
Pitichedda	35.0 F-J	43.8 E-K	56.5 A	3.04 H-K	1.54 A-G	50.6 D	45.5 A	0.00 B	2.53 I J	2.18 B-F	0.86 A-D	1.93 OP	1.40 C-K	0.73 AB
Provvista	34.5 G-J	41.0 G-K	60.8 A	5.73 A-G	1.45 A-I	25.5 G-L	21.1 B-F	1.75 B	3.39 B-F	2.43 A-F	0.73 C-H	2.44 C-K	1.47 A-I	0.60 F-L
Riu Loi	41.5 A-I	46.5 C-K	63.3 A	4.76 A-J	1.23 B-J	25.8 G-L	4.3 E F	1.50 B	2.94 D-J	2.31 B-F	0.79 A-G	2.18 F-O	1.47 A-I	0.68 A-F
Schina de porcu	40.3 A-I	47.0 B-K	62.8 A	4.45 A-J	1.55 A-G	35.0 EF	4.0 E F	0.25 B	3.19 C-I	2.56 A-E	0.81 A-E	2.43 C-K	1.60 A-D	0.66 B-I
Stampasaccusu	22.8 J	35.3 K	53.5 A	5.33 A-I	1.15 E-J	22.1 I-L	8.3 C-F	0.75 B	3.17 C-J	2.06 C-F	0.65 E-H	2.47 C-J	1.17 KL	0.48 MN
Sunda G.	52.5 A-D	59.3 A-D	69.5 A	6.01 A-F	1.33 A-I	22.3 I-L	4.0 E F	0.75 B	3.40 A-E	2.83 AB	0.83 A-E	2.40 C-K	1.61 A-D	0.67 A-H
Sunda N.	56.5 A	62.3 A	72.5 A	5.33 A-I	1.25 B-J	23.5 H-L	6.5 D-F	2.75 B	3.13 C-J	2.30 B-F	0.74 B-H	2.36 C-N	1.40 C-K	0.59 G-L
Troito A	55.8 AB	61.3 A-C	70.3 A	6.65 A-D	1.29 B-J	19.6 L	4.5 D-F	0.50 B	3.38 B-G	2.63 A-D	0.78 A-G	2.38 C-L	1.55 A-F	0.65 B-I
Troito B	50.3 A-G	56.0 A-F	72.0 A	3.90 E-K	1.28 B-J	32.9 E-G	1.2 E F	1.00 B	2.84 D-J	2.29 B-F	0.81 A-E	2.11 H-P	1.38 D-L	0.66 B-I
Vargiu	41.8 A-I	48.3 A-K	63.3 A	4.77 A-J	1.15 F-J	24.4 H-L	1.5 E F	2.25 B	2.77 D-J	2.25 B-F	0.81 A-E	2.08 J-P	1.44 A-J	0.69 A-E
Vavani Perra	31.3 I J	41.8 F-K	64.5 A	4.69 A-J	1.16 E-J	24.7 H-L	6.5 D-F	0.00 B	2.63 G-J	2.12 B-F	0.81 A-E	1.98 M-P	1.35 E-L	0.69 A-F
Jordanolo	46.8 A-I	55.3 A-G	66.0 A	2.86 I-K	1.91 A	66.9 AB	3.0 E F	0.00 B	4.41 A	1.91 D-F	0.43 H	3.80 A	1.43 A-J	0.38 N
Ne plus ultra	38.5 C-J	43.8 E-K	61.5 A	2.81 I-K	1.66 A-F	59.3 BC	21.8 A-F	1.00 B	4.14 A	2.26 B-F	0.55 H	2.88 B	1.35 E-L	0.47 MN
Nonpareil	50.5 A-G	58.8 A-D	70.5 A	1.54 K	1.15 E-J	74.5 A	1.0 E F	8.02 A	3.12 C-J	1.79 F	0.58 G-H	2.43 C-K	1.27 H-L	0.51 LM
Picantili	54.5 A-C	61.8 AB	72.8 A	2.91 I-K	1.64 A-G	56.2 CD	8.0 C-F	3.00 AB	3.46 A-E	2.39 A-F	0.69 D-H	2.52 B-G	1.54 A-F	0.61 E-I
Mean ± std.dev	**41.4** ± 7.71	**48.6** ± 7.09	**63.7** ± 5.08	**5.09** ± 1.42	**1.41** ± 0.24	**31.0** ± 13.7	**9.5** 11.0	**1.1** ± 1.39	**3.16** ± 0.38	**2.38** ± 0.27	**0.76** ± 0.10	**2.37** ± 0.32	**1.46** ± 0.14	**0.62** ± 0.08

(a) Levels not connected by same letters are significantly different

Tab. 2.8 One-way ANOVA, x = genetic cluster from STRUCTURE

Sardinian + three american Commercial					*Sardinian + Picantili*			
Phenotypic trait [(a)]	**DF**	**SS**	**F**	**P**	**DF**	**SS**	**F**	**P**
bloom starting (days from Jan 1st)	4	318.82	1.41	N.S	3	274.63	1.59	N.S
bloom peak (days from Jan 1st)	4	329.86	1.73	N.S	3	272.74	1.93	N.S
bloom end (days from Jan 1st)	4	104.25	1.00	N.S	3	90.38	1.15	N.S
nut weight (gr)	4	27.43	4.50	<0.01	3	4.17	0.88	N.S
kernel weight (gr)	4	0.22	0.99	N.S	3	0.14	0.90	N.S
kernel %	4	0.62	14.55	<0.0001	3	0.06	1.80	N.S
double kernels (%)	4	0.08	1.65	N.S	3	0.08	2.21	N.S
nuts with kernel failed (%)	4	0.0011	1.52	N.S	3	0.00003	0.12	N.S
nut length (cm)	4	2.13	5.21	<0.01	3	0.42	1.71	N.S
nut width (cm)	4	0.76	3.04	<0.05	3	0.25	1.34	N.S
Nut width/length	4	0.19	7.72	<0.001	3	0.003	0.15	N.S
kernel length (cm)	4	1.69	6.23	<0.001	3	0.28	2.04	N.S
kernel width (cm)	4	0.13	1.77	N.S	3	0.09	1.55	N.S
kernel width/length	4	0.09	4.92	<0.01	3	0.004	0.28	N.S

[(a)] parameter expressed in % were transformed in arcsen for statistical elaborations

Tab. 2.9 Tukey significant differences among means of one-way ANOVA, x= genetic cluster [(a)]

Genetic cluster	Nut weight (gr)	Kernel %	Nut length (cm)	Nut width (cm)	Nut width/length	Kernel length (cm)	Kernel width/length
1A1	4.79 A	34.8 B	3.1 B	2.41 AB	0.77 A	2.38 B	0.63 A
1A2	5.15 A	25.9 B	2.9 B	2.25 AB	0.78 A	2.15 B	0.65 A
1A3	5.40 A	27.3 B	3.1 B	2.46 AB	0.79 A	2.36 B	0.63 A
1B	5.73 A	25.1 B	3.2 B	2.46 AB	0.78 A	2.36 B	0.65 A
2	2.40 B	67.2 A	3.9 A	1.99 B	0.52 B	3.04 A	0.46 B

[(a)] Levels not connected by same letters are significantly different

2.3.4b Multivariate cluster analysis and MANOVA

Phenotypical traits underwent also a multivariate cluster analysis, in order to verify the presence of phenotypic grouping among the accessions: the output dendrogram according to Ward, is shown in fig. 2.4. From the successive one-way MANOVA analysis, the first partition ($F=8.406$, $p<10^{-4}$) separated the violet group (I) from the remaining (black arrows cuts); from the second partition, (the most probable, with $F = 11.243$, $p<10^{-4}$) the yellow (II) and blue (III) groups separated (red arrows cuts).

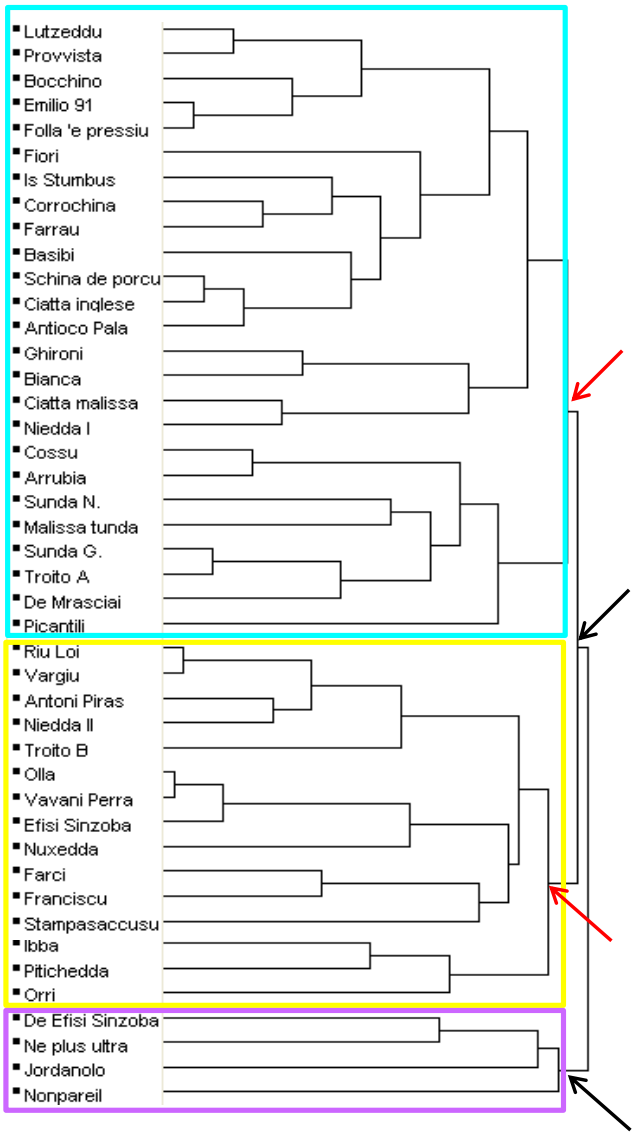

Fig. 2.4 *Phenotypic clustering of accessions*

Table 2.10 shows the ANOVA means grouped for the phenotypic groups. Among the three groups, n. II has the earliest bloom; that exposes those accessions to a high loss of flowers, due to the frost, and thus to a low production. Groups I and III have a similar bloom period, but the kernel percentage value, parameter that strongly influences productivity, is on average much higher in group I (about 65% versus 27%). Group I also includes the Sardinian accession De Efisi Sinzoba and three commercial varieties (the American ones); commercial Picantili grouped within group III, confirming its separation from the other commercial variety already highlighted in genetic clustering.

Tab 2.10 Means of phenotypic groups for phenotypic traits

ACCESSION	Bloom starting (days from Jan 1st)	Bloom peak (days from Jan 1st)	Bloom end (days from Jan 1st)	nut weight (gr)	Kernel weight (gr)	kernel % or shelling %	double kernels %	Nuts with kernel failed %	Nut length (cm)	Nut width (cm)	Nut shape (nut width/ length)	Kernel length (cm)	Kernel width (cm)	Kernel shape (nut width/ length)
De Efisi Sinzoba	40.0	48.0	65.0	2.28	1.39	60.7	14.3	0.25	3.11	1.81	0.58	2.45	1.21	0.49
Jordanolo	47.0	55.0	66.0	2.86	1.91	66.9	3.0	0.00	4.41	1.91	0.43	3.80	1.43	0.38
Nonpareil	50.0	59.0	70.0	1.54	1.15	74.6	1.0	8.00	3.12	1.79	0.58	2.43	1.27	0.53
Ne Plus Ultra	38.0	44.0	61.0	2.80	1.66	59.2	21.8	1.00	4.14	2.26	0.55	2.88	1.35	0.47
mean phen. gr I	**43.8±5.7**	**51.5±6.8**	**65.5±3.7**	**2.37±0.61**	**1.53±0.33**	**65.4±7.0**	**10±9.8**	**2.31±3.8**	**3.69±0.68**	**1.94±0.22**	**0.53±0.07**	**2.89±0.64**	**1.31±0.10**	**0.47±0.06**
Riu Loi	41.0	46.0	63.0	4.76	1.23	25.8	4.3	1.5	2.9	2.3	0.8	2.2	1.5	0.7
Olla	31.0	40.0	63.0	4.44	1.13	25.4	10.5	0.3	2.6	2.1	0.8	2.0	1.3	0.7
Ibba	31.0	40.0	58.0	3.86	1.41	36.4	31.5	1.3	2.9	2.1	0.7	2.3	1.3	0.5
Vargiu	42.0	48.0	63.0	4.77	1.15	24.1	1.5	2.3	2.8	2.3	0.8	2.1	1.4	0.7
Farci	40.0	47.0	62.0	4.31	0.98	22.6	0.5	1.5	3.1	2.3	0.7	2.2	1.3	0.6
Antoni Piras	44.0	52.0	60.0	5.65	1.31	23.2	4.5	1.5	3.2	2.3	0.7	2.3	1.4	0.6
Stampasaccusu	23.0	35.0	53.0	5.33	1.15	21.6	8.3	0.8	3.2	2.1	0.7	2.5	1.2	0.5
Nuxedda	41.0	47.0	65.0	4.14	0.81	19.6	0.0	1.8	2.4	2.2	0.9	1.7	1.3	0.7
Troito B	50.0	56.0	72.0	3.90	1.28	32.7	1.3	1.0	2.8	2.3	0.8	2.1	1.4	0.7
Niedda II	42.0	47.0	69.0	5.94	1.31	22.0	1.5	0.8	3.1	2.3	0.8	2.3	1.5	0.6
Efisi Sinzoba	32.0	40.0	63.0	4.62	1.18	25.6	9.0	0.5	2.7	2.2	0.8	2.1	1.4	0.7
Vavani Perra	31.0	42.0	64.0	4.69	1.15	24.6	6.5	0.0	2.6	2.1	0.8	2.0	1.3	0.7
Orri	39.0	45.0	59.0	4.51	1.30	28.8	26.5	4.5	2.6	2.2	0.9	2.0	1.4	0.7
Pitichedda	35.0	44.0	56.0	3.04	1.54	50.5	42.5	0.0	2.5	2.2	0.9	1.9	1.4	0.7
Franciscu	41.0	47.0	64.0	3.74	1.01	26.9	9.8	0.3	2.8	1.9	0.7	2.1	1.1	0.5
mean phen gr II	**37.5±6.9**	**45.1±5.2**	**62.3±4.7**	**4.51±0.75**	**1.19±0.18**	**27.3±7.7**	**10.5±12.8**	**1.18±1.14**	**2.83±0.24**	**2.19±0.10**	**0.78±0.07**	**2.11±0.18**	**1.34±0.10**	**0.6±0.08**

continues….

ACCESSION	Bloom starting (days from Jan 1st)	Bloom peak (days from Jan 1st)	Bloom end (days from Jan 1st)	nut weight (gr)	Kernel weight (gr)	kernel % or shelling %	double kernels %	Nuts with kernel failed %	Nut length (cm)	Nut width (cm)	Nut shape (nut width/ length)	Kernel length (cm)	Kernel width (cm)	Kernel shape (nut width/ length)
Lutzeddu	33.0	41.0	57.0	7.04	1.69	24.1	16.8	1.5	3.5	2.5	0.7	2.5	1.5	0.6
Fiori	35.0	43.0	60.0	3.27	1.70	52.1	3.8	0.3	3.2	2.3	0.7	2.5	1.5	0.6
Provvista	34.0	41.0	61.0	5.73	1.45	25.3	21.0	1.8	3.4	2.4	0.7	2.4	1.5	0.6
Ghironi	44.0	51.0	64.0	6.97	1.60	22.9	29.5	0.5	3.3	2.7	0.8	2.5	1.7	0.7
Is Stumbus	39.0	45.0	63.0	5.33	1.49	27.9	0.5	0.0	3.5	2.4	0.7	2.8	1.5	0.5
Bocchino	32.0	38.0	55.0	6.32	1.62	25.6	3.3	1.8	3.1	2.4	0.8	2.4	1.6	0.7
Cossu	51.0	59.0	73.0	5.93	1.67	28.1	6.8	1.8	3.2	2.6	0.8	2.4	1.7	0.7
Sunda N.	56.0	62.0	72.0	5.33	1.25	23.5	6.5	2.8	3.1	2.3	0.7	2.4	1.4	0.6
Sunda G.	52.0	59.0	69.0	6.01	1.33	22.2	4.0	0.8	3.4	2.8	0.8	2.4	1.6	0.7
Basibi	38.0	45.0	63.0	6.86	1.40	20.3	0.5	1.0	3.2	2.6	0.8	2.4	1.7	0.7
Arrubia	52.0	59.0	72.0	6.17	1.51	24.5	0.8	0.8	3.2	2.6	0.8	2.5	1.6	0.6
Corrochina	37.0	45.0	63.0	4.93	1.43	29.0	0.8	1.0	3.4	2.6	0.8	2.7	1.4	0.5
Bianca	49.0	54.0	70.0	6.98	1.82	26.1	40.0	0.3	3.3	2.6	0.8	2.4	1.6	0.7
De Mrasciai	51.0	58.0	67.0	6.00	1.25	20.8	0.0	0.0	2.8	2.7	1.0	2.1	1.5	0.7
Troito A	56.0	61.0	70.0	6.65	1.29	19.4	4.5	0.5	3.4	2.6	0.8	2.4	1.5	0.6
Emilio 91	34.0	39.0	55.0	6.48	1.71	26.4	4.8	0.5	3.3	2.6	0.8	2.4	1.6	0.7
Schina de porcu	40.0	47.0	63.0	4.45	1.55	34.8	4.0	0.3	3.2	2.6	0.8	2.4	1.6	0.7
Ciatta malissa	41.0	48.0	62.0	6.77	1.50	22.1	9.3	1.0	3.0	2.8	0.9	2.2	1.7	0.7
Ciatta inglese	42.0	48.0	64.0	5.24	1.59	30.4	5.3	1.5	3.2	2.5	0.8	2.4	1.6	0.7
Niedda I	42.0	48.0	59.0	6.59	1.59	24.2	17.3	0.5	3.4	3.1	0.9	2.4	1.6	0.7
Antioco Pala	43.0	48.0	62.0	5.40	1.51	27.9	17.8	0.8	3.1	2.5	0.8	2.5	1.5	0.6
Farrau	39.0	52.0	60.0	6.30	1.47	23.3	0.0	0.3	3.3	2.5	0.7	2.5	1.5	0.6
Folla 'e pressiu	35.0	42.0	59.0	6.52	1.68	25.8	8.8	1.3	3.2	2.6	0.8	2.4	1.6	0.7
Malissa tunda	50.0	55.0	66.0	6.55	1.35	20.6	2.5	0.8	3.7	2.5	0.7	2.6	1.5	0.6
Picantili	54.0	62.0	73.0	2.91	1.63	56.2	8.0	3.0	3.5	2.4	0.7	2.5	1.5	0.6
mean phen gr III	**43.2±7.8**	**50±7.6**	**64.1±5.4**	**5.87±1.08**	**1.52±0.15**	**27.3±8.79**	**8.6±9.97**	**0.97±0.79**	**3.3±0.18**	**2.6±0.17**	**0.79±0.07**	**2.4±0.13**	**1.6±0.08**	**0.6±0.05**

To verify if, for the combination of all the fourteen traits, the *Structure* groups are significantly diverse, a one-way MANOVA analysis was carried out. This multivariate analysis indicated that the genetic groups were differentiated when all traits were considered together (F=2.70, Wilks p<0.0001). However the one-way MANOVA carried out on *Structure* groups including only Sardinian cvs. was not significant, confirming the ANOVA results.

2.3.4c Contingency analysis

Finally, the contingency analysis, through Pearson Chi-squared test, was carried out in order to evaluate the distribution of genetic clusters within each phenotypic cluster: phenotypic and genetic clusters showed a significant association (Pearson 38.07, p <0.0001) as reported in fig. 2.5. Phenotypic group I almost coincides with CL 2 genetic cluster, as three of its four accessions are the three American cvs.; the forth one is the Sardinian "De Efisi Sinzoba", that therefore is the only Sardinian that can be compared to the American varieties. Phenotypic group II is more associated to 1A2 and 1A3 genetic clusters; and within phenotypic group II there is a predominance of 1A3 genetic cluster.

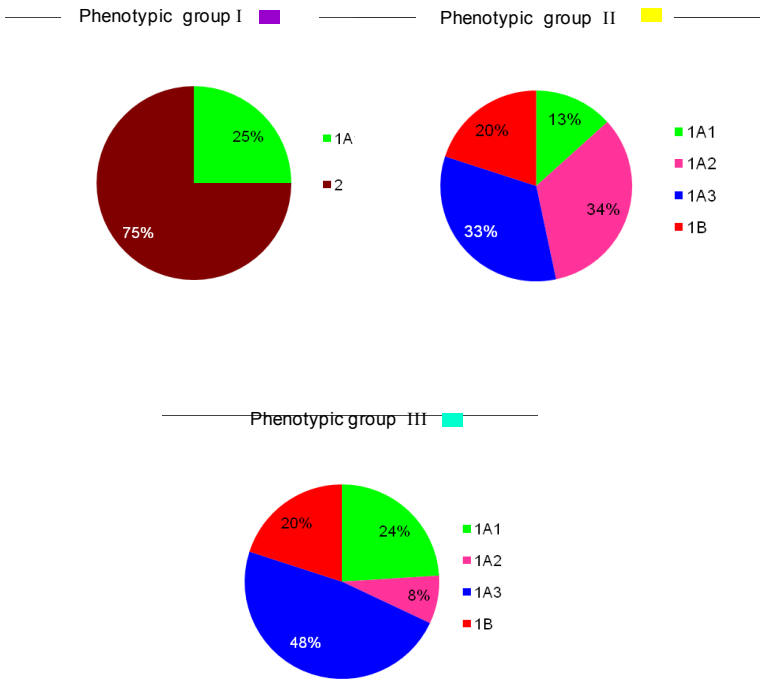

Fig. 2.5 *Distribution of genetic clusters within each phenotypic group*

2.3.5 Conclusions

In conclusion, from the comparison between Sardinian and commercial accessions it is important to underline the finding that both in terms of allelic richness and H_o the Sardinian genotypes present higher level of genetic diversity than the commercial varieties. An interesting case was represented by Picantili, imported from Ukraine, that is included in cluster 1A together with the majority of Sardinian cultivars. This finding is in agreement with observations reported in chapter 1 that suggested an exchange between sardinian and continental almond genotypes. At phenotypic level Sardinian germplam was clearly distinguished from the commercial varieties. On the other hand it was very interesting to note that the differences among Sardinian model based groups were not significant for any of the fourteen analyzed traits. A possible interpretation for this observation is that the sardinian genotypes are highly adapted to the enviromental conditions of the island and thus although genetically different they have a remarkable canalization of phenotypic traits.

2.4 Bibliography

Amirbakhtiar N., Shiran B., Moradi H. and Sayed-Tabatabaei B.E. (2006). Molecular characterization of Almond Cultivars Using Microsatellite Markers Proc. IVth IS on Pistachios and Almonds. Acta Hort. 726, ISHS 2006.

Aranzana M.J., Garcia-Mas J., Carboa J. and Arus P. (2002) Development and variability of micro satellite 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 Arus P. (2003). A set of simple-sequence repeat (SSR) markers covering the Prunus genome. Theor Appl Genet 106: 819–825.

Avanzato D. (2006). In "Following Almond Footprints *(Amygdalus communis* L.*)* Scripta Horticulturae n. 4, pg. 80 (a publication of ISHS).

Cipriani G., Lot G., Huang WG., Marrazzo M.T., Peterlunger E. and Testolin R. (1999). AC/GT and AG/CT microsatellite repeats in peach [Prunus persica (L) Batsch]: isolation, characterisation and cross-species amplification in Prunus. Theor. Appl. Genet. 99: 65-72.

Dangl G.S., Yang J., Golino D.A. and Gradziel T. (2009). A practical method for almond cultivar identification and parental analysis using simple sequence repeat markers. Euphytica 168: 41–48.

De Giorgio D. and Polignano G.B. (1999). Evaluating the biodiversity of almond cultivars from a germplasm collection field in Southern Italy. In: Stott D.E., Mothar R.H., Steinhardt G.C. (eds.) 2001. Sustaining the global farm: 305-311.

Dirlewanger E., Cosson P., Tavaud M., Aranzana M.J., Poizat C., Zanetto A., Arús P. and Laigret F. (2002). Development of microsatellite markers in peach [*Prunus persica*

(L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.) Theor. Appl. Genet. 105: 127-138.

Evanno G., Regnaut S. and Goudet J. (2005). How to use the software Structure: a simulation study. Molecular Ecology 14: 2611-2620.

Goudet J. (2002). FSTAT vers 2.9.3.2 Available from http://www .unil.ch/izea/softwares/fstat.html. Updated from Goudet J. (1995). FSTAT (vers. 1.2): a computer program to calculate F-statistics. J. Hered. 86: 485-486.

JMP ver. 7 . Copyright © 2007 by SAS Institute Inc., Cary, NC, USA

Kester D.E. and Ross N.W. (1996) in "Almond production manual". Publication 3364, University of California, Division of Agriculture and Natural Resources (ANR). Pg. 1-2.

Ladizinsky G. (1999). On the origin of almond. Genetic Resources and Crop Evolution 46: 143-147.

Miller M.P. (1997). Tools for population genetic analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author. Copright (C) 1997 Mark P. Miller. Department of Biological Sciences, Nothern Arizona University, Box 5640, Flagstaff, AZ 86011-5640. E-mail: Mark.Miller@usu.edu.

Nei M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583-590.

Pritchard J.K., Wen X. and Falush D. (2007). Documentation for Structure software, Version 2.2. In: http://pritch.bsd.uchicago.edu/software.

PSR 2007-2013 available on web at http://www.regione.sardegna.it/documenti/14_43_20071121202527.pdf , pg 9-10.

Shiran B., Amirbakhtiar N., Kiani S., Mohammadi Sh., Sayed-Tabatabaei B.E., and Moradi H. (2007). Molecular characterization and genetic relationship among almond cultivars assessed by RAPD and SSR markers. Scientia Horticulturae 111: 280–292.

Sneath P.H.A. and Sokal R.R. (1973). In: Numerical taxonomy. Freeman WH and Company (eds) San Fransisco, USA.

Sorkheh K., Shiran B., Gradziel T.M., Epperson B.K., Martinez-Gomez P. and Asadi E. (2007). Amplified fragment length polymorphism as a tool for molecular characterization of almond germplasm: genetic diversity among cultivated genotypes and related wild species of almond, and its relationships with agronomic traits. Euphytica 156: 327–344.

Testolin R., Marrazzo T., Cipriani G., and Quarta R. (2000). Microsatellite DNA in peach and its use in fingerprinting and testing the genetic origin of cultivars. Genome, 43 (3): 512-520.

Ward J.H. (1963). Hierarchical Grouping to optimize an objective function. Journal of American Statistical Association 58 (301): 236-244.

Xie H., Sui Y., Chang F.Q, Xu Y. and Ma R.C. (2006). SSR allelic variation in almond (Prunus dulcis Mill.). Theor. Appl. Genet. 112: 36.

CHAPTER 3

STUDY ON SELF-INCOMPATIBILITY ALLELES OF SARDINIAN ACCESSIONS

3.1 Introduction

3.1.1 Self-incompatibility in plants

Most of flowering plants (Angiosperms) produce hermaphrodite flowers. Due to the proximity of the anther and pistil in such flowers, there is a great tendency for pollen to land on the stigma of the same flower (Kao *et al.* 1996). This condition, promoting self-fertilization and inbreeding, could be unfavourable in evolutive terms. Nevertheless Angiosperms include the most successful species of extant plants. That reproductive success is partially due to the evolution of different mechanisms that limit and/or prevent self fertilization.

Self-incompatibility (SI) is the most used reproductive strategy in flowering plants: it allows the pistil to recognize and reject self-pollen or pollen from related individuals, thereby preventing self-fertilization.

The SI response is based on the self/non-self-recognition process between pollen and pistil that is followed by the selective pollen-tube development inhibition.

Classical genetic studies have classified SI into two systems (see fig. 3.1), gametophytic SI (GSI) and sporophytic SI (SSI), depending on whether the SI phenotype in pollen is determined by the S-genotype of haploid pollen or by the S-genotypes of the diploid anther (Silva and Goring, 2001; Takayama and Isogai, 2005; Iwano and Takayama, 2012).

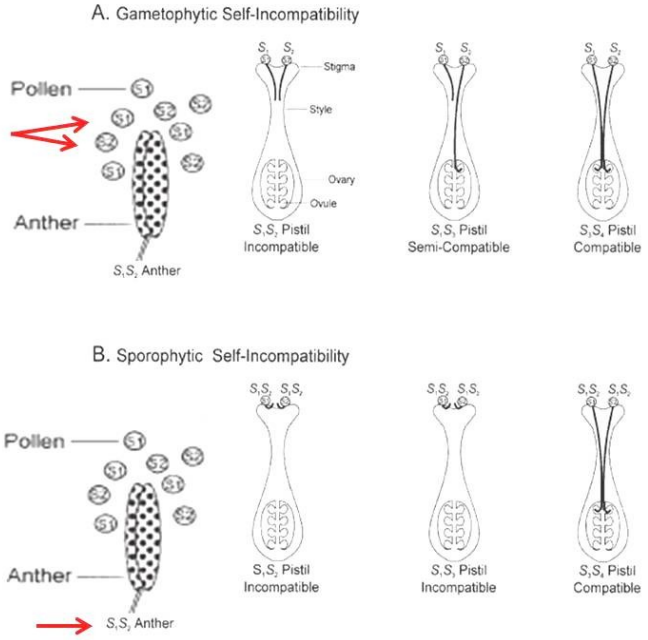

Fig 3.1. *(A) Gametophitic self-incompatibility and (B) sporophitic self-incompatibilit (from Silva and Goring, 2001)*

Fertilization will only be obtained if parents do not share identical SI alleles (cross-compatibility) and as a consequences the progeny is heterozygous for the SI alleles.

In GSI, pollen grains, with the S-allele in common with an S-allele of pistil, are rejected (that is, after germination, pollen-tube elongation is terminated within the style), whereas the others grow through the pistil until carrying out the oosphere fertilization. Therefore three cases can occur: total incompatibility (no fertilization), semi-compatibility (partial fertilization only by the pollen grains that possess the S-allele not in common with the pistil) and compatibility (fertilization).

3.1.2 Genes responsible for self-incompatibility

In most species SI is controlled by a single locus, namely the *S*-locus, which contains at least two genes: one encoding the male determinant that is carried by the pollen grain, and the other encoding the female determinant that is expressed in the pistil. Both the male and female determinants are multiallelic and inherited as one segregating unit. The multiple variants of this gene complex are called *S*-haplotypes (Takayama and Isogai, 2005). Self/non-self discrimination between male and female operates at the level of the protein-

protein interactions between the two determinants. An incompatible response occurs when both determinants are issued from the same *S*-haplotype (fig. 3.2).

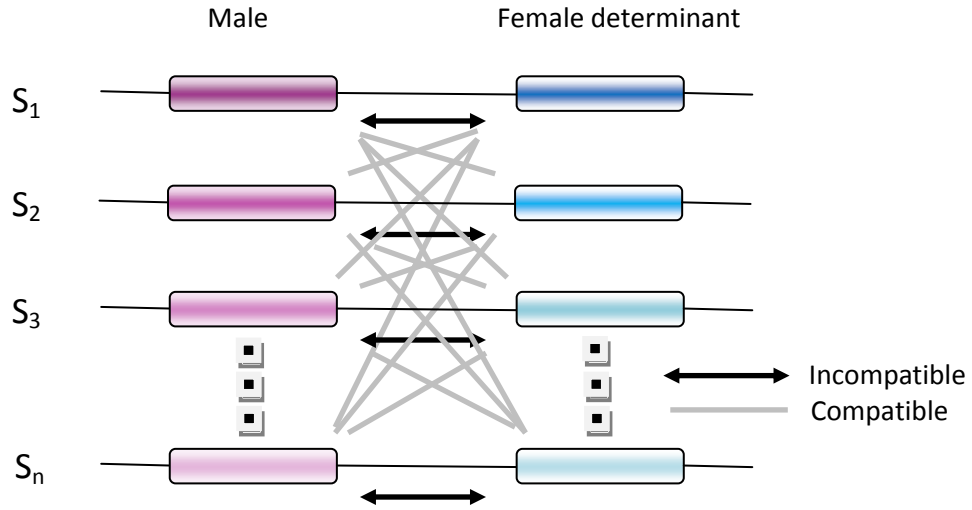

Fig 3.2. *organization of locus-S genes (from Takayama and Isogai, 2005)*

So far, both determinants have been identified in the Brassicaceae, Solanaceae, Rosaceae and Scrophulariaceae, whereas in Papaveraceae only the female determinant is known.

Recent molecular analyses have revealed that GSI and SSI systems contain a few diverged molecular mechanisms.

For example, in SSI, Brassicaceae operate a self-recognition (interaction between male-determinants and female-determinants derived from the same S-haplotype) utilizing pistil-expressed receptor kinase. The determinant genes encode a pollen ligand and its stigmatic receptor kinase; their interaction induces incompatible signaling (s) within the stigma papilla cells (Takayama and Isogai, 2005)

In GSI, Papaveraceae use a self-recognition process, through the intervention of a transmembrane-protein mediated Ca^{2+} signaling system (Wheeler *et al.*, 2009) and the reaction occurs in pollen grain, whereas Solanaceae, Rosaceae and Scrophulariaceae use pistil-expressed S-RNase based non-self recognition system which involves the recognition of non-self partners and disregard of the self partner (Meng *et al.*, 2010, Ushijima *et al.*, 2003; Lai *et al.*, 2002).

3.1.3 The Solanaceae-type self-incompatibility

3.1.3a Female determinant

The Solanaceae (Ai *et al.*, 1990; Tsai *et al.*, 1992; Kaufmann *et al.*, 1991), Rosaceae (Sassa *et al.*, 1993; Tao *et al*., 1997), and Scrophulariaceae (Xue *et al.*, 1996) families all share the same female determinant molecule. The pistil S-protein was first identified, as an S-RNase, in *Nicotiana alata* (Solanaceae), so we refer to this S-RNase-mediated SI as Solanaceae-type SI. The potential function of the pistil S-protein was initially inferred considering the striking degree of sequence similarity with the catalytic site of the RNase T2 of *Aspergillus oryzae* mould (Kawata *et al.*, 1988) (fig. 3.3).

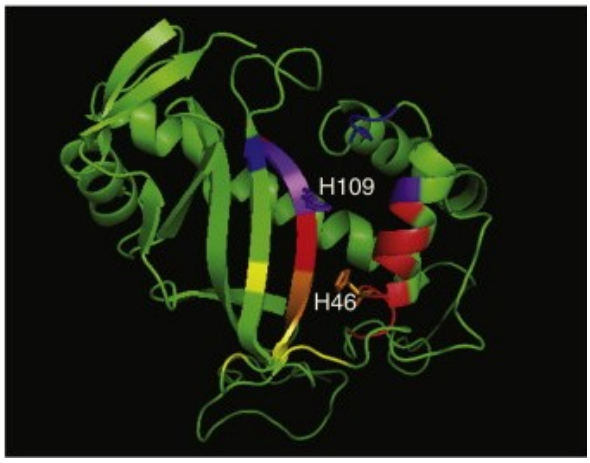

Fig 3.3. *3D structure of RNase T2found in A. oryzae*

Subsequent molecular work confirmed that the pistil S-protein of Solanaceae indeed possess ribonuclease activity (McClure *et al.*, 1989; Singh *et al.*, 1991.

S-RNases were found to be basic glycoproteins of approximately 30 kDa in size with RNase activity, abundantly expressed in the pistil and localized primarily to the stigma and transmitting tract of the style (Cornish *et al.*, 1987).

A comparison of the predicted amino acid sequences of Solanaceae S-RNases from different species revealed an unusually high degree of sequence diversity (Tsai *et al.*, 1992; Clark *et al.*, 1991). However, five highly conserved regions (fig. 3.4), three of which (C1, C4

and C5) primarily composed of hydrophobic amino acids, are embedded within a highly polymorphic protein sequence.

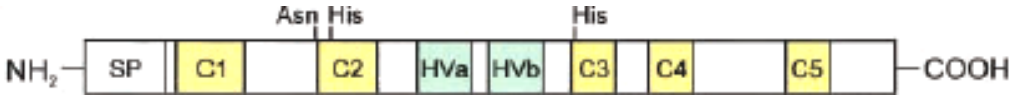

Fig. 3.4. *Schematic representation of the primary structure of S-RNases (adaptation from Silva and Goring, 2001)*

The two remaining conserved regions, C2 and C3, are similar to the active site of the fungal ribonucleases RNase T2 (Kawata *et al.*, 1988). There are two hypervariable regions in the S-RNases from the Solanaceae, termed HVa and HVb, whereas only one hypervariable region, corresponding to HVa, has been detected in the rosaceous S-RNases (Takayama and Isogai, 2005). The hypervariable regions are highly hydrophilic and they probably mediate interactions with pollen S-allele products (Silva and Goring, 2001).

3.1.3b Male determinant

The first clues for the male determinant was obtained from sequence analysis of the S-locus region of *Antirrhinum hispanicum*, a member of the Scrophulariaceae. The region of the S2-haplotype contained a novel F-box protein gene, AhSLF-S2 (*A. hispanicum* S-locus F-box of S2-haplotype), which is specifically expressed in anther and pollen grains of S2-haplotype (Lai *et al.*, 2002).

Genomic analysis of the S-locus of some *Prunus* spp. (Entani *et al.*, 2003), included almond, (Ushijima *et al.*, 2003) revealed polymorphic F-box protein genes that fulfilled the conditions of a pollen S-determinant gene: (a) it is located within the highly divergent genomic region of the S-locus, (b) it exhibits S-haplotype-specific diversity and (c) it is specifically expressed in pollen. They were named S-locus F box (SLF or SFB).

3.1.3c Mechanisms of pollen inhibition in Solanaceae-type self-incompatibility

In spite of the fact that both female and male determinants have been identified, the molecular mechanisms regulating how the encoded molecules interact and inhibit self-pollen growth is not yet clear.

The observations that the ribonuclease activity of S-RNases is essential for pollen rejection (Huang *et al.*, 1994) and pollen RNA is degraded specifically after incompatible pollination (McClure *et al.*, 1990) inspired the "inhibitor model": according to this model the pollen S-determinant would inhibit all S-RNases with the exception of the same aplotype S-RNase. Thus, S-RNases would function as highly specific cytotoxins inhibiting the growth of incompatible pollen. Immunocytochemical experiments indicated that S-RNases enter the pollen tube. Thus the specific inhibition of non-self RNase would occur after importation within the cell (see fig. 3.5).

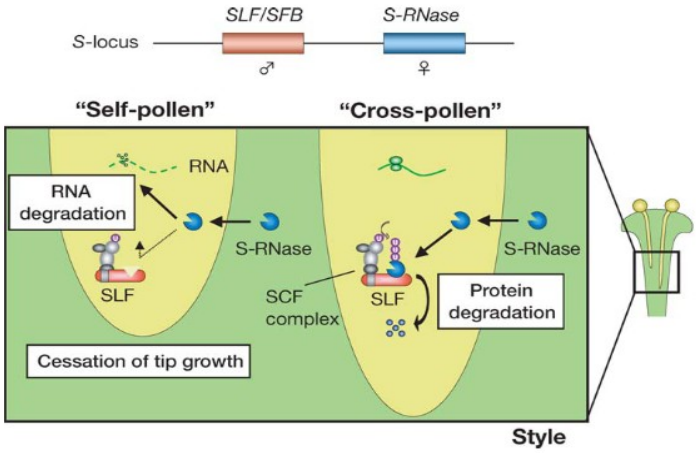

Fig. 3.5. *Proposed molecular mechanism of self-incompatibility response in Solanaceae-type SI* (*from Takayama and Isogai, 2005)*

An SLF allelic variant specifically recognizes its non-self S-RNases and mediates their degradation by the ubiquitin–26S-proteasome system: F-box, with three other proteins, make up the SFC E3 ubiquitin ligase complex (Gagne *et al.*, 2002). The SFC E3 ubiquitin ligases would act in conjunction with the E1, E2 and E3 enzymes in facilitating the transfer of ubiquitin to appropriate target proteins, predisponing them for proteasome mediated degradation. However, SLF shows much lower allelic sequence diversity than S-RNase, and it was amazing how an SLF allelic product could recognize a large number of divergent non-self S-RNases to allow cross-compatible pollinations.

According to a recent 'collaborative non-self recognition' model in *Petunia* (Kubo *et al.* 2010) the S-locus encodes a single female and multiple male S-determinants (fig. 3.6).

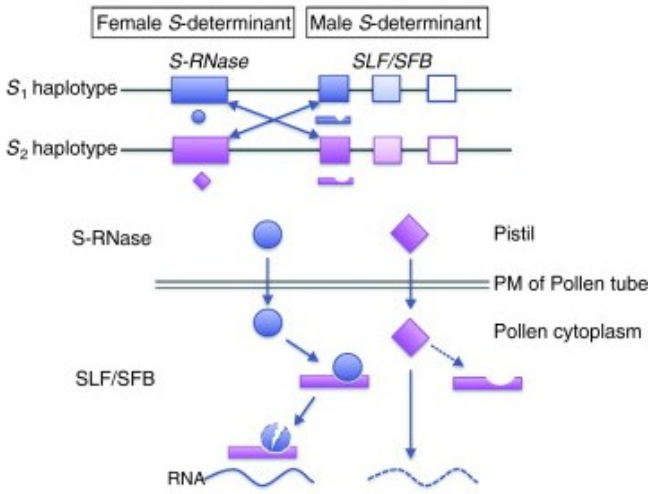

Fig 3.6. *Non-self recognition system in Solanaceae* (*from Iwano and Takayama., 2012)*

In the self-incompatible pollination no SLF would interact with self S-RNase, which therefore could degrade the self-pollen RNA; in self-compatible pollination, the products of multiple types are required for the entire group of non-self S-RNases to be collectively recognized and detoxified, allowing the growth of pollinic tube (Iwano and Takayama, 2012).

In this SI model, increasing the repertoire of SLFs would be advantageous, as this would increase the number of potential mating partners. Other SI species also have a single S-RNase and multiple SLFs in the S-loci (Zhou *et al.* , 2003; Okada *et al.*, 2011). Whether these other species also adopt similar 'collaborative non-self recognition' SI, and how these multiple types of SLF genes have emerged in the S-locus is still unclear.

3.1.4 Self-incompatibility in almond

Most of almond cultivars (*Prunus amygdalus* Batsch) are self-incompatible, although self-compatible cultivars exist (Lopez *et al.*, 2006). Almond, belonging to *Rosaceae*, possesses Solanaceae-type SI (GSI, pistil-expressed S-RNase based non-self recognition system). Therefore its multiple, codominant, S-RNases encoding alleles possess, within their coding sequences, only one hypervariable region (HV or RHV), that corresponds to HVa of Solanaceae-type S-RNases; and five conserved regions, with C4 named RC4, which is Rosaceae-specific (Ushijima *et al.*, 1998). They also contain two polymorphic introns: one

within the RHV region, like the other *Rosaceae*, and another located upstream the C1 region, between the regions coding for the signal peptide and the mature protein, proper of *Prunus* genus (fig. 3.7; Igik and Kohn, 2001; Sonneveld *et al.*, 2003).
Polymorphisms of introns are primarily responsible for the S-allele variability.

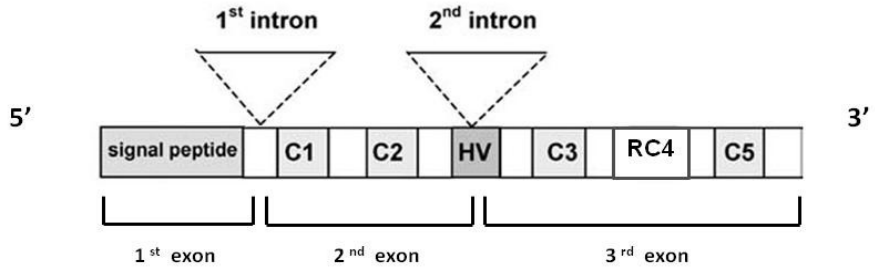

Fig 3.7. *S-RNase gene structure in Prunus spp. (adaptation from Sonneveld et al., 2003)*

Incompatibility genotypes were traditionally identified by controlled pollination; the first identified ones were S_1, S_2, S_3, S_4, S_7, S_8 self-incompatible and S_f compatible alleles (Crossa-Raynaud and Grasselly, 1985); the last one is characteristic of some European cultivars, including the Italian cultivar from Apulia "Tuono".
In 1994, Kester *et al.* identified four alleles in some American cultivars (S_a, S_b, S_c, S_d). Three years later Boskovic *et al.* (1997) demonstrated that self-incompatibility alleles code for stylar ribonucleases and determined the S genotype of 29 cultivars through IEF (isoelectric focusing) of proteins. Such study confirmed the results from Crossa-Raynaud and Grasselly and from Kester, because the American S_b coincided with the French allele S_1, S_c with S_7, S_d with S_8, and S_a , not having a corresponding allele, was redesignated S_5; moreover two new alleles were identified (S_6 and S_9). Recent studies were based on PCR with consensus primers designed: I) on conserved regions within Rosaceae s-RNases (Tamura *et al*., 2000; Sonneveld *et al.*, 2003; Sutherland *et al.*, 2004; Ortega *et al.*, 2005; Halsaz *et al.*, 2008); II) on specific sequence of S-alleles (Ma and Oliveira, 2001; Channuntapipat *et al*., 2003).

So far, thirty-seven S-incompatibility alleles, in addition to the S_f self-compatibility allele, and twenty-seven cross-incompatible groups have been identified (Kodad *et al.*, 2010a).

The present study aimed to identify the SI genotypes of the sweet almond genotypes of the collection field already characterized through SSR markers and phenotypic traits (see second chapter).

3.2 Materials and methods

3.2.1 Plant material

Plant material came from eighteen cultivars of Sardinian almond (Tab. 3.1) conserved in the collection field of AGRIS Sardegna (see second chapter) .

DNA was extracted from young leaves as reported in the first and second chapter.

Tab. 3.1 Accessions addressed in this study

Accession No.	cultivar	collection site/origin
1	ANTONI PIRAS	Sardinia
2	ARRUBIA	Sardinia
3	CIATTA INGLESE	Sardinia
4	DE EFISI SINZOBA	Sardinia
5	FARCI	Sardinia
6	FARRAU	Sardinia
7	FIORI	Sardinia
8	FOLLA 'E PRESSIU	Sardinia
9	FRANCISCU	Sardinia
10	IBBA	Sardinia
11	IS STUMBUS	Sardinia
12	NIEDDA II	Sardinia
13	NUXEDDA	Sardinia
14	OLLA	Sardinia
15	PITICHEDDA	Sardinia
16	RIU LOI	Sardinia
17	SCHINA DE PORCU	Sardinia
18	STAMPASACCUSU	Sardinia

3.2.2. Amplification of almond self-incompatibility alleles through PCR

Ortega *et al.* (2005) developed a reliable PCR approach for determining incompatibility S- genotypes, based on the lengths of first and second introns.

We amplified the genic fragment including first and second introns through the following consensus primers developed for Rosaceae: Pa-cons1F (5'- (C/A)CT TGT TCT TG(C/G) TTT (T/C)GC TTT CTT C -3', Sonneveld *et al.*, 2003); and EM-PC3consRd (5'-AWS-

TRC-CRT-GYT-TGT-TCC-ATT-C-3', Sutherland *et al.*, 2004). More in details the amplification included: the first exon, (substantially constituted by the signal peptide), the first intron, the second exon with C1, C2 regions; the RHV region, that contains the second intron; and the beginning of the third exon, with part of C3 (see fig 3.8).

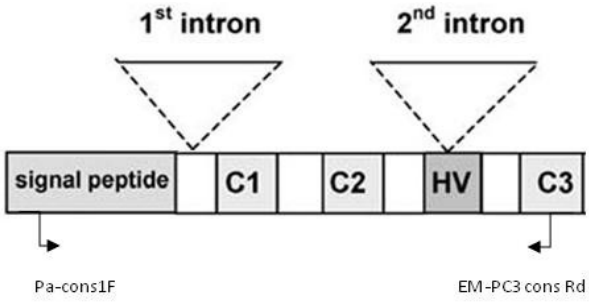

Fig. 3.8. *Primer ubications within Prunus S-RNase gene*

PCR reaction was performed in a 25 µl volume containing: Buffer 1/10 v/v, $MgCl_2$ 2.5 mM, dNTP mix 0.2 mM, primers 0.3 µM each, DNA 50 ng and 0.625 U Taq polymerase (*Platinum, InVitrogen*: it is endowed with terminaltranspherase activity, that adds a single dA nucleotide at the 3' ends of amplification product). Amplification was carried out in a Gene Amp PCR System 9700 Thermal Cycler (Applied Biosystem), with the following cycles:

Step 1:	94°C	2'	
Step 2:	94°C	10''	x 35 cycles
Step 3:	58°C	2'	
Step 4:	72°C	2'	
Step 5:	72°C	10'	

Amplification results were checked through electrophoresis on 1.5% agarose gel in TAE (Trizma Base 2 M, Acetic Acid 5.7%, EDTA 0.5 M), stained with ethidium bromide.

3.2.3 S-gene sequence amplification and cloning

3.2.3a Plasmid-insert ligation

PCR products were cloned, as cloning gives the possibility of studing the S-haplotype of each accession by getting clones containing each of the two alleles. Cloning was carried out through the "TOPO TA Cloning Kit® for Sequencing" in the pCR®4-TOPO® vector (both from InVitrogen). TA cloning methodology exploit the terminaltranspherase activity of several Taq (such "Platinum"), allowing the direct insertion of PCR products in vectors with a single dT nucleotide at the *3'* end.

A 2.5 µL aliquot from each dA template (one per cultivar) was added to each reaction of TOPO TA Cloning mix, constituted by vector with dT end plus Topoisomerase I enzyme (from *Vaccinia*) bound to the cloning vector (Fig. 3.9).

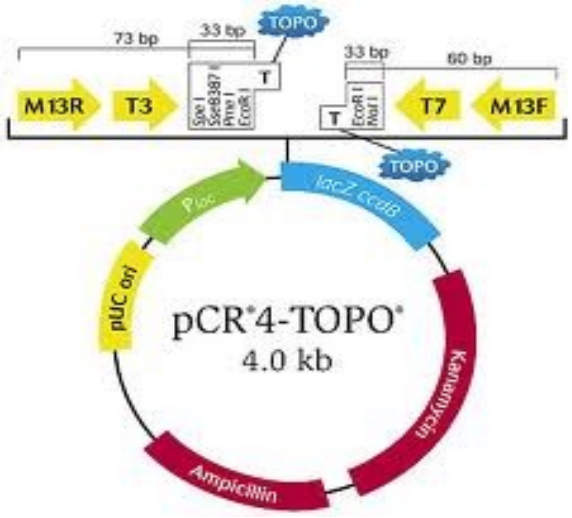

Fig. 3.9. *pCR®4-TOPO® vector characteristics*

3.2.3b Competent cells transformation

Competent bacterial cells (*E.coli* strain DH5α or TOP10, included in the kit), were transformed with the recombinant plasmidic vector pCR®4-TOPO® (see par. 2.3.1.) through thermal shock, according to the kit manual protocol, as follows: each competent cell aliquot was added with 2 µL of a TOPO TA Cloning® reaction, then incubated 20' in ice, followed by 30'' in water bath at 42 ° C and by a final immediate incubation in ice.

Each aliquot of putatively transformed bacterial cells was added with 300 µL of SOC (Triptone 2%, yeast extract 0.5%, NaCl 0.05%, KCl_2 1.86%, $MgCl_2$ 9.5%) and shaken 1 h at 37 °C. After that, each bacterial culture was poured, in a Petri dish, on LB media (Triptone 1%, yeast extract 0.5%, NaCl 1%, Bacto-agar 1.5%) added with kanamicin (100 µg/mL) and grown over night at 37 °C. Only transformed cells (containing the plasmid that confers kanamicin resistance) grew and formed colonies.

In order to check the transformation occurrence, colony PCRs were carried out: by means of sterile tips some colonies (5 to 10 per accession) were picked from each Petri dish and put in PCR tubes for PCR amplification with M13 primers (one colony in each tube, with a 10 µL volume of PCR mix containing: Buffer 1/10 v/v, $MgCl_2$ 1.5 mM, dNTP mix 0.4 mM, primers 0.2 µM each, and 1U/µL Taq polymerase (*Recombinant, InVitrogen*). The amplification protocol, carried out in the Gene Amp PCR System 9700 Thermal Cycler (Applied Biosystem), was the following:

Step 1:	94°C	45’	
Step 2:	94°C	40’’	x 35 cycles
Step 3:	50°C	40’’	
Step 4:	72°C	2’	
Step 5:	72°C	7’	

Transformants were identified on 1.5% agarose gel in TAE (Trizma Base 2 M, Acetic Acid 5.7%, EDTA 0.5 M) stained with ethidium bromide.

3.2.3c Plasmid extraction

Transformed plasmids were extracted through the “GenElute Plant Genomic DNA miniprep kit” (SIGMA) and DNA quality was checked on 1% agarose gel stained with ethidium bromide.

Plasmid DNAs were sent to the “Macrogen on-line Sequencing order system” (Seoul, Korea) for double strand sequencing.

3.2.4 Sequence analysis

Sequences obtained were first analyzed through Vec-screen software, available on web at http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html, that eliminated the vector regions upstream and downstream the M13 primers.

In order to study nucleotide homology, vector-free sequences (5 to 10 per accession) were compared to the GeneBank/EMBL/DDBJ data base through the BLAST Nucleotide software (Basic Local alignment Search Tool, available on web at http://www.ncbi.nlm.nih.gov/Blast).

All the sequences sharing homology to *Prunus* S-alleles were downloaded and aligned through CLUSTALW (Thompson *et al*. 1994), available on web at (www.ebi.ac.uk/Tools/msa/clustalw2/). For each accession, from the alignment of its 5 to 10 clones with their reference sequence, two putative alleles were chosen, whose consensus sequences were obtained. Each consensus sequence is constituted by the connection of the contig sequences coming from forward and reverse alignments of each clone, as shown in fig. 3.10 (example: clone 6-1 of "Fiori"):

```
6-1F     ACTTGTTCTTGGTTTTGCTTTCTTCTTTTGTTACGTTATGAGCAGTGGTGGGTTGCATTA 60
6-1cons  ACTTGTTCTTGGTTTTGCTTTCTTCTTTTGTTACGTTATGAGCAGTGGTGGGTTGCATTA 60
6-1R     ------------------------------------------------------------

6-1F     CAATCTTTCCCTATATATCCTATATGTATATAATCAGCATTGCGTTTCTCTACTTGTATT 120
6-1cons  CAATCTTTCCCTATATATCCTATATGTATATAATCAGCATTGCGTTTCTCTACTTGTATT 120
6-1R     ------------------------------------------------------------

6-1F     TTTTGTTCAGAGAAACTGTTGTGTGTGTTGGATGATATATCACATGACATGCGGTGTATT 180
6-1cons  TTTTGTTCAGAGAAACTGTTGTGTGTGTTGGATGATATATCACATGACATGCGGTGTATT 180
6-1R     ------------------------------------------------------------

6-1F     GATTTCACCCACATATTTGGCATTTAATCTAACGCACAACTTTTTTTGGATGAGTATTTG 240
6-1cons  GATTTCACCCACATATTTGGCATTTAATCTAACGCACAACTTTTTTTGGATGAGTATTTG 240
6-1R     ----------------------------------------------------GGGT-TTTG 8
                                                              * ** ****
6-1F     GTGATTGTTTTTTCTGCATGTGCTCTTTTTATTTTATTAATCTTTTGTTTATTTTGATGAT 300
6-1cons  GTGATTGTTTTTTCTGCATGTGCTCTTTTTATTTTATTAATCTTTTGTTTATTTTGATGAT 300
6-1R     GGGATTGTTTTCC-GCAAGGGCTCTTTTTATTT-AATAATCTTTTGTTTATTTTGAAGAT 66
         * ********* * *** * ************ * ******************** ***

6-1F     TGTTTGCAATAAGTGGAGTCTATTCATCACAATAATTTTGGCAGGATCTTATGGGTATTT 360
6-1cons  TGTTTGCAATAAGTGGAGTCTATTCATCACAATAATTTTGGCAGGATCTTATGGGTATTT 360
6-1R     TGTTTGCAATAAGTGGAGTCTATTCATCACAATAATTTTGGCAGGATCTTATGGGTATTT 126
         ************************************************************

6-1F     TCAATTTGTGCAACAATGGCCACCGACCAACTGCAGAGTTCGCATCAAGCGACCTTGCTC 420
6-1cons  TCAATTTGTGCAACAATGGCCACCGACCAACTGCAGAGTTCGCATCAAGCGACCTTGCTC 420
6-1R     TCAATTTGTGCAACAATGGCCACCGACCAACTGCAGAGTTCGCATCAAGCGACCTTGCTC 186
         ************************************************************
```

continues..

```
6-1F     CAAACCTCGGCCACTACAAAATTTCACCATCCATGGCCTTTGGCCAAGTAATTATTCAAA 480
6-1cons CAAACCTCGGCCACTACAAAATTTCACCATCCATGGCCTTTGGCCAAGTAATTATTCAAA 480
6-1R     CAAACCTCGGCCACTACAAAATTTCACCATCCATGGCCTTTGGCCAAGTAATTATTCAAA 246
         ************************************************************

6-1F     CCCAACGAAGCCCAGTAATTGCAATGGGGCAAAATATGAGGACAGGAAAGTGGTATGTCA 540
6-1cons CCCAACGAAGCCCAGTAATTGCAATGGGGCAAAATATGAGGACAGGAAAGTGGTATGTCA 540
6-1R     CCCAACGAAGCCCAGTAATTGCAATGGGGCAAAATATGAGGACAGGAAAGTGGTATGTCA 306
         ************************************************************

6-1F     TTATTTTTTTATTTTCTCTTTAGTTTTTAGAAAATTAAATTGTCATGTGAAGATAATATA 600
6-1cons TTATTTTTTTATTTTCTCTTTAGTTTTTAGAAAATTAAATTGTCATGTGAAGATAATATA 600
6-1R     TTATTTTTTTATTTTCTCTTTAGTTTTTAGAAAATTAAATTGTCATGTGAAGATAATATA 366
         ************************************************************

6-1F     CTTTCAATGAATCTTTGGGTGTCCTAAAATTTCGATGTCGGTCCTTGTTAGACACATTAT 660
6-1cons CTTTCAATGAATCTTTGGGTGTCCTAAAATTTCGA--TCGGTCCTTGTTAGACACATTAT 658
6-1R     CTTTCAATGAATCTTTGGGTGTCCTAAAATTTCGATGTCGGTCCTTGTTAGACACATTAT 426
         ********************************** ************************

6-1F     TTTGAATAAATAACTACCACGTAGATATTACTTTTATTGAACCACGTAGATATTATGATA 720
6-1cons TTTGAATAAATAACTACCACGTAGATATTACTTTTATTGAACCACGTAGATATTATGATA 718
6-1R     TTTGAATAAATAACTACCACGTAGATATTACTTTTATTGAACCACGTAGATATTATGATA 486
         ************************************************************

6-1F     TCCTCCAATTGAAGGACCTGCTATTATTCTGCATGTATACATTCAAAATACTATACCAAA 780
6-1cons TCCTCCAATTGAAGGACCTGCTATTATTCTGCATGTATACATTCAAAATACTATACCAAA 778
6-1R     TCCTCCAATTGAAGGACCTGCTATTATTCTGCATGTATACATTCAAAATACTATACCAAA 546
         ************************************************************

6-1F     ATCAAATCAGAAAATGATAAATAAAAAAAGTCATTCACACAAGACAAAATGTTAAAGTGG 840
6-1cons ATCAAATCAGAAAATGATAAATAAAAAAAGTCATTCACACAAGACAAAATGTTAAAGTGG 838
6-1R     ATCAAATCAGAAAATGATAAATAAAAAAAGTCATTCACACAAGACAAAATGTTAAAGTGG 606
         ************************************************************

6-1F     TAAATTGGACAAATCGGTCCAGTCATTGTCCGACAATGACATGAAATAACTCTTATGTCA 900
6-1cons TAAATTGGACAAATCGGTCCAGTCATTGTCCGACAATGACATGAAATAACTCTTATGTCA 898
6-1R     TAAATTGGACAAATCGGTCCAGTCATTGTCCGACAATGACATGAAATAACTCTTATGTCA 666
         ************************************************************

6-1F     CTGTCGGACAGTGACCCCAGAGTTGTTCAAAGTCACTCTCAGTGAGTGACTTAAGAGTTT 960
6-1cons CTGTCGGACAGTGACCCCAGAGTTGTTCAAAGTCACTCTCAGTGAGTGACTTAAGAGTTT 958
6-1R     CTGTCGGACAGTGACCCCAGAGTTGTTCAAAGTCACTCTCAGTGAGTGACTTAAGAGTTT 726
         ************************************************************

6-1F     ATTTCATGTCACTGTCTTACAAGGATTGGACCGTTTTGTCCAATTTTTGAAGGACTAGAC 1020
6-1cons ATTTCATGTCACTGTCTTACAAGGATTGGACCGTTTTGTCCAATTTTTGAAGGACTAGAC 1018
6-1R     ATTTCATGTCACTGTCTTACAAGGATTGGACCGTTTTGTCCAATTTTTGAAGGACTAGAC 786
         ************************************************************

6-1F     CATTTTGTCTCGAT------------------------------------------------ 1034
6-1cons CATTTTGTCTCGATGCCCCCCTATTTGTACTTATATGATAAAGCCCGGGCCCCAAATTAA 1078
6-1R     CATTTTGTCTCGATGCCCCCCTATTTGTACTTATATGATAAAGCCCGGGCCCCAAATTAA 846
         **************

6-1F     ------------------------------------------------------------
6-1cons AATTCAATTTAATAATCAGGGTTAACGAAGAGAAAAACAAATTTATCAAATAATGAAATC 1138
6-1R     AATTCAATTTAATAATCAGGGTTAACGAAGAGAAAAACAAATTTATCAAATAATGAAATC 906

6-1F     ------------------------------------------------------------
6-1cons TATCTATCCTTTTTTACCCCTAAAAAAAATTTTAAAAAATCTAACTATCCCTTAAGGTTT 1198
6-1R     TATCTATCCTTTTTTACCCCTAAAAAAAATTTTAAAAAATCTAACTATCCCTTAAGGTTT 966

6-1F     ------------------------------------------------------------
6-1      TGCTTTTTCTGAAAATATGTCTACATTGTTTGCATGTCTCAGTACCCTAAATTGCGATCC 1258
6-1R     TGCTTTTTCTGAAAATATGTCTACATTGTTTGCATGTCTCAGTACCCTAAATTGCGATCC 1026
```

continues..

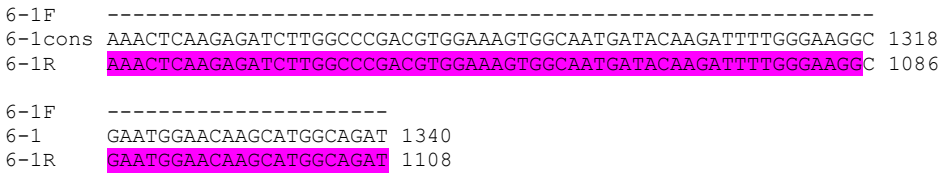

Fig 3.10. *Example of "Fiori" cl 6-1: its consensus sequences is composed by the portion of M13 F sequence (in yellow) plus the portion of M13 R sequence (in pink) plus the overlapping sequence (in green)*

In order to confirm the finding of new putative alleles in the studied accessions, the sequences were analyzed through Genewise version 2.1.20 (by Birney and Copley), available at http://www.ebi.ac.uk/Tools/psa/genewise/. This software aligns DNA sequences to protein and predicts exon-intron junctions.

3.3 Results and discussion

3.3.1 Classification of amplicons

S-locus genotyping was obtained by sequencing PCR amplicons produced with consensus primers Pa-cons1F and EM-PC3consRd. The length of amplified fragments ranged between 800 and 1400 bp (see an example of amplicon electrophoresis in fig. 3.11).

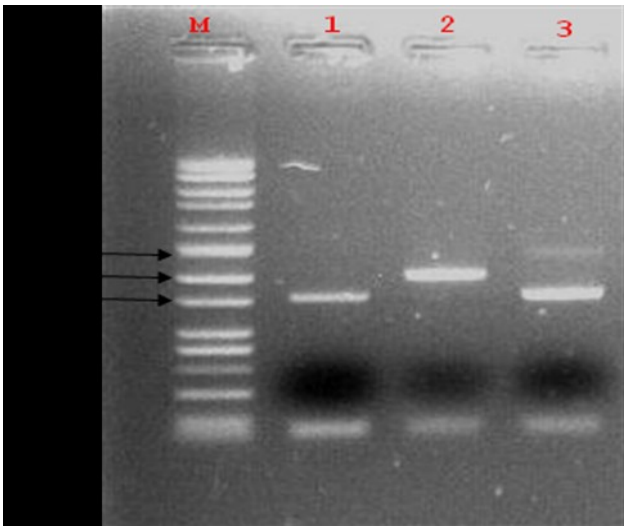

Fig.3.11. *Example of amplicon electrophoresis: lane 1=CIATTA INGLESE, lane 2= FOLLA 'E PRESSIU, lane 3= OLLA. M= DNA size marker (50-10000 bp Sigma D7058)*

This size range is consistent with data obtained for other almond varieties or other *Prunus* spp. (Sonneveld *et al*., 2003). From double strand sequencing of the cloned amplicons and sequence alignment, consensus sequences were obtained. Sequence comparison through the Blast N algorithm confirmed the homology with SI allele sequences deposited in GeneBank/EMBL/DDBJ database.

Amplicon sequences were classified as copies of known alleles based on the length of intron-exon sequences. Intron-exon junctions were identified using GENEWISE software. The details of intron and exon lengths are reported in table 3.2.

Tab. 3.2 Lengths of amplified fragments for the clones of local almond accessions. Ref. sequences are in bold

Accession	Clone	Genotipo SI	N Rif. Subject	1st exon	1st intron	2nd exon	2nd intron	3rd exon
	11-1					156	253	96
	11-2					156	253	96
IBBA	11-3					156	253	96
	21-1					156	302	51
NUXEDDA	21-9					156	253	96
	19-5					156	253	99
ANTONI PIRAS	19-8					156	253	96
		S2	**AF510416**			**156**	**253**	**381**
SCHINA DE PORCU	32-6			94	285	223	375	68
	32-7			46	276	182	363	36
	34-3			46	276	182	359	99
CIATTA INGLESE	34-7			46	277	182	359	99
	26-6			52	186	179	252	99
NIEDDA II	26-9			52	186	176		
		S6	**AM231657**	**46**	**276**	**182**	**359**	**327**
	42-6			46	302	185	373	99
PITICHEDDA	42-10			46	302	185	904	99
		Sg	**DQ156218**	**76**	**300**	**185**	**830**	**399**
DE EFISI SINZOBA	36-1			46	215	188	129	99
	36-7			46	215	184	134	99
		S18	**AM231667**	**46**	**215**	**188**	**129**	**339**

continues…

Accession	Clone	Genotipo SI	N Rif. Subject	1st exon	1st intron	2nd exon	2nd intron	3rd exon
	14-7			46	346	191	328	99
FARCI	14-10			46	346	194	387	62
		S21	**AM2131670**	**46**	**346**	**191**	**326**	**348**
RIU LOI	7-3			49	297	187	488	69
	40-6			46	297	188	489	81
FOLLA 'E PRESSIU	40-11			46	297	188	588	
		S23	**AB488496**	**76**	**297**	**188**	**489**	**411**
	6-1			46	297	192	744	62
FIORI	6-7			46	297	192	693	62
	47-6			36	275	213	710	69
FRANCISCU	47-9			69	262	213	699	95
		S24	**AM231672**	**46**	**297**	**188**	**710**	**333**
	12-1			46	241	185	497	96
IS STUMBUS	12-6			46	241	185	493	96
		S27	**AM231675**	**46**	**295**	**182**	**1176**	**336**
	20-1					174	553	403
STAMPASACCUSU	20-8					174	613	432
		S33	**EF690375**			**174**	**617**	**374**
	23-2			46	110	185	295	99
ARRUBIA	23-7			46	110	185	294	96
		S40	**HQ622703**	**46**	**110**	**185**	**296**	**327**
	10-5			52	328	182	431	93
OLLA	10-10			52	328	173	194	72
		S 41	**HQ622704**	**46**	**303**	**176**	**440**	**372**
FARRAU	39-8			46	321	188	341	99
		S2 persica	**AB597186**	**76**	**321**	**188**	**341**	**417**

3.3.2 Alignment of amino acid sequences and matrix of identity

On the whole, twenty-four almond S-incompatibility alleles were identified from the classification of the amplicons for the eighteen Sardinian almond accessions, plus the S_2 allele of *Prunus persica* found in cultivar "Farrau". This result was in agreement with Ortega *et al*. (2006), who claimed that several almond S-alleles showed identities with S-alleles of other *Prunus,* exceeding even 97%.

Among the twenty-four almond S-alleles, six coincided with reference sequences deposited in GeneBank/EMBL/DDBJ database; three could be considered S-like, as they have a big difference in second intron length, but are identical in the second exon length (S_2-like S_g-like, S_{23}-like); fifteen are putative new alleles, as they were not classified as copies of previously characterized S alleles.

To understand whether those putative new alleles encode for novel proteins we studied the alignment of deduced aminoacid sequences. Figure 3.12 shows the alignment with the respective reference sequences. The presence of the differences shown was supported by the percentage of identity between amino acid sequences in the identity matrix (see table 3.3).

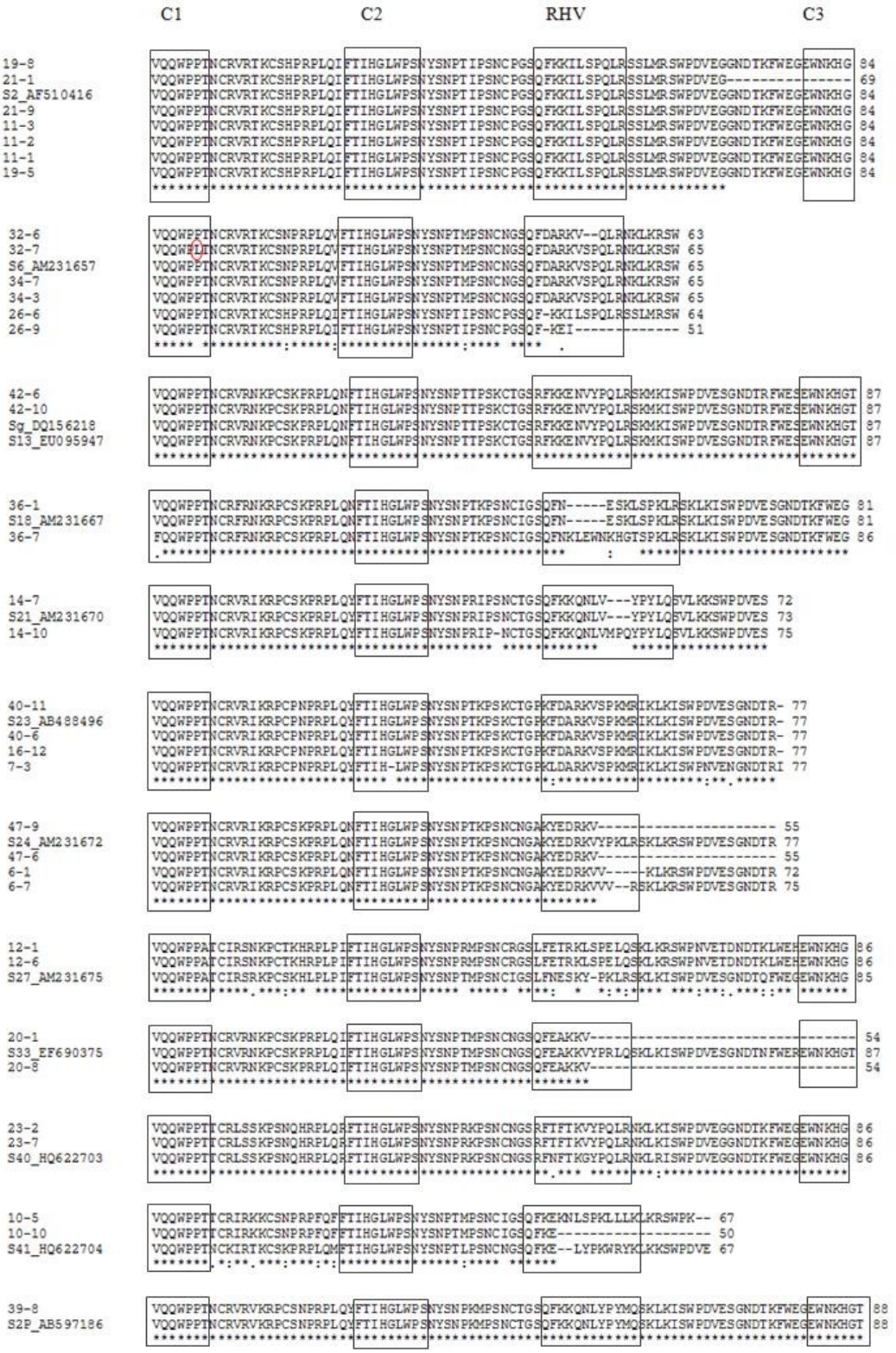

Fig. 3.12 . *Alignment of partial deduced amino acid sequences of the S alleles identified in the studied accessions. The C1, C2, RHV and partial C3 regions are boxed.*

Tab. 3.3 Identity matrix of aminoacid sequences of clones and reference alleles

Clones and ref. alleles	26-6	26-9	32-6	32-7	S_6	36-1	36-7	S_{18}	14-7	14-10	S_{21}	7-3	S_{23}	6-1	6-7	47-6	47-9	S_{24}	12-1	12-6	S_{27}	20-1	20-8	S_{33}	23-2	23-7	S_{40}	10-5	10-10	S_{41}
26-6	ID	0.781	0.687	0.661	0.676	0.172	0.139	0.172	0.205	0.2	0.205	0.197	0.207	0.208	0.213	0.234	0.234	0.194	0.093	0.093	0.105	0.203	0.203	0.16	0.127	0.127	0.127	0.611	0.64	0.641
26-9		ID	0.698	0.661	0.676	0.16	0.139	0.16	0.191	0.186	0.191	0.197	0.207	0.208	0.2	0.272	0.272	0.194	0.104	0.104	0.094	0.259	0.259	0.16	0.127	0.127	0.127	0.626	0.823	0.656
32-6			ID	0.815	0.83	0.209	0.162	0.209	0.246	0.186	0.246	0.197	0.207	0.208	0.213	0.238	0.238	0.22	0.093	0.093	0.105	0.222	0.222	0.16	0.139	0.139	0.139	0.656	0.666	0.731
32-7				ID	0.984	0.172	0.151	0.172	0.205	0.186	0.205	0.184	0.194	0.194	0.28	0.215	0.215	0.194	0.081	0.081	0.176	0.2	0.2	0.149	0.127	0.127	0.139	0.746	0.63	0.641
S_6					ID	0.185	0.162	0.185	0.219	0.2	0.219	0.197	0.207	0.208	0.293	0.23	0.23	0.207	0.093	0.093	0.188	0.215	0.215	0.16	0.139	0.139	0.151	0.761	0.646	0.656
36-1						ID	0.604	1	0.555	0.469	0.555	0.296	0.753	0.58	0.567	0.555	0.555	0.79	0.104	0.104	0.117	0.16	0.16	0.149	0.116	0.116	0.116	0.172	0.148	0.197
36-7							ID	0.604	0.523	0.43	0.523	0.302	0.476	0.523	0.5	0.5	0.5	0.534	0.127	0.127	0.081	0.151	0.151	0.195	0.139	0.139	0.127	0.151	0.127	0.151
S_{18}								ID	0.555	0.469	0.555	0.296	0.753	0.58	0.567	0.555	0.555	0.79	0.104	0.104	0.117	0.16	0.16	0.149	0.116	0.116	0.116	0.172	0.148	0.197
14-7									ID	0.586	1	0.355	0.558	0.63	0.586	0.589	0.589	0.558	0.093	0.093	0.094	0.191	0.191	0.16	0.116	0.116	0.116	0.178	0.164	0.205
14-10										ID	0.586	0.407	0.519	0.546	0.546	0.533	0.533	0.532	0.186	0.186	0.164	0.253	0.253	0.229	0.186	0.186	0.186	0.2	0.16	0.16
S_{21}											ID	0.355	0.558	0.63	0.586	0.589	0.589	0.558	0.093	0.093	0.094	0.191	0.191	0.16	0.116	0.116	0.116	0.178	0.164	0.205
7-3												ID	0.376	0.355	0.342	0.342	0.342	0.337	0.43	0.43	0.27	0.394	0.394	0.505	0.476	0.476	0.453	0.171	0.171	0.184
S_{23}													ID	0.61	0.61	0.597	0.597	0.831	0.093	0.093	0.105	0.168	0.168	0.149	0.116	0.116	0.116	0.194	0.181	0.207
6-1														ID	0.746	0.763	0.763	0.74	0.116	0.116	0.105	0.18	0.18	0.16	0.127	0.127	0.127	0.194	0.18	0.208
6-7															ID	0.733	0.733	0.727	0.093	0.093	0.294	0.173	0.173	0.149	0.127	0.127	0.139	0.266	0.173	0.213
47-6																ID	1	0.714	0.093	0.093	0.105	0.236	0.236	0.149	0.116	0.116	0.116	0.194	0.236	0.223
47-9																	ID	0.714	0.093	0.093	0.105	0.236	0.236	0.149	0.116	0.116	0.116	0.194	0.236	0.223
S_{24}																		ID	0.093	0.093	0.117	0.168	0.168	0.149	0.116	0.116	0.116	0.194	0.168	0.22
12-1																			ID	1	0.523	0.465	0.465	0.724	0.697	0.697	0.686	0.127	0.116	0.116
12-6																				ID	0.523	0.465	0.465	0.724	0.697	0.697	0.686	0.127	0.116	0.116
S_{27}																					ID	0.458	0.458	0.448	0.43	0.43	0.441	0.211	0.105	0.105
20-1																						ID	1	0.62	0.453	0.453	0.441	0.179	0.222	0.179
20-8																							ID	0.62	0.453	0.453	0.441	0.179	0.222	0.179
S_{33}																								ID	0.747	0.747	0.724	0.149	0.137	0.149
23-2																									ID	1	0.965	0.162	0.151	0.127
23-7																										ID	0.965	0.162	0.151	0.127
S_{40}																											ID	0.174	0.151	0.127
10-5																												ID	0.746	0.656
10-10																													ID	0.626
S_{41}																														ID

A detail of those amino acid sequences, not showing 100% identity with the respective reference sequence (in red on table 3.3) is represented in fig 3.13. The aligned deduced proteins present several polymorphisms, suggesting that they potentially encode for novel S-RNases.

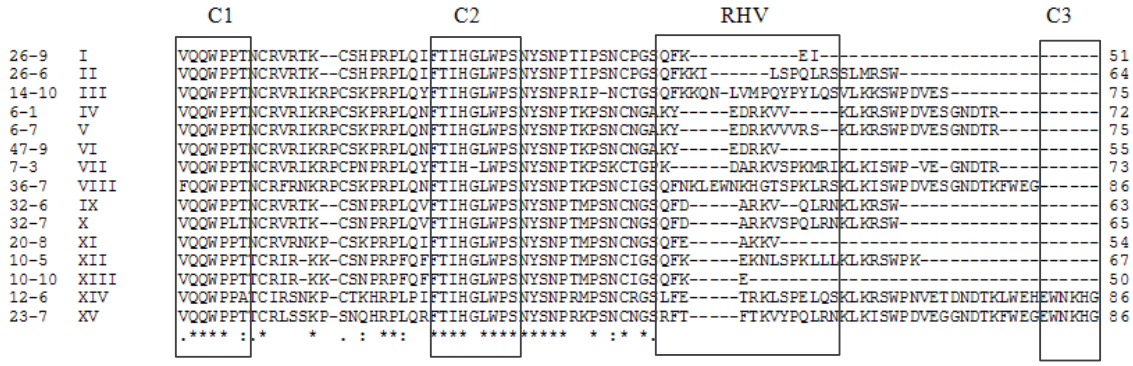

Fig. 3.13. *Multialignment of putative new S alleles of Sardinian cv.(from SS_I to SS_{XVI}). The C1, C2, RHV and partial C3 regions are boxed.*

In some cases the percentage of identity to the reference sequences was very high (e.g. 32-7 of "Schina de porcu" to S_6, with 98.4%), the only difference being a leucine instead of a proline residue in C1 region (see fig 3.12). Wünsch and Hormaza (2004) suggested that if two sequences have a similar intron length and are exceptionally similar both at nucleotide and amino-acid levels, the two alleles could have derived from a common ancestral allele or one could have evolved from the other after several mutations. However, the high similarity or identity of two alleles does not mean that they have the same functionality (Kodad *et al.,* 2010b).

3.3.3 New S-incompatibility alleles

The S-incompatibility alleles found are listed in table 3.4. The fifteen putative new alleles are indicated with the prefix SS (S=self-incompatibility, S=Sardinia).

Tab. 3.4 Putative S-genotype of the studied accessions

Accession	Clone	Genotipo SI	N Rif. Subject	alleles based on intron lengths	based on aminoacidic sequence	Putative S genotipe
	11-1			S_2	S_2	
	11-2			S_2	S_2	
IBBA	11-3			S_2	S_2	S_2/?
	21-1			S_2-like	S_2	
NUXEDDA	21-9			S_2	S_2	**$S2/S_2$-like**
	19-5			S_2	S_2	
ANTONI PIRAS	19-8			S_2	S_2	S_2/?
		S2 partial	**AF510416**			
	32-6				SS_{IX}	
SCHINA DE PORCU	32-7				SS_X	SS_{IX}/SS_{IX}
	34-3			S_6	S_6	
CIATTA INGLESE	34-7			S_6	S_6	S_6/?
	26-6				SS_I	
	26-9				S_{II}	SS_I/SS_{II}
NIEDDA II		**S6**	**AM231657**			
	42-6			S_g-like	S_g/S_{13}	$S_g S_{13}/S_{g\text{-like}}$
PITICHEDDA	42-10			S_g-like	S_g	
		Sg	**DQ156218**			
	36-1			S_{18}	S_{18}	
DE EFISI SINZOBA	36-7				SS_{VIII}	S_{18}/SS_{VIII}
		S18	**AM231667**			
	14-7			S_{21}	S_{21}	
FARCI	14-10			S_{III}	S_{III}	S_{21}/S_{III}
		S21	**AM2131670**			

continues…

Accession	Clone	Genotipo SI	N Rif. Subject	alleles based on intron lengths	based on aminoacidic sequence	Putative S genotipe
RIU LOI	7-3			S_{23}	SS_{VII}	$S_{23}SS_{VII}$/?
FOLLA 'E PRESSIU	40-6			S_{23}	S_{23}	
	40-11			S_{23}-like	S_{23}	S_{23}/S_{23}-like
		S23 cds	**AB488496**			
FIORI	6-1				SS_{IV}	
	6-7				SS_{V}	SS_{IV}/SS_{V}
FRANCISCU	47-6				SS_{VI}	SS_{VI}/?
	47-9				SS_{VI}	
		S24	**AM231672**			
IS STUMBUS	12-1				SS_{XIV}	
	12-6				SS_{XIV}	SS_{XIV}/?
		S27	**AM231675**			
STAMPASACCUSU	20-1				SS_{XI}	
	20-8				SS_{XI}	SS_{XI}/?
		S33	**EF690375**			
ARRUBIA	23-2				S_{XV}	
	23-7				S_{XV}	S_{XV}/?
		S40	**HQ622703**			
OLLA	10-5				S_{XII}	
	10-10				S_{XIII}	S_{XII}/S_{XIII}
		S 41	**HQ622704**			
FARRAU	39-8				S_2P/?	
		S2 persica	**AB597186**			S_2P/?

3.4 Conclusions

While the isolation of the full-length sequence of these alleles will allow a deeper genomic characterization, controlled crosses with known incompatibility genotypes are needed to attribute them to incompatibility groups. In absence of these data it will be difficult to define the exact entity of S-allele functional diversity. However if we take into consideration the degree of molecular diversity it is tempting to conclude that these Sardinian almond genotypes have a high S-allele richness. This finding may be explained considering that, due to the low level of introgression of new almond gene pools in Sardinia, the Sardinian almond gene pools have been subjected to an exceptional selective pressure for new S-alleles.

Moreover, we would like to recall to the attention that almond plantation has been often started from seeds, with a low selection by breeders and this may have allowed the transmission of new incompatibility alleles.

3.4 Bibliography

Ai Y., Singh A., Coleman C.E., Ioerger T.R., Kheyr-Pour A. and Kao T.-H.(1990). Self-incompatibility in *Petunia inflate*: isolation and characterization of cDNAs enconding three S-allele-associated proteins. Sex. Plant Reprod. 3: 130-138.

Birney E. and Copley R. (1999). Wise2 documentation (version 2.1.20 stable).http://www.sanger.ac.uk/Software/Wise2. Available on web at http://www.ebi.ac.uk/Tools/psa/genewise/

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

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

Clark A.G. and Kao T.H. (1991). Excess nonsynonymous substitution at shared polymorphic sites among self-incompatibility alleles of Solanaceae. Proc. Natl. Acad. Sci. USA 88: 9826–9827.

Cornish E.C., Pettitt J.M., Bonig I. and Clarke A.E. (1987). Developmentally controlled expression of a gene associatedwith self-incompatibility in *Nicotiana alata*. Nature 326: 99–102.

Crossa-Raynaud P. and Grasselly C.(1985). Existence de groups d'interstèrilitè chez l'amandier. Options Mèditerranèennes. Serie Etudes 1: 43-45.

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

Gagne J.M., Downes B.P., Shiu S.H., Durski A. and Vierstra R.D. (2002). The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. Proc. Natl. Acad. Sci. USA 99:11519–11524.

Huang S., Lee H.-S., Karunanandaa B. And Kao T.-H. (1994). Ribonuclease activity of *Petunia inflate* S proteins is essential for rejection of self-pollen. Plant Cell 6: 1021–1028.

Halasz J., Agota Fodor A., Hegedus A. and Pedryc A. (2008). Identification of a new self-incompatibility allele (S31) in a Hungarian almond cultivar and its reliable detection. Scientia Horticulturae 116 (2008) 448–451.

Igic B., and Kohn J.R. (2001). Evolutionary relationships among self-incompatibility RNases. Proc. Natl. Acad. Sci. 98: 13167—13171.

Iwano M. and Takayama S. (2012). Self/non-self discrimination in angiosperm self-incompatibility. Current Opinion in Plant Biology 15: 78–83.

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

Kaufmann H., Salamini F. and Thompson R.D. (1991). Sequence variability and gene structure at the self-incompatibility locus of *Solanum tuberosum*. Mol. Gen. Genet. 226: 457-466.

Kawata Y., Sakiyama F. and Tamaoki H. (1988). Amino acid sequence of ribonuclease T2 from *Aspergillus oryzae*. Eur. J. Biochem. 176: 683–697.

Kester D.E., Gradziel T.M. and Micke W.C. (1994). Identifying pollen incompatibility groups in California almond cultivars. J. Am .Soc. Hort. Sci. 119: 106-109.

Kodad O., Alonso J.M.., Fernandez i Martì A., Oliveira M.M. and Socias i Company R. (2010a). Molecular and physiological identification of new *S* alleles associated with self-incompatibility in local Spanish almond cultivars. Scientia Hort. 123: 308-311.

Kodad O., Sanchez A., Saibo N., Oliveira M.M. and Socias i Company R. (2010b). Molecular characterization of five new *S* alleles associated with self-incompatibility in local Spanish almond cultivars. Options Méditerranéennes – XIV GREMPA Meeting on Pistachios and Almonds 94: 105-109.

Kubo K., Entani T., Takara A., Wang N., Fields A.M., Hua Z., Toyoda M., Kawashima S., Ando T., Isogai A. *et al.* (2010). Collaborative non-self recognition system in S-RNase-based self-incompatibility. Science 330: 796-799.

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

Lopez M., Vargas F.J. and Batlle I. (2006). Self-(in)compatibility almond genotypes: a review. Euphytica 150: 1–16.

Ma R.-C. and Oliveira M.M. (2001). Molecular cloning of the self-incompatibility genes S1 and S3 from almond (*Prunus dulcis* cv. Ferragnès). Sex. Plant Reprod. 14: 163–167.

McClure B.A., Haring V., Ebert P.R., Anderson M.A., Simpson R.J., Cornish E.C. *et al*. (1989). Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature 342: 955–957.

McClure B.A., Gray J.E. and Anderson M.A., Clarke A.E. (1990). Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature 347: 757–760.

Meng X., Sun P. and Kao T.-H. (2010). S-RNase-based self-incompatibility in *Petunia inflata*. Ann Bot doi: 10.1093/aob/mcq253.

Okada K., Tonaka N., Taguchi T., Ichikawa T., Sawamura Y., Nakanishi T. and Takasaki-Yasuda T. (2011). Related polymorphic F-box protein genes between haplotypes clustering in the BAC contig sequences around the S-Rnase of Japanese pear. J. Exp Bot 62: 1887-1902.

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

Ortega E., Bošković R., Sargent D.J. and Tobutt K.R. (2006). Analysis of *S*-RNase alleles of almond (*Prunus dulcis*): Characterization of new sequences, resolution of synonyms and evidence of intragenic recommendation. Mol. Gen. Genet. 276: 413-426.

Sassa H., Nishio T., Kowyama Y., Hirano H., Koba T. and Ikehashi H. (1993). Self-incompatibility (*S*) alleles of the Rosaceae encode members of a distinct class of the T2/S ribonuclease superfamily. Mol. Gen. Genet. 250: 547–557.

Silva N.F. and Goring D.R. (2001). Mechanisms of self-incompatibility in flowering plants. Cell. Mol. Life Sci. 58: 1988–2007.

Singh A., Ai Y. and Kao T.-H. (1991). Characterization of ribonuclease activity of three S-allele-associated proteins of *Petunia inflata*. Plant Physiol. 96: 61–68.

Sonneveld T., Tobutt K.R. and Robbins T.P. (2003). Allele-specific PCR detection of sweet cherry self-incompatibility (S) alleles S1 to S16 using consensus and allele-specific primers. Theor. Appl. Genet. 107:1059–1070.

Sutherland B.G., Robbins T.P. and Tobutt K.R. (2004). Primers amplifying a range of *Prunus* S-alleles. Plant Breeding 123, 582—584.

Takayama S. and Isogai A. (2005). Self-incompatibility in plants. Annu. Rev. Plant Biol. 56: 467-489.

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

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

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). Nucleic Acids Res. 22, 4673–4680) available on web at (www.ebi.ac.uk/Tools/msa/clustalw2/).

Tsai D.-S., Lee H.-S., Post L. C., Kreiling K.M. and Kao T.-H. (1992). Sequence of an S-protein of *Lycopersicon peruvianum* and comparison with other Solanaceous S-proteins. Sex. Plant Reprod. 5: 256–263.

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

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

Wheeler M.J., de Graaf B.H., Hadjiosif N., Perry R.M., Poulter N.S., Osman K., Vatovec S., Harper A., Franklin F.C., Franklin-Tong V.E. (2009). Identification of the pollen self-incompatibility determinant in Papaver rhoeas. Nature 459:992-995.

Wünsch A. and Hormaza J.I. (2004). S-allele identification by PCR analysis in sweet cherry cultivars. In: Plant Breed., 123, p. 327-331.

Xue Y., Carpenter R., Dickinson H. G. and Coen E. S. (1996). Origin of allelic diversity in *Antirrhinum* S locus RNases. Plant Cell 8: 805–814.

Zhou J., Wang F., Ma W., Zhang Y., Han B. and Xue Y. (2003). Structural and transcriptional analysis of S-locus F-box genes in Antirrhinum. Sex. Plant Reprod. 16:165-177.

CHAPTER 4

MAIN FATTY ACIDS AND α-TOCOPHEROL CONCENTRATIONS IN OIL OF SARDINIAN ALMOND ACCESSIONS

4.1 Introduction

4.1.1 Quality characterization of almond production

In almond (*Amygdalus communis* L., syn *Prunus amygdalus* Batsch., syn *Prunus dulcis* Miller), the kernel is the edible part of the fruit. Almond quality comprises commercial, industrial, organoleptic and nutritional aspects related to the nut and kernel.

Commercial quality includes characteristics regarding the external presentation, like size, texture, absence of double kernels etc., as well as marketable yield. Industrial quality refers to the cultivar's attitude to handling, transportation, processing and storage. Organoleptic quality is highly variable and subjective as it consists of those parameters related to consumer preferences. Nutritional quality refers to the specific nutrients provided and the contribution to consumer health on the whole.

Until recently, as far as fruit is concerned, almond breeding has been focused on selecting fruits of a high physical quality (mainly related with commercial quality). For this reasons, information about the chemical composition of the almond oil kernel and their variability are scarce. Including such analyses in the evaluation of almond varieties would be of great interest to explore the possible utilizations of the product. Almond have been used in different ways: they are consumed raw, roasted, peeled or unpeeled; processed into food items, such as marzipan, nougat (*"torrone"*) and other traditional sweets, typical in the Mediterranean basin, and into almond milk. Additionally almond is used in the pharmaceutical and cosmetics industries. Thus the same kernel trait may be considered positive or negative depending on its final utilization.

4.1.2 Health and beauty benefits of almond

Almond oil has been used for its numerous health and beauty benefits since the ancient civilizations of India, China and Greece (Zohary and Hopf, 2000; Puri, 2003) and many of those uses have nowadays confirmed their effectiveness: almond oil is used today by beauticians, aromatherapists and massage therapists for its many natural emollient and skin-rejuvenating properties (Kuriyama *et al.*, 2005). Moreover, due to their nutritional and food quality, almonds become over the time an important component of many diets.

The high nutritive value of almond kernels arises mainly from their high lipid content (more than 50% of the almond kernel dry weight), on which our study concentrated. Almond lipids (oil) constitute an important caloric source but does not contribute to cholesterol formation in humans: this is due to its high level of unsaturated fatty acids (UFAs), mainly monounsaturated fatty acids, (MUFAs), that are associated with reduced incidence of cardiovascular diseases (Hyson *et al*., 2002; Jenkins *et al*. 2002). Moreover, according to preliminary results (Ahmad, 2010) almond oil has many properties, including anti-inflammatory, immunity-boosting and anti-hepatotoxicity effects (Hyson *et al.*, 2002). Association has also been made between almond oil consumption and reduction in the incidence of colonic cancer (Davis and Iwahashi, 2001).

4.1.3 Composition of almond oil

Almond oil includes five major fatty acids: three UFAs (oleic, monounsaturated, palmitoleic and linoleic, polyunsaturated), that account for the 90% of the total lipids; and two SFA (palmitic and stearic).

Almond oil also contains eight minor fatty acids (Martin-Carratalà *et al.*, 1998). Among them it is worth mentioning the polyunsaturated α-linolenic acid (= ALA) that, like linoleic acid, is an essential fatty acids: they are not synthesized by the human organism, so they must be taken in through the diet. They are the starting point, respectively, of omega-3 and omega-6 fatty acid families. Among the several functions of these compounds, the omega-3s have anti-inflammatory and anti-thrombotic effects and the omega-6s reduce the blood cholesterol concentration (Djousse *et al.*, 2005). As nowadays' diet is poor in fish, the primary source of omega-3 fatty acid family, their intake is scarce if compared to omega-6 fatty acid family. Thus some alternative, vegetable, sources of omega-3s are precious.

Finally, the unsaponifiable fraction of almond oil contains sterols, aliphatic alcohols hydrocarbons and liposoluble vitamins (Schirra *et al.*, 1997).

Besides the nutritional value, lipid content and composition is also important for determination of oil stability, since component fatty acids differ in their vulnerability to oxidation (Senesi *et al.*, 1996), thus influencing the resistance to rancidity. As polyunsaturated fatty acids (PUFAs) are more susceptible to oxidation than MUFAs, Kester *et al.* (1993) suggested that oleic acid/linoleic acid ratio is a good index of resistance to oil rancidity, with higher ratios indicating better resistance.

Resistance to oil rancidity also depends on the presence of natural antioxidant, such as tocopherols, that protect PUFAs against peroxidation. In almond the tocopherols lengthen the storage time by playing an important role against fat oxidation. Several studies have shown that almond nuts can be stored for nine months with little quality compromission (Zacheo *et al.*, 2000); storage is possible up to one year for cultivars with high concentration of natural antioxidants such as α-tocopherol (more than 400 mg/ kg oil, Rizzolo *et al*., 1994).

Tocopherols are also important for human health: α-tocopherol is the form of vitamin E that is most efficiently used by the human body, yet it is often deficient in modern diets (Krings and Berger, 2001). Vitamin E along with the antioxidant polyphenols and fibers may help to prevent heart diseases and cancer (Jenkins *et al.*, 2002; Davis and Iwasashi, 2001). Almond is the nut with the highest α-tocopherol content. For this reason almonds was included in the recommendations of The Dietary Guidelines for Americans (USDA, 2005) in the context of enhancing the intake of this vitamin (Lopez-Ortis *et al*., 2008).

4.1.4 Survey on kernel composition of Sardinian almond collection

The collection characterized in this study is a good representation of the almond germplasm that constitute most of the small-scale production in Sardinia. Those varieties, in general, are not competitive from the productive point of view. Their low production is caused by self-incompatibility and/or early-blooming, and by the application of traditional agronomic techniques. The identification of some cultivars endowed with particular qualitative characteristics could stimulate their cultivation for purposes other than massive production (e.g.: nutritional and cosmetic/pharmaceutical uses, characterization of typical food productions etc.).

The objective of this study was to determine the oil content and its principal components for the Sardinian almond collection, possibly in order to establish the chemical profiles helpful for cultivar characterization and to find a selection criterion for their quality evaluation.

4.2 Materials and methods

4.2.1 Plant material and analyzed parameters

Forty-three almond accessions were analyzed in this study. They came from the AGRIS collection field located in southern Sardinia (see tab. 2.1 chapter 2). It included forty sweet almond Sardinian accessions, and three commercial varieties (Nonpareil and Ne Plus Ultra from California; Picantili, from Ukraine).

4.2.2 Sample preparation and determination of kernel humidity

Data were collected on 2011. The following parameters were analyzed: oil content, as % on kernel dry weight (DW); percentages on total lipids of the major fatty acids (oleic, linoleic, palmitic, stearic, palmitoleic) plus ALA; and α–tocopherol (T) content (mg/kg oil). From the fatty acid percentages, the ratios oleic acid % / linoleic acid % (O/L) and UFAs % / SFAs % the were determined.

Mature nuts were randomly harvested when the hull was fully dried and opened along the suture. After cracking, seed (kernel) teguments were removed by pouring in warm water for about a min. The peeled kernel were wiped and stored in plastic bags at -20°C until further use.

For chemical analyses, three replicates of 15 kernel/each per genotype were defrost, dried at room temperature over night, then ground in a coffee mill. Flour from each replicate was sieved through a 18 mesh (1 mm)-diameter sieve. Then it was split into two aliquots of about 5 g each, to be used respectively for oil extraction and for humidity determination, since the two processes are technically incompatible.

Kernel humidity was calculated, for each replicate, on a 5 g aliquot of the kernel flour obtained, which was by drying the sample in oven at 105 °C over night.

4.2.3 Oil extraction

Oil was extracted by treating the samples with 80-100 ml of petroleum ether (30-50°C) in a Soxlet apparatus for 4 h; then petroleum extracts containing lipids were recovered in a flask and subjected to vacuum distillation in a rotovapor at 40°C. The residual petroleum ether was eliminated by blowing N_2 in the flask, and lipid weight was determined. Total lipid

percentage was calculated on dry weight (DW), obtained after subtraction of the humidity calculated on the first flour aliquot. Oil extracts were transferred in vials, the empty space filled with N_2 to avoid exposition to oxygen, and stored at -20°C until further analysis were performed (the analysis of oil composition was completed within five months from the peeled kernel storage).

4.2.4 Fatty acids extraction

The following step was performed to transform the oil fatty acids to the corresponding methyl esters (FAMEs): 0.5 g aliquots of oil, were dissolved in 6 ml hexane, and 250 µL of KOH 2N in methanol was added. After moderate shaking, tubes were centrifuged at 2000 x g for 10 min. Surnatants were recovered and transferred in vials for gas chromatography (GC) analysis.

GC analysis was carried out through a Clarus GC 680 gas chromatograph, equipped with a Supelco 2380 column (60 m x 0.25 mm i. d., 0.2 µm film thickness) and a flame ionization detector (FID). Injection volume was 0.5 µL and helium was used as carrier gas at a flow rate of 0.37 mL min^{-1}. Both temperatures of injector and detector were 220°C. The initial column temperature was set at 185°C for 25 min. The oven temperature was then increased up to 200°C at 10°C/min ramp rate, and maintained at 200°C for 10 min, then increased up to 220°C at 10°C/min ramp rate, and maintained at 220°C for 20 min, for a total run time of 58 min 30 sec. The identification of the FAMEs was achieved by comparison with relative retention times in the reference samples of each methyl ester standard (from Sigma-Aldrich). Results were recorded and processed by Total Chrome Work Station.

4.2.5 α-tocopherol extraction

Samples for T extraction were prepared by dissolving 0.1 g of oil in 1.9 mL acetone (about 2 ml of total sample), shaking and filtering through a 0.22 µm syringe cellulose filter. T determination was performed using a Waters HPLC, equipped with a Waters 600 Controller pump unit and a 717 plus auto sampler.

The chromatography column, kept at 25°C, was a Spherisorb ODS2 250 x 4.6 x 5 µm, with a Phenomenex pre-column (cartridge C18 AJO-4287). The mobile phase was a 1:1 mixture of acetonitrile and methanol, at a flow rate of 1 mL/min. Detection of T was carried out using a

Waters 996 photodiode array detector (PDA) at 295 nm wave length, within a run time of 18 min. Results were recorded and processed by Enpower 2 Work Station.

T concentration was quantified in mg/l basing on a calibration curve with α-tocopherol (from Sigma-Aldrich) as external standard. From the oil weight in the 2 ml sample (see above), T value was then expressed as mg/kg oil.

4.2.6 Statistical analyses

Statistical analyses on oil content and composition data were carried out by *JMP ver.7* software (SAS Institute Inc. 2007) as follows:

- a one-way ANOVA with Tukey test to detect significant differences among accessions;
- a multivariate cluster analysis (according to Ward method), followed by a MANOVA analysis that identified the most probable number of clusters (groups of accessions more similar for overall oil quality performances).

4.3 Results and discussions

4.3.1 ANOVA for oil parameters

Significant differences of lipid content and its composition was detected in Sardinian and commercial accessions (one-way ANOVA $p < 10^{-4}$, see tab. 4.1). The number of classes obtained for each parameter was quite high, indicating large and continuous variations of each parameter among accessions (tab. 4.2).

Tab. 4.1 One-way ANOVA, x = accession [(a)]

	Sardinian and Commercial				*Sardinian*			
Oil content and composition [(a)]	**DF**	**SS**	**F**	**P**	**DF**	**SS**	**F**	**P**
Oil content (% of kernel dw)	42	0.14639374	19.5695	<.0001	39	0.13109632	18.176	<.0001
Oleic acid (%)	42	0.28270084	34.4662	<.0001	39	0.24368136	32.8688	<.0001
Linoleic acid (%)	42	0.11193158	34.861	<.0001	39	0.09095549	31.9972	<.0001
Palmitoleic acid (%)	42	0.00010245	10.4415	<.0001	39	0.00009636	10.0183	<.0001
α-linolenic acid (%)	42	1.66E-06	28.4085	<.0001	39	1.44E-06	25.5739	<.0001
Palmitic acid (%)	42	0.0024684	20.8186	<.0001	39	0.00237786	21.7759	<.0001
Stearic acid (%)	42	0.00083233	16.2295	<.0001	39	0.0008091	19.1144	<.0001
UFAs % / SFAs %	42	70.1596	18.8285	<.0001	39	67.69533	19.6588	<.0001
O/L	42	84.556802	36.5681	<.0001	39	74.934397	33.7917	<.0001
α-Tocopherol (mg/kg of oil)	42	484320.68	18.4401	<.0001	39	455469.23	18.3403	<.0001

[(a)] parameter expressed in % were transformed in arcsen for statistical elaborations

Tab. 4.2 Tukey mean separations for oil parameters: Sardinian and commercial accessions

ACCESSION/CULTIVAR	Lipid content (% of kernel dry weight)	Oleic acid (%)	Linoleic acid (%)	Palmitoleic acid (%)	Linolenic acid (%)	Palmitic acid (%)	Stearic acid (%)	% SFAs/ %UFAs	% oleic acid/ %linoleic acid	Tocopherol (mg/kg of oil)
Antioco Pala	55.9 M-O	77.8 A	14.25 LM	0.437 D-K	0.043 B-G	5.69 L-O	1.62 H-K	12.7 A-C	5.5 AB	357.9 N-P
Antoni Piras	58.3 G-O	70.7 G-L	19.64 D-H	0.537 A-H	0.020 M-O	6.57 C-G	2.27 A-C	10.3 LM	3.6 F-I	393.1 H-O
Arrubia	56.3 K-O	71.6 E-K	19.21 E-I	0.603 A-C	0.039 E-I	6.62 B-F	1.74 D-K	10.9 G-M	3.7 F-I	376.4 L-P
Basibi	59.3 D-M	72.1 E-J	19.18 E-I	0.197 L	0.020 M-O	6.01 G-N	2.03 B-F	11.4 E-J	3.8 F-I	417.1 G-N
Bianca	64.3 AB	73.0 D-H	18.52 F-J	0.416 E-K	0.033 G-L	5.85 K-O	2.05 B-D	11.6 D-I	3.9 E-H	493.9 B-G
Bocchino	58.5 F-O	70.1 H-L	20.65 C-G	0.556 A-F	0.024 L-N	6.91 BC	1.65 G-K	10.7 I-M	3.4 G-J	383.3 K-P
Ciatta inglese	60.1 D-J	74.3 B-E	17.10 H-L	0.694 A	0.040 E-I	6.22 D-L	1.49 J-L	12.0 A-F	4.4 D-F	452.7 D-L
Ciatta malissa	57.7 H-O	70.9 G-L	19.98 C-G	0.598 A-C	0.055 AB	6.54 C-H	1.73 D-K	11.1 F-M	3.5 G-I	431.5 F-N
Corrochina	59.0 E-N	69.8 H-L	21.28 C-F	0.520 B-J	0.032 G-M	6.54 C-H	1.68 E-K	11.2 F-L	3.3 G-K	547.7 AB
Cossu	59.0 E-N	71.1 F-L	19.41 E-H	0.462 C-K	0.036 F-L	6.29 D-K	2.57 A	10.3 LM	3.7 F-I	433.2 E-N
De Efisi Sinzoba	62.4 A-E	66.2 MN	24.36 AB	0.525 B-H	0.046 B-F	6.73 B-E	2.00 B-G	10.4 J-M	2.7 JK	390.1 I-O
De Mrasciai	57.4 I-O	78.0 A	13.97 M	0.360 JK	0.024 K-N	6.01 H-O	1.43 KL	12.4 A-D	5.6 A	387.9 J-P
Efisi Sinzoba	57.9 H-O	75.5 A-D	15.99 J-M	0.523 B-I	0.010 O	6.06 G-M	1.78 D-K	11.7 C-H	4.7 B-E	468.8 B-J
Emilio 91	57.3 I-O	68.0 L-N	22.37 B-D	0.593 A-D	0.020 M-O	6.87 BC	1.92 C-I	10.4 K-M	3.0 I-K	322.8 OP
Farci	58.2 G-O	70.6 G-L	20.27 C-G	0.530 B-H	0.029 I-N	7.16 AB	1.25 L	10.9 H-M	3.5 G-J	623.6 A
Farrau	55.6 N-P	70.4 H-L	19.89 C-H	0.545 A-G	0.033 G-L	7.47 A	1.50 J-L	10.1 M	3.5 G-I	474.1 B-H
Fiori	60.2 C-J	68.5 K-N	22.35 B-D	0.528 B-H	0.034 F-L	6.75 B-D	1.66 F-K	10.9 H-M	3.1 I-K	493.3 B-G
Folla 'e pressiu	56.9 I-O	71.7 E-K	19.42 E-H	0.457 C-K	0.023 L-N	6.27 D-K	1.90 D-I	11.2 E-L	3.7 F-I	365.8 M-P
Franciscu	61.1 B-H	72.2 E-I	19.33 E-H	0.463 C-K	0.020 M-O	6.16 F-M	1.65 G-K	11.8 B-H	3.7 F-I	455.5 C-L
Ghironi	63.9 AB	74.2 B-F	17.11 H-L	0.466 C-K	0.035 F-L	5.71 L-O	2.36 AB	11.4 E-J	4.3 D-F	507.2 B-F
Ibba	65.1 A	69.6 I-M	21.60 B-E	0.420 E-K	0.020 M-O	6.21 D-L	2.00 B-G	11.2 F-L	3.2 H-K	420.1 G-N

continues…

ACCESSION/CULTIVAR	Lipid content (% of kernel dry weight)	Oleic acid (%)	Linoleic acid (%)	Palmitoleic acid (%)	Linolenic acid (%)	Palmitic acid (%)	Stearic acid (%)	% SFAs/ %UFAs	% oleic acid/ %linoleic acid	Tocopherol (mg/kg of oil)
Is Stumbus	59.8 D-K	70.9 G-L	20.40 C-G	0.487 B-J	0.034 F-L	6.40 C-J	1.71 D-K	11.3 E-K	3.5 G-I	453.7 C-L
Lutzeddu	61.2 B-H	73.7 C-G	18.36 G-H	0.455 C-K	0.024 L-N	5.65 M-O	1.67 F-K	12.7 A-C	4.0 E-G	421.5 G-N
Malissa tunda	51.9 P	70.2 H-L	21.89 B-E	0.318 KL	0.020 M-O	6.02 G-N	1.43 KL	12.4 A-D	3.2 H-K	475.5 B-H
Niedda I	56.1 L-O	70.2 H-L	21.50 C-E	0.381 H-K	0.050 A-E	5.99 I-O	1.74 D-K	11.9 A-G	3.3 G-K	515.3 B-E
Niedda II	55.6 N-P	76.1 A-C	15.59 K-M	0.532 B-H	0.042 D-I	6.02 G-N	1.56 I-K	12.2 A-E	4.9 A-D	434.2 E-N
Nuxedda	56.8 J-O	76.4 A-C	14.99 LM	0.496 B-J	0.060 A	6.19 E-M	1.74 D-K	11.6 D-I	5.1 A-D	401.0 H-O
Olla	60.4 C-I	75.8 A-D	15.72 J-M	0.520 B-J	0.020 M-O	5.74 K-O	1.99 B-G	11.9 A-G	4.8 A-D	469.1 B-J
Orri	59.6 D-L	69.7 I-L	21.20 C-G	0.398 F-K	0.037 F-K	6.52 C-H	1.99 B-G	10.7 I-M	3.3 G-K	436.8 D-N
Pitichedda	58.9 E-O	77.7 A	14.05 M	0.398 F-K	0.035 F-L	5.90 J-O	1.78 D-K	12.0 A-F	5.5 A	305.7 P
Provvista	62.0 A-F	75.7 A-D	16.43 I-M	0.458 C-K	0.042 D-I	5.45 O	1.81 D-J	12.7 AB	4.6 C-E	536.8 BC
Riu Loi	61.6 B-G	71.8 E-K	19.67 D-H	0.424 E-K	0.026 J-M	6.04 G-N	1.92 C-I	11.6 D-I	3.7 F-I	395.4 H-O
Schina de porcu	59.5 D-L	70.2 H-L	20.82 C-G	0.627 AB	0.030 H-M	6.36 C-J	1.78 D-K	11.3 E-L	3.4 G-J	402.7 H-O
Stampasaccusu	56.3 K-O	69.4 I-M	21.06 C-G	0.463 C-K	0.038 E-J	6.77 B-D	2.07 B-D	10.3 LM	3.3 G-K	406.5 H-N
Sunda G.	60.0 D-K	69.9 H-L	20.78 C-G	0.570 A-E	0.042 C-H	6.52 C-H	1.98 C-H	10.7 I-M	3.4 G-J	496.2 B-G
Sunda N.	57.7 H-O	69.9 H-L	21.30 C-F	0.512 B-J	0.053 A-D	6.05 G-N	2.04 B-E	11.3 E-K	3.3 G-K	493.8 B-G
Troito A	63.6 A-C	77.5 A	14.44 LM	0.394 G-K	0.040 E-I	5.78 K-O	1.65 G-K	12.4 A-D	5.4 A-C	416.1 G-N
Troito B	58.5 F-O	69.6 I-L	21.18 C-G	0.514 B-J	0.043 B-G	6.42 C-I	2.04 B-E	10.8 H-M	3.3 G-K	471.2 B-I
Vargiu	59.5 D-M	76.9 AB	15.31 LM	0.420 E-K	0.036 F-K	5.49 NO	1.71 D-K	12.9 A	5.1 A-D	378.4 K-P
Vavani Perra	58.0 H-O	75.7 A-D	15.60 K-M	0.541 A-H	0.037 F-J	6.42 C-I	1.59 I-K	11.5 D-I	4.9 A-D	436.1 E-N
Ne plus ultra	62.7 A-D	65.4 N	25.59 A	0.402 F-K	0.017 NO	6.66 B-F	1.78 D-K	10.8 H-M	2.6 K	493.1 B-G
Nonpareil	61.2 B-H	69.3 I-M	22.38 B-D	0.465 C-K	0.029 I-N	6.11 F-M	1.55 I-K	12.0 A-F	3.1 I-K	519.3 B-D
Picantili	55.2 OP	68.9 I-M	22.72 BC	0.363 I-K	0.054 A-C	5.91 I-O	1.92 C-I	11.8 C-H	3.0 I-K	461.5 C-K
Mean ± std.dev	**59.1** ± 2.74	**72.0** ± 3.25	**19.23** ± 2.93	**0.478** ± 0.09	**0.034** 0.01	**6.26** ± 0.44	**1.81** ± 0.26	**11.4** ± 0.75	**3.87** ± 0.82	**442.2** ± 62.14

(a) Levels not connected by same letters are significantly different

Lipid content percentages on kernel dry weight ranged from 65.1% ("Ibba") to 51.9% ("Malissa Tunda"). "De Efisi Sinzoba", which is comparable to the commercial varieties for productive performances (see chapter 2), reached 62.3%. Oil content values were comparable with data reported for European genotypes (54.8% to 61.7%, Garcia-Lopez *et al.*, 1996; 50.7% to 67.5%, Kodad and Socias i Company, 2008; 57.9% to 65.4%, Kodad and Socias i Company, 2009). As to reference cvs., "Picantili" exhibited 55.2% (the last but one percentage), "Nonpareil" and "Ne Plus Ultra" 61.8% and 62.7% respectively: those values differed from what was reported about Californian cultivars (48% and 50% respectively for "Nonpareil" and "Ne Plus Ultra" in California, Abdallah *et al.*, 1998; 53.1% for "Nonpareil" in Spain, Garcia-Lopez *et al.*, 1996), but were comparable to the highest oil contents of Sardinian cvs: that could suggest that also location, and thus climatic conditions, may influence oil content.

Oleic acid ranged from 78% of total lipids for "De Mrasciai" to 66.2% for "De Efisi Sinzoba", the latter being comparable to Californian "Ne Plus Ultra" (65.4%, the lowest value in the collection); also "Nonpareil" performed one of the lowest percentages of oleic acid (69.3%): oleic acid percentages of those reference cultivars are consistent with literature data (e.g.: respectively 65.5% and 68.1% reported by Abdallah *et al.*, 1998). Eleven Sardinian accessions, out of forty, exhibited more than 75% of oleic acid over total lipids, and nineteen more performed higher than 70%. Those high percentages are comparable with contents of some Spanish cvs. (Kodad and Socias i Company, 2008) and Turkish selections (Askin *et al.*, 2007).

Linoleic acid exhibited a strong inverse correlation ($r^2 \approx 0.98$) with oleic acid content (fig. 4.1), as repeatedly reported in bibliography (Abdallah *et al., 1998;* Askin *et al.*, 2007; Kodad and Socias i Company, 2008).

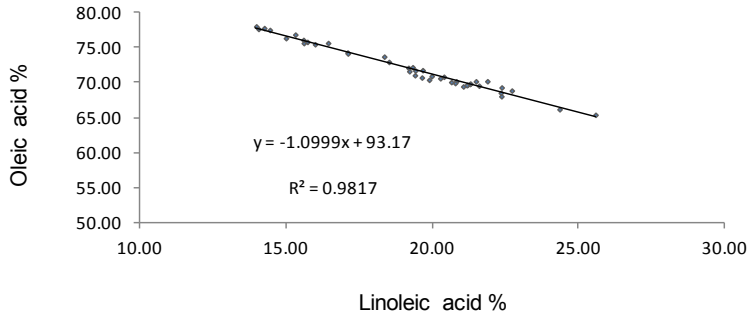

Fig. 4.1. *Correlation between oleic acid and linoleic acid*

That is why oleic acid appears to be controlled by its conversion to linoleic acid (Abdallah *et al.,1998*). Thus the accessions, over the whole collection, ranked almost in the opposite direction as for oleic acid: for example, the lowest oleic percentages ("Ne Plus Ultra" and "De Efisi Sinzoba") corresponded the highest linoleic ones (25.29% and 24.36% respectively); the accession with the highest oleic contents ("De Mrasciai") showed the lowest linoleic percentage (14%) (see tab. 4.2).

The other main fatty acids, the values ranged: from 0.69 % ("Ciatta Malissa") to 0.2% ("Basibi") for palmitoleic acid; from 7.47% ("Farrau") to 5.45% ("Provvista") for palmitic acid; from 2.57% ("Cossu"), to 1.25% ("Farci") for stearic acid. All those contents were comparable with data reported for European genotypes (Garcia-Lopez *et al*., 1996; Kodad and Socias i Company, 2008; Kodad and Socias i Company, 2009) and Californian varieties (Abdallah *et al.*, 1998; Sathe *et al*., 2008).

The α-linolenic acid, it ranged from 0.06% ("Nuxedda") to 0.01% ("Efisi Sinzoba"), with an average of 0.034%. Those percentages differ from data found in bibliography: for eight Californian cultivars Sathe *et al.* (2008) recorded higher mean values (0.07%), in particular, for "Nonpareil" it is in the order of 0.07% average content, among different region of California, vs. 0.029% in "Nonpareil" of our collection; Carratalà *et al.* (1998), reported lower values in a set of varieties cultivated in Spain (seven Spanish, three Italian, one French, one Australian, four American, one Tunisian, one Caucasic), ranging from 0.02% of the Italian "Cristomorto" to 0.006% of Spanish "Del Cid", with the American "Nonpareil" at 0.017%

Trend of oleic/linoleic ratio (the higher is the ratio, the longer is the storage time) reflected the oleic acid trend, but with fewer classes, varying from 5.59 in "De Mrasciai" to 2.56 in "Ne Plus Ultra".

UFA/SFA ratio (higher ratios indicating better dietary characteristics) ranged from 12.9 in "Vargiu" to 10.1 of "Farrau"; it was in part a consequence of the oleic + linoleic contents (strong correlation $r^2 \approx 0.99$), but it was also influenced by the percentages of palmitic acid (inverse correlation, with $r^2 \approx 0.71$):

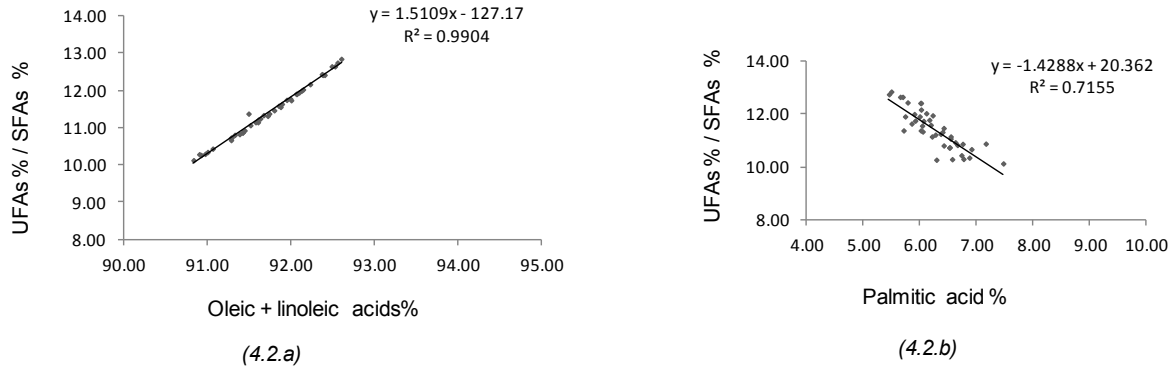

Fig. 4.2 *Correlation between UFAs % / SFAs % and O+L acids (4.2 a) and palmitic acid (4.2 b)*

Content of α–tocopherol ranged from 623.6 mg/kg of oil in "Farci" to 305.7 mg/kg in "Pitichedda", with an average over the whole collection of 442.2 mg/kg. Besides "Farci", four more Sardinian accessions plus reference cv. "Nonpareil" exhibited tocopherol values higher than 500 mg/kg. These values are abundantly above those reported in literature: α–tocopherol ranged from 187.3 to 490.3 g/kg within eight European cultivars and 48 selections (Kodad *et al.*, 2006); in fourteen Turkish accessions α–tocopherol ranged from 144 to 463 mg/kg oil (Yildirim *et al.* 2010).

4.3.2 The best Sardinian accessions for oil parameters (preliminary results)

From this first-year preliminary data, some indications came out about Sardinian cultivars that are worth being considered for their oil composition; in any case observations need to be repeated at least for another season, in order to investigate how oil parameters are influenced by climatic trends.

Table 4.3 reports the lists of the best cultivars for each of the main observed parameters of kernel quality. They can be associated to some kernel characteristics: total oil content is related to the attitude to confectionery industry (that is the main utilization of that local germplasm); UFAs % / SFAs % ratio, ALA and T contents are related to nutritional characteristics; T and O/L ratio is related to storage length.

First of all, a higher oil content is desired to produce nougat and sweets, as it results in less water absorption by the almond paste; besides confectionery industry, kernels with high-oil percentage are better also for cosmetic industry (Socias i Company *et al.*, 2008). Cultivars that exhibited more than 60% lipid content are reported.

The % UFAs / % SFAs ratio is one of the parameter used in nutritional quality evaluation of fatty foods (Socias i Company *et al.* 2008), as UFA are inversely correlated with serum cholesterol levels (Kodad and Socias i Company, 2008; Sabate *et al.*, 2002). In table 4.3, cultivars with ratios ≥ 12 are listed.

As to ALA, almond oil is a minor provider, if compared to the main vegetables rich in that fatty acid (mainly flax seeds and nuts, secondarily olive oil, peanuts and pistachio seeds), but yet some accessions in our collection are worth being considered for this parameter. Table 4.3 reports cv. with ALA content ≥ 0.04 %.

Regarding T, some Sardinian accessions have a much higher content than data from literature: in table 4.3, the maximum value of 490 mg/kg oil reported by Kodad *et al*., 2006, was the minimum threshold considered.

From the storage length point of view, again, α–tocopherol plays an important role, as an antioxidant toward fatty acids (the higher it is, the longer is the storage time), but also a high oleic/linoleic acids ratio indicate a good stability. This parameter in our collection (ranging from 2.56 to 5.59 are slightly higher than literature reported values (Abdallah *et al.*, 1998: from 2.5 to 4.9). Thus in table 4.3 the cvs. with ratio ≥ 4 are listed.

Tab. 4.3 List of the best cultivars for each of the principal parameters of kernel quality

OIL ≥ 60%) (for confectionery industry)		**UFA % / SFA % ≥ 12 (for nutritional characteristics)**		**α-LINOLENIC ACID (ALA) ≥ 0.04 % (for nutritional characteristics)**		**α-TOCOFEROL (T) ≥ 490 mg/kg oil (for nutritional characteristics and storage length)**		**OLEIC / LINOLEIC RATIO (O/L) ≥ 4 (for storage length)**	
Ibba	65.1	Vargiu	12.9	Nuxedda	0.060	Farci	623.6	De Mrasciai	5.6
Bianca	64.3	Provvista	12.7	Ciatta Malissa	0.055	Corrochina	547.7	Pitichedda	5.5
Ghironi	63.9	Lutzeddu	12.7	Picantili	0.054	Provvista	536.8	Antioco Pala	5.5
Troito A	63.6	Antioco Pala	12.7	Sunda N.	0.053	Nonpareil	519.3	Troito A	5.4
Ne Plus Ultra	62.7	Troito A	12.4	Niedda I	0.050	Niedda I	515.3	Nuxedda	5.1
De Efisi Sinzoba	62.4	Malissa Tunda	12.4	De Efisi Sinzoba	0.046	Ghironi	507.2	Vargiu	5.1
Provvista	62.0	De Mrasciai	12.4	Antioco Pala	0.043	Sunda G.	496.2	Niedda II	4.9
Riu Loi	61.6	Niedda II	12.2	Troito B	0.043	Bianca	493.9	Vavani Perra	4.9
Lutzeddu	61.2	Nonpareil	12.0	Sunda G.	0.042	Sunda N.	493.8	Olla	4.8
Nonpareil	61.2	Pitichedda	12.0	Niedda II	0.042	Fiori	493.3	Efisi Sinzoba	4.7
Franciscu	61.1	Ciatta inglese	12.0	Provvista	0.042	Ne Plus Ultra	493.1	Provvista	4.6
Olla	60.4			Ciatta inglese	0.040			Ciatta inglese	4.4
Fiori	60.2			Troito A	0.040			Ghironi	4.3
Ciatta inglese	60.1							Lutzeddu	4.0

4.3.3 Multivariate cluster analysis and MANOVA

Oil parameters were also analyzed by multivariate cluster analysis, in order to verify the presence of phenotypic clusters among the accessions: the output dendrogram according to Ward, is shown in fig.4.3. From the following MANOVA analysis, the most significant partition of the dendrogram ($p<10^{-4}$, with F value = 21.7947), indicated by the red arrows, grouped the accessions in two phenotypic clusters (frames of different colors).

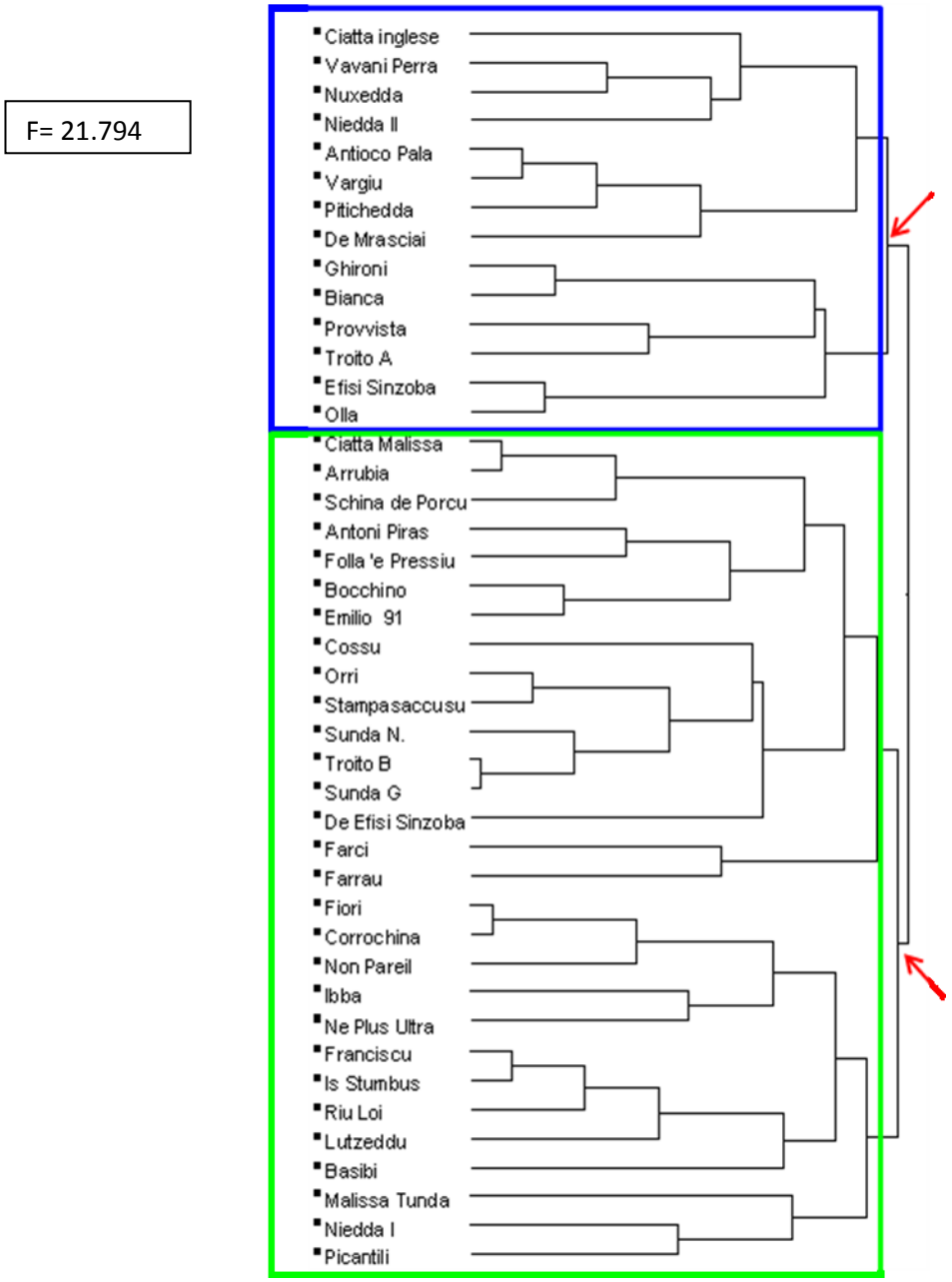

Fig. 4.3. *Clusters of accessions based on oil parameters*

The blue phenotypic cluster comprises almost all the best accessions for % UFA / % SFA and O/L ratios, meaning that cluster grouping was influenced mostly by those two parameters.

4.3.4 Conclusions

In conclusion, 31 accessions out of 43 outstand for at least one of the main parameters of kernel quality: 28 Sardinian and 3 commercial. However, none of the reference cultivars surpassed any of the best Sardinian accessions in any of the principal parameters of kernel quality (tab 4.3), although Sardinians could suffer by the competition of the much higher production of the first ones (see chapter two). The only Sardinian accession with productive level comparable to reference varieties ("De Efisi Sinzoba") can compete with them as to oil content (and thus as to attitude to confectionery industry) and overcome them as to α –linolenic acid. But "De Efisi Sinzoba" has one of the lowest % UFA / % SFA and O/L ratios, that means lower storage length and lower nutritional quality, again reflecting the characteristics of commercial varieties.

Among the 28 outstanding Sardinian accessions "Provvista" is among the best for all the five parameters; "Ciatta Inglese" and "Troito A" for all except α –tocopherol; "Antioco Pala" for % UFA / % SFA ratio, α –linolenic acid and α –tocopherol (all nutritional parameters); "Vargiu", "De Mrasciai" and "Pitichedda", for % UFA / % SFA and O/L ratios (nutritional quality and storage length); "Fiori", for oil and α –tocopherol contents (confectionery industry attitude; nutritional quality and storage). "Ibba", "Nuxedda" and "Farci" are at the first positions each one for only one parameter (respectively oil, α–linoleic and α –tocopherol contents).

All the Sardinian accessions have a low productive level compared to reference cultivars. However these accessions, well adapted to the local ecosystem, could be exploited for their particular qualitative characteristics (e.g.: distinctive food productions of local confectionary industry, nutritional and cosmetic/pharmaceutical uses, etc.). This is an interesting perspective as, in recent years, food and health aspects are receiving special attention from the general public.

4.4 Bibliography

Ahmad Z. (2010) The uses and properties of almond oil. Complementary Therapies in Clinical Practice 16: 10–12.

Abdallah A., Ahumada M.H. and Gradziel T.M. (1998). Oil content and fatty acid composition of almond kernels from different genotypes and California production regions. J. Am. Soc. Hort. Sci. 123:1029–33.

Askin M.A., Baltab M.F., Tekintasc F.E., Kazankayab A. and Baltab F. (2007). Fatty acid composition affected by kernel weight in almond [*Prunus dulcis* (Mill.) D.A. Webb.] genetic resources. Journal of Food Composition and Analysis 20: 7–12.

Davis P.A. and Iwahashi C.K. (2001). Whole almonds and almond fractions reduce aberrant crypt foci in a rat model of colon carcinogenesis. Cancer. Lett. 165 (1): 27–33.

Dietary guidelines for Americans (2005). Available on web at http://www.health.gov/dietaryguidelines/dga2005/document/.

Djousse L., Arnett D.K., Pankow J.S., Hopkins P.N., Province M.A. and Ellison R.C. (2005) Hypertension 45 (3): 368-73.

Garcìa-Lopez C., Grané-Teruel N., Berenguer-Navarro V., Garcìa-Garcìa J.E. and Martìn-Carratalà M.L. (1996). Major fatty acid composition of 19 almond cultivars of different origins. A chemometric approach. J. Agr. Food Chem. 44: 1751–1755.

Garcìa-Pascual P., Mateos M., Carbonell V. and Salazar D.M. (2003). Influence of storage conditions on the quality of shelled and roasted almonds. Biosystem Engineering 84: 201-209.

Hyson D.A., Schneeman B.O. and Davis P.A. (2002). Almonds and almond oil have similar effects on plasma lipids and LDL oxidation in healthy men and women. J. Nutr. 132 (4): 703–707.

Jenkins D.J., Kendall C.W., Marchie A., Parker T.L., Connelly P.W., Qian W. *et al.* (2002). Dose response of almonds on coronary heart disease risk factors: blood lipids, oxidized low-density lipoproteins, lipoprotein a, homocysteine, and pulmonary nitric oxide: a randomized, controlled, crossover trial. Circulation 106 (11): 1327–1332.

JMP ver. 7 . Copyright © 2007 by SAS Institute Inc., Cary, NC, USA

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

Kodad O. and Socias i Company R. (2008). Variability of oil content and of major fatty acid composition in almond (*Prunus amygdalus* batsch) and its relationship with kernel quality. J. Agric. Food Chem. 56: 4096–4101.

Kodad O. and Socias i Company R. (2009). Fatty acid and tocopherol concentrations in almond oil and its implicatione in a breeding programme. Acta Horticultureae 814, ISHS.

Kodad O., Socias i Company R.,. Prats M.S. and López Ortiz M.C. (2006). Variability in tocopherol concentrations in almond oil and its use as a selection criterion in almond breeding. Journal of Horticultural Science & Biotechnology 81 (3): 501–507.

Krings U. and Berger R.G. (2001) Antioxidant activity of some roasted foods. Food Chem. 72 (2): 223-229.

Kuriyama H., Watanabe S., Nakaya T., Shigemori I., Kita M., Yoshida N. *et al.* (2005). Immunological and psychological benefits of aromatherapy massage. Evid. Based Complement. Alternat .Med. 2(2): 179–84.

Lòpez-Ortiz C.M., Prats-Moya S., Beltràn-Sanahuja A., Maestre-Pèrez S.E., Granè-Teruel N. and Martìn-Carratalà M.L. (2008). Comparative study of tocopherol homologue content in four almond oil cultivars during two consecutive years. Journal of Food Composition and Analysis 21: 144–151.

Martìn-Carratalà M.L., Garcìa-Lòpez C., Merengue-Navarro V., Granè-Teruel N., Garcìa-Garcìa J.E. (1998). New contribution to the chemometric characterization of almond cultivars on the basis of their fatty acid profiles. J. Agric. Food Chem. 46: 963-967.

Rizzolo A., Senesi E. and Colombo C. (1994). Studies on the storage of shelled and in-shelled almonds. Acta Hort. 373: 259-264.

Sabate J., Haddad E., Tanzman J.S., Jambazian P. and Rajaram S. (2003). Serum lipid response to the graduated enrichment of a Step I diet with almonds: A randomized feeding trial. Am. J. Clin. Nutr. 77: 1379–1384.

Sathe S.K., Seeram N.P., Kshirsagar H.H.,. Heber D, and Lapsley K.A. (2008). Fatty Acid Composition of California grown almonds. Journal of Food Science - C: Food Chemistry 73 (9): 607-614.

Schirra M. (1997). Postharvest technology and utilization of almond. Hort. Rev. 20: 267-292.

Senesi E., Rizzolo A., Colombo C. and Testoni A. (1996). Influence of pre-processing storage conditions on peeled almond quality. Ital. J. Food. Sci. 2: 115–125.

Socias i Compani R., Kodad O., Alonso J.M. and Gradziel T.M. (2008). Almond quality: a breeding perspective. In: Horticultural Reviews 34: 197-238.

Yildirim A.N., San B., Koyuncu F. and Yıldırım F. (2010). Variability of phenolics, α-tocopherol and amygdalin contents of selected almond (*Prunus amygdalus* Batsch.) genotypes. Journal of Food, Agriculture & Environment 8 (1): 76-79 .

Zacheo G., Cappello M.S.,. Gallo A, Santino A., and Cappello A.R. (2000). Changes associated with postharvest ageing in almond seeds. Lebensm.-Wiss. Technol. 33: 415–423.

Zohary D., Hopf M. (2000). Domestication of plants in the old world. 3rd ed. London-Oxford University Press pg. 186.