Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo XXXIII

Settore Concorsuale: 07/B2

Settore Scientifico Disciplinare: AGR/03

CHARACTERIZATION OF THE CHESTNUT GERMOPLASM OF EMILIA-ROMAGNA REGION AND ANCIENT FRUIT TREE VARIETIES VALORIZATION

Presentata da: Sara Alessandri

Coordinatore Dottorato — **Chiar.mo Prof. Massimiliano Petracci**

Supervisore — **Chiar.mo Prof. Luca Dondini**

Esame finale anno 2020

Biodiversity is our most valuable but least appreciated resource

(Edward Osborne Wilson, The diversity of life 1992)

TABLE OF CONTENTS

TABLE OF CONTENTS

Abstract

Molecular characterization represents a valid support for the recovery of germoplasm and is also motivated by the interest for the valorization of local productions in order to make their traceability possible. Molecular characterization is also fundamental for the individuation of misnomers in collection fields in which the different varieties are preserved.

In particular, microsatellites have been used in this research to investigate the genetic diversity, inside a population and at an individual level, and the correct varietal correspondence.

The research is mainly based on the study of European chestnut (*Castanea sativa* Mill.) cultivars to evaluate the genetic diversity and relationships in Emilia Romagna region of Italy.

Subsequently, a STRUCTURE analysis was carried out at European level with the allelic frequencies of the samples collected in Emilia-Romagna. Variation found at group and subgroup level may reflect a combination of historical migration/selection processes and adaptive factors to different environments between Italian and Spanish regions.

In addition, a case study for the valorization of an old local variety and its re-introduction in the cultivation areas was proposed. This research was carried out by a morphological and molecular characterization of the local apple variety 'Rosa Romana'. The conservation of this variety entails the discrimination of different accessions with very similar phenotype that are present in the original cultivation area. The identification of historical trees and most adequate reference plants are fundamental steps for the correct propagation of this old variety and for the development of nursery activities. This will also promote and re-evaluate the exploitation and protection of such ancient Italian apple cultivars. This model could be in future also carried out for chestnut varieties.

In conclusion, analysis with molecular markers is of fundamental importance for the protection and the maintenance of local and ancient varieties which allow to increase the allelic variability available for breeding programs.

CHAPTER 1: INTRODUCTION

1.1 Biodiversity conservation and valorization

The term “biodiversity” was coined by the entomologist Edward Osborne Wilson in 1986 with the view to refer to the variety of life forms on earth.

In 1986, the word was officially presented in the report of the first National Forum on Biodiversity organized in Washington by the National Academy of Sciences and the Smithsonian Institution and addressed to the United States Congress.

However, the concept of biodiversity became part of the official scientific language only in 1992 with the publication of the essay "The Diversity of Life" (Wilson, 1992). Although this definition was not very clear, it introduced for the first time a subject destined to become of particular importance in different fields (academic, social, political, cultural) as it involves issues of environment protection.

During the 1992 World Summit in Rio de Janeiro, the United Nations Convention on Biological Diversity established that biodiversity is: «*the variability among all living organisms, including of course, those of the subsoil and air, aquatic and terrestrial ecosystems and the ecological complexes of which they are part. Biodiversity so includes the diversity within species, between species and between ecosystems*».

There are three main levels of biodiversity:

- **between ecosystems**. The diversity of ecosystems concerns the variety of interactions between different ecosystem and how these are affected by the physical environment.

- **between species.** The diversity of the species represents the taxonomic variability, that is to say the number of different species within a given geographical area.

- **within species**. The diversity within species is related to the internal variability of an individual population or to the populations that make up a species. Within a species, groups of individuals may reproduce in isolation from the original group due to geographical or behavioral factors. As a result of different selective pressures, these populations may develop different characteristics from those of the main population and thus form a distinct subspecies.

In the last decades an alarming increase in the number of species at risk of disappearance from the spontaneous vegetation has been documented. This phenomenon is mainly due to the environmental and human pressure affecting vitality and number of species. This results in a decrease of the biological possibility of propagation of species (Gulati, 2018). Among the causes of the biodiversity decline, we might count partial destruction and degradation of the natural habitat; destabilization of the ecosystems due to climate change, pollution, increase in the number of invasive species, and human factors such as industrial agriculture practices, the diffusion of modern varieties and the selection of a few wild species (Gulati, 2018).

In particular, the integration of agricultural markets, the industrialization of agriculture and the use of high-yield seeds have led to the use of an extremely limited number of plant varieties resulting in the progressive loss of most of the genetic varieties characteristic of the different geo-pedoclimatic areas (Silvanini et al., 2011).

The primary problem has been the high utilization of modern high-yielding varieties only. On the contrary, local and traditional varieties, cultivated since millennia, have been abandoned.

The extinction of all these natural varieties has a negative impact not only from a naturalistic point of view but also from an applicative one.

Since the birth of agriculture in the Neolithic about 10,000 years ago, there has been a steady growth in the number of cultivated varieties opposed to a reduction in wild species. In recent periods, in addition, a reduction in cultivated varieties also occurred: the spread of modern agriculture techniques, such as intensive farming and mechanization of agricultural practices, resulted in the selection of the most productive varieties selected for organoleptic traits, aesthetic standards and resistance to diseases (Marconi et al., 2018). As a consequence, a lot of traditional and local varieties with less productive appeal have been abandoned.

The massive use of few and related cultivars has dramatically reduced the genetic diversity of local varieties, confined to marginal, niche areas. In Central Italy a limited spread of intensive fruit orchards has made it possible to preserve much of the local genetic diversity.

Many of the local varieties, although of low productivity, were relatively stable under extreme environmental conditions, and their high genetic variability ensured reliable harvesting for local communities in the past (Albertini et al., 2015).

Over time, the mixture of modern introduced varieties with the local ones, increased the genetic variability available for breeders, but generated some confusion regarding local genetic resources and their correct denomination. Consequently, the need for characterization and clarification of possible synonyms, homonyms, and/or labeling errors in these old and local genetic resources is a fundamental and necessary step for the conservation and management of germoplasm collections (Cipriani et al., 2010; Marconi et al., 2018).

1.2 The importance of germoplasm collections

Germoplasm is a living genetic resource (as DNA, tissues, seeds or trees) that is fundamental for the purpose of animal and plant preservation, breeding, and other research uses (Williams, 1991).

Germplasm collections can range from collections of local species to domesticated breeding lines that have undergone extensive human selection. The main objective of germplasm collection is the preservation of genetic diversity of a particular plant or of a genetic stock for future exploitation (Peefers and Calwey, 1988).

Reliable yields and high-quality fruits are the features upon which the success of growers and retailers depends on. As different cultivars respond differently to pests and diseases, it is important to identify cultivars and their characteristics. Therefore, varietal identification as well as clear understanding of the genetic structure within a gene pool is a crucial factor for the establishment of efficient strategies in breeding programs (McCleary et al., 2013; Torello Marinoni et al., 2013).

In addition, genetic variability and allelic diversity in old accessions could be of extreme interest in terms of selection and adaptation toward a changing environment (Caballero and García-Dorado, 2013). Therefore, even though such varieties are characterized by low fruit quality and yield, their allelic diversity could be essential for crop improvement (*e.g.* by providing interesting traits for the development of new varieties).

The main problems of germplasm collections are the high management costs necessary to maintain a large number of varieties and to identify accessions. Further difficulties are caused by the lack of recognized varietal standards and by the strong environmental variability which hinders accession identification on a phenotypic basis only (Cipriani et al., 2010). In addition, high redundancy levels are linked to rough naming criteria by farmers (Cipriani et al., 2010; Liang et al., 2015).

To date core collection is a standard approach to improve evaluation, management, and use of germplasms, (see Global Action Plan for the Conservation and Sustainable Use of Plantogenetic Resources for Food and Agriculture; FAO, 1996).

A core collection is a representative sample of the whole germoplasm collection that minimizes redundancy and maximizes genetic diversity. Following Frankel and Brown's proposal in 1984 the core collection approach has been applied in many genetic banks.

An efficient core collection should be smaller than the whole germoplasm collection (from 5% to 20% - Brown, 1989; Yonezawa, 1995; van Hintum et al., 2000; Wang et al., 2011).

With the core collection approach, molecular markers can provide a valid tool to assess genetic diversity, detecting duplicates, synonymies or homonymies and supporting plant genetic resources manegement (Marconi et al., 2018). The molecular characterization of germplasm collections aims to reduce the number of genotypes to collect, removing redundancies in order to unique genotype identification. (Hummer et al., 2015; Linag et al., 2015; Urrestarazu et al.,2016; Pereira-Lorenzo et al., 2017).

Genotypes accessions in germplasm collections can be also used in varietal certification to determine and verify cultivar pedegree. Access to genetic information from different germplasm collections is a valuable resource for validation and reconstruction of plant pedigrees.

The increasing availability of genetic information relating to germplasm collections allowed the understanding of the evolutionary relationships and origins of domesticated species through phylogeny reconstruction and population genomics (Urrestarazu et al. 2012 and 2016; Pereira-Lorenzo et al., 2017).

1.3 Molecular characterization (fingerprinting with microsatellites)

The growth of the agricultural market in recent decades led breeders to focus on research of new techniques for varietal characterization, genetic improvement and enhancement of local varieties traits.

Previously varietal identification was based on the phenotypical traits only, i.e. visual examination of pomological characteristics (ripening time, shape, qualitative descriptors - Sansavini et al., 1997). The main weakness of this identification method is that morphological characters could be influenced by environmental variability.

The main limits of such morphological characterization are the following: polygenic control of morphological and phenological traits; strong influence of pedoclimate conditions; orchard management and the need of large characters number in order to draw up a suitable description card (Sansavini et al., 1997; Silvanini et al., 2011).

Therefore, the use of phenological traits alone does not consent accurate varietal identification. For this reason, also molecular markers are currently used for varietal identification purposes.

Molecular techniques allow the study of the genome of plant species by tracing genetic imprints for each genotype and analyzing particular areas of DNA called molecular markers.

Molecular markers are therefore based on detection of differences (polymorphisms) in the DNA nucleotide sequence.

Numerous molecular markers are used for different purposes other than cultivar identification, ranging from population genetics, linkage analysis and assisted selection *e.g.* RAPD (Random Amplified Polymorphic DNA), I-SSR (Inter-simple Sequence Repeats or Inter-microsatellites), SSR (Simple Sequence Repeats or Microsatellites), SAMPL (Selective Amplification of Microsatellite Polymorphic Loci), AFLP (Amplified Fragment Length Polymorphism) and more recently SNP (Single Nucleotide Polymorphism).

Microsatellites (SSRs) are particularly important for varietal identification. They highlight polymorphisms at repeated DNA sequences level by using specific primers which are complementary to the sequences flanking the microsatellites.

The number of microsatellite repetitions can vary from one individual to another and also within the same individual in homologous chromosomes, allowing the detection of different alleles of a heterozygous individual. Microsatellites are, therefore, codominant markers. The bases repeated in tandem can be from 1 to 6, constituting mono-nucleotide, di-nucleotide, tri-nucelotide, tetra-nucleotide microsatellites (Silvanini et al., 2011).

Compared to other types of markers, SSRs offer many advantages. They are ubiquitous in the genome, not influenced by the environment at development stage of the plant. They also have a simple Mendelian inheritance which is ideal for genetic studies.

The SSRs molecular analysis data were collected in numerical matrices that show the analyzed genotype and the electrophoretic profile obtained for each molecular marker using the ABI 3730 XL Analyzer sequencer model (Applied Biosystems). The data are binary (0 and 1): presence or absence of DNA fragments is recorded.

Data matrices are analyzed with appropriate software, such as CERVUS by Kalinowski et al. (2007) and NTSYS 2.0 that consent to highlight equal or similar molecular profiles; perform paternity tests; hypothesize possible kinship. The level of data confidence is calculated according to the number and polymorphism of the markers used in the analysis. Cases of synonyms can be easily highlighted.

Using appropriate Bayesian analyses, it is also possible to establish, with a good degree of confidence, the membership of an anonymous sample to a certain variety. The latter type of analysis is performed by creating a reference molecular data archive for comparison between the molecular profiles of the unknown sample with those known and already recorded (fingerprinting).

Molecular analysis with microsatellites provides an important tool in distinguishing very similar varieties and detecting errors in the propagation phase.

Due to their high polymorphism and co-dominant inheritance, microsatellites are considered the best marker to investigate genetic diversity within a population and at an individual level (Blair et al., 2009; Zhang et al., 2011; Patzak et al., 2012; Emanuelli et al., 2013; Liang et al., 2015; Urrestarazu et al., 2016; Mousavi et al., 2017; Pereira-Lorenzo et al., 2017 and 2020).

With the study of SSR allele frequencies, it is also possible to investigate population structure among groups of individuals (Liang et al., 2015; Urrestarazu et al.,2016; Pereira-Lorenzo et al., 2019). These analyses use a model-based Bayesian approach implemented in the software STRUCTURE (Pritchard et al., 2000). This approach does not require any prior information to assign individuals to different populations (Pritchard et al., 2000).

The STRUCTURE analysis has been extensively used to assess the genetic structure of several fruit tree species; such as pear (Volk et al., 2006; Miranda et al., 2010; Ferreira dos Santos et al., 2011; Baccichet et al., 2020), peach (Aranzana et al., 2010), apple (Liang et al., 2015; Urrestarazu et al., 2016) and sweet cherry (Mariette et al. 2010).

SSRs varietal identification, as well as providing a clear understanding of genetic structure within a gene pool, is a crucial factor for the establishment of efficient strategies in breeding programs (Torello Marinoni et al., 2013; Urrestarazu et al., 2016; Pereira-Lorenzo et al., 2017).

Microsatellites have been also widely used to assess the genetic diversity in core collections of fruit trees (Yun et al., 2015; Lassois et al., 2016; Urrestarazu et al., 2016, Pereira-Lorenzo et al., 2017) for cultivar characterization (Goulão and Oliveira, 2001; Patzak et al., 2012; Pérez-Romero et al., 2015) and parentage analysis (Kitahara et al., 2005; Moriya et al., 2011; Lassois et al., 2016, Urrestarazu et al., 2012 and 2016; Pereira-Lorenzo et al., 2017).

1.4 Germoplasm conservation: *in situ* and *ex situ*

Living collections are important tools for preserving germplasm as repositories in orchards, plantations, and vineyards. These resources promote plant genetic variability as a living source and include almost of the current commercial cultivars, ancient varieties, breeding material and wild trees. Germplasm collections are essential to improve and better understand research in plant biology, crop improvement and biodiversity conservation. (Figure 1.1).

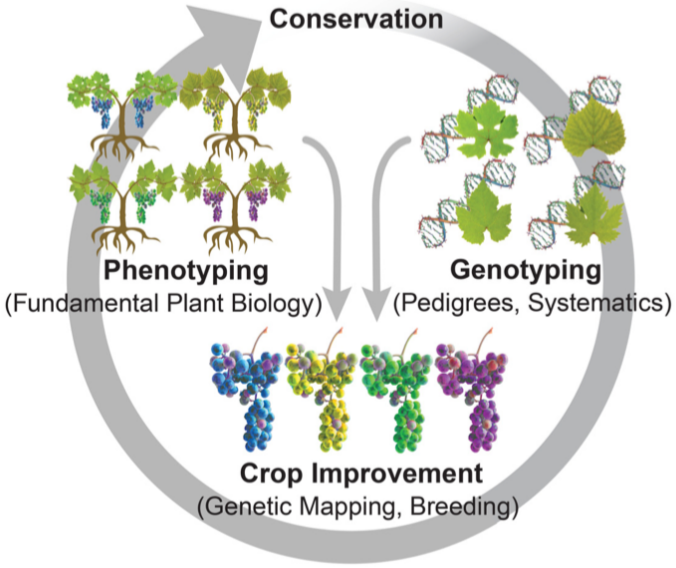

Figure 1.1: Three primary benefits of germplasm collections: phenotyping, genotyping, and crop improvement.

Germplasm collections are used also to describe phenotypic and interannual variations in large numbers of accessions living under common conditions.

Using germplasm collections to study phenotypic variation not only improves the knowledge on plant biology but can be also used for studying phenotypic plasticity in plants under changing climatic conditions.

One mechanism for adapting agricultural species so as to respond to climate change is the introgression of traits from related wild species with features that better suit to current environmental conditions (Warschefsky et al., 2014). It is essential to compare the different germplasm collections to identify those genotypes that better withstand such conditions.

Germplasm collections are also a valuable resource for the study of somatic variation across clones. Somatic mutation generates genetic variation in perennial fruit crops; thus, this results in phenotypic differences which can be identified and are indeed often retained as potential new cultivars or germplasm.

The protection of germplasm can be pursued with two different methodologies: *in situ* and *ex situ* conservation:

1. *In situ* conservation means the conservation of species in the areas of origin, with the aim of protecting biological resources by studying the growth and development of the species in their natural habitat (Figure 1.2). It involves the creation of protected natural areas such as genetic and biosphere reserves, parks and oases.

Figure 1.2 Chestnut *in situ* conservation at the field of La Martina, Monghidoro (BO).

The goals are:

• To evaluate the actual genetic biodiversity of each species (both local and commercial);

• To study the mechanisms that regulate the maintenance of genetic diversity (such as characteristics of habitat, climate, reproductive self-incompatibility).

Effective *in situ* conservation requires continuous characterization of species and population analyses based on genetic frequency. This consents to assess variations in biodiversity which are useful for *ex situ* conservation strategies.

2. *Ex situ* conservation means the maintenance of species outside their natural habitat with the aim of ensuring biodiversity and making plants available for human activities. *Ex situ* conservation has the disadvantage of providing a partial conservation only of the different species. *Ex situ* conservation has the following purposes:

• To develop new cultivars through breeding;
• To provide reserve populations to promote survival of species during the reintroduction/restocking stage or to favor habitat recovery;
• To provide research plant material.

To achieve these objectives, databases containing genetic and taxonomic information have been set up (germplasm banks, *in vivo* plant collections, DNA and seed banks, tissue and cell cultures). Both synthetic and detailed descriptive morphological cards have been developed by research activity of both national (List of fruit cultivars found in Italy, Florence 1998; Fideghelli, 2016), and European working groups (European Cooperative Program for Plant Genetic Resources).

The germplasm collections are distributed throughout Europe, but mainly in the regions with the longest fruit harvesting tradition: Belgium, France, Spain and Italy. Following the alarm-bell rang by the scientific community on the risks related to the loss of biodiversity and the extinction of many varieties that are no longer cultivated, the European Community has developed a regulatory framework for the safeguard and protection of germplasms. Law 194/2015 was recently issued in Italy "for the protection and enhancement of biodiversity". This law provides for the implementation and updating of the national biodiversity archive and establishes guidelines for the conservation and characterization of plant biodiversity of interest to agriculture.

On a regional level, Emilia-Romagna introduced Law No. 1 d.d. 29/01/2008 which supports recovery and reintroduction projects of ancient native varieties at risk of extinction. Currently in the "Repertoire" introduced by the aforesaid regional law, 107 varieties at risk of extinction are registered and reported by research and scientific institutes and agricultural farmers. The ancient varieties reproduced and planted in the collections are mainly chosen of risk of extinction. In some cases, there are only few ancient trees absent which the variety it would be lost forever. Another criterion pursued is to save the varieties that have agronomic and organoleptic characteristics interesting for relaunch and commercial exploitation on local markets.

1.4.1 Chestnut (Castanea sativa Mill.) germoplasm collections

In last decades, chestnut cultivation declined in most of Italian chestnut growing regions for several factors such as climate change, spread of diseases (chestnut blight, gall wasp and ink disease), lack of skilled growers, rural decline and lack of genetic data. All of which are crucial for improving conservation and management skills (Mellano et al., 2012; Fedrigotti and Fischer, 2015).

The long history of cultivation of chestnuts in Italy is characterized by clonal propagation and selection of the best genotypes. We can recognize two different sub-groups: chestnut-type and Marroni-type varieties. On a botanical level, chestnuts are produced by wild individuals of *Castanea sativa* and by varieties of grafted chestnuts. The Marroni-type varieties are produced by individuals of a macrocarpa (which means "large fruit") subspecies of *C. sativa*.

In Italy, the Marroni-type is considered a variety of excellent quality. ‘Marrone di Castel del Rio’ which originates from Emilia Romagna region has been awarded a PGI certification, being one of the most valuable and well-known chestnut cultivars in Italy and abroad.

Chestnut fruits contain a large quantity of bioactive compounds such as monoterpenes, phenols and vitamin C useful for human diet. The compounds levels of chestnut were equal or higher with respect to the main hazelnut, walnut and almond varieties (Beccaro et al., 2020).

To date, the interest in and demand for sustainable and local products is considerably increased, contributing to a global process of valorization of chestnut cultivars with particular reference to the Italian market (Fedrigotti and Fischer, 2015). The rediscovery of local fruit crops encourages the in-depth study of the chestnut tree and the enlargement of local germplasms to preserve the existing biodiversity and eventually identify desirable traits that could be potentially useful to chestnut industry.

The above positive trend is however threatened by rural decline and abandonment of mountain areas as well as the decay of ancient chestnut orchards. This risk can be prevented by the identification and conservation of local varieties at risk of extinction. It is of fundamental importance to protect and conserve local germplasm in collection fields in order to avoid further loss of genetic biodiversity.

In particular, in the Emilia-Romagna region there are three different collection fields (Figure 1.3): Didactic and Experimental Park of Granaglione (Bologna); Zocca (Modena) and Paloneta Brisighella (Ravenna).

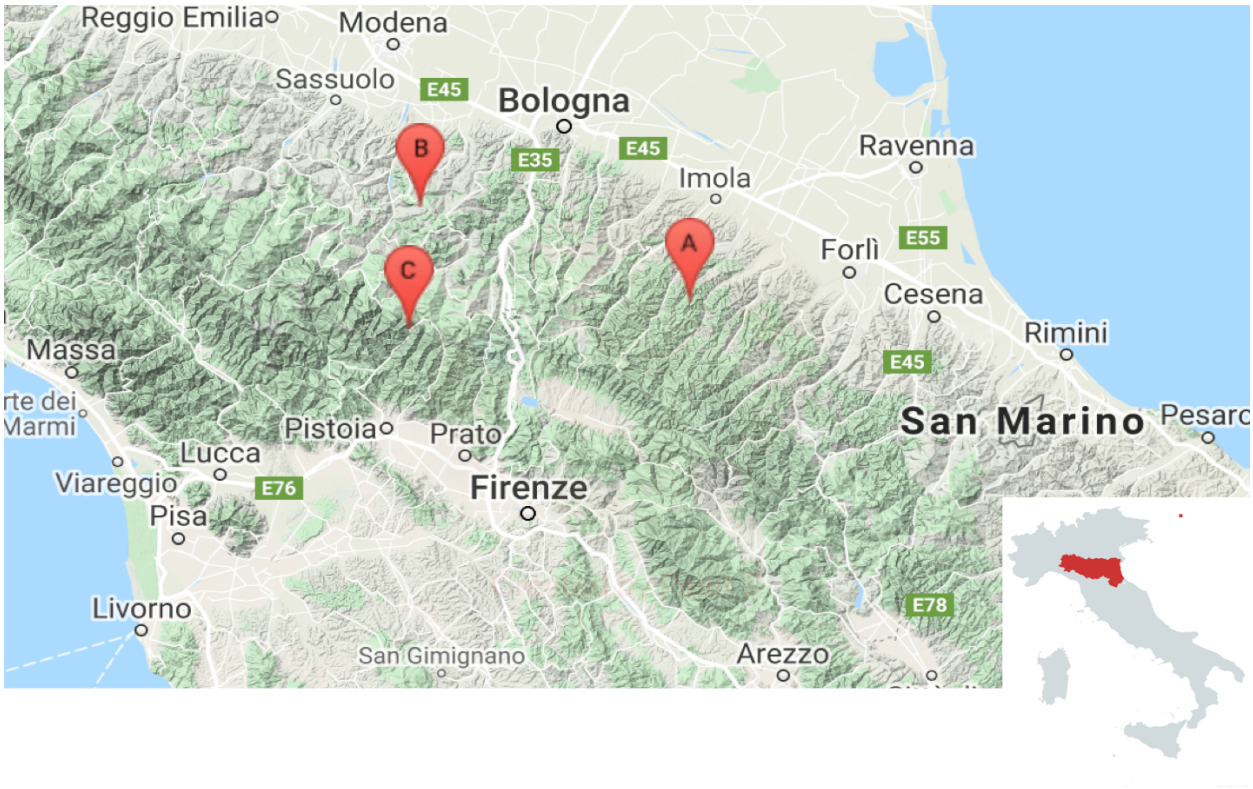

Figure 1.3: Maps of the three collection fields analyzed in Emilia-Romagna region, Italy. A) Paloneta (FA); B) Zocca (MO); and C) Didactic and Experimental Park of Granaglione (BO).

These collection camps are located in the "*castanetum*" phytoclimatic area (de Philippis, 1937), which are typically characterized by sub-acid soil (pH around 5-6.5); annual rainfalls of 600-1500 mm; mean annual air temperatures of 9-13 °C (Heiniger and Conedera, 1992; Bonuous, 2002; Gomes-Laranjo et al., 2008; Perulli et al., 2020)

The Didactic and Experimental Park of Granaglione is an important center for the conservation of chestnut biodiversity (Figure 1.4). The site contains most of regional varieties such as the Marroni-types ('Marrone di Zocca' and 'Castel del Rio') and other Italian Marroni such as the 'Roncegno', 'Centa di S. Nicolò' and 'Drena' cultivars. Within the collection field there are also several local chestnut cultivars (Tuscany-Emilian Appenines).

Figure 1.4: Didactic and Experimental Park of Granaglione (BO).

On the other hand, the collection field of Zocca is divided into reference and scion fields. Inside of both collection fields there are numerous varieties of chestnuts from Emilia-Romagna region and several clones of the 'Marrone di Zocca'. In particular, in the scion field the plants are kept in a juvenile stage in order to be able to collect scions annually to be distributed to chestnut growers (Figure 1.5).

Figure 1.5: Chestnut trees maintained in juvenile stage in Zocca collection field.

The Paloneta collection was created in 1986 by Professor Bellini of the University of Florence (Figure 1.6). About 80 accessions were selected among the Marroni-type and chestnuts varieties mostly consumed as fresh products and used for flour and wood. Different genotypes were selected and preserved therefore contributing to the conservation and maintenance of agrobiodiversity of the chestnut germplasm.

Figure 1.6: Collection field of Paloneta, Brisighella (RA)

1.5 Molecular review of European Chestnuts.

1.5.1 Origin and Geographical Distribution

The various species of Castanea spp., a member of the Fagacae family, have morphological and ecological traits well differentiated with regard to the vegetative habitus, the size, the characteristics of the fruit and wood, the adaptability and resistance to biotic and abiotic factors, tracts that have greatly influenced their spread over time. Three different species with different diffusion areas have been identified: *Castanea sativa* Mill in the Mediterranean area; *Castanea crenata* Sieb. and Zucc. and *Castanea mollissima* Bl. in Asia; *Castanea dentata* Borkh in North America (Bounous, 2002).

The only native species spread in Europe is *Castanea sativa* Mill, known as a multipurpose species for its high-quality nuts, timber and flour production, widely cultivated throughout Europe since the ancient times.

Natural events, such as glaciation, and human influences, played an important role in the geographical and genetic distribution of this species (Conedera et al., 2004; Poljak et al., 2017). For this reason, it is very difficult to trace the natural spread of chestnut in Europe, particularly in the Mediterranean area (Conedera et al., 2004).

The palynological data (Conedera et al.,2004) indicate that the first real expansion of this species due to human influence dates back to 37,000 years ago in Turkey, north-east Greece and south-east Bulgaria (Figure 1.7).

The chestnut tree has been at first introduced in South Italy by Greeks about 5,000 years ago; later by Romans from Turkey and the Caucasus, and then it was spread in Europe throughout the Middle Ages mainly by the Benedictine monks and thanks to the interest of Matilde of Canossa (Huntley and Birks, 1983; Bernetti, 1995; Krebs et al., 2004 and 2019; Mattioni et al., 2013 and 2017; Roces-Diaz et al., 2018). Roughly 1,000 years ago, the chestnut trees attained the current distribution, with the exception of England, which was colonized later (Conedera et al., 2004; Jarman et al., 2019 – Figure 1.7).

In that period, also known as the 'proper chestnut civilization period' in Western Europe, chestnuts became a vital part of people's diet (Conedera et al., 2004). Later in Central-North Italy and in France, the Marroni cultivar was selected for commercial purposes, given the good-quality of fruits (Conedera et al.,2004; Gobbin et al., 2007).

The decline began at the end of the nineteenth century, conditioned by major social and economic changes, in particularly at the beginning of the twentieth century (di Trento, P. A., 2008).

Nowadays the cultivation is present mainly in Italy, especially in the Mediterranean area (Sardinia, Sicily, Tuscany, Calabria and Campania), in the Apennines and Alps (Piedmont and Trentino-Alto Adige), in France, Spain and Portugal. There are also small chestnut populations in southern England and in central-north Germany and Switzerland.

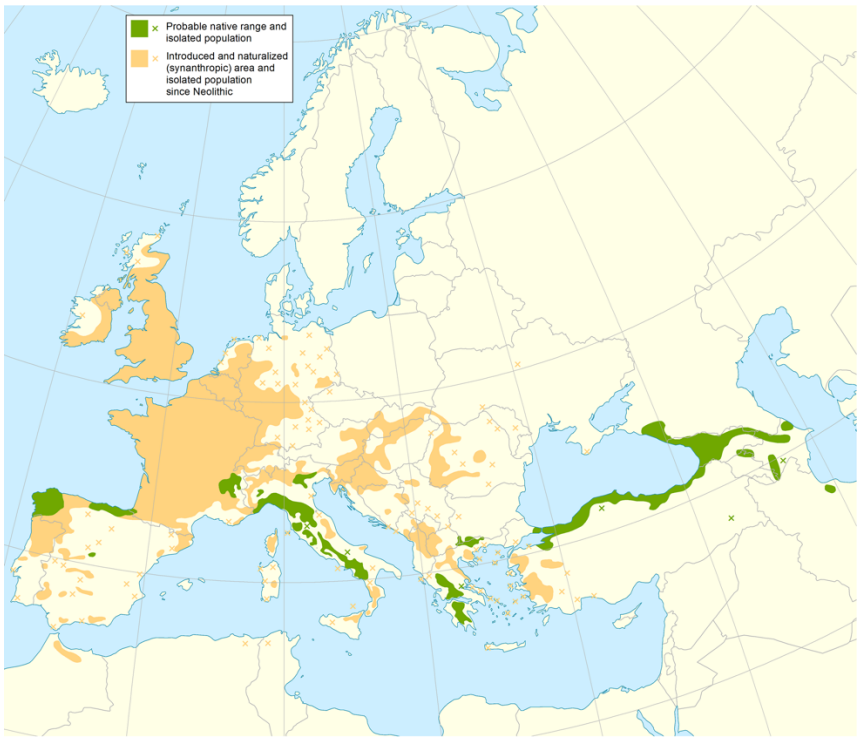

Figure 1.7: European chestnut distribution (European Union, 2017)

A great number of chestnut varieties have been described in Europe since the beginning of the last century based on morphological evaluation (Lavialle, 1906; Vigiani, 1908; Breviglieri, 1951; Valle, 1959; Bergamini, 1975). Nowadays, the advances in molecular techniques offer powerful new tools, allowing the conservation of the genetic resources, the protection of the qualities of the different varieties, as well as the implementation of management strategies. *Castanea sativa* Mill can be divided into two different groups: chestnuts and sweet chestnuts (or Marroni) groups (Bounous and De Guarda, 1999; Gobbin et al., 2007; Mellano et al., 2012; Marinoni et al., 2013).

In particular, sweet chestnuts are the results of selection by growers, in search of high-quality nuts and wood, thus this variety was distributed in Italy and France through propagation of plants from the Emilian-Tuscany Apennines thanks to Matilde di Canossa (Conedera et al.,2004). However, clonality in the Marroni group was the result of the cultivar importance in a specific region and thus it represented a low risk strategy to maintain the local populations with superior traits (Pereira-Lorenzo et al., 2011; Marinoni et al., 2013).

1.5.2 Molecular markers

Castanea spp. has 2n = 24 chromosomes (Jaynes, 1972), the number characteristic of most of the Fagaceae (Mehra et al., 1972; D'Emerico et al., 1995). The first markers used to describe the genetic diversity of *Castanea sativa* Mill. were isozymes (Sawano et al., 1984; Fineschi et al., 1994; Ramos-Cabrer and Pereira-Lorenzo, 2005), followed by random amplified polymorphic DNAs (RAPDs) (Fineschi et al., 1994; Italia Galderisi et al.,1998; Ponchia et al., 2001) and by expressed sequence tags (EST) (Scott et al., 2000; Krutovskii and Neale, 2001; Kalia et al., 2011; Martin et al., 2010 and 2017). These last markers allow a better characterization of functional diversity in relation to adaptive variation and interspecific transferability (Varshney et al., 2005; Yatabe et al., 2007). In the last twenty years microsatellite markers (SSRs) were mostly used for a more accurate estimate of the genetic diversity of the population under study (Buck et al., 2003; Marinoni et al., 2003).

The evaluation of the genetic diversity is essential for planning a conservation strategy and breeding programs, in order to create cultivars with a pathogeny resistance (for example against pathogenic fungus *Gnomoniopsis pascoe* and the Chinese wasp *Dryocosmus kuriphilus*) and with the best quality of nuts. The first to be studied were: a) the interspecific hybrid 'Bouche de Bétizac' (*C. sativa* 'Bouche Rouge' × *C. crenata* 'CA04') from France (Sartor et al., 2009; Dini et al., 2012; Botta et al., 2012); b) the 'Pugnenga' cultivar of *C. sativa* Mill., an Italian cultivar native of the Cuneo Province (Piedmont); c) the 'Savoye', a cultivar of *C. sativa* Mill native of France, from the Midi-Pyrenees Region (Sartor et al., 2015).

Molecular markers have become an effective way to address the genetic characterization of a plant population, estimating genetic diversity and determining the genetic relationships between the accessions. These tools allow to create an identity card (fingerprint) of any individual variety/cultivar.

Microsatellites (or SSR, simple sequence repeats), in particular, are considered the most suitable markers for exploring genetic diversity. The advantages of microsatellites lie in their high reproducibility and their high degree of polymorphism due to the high rate of mutation of this type of sequences. For these reasons the SSR markers are useful for characterizing demographic patterns of variation (migration and drift), studies of gene flow, introgressive hybridization.

The SSR markers that have been commonly used in *Castanea sativa* Mill., include the following series: CsCAT as di-nucleotide (Botta et al., 2001; Marinoni et al., 2003), EMCs being tri-nucleotide repeats (Buck et al., 2003), OCI, OAL, RIC and CIO (Gobbin et al., 2007) obtained from Castanea species, but also SSR obtained from the Quercus gene, such as QpZAG and QrZAG (Barreneche et al., 2004).

Nowadays, new SSR are more and more developed, not only for cultivar identification (fingerprinting) but also for more accurate genome mapping (Casasoli et al., 2001; Barreneche et al., 2004; Torello-Marinoni et al., 2017 and Staton et al., 2019) and QTL analysis (LaBonte et al., 2018; Nishio et al., 2018 and Ji et al., 2018).

1.5.3 Characterization of the European populations with SSRs

The published studies investigating on chestnut genetic diversity distribution, have drawn a quite complex picture of the tree populations in Europe.

Pereira-Lorenzo et al. (2010) performed the first large-scale molecular study using 10 SSRs from clonally propagated cultivars (574 accessions) of *C. sativa* Mill. and hybrids (71 accessions). The authors showed that there were two main sources of variability in chestnuts trees in the Iberian Peninsula: one in the North (Asturias and Galicia) and a second in the Centre of the peninsula (Tras-Os-Montes, Salmanca, Cáceres and Ávila). The results indicated, in addition, that the Southern and Canary Islands cultivars were inside the main gene pools found in the Iberian Peninsula, indicating a common origin.

Another study conducted by Pereira-Lorenzo et al. (2011), analysed grafted chestnut trees from Portugal and Spain, analysing accessions from the northern Iberian Peninsula to the Canary Islands and the Azores, using 10 SSRs. Ten principal cultivar groups were identified: 4 in northern Spain, 5 in the centre of the Iberian Peninsula and one in southern Spain, more related with the gene pool of the centre of Iberian Peninsula.

The Northern area presented the higher variability of alleles and gene diversity, while the Central-Southern area were the place of origin of the main cultivar groups. Additionally, the gene pool of southern Spain showed a close relationship with both main clusters found in the Iberian Peninsula, indicating that the main cultivars, as 'Longal', were selected more recently for the establishment of new orchards in Andalucía (Southern Iberian Peninsula) and in the Canary Islands. This contributed to the development of new important cultivars: 'Laga', 'Pilonga', 'Temprana' and 'Pelona' in the South; 'Injerta', 'Verata', and 'Mondarina' in Extremadura.

The study demonstrated that the cultivar origin and the diversification process were a result of clonal propagation of selected seedlings, hybridization and possible mutations, that contributed to the high level of genetic diversity, as supported Martin et al. (2009).

Martín et al. (2012), studied chestnuts accessions mainly of Spanish origin. Their results revealed a high genetic diversity within the samples: three clusters were identified, each cluster representing one geographical region (southeast, northwest and northeast Spain).

A similar study was also performed by Mattioni et al. (2013), comparing chestnut populations from Spain, Greece, Turkey and Italy. The STRUCTRE results were congruent with the hypothesized glacial refugia, proving the migration of the chestnut from Turkey and Greece to Italy, demonstrating the genetic divergence between the eastern (Greek and Turkish) and western (Italian and Spanish) populations.

Torello-Marinoni et al. (2013) explored samples from Piedmont, defining four populations strongly linked to geographical origin and prevalent use. These authors also found a great homogeneity between the cultivars with the denomination 'Marroni', that suggested a monoclonal origin, maintained through grafting techniques by growers that selected them for their high-quality fruit traits.

Quintana et al. (2015) measured a high genetic diversity concentration in the glacial refugium of El Bierzo, in the Castilla y León region north-west. Their data supports the elevated levels of genetic diversity present in Spanish chestnut populations. Similar results were previously obtained by Martín et al. (2012), highlighting that El Bierzo region represents a hot spot with high genetic variability, quite separated from any other European population.

These results were in line with the previous results by the study of Mattioni et al. (2008), in which the authors identified the existence of a small additional gene pool located in the northwest of Spain, diverging

from the main European pool. The likely explanation was the dispersal of genetic material from one or more source populations over time by human activities.

Mattioni et al. (2017) conducted a large-scale structure for characterization of 1,608 wild trees samples in 73 European sites. The authors identified three main gene pools: the first gene pool included populations from eastern Turkey, Azerbaijan, Georgia, Russia and one population from Romania; the second, populations from western Turkey, Greece and Bulgaria and the third cluster from all the western European populations. Furthermore, the authors identified a significant structure barrier dividing the eastern from the central and western European populations.

Poljak et al. (2017) studied samples from central Europe and western part of the Balkan peninsula. The authors assessed the presence of three clusters: two mainly located in northern and western parts of the region and the other placed toward the Greek border. It was hypothesized that the two different groups could have originated from glacial refugia in Central Europe and the southern parts of the Balkan peninsula respectively. It was also stated that the southern group could as well have originated via migrations from Asia Minor. The results suggested that the regular geographic distribution of the populations was mainly due to natural diffusion, rather than human intervention.

In addition, with 24 highly polymorphic SSRs, Pereira-Lorenzo et al. (2017) selected 118 European cultivars, out of which 96 were from Spain, 16 from Italy, 4 from France and 2 from Portugal. Two major clusters were identified: Spanish and Italian cultivar clusters, where Italian cluster demonstrated a higher genetic diversity. The results are also in accordance with the research previously carried out by Mattioni et al. (2016), where the cultivars from north-west Spain were divided from central and southern Spain, as well as from southern Italy. Furthermore, the authors identified the additional genetic substructure of five different groups of cultivars.

Furthermore, Pereira-Lorenzo et al. (2019) carried out the first genetic analysis on ancient giant chestnuts in Italy and the Iberian Peninsula to better understand the effect of graft on the chestnut domestication process, and to study the impact of cultivars selection on maintained genetic diversity. For this reason, the sampling was done by the shoots and the base of the trunk in order to distinguish wild and grafted trees. Evidence of "instant domestication" was obtained only recently in the same areas where cultivars were spread. The results showed no distinct genetic structure among wild and cultivated chestnut trees. The two recent woks of Bouffartigue et al. (2019 and 2020) were in line with these findings and added more

information about the European gene pool, inasmuch it recognized a genetic structure affected by natural events, as the recolonization after the last glaciation, and by historical human processes, where it is possible to recognize a common origin of the most part of French varieties with the Iberian Peninsula and the association of the Italian gene pool with the South-East France.

Jarman et al. (2019) collected new evidence from genetic, dendrochronological, archaeological and historical analyses in England and Wales chestnut trees compared to Europe. Genetic analysis showed that the sweet chestnut trees of England derived from France, Spain, Portugal, Italy and Romania. These areas were important refugia during the Last Glacial Maximum for sweet chestnut.

A recent study investigated the genetic diversity of Switzerland chestnut cultivars (Pereira-Lorenzo et.al., 2020); a STRUCTURE analysis identified two different main clusters: the first was mostly related to the European Chestnut Genetic Dataset varieties and the second represented a specific genetic group in Switzerland, possibly a survival of abandoned chestnut cultures of the Medieval times (Pereira-Lorenzo et al., 2017).

1.5.4 Conclusions/Summary

Chestnut growing lost diffusion in the recent years; but in the past it was the most widespread cultivation in mountain areas, where chestnuts were the main source of sustenance, called the “bread tree”. This contributed to originate a large number of varieties through graft or seedling. This ancient practice nowadays creates difficulties in distinguishing the different varieties based only on morphological characters, as common practice in the past. Nowadays, molecular analysis has been recognized as essential in order to identify and characterize the varieties cultivated in the different areas of diffusion of the species.

The aim of this review was to collect chronologically the main molecular characterization studies carried out with microsatellites (SSR) in order to investigate the richness of the germplasm of Europe chestnuts.

Verification of varietal responsiveness and characterization of cultivars is not aimed only at the genetic characterization for the recovery and preservation of germplasm, but also at the enhancement of local productions, to make traceability possible (*i.e.* products protection by national or European certifications). Moreover, the detection of errors in the collection fields is a fundamental tool for *ex-situ* conservation of the germplasm studied, for the identification of cases of homonym/synonym and resolution of legal disputes for the certification of productions.

The revival of modern chestnut cultivation involves the preliminary acquisition of information on local germplasm and the cataloguing of accessions through the use of morphological descriptors and molecular markers. However, since the experiences conducted so far have followed heterogeneous methodologies, it would be useful to standardize the methods and the use of a common description card that is the starting point in the work of varietal characterization, like, a common set of primers (12-15 SSRs) shared at national and international level.

There are currently two main issues: a) few authors publish the allelic profile for each cultivar analyzed, causing the inability of researchers to compare the data obtained; b) when allelic profiles are published it is not always easy to compare the data as there is no standardization in reading the size of alleles.

These issues could be overcome by the use of dinucleotides (CsCAT series) along with trinucleotides primers (EMCs series). The latter, being less polymorphic, have a low rate of mutations that would allow for greater ease in interpretation and comparison with data produced in other laboratories in order to create a common dataset.

1.6 Aim of the study

The use of molecular markers in varietal characterization of fruit species is a widely used and well-established tool that integrates morphological characterization. In particular, microsatellites (SSRs) have been used in this research to investigate the genetic diversity both at population and individual level so as to identify accurately plant varieties.

This thesis is divided into four chapters: this Introduction, *i.e*. Chapter 1, and the three following chapters the content of which is briefly summed up below.

- **Chapter 2**

This chapter is mainly based on the study of chestnut (*Castanea sativa* Mill.) cultivars in the Italian region of Emilia Romagna in order to evaluate their genetic diversity and relationships. In particular, this chapter focuses on the mountain area of the Tuscan-Emilian Apennines (about 600 m.s.l.).

By adopting molecular characterization techniques, this research aims to preserve biodiversity of the existing ecotypes and chestnut heritage from further genetic erosion. The main goal is to establish a standard reference database that will be used as a starting point for future selection programs promoting the chestnut fruit cultivation.

- **Chapter 3**

This chapter is focused on the importance of the correct identification of the chestnut varieties collected in the germoplasm collection Didactic and Experimental Park of Granaglione, promoted by the "BIODIVERSAMENTE CASTAGNO" (project with the contribution of the Emilia-Romagna Region) and by the National Academy of Agriculture (with the support of the Fondazione Cassa di Risparmio di Bologna). The identification of synonym accessions and the related morphological characterization

underlined the importance of verifying germplasm collections with powerful tools such as molecular markers.

- **Chapter 4**

This chapter analyses the chestnut trees at European level. By means of a STRUCTURE analysis the molecular data related to Emilia-Romagna chestnut cultivars have been compared with molecular data related in particular to Spanish cultivars. In this way, it has been possible to better investigate the genetic connections between Spanish and Italian chestnut accessions.

- **Chapter 5**

In addition, this work also aims to confirm the identity of an ancient local variety of apple to determine the possibility of promoting old local varieties and their re-introduction in the regional markets. In this case the genetic diversity analysis by SSR markers represents a tool for the certification of the genetic identity of the old varieties. In this Chapter the case of study of the 'Rosa Romana' apple is presented. A molecular assay was carried out for identifying the true-to-typeness of 'Rosa Romana' accessions. Therefore, the ancient varieties still present *in situ,* residual orchards and collections are studied. The identification of the genetic variability present among the ancient 'Rosa Romana' genotypes is one of the first steps for promoting and supporting the nursery activities that are fundamental for the availability of trees for the new orchards. This case study was also planned for defining the reference trees that could be used for producing certified plants by grafting. The characterization of the accessions of 'Rosa Romana', therefore, contributes to the recovery and valorization of the local genetic heritage which was lost over time, in order to requalify it.

1.7 Bibliography

Albertini, E.; Dalla Ragione, I.; Falcinelli, M.; Ferradini, N.; Gramaccia, M.; Lorenzetti, S. (2015). Il contributo delle analisi genetiche allo studio ed alla conoscenza della diversità delle varietà locali di melo, in La Biodiversità di Interesse Agrario Della Regione Umbria - Specie Arboree da Frutto, Vol. 2, (Todi PG: 3A-PTA), 12–57.

Aranzana, M.J.; Abbassi, E-K; Howad, W.; Arús, P. (2010) Genetic variation, population structure and linkage disequilibrium in peach commercial varieties. BMC Genet 11:69

Baccichet, I.; Foria, S.; Messina, R.; Peccol, E.; Losa, A.; Fabro, M.: Gori, G.: Zandigiacomo, P.; Cipriani, G.; Testolin, R., (2020) Genetic and ploidy diversity of pear (Pyrus spp.) germplasm of Friuli-Venezia Giulia, Italy. Genet Resour Crop Evol, 67, 83–96.

Barreneche, T.; Casasoli, M.; Russell, K.; Akkak, A.; Meddour, H.; Plomion, C.; Villani, F.; Kremer, A. (2004) Comparative mapping between Quercus and Castanea using simple-sequence repeats (SSRs). Theor. Appl. Genet. 108: 558–566. doi: 10.1007/ s00122-003-1462-2. PMID:14564395.

Bergamini, A. (1975). Osservazioni sulla morfologia fiorale di alcune cultivar di castagno. Riv. Ortoflorofruttic. Ital. 59: 103–108.

Bernetti, G. (1995) – Selvicoltura speciale. UTET – Torino: 415 pp

Blair, M.W.; Díaz, L.M.; Buendía, H. F.; Duque, M. C. (2009). Genetic diversity, seed size associations and population structure of a core collection of common beans (*Phaseolus vulgaris* L.). Theor. Appl. Genet. 119, 955–972. doi: 10.1007/s00122-009-1064-8

Botta, R.; Sartor, C.; Dini, F.; Costa, R.; Santos, C.; Quacchia, A.; ... & Torello Marinoni, D. (2012). Susceptibility to *Dryocosmus kuriphilus* Yasumatsu and molecular analyses of chestnut genotypes. In Genetics of Fagaceae and Nothofagaceae-IUFRO 2012 Conference (pp. 1-pagina). INRA.

Botta, R.; Marinoni, D.; Beccaro, G.; Akkak, A.; Bounous, G. (2001). Development of a DNA typing technique for the genetic certification of chestnut cultivars. For. Snow Landsc. Res. 76: 425–428.

Bouffartigue, C.; Debille, S.; Fabreguette, O.; Ramos Cabrer, A.; Pereira-Lorenzo, S.; Flutre, T.; Harvengt, L, (2019). Genetic diversity and structure of sweet chestnut (*Castanea sativa* Mill.) in France: At the intersection between Spain and Italy. bioRxiv 792259; doi: https://doi.org/10.1101/792259

Bouffartigue, C.; Debille, S.; Fabreguettes, O.; Cabrer, A. R.; Pereira-Lorenzo, S.; Flutre, T., & Harvengt, L. (2020). High admixture between forest and cultivated chestnut (*Castanea sativa* Mill.) in France. bioRxiv, 792259.

Bounous, G. (2002) 'Frutteto di castagno', in: Bounous G. (ed.) Il Castagno. 1st edn. Bologna: Edagricole, pp. 1-30.

Bounous, G.; De Guarda-Bounous, A. (1999). Among the chestnut trees in Cuneo Province. Edizioni Metafore, Cuneo.

Breviglieri, N. (1951). Ricerche sulla biologia fiorale e di fruttificazione della *Castanea sativa e Castanea crenata* nel territorio Vallombroso (Centro Studi sul Castagno 1). La Rivista Scientifica, 21(Suppl.): 15–50.

Brown, A.H.D. (1989) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Buck, E.J.; Hadonou, M.; James, C.J.; Blakesley, D.; Russell, K. (2003) Isolation and characterization of polymorphic microsatellites in European chestnut (*Castanea sativa Mill*.) Mol Ecol Notes 3:239–241. https://doi.org/10.1046/j.1471-8286.2003.00410.x

Caballero, A., García-Dorado, A. (2013). Allelic diversity and its implications for the rate of adaptation. Genetics 195, 1373–1384. doi: 10.1534/genetics.113. 158410/-/DC1

Cipriani, G.; Spadotto, A.; Jurman, I.; Di Gaspero, G.; Crespan, M.; Meneghetti, S.; Frare, E.; Vignani, R.; Cresti, M.; Morgante, M.; Pezzotti, M.; Pe, E.; Policriti, A.; Testolin, R. (2010) The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. Theor Appl Genet 121:15691585

Conedera, M.; Krebs, P.; Tinner, W.; Pradella, M.; & Torriani, D. (2004). The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. Vegetation History and Archaeobotany, 13(3), 161–179. https://doi.org/10.1007/s00334-004-0038-7

D'Emerico, S.; Bianco, P.; Medagli, P.; Schirone, B. (1995). Karyotype analysis in Quercus spp. (Fagaceae). Silvae Genet. 44: 66-70.

de Philippis, A. (1937) Classificazione ed indici del clima in rapporto alla vegetazione forestale italiana. Firenze: Ricci. doi: 10.1080/11263503709438750

di Trento, P. A. (2008). Il castagno e il noce nella provincia di Trento. Provincia autonoma di Trento, dipartimento agricoltura e alimentazione e servizio promozione delle attività agricole, Trento

Dice, L.R. (1945) "Measures of the amount of ecologic association between species". Ecology 26:297–302.

Dini, F.; Sartor, C.; Botta, R. (2012). Detection of a hypersensitive reaction in the chestnut hybrid 'Bouche de Bétizac' infested by *Dryocosmus kuriphilus* Yasumatsu. Plant Physiol. Bioch. ISSN 0981-9428, 60: 67-73.

Emanuelli, F.; Lorenzi, S.; Grzeskowiak, L.; Catalano, V.; Stefanini, M.; Troggio, M. (2013). Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grapes. BMC Plant Biol. 13:39. doi: 10.1186/1471-2229-13-39

Ferreira dos Santos, A.R.; Ramos-Cabrer, A.M.; Díaz-Hernández, M.B.; Pereira-Lorenzo, S. (2011) Genetic variability and diversification process in local pear cultivars from northwestern Spain using microsatellites. Tree Genet Genomes 7:1041–1056

Fedrigotti, V. B.; Fischer, C. (2015). Sustainable development options for the chestnut supply chain in South Tyrol, Italy. Agriculture and Agricultural Science Procedia, 5, 96-106.

Fideghelli, C. Atlante dei Fruttiferi Autoctoni Italiani; Crea—Centro di frutticoltura: Roma, Italy, 2016.

Fineschi, S.; Taurchini, D.; Muller-Starck, G.; Conedera, M. (1994). Genetic characterization of cultivated varieties of European chestnut (*Castanea sativa* Mill.) in Southern Switzerland. Analysis of RAPD's molecular markers. Istituto di Coltivazioni Arboree, University of Perugia, Spoleto, Italy. pp. 309–313.

Galderisi, U.; Cipollaro, M.; Di Bernardo, G.; De Masi, L.; Galano, G., & Cascino, A. (1998). Molecular typing of Italian sweet chestnut cultivars by random amplified polymorphic DNA analysis. The Journal of Horticultural Science and Biotechnology, 73(2), 259-263.

Gobbin, D.; Hohl, L.; Conza, L.; Jermini, M.; Gessler, C.; Conedera, M. (2007) Microsatellite-based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. Genome 50:1089–1103. doi: 10.1139/G07-086

Gomes-Laranjo, J., Paulo Coutinho, J. and Peixoto, F. (2008) 'Ecophysiological characterization of *C. sativa* Mill', Acta Horticulturae, 784, pp. 99-106. doi:10.17660/actahortic.2008.784.14

Goulão, L.; Oliveira, C.M. (2001). Molecular characterization of cultivars of apple (*Malus* × *domestica* Borkh.) using microsatellite (SSR and ISSR) markers. Euphytica 122, 81–89. doi: 10.1023/A:1012691814643

Gulati, R. (2018). Strategies for sustaining plant germplasm evaluation and conservation—a review. Life Sci Inform, 4, 313-320.

Heiniger, U and Conedera, M. (1992) 'Chestnut forests and chestnut cultivation in Switzerland', International Chestnut Conference. Morgantown, West Virginia (USA), 10-14 July, pp.175-178.

Hummer, K.E.; Dempewolf, H.; Bramel, P.; Markham, R.; Stover, E. (2015) Status of global strategies for horticultural fruit crops. Acta horticulturae

Huntley, B.; Birks, H. J. B. (1983). Atlas of past and present pollen maps for Europe, 0-13,000 years ago. Cambridge University Press.

Jarman, R.; Hazell, Z.; Campbell, G.; Webb, J.; & Chambers, F. M. (2019). Sweet Chestnut (*Castanea sativa* Mill.) in Britain: Re-assessment of its Status as a Roman Archaeophyte. Britannia. https://doi.org/10.1017/S0068113X19000011

Jaynes, R.A. (1972). Genetics of Chestnut. USDA Forest Serv. Res.Pap. WO-17. 13 p.

Ji, F.; Wei, W.; Liu, Y.; Wang, G.; Zhang, Q.; Xing, Y.; ... & Qin, L. (2018). Construction of a SNP-based high-density genetic map using genotyping by sequencing (GBS) and QTL analysis of nut traits in Chinese chestnut (*Castanea mollissima* Blume). Frontiers in plant science, 9, 816.

Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. (2011). Microsatellite markers: an overview of the recent progress in plants. Euphytica 177:309–334

Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol. Ecol. 16:1099–1106.

Kitahara, K., Matsumoto, S., Yamamoto, T., Soejima, J., Kimura, T., Komatsu, H., et al. (2005). Parent identification of eight apple cultivars by S-RNase analysis and simple sequence repeat markers. HortScience 40, 314–317.

Krebs, P.; Conedera, M.; Pradella, M.; Torriani, D.; Felber, M.; Tinner W. (2004) Quaternary refugia of the sweet chestnut (*Castanea sativa* Miller): an extended palynological approach. Vegetation History and Archaeobotany, 13, 145–160.

Krutovskii, K.V.; Neale, D.B. (2001). Forest genomics for conserving adaptive genetic diversity. In: Geburek T, Turok J (eds) Conservation and management of forest genetic resources in Europe. Arbora Publishers, Zvolen, pp 369–390 http://www.fao.org/DOCREP/003/X6884E/ X6884E00.html

LaBonte, N. R.; Zhao, P.; & Woeste, K. (2018). Signatures of selection in the genomes of Chinese chestnut (*Castanea mollissima* Blume): the roots of nut tree domestication. Frontiers in plant science, 9, 810.

Lassois, L.; Denancé, C.; Ravon, E.; Guyader, A.; Guisnel, R.; Hibrand-SaintOyant, L. et al. (2016). Genetic diversity, population structure, parentage analysis, and construction of core collections in the French apple germplasm based on SSR markers. Plant Mol. Biol. Rep. 34, 827–844. doi: 10.1007/s11105015-0966-7

Lavialle, G.B. (1906). Le chataignier. Vigot Freres, Paris.

Liang, W.; Dondini, L.; De Franceschi, P.; Paris, R.; Sansavini, S.; Tartarini, S. (2015) Genetic diversity, population structure and construction of a core collection of apple cultivars from Italian germplasm. Plant Mol Biol Rep 33:458–473

Marconi, G.; Ferradini, N.; Russi, L.; Concezzi, L.; Veronesi, F.; Albertini, E. (2018). Genetic characterization of the apple Germplasm collection in Central Italy: the value of local varieties. Frontiers in plant science, 9, 1460.

Marinoni, D.; Akkak, A.; Bounous, G.; Edwards, K.J.; Botta, R. (2003). Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). Mol. Breed. 11 (2), 127–136 https://doi.org/10.1023/A:1022456013692.

Marinoni, D. T.; Akkak, A.; Beltramo, C.; Guaraldo, P.; Boccacci, P.; Bounous, G.; ... & Botta, R. (2013). Genetic and morphological characterization of chestnut (*Castanea sativa* Mill.) germplasm in Piedmont (north-western Italy). Tree genetics & genomes, 9(4), 1017-1030.

Martin, M.A.; Alvarez, J.B.; Mattioni, C.; Cherubini, M.; Villani, F.; Martin, L.M. (2009) Identification and characterization of traditional chestnut varieties of southern Spain using morphological and simple sequence repeat (SSRs) markers. Ann Appl Biol 154:389–398

Martín, M.A.; Mattioni, C.; Cherubini, M.; Taurchini, D.; Villani, F. (2010) Genetic characterisation of traditional chestnut varieties in Italy using microsatellites (SSRs) markers. Ann Appl Biol 157:37–44. https://doi.org/10.1111/j.1744-7348.2010.00407.x

Martín, M.A.; Mattioni, C.; Molina, J.R.; Alvarez, J.A.; Cherubini, M.; Herrera, M.A.; Villani, F.; Martín, L.M. (2012) Landscape genetics structure of chestnut (*Castanea sativa* Mill.) in Spain. Tree Genet Genomes 8:127–136. https://doi.org/10.1007/s11295-011-0427-x

Martín, M. A.; Monedero, E.; & Martín, L. M. (2017). Genetic monitoring of traditional chestnut orchards reveals a complex genetic structure. Annals of Forest Science, 74(1). https://doi.org/10.1007/s13595-016-0610-1

Martín, M. A.; Mattioni, C.; Cherubini, M.; Villani, F.; & Martín, L. M. (2017). A comparative study of European chestnut varieties in relation to adaptive markers. Agroforestry systems, 91(1), 97-109.

Mattioni, C.; Martin, M.A.; Pollegioni, P.; Cherubini, M.; & Villani, F. (2013). Microsatellite markers reveal a strong geographical structure in European populations of *Castanea sativa* (Fagaceae): Evidence for multiple glacial refugia. American Journal of Botany, 100(5), 951–961. https://doi.org/10.3732/ajb.1200194

Mattioni, C.; Cherubini, M.; Micheli, E.; Villani, F.; Bucci, G. (2008) "Role of domestication in shaping *Castanea sativa* genetic variation in Europe". Tree Genetics & Genomes 4(3):563574.

Mattioni, C.; Martin, M. A.; Chiocchini, F.; Cherubini, M.; Gaudet, M.; Pollegioni, P.;… Villani, F. (2017). Landscape genetics structure of European sweet chestnut (*Castanea sativa* Mill): indications for conservation priorities. Tree Genetics and Genomes, 13(2). https://doi.org/10.1007/s11295-017-1123-2

McCleary, T.; McAllister, M.; Coggeshall, M.; Romero-Severson, J. (2013). EST-SSR markers reveal synonymies, homonymies and relationships inconsistent with putative pedigrees in chestnut cultivars. Genetic Resources and Crop Evolution, 60(4), 12091222. https://doi.org/10.1007/s10722-012-9912-9

Mehra, P. N.; Hans, A. S.; & Sareen, T. S. (1972). Cytomorphology of Himalayan Fagaceae. Silvae Genetica, 21, 102-109.

Mellano, M.; Beccaro, G.; Donno, D.; Marinoni, D.T.; Boccacci, P.; Canterino, S.; Cerutti, A.; Bounous, G. (2012) Castanea spp. biodiversity conservation: Collection and characterization of the genetic diversity of an endangered species. Genet. Resour. Crop Evol. 59, 1727–1741.

Miranda, C.; Urrestarazu, J.; Santesteban, L.G.; Royo, J.B.; Urbina, V. (2010) Genetic diversity and structure in a collection of ancient Spanish pear cultivars assessed by microsatellite markers. J Amer Soc Hortic Sci 135:428–437

Moriya, S.; Iwanami, H.; Yamamoto, T; Abe, K. (2011). A practical method for apple cultivar identification and parent-offspring analysis using simple sequence repeat markers. Euphytica 177, 135–150. doi: 10.1007/s10681-0100295-8

Mousavi, S.; Mariotti, R.; Regni, L.; Nasini, L.; Bufacchi, M.; Pandolfi, S. (2017). The first molecular identification of an olive collection applying standard simple sequence repeats and novel expressed sequence tag markers. Front. Plant Sci. 8:1283. doi: 10.3389/fpls.2017.01283

Nishio, S.; Terakami, S.; Matsumoto, T.; Yamamoto, T.; Takada, N.; Kato, H., ... & Saito, T. (2018). Identification of QTLs for agronomic traits in the Japanese chestnut (*Castanea crenata* Sieb. et Zucc.) breeding. The Horticulture Journal, 87(1), 43-54.

Patzak, J.; Paprštein, F.; Henychová, A.; Sedlák, J. (2012). Comparison of genetic diversity structure analyses of SSR molecular marker data within apple (Malus × domestica) genetic resources. Genome 55, 1–19. doi: 10.1139/g2012054

Peefers, J. P.; Calwey, N. W. (1988). "Germplasm Collection and Breeding Needs in Europe." Economic Botany 42: 503521.

Pereira-Lorenzo, S.; Bischofberger, Y.; Conedera, M.; Piattini, P.; Crovadore, J.; Chablais, R.; Rudow, A.; Hatt, S.; Ramos-Cabrer, A.M.; Barreneche, T. (2020) Reservoir of the European chestnut diversity in Switzerland. Biodivers. Conserv.

Pereira-Lorenzo, S.; Ramos-Cabrer, A. M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, B.; Martín, A. (2019). Instant domestication process of European chestnut cultivars. Annals of Applied Biology, 174(1), 74–85. https://doi.org/10.1111/aab.12474

Pereira-Lorenzo, S.; Ramos-Cabrer, A.M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, M.B.; Martín, L.M.; Martín, Á. (2017) Database of European chestnut cultivars and definition of a core collection using simple sequence repeats. Tree Genet. Genomes 13, 114.

Pereira-Lorenzo, S., Lourenço Costa, R.M., Ramos-Cabrer, A.M., Ciordia-Ara, M., Marques Ribeiro, C.A., Borges, O., Barreneche, T. (2011) Chestnut cultivar diversification process in the Iberian Peninsula, Canary Islands, and Azores. Genome 54:301–315. https://doi.org/10.1139/g10-122

Pereira-Lorenzo, S.; Costa, R.; Ramos-Cabrer, A.M.; Ribeiro, C.; Serra da Silva, C.; Manzano, G.; Barreneche, T. (2010) Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. Tree Genet Genomes 6:701–715. https://doi.org/10.1007/s11295-010-0285-y

Pérez-Romero, L. F.; Suárez, M. P.; Dapena, E.; Rallo, P. (2015). Molecular and morphological characterization of local apple cultivars in Southern Spain. Genet. Mol. Res. 14, 1487–1501. doi: 10.4238/2015.February.20.4

Perulli, G. D., Boini, A., Bresilla, K., Morandi, B., Grappadelli, L. C., & Manfrini, L. (2020) Growth analysis of sweet chestnut burr in two seasons with differing weather conditions. Italus Hortus, 27: 31-39. doi: 10.26353/j.itahort/2020.1.3139

Poljak, I.; Idžojtić, M.; Šatović, Z.; Ježić, M.; Ćurković-Perica, M.; Simovski, B.; … Liber, Z. (2017). Genetic diversity of the sweet chestnut (*Castanea sativa* Mill.) in Central Europe and the western part of the Balkan Peninsula and evidence of marron genotype introgression into wild populations. Tree Genetics and Genomes, 13(1). https://doi.org/10.1007/s11295-017-1107-2

Ponchia, G.; Aguanno, M.; Meneghetti, S. (2001). Osservazioni morfologiche, carpo metriche e molecolari su alcune piante di castagno del Feltrino. Atti del Convegno Nazionale Castagno 2001. Marradi (Firenze) 25-27 ottobre:103-108.

Pritchard, J.K.; Stephens, M.; Donnelly, P. (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Quintana, J.; Contreras, A.; Merino, I.; Vinuesa, A.; Orozco, G.; Ovalle, F., & Gomez, L. (2015). Genetic characterization of chestnut (*Castanea sativa* Mill.) orchards and traditional nut varieties in El Bierzo, a glacial refuge and major cultivation site in northwestern Spain. Tree Genetics and Genomes, 11(1). https://doi.org/10.1007/s11295-014-0826-x

Ramos-Cabrer, A.M.; Pereira-Lorenzo, S. (2005) Genetic relationship between *Castanea sativa* Mill. trees from North-western to South Spain based on morphological traits and isoenzymes. Genet Resour Crop Ev 52:879–890. doi:10.1007/s10722-003-6094-5

Roces-Díaz, J. V.; Jiménez-Alfaro, B.; Chytrý, M.; Díaz-Varela, E. R.; & Álvarez-Álvarez, P. (2018). Glacial refugia and mid-Holocene expansion delineate the current distribution of *Castanea sativa* in Europe. Palaeogeography, Palaeoclimatology, Palaeoecology, 491, 152–160. https://doi.org/10.1016/j.palaeo.2017.12.004

Rohlf, F.J. (1994) NTSYS-pc Numerical taxonomy and multivariate analysis system (version 1.80). State University of New York.

Sansavini, S.; Pancaldi, M.; Vinatzer, B. (1997) Il 'Fingerprinting' nelle specie da frutto: progressi nelle tecnologie molecolari e trasferibilità dei metodi nell'identificazione varietale e nella distinzione dei cloni. Relazione Convegno Agro-Bio-Fruit, Cesena.

Sartor, C.; Dini, F.; Torello Marinoni, D.; Mellano, M.G.; Beccaro, G.L.; Alma, A.; Quacchia, A.; Botta, R. (2015). Impact of the Asian wasp *Dryocosmus kuriphilus* (Yasumatsu) on cultivated chestnut: Yield loss and cultivar susceptibility. Sci. Hortic. http://dx.doi.org/10.1016/j. scienta.2015.10.004.

Sartor, C.; Torello Marinoni, D.; Beccaro, G. L.; Mellano, M. G.; Bounous, G.; Botta, R., ... & Alma, A. (2009). Evaluation of susceptibility to gall wasp in chestnut cultivars and molecular mechanisms of plant response to infestation. In Castanea 2009 (pp. 69-69). Università di Torino.

Sawano, M.; Ichii, T.; Nakanishi, T.; Kotera, Z. (1984). Studies on identification of chestnut species and varieties by isozyme analysis. Sci. Rep. Fac. Agric. Kobe Univ. 16: 67–71.

Scott, K.D.; Eggler, P.; Seaton, G.; Rossetto, M.; Ablett, E.M.; Lee, L.S.; Henry, R.J. (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723–726

Silvanini, A.; Marinoni, D. T.; Beccaro, G. L.; & Ganino, T. (2011). La caratterizzazione varietale del germoplasma di Castanea sativa Mill. Italus Hortus, 18, 47-61.

Staton, M.; Addo-Quaye, C.; Cannon, N.; Sun, Y.; Zhebentyayeva, T.; Huff, M.; ... & Henry, N. (2019). The Chinese chestnut genome: a reference for species restoration. bioRxiv, 615047.

Torello Marinoni, D.; Akkak, A.; Beltramo, C.; Guaraldo, P.; Boccacci, P.; Bounous, G.; Botta, R. (2013). Genetic and morphological characterization of chestnut (*Castanea sativa* Mill.) germplasm in Piedmont (north-western Italy). Tree Genetics and Genomes, 9(4), 1017–1030. https://doi.org/10.1007/s11295-013-0613-0

Urrestarazu, J.; Denancé, C.; Ravon, E.; Guyader, A.; Guisnel, R.; Feugey, L.; Poncet, C.; Lateur, M.; Houben, P.; Ordidge, M.; Fernandez-Fernandez, F.; Evans, K.M.; Paprstein, F.; Sedlak, J.; Nybom, H.; Garkava-Gustavsson, L.; Miranda, C.; Gassmann, J.; Kellerhalls, M.; Suprun, I.; Pikunova, A.V.; Krasova, N.G.; Torutaeva, E.; Dondini. L.; Tartarini, S.; Laurens, F.; Durel C.E. (2016) Analysis of the genetic diversity and structure across a wide range of germplasm reveals prominent gene flow in apple at the European level. BMC Plant Biology 16:130

Urrestarazu, J.; Miranda, C.; Santesteban, L.G.; Royo, J.B. (2012) Genetic diversity and structure of local apple cultivars from Northeastern Spain assessed by microsatellite markers. Tree Genet Genomes 8: 1163–1180

Valle, G. (1959). Osservazioni sulla morfologia dei rami di un anno in abito invernale in alcune cultivar di *Castanea sativa* Mill. Centro di studio sul castagno, Pubblicazione, 5: 5–94.

van Hintum, T.J.L.; Brown, A.H.D.; Spillane, C.; Hodgkin, T. (2000) Core collections of plant genetic resources IPGRI. (Technical Bulletin No 3). International Plant Genetic Resources Institute, Rome, Italy

Varshney, R. K.; Graner, A.; & Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. TRENDS in Biotechnology, 23(1), 48-55.

Vigiani, D. (1908). Le varietà di castagno ed i criteri da seguirsi per classificarle. Atti. R. Accad. Georgofili, 16: 38–73.

Volk, G.M.; Richards, C.M.; Henk, A.D.; Reilley, A.A.; Bassi, N.V.; Postman, J.D. (2006) Diversity of wild *Pyrus communis* based on microsatellite analyses. J Am Soc Hortic Sci 131:408–417

Wang, M. L.; Sukumaran, S.; Barkley, N.A.; Chen, Z.; Chen, C.Y.; Guo, B.; Yu, J. (2011). Population structure and marker–trait association analysis of the US peanut (*Arachis hypogaea* L.) mini-core collection. Theoretical and applied genetics, 123(8), 1307-1317.

Warschefsky, E.; Penmetsa, R.V.; Cook, D.R.; von Wettberg, E.J. (2014). Back to the wilds: Tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. Am. J. Bot. 101:1791–1800. doi:10.3732/ajb.1400116

Wilson, E. O., The Diversity of Life, The Belknap Press of Harvard University Press, Cambridge, MA, 1992, p. 393

Williams, J. T. (1991). Plant Genetic Resources: Some New Directions. Advances in Agronomy 45: 61-91.

Yatabe, Y.; Kane, N. C.; Scotti-Saintagne, C.; & Rieseberg, L. H. (2007). Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, Helianthus annuus and H. petiolaris. Genetics, 175(4), 1883-1893.

Yonezawa, K. (1995). Sampling strategies for use in stratified germplasm collections. Core collections of plant genetic resources.

Yun, W.H.; Ban, S.H.; Kim, G.H.; Kim, J.H.; Kwon, S.I.; Choi, C. (2015). Assessment of apple core collections constructed using phenotypic and genotypic data. Genet. Mol. Res. 14, 6453–6464. doi: 10.4238/2015.June.11.21

Zhang, P.; Li, J.; Li, X.; Liu, X.; Zhao, X.; Lu, Y. (2011). Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR markers. PLoS One 6:e27565. doi: 10.1371/journal.pone.0027565

CHAPTER 2: Genetic diversity of Castanea sativa Mill. accessions from Tuscan-Emilian Apennines and Emilia Romagna region (Italy)

2.1 Abstract

This work investigated the genetic diversity of 134 *Castanea sativa* Mill. accessions present in Emilia-Romagna. Samples were taken from three collection fields (Granaglione, Zocca and Paloneta) in the Tuscan-Emilian Apennines. The accessions were analyzed by using 16 microsatellite markers (SSR). Genetic distances among accessions, calculated through the DICE coefficient, were used to construct an UPGMA cluster analysis. One major genotype (named "Marroni") was identified across the three investigated collection fields; this variety corresponds to a sweet chestnut cultivar that has been propagated and widely diffused in the Emilia-Romagna region. Other genotypes were represented by different varieties of Italian chestnuts. The results of this study will be used to define and share guidelines for the characterization and varietal certification of the chestnut varieties in the Emilia-Romagna region.

Keywords: Chestnut, genetic diversity, local germplasm, SSR, cluster analysis

2.2 Introduction

To date, the natural distribution area of the European chestnut (*Castanea sativa* Mill.) mainly includes southern Europe and southwestern Asia. In particular, the European distribution area extends from the northwestern part of the Iberian Peninsula to Caucasia and the Caspian Sea (Conedera et al., 2016).

According to palaeobotanic data, the current biodiversity of the chestnut tree originates from glacial refugia located in Transcausasia and in the Italian and Iberian peninsulas, where chestnut trees probably found a favorable habitat. During the Holocene, chestnut trees spread to the surrounding areas as a result of post-glacial climate conditions and human activities (Krebs et al., 2019).

The first unambiguous evidence of chestnut cultivation was reported in the Middle East and Greece and dates back to about 4000 B.C., although chestnut use was reported during the Neolithic (6000 BP)(

Kaltenrieder et al.,). Subsequently, in the Greek and pre-Christian world, chestnut tree cultivation was a minor activity.

The role of the chestnut changed at the beginning of the Christian era, when the versatility of this tree was better understood. In Italy, and thereafter in Europe, chestnut cultivation might have been introduced by the Romans, although there is no clear evidence of systematic tree planting in the Italian territory (Conedera et al., 2004).

During the Middle Ages, the cultivation of chestnut in the Italian Apennines intensified thanks to Matilde di Canossa around the year 1110 A.D. To render the territory self-sufficient, she strongly encouraged the cultivation of chestnut in this region, which is why many old and monumental trees in this area are named "Matildici" (Conedera et al., 2004).

The increase of chestnut cultivation led to the birth of the idea of tree/fruit selection: productivity, size and flavor of the fruit started to be taken into account (Bassi and Marangoni, 1984). An example of such development is given by the Marroni genotype.

In the 16th century, in a region between Tuscany and Emilia-Romagna, a cultivar called ‘Marrone Fiorentino’ was selected and propagated throughout different regions of central and northern Italy (Bassi and Marangoni, 1984; Borghetti et al., 1983; Breviglieri, 1955).

The Marroni genotype (or sweet chestnut) was selected for its excellent characteristics: i) weight of the fruit above average (maximum 70 fruits per kg); ii) one to three fruits per burr; iii) monoembryonic nuts; iv) epicarp of bright light color, marked with accentuated grooves of darker coloring; v) thin and easy-to-remove episperm (cuticle), not deep in the cotyledons; vi) floury paste, sugary, consistent, resistant to cooking without breaking up (Breviglieri, 1955). Another feature among the Marroni accessions is that they are androsterile.

The genetic uniformity among Marroni group accessions is the result of clonal propagation carried out by growers to maintain the desired characteristics (Martín et al., 2010; Mellano et al., 2012). Subsequently, the Marroni genotype was planted in various areas, where it was given different names, such as Marrone di Castel del Rio, Marrone di Zocca, Marrone Buono of Marradi, Marrone Biondo di Monghidoro and others (Gallesio, 1817; Breviglieri, 1955; Borghetti et al., 1983; Bassi and Marangoni, 1984).

Other than the Marroni genotype, in the Tuscan-Emilian Apennines, the other dominant varieties of chestnut, which are mainly used for the production of flour and other derivatives, are the following: 'Carpinese' or 'Carrarese', 'Pastanese', 'Pistolese', 'Piusela', 'Molana', 'Ceppa' and 'Loiola' (Bassi and Marangoni, 1984; Antonaroli et al., 1984; Bagnaresi et al., 1977). These chestnut cultivars are characterized by variable fruit weight (each cultivar presenting a specific weight range), polyembrionc nuts with an adherent and intrusive episperm and lower fruit sweetness compared to the Marroni group (Breviglieri, 1955; Bagnaresi et al., 1977).

To date, considering Italy as a whole, chestnut trees are mainly present in six regions of the country (Campania, Lazio, Tuscany, Emilia-Romagna, Piedmont, Veneto). Italian varieties are characterized by a wide genetic variability resulting from a tradition of multiplying the varieties by seed. This tradition contributed both to a high number of native ecotypes throughout the country and to the subsequent selection of cultivars that, over time, have adapted to different areas. Each has specific characteristics that are regulated by the PGI (Protected Geographical Indication) issued by the European Union (Neri et al., 2010; Fideghelli, 2016). In particular, Marrone of 'Castel del Rio' has been awarded PGI certification, being one of the most valuable and known chestnut cultivars in Italy and abroad, originating from the Emilia Romagna region.

The highest number of varieties is cultivated in Tuscany (26.9%), followed by Piedmont (15.2%), Campania (12.8%), Emilia Romagna (8.8%) and Calabria (7.5%; Fideghelli, 2016).

Nowadays, there are many challenges that threaten chestnut production in the Tuscan-Emilian Apennines, e.g. the diffusion of pathogens and pests such as the *Gnomoniopsis ascoe* fungus and the Oriental chestnut gall wasp, *Dryocosmus kuriphilus* (Lucchi et al., 2016). In addition, there are socio-economic problems related to the market and to a rapidly changing environment (Pezzi et al., 2011).

Chestnuts should be also considered for their phytochemical and nutritional composition (Beccaro et al., 2020). This research confirms the presence of very important molecules for human nutrition (monoterpenes, polyphenols and vitamin C) related to many biological properties, such as anticancer, anti-atherosclerotic, anti-inflammatory, antihepatotoxic, and antioxidant activities (Landete, 2011). In addition, the higher level of fructose found in fruits define chestnuts as a potential functional food for persons suffering from diabetes type 2 (Beccaro et al., 2019).

These elements encourage the in-depth study of the chestnut tree and the enlargement of local germplasms to preserve the existing biodiversity and eventually identify desirable traits, such as resistance to pests or features that could be potentially useful to the Italian chestnut industry.

The identification of redundant accessions (identical genetic profile but with a different name) represents a fundamental preliminary step to undertake a genetic characterization of the germplasm, since most of the accessions have been found in the fields and initially identified with their local names (Pereira-Lorenzo et al., 2010; Cipriani et al., 2010).

In such cases, it is necessary to support a further phenotypic analysis, using pomological charts to verify the presence of a true state of synonymy, if known, or to identify different phenotypes probably due to point genetic mutations, structural genome changes or epigenetics (Cipriani et al., 2010).

Molecular markers, such as Simple Sequence Repeats (SSRs) or microsatellites, can support pomological analyses and have been used for genetic diversity analysis (and for structure analysis) in several fruit tree species (Cipriani et al, 2010 in grapevine; Liang et al., 2015 and Urrestarazu et al., 2016 in apple; Bhattarai and Mehlenbacher , 2017 in hazelnut, Ferradini et al. 2017 and Baccichet et al. 2020 in pear). The related datasets have provided a useful support for varietal identification. The same approach can be used for the analysis of the genetic diversity of chestnuts as well. This approach was used to characterize germplasm collections (Martín et al 2010; Mellano et al. 2012; Pereira-Lorenzo et al., 2020) and to describe the existing relationships among Italian and European varieties (Beghè et al. 2013; Quintana et al. 2015; Pereira-Lorenzo et al. 2017 and 2019; Bouffartigue et al., 2019). The use of SSRs allowed the identification and characterization of traditional varieties from southern Spain (Martín et al., 2009 and 2010). These studies could be used as a model in order to extend the analysis to other regional germplasms in Italy and Europe (Pereira-Lorenzo et al., 2017 and 2020; Bouffartigue et al., 2019), to characterize the collections and to provide tools for varietal certification.

Currently, the characterization of chestnut biodiversity in Emilia Romagna has been mainly performed by means of pomological and morphological analyses. The genetic information available is still limited. These morphological descriptions are available in the regional repertoire of the varieties at risk of genetic erosion.

Therefore, the main objectives of this study are a) to describe the biodiversity of the existing ecotypes and to preserve the existing chestnut heritage from further genetic erosion; b) to provide a genetic database of the main cultivars in Emilia Romagna for traceability and conservation purposes.

2.3 Materials and methods

2.3.1 The origin of the biological material

A panel of 134 accessions were collected in the area of the Tuscan-Emilian Apennines. In particular the sampling was carried out in three collections fields: Didactic and Experimental Park of Granaglione, the Collection of Zocca and Paloneta (created in Emilia Romagna by the University of Florence). These fields are characterized by the presence of several grafted replicates of varieties known only at phenotypic level so far (Table 2.1).

Table 2.1: list of the 134 varieties sampled in tree collection camps of the Emilia Romagna region. *in bold the hybrid sample from *C. sativa* x *C. crenata*.

Name Variety	Number of analyzed accessions	Collection camp
Biancherina	1	Zocca
Bovalghe	2	Granaglione e Zocca
Calarese	2	Zocca
Caprarola	1	Brisighella
Carrarese	2	Zocca
Castel del Rio	7	Granaglione
Castione	5	Granaglione
Centa di S. Nicolò	2	Granaglione
Ceppa	5	Granaglione e Zocca
Chiusa Pesio	2	Brisighella
Città di Castello	2	Brisighella
Drena	3	Granaglione
Gaggio Montano	2	Brisighella
Garfagnina	2	Zocca
Gavignano	2	Brisighella
Lisanese	11	Granaglione e Zocca
Locale Paloneta	4	Brisighella
Loglia	1	Brisighella
Loiola	2	Zocca

Madonna	4	Zocca
Marrone di Marradi	1	Brisighella
Marrone dell'Isola d'Elba	1	Brisighella
Mascherina	2	Zocca
Massangaia	2	Zocca
Matildici	4	Granaglione
Molana	2	Zocca
Montemarano	3	Brisighella
Monzone	2	Brisighella
Napoletana	1	Zocca
Palazzo del Pero	1	Brisighella
Pastanese	8	Granaglione e Zocca
Pastonese	3	Granaglione e Zocca
Pelosa	4	Granaglione
Pistolese	1	Zocca
Pitigliano	3	Brisighella
Piusela	2	Zocca
Precoce Migoule*	1	Zocca
Riggiolana	2	Brisighella
Roccamonfina	1	Brisighella
Roncegno	2	Granaglione
Sborgà	6	Granaglione e Zocca
Svizzera	4	Granaglione e Zocca
Tempurina	1	Brisighella
Tosca	4	Zocca
Zocca	11	Granaglione e Zocca

2.3.2 Molecular and genetic diversity analyses

For each accession, genomic DNA was extracted from 50 mg of young freeze-dried leaves following the standard CTAB protocol (Maguire et al., 1994). Genomic DNA was quantified using the Nanodrop™ ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and diluted to 10 ng/µL.

The PCR reactions were performed with the thermal cycler 2700 GeneAmp PCR System (ABI Prism) and carried out with 9 µL of master mix and 1 µL of DNA template. The PCR reactions followed this amplification protocol: an initial denaturation step of 10 min at 95°C, followed by 35 cycles for 30" at

95°C, and 30" at specific annealing temperature (Table 2.2), and 30" at 72°C, with a final extension step of 7' at 72°C.

Table 2.2: SSRs markers used in amplifications (Pereira-Lorenzo et al., 2017). LG: Linkage group; FAM; VIC; NED; PET indication of the corresponding dye (fluorochromes used in PCR analysis) and multiplex used.

Locus	Repetition motif	LG	Annealing Temperature (°C)	Multiplex	Allele size range (bp)
CsCAT1	$(TG)_5TA(TG)_{24}$	8	50	2 VIC	186-258
CsCAT2	$(AG)_{16}$	10	55	4 VIC	124 – 148
CsCAT3	$(AG)_{20}$	12	50	2 NED	138 – 198
CsCAT6	$(AC_{)24}$	1	50	2 FAM	174 – 225
CsCAT8	$(GT)_7(GA)_{20}$	6	50	5 VIC	186 – 268
CsCAT14	$(CA)_{22}$	2	58	4 NED	145 – 167
CsCAT15	$(TC)_{12}$	8	50	6 FAM	79 – 93
CsCAT16	$(TC)_{20}$	6	50	1 NED	232 – 276
CsCAT17	$(CA)_{19}A(CA)_2AA(CA)$	2	58	4 PET	152 – 166
CsCAT41B	$(AG)_{20}$	8	50	1 VIC	196 – 258
EMCs2	$(GGC)_7$	6	55	3 NED	128 – 150
EMCs15	$(CAC)_9$	9	55	3 VIC	133-161
EMCs22	$(GA)_{19}$	2	60	4 VIC	132-163
EMCs38	$(GA)_{31}$	4	56	3 PET	122 – 158
OAL	$(CT)_{16}AGT(CT)_2$	OAL	60	Single PCR VIC	178-227
QrZAG96	$(TC)_{20}$	10	55	Single PCR FAM	291 – 332

Preliminary phases of genetic characterization focused on the estimation of genetic diversity and on the determination of genetic relationships within the studied germplasm. Molecular markers (SSR) allowed to create a fingerprint for each single variety.

The microsatellites used were selected by the series (CsCAT and EMCs) and OAL elaborated on the chestnut (Marinoni et al., 2003; Buck et al., 2003 and Gobbin et al., 2007) and QrZAG developed from *Quercus robur* (Kampfer et al., 1998).

In order to characterize regional varieties, the samples were amplified by 16 pairs of labeled primers which were found to be the most polymorphic. The primers were used by multiplex set according to Pereira-Lorenzo et al., 2017(Table 2.1.2). In order to estimate the size of DNA fragments, the samples were aligned with the European dataset (Pereira- Lorenzo et al., 2017).

2.3.3 Genetic and Cluster analysis

The number of alleles per locus (k), the expected (He) and the observed heterozygosities (Ho) and the polymorphism information content (PIC) of the unique genotypes were estimated using the CERVUS Software Version 3.0.3 (Kalinowski et al., 2007). A PIC value greater than 0.7 was considered to be highly polymorphic and informative for a certain locus. A Parentage analysis on unique diploid genotypes with the CERVUS software (Marshall et al., 1998; Kalinowski et al 2007) was carried out. Two criteria were considered to establish parental relationships: a LOD confidence interval and the Delta value with a threshold of 95%.

Using all the obtained data, a cluster analysis was carried out with the construction of the dendrogram relative to genetic distances, elaborated using the Unweighted Pair-Group Method (UPGMA). The genetic distance between the cultivars was calculated using the DICE coefficient (Dice, 1945) with the SimQual NTSYSpc 2.0 (Rohlf, 1994).

To have further confirmation on the genetic similarities previously observed with the cluster analysis, the R software (Project for Statistical Computing, version 3.2.2, ,Copyright (C) 1989, 1991 Free Software Foundation, Inc. 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA) was used to perform the principal component analysis (PCA) on the 21 unique genotypes identified.

2.4 Results

2.4.1 SSR Profiles

The 16 selected molecular markers allowed the analysis of the genetic diversity and provided useful support for the direct analysis of varietal identification. In general, allele frequencies were not uniformly distributed within the investigated loci. The unique genotypes identified showed frequencies ranging from very low (as for the EMCs2 locus) to very high for the CSCAT3 locus, with 3 and 16 alleles, respectively. The 16 SSRs used in this study revealed a total of 132 alleles, with an average of 8.2 alleles per locus. Comparing the size of the DNA fragment with the chestnut EU dataset (Pereira-Lorenzo et al., 2017), 6 unique alleles were found in 21 unique genotypes: CSCAT 16 – 128; CSCAT3 – 227 and 257; CSCAT1 – 179, QrZAG96 – 163 and EMCs38 – 234 (in bold in Table 2.3).

The absence of amplification of EMCs38 on one genotype (‘Madonna, Table 2.3) may be due to the presence of null alleles. For this reason, the ‘Madonna’ genotype was not considered in the heterozygosity analysis, which was carried out with 20 unique varieties.

CSCAT3 and EMCs38 with a PIC value around 0.885 and 0.801 appeared to be the highest informative loci. Conversely, OAL and EMCs15 with a PIC value=0.300 and 0.473 respectively are the least informative.

The high value of expected heterozygosity directly reflects the high level of genetic diversity present in chestnut trees derived from cross-pollination: the value of observed heterozygosity ranged between 0.350 for OAL to 1 for CsCAT14, whereas the expected heterozygosity ranged between 0.319 for OAL to 0.917 for CsCAT3 (Table 2.4).

Table 2.3: Allelic profiles of 21 varieties (prime name, synonyms and number of accessions analyzed) from the Emilia Romagna region for 16 SSR (-1 for missing value); *in bold unique allele.

Varieties	N° accession	Synonyms	Cscat 41b	Cscat 41b	Cccat 16	Cscat 16	Cscat 6	Cscat 6	Cscat 1	Cscat 1	Cscat 3	Cscat 3	QrZag 96	QrZag 96	EMCs15	EMCs15	EMCs38	EMCs38	EMCs2	EMCs2	EMCs22	EMCs22	CsCAT2	CsCAT2	CsCAT17	CsCAT17	CsCAT14	CsCAT14	CsCAT15	CsCAT15	CsCAT8	CsCAT8	OAL	OAL
Marrone Fiorentino	66	Castel del Rio, Castione, Centa di S. Nicolò, Chiusa Pesio, Città di castello, Drena, Gaggio Montano, Gavignano, Locale Paloneta, M. di Marradi, M. Isola d'Elba, Montemarano, Monzone, Napoletana, Palazzo del Pero, Pastonese, Pitigliano, Riggiolana, Roccamonfina, Roncegno, Sborgà, Tempurina e M. di Zocca	228	234	126	132	159	173	215	223	225	239	153	155	91	91	240	244	160	160	132	134	227	227	149	155	133	150	124	134	203	208	297	309
Biancherina	1		212	216	**128***	141	179	194	194	208	223	223	153	165	91	91	272	272	160	163	134	147	209	217	141	141	141	150	134	134	189	208	297	297
Bovalghe	2		212	212	141	141	159	177	194	194	223	253	161	165	85	91	240	244	163	166	134	145	209	233	141	141	150	161	124	134	201	208	297	297
Calarese	2		212	216	141	148	159	179	194	208	237	260	153	161	91	91	244	258	163	163	134	145	209	233	149	163	133	161	124	134	201	208	297	305
Carrarese	2		212	212	130	148	192	192	194	194	**227***	239	161	165	85	91	**234***	258	160	163	134	145	231	233	141	163	141	150	134	134	189	189	297	297
Ceppa	5		212	216	141	141	165	194	194	223	**227**	237	153	165	91	91	**234**	**234**	166	166	128	134	209	229	157	163	133	141	124	134	189	201	297	301
Lisanese	11		212	212	130	143	179	179	194	194	239	251	153	161	91	91	258	272	163	166	132	147	219	227	139	163	141	152	124	134	189	208	297	297
Loglia	2		212	226	**128**	141	177	194	194	219	235	249	**163***	**163**	82	91	258	272	160	166	128	147	211	215	141	149	141	150	124	124	203	212	297	297
Loiola	2		212	216	**128**	141	177	192	194	208	223	223	153	165	91	91	248	272	160	163	134	147	209	217	145	149	141	150	134	134	189	189	297	297
Madonna	2		235	235	130	143	159	184	206	223	223	231	153	165	91	91	-1	-1	160	163	132	145	217	217	141	149	141	150	124	134	199	212	301	309
Mascherina	2		223	233	141	143	192	194	217	223	223	223	153	**163**	85	91	258	258	160	160	132	147	211	211	145	149	141	150	124	155	189	201	297	303
Massangaia	2		212	220	141	148	177	179	194	194	231	239	153	159	82	91	238	244	163	166	145	147	209	209	141	141	133	161	130	134	208	212	297	303
Molana	2		216	216	130	145	159	177	215	219	223	231	161	161	88	91	232	242	166	166	132	147	209	229	141	149	133	141	124	134	203	212	297	297
MontemaranoM20	3		216	220	**128**	141	159	177	194	219	208	231	153	153	88	91	242	246	160	166	134	134	209	219	141	149	133	150	130	134	199	201	297	297
Pastanese	11	Pastonese, Matildico	212	216	143	148	159	194	194	208	239	**257***	153	165	85	91	234	258	163	166	134	134	209	229	157	163	133	150	134	134	189	201	297	305
Pelosa	4		212	216	143	148	179	192	**179***	194	237	260	153	161	91	91	258	258	163	166	134	145	209	229	141	163	133	161	124	134	201	208	297	297
Pistolese	1		212	220	130	148	179	192	194	208	239	255	153	153	85	91	244	262	160	163	147	147	209	213	157	159	141	150	134	158	189	212	297	297
Piusella	2		212	216	132	145	177	192	194	208	247	255	153	155	85	91	258	258	163	166	128	134	209	209	157	163	150	150	134	134	189	208	297	297
PrecoceMigoule	2		228	228	145	145	138	192	188	194	202	247	153	159	79	82	244	244	163	166	128	136	196	211	139	141	135	141	124	134	178	201	297	301
Svizzera	4		216	228	130	148	159	192	208	223	239	260	153	161	85	91	234	244	160	163	132	134	217	229	149	149	133	150	134	134	201	208	297	297
Tosca	6	Garfagnina	212	226	141	148	177	192	194	194	231	253	159	165	91	91	258	258	163	163	134	134	209	211	145	163	133	141	124	134	208	212	297	301

Table 2.4: The number of individuals (N), the number of alleles (k), the observed (Ho) and expected (He) heterozygosity and the polymorphic information content (PIC) are reported for each SSR locus in *C. sativa* accessions.

Locus	k	N	Ho	He	PIC
CsCAT41	10	20	0.750	0.763	0.710
CsCAT16	9	20	0.900	0.840	0.795
CsCAT6	8	20	0.900	0.823	0.773
CsCAT1	9	20	0.750	0.710	0.662
CsCAT3	16	20	0.850	0.917	0.885
QrZAG96	6	20	0.750	0.731	0.674
EMCs15	6	20	0.600	0.521	0.473
EMCs38	13	20	0.750	0.842	0.801
EMCs2	3	20	0.750	0.668	0.577
EMCs22	7	20	0.750	0.760	0.703
CsCAT2	11	20	0.800	0.831	0.794
CsCAT17	8	20	0.750	0.827	0.780
CsCAT14	7	20	1.000	0.787	0.730
CsCAT15	5	20	0.650	0.573	0.499
CsCAT8	9	20	0.900	0.827	0.779
OAL	5	20	0.350	0.319	0.300

2.4.2 Cluster analysis

The dendrogram derived from the analysis of the molecular profiles allowed the identification of the similarities and / or identity among the studied samples (134 accessions in total; Table 2.1), highlighting, in particular, the distinction between the varieties of sweet chestnut (Cluster 1) and chestnut (Cluster 2; Figure 2.1).

Cluster 1 included 66 accessions of sweet chestnut with a uniform molecular profile even if the samples had been classified with different names, confirming synonymy among the Marroni group: 'Caprarola', 'Castel del Rio', 'Castione', 'Centa di S. Nicolò', 'Chiusa Pesio', 'Città di Castello', 'Drena', 'Gaggio Montano', 'Gavignano', 'Locale di Paloneta', 'Marron Buono di Marradi', 'Marrone dell'Isola d'Elba', 'Montemarano', 'Monzone', 'Napoletana', 'Palazzo del Pero', 'Pitigliano', 'Roccamonfina', 'Riggiolana', 'Roncegno', 'Sborgà', 'Tempurina' and 'Zocca'. Our results therefore indicated that the Marroni group is represented by a single genotype named 'Marrone Fiorentino' described in the EU chestnut dataset (Pereira-Lorenzo et al., 2017). This cluster also included accessions of Marroni called 'Pastonese', which should not be confused with the 'Pastanese' chestnut variety, as well as a Marroni accession called 'Madonna' that differs from the 'Madonna' chestnut variety. This was also observed for the accession known as 'Montemarano'. In addition, an old 'Matildico' tree was found in the Marroni group.

Conversely, Cluster 2 showed higher variability, forming numerous sub-clusters. 20 different chestnut genotypes were identified in a total of 68 accessions (Figure 2.2; Table 2.3). The dendrogram showed solid sub-clusters of accessions labelled: 'Lisanese', 'Pastonese', 'Mascherina', 'Calarese', 'Pelosa', 'Svizzera', 'Ceppa', 'Carrasere', 'Bovalghe', 'Massangaia', 'Piusela', 'Loglia', 'Molana' and 'Tosca'. This indicated a good propagation of the chestnut varieties in the Tuscan-Emilian Apennines area (Figure 2.2). These genotypes were separated in the dendrogram from the chestnut cultivars from southern Italy, such as 'Montemarano' (mainly cultivated in the Campania region).

As shown in Figure 2.1, the 'Precoce Migoule' variety, a hybrid cultivar deriving from *Castanea sativa* x *Castanea crenata* (Fideghelli, 2016), turned out to be very distant from the local chestnut cultivars, as most of the informative loci have different alleles (dataset in Table 2.3).

The dendrogram for Cluster 2 also revealed the presence of synonymous accessions (identical SSR profile but different cultivar name) such as 'Garfagnina' and ‘Tosca'. Furthermore, the ‘Pastanese’ accessions were grouped together with ‘Pastonese’ accessions and several ‘Matildico’ trees (4-8-15).

Occasional misnomers have been found by SSR analyses, such as an accession called ‘Garfagnina’ in the group of the ‘Carrarese’ cultivar, the accession named ‘Z21’ in the ‘Tosca’ genotype group and an accession named ‘Pastonese’ with an allelic profile identical to ‘Precoce Migoule’.

In conclusion, the 134 accessions analyzed showed 21 different genotypes representative of the Emilia-Romagna biodiversity (Table 2.3), with a clear separation between the Marroni group accessions (Cluster 1) and all chestnut varieties from central and southern Italy (Cluster 2). In addition, a Principal Coordinate Analysis (PCoA) on the 21 previously identified unique varieties was conducted with the R software. Figure 2.2 shows that the 'Precoce Migoule' varieties, a hybrid cultivar, differ considerably from the varieties present in the Tuscan-Emilian Apennines, which formed a small cluster. Furthermore, the 'Madonna' genotypes and Marroni group were found to be more similar to each other but separated from all the other chestnut varieties.

A parentage analysis was carried out by CERVUS and was performed excluding the locus EMCs38, which may be present null alleles. The parentage analysis did not reveal possible parental relationships (data not shown).

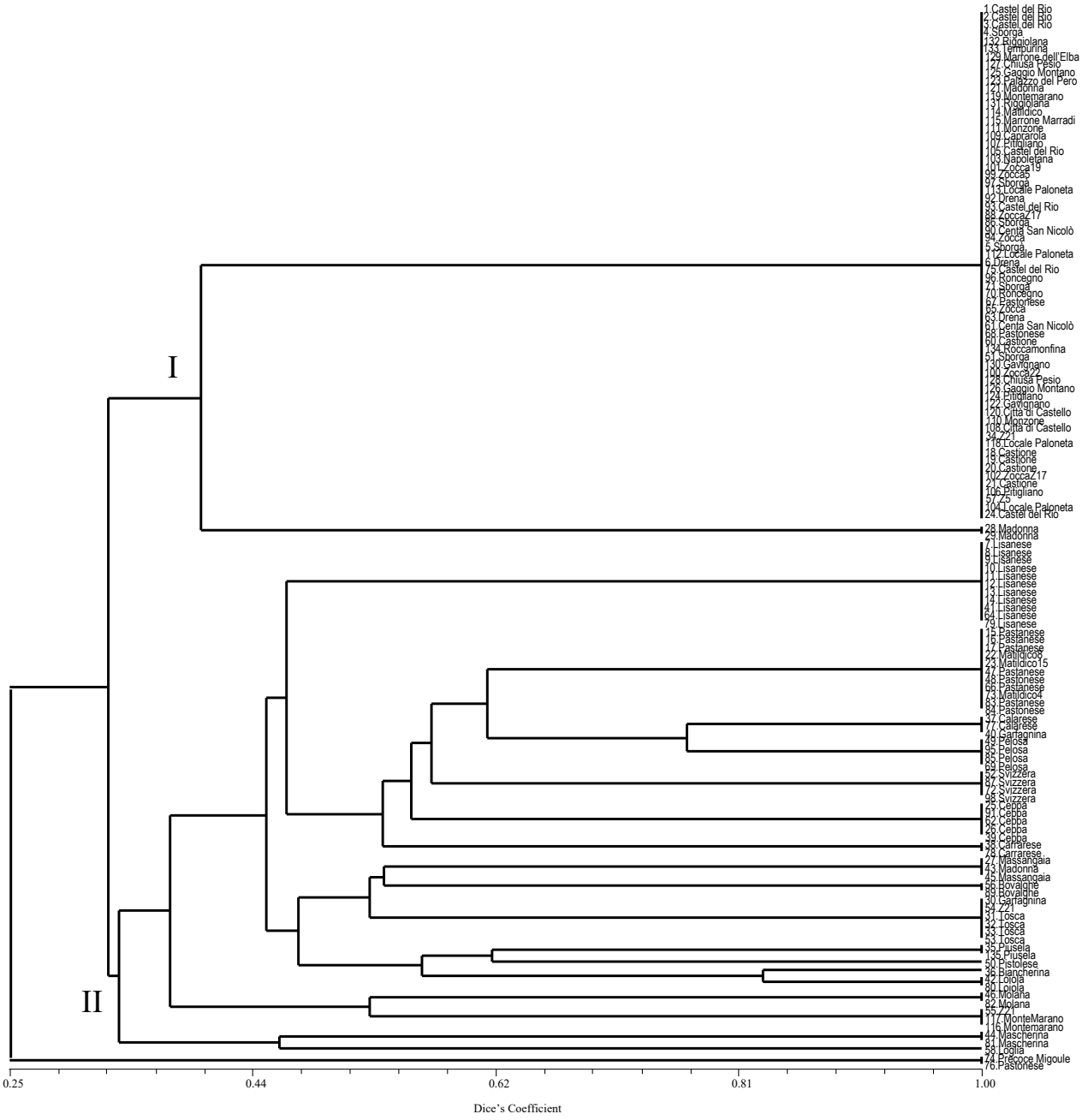

Figure 2.1: Dendrogram based on 16 SSRs using DICE coefficient demonstrating genetic characterization and relationship of 134 chestnut accessions from the Tuscan-Emilian Apennines, divided in two clusters.

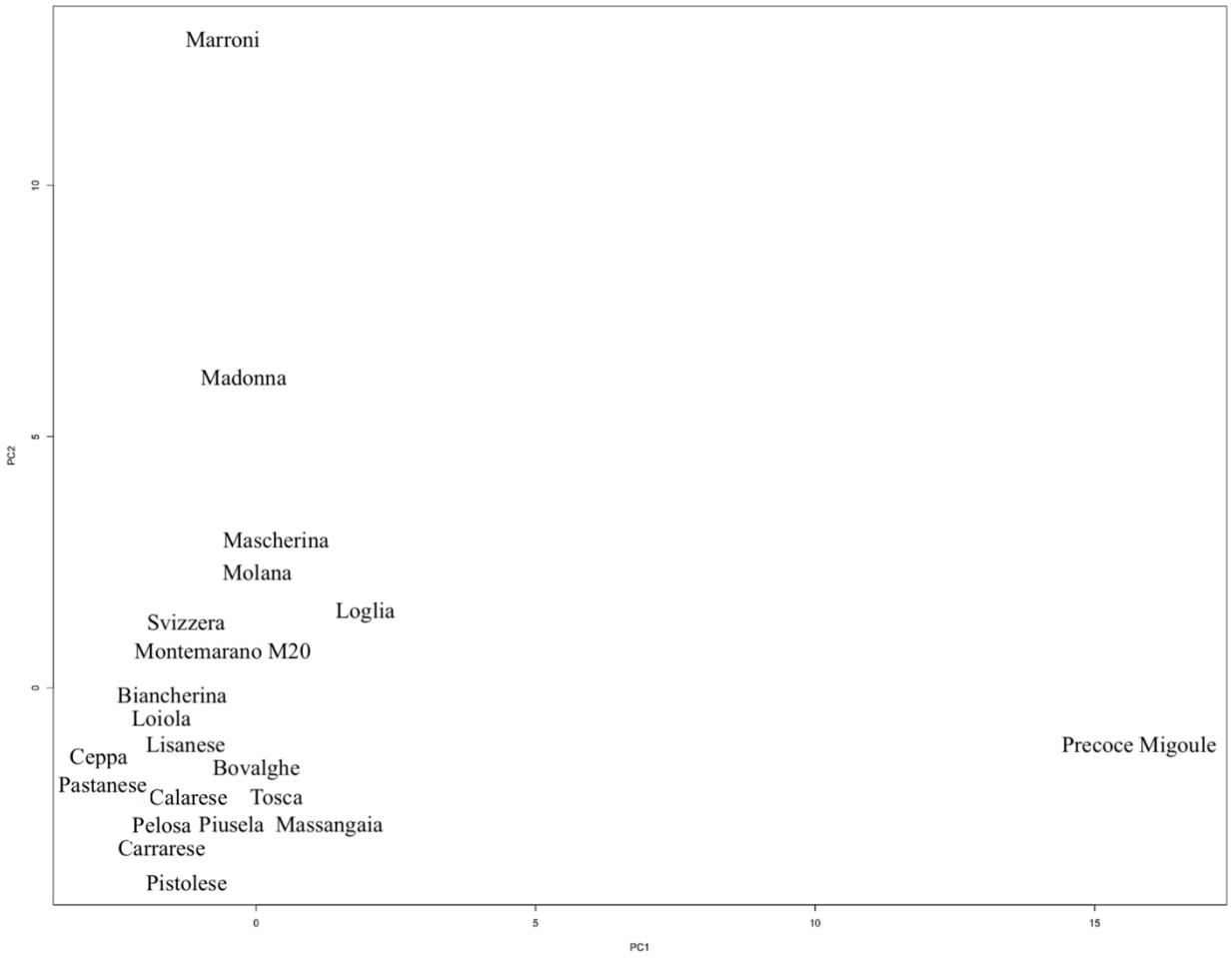

Figure 2.2: Principal Coordinate Analysis (PCoA) of the 21 chestnut unique genotypes based on the 16 SSR data. The first component explains 15% of the variation and the second component 11%.

2.5 Discussion

In this study, we performed the molecular characterization of a collection of 134 grafted chestnut and sweet chestnut (Marroni group) accessions from different collections in the Emilia-Romagna region, which corresponded to 21 representative varieties. The relatively high number of accessions of the dataset (with varieties that are well distributed in the regional territory and also include commercially used varieties) provided a good overview of the distribution of grafted chestnut varieties in the region. The set of SSRs used in this study was chosen mainly on the basis of their distribution throughout the chestnut genome, in

order to reach a high value of genomic coverage to estimate the population's genetic diversity. This marker set was also used in the genetic diversity study of Spanish chestnut germplasm described by Pereira Lorenzo et al., 2017. A work by Urrestarazu et al., 2015 studied the variations in the results of genetic diversity analysis in relation to the number of markers used. This work identified that 15-16 is the ideal number of markers for this type of analysis and asserted that a higher number of markers does not positively influence the statistical stability of the results.

The present study was based on a molecular analysis using 16 specific SSRs. Their high variability made it possible to amplify and visualize numerous alleles (the mean of 8.2 alleles). The high degree of polymorphism and high discriminating power among the analyzed samples was expected for a cross-pollination species, such as *C. sativa*.

The presence of unique alleles was found in five SSRs tested. This evidences a relevant genetic diversity among the *C. sativa* species due to the high discriminant power of the molecular marker set used. Our molecular marker set was picked with the intent of creating an effective varietal identification tool for future use, as many other crops have.

In particular, the CsCAT3 primer was found to be one of the most discriminating loci (PIC -0.885), as already confirmed by other studies (Pereira- Lorenzo et al., 2010 and 2011; Martín et al., 2012). These markers should be checked as a first step to identify varieties in Piedmont with the EU database (Pereira-Lorenzo et al., 2017).

Conversely, the EMCs series of loci, being trinucleotide SSRs, mutate at a lower rate than dinucleotide SSRs (CSCAT series), resulting in lower polymorphism (Beghè et al., 2013), as was the case for EMCs15 (PIC- 0.473). In addition, the OAL marker (Gobbin et al., 2007) presented the lowest capacity for discrimination (PIC -0.300), further emphasizing the lower values of heterozygosity (Ho=0.350; He=0.319).

The cluster analysis showed an overall high genetic diversity, which demonstrated the importance of characterizing the chestnut trees present in this territory. The traditional cultivars are frequently called according to geographic origin, ripening period and traits of the nut, creating difficulties in their classification (Gobbin et al., 2007; Martín et al., 2009). For example, the name 'Pelosa' is a cultivar known

in Emilia-Romagna and also in Piedmont for the big nut size and the presence of hairiness on the epicarp of the nut, as suggested by its name (Marinoni et al., 2013).

The study evidences that each area presented its own specific chestnut genotype (represented by Cluster 2): 'Piusela' varieties in the Reggio-Emilia area, 'Pelosa', 'Lisanese' and 'Pastanese' in the Tuscany Apennines and 'Montemarano' in Campania (Bagnaresi et al., 1977; Martín et al., 2010; Fideghelli, 2016). This was also confirmed by the Principal Component Analysis in which chestnut varieties from the Tuscan-Emilian Apennines were found to be close to each other and separated from the varieties of southern Italy and from the 'Precoce Migoule' hybrid cultivar.

Furthermore, the 'Pastanese' cultivar and 'Matildico' trees were found to belong to the same genotype which is known for the production of high-quality flour. It is at least arguable, therefore, that the 'Matildici' cultivars could be the very cultivars planted by 'Matilde of Canossa' in the Middle Ages (Piccioli, 1922; Breviglieri, 1955; Bagnaresi et al., 1977; Antonaroli and Bassi, 1999).

The presence of ancient trees and known varieties in the same cluster had already been described in Italy and Spain (Pereira-Lorenzo et al. 2019) and in Switzerland (Pereira-Lorenzo et al. 2020). On the contrary, the molecular results from sweet chestnut trees (Marroni group, Cluster 1) showed a uniform profile sharing the same allelic profile as a result of clonal propagation. This is because they were selected by growers to maintain the desired characteristics, such as high quality monoembryonic nuts with high nut weight and thin episperm (cuticle) with a floury and sweet taste (Breviglieri, 1955). These results are further confirmed by pomological characterization evidencing a high rate of homogeneity in the Marroni group (Breviglieri, 1955; Borghetti et al., 1983; Bassi and Marangoni, 1984). The selection and cultivation of these clones led to the spread of the Marroni group in distinct geographical areas. Later on, environmental factors affected the nuts' morphological aspects (Borgetti et al., 1983), leading to different denominations such as 'Marrone di Castel del Rio', 'M. di Zocca', 'Centa di San Nicolò', 'Roncegno', 'Drena', 'Marrone di Gaggio Montano', which are synonyms of the Marroni Fiorentino described in the EU chestnut database (Pereira-Lorenzo et al., 2017 and 2019). Further Marroni groups with the same molecular profile, such as cv. 'Marrone di Cuneo', 'Marrone di Combai' and 'Chiusa Pesio', were also described in Piedmont (Martín et al.,2010; Mellano et al., 2012; Marinoni et al., 2013) and showed a different genetic profile compared to the 'Marrone di Cuneo' (genetic synonym of 'Marrone Gambarogno') found in Switzerland (Pereira-Lorenzo et al., 2020).

Summarizing, the results obtained from Tuscan-Emilian Apennines varieties confirmed the close relationship between the diffusion of the genotypes and local population. Where farmers focus on clonal propagation for production purposes, such as for the Marroni group (Cluster 1), the genetic diversity of the crop is reduced. By contrast, the chestnut group (Cluster 2) featured a higher genetic diversity between distinct gene pools due to the selection of trees originated by seeds and propagated by grafting among a broad genetic base which led to a reduction of differences between wild and cultivated chestnut trees (Pereira-Lorenzo et al., 2019).

Finally, this research work points out the importance of ex situ collections so as to provide plant material for breeding programs and for nursery propagation. The availability of the molecular profile for several varieties will support the varietal classification activity, which is currently more difficult, as many genotypes were cultivated in different regions with different denominations.

2.6 Conclusions

In conclusion, the performed molecular characterization allowed the correct identification of the varieties mainly cultivated in the area of the Tuscan-Emilian Apennines. The identification of synonymous accessions emphasized the importance of verifying collections of germplasm with powerful tools such as molecular markers. These tools are fundamental to avoid both redundancy and possible issues of varietal certification for propagation in nurseries.

Furthermore, this research promotes the diffusion of ecotypes to evaluate the preservation of chestnut biodiversity with the inclusion of varieties at risk of genetic erosion. The involvement of local farmers as project partners increased their awareness of underlying matters and their availability to host and guard plants at risk of genetic erosion. Genotypes at risk, e.g. the Marroni group, must be reintroduced taking into account soil and climate characteristics.

This research also analyzed the genetic diversity with the aim of enriching collection fields in the Emilia-Romagna region through identified unique varieties. The results confirm that the Italian chestnut germplasm is an important source of genetic biodiversity and contributes to the preservation and enhancement of the entire chestnut genetic heritage.

2.7 Bibliography

Antonaroli, R.; Bagnaresi, U.; Bassi, D. (1984) Indagini sulla variazione di alcuni caratteri morfologici in popolazioni di castagno da frutto nella provincia di Bologna. Monti e Boschi, 2, 47-50.

Antonaroli, R.; Bassi, D. (1999) Le varietà dell'Emilia-Romagna. Il divulgatore XXII, 10, 5-34.

Baccichet, I.; Foria, S.; Messina, R.; Peccol, E.; Losa, A.; Fabro, M.; Gori, G.; Zandigiacomo, P.; Cipriani, G.; Testolin, R. Genetic and ploidy diversity of pear (Pyrus spp.) germplasm of Friuli-Venezia Giulia, Italy. Genet Resour Crop Evol, 2020, 67, 83–96.

Bagnaresi, U.; Bassi, D.; Casini, E.; Conticini, L.; Magnani, G.P. (1977) Contributo alla individuazione delle cultivar di castagno tosco-emiliane. Atti del Convegno "Giornata del Castagno", Caprese Michelangelo (AR), 165-234.

Bassi D.; Marangoni B. (1984) Contributo allo studio varietale del castagno da frutto (*Castanea sativa* Mill.): caratteri biometrici e analisi chimico-fisiche dei frutti. Rivista di Frutticoltura, 6, 43–46.

Beccaro, G. L.; Donno, D.; Lione, G. G.; De Biaggi, M.; Gamba, G.; Rapalino, S.; Riondato, I.; Gonthier, P.; Mellano, M. G. (2020). Castanea spp. Agrobiodiversity Conservation: Genotype Influence on Chemical and Sensorial Traits of Cultivars Grown on the Same Clonal Rootstock. Foods, 9(8), 1062.

Beccaro, G.; Bounous, G.; De Biaggi, M.; Donno, D.; Marinoni, D.T.; Zou, F.; Mellano, M.G. (2019) Botany, Anatomy, and Nut Composition. In the Chestnut Handbook; CRC Press: Boca Raton, FL, USA, pp. 5–34.

Beghè D.; Ganino T.; Dall'Asta C.; Silvanini A.; Cirlini M.; Fabbri A. (2013) Identification and characterization of ancient Italian chestnut using nuclear microsatellite markers. Scientia Horticulturae, 161, 50-57.

Bhattarai, G. and Mehlenbacher, S. A. (2017) In silico development and characterization of tri-nucleotide simple sequence repeat markers in hazelnut (*Corylus avellana* L.). PloS one, 12(5), e0178061

Borghetti, M.; Giannini, R.; Nocentini C. (1983) Indagini preliminari sulla variazione di alcuni caratteri del frutto in popolazioni di "Marrone Fiorentino". Monti e Boschi, 1, 49-52.

Bouffartigue, C.; Debille, S.; Fabreguette, O.; Ramos Cabrer, A.; Pereira-Lorenzo, S.; Flutre, T.; Harvengt, L. (2019) Genetic diversity and structure of sweet chestnut (*Castanea sativa* Mill.) in France: At the intersection between Spain and Italy bioRxiv, 792259; doi: https://doi.org/10.1101/792259

Breviglieri N. (1995) Indagini ed osservazioni sulle migliori varietà italiane di castagno (*Castanea sativa* Mill.). Centro di Studio sul Castagno, pubblicazione, 2, 27-166.

Buck, E.J.; Hadonou, M.; James, C.J.; Blakesley, D.; Russell, K. (2003) Isolation and characterization of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill.) Mol Ecol Notes, 3, 239–241. https://doi.org/10.1046/j.1471-8286.2003.00410.x

Cipriani, G.; Spadotto, A.; Jurman, I.; Di Gaspero, G.; Crespan, M.; Meneghetti, S.; Frare, E.; Vignani, R.; Cresti, M.; Morgante, M.; Pezzotti, M.; Pe, E.; Policriti, A.; Testolin, R. (2010) The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages and reveals a large admixture amongst varieties of different geographic origin. Theor Appl Genet, 121, 1569-1585. https://doi.org/10.1007/s00122-010-1411-9

Conedera, M.; Krebs, P.; Tinner, W.; Pradella, M.; Torriani, D. (2004) The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. Vegetation History and Archaeobotany, 13.3, 161–179. https://doi.org/10.1007/s00334-004-0038-7

Conedera, M.; Tinner, W.; Krebs, P. (2016) *Castanea sativa* in Europe: distribution, habitat, usage and threats.

Dice, L.R. (1945) Measures of the amount of ecologic association between species. Ecology, 26, 297–302.

Ferradini, N.; Lancioni, H.; Torricelli, R.; Russi, L.; Dalla Ragione, I.; Cardinali, I.; Marconi, G.; Gramaccia, M.; Concezzi, L.; Achilli, A.; Veronesi, F., Albertini, E. (2017) Characterization and phylogenetic analysis of ancient Italian landraces of pear. Frontiers in plant science, 8, 751.

Fideghelli, C. (2016) Aspetti pomologici e qualitativi dei materiali di propagazione. I Georgofili, 2.2, 55-61.

Fideghelli, C. (2016) Atlante dei fruttiferi autoctoni italiani. Crea - Centro di frutticoltura, Roma.

Gallesio, G. (1817-1839) Pomona Italiana, ossia Trattato degli alberi fruttiferi. Niccolò Capurbo, Pisa.

Gobbin, D.; Hohl, L.; Conza, L.; Jermini, M.; Gessler, C.; Conedera, M. (2007) Microsatellite-based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. Genome, 50.12, 1089–1103. https://doi.org/10.1139/G07-086

Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol, 16, 1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Kaltenrieder, P; Procacci, G.; Vannière, B.; Tinner, W. (2010) Vegetation and fire history of the euganean Hills (Colli Euganei) as recorded by Late glacial and Holocene sedimentary series from Lago della Costa (northeastern Italy). Holocene, 20, 679e695.

Kampfer, S.; Lexer, C.; Glossl, J.; Steinkellner, H. (1998) Characterization of (GA)(n) microsatellite loci from Quercus robur. Hereditas, 129.2, 183–186. https://doi.org/10.1111/j.1601-5223.1998.00183.x

Krebs, P.; Pezzatti, G. B.; Beffa, G.; Tinner, W.; Conedera, M. (2019) Revising the sweet chestnut (*Castanea sativa* Mill.) refugia history of the last glacial period with extended pollen and macrofossil evidence. Quaternary science reviews, 206, 111-128.

Landete, J.M. (2011) Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health. Food Res. Int. 44, 1150–1160.

Liang, W.; Dondini, L.; De Franceschi, P.; Paris, R.; Sansavini, S.; Tartarini, S. (2015) Genetic diversity, population structure and construction of a core collection of apple cultivars from Italian germplasm. Plant Mol Biol Rep, 33, 458-473

Lucchi, E.; Frascaroli, F.; Maresi, G.; Ferretti, F.; Viaggi, D.; Pezzi, G. (2016) Gestione dei castagneti, realtà e prospettive. Ecoscienza, 1, 72-73

Maguire, T. L.; Collins, G. G.; Sedgley, M. (1994) A modified CTAB DNA extraction procedure for plants belonging to the family proteaceae. Plant Molecular Biology Reporter, 12.2, 106–109. https://doi.org/10.1007/BF02668371

Marinoni, D. T.; Akkak, A.; Beltramo, C.; Guaraldo, P.; Boccacci, P.; Bounous, G.; Ferrara, A.M.; Ebone, A.; Viotto, E.; Botta, R. (2013) Genetic and morphological characterization of chestnut (*Castanea sativa* Mill.) germplasm in Piedmont (north-western Italy). Tree genetics & genomes, 9.4, 1017-1030.

Marinoni, D.; Akkak, A.; Bounous, G.; Edwards, K.J.; Botta, R. (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). Mol. Breed, 11.2, 127–136 https://doi.org/10.1023/A:1022456013692.

Marshall, T.C.; Slate, J.; Kruuk, L.E.B.; Pemberton, J.M. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol, 7, 639–655.

Martín, M.A.; Alvarez, J.B.; Mattioni, C.; Cherubini, M.; Villani, F.; Martín, L.M. (2009) Identification and characterization of traditional chestnut varieties of southern Spain using morphological and simple sequence repeat (SSRs) markers. Ann Appl Biol, 154, 389–398. https://doi.org/10.1111/j.1744-7348.2008.00309.x

Martín, M.A.; Mattioni, C.; Cherubini, M.; Taurchini, D.; Villani F. (2010) Genetic characterisation of traditional chestnut varieties in Italy using microsatellites (SSRs) markers. Ann Appl Biol, 157, 37–44. https://doi.org/10.1111/j.1744-7348.2010.00407.x

Martín, M.A.; Mattioni, C.; Molina, J.R.; Alvarez, J.A.; Cherubini, M.; Herrera, M.A.; Villani, F.; Martín, L.M. (2012) Landscape genetics structure of chestnut (*Castanea sativa* Mill.) in Spain. Tree Genet Genomes, 8:127–136. https://doi.org/10.1007/s11295-011-0427-x

Mellano, M.G.; Beccaro, G.L.; Donno, D.; Torello Marinoni, D.; Boccacci, P.; Canterino, S.; Cerutti, A. K.; Bounous P. (2012) Castanea spp. biodiversity conservation: collection and characterization of the genetic diversity of an endangered species. Genet Resour Crop Evol, 2012, 59, 1727–1741.

Neri, L.; Dimitri, G.; Sacchetti G. (2010) Chemical composition and antioxidant activity of cured chestnuts from three sweet chestnut (*Castanea sativa* Mill.) ecotypes from Italy. Journal of Food Composition and Analysis, 23.1, 23–29. https://doi.org/10.1016/j.jfca.2009.03.002

Pereira-Lorenzo, S.; Bischofberger, Y.; Conedera, M.; Piattini, P.; Crovadore, J.; Chablais, R.; Rudow, A.; Hatt, S.; Ramos-Cabrer, A.M.; Barreneche, T.; Lefort, F. (2020) Reservoir of the European chestnut diversity in Switzerland. Biodivers Conserv. https://doi.org/10.1007/s10531-020-01970-2

Pereira-Lorenzo, S.; Costa, R.; Ramos-Cabrer, A.M.; Ribeiro, C.; Serra da Silva, C.; Manzano, G.; Barreneche, T. (2010) Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. Tree Genet Genomes, 6, 701–715. https://doi.org/10.1007/s11295-010-0285-y

Pereira-Lorenzo, S.; Lourenço Costa, R.M.; Ramos-Cabrer, A.M.; Ciordia-Ara, M.; Marques Ribeiro, C.A.; Borges, O.; Barreneche, T. (2011) Chestnut cultivar diversification process in the Iberian Peninsula, Canary Islands, and Azores. Genome, 54:301–315. https://doi.org/10.1139/g10-122

Pereira-Lorenzo, S.; Ramos-Cabrer, A.M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, M.B.; Martín, L.M.; Martín, Á. (2017) Database of European chestnut cultivars and definition of a core collection using simple sequence repeats. Tree genetics & genomes, 13.5: 114. https://doi.org/10.1007/s11295-017-1197-x

Pereira-Lorenzo, S.; Ramos-Cabrer, A. M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, B.; Martín, L.M.; Robles-Loma, A.; Cáceres, Y.; Martín, A. (2019) Instant domestication process of European chestnut cultivars. Annals of Applied Biology, 174.1, 74-85.

Pezzi, G.; Maresi, G.; Conedera, M.; Ferrari, C. (2011) Woody species composition of chestnut stands in the Northern Apennines: the result of 200 years of changes in land use. Landscape Ecology, 2011, 26, 1463–1476.

Piccioli, L. (1902) Monografia del castagno.

Quintana, J.; Contreras, A.; Merino, I.; Vinuesa, A.; Orozco, G.; Ovalle, F.; Gomez, L. (2015) Genetic characterization of chestnut (*Castanea sativa* Mill.) orchards and 75 traditional nut varieties in El Bierzo, a glacial refuge and major cultivation site in northwestern Spain. Tree Genetics and Genomes, 11.1 https://doi.org/10.1007/s11295-014-0826-x

Rohlf, F.J. (1992) NTSYS-pc Numerical taxonomy and multivariate analysis system (version 1.80). Applied Biostatistics.

Urrestarazu, J.; Denancé, C.; Ravon, E.; Guyader, A.; Guisnel, R.; Feugey, L.; Poncet, C.; Lateur, M.; Houben, P.; Ordidge, M.; Fernandez-Fernandez, F.; Evans, K.M.; Paprstein, F.; Sedlak, J.; Nybom, H.; Garkava-Gustavsson, L.; Miranda, C.; Gassmann, J.; Kellerhalls, M.; Suprun, I.; Pikunova, A.V.; Krasova,

N.G.; Torutaeva, E.; Dondini, L.; Tartarini, S.; Laurens, F.; Durel, C.E. (2016) Analysis of the genetic diversity and structure across a wide range of germplasm reveals prominent gene flow in apple at the European level. BMC Plant Biology, 16.1, 130. https://doi.org/10.1186/s12870-016-0818-0

Urrestarazu, J.; Royo, J.B.; Santesteban, L.G.; Miranda, C. (2015) Evaluating the Influence of the Microsatellite Marker Set on the Genetic Structure Inferred in *Pyrus communis* L., PLoS One, 10.9 doi: 10.1371/journal.pone.0138417.

CHAPTER 3: Pomological characterization and fingerprinting of chestnut trees from the Didactic and Experimental Park of Granaglione

3.1 Abstract

As part of the regional project "BIODIVERSAMENTE CASTAGNO" and with the support of the National Academy of Agriculture, a panel of 118 accessions of *Castanea sativa* Mill., preserved into Didactic and Experimental Park of Granaglione collections located in the Tuscan-Emilian Apennines was analyzed at molecular level by means of microsatellite markers. The genetic distance between accessions has been calculated through the DICE coefficient and a dendrogram has been constructed using the clustering method (UPGMA). The results showed the relationships between the accessions analyzed (Marroni group and chestnuts); a very uniform molecular profile was produced for the grafted varieties while for the ancient local varieties an allelic variability was observed in specific loci. This information is useful to characterize and to define the main genotypes present in this territory for the conservation of their biodiversity.

3.2 Introduction

Chestnut cultivation has always played a key role in the Italian economy, in particular mountain areas, such as the Apennines Tosco-Emiliano. The European chestnut (*Castanea sativa* Mill.) shows a high genetic variability, probably due to the vast area of cultivation and the adaptation to the various geo-pedoclimatic situations in which it grows, situations not always favorable. Unfortunately, there is little information on genetic diversity of the chestnut varieties present in Emilia-Romagna.

Currently, most of the cultivated varieties of chestnut are the result of a selection of local genotypes by farmers, that has led to a remarkable enrichment of the genetic biodiversity of this species. Nevertheless, this situation has made varietal classification difficult: these genotypes were cultivated in different regions with different denominations. Today there are varieties that are usually referred to by geographical origin, maturing period and/or type of use, making cataloguing very difficult (Fineschi, 1988).

The Didactic and Experimental Park of Granaglione is an important Centre for the conservation of the biodiversity of the species, and a great opportunity for an analysis of genetic diversity to be conducted with molecular markers.

Molecular marker analyses (SSRs or Simple Sequence Repeat) are able to effectively analyze genetic diversity within populations and provide useful support for analyses aimed at the varietal identification of that species within collections and in the territory.

The Park has an area of 9 hectares at an altitude of about 600 m.s.l.m. It was acquired in 2003 by the Foundation Cassa di Risparmio in Bologna as part of the Apennines Project for the enhancement of chestnut culture by Prof. Umberto Bagnaresi, in collaboration with the Experimental Centre for the Study and Analysis of the Soil of the University of Bologna.

The Park territory was divided into six areas of interest (Figure 3.1): A) traditional chestnuts, in order to preserve the traditional cultural examples of the beginning of the last century; B) wood-specialized chestnuts, obtained by seeds or grafting from four selected varieties ('Cardaccio', 'Mozza', 'Perticaccio', 'Politora') were included. These varieties were selected for the rapid growth, the absence of "ring shake" and the superior technological characteristics of wood; C) specialized fruit chestnuts, with 14 distinct local varieties of chestnuts and Marroni, including the cultivars 'Ceppa' and 'Pastinese', that appear among the 13 types of chestnut at risk, as reported in the catalogue of the Emilia-Romagna region; D) monumental chestnuts, characterized by majestic ancient plants with narrow canopy and large space in the between, according to the traditional 'sesto matildico'; E) natural areas, with spontaneous tree species such as poplar, cherry, oak, birch, developed over time together with some sporadic conifers. The aim was to re-naturalize the mixed forest area by promoting the development of plant biodiversity; F) propagation areas, for the 'ceduo' forest.

Figure 3.1: Map of Didactic Experimental Park of Granaglione subdivided into six areas: traditional chestnuts; wood-specialized chestnuts; specialized fruit chestnuts; monumental chestnuts; natural areas; and propagation areas.

The main objective of the Didactic Experimental Park of Granaglione was to preserve the biodiversity of the present ecotypes and the existing chestnut heritage from further genetic erosion, in order to create the starting point for future selection programs useful for the revival of chestnuts as a fresh product or for the production of flours, as many problems in recent decades have plagued Italian chestnut cultivation.

3.3 Plant materials and cluster analysis

The leaf material for molecular analysis was taken in the Didactic Exerimental Park of Granaglione (BO)(Figure 3.2). The varieties were collected in Table 3.1 for a total of 118 accessions.

For each fruit sample withdrawn, two replicates (A and B) were initially collected from different plants with the same name. The comparison between the samples was used to define varietal references. If different, all the plants classified with the same name were taken in order to confirm their identity.

The Matildici samples, which are hypothesized to be the varieties spread by Matilde of Canossa (Mantova, 1046-1115) during the Middle Ages with the help of the Benedictine monks, were collected from the "monumental chestnut", as a historical testimony of the chestnut crop.

For proper identification of samples collected from monumental chestnuts, the sample code was associated with the GPS coordinate reference (Table 3.1). In this way it was possible to reach two objectives: a) put together a list of trees and their position in the sites representative for the different pedologic environments; b) define a sampling protocol.

Genomic DNA was extracted from young leaves following the standard CTAB protocol (Maguire et al., 1994) and the samples were analyzed with 16 microsatellites markers (SSRs), chosen based on their high polymorphism (Pereira-Lorenzo et al., 2017).

An ABI 3730 XL Analyzer (Applied Biosystems) was used to analyze the fragments available at the Medical Genetics Laboratory at the Sant'Orsola Hospital in Bologna. The collected data were analyzed using Cluster analysis conducted with the SimQual NTSYSpc 2.0 procedure, protocol which confirms or disproves the differences between genotypes (Rohlf et al., 2004). The genetic distance between accessions was calculated by the DICE coefficient (Dice 1945).

Figure 3.2: Map of the Emilia-Romagna region with particular reference to the Didactic and Experimental Park of Granaglione (BO).

Table 3.1: List of the 118 accessions collected (86 chestnut varieties and 32 wood chestnuts), their origin area, GPS coordinates and use.

Samples number	Name Variety	Tree number	GPS Coordinate	Samples number	Name Variety	Tree number	GPS Coordinate
1	Bovalghe	106		60	Matildico 15		N 44°08.432' E 010°57.469'
2	Bovalghe	107		61	Pastinese	5	
3	Castel del Rio	1		62	Pastinese	3	
4	Unknown 1	2		63	Unknown 17	75	
5	Unknown 2	7	N 44°29.715' E 011°20.206'	64	Unknown 18	78	
6	Unknown 3	124		65	Pastanese (biffoni)	79	
7	Unknown 4	125		66	Pastanese (biffoni)	80	
8	Castel del Rio	128		67	Pastanese (biffoni)	81	
9	Castel del Rio	129		68	Pastanese (biffoni)	83	
10	Unknown 5	9		69	Unknown 19	85	
11	Castel del Rio	137		70	Pelosa	43	
12	Unknown 6	64		71	Pelosa	45	
13	Unknown 7	138		72	Roncegno	140	
14	Castione	61		73	Roncegno	147	
15	Castione	60		74	Unknown 20	13	
16	Unknown 8	63		75	Unknown 21	14	
17	Castione	59		76	Unknown 22	20	
18	Unknown 9	58		77	Unknown 23	27	
19	Castione	57		78	Unknown 24	28	
20	Castione	54		79	Sborgà	116	
21	Unknown 10	55		80	Sborgà	136	
22	Centa S. Nicolò	67		81	Sborgà	15	
23	Centa S. Nicolò	70		82	Sborgà	16	
24	Ceppa	32		83	Svizzera	17	
25	Ceppa	40		84	Svizzera	135	
26	Drena	30		85	Zocca	22	
27	Unknown 11	31		86	Zocca	24	
28	Unknown 12	33		87	Legno 1	303	N 44°08.408' E 010°57.483'
29	Unknown 13	34		88	Legno 2	419	N 44°08.407' E 010°57.488'
30	Unknown 14	111		89	Legno 3	414	N 44°08.404' E 010°57.487'
31	Unknown 15	112		90	Legno 4	411	N 44°08.400' E 010°57.494'
32	Drena	29		91	Legno 5	340	N 44°08.396' E 010°57.516'
33	Drena	108		92	Legno 6	385	N 44°08.395' E 010°57.503'
34	Lisanese	41b	N 44°08.407' E 010°57.375'	93	Legno 7	371	N 44°08.395' E 010°57.508'
35	Lisanese	42b	N 44°08.402'' E 010°57.370'	94	Legno 8	82	N 44°08.388' E 010°57.527'
36	Lisanese	47b		95	Legno 9	107	N 44°08.389' E 010°57.530'
37	Lisanese	48	N 44°08.400' E 010°57.380'	96	Legno 10	109	N 44°08.389' E 010°57.531'
38	Lisanese	49		97	Legno 11	108	N 44°08.391' E 010°57.532'
39	Lisanese	41c	N 44°08.394' E 010°57.375'	98	Legno 12	102	N 44°08.392' E 010°57.535'
40	Lisanese	71	N 44°08.393' E 010°57.375'	99	Legno 13	99	N 44°08.389' E 010°57.538'
41	Unknown 16	72		100	Legno 14	97	N 44°08.386' E 010°57.534'
42	Lisanese		N 44°08.402' E 010°57.369'	101	Legno 15	288	N 44°08.414' E 010°57.468'
43	Lisanese	42		102	Legno 16	284	N 44°08.410' E 010°57.463'
44	Lisanese	47		103	Legno 17	282	N 44°08.406' E 010°57.467'
45	Lisanese	41		104	Legno 18	269	N 44°08.399' E 010°57.461'
46	Matildico 1		N 44°08.361' E 010°57.502'	105	Legno 19	273	N 44°08.397' E 010°57.465'
47	Matildico 2		N 44°08.358' E 010°57.514'	106	Legno 20	226	N 44°08.390' E 010°57.481'
48	Matildico 3		N 44°08.349' E 010°57.537'	107	Legno 21	230	N 44°08.387' E 010°57.484'
49	Matildico 4		N 44°08.350' E 010°57.547'	108	Legno 22	239	N 44°08.388' E 010°57.494'
50	Matildico 5		N 44°08.353' E 010°57.555'	109	Legno 23	353	N 44°08.389' E 010°57.497'
51	Matildico 6		N 44°08.341' E 010°57.553'	110	Legno 24	350	N 44°08.388' E 010°57.503'
52	Matildico 7		N 44°08.341' E 010°57.543'	111	Legno 25	345	N 44°08.389' E 010°57.509'
53	Matildico 8		N 44°08.317' E 010°57.511'	112	Legno 26	140	N 44°08.385' E 010°57.511'
54	Matildico 9		N 44°08.295' E 010°57.513'	113	Legno 27	124	N 44°08.380' E 010°57.514'
55	Matildico 10		N 44°08.296' E 010°57.514'	114	Legno 28	34	N 44°08.380' E 010°57.519'
56	Matildico 11		N 44°08.393' E 010°57.456'	115	Legno 29	60	N 44°08.376' E 010°57.518'
57	Matildico 12		N 44°08.385' E 010°57.441'	116	Legno 30	65	N 44°08.374' E 010°57.524'
58	Matildico 13		N 44°08.377' E 010°57.440'	117	Legno 31	19	N 44°08.372' E 010°57.509'
59	Matildico 14		N 44°08.425' E 010°57.464'	118	Legno 32	10	N 44°08.372' E 010°57.506'

3.4 Pomological characterization

In order to describe different tree characteristics in the collection field, the morphological descriptors were determined using mainly UPOV (1989) and Bolvansky & Mendel (2001) descriptors for *Castanea sativa* (Table 3.2). In this way, it was possible to confirm the "true-to-type" or "not-true-to-type" chestnut accessions analyzed at a phenotypic level in the collection field, in order to create maps of genetic identity of the main varieties present in the territory. A genetic profile was added to the identity cards for future varietal certification.

Table 3.2. Descriptors list for Castanea sativa Mill by UPOV (1989) and Bolvansky & Mendel (2001).

1.	**Length of spines (mm)**	short (until 7mm), medium (7.1-14.9), long (15-25mm)
2.	**Number of nuts per burr**	calculated on 2 fruits
3.	**Ripening time**	very early (before 15/9), early (15-30/9), medium (1-15/10), late (16-31/10), very late (after 01/11)
4.	**Size (number of nuts per kg)**	very big <60/kg, big 61-80kg, medium 81-100kg, small 101-120kg, very small >120/kg
5.	**Color**	light brown, brown, dark brown, reddish brown, blackish brown
6.	**Shape**	ovoid, broad ovoid, globose, transverse ellipsoid, transverse - broad ellipsoid
7.	**Hairiness**	absent, present (only around the torch), present (only around the torch and downward), present (spread all over the nut)
8.	**Embryony**	mono-embryonic or poly-embrionic
9.	**Degree of penetration of seed coat into embryo**	weak, medium, strong
10.	**Hilum**	small, medium, large

3.5 Results and discussion

Molecular analysis showed that the different varieties of chestnuts in the Didactic and Experimental Park of Granaglione have very different molecular profiles, while the Marroni group had a unique allelic profile (Figure 3.2).

The Marroni of Castel del Rio and of Zocca at were confirmed true-to-type, belonging to the Marroni group. In addition, also to the Trentino Marroni group with the cultivars ‘Roncegno’, ‘Drena’, ‘Centa San Nicolò’ and ‘Castione’ shared the same allelic profile with ‘Castel del Rio’ and ‘Zocca’. Furthermore, ‘Sborgà’, ‘Bovalghe’ and a group of ‘Pastinese’ accessions showed a Marroni allelic profile.

Marroni, although they are called cultivars, could be defined as populations with a high degree of genetic homogeneity derived from the propagation by graft. The selection and cultivation of clones led to the spread of this genotype in several geographical areas with the development of distinctive morphological aspects, leading to the confusing emergence of different denominations.

The varieties ‘Ceppa’, ‘Lisanese’, ‘Svizzera’, ‘Pelosa’, ‘Calarese’, ‘Carrarese’, ‘Piusela’, ‘Pastanese (Biffoni)’ were all well distinguishable with molecular markers.

After genetic characterization, the true-to-type chestnut accessions were confirmed at a phenotypic level. Figures 3.3 shows the reference varieties 'Castel del Rio', 'Castione', 'Drena' and 'Sborgà' (Marroni-type) and 'Pastanese (biffoni)' (chestnut-type) in relation to a panel of not-true-type accessions. In this case these putative misnomers seem to be due to the characteristic of the chestnuts trees of producing suckers from the stumps. The allele variability in fact seems to confirm that they are derived from natural hybridization of the chestnuts and they should be considered, from a genetic point of view, as seedlings (data not shown). For these reasons all those genotypes have been indicated as “Unknown” in Figure 3.2. The tendency of suckering of the chestnuts should be considered among the problem that are expected during the maintenance of the chestnut germplasm collections and this situation needs a constant pomological and molecular control of the plants that is indispensable for verifying the true varietal correspondence.

Chestnut fruit characteristics for the 14 verified chestnut varieties present in the Didactic and Experimental Park of Granaglione following the UPOV (1989) and Bolvansky and Mendel (2001) descriptors for *Castanea sativa*, are shown in Table 3.3.

The analysis finally focused on the study of 15 Matildici specimens, to extend analysis on genetic variability of *C. sativa*.

The Matildici samples 4, 15 and 8 showed equal allelic profile of the 'Pastanese (Biffoni)' variety, confirming the great rooting in the territory of this variety. The Matildici 11, 6 and 12 are also very similar to the genetic profile of 'Pastanese', which nevertheless have slight differences at the level of the allelic profile. This finding is very common, however, when analyzing ancient varieties with molecular markers, as microsatellite markers, subject to mutation variability. The presence of very ancient trees with this allelic profile seems to demonstrate that they represent the true-to-type 'Pastanese'.

Other Matildici have shown greater allelic variability and in-depth molecular and pomological analysis will be needed to identify the variety to which they should be ascribed. Point-like genetic mutations, structural variations genomics and epigenetics are difficult to identify with a reduced number of SSR markers (Gross et al., 2012), while they are morphologically observable (Cipriani et al., 2010).

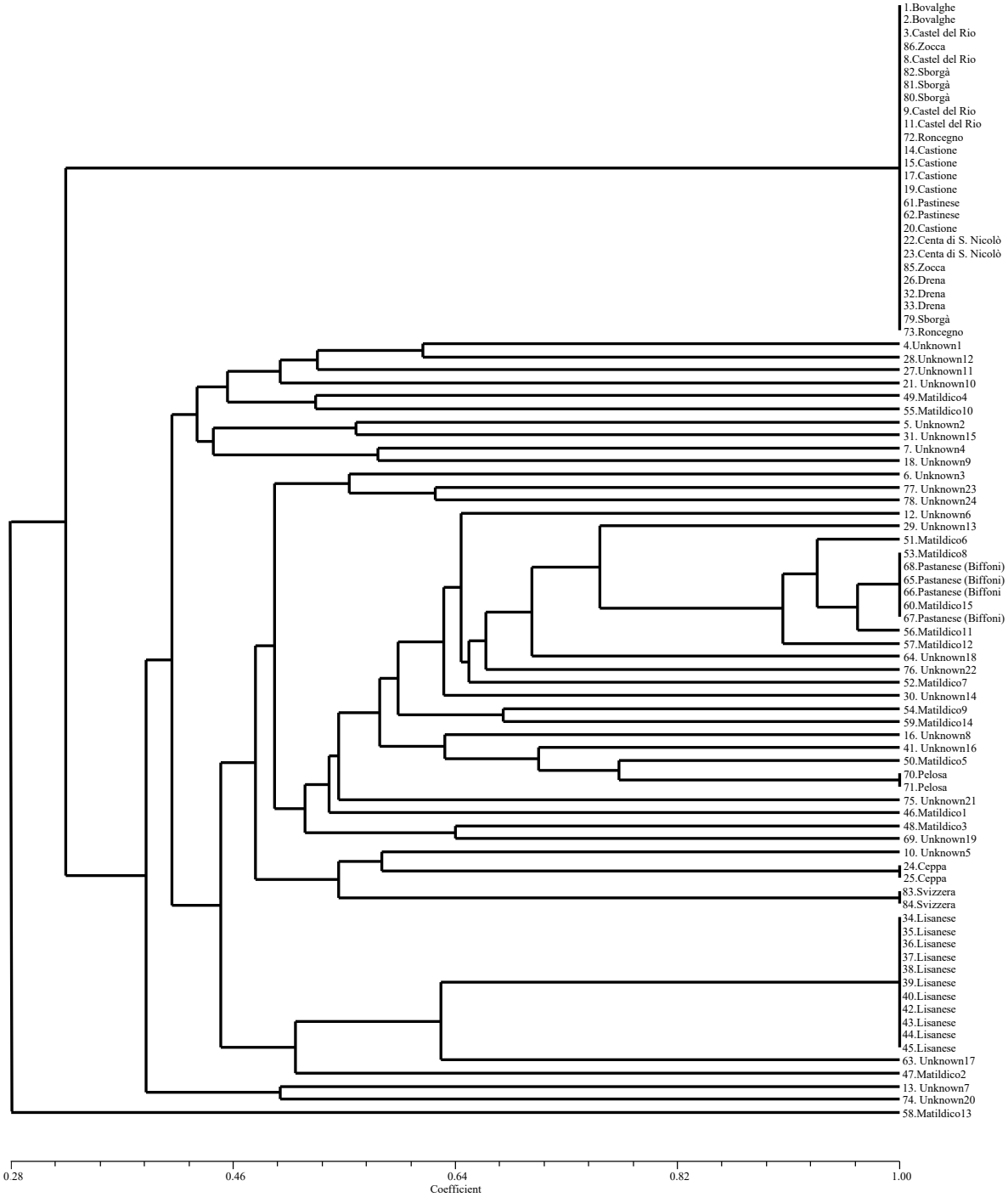

Figure 3.2: UPGMA elaborated with NTYSYS of 86 chestnut samples

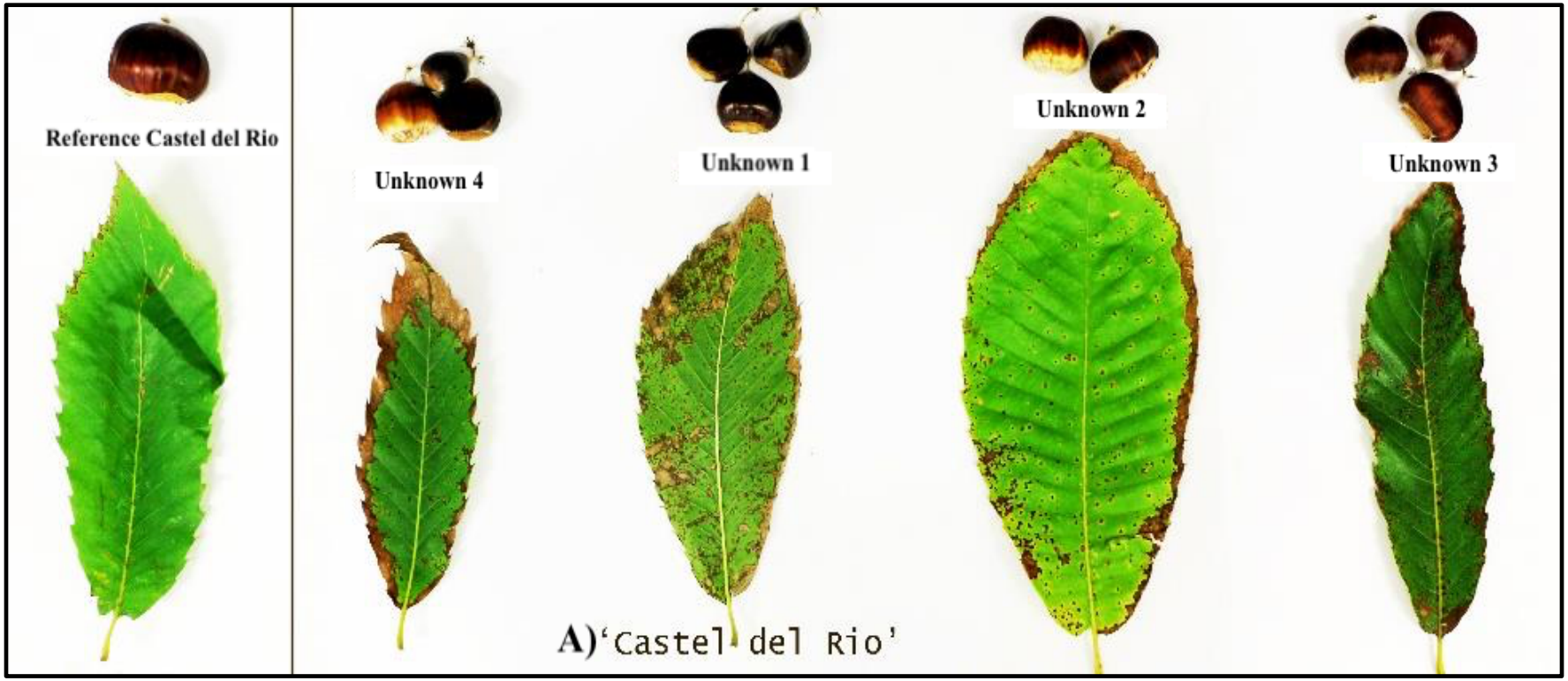

Figure 3.3 A: Comparison of the reference Castel del Rio Marrone with all the other cultivars represent not-true-to-type identity.

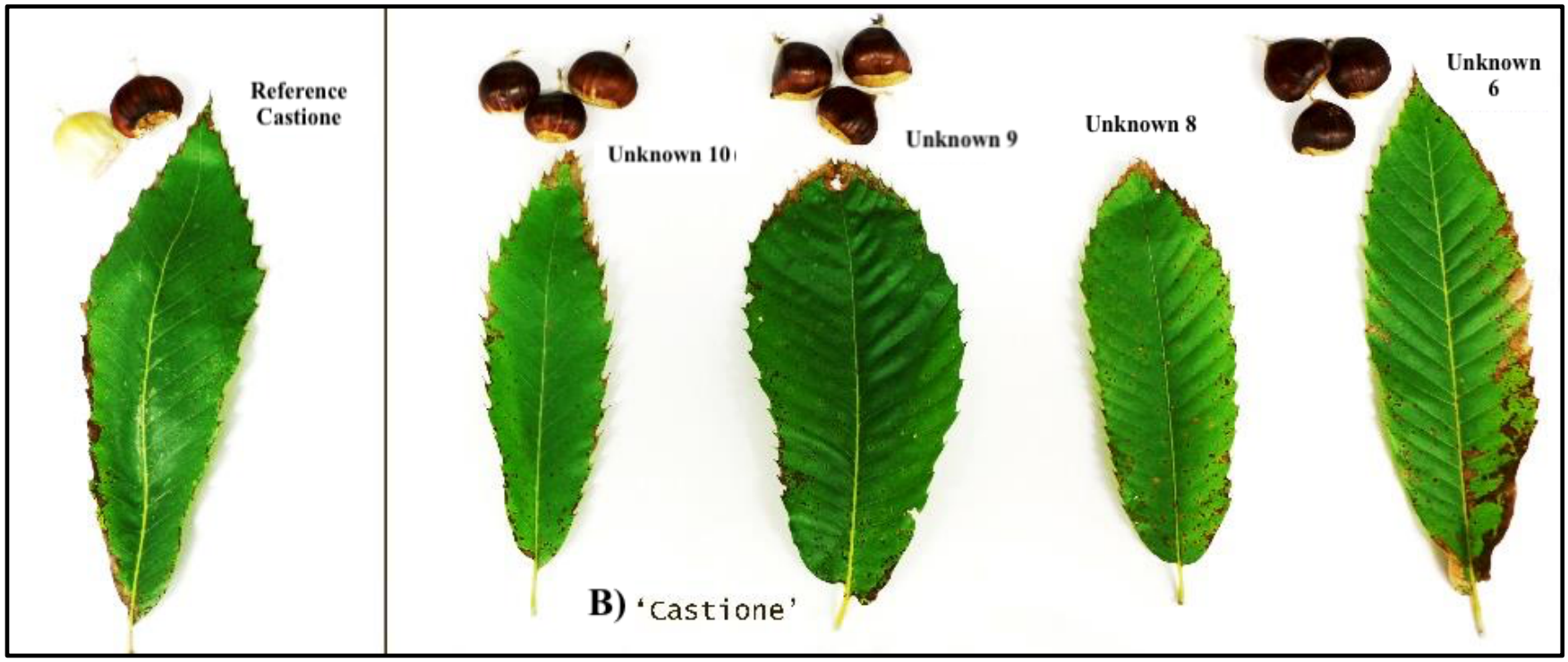

Figure 3.3 B: Comparison of the reference Castione Marrone with all the other cultivars represent not-true-to-type identity.

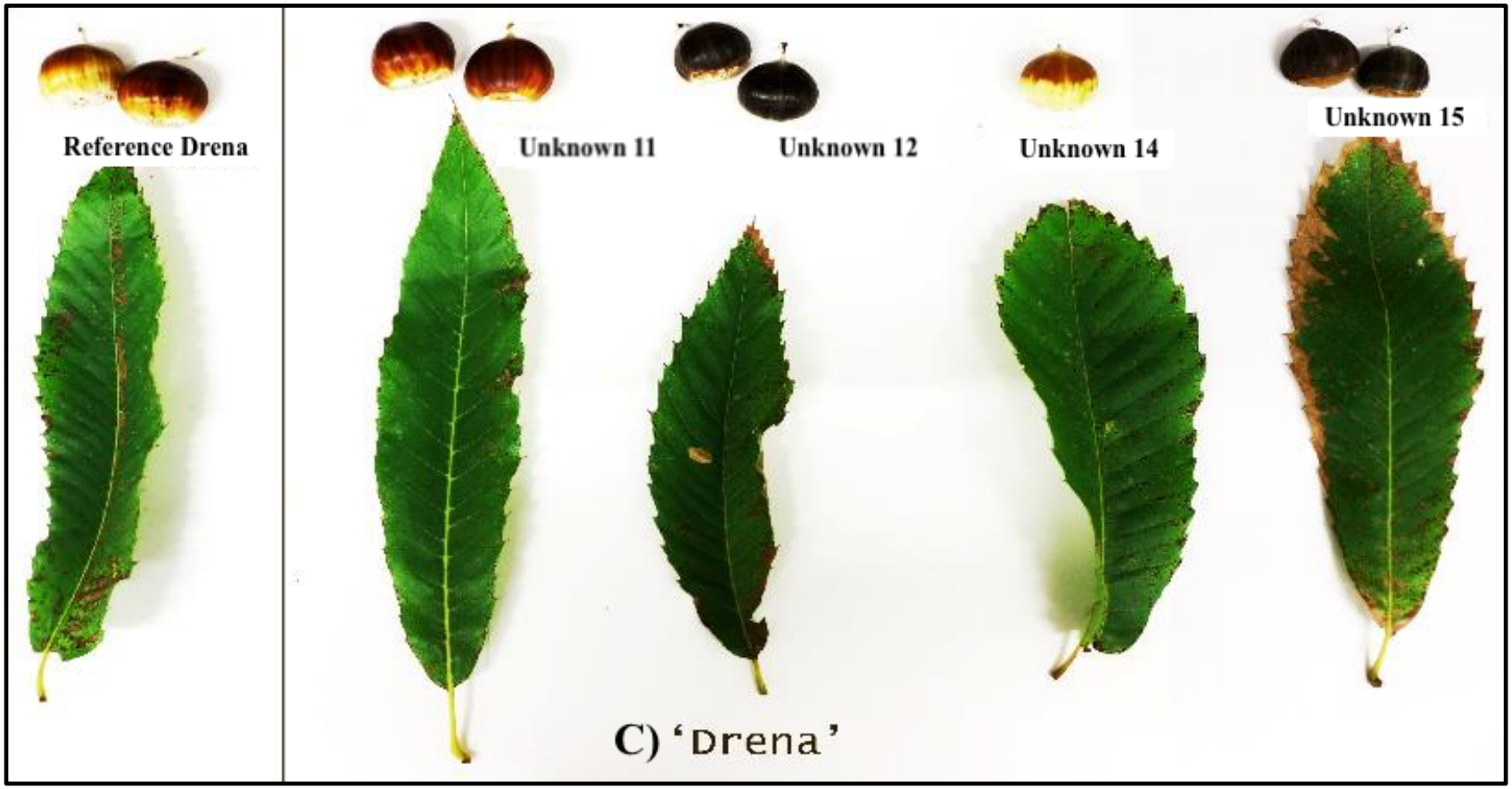

Figure 3.3 C: Comparison of the reference Drena Marrone with all the other cultivars represent not-true-to-type identity.

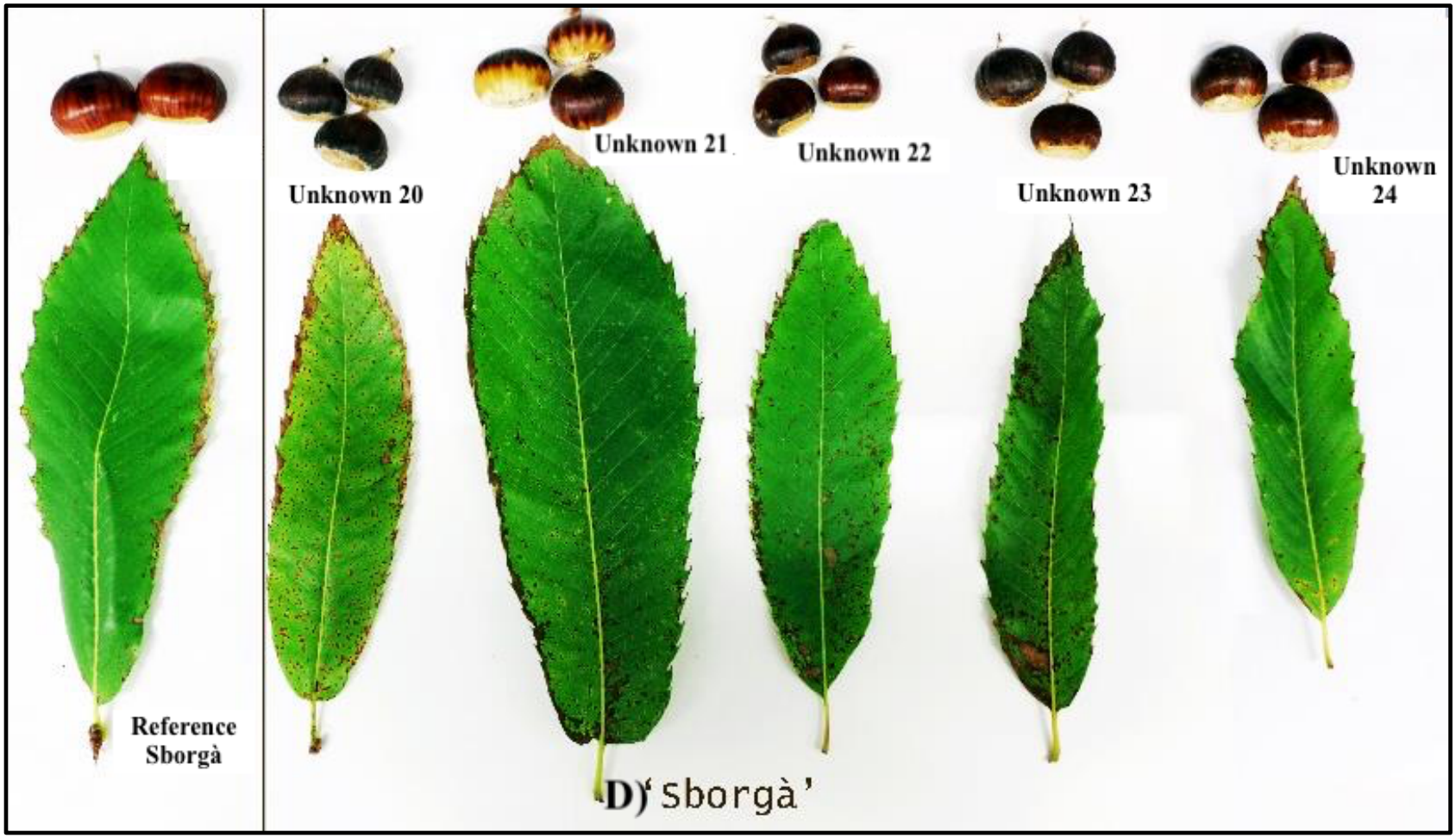

Figure 3.3 D: Comparison of the reference Sborgà Marrone with all the other cultivars represent not-true-to-type identity.

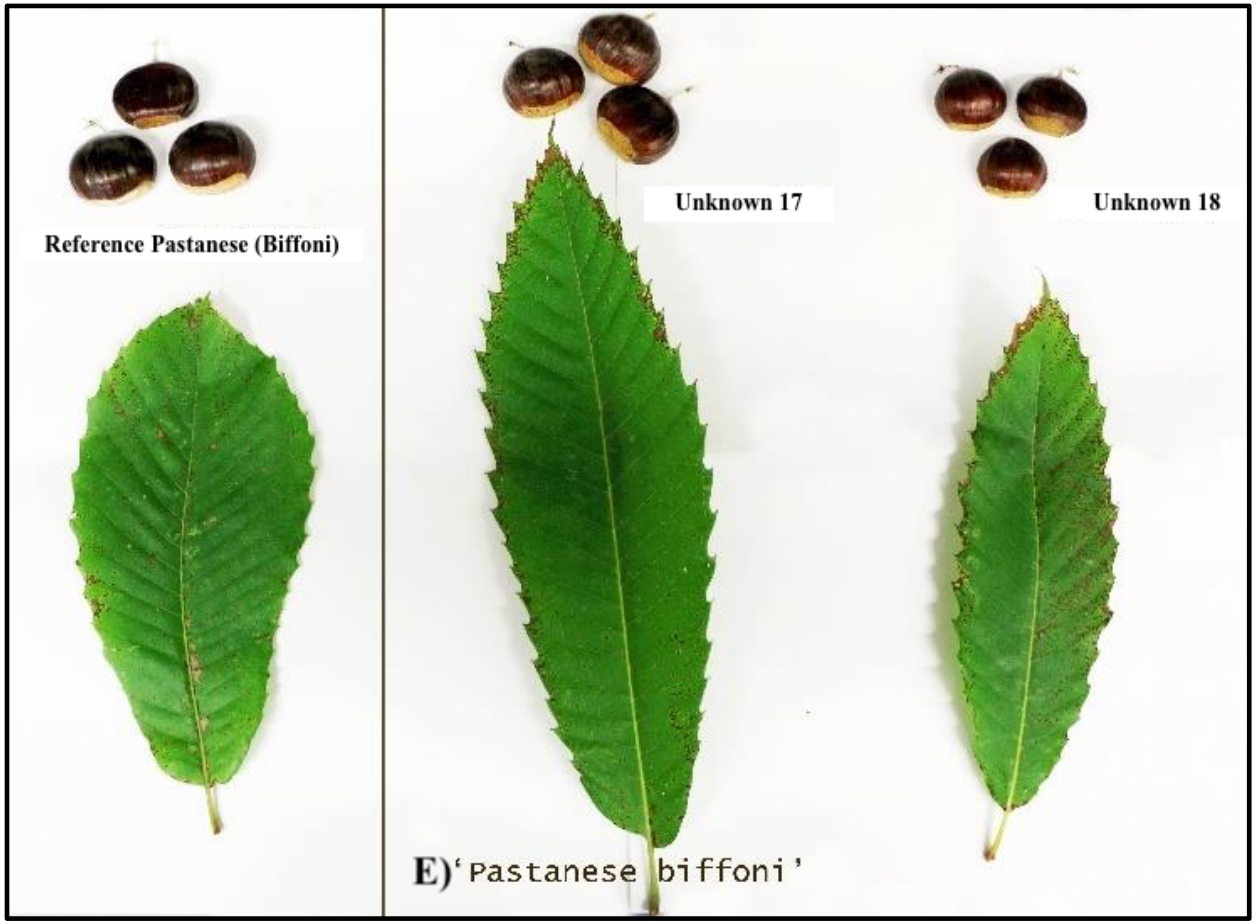

Figure 3.3 E: Comparison of the reference Chestnuts Pastanese with all the other cultivars represent not-true-to-type identity.

Table 3.3: Descriptors list for the fruit phenotypic assessment of the 14 chestnut varieties from the Didactic and Experimental Park of Granaglione.

Bovalghe

Plant

High vigor, expanded habitus, intermediate sprouting, medium and constant productivity.

Fruit

Length of spines (mm)	medium (10 mm)
Number of nuts per burr	3
Ripening time	early
N° fruits per KG and SIZE (number of nuts per kg)	140-160 frutti/kg very small
Color	brown
Shape	trasverse - brad ellipsoid
Hairiness	present (torch and downward)
Embryony	poly- embrionic
Degree of penetration of seed coat into embryo	strong
Hilum	medium

Castel del Rio

Plant

Medium vigor, expanded habitus, astamine (wild and other local chestnut varieties pollinizer), high productivity.

Fruit

Length of spines (mm)	medium (12mm)
Number of nuts per burr	3
Ripening time	early - medium
N° fruits per KG and SIZE (number of nuts per kg)	70-75 frutti/kg big
Color	reddish - brown
Shape	transverse - broad ellipsoid
Hairiness	present (torch and downward)
Embryony	mono- embryonic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Castione

Plant

High vigor, expanded habitus, sprouting between 15 April and 30 April, variable productivity.

Fruit

Length of spines (mm)	medium (14.9mm)
Number of nuts per burr	3
Ripening time	medium -late
N° fruits per KG and SIZE (number of nuts per kg)	45-120 frutti/kg medium
Color	brown - dark brown
Shape	trasverse- ellipsoid
Hairiness	present (torch)
Embryony	mono- embryonic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Centa S. Nicolò

Plant

High vigor, expanded habitus, sprouting between 15 April and 30 April, variable productivity.

Fruit

Length of spines (mm)	medium (10 mm)
Number of nuts per burr	3
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	45-120 frutti/kg medium
Color	reddish light brown
Shape	trasverse- broad ellipsoid
Hairiness	present (torch)
Embryony	mono-embrionic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Ceppa

Plant

High vigor, expanded habitus, intermediate sprouting, constant productivity.

Fruit

Length of spines (mm)	long (15 mm)
Number of nuts per burr	2 to 3
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	90 frutti/kg
	small
Color	dark brown
Shape	globose
Hairiness	present (torch)
Embryony	mono-embryonic
Degree of penetration of seed coat into embryo	strong
Hilum	small

Drena

Plant

High vigor, expanded habitus, sprouting between 15 April and 30 April, variable productivity.

Fruit

Length of spines (mm)	medium (14.5mm)
Number of nuts per burr	3
Ripening time	medium -late
N° fruits per KG and SIZE (number of nuts per kg)	45-120 frutti/kg medium
Color	reddish - brown
Shape	trasverse- ellipsoid
Hairiness	present (torch)
Embryony	mono- embryonic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Lisanese

Plant

High vigor, expanded habitus, intermediate sprouting, variable productivity.

Fruit

Length of spines (mm)	medium (14.5mm)
Number of nuts per burr	3
Ripening time	medium -late
N° fruits per KG and SIZE (number of nuts per kg)	45-120 frutti/kg medium
Color	dark - brown
Shape	globose
Hairiness	present (torch)
Embryony	mono- embryonic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Pastanese (Biffoni)

Plant

High vigor, expanded habitus, intermediate sprouting, normal productivity.

Fruit

Length of spines (mm)	medium (14.5mm)
Number of nuts per burr	4
Ripening time	medium - late
N° fruits per KG and SIZE (number of nuts per kg)	110-150 frutti/kg very small
Color	dark brown
Shape	globose
Hairiness	present (spread all over)
Embryony	mono-embryonic
Degree of penetration of seed coat into embryo	strong
Hilum	medium

Pastinese

Plant

High vigor, expanded habitus, intermediate sprouting, high productivity.

Fruit

Length of spines (mm)	medium (14 mm)
Number of nuts per burr	3
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	70-80 frutti/kg medium
Color	brown
Shape	trasverse - broad ellipsoid
Hairiness	present (torch and downward)
Embryony	poly-embrionic
Degree of penetration of seed coat into embryo	medium
Hilum	big

Pelosa

Plant

High vigor, expanded habitus, intermediate sprouting, low productivity.

Fruit

Length of spines (mm)	medium (11.9mm)
Number of nuts per burr	3 to 4
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	140-160 frutti/kg small
Color	blackish brown
Shape	globose
Hairiness	present (spread all over)
Embryony	mono-embryonic
Degree of penetration of seed coat into embryo	strong
Hilum	large

Roncegno

Plant

High vigor, expanded habitus, sprouting between 15 April and 30 April, variable productivity.

Fruit

Length of spines (mm)	medium (14.9 mm)
Number of nuts per burr	3
Ripening time	medium -late
N° fruits per KG and SIZE (number of nuts per kg)	45-120 frutti/kg medium
Color	reddish brown
Shape	trasverse-ellipsoid
Hairiness	present (torch)
Embryony	mono-embrionic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Sborgà

Plant

High vigor, expanded habitus, intermediate sprouting, medium productivity.

Fruit

Length of spines (mm)	medium (9.3mm)
Number of nuts per burr	3
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	95 frutti/kg medium
Color	brown
Shape	trasverse- ellipsoid
Hairiness	present (torch)
Embryony	mono- embryonic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

Svizzera

Plant

High vigor, expanded habitus, intermediate sprouting, low productivity.

Fruit

Length of spines (mm)	medium (12mm)
Number of nuts per burr	3
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	130 frutti/kg small/medium
Color	Dark-brown or brown
Shape	ovoid
Hairiness	present (spread all over)
Embryony	poly-embrionic
Degree of penetration of seed coat into embryo	strong
Hilum	small

Zocca

Plant

Medium vigor, expanded habitus, astamine (wild and other local chestnut varieties pollinizer), high productivity.

Fruit

Length of spines (mm)	long (15.2 mm)
Number of nuts per burr	2 to 3
Ripening time	medium
N° fruits per KG and SIZE (number of nuts per kg)	70-80 frutti/kg medium
Color	reddish brown
Shape	trasverse - broad ellipsoid
Hairiness	present (torch)
Embryony	mono-embrionic
Degree of penetration of seed coat into embryo	weak
Hilum	medium

3.5.1 Cluster analysis of wood chestnuts varieties

Finally, a cluster analysis was carried out on 32 wood samples collected from the Didactic and Experimental Park of Granaglione. In this respect, ‘Perticaccio’ cultivar was added at the analysis as reference from Pereira-Lorenzo et al., 2017 and ‘Mozza’ from the Paloneta collection (RA).

In particular, Figure 3.4 highlights the presence of four clusters (A, B, C, D) that should correspond to the varieties 'Cardaccio', 'Mozza', 'Politora' and 'Perticaccio'.

There were only a few individuals with an identical allelic profile (‘Legno 5,7,18,19’ and ‘Legno 8,16’) and only samples ‘Legno 5,7, 18,19’ resulted similar to reference ‘Perticaccio’, that is present in cluster D. They shared a similarity rate of about 85% (Figure 3.4). In particular, most of the polymorphic alleles differ from the reference ‘Perticaccio’ of a few bases, suggesting the presence of genetic variability among the trees belonging to this variety.

In Figure 3.4, ‘Legno 23’ and ‘Legno 29’ shared 50% of the allelic profile with the reference ‘Mozza’, supporting the indication that part of this plant material was derived from seeds.

These varieties were introduced in the Didactic and Experimental Park of Granaglione for their high growth capacity, good wood quality and the absence of “ring shake”. In particular, ‘Politora’ cultivar presents a high growth rate and has stems of good quality: straight, cylindrical and with small branches. ‘Cardaccio’ and ‘Perticaccio’ cultivars also have excellent growth but compared to ‘Politora’ have a lower apical dominance; on the contrary they have a more consistent and vigorous branchiness (Giannini et al., 2012).

Unfortunately, there are only a few papers in literature that investigated phenotypic characters and there is no reference work for molecular genetic analysis on the collected four wood varieties. The possibility that different genotypes with very similar characteristics may have be ascribed to a single common variety cannot be excluded.

The high genetic variability of the analyzed samples should derive from the selection of seeds for wood production purposes. To date the specific traits for each genotype are maintained. A comparison with certified reference accessions is needed to clarify this situation. In spite of this most of these trees have all the characteristics required for wood production. A challenge could be the selection of the best genotype as new promising clone for wood production.

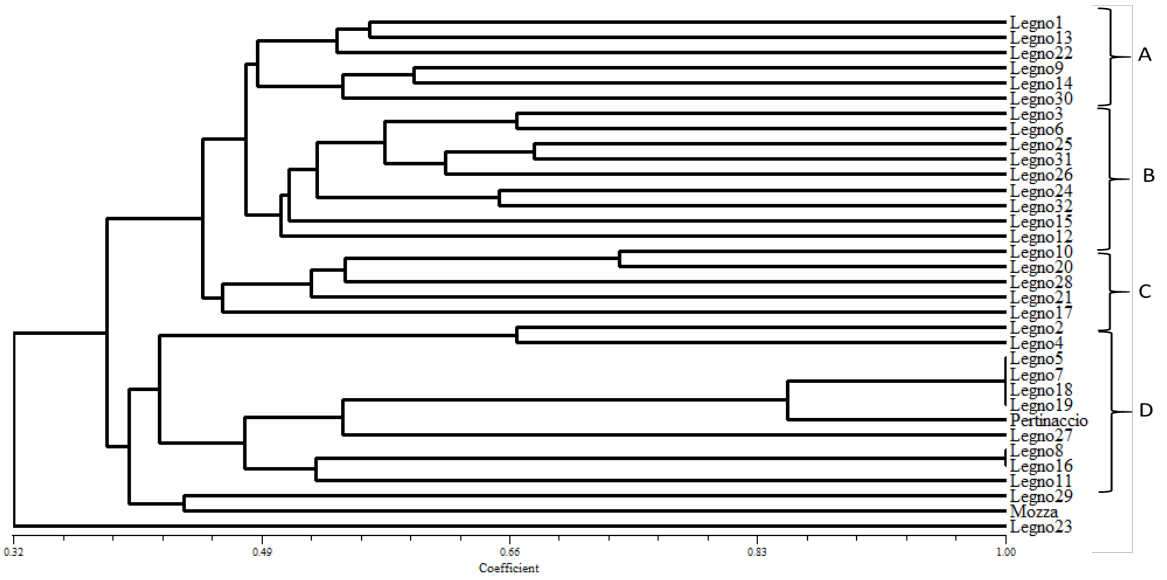

Figure 3.4: UPGMA elaborated with NTYSYS of 32 woody chestnut samples and 'Perticaccio' and 'Mozza' reference.

3.6 Conclusions

The molecular characterization carried out allowed the correct identification of the varieties mainly cultivated in the Didactic and Experimental Park of Granaglione. The identification of synonym accessions underlined the importance of verifying germplasm collections with powerful tools such as molecular markers. These tools are key to avoid redundancy in collections and the to respond to challenges of varietal certification for propagation in nurseries. This work also highlighted the techniques and practices necessary to promote the maintenance of chestnut biodiversity by inserting varieties at risk of genetic erosion. In addition, the presence of high genetic variability among the wood-specialized chestnut trees should be use in future for the selection of new promising clones for a sustainable production of wood from chestnuts.

Project partners made themselves available to house and store plants at risk of genetic erosion that will be reintroduced taking into account the pedoclimatic characteristics of the company's fields.

3.7 Bibliography

Bolvansky, M.; Mendel, L. (2001). Revised descriptor list for the evaluation of genetic resources of European chestnut (*Castanea sativa* Mill.). Forest Snow and Landscape Research, 76(3), 439–444.

Cipriani, G.; Spadotto, A.; Jurman, I.; Gaspero, G.; Di Crespan, M.; Meneghetti, S.; Testolin, R. (2010). The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages and reveals a large admixture amongst varieties of different geographic origin. Theoretical and Applied Genetics, 121(8), 1569–1585. https://doi.org/10.1007/s00122-010-1411-9

Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. Ecology, 26(3), 297–302. https://doi.org/10.2307/1932409

Fineschi, S. (1988) - Genetics of Chestnut (*Castanea sativa* Mill.) II. Uniformity of enzyme phenotypes in grafted tree orchards. Silvae Genetica 37, 82-83.

Giannini, R.; Maltoni, A.; Mariotti, B.; Paffetti, D.; Tani, A.; Travaglini, D. (2014) Valorizzazione della produzione legnosa dei boschi di castagno. L'Italia Forestale e Montana, 69 (6):307-317. http://dx.doi.org/10.4129/ifm.2014.6.01

Gross, B.L.; Volk, G.M.; Richards, C.M.; Forsline, P.L.; Fazio, G.; Chao, C.T. (2012) Identification of "duplicate" accessions within the USDA-ARS National Plant Germplasm System Malus collection. J Am Soc Hortic Sci 137:333–342

Maguire, T. L.; Collins, G. G.; & Sedgley, M. (1994). A modified CTAB DNA extraction procedure for plants belonging to the family proteaceae. Plant Molecular Biology Reporter, 12(2), 106–109. https://doi.org/10.1007/BF02668371

Pereira-Lorenzo, S.; Ramos-Cabrer, A. M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, M. B.; … Martín, Á. (2017). Database of European chestnut cultivars and definition of a core collection using simple sequence repeats. Tree Genetics and Genomes, 13(5). https://doi.org/10.1007/s11295-017-1197-x

Rohlf, F. J. (2004). NTSYSpc-Numerical Taxonomy and Multivariate Analysis System. 44. Retrieved from http://www.exetersoftware.com/downloads/ntsysguide21.pdf

UPOV. (1989). Guidelines for the conduct of tests for distinctess, homogeneity and stability: Chestnut (*Castanea sativa* Mill.). International Union for the Protection of New Varieties of Plants, 3–4. https://doi.org/10.1017/CBO9781107415324.004

CHAPTER 4: Genetic diversity among European chestnuts and estimation of the related gene flow

4.1 Abstract

The only native species of the *Castanea* genus in Europe is C*astanea sativa* Mill., a widely spread and important multipurpose tree in the Mediterranean area (fruit, wood, shelter for hives). A total of 319 unique genotypes were analyzed with 16 SSRs with the aim of expanding the genetic knowledge of chestnut trees and to promote the traceability of the local products. A Bayesian approach combined with the Markov Chain Monte Carlo (MCMC) simulation method was applied. The study revealed the existence of two genetically and, to a large extent, geographically distinct groups of chestnut populations (C1 and C2). Analysis of molecular variance (AMOVA) showed a high level of genetic diversity within populations (92%), rather than among populations (8%).

The STRUCTURE analysis revealed also a subdivision for K=3 and K=4, with a clear separation between the cultivars of the North and South of Spain from the Italian varieties. Our results confirmed a common genetic structure between chestnut populations from South of Spain and South of Italy that was the result of historical events and long human impact.

Moreover, the gene flow between cultivated chestnut trees and wild chestnuts affected their genetic structure: our results revealed that wild-grown introgression in chestnuts depends on the proximity of chestnut orchards and natural populations. The percentage of molecular variance was 99% within populations, indicating an absent genetic differentiation between wild and cultivated chestnut trees.

Keyword: *Castanea sativa* Mill., SSR, Genetic diversity, Germoplasm conservation, Structure analysis

4.2 Introduction

The *Castanea sativa* Mill. is an autochthonous species in Europe that has seen its geographical distribution influenced at first by the climatic conditions of a glacial period, and then by the interest that humans have placed on the species (Conedera et al., 2004).

Genotypes differing for numerous characters of fruit resistance to biotic and abiotic factors have been selected over the centuries based on phenotypic factors (Silvanini et al., 2011).

In the last 20 years, numerous studies have explored chestnut genetics (Villani et al., 1999; Martin et al., 2012; Mattioni et al., 2013; Lusini et al., 2014; Mattioni et al., 2013; Pereira-Lorenzo et al., 2010, 2011, 2020; Bouffartigue et al., 2019). In Italy and France, a particularly good-quality group of cultivars called Marroni was selected and cultivated for commercial purposes. This has given rise to a very complex structure of the chestnut culture with a fragmentary distribution throughout Europe (Conedera et al. 2004), characterized by the existence of a considerable number of different chestnut cultivars over the centuries (Pitte, 1986; Conedera et al., 1996).

The progressive molecular characterization of germplasm aims to reduce the number of genotypes, eliminating synonymies/homonymies, to identify the unique genotypes (core collection) (Ciocchini et al., 2016; Pereira-Lorenzo et al., 2017; Mellano et al., 2018).

Moreover, expanding the tree genetic knowledge in each country allows to create a common dataset to evaluate the traceability of the local products, regulated by PGI (Protected Geographical Indication) rules emanated by the European Union (Neri et al., 2010; Fideghelli, 2016). This knowledge is essential to preserve the genetic resources indefinitely, and to manage and exchange plant material at the international level (Hummer et al., 2015).

The main aims of this research were: a) to contributed to the expansion of the European Genetic Dataset based on the reference SSR for the identification of chestnut cultivars; b) to study the genetic structure in order to define the historical connections that occurred in the past. This is important to guarantee their identification and preservation; c) to define the fixation index between wild chestnuts and cultivated local varieties.

Previous studies (Fernández-Cruz and Fernández-López, 2016; Mattioni et al., 2013 and 2017) reported a strong geographical structure in wild populations of Southern Europe (in Italy, Spain, Greece and Turkey). These findings agree with evidence of spontaneous establishment originating from the Last Glacial Maximum refugia in the north of the Iberian, Italian and Balkan peninsulas (Krebs et al., 2004). In addition (Bouffartigue et al.; 2020; Pereira-Lorenzo et al., 2017 and 2019; Mattioni et al., 2008), compared naturalized, woody plant (coppice) and cultivated populations in Italy, Greece, Spain, the UK and France. The results showed low differences in within-population genetic parameters between cultivated varieties and wild chestnuts which may be the result of a long-term management techniques.

Recently, Pereira-Lorenzo et al. 2019 reported the index of fixation of genotypes by grafting from spontaneous chestnuts, and the results suggested a possible lack of genetic structure between wild and cultivated chestnuts; however, we they compared cultivated chestnut with the oldest wild populations (no grafted giant chestnuts), no differences were found between both genetic groups, with domesticated chestnut keeping most of the diversity found in the wild. Moreover, the results found in France (Bouffartigue et al., 2020) indicated a genetic structure affected by natural events, such as the recolonization after the last glaciation and by historical human processes, concluding that it is possible to suggest a common origin of the most part of French varieties with the Iberian Peninsula and the association of the Italian gene pool with the South-East France.

4.3 Materials and methods

4.3.1 Plant material

A total of 500 wild and grafted chestnut trees, corresponding to 319 chestnut unique genotypes, were sampled in 4 different countries: 155 in Italy, 154 in Spain, 4 in Portugal and 6 in France.

A total of 138 out of the 319 unique genotypes were from northern Italy, mainly in the area of the Tuscan-Emilian Apennines in the Emilia-Romagna. In particular, 111 samples were collected from different chestnuts collection camps of the Emilia-Romagna region (Didattic and Experimental Park of Granaglione, Collection of Zocca and Collection of Paloneta – Brisighella), 3 were harvested in the Trentino Alto Adige region and 24 from four private companies specialized in chestnut production (La Martina in Monghidoro, Bologna; and Tizzano in Zocca, Modena; Canovi and Teggiolina in Reggio-Emilia). The other 21 cultivars were previously analyzed by the University of Bologna (Alessandri et al., 2020).

In addition, 50 samples were from different regions of Spain (Galicia, Andalucía, Asturias, Castilla-León, Canary Islands, Extremadura, Cantabria) and 2 from Portugal.

Furthermore, 4 representative samples of Castanea pumila, C. crenata, C. mollissima and Volos cultivars were added to the analysis to get information on interspecific hybrids (Euro-Japanese) versus the European varieties.

The 138 unique genotypes identified in northern Italy were standardized with the SSR profiles of the unique accessions taken from the work of Pereira-Lorenzo (2017) for a total of 319 samples.

From the European dataset published by Pereira-Lorenzo et al. (2017), a total of 128 European cultivars were added to the analysis. All accessions analyzed are presented in the Table 4.1.

Table 4.1: Unique genotypes analyzed for a total of 319 samples.

Cultivar Prime Name	Synonyms	Number of accessions	Accessions analyzed in previous studies	Region	Country
125-10 TG		**1**	i	Galicia	Spain
324-29 TG		**1**	i	Galicia	Spain
Abada		**1**	a	Galicia	Spain
Alcobilla 6		**1**	u	Castilla-León	Spain
Alcobilla 8		**1**	u	Castilla-León	Spain
Alcobilla1		**1**	u	Castilla-León	Spain
Alcobilla10		**1**	u	Castilla-León	Spain
Alcobilla12		**1**	u	Castilla-León	Spain
Amarela		**4**	a	Galicia	Spain
Amarelante1		**5**	a	Galicia	Spain
Amola 1		**1**	u	Galicia	Spain
Antes De Villasumil 1		**1**	u	Castilla-León	Spain
Antigua		**1**	l	Extremadura	Spain
Arial		**1**	a	Galicia	Spain
Armeiriz2		**1**	u	Galicia	Spain
Armentina		**1**	l	Canary Islands	Spain
Baamonde 1 No Capilla		**1**	u	Galicia	Spain
Beira Valente 2		**1**	u	Trás-os-Montes	Portugal
Beira Valente 3		**1**	u	Trás-os-Montes	Portugal
Bermello		**2**	a	Galicia	Spain
Biancherina1		**1**	u	Emilia-Romagna	Italy
Biancherina2		**1**	u	Emilia-Romagna	Italy
Bianchina3		**1**	u	Emilia-Romagna	Italy
Blanca Canarias		**1**	l	Canary Islands	Spain
Blanca Galicia		**1**	a	Galicia	Spain
Boroñona		**1**	a	Asturias	Spain
Bouche de Betizac		**1+1**	i , u	Corrèze	France
Bracalla		**3**	b	Piemonte	Italy
C. Crenata		**1**	u	Emilia-Romagna	Giappone
C. Mollissima		**1**	u	Emilia-Romagna	Cina
C. Pumila		**1**	u	Emilia-Romagna	America
CA-15		**1**	i	INRA- Bordeaux	France
Cachero		**2**	l	Canary Islands	Spain
Calambres		**1**	l	Castilla-León	Spain
Calarese	Garfagnina	**2+1**	s , u	Emilia-Romagna	Italy
Calva2		**1**	a	Galicia	Spain
Calva3		**1**	a	Galicia	Spain
Calvotera		**1**	l	Extremadura	Spain
Camberoune		**1**	i	Dordogne	France
Campano		**4**	l	Castilla-León	Spain
Campilla		**1**	a	Galicia	Spain
Capilla1		**1**	a , e	Andalucía	Spain
Capilla2		**1**	a , e	Andalucía	Spain
Caprarola		**1**	u	Emilia-Romagna	Italy
Carrarese		**2**	s	Emilia-Romagna	Italy
Carrelao 12		**1**	u	Galicia	Spain
Castagna Venostana S1		**1**	u	Trentino Alto Ac	Italy
Castiglion dei Pepoli		**1**	u	Emilia-Romagna	Italy
Cedo		**1**	a	Galicia	Spain
Ceppa		**5**	s	Emilia-Romagna	Italy
Cerdedelo B1		**1**	u	Galicia	Spain
Cerdedelo B3		**1**	u	Galicia	Spain
Cerdedelo2 Das Viñas 1		**1**	u	Galicia	Spain
Cerredo		**3**	a	Galicia	Spain
Chaguazoso Cementerio 2		**1**	u	Galicia	Spain
Chaguazoso Fonte 1		**1**	u	Galicia	Spain
Chamberga1		**14**	a	Asturias	Spain
Chiusa Pesio1		**3**	b	Piemonte	Italy
Chiusa Pesio2		**1**	b	Piemonte	Italy
Colorada		**1**	e	Andalucía	Spain
Comisaria Pelona	Mojinegra	**2**	e	Andalucía	Spain
Comisaria1		**2**	a , b	Andalucía	Spain
Comisaria2		**3**	a , b	Andalucía	Spain
Courela		**1**	a	Galicia	Spain
De Pablo		**1**	l	Extremadura	Spain
Denia De Onís		**1**	u	Asturias	Spain
Doney 1 Sanabria		**1**	u	Castilla-León	Spain
Entrambosrios 1		**1**	u	Galicia	Spain
Famosa		**2**	a	Galicia	Spain
Galega		**1**	a	Asturias	Spain
Gallega		**1**	l	Andalucía	Spain
Galliciana		**1**	a	Asturias	Spain
Garfagnina	Tosca, Zocca21, Legno17	**6 + 5**	s , u	Emilia-Romagna	Italy
Garrida		**1**	a	Galicia	Spain
Garrone		**1**	u	Emilia-Romagna	Italy
Garrone Rosso		**1**	u	Piemonte	Italy
Grua		**1**	a	Asturias	Spain
Injerta Bierzo		**1**	a	Castilla-León	Spain
Injerta Gorda		**1**	l	Extremadura	Spain
Injerta Guadalupe		**1**	l	Extremadura	Spain
Injerta Roja		**1**	l	Extremadura	Spain
Injerta Tío Sabino		**1**	l	Extremadura	Spain
Inserta	Curciaspeciale	**4**	b	Calabria	Italy
Inxerta		**2**	a	Galicia	Spain
Judia		**1**	c	Trás-os-Montes	Portugal
La Pesanca Riofabar 2		**1**	u	Asturias	Spain
Laga		**5**	c	Andalucía	Spain
Laguilla		**1**	e	Andalucía	Spain
Las Caldas, Caces		**1**	u	Asturias	Spain
Lebre		**1**	a	Galicia	Spain
Legno 1		**1**	u	Emilia-Romagna	Italy
Legno 10		**1**	u	Emilia-Romagna	Italy
Legno 11		**1**	u	Emilia-Romagna	Italy
Legno 12		**1**	u	Emilia-Romagna	Italy
Legno 13		**1**	u	Emilia-Romagna	Italy
Legno 14		**1**	u	Emilia-Romagna	Italy
Legno 15		**1**	u	Emilia-Romagna	Italy
Legno 16		**1**	u	Emilia-Romagna	Italy
Legno 2		**1**	u	Emilia-Romagna	Italy
Legno 20		**1**	u	Emilia-Romagna	Italy
Legno 21		**1**	u	Emilia-Romagna	Italy
Legno 216		**1**	u	Emilia-Romagna	Italy
Legno 22		**1**	u	Emilia-Romagna	Italy
Legno 23		**1**	u	Emilia-Romagna	Italy
Legno 24		**1**	u	Emilia-Romagna	Italy
Legno 25		**1**	u	Emilia-Romagna	Italy
Legno 26		**1**	u	Emilia-Romagna	Italy
Legno 27		**1**	u	Emilia-Romagna	Italy
Legno 28		**1**	u	Emilia-Romagna	Italy
Legno 29		**1**	u	Emilia-Romagna	Italy
Legno 3		**1**	u	Emilia-Romagna	Italy
Legno 30		**1**	u	Emilia-Romagna	Italy
Legno 31		**1**	u	Emilia-Romagna	Italy
Legno 32		**1**	u	Emilia-Romagna	Italy
Legno 4		**1**	u	Emilia-Romagna	Italy
Legno 6		**1**	u	Emilia-Romagna	Italy
Legno 7		**1**	u	Emilia-Romagna	Italy
Legno 9		**1**	u	Emilia-Romagna	Italy
Lisanese		**11**	s	Emilia-Romagna	Italy
Llanisca		**1**	a	Asturias	Spain
Loglia		**1**	s	Emilia-Romagna	Italy
Loiola		**2**	s	Emilia-Romagna	Italy
Longal	Injerta from Hervás	**5**	a , c	Galicia, Extrema	Spain
Loura		**1**	a	Galicia	Spain
Lucente1		**1**	b	Campania	Italy
Lucente2		**1**	b	Campania	Italy
Luguesa		**6**	a	Galicia	Spain
Madonna		**1**	u	Emilia-Romagna	Italy
Madonna		**4+4**	b , u	Piemonte, Camp	Italy
Majadas		**1**	l	Castilla-León	Spain
Mamma	Inserta, Tempestiva	**7**	b	Calabria	Italy
Mand al Broc		**1**	u	Emilia-Romagna	Italy
Marrone Comballe		**2**	l	South-East of Fr	France
Marrone di Knoll		**1**	u	Trentino Alto Ac	Italy
Marrone di Melfi	Marrone Roccadaspide, Mercogliana	**8**	b	Basilicata, Camp	Italy
Marrone di Unterganzner		**1**	u	Trentino Alto Ac	Italy
Marrone Fiorentino	Borgovelino, Caprese Michel, Caprese Michelangelo, Castel del Rio, Castione, Centa di S. Nicolò, Chiusa Pesio, Città di Castello, Drena, Gaggio Montano, Gavignano, Locale Paloneta, M. Isola d'Elba, Marradi, Marrone Val Susa, Montemarano, Napoletana, Palazzo del pero, Pastonese, Pitigliano, Riggiolana, Roccamonfina, Roncegno, Sborgà, Tempurina, Zocca	**23+67+40**	b , s , u	Lazio, Toscana, I	Italy
Marrone Roccadaspide		**2**	b	Campania	Italy
Martahiña		**1**	c	Trás-os-Montes	Portugal
Martina 5		**1**	u	Emilia-Romagna	Italy
Mascherina		**2**	s	Emilia-Romagna	Italy
Massangaia	Madonna, Mand al broc	**2 + 4**	s , u	Emilia-Romagna	Italy
Matildico 1		**1**	u	Emilia-Romagna	Italy
Matildico 1.1		**1**	u	Emilia-Romagna	Italy
Matildico 10		**1**	u	Emilia-Romagna	Italy
Matildico 11		**1**	u	Emilia-Romagna	Italy
Matildico 12		**1**	u	Emilia-Romagna	Italy
Matildico 13		**1**	u	Emilia-Romagna	Italy
Matildico 2		**1**	u	Emilia-Romagna	Italy
Matildico 2.1		**1**	u	Emilia-Romagna	Italy
Matildico 3		**1**	u	Emilia-Romagna	Italy
Matildico 3.1		**1**	u	Emilia-Romagna	Italy
Matildico 4		**1**	u	Emilia-Romagna	Italy
Matildico 5		**1**	u	Emilia-Romagna	Italy
Matildico 5.1		**1**	u	Emilia-Romagna	Italy
Matildico 6		**1**	u	Emilia-Romagna	Italy
Matildico 7		**1**	u	Emilia-Romagna	Italy
Matildico 9		**1**	u	Emilia-Romagna	Italy
Mendoia-2		**1**	u	Galicia	Spain
Mercogliana		**1**	b	Campania	Italy
Miguelina		**2**	a	Asturias	Spain
Molana		**2**	s	Emilia-Romagna	Italy

Cultivar Prime Name	Synonyms	Number of accessions	Accessions analyzed in previous studies	Region	Country
Mollar1		**2**	d	Canary Islands	Spain
Mollar2		**1**	d	Canary Islands	Spain
Mollar3		**1**	d	Canary Islands	Spain
Mondistollo		**1**	u	Emilia-Romagna	Italy
Monfortina		**1**	a	Galicia	Spain
Montagne		**1**	i	Dordogne	France
Montemarano	Montella, Zocca21	**2+2**	s , u	Emilia-Romagna	Italy
Mozza		**1**	u	Emilia-Romagna	Italy
Mulata1		**1**	d	Canary Islands	Spain
Mulata2		**1**	d	Canary Islands	Spain
Nacerona 1 Ocejo		**1**	u	Cantabria	Spain
Necas 2		**1**	u	Galicia	Spain
Necas 3		**1**	u	Galicia	Spain
Negral		**5**	a	Galicia, Castilla-l	Spain
Negrera		**1**	a	Asturias	Spain
Nespereira4		**1**	u	Galicia	Spain
País		**4**	a	Galicia	Spain
Parede		**6**	a	Asturias, Castilla	Spain
Parruquina		**1**	a	Asturias	Spain
Pastanese	Matildico, Pastonese	**8 + 3**	s , u	Emilia-Romagna	Italy
Peixeroos 1 Raiz		**1**	u	Galicia	Spain
Peixeroos 2		**1**	u	Galicia	Spain
Pelado		**1**	a	Galicia	Spain
Pelona Andalucia		**2**	a , b	Andalucía	Spain
Pelona Asturias		**1**	a	Asturias	Spain
Pelona Avila		**1**	i	Castilla-León	Spain
Pelosa		**4**	s	Emilia-Romagna	Italy
Peluda Tardía2		**1**	e	Andalucía	Spain
Pertinaccio	Cardaccio, Garrone Rosso	**8**	e	Toscana, Emilia	Italy
Pesaguero 1		**1**	u	Cantabria	Spain
Pesaguero 5		**1**	u	Cantabria	Spain
Pesaguero 8		**1**	u	Cantabria	Spain
Petra		**1**	u	Emilia-Romagna	Italy
Picona		**1**	a , c	Galicia	Spain
Pilonga		**1**	a , e	Andalucía	Spain
Pilonga2		**2**	e , g	Andalucía	Spain
Pistolese		**1**	s	Emilia-Romagna	Italy
Piusela 1		**1**	u	Emilia-Romagna	Italy
Piusella		**2 + 3**	s , u	Emilia-Romagna	Italy
Planta Alajar	Comisaria Rubia	**6**	a , e	Andalucía	Spain
Planta Alajar		**1**	u	Andalucía	Spain
Pollayo 1		**1**	u	Cantabria	Spain
Pollayo 5		**1**	u	Cantabria	Spain
Porteliña		**1**	a	Galicia	Spain
Portuguesa		**1**	e	Andalucía	Spain
Precoce Migoule	Pastonese(Z21)	**2+1**	s , u	Emilia-Romagna	Italy
Puga		**1**	a	Galicia	Spain
Punghenta		**1**	u	Emilia-Romagna	Italy
Raigona1		**1**	a	Galicia	Spain
Raigona2		**1**	a	Galicia	Spain
Rapada		**4**	a , c	Galicia	Spain
Rapuca1		**1**	a	Asturias	Spain
Rapuca2		**1**	a	Asturias	Spain
Ribeira 1		**1**	u	Galicia	Spain
Ribeira 3		**1**	u	Galicia	Spain
Riggiola1		**2**	b	Calabria	Italy
Riggiola2		**3**	b	Calabria	Italy
Rossola		**1**	u	Emilia-Romagna	Italy
San Roman De Sanabria 1		**1**	u	Castilla-León	Spain
Santa Eufemia (Baños Molgás) 1		**1**	u	Galicia	Spain
Sergude		**1**	a	Galicia	Spain
Serodia		**1**	a	Galicia	Spain
Sfronzola		**1**	u	Emilia-Romagna	Italy
Sietepernadas		**1**	d	Canary Islands	Spain
Stanco 2		**1**	u	Emilia-Romagna	Italy
Super		**1**	l	Extremadura	Spain
Svizzera		**4**	s	Emilia-Romagna	Italy
Tardía Clara		**1**	e	Andalucía	Spain
Tardía Oscura		**1**	e	Andalucía	Spain
Temprana Genalguacil		**1**	e	Andalucía	Spain
Temprana Jubrique		**1**	e	Andalucía	Spain
Temprana1		**1**	a , e	Andalucía	Spain
Tempuriva	Montemarano	**4**	b	Campania, Piem	Italy
TG 90025		**1**	i	Galicia	Spain
Tizzano 4		**1**	u	Emilia-Romagna	Italy
Tizzano 5		**1**	u	Emilia-Romagna	Italy
Tomasa		**1**	a , e	Andalucía	Spain
Tosca		**1**	u	Emilia-Romagna	Italy
Trogais 1		**1**	u	Galicia	Spain
Unknow 28		**1**	u	Emilia-Romagna	Italy
Unknow 31		**1**	u	Emilia-Romagna	Italy
Unknow 50		**1**	u	Emilia-Romagna	Italy
Unknown 1		**1**	u	Emilia-Romagna	Italy
Unknown 10		**1**	u	Emilia-Romagna	Italy
Unknown 11		**1**	u	Emilia-Romagna	Italy
Unknown 12		**1**	u	Emilia-Romagna	Italy
Unknown 13		**1**	u	Emilia-Romagna	Italy
Unknown 14		**1**	u	Emilia-Romagna	Italy
Unknown 15		**1**	u	Emilia-Romagna	Italy
Unknown 16		**1**	u	Emilia-Romagna	Italy
Unknown 17		**1**	u	Emilia-Romagna	Italy
Unknown 18		**1**	u	Emilia-Romagna	Italy
Unknown 19		**1**	u	Emilia-Romagna	Italy
Unknown 2		**1**	u	Emilia-Romagna	Italy
Unknown 20		**1**	u	Emilia-Romagna	Italy
Unknown 21		**1**	u	Emilia-Romagna	Italy
Unknown 22		**1**	u	Emilia-Romagna	Italy
Unknown 23		**1**	u	Emilia-Romagna	Italy
Unknown 24		**1**	u	Emilia-Romagna	Italy
Unknown 25		**1**	u	Emilia-Romagna	Italy
Unknown 26		**1**	s	Emilia-Romagna	Italy
Unknown 27		**1**	u	Emilia-Romagna	Italy
Unknown 29		**1**	u	Emilia-Romagna	Italy
Unknown 3		**1**	u	Emilia-Romagna	Italy
Unknown 30		**1**	u	Emilia-Romagna	Italy
Unknown 32		**1**	u	Emilia-Romagna	Italy
Unknown 33		**1**	u	Emilia-Romagna	Italy
Unknown 34		**1**	u	Emilia-Romagna	Italy
Unknown 35		**1**	u	Emilia-Romagna	Italy
Unknown 36		**1**	u	Emilia-Romagna	Italy
Unknown 37		**1**	u	Emilia-Romagna	Italy
Unknown 38		**1**	u	Emilia-Romagna	Italy
Unknown 39		**1**	u	Emilia-Romagna	Italy
Unknown 4		**1**	u	Emilia-Romagna	Italy
Unknown 40		**1**	u	Emilia-Romagna	Italy
Unknown 41		**1**	u	Emilia-Romagna	Italy
Unknown 42		**1**	u	Emilia-Romagna	Italy
Unknown 43		**1**	u	Emilia-Romagna	Italy
Unknown 44		**1**	u	Emilia-Romagna	Italy
Unknown 45		**1**	u	Emilia-Romagna	Italy
Unknown 46		**1**	u	Emilia-Romagna	Italy
Unknown 47		**1**	u	Emilia-Romagna	Italy
Unknown 48		**1**	u	Emilia-Romagna	Italy
Unknown 49		**1**	u	Emilia-Romagna	Italy
Unknown 5		**1**	u	Emilia-Romagna	Italy
Unknown 6		**1**	u	Emilia-Romagna	Italy
Unknown 7		**1**	u	Emilia-Romagna	Italy
Unknown 8		**1**	u	Emilia-Romagna	Italy
Unknown 9		**1**	u	Emilia-Romagna	Italy
Unknown1.s		**1**	d	Canary Islands	Spain
Unknown2.s		**1**	g	Galicia	Spain
Unknown3.s		**1**	g	Galicia	Spain
Unknown4.s		**1**	g	Galicia	Spain
Unknown5.s		**1**	f	Asturias	Spain
Unknown6.s		**1**	f	Asturias	Spain
Unknown7.s		**1**	f	Asturias	Spain
Unknown8.s		**1**	g	Galicia	Spain
Valduna1		**4**	a	Asturias	Spain
Valduna2		**1**	a	Asturias	Spain
Vazqueña		**1**	e	Andalucía	Spain
Vega Selorio Villaviciosa		**1**	u	Asturias	Spain
Vegamesada		**1**	a	Asturias	Spain
Ventura		**1**	a	Galicia	Spain
Verata		**7**	a , c , l	Extremadura	Spain
Verdale		**1**	i	Dordogne	France
Verdial		**1**	a , c , f	Asturias	Spain
Verdina		**1**	a , c , f	Asturias	Spain
Verduenga 1		**1**	u	Castilla-León	Spain
Verduenga 3 (Raiz)		**1**	u	Castilla-León	Spain
Verduenga 6 (Raiz)		**1**	u	Castilla-León	Spain
Villanueva Iglesia		**1**	u	Asturias	Spain
Villaorille, Souto Quirós 1		**1**	u	Asturias	Spain
Villaorille, Souto Quirós 2		**1**	u	Asturias	Spain
Villasumil 1		**1**	u	Castilla-León	Spain
Vime De Sanabria 1		**1**	u	Castilla-León	Spain
Volos		**1**	u	Emilia-Romagna	Grecia
Xabrega		**1**	a	Galicia	Spain

[a] Pereira Lorenzo S, Díaz Hernández MB, Ramos Cabrer AM. Use of highly discriminating

[b] Martín MA, Mattioni C, Cherubini M, Taurchini D, Villani F. Genetic characterisation of

[c] Pereira-Lorenzo S, Lourenço Costa RM, Ramos-Cabrer AM, Ciordia-Ara M, Marques Ribeiro CA,

[d] Pereira-Lorenzo S, Ríos-Mesa D, González-Díaz AJ, Ramos-Cabrer AM. Los castañeros de

[e] Martín MA, Alvarez JB, Mattioni C, Cherubini M, Villani F, Martín LM. Identification and

[f] Pereira Lorenzo S, Ramos Cabrer AM, Díaz Hernández MB, Ciordia M Características morfológicas

[g] Pereira-Lorenzo S, Pereira-Lorenzo S, Fernández-López J. Los cultivares autóctonos de castaño

i Pereira-Lorenzo S, Lourenço Costa RM, Ramos-Cabrer AM, Marques Ribeiro CA, Serra da Silva M,

l Pereira-Lorenzo S, Ramos-Cabrer AM, Barreneche T, Mattioni C, Villani F, Díaz Hernández MB,

s Alessandri S, Krznar M, Ajolfi D, Dondini L (2020) Genetic diversity of Castanea sativa accessions

[u] unpublished data

4.3.2 DNA extraction and PCR amplification of microsatellites (SSRs)

Young leaves used for DNA extraction were frozen in liquid nitrogen and stored at -80 ° C or lyophilized. The extraction was carried out on samples of 0.1-0.5 grams of fresh leaves milled in liquid nitrogen, or on 5 mg of grinder lyophilized leaves. The DNA were extracted with the CTAB protocol (Maguire et al., 1994). Genomic DNA was quantified by Nanodrop™ ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and diluted to 10 ng/µl.

Based on the work of Pereira-Lorenzo et al. 2017, genomic SSR loci belonging to four series were tested divided into 5 multiplex and 2 singleplex PCRs: CsCAT (Marinoni et al., 2003), EMCs (Buck et al., 2003) and OAL (Gobbin et al., 2007) developed from *C. sativa* Mill., and QrZAG (Kampfer et al., 1998) which was developed from *Quercus robur* L.

Multiplex PCR is a system in which several markers are simultaneously amplified in the same reaction (Sint et al., 2012). The PCR reaction was performed in 10 µl final volume containing 6.45 µl of sterile H_2O, 1 µl of GeneAmp® 10X reaction buffer, 0.8 µl of $MgCl_2$, 0.25 of dNTPs and 0.1 units AmpliTaq Gold™ DNA polymerase. Amplification products were analyzed on a 3130 Genetic Analyzer capillary sequencer (Applied Biosystems, USA). The internal GeneScan TM size standard 500 LIZ (-250) was included in each sample. The allele sizes were detected using Peak Scanner TM software (Applied Biosystems). The samples collected were suitably standardized with the alleles found in the European dataset for the 16 SSR under exams. The dimensions of each alleles in the loci selected were checked with the software Peak Scanner first mentioned.

4.3.3 Genetic diversity and population genetic structure

The biodiversity of the study population was assessed by the Cervus software version 3.0.3 (Kalinowski et al., 2007): the number of alleles per locus (k), the expected and the observed heterozygosity (H_e and H_o), polymorphism information content (PIC) and the probability of allele null (F-null) were estimated. A PIC value greater than 0.7 was considered to be highly polymorphic and informative for a certain locus. The frequency of the null alleles (F-null) for each locus was calculated using the maximum likelihood (ML) estimator of Kalinowski (2007) implemented in Cervus software.

The data were organized into a square matrix where code '0' and '1' were used respectively for the absence and presence of a certain allele (the code for the missing data was 9). The genetic distance between cultivars

was calculated with Jaccard coefficient (JC) using the R software. The construction of the genetic distance dendrogram was elaborated using the Unweighted Pair-Group Method (UPGMA) by R software with the function 'hclust, method = 'average", packages ‘adegenett’

The construction of dendrogram allowed to identify the synonymies/homonymies among the accessions and defined the cultivated chestnut from the wild samples of Italian dataset.

The software STRUCTURE 2.3.3 (Pritchard et al., 2000) was used to evaluate the population structure and to calculate the estimated membership coefficients (Q-value) that indicates the membership of each individual in each cluster. This analysis was conducted with a Bayesian approach combined with the Markov Chain Monte Carlo (MCMC) simulation method and was performed using an "admixture model" and correlated allele frequencies. 30 replicate runs of STRUCTURE were performed by setting the number of clusters (K) from 1 to 15. Each run consisted of a burning period of 200.000 steps followed by 200.000 MCMC replicates, with the options use locprior=0, popinfo = 0, popflag = 0 (Pereira-Lorenzo et al., 2019; Porras-Hurtado et al, 2013).

306 samples (hybrid individuals were not considered, for a total of 13 samples) were tested for an analysis of STRUCTURE with the same condition beforementioned. A Q threshold of 0.8 was used to infer an accession to a specific cluster. The ΔK value (defined as the most probable number of clusters in the population), was calculated through Structure Harvester v.09.93 (Earl, 2012) by testing the change of the log-likelihood between K values (ΔK) as described by Evanno (2005).

4.3.4 Genetic differentiation

To validate the genetic structure revealed by the Bayesian model-based clustering, a multivariate Principal Coordinate Analysis (PCoA) was elaborated with GenAlEx (Peakall and Smouse, 2006). The PCoA representation was based on the distance measures elaborated with Jaccard coefficient two major groups and on the second subdivision of the population (K=3 and K=4), both defined by the estimates of ΔK from STRUCTURE.

A set of analysis to estimate the population differentiation was conducted under four scenarios: a) the two main groups (Cluster 1 vs Cluster 2) resulting by the Structure analysis; b) the sub- groups (K=3, K=4), c) the cultivated varieties versus the wild samples (200 vs 106, respectively) and d) the cultivated varieties and the wild samples separated in the two main cluster (C1 and C2).

Pairwise FST values were estimated for the different partitioning levels considered using GeneAlEx software; missing data were coded as 0. FST value assumes values from -1 (absent inbreeding, excess of heterozygous) to 1 (non-random reproduction, excess homozygous).

Hierarchical analysis of molecular variance (AMOVA) was implemented in the GeneAlEx program (Peakall and Smouse, 2006) to evaluate the genetic variation among and within Clusters. Tests of significance were performed using 9999 permutations within the total dataset (306 samples).

4.4 Results

4.4.1 Genetic diversity

The 16 SSRs selected showed a high level of polymorphism and discriminating power and revealed a total of 212 alleles, with an average number of 13.25 alleles per locus (Table 4.2). The average PIC was 0.735, swinging between 0.879 for CsCAT3 and 0.593 for EMCs 2. These markers appeared to be the most and least informative loci.

CsCAT 41 was known to amplify two different sites (A and B), for this reason, the CsCAT41A locus was removed from the data before the analyses (Pereira-Lorenzo et al., 2010).

CsCAT2 and EMCs38 showed a high frequency values of null alleles (0.2090 and 0.1181, respectively).

Table 4.2: The number of individuals (N), the number of alleles (k), the observed (Ho) and expected (He) heterozygosity, the polymorphic information content (PIC) and null alleles frequencies are reported for each SSR locus in *C. sativa* accessions.

Locus	k	N	Ho	He	PIC	F(Null)
Cscat 41B	12	319	0.690	0.812	0.793	0.0719
Cccat 16	12	319	0.777	0.812	0.785	0.0181
Cscat6	19	319	0.846	0.872	0.858	0.0133
Cscat1	16	319	0.596	0.619	0.601	0.0224
Cscat3	26	319	0.803	0.889	0.879	0.0509
QrZag 96	12	319	0.596	0.718	0.694	0.0828
EMCs15	7	319	0.599	0.661	0.605	0.0539
EMCs38	19	319	0.693	0.883	0.871	0.1181
EMCs2	6	319	0.624	0.661	0.593	0.0303
EMCs22	10	319	0.652	0.682	0.654	0.0202
CsCAT2	16	319	0.555	0.855	0.841	0.2090
CsCAT17	13	319	0.771	0.844	0.827	0.0426
CsCAT14	10	319	0.721	0.753	0.711	0.0153
CsCAT15	11	319	0.665	0.666	0.605	-0.0037
CsCAT8	11	319	0.727	0.843	0.821	0.0687
OAL	12	319	0.586	0.651	0.629	0.0481

4.4.2 Population genetic structure with hybrids

The genetic structure of the 319 unique genotypes was evaluated with 14 loci out of the 16 under examination; the two loci with presence of null alleles (loci: EMCs38 and CsCAT2) were removed from the STRUCTURE analysis.

In the first set of STRUCTURE analysis (Figure 4.1), the ΔK statistics gave a maximum value of K=2 (ΔK = 78.61, although a small peak of ΔK were observed for K=3 (ΔK=18.42); K =6 (ΔK=6.21) and for K =8 (ΔK=6.05).

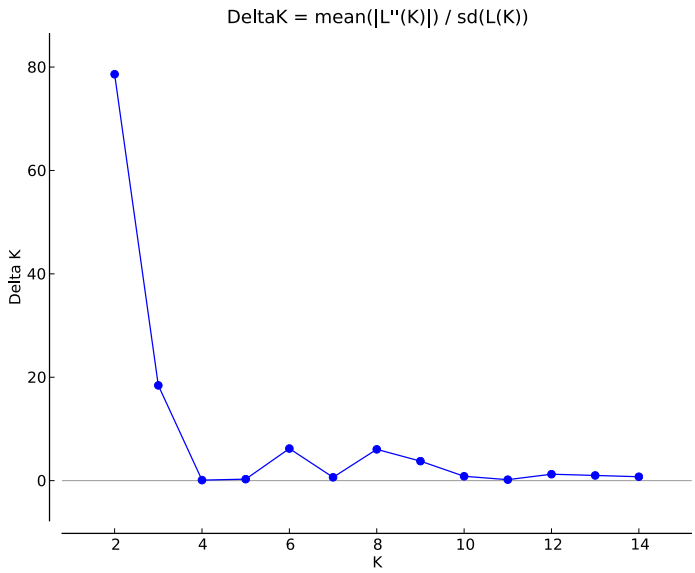

Figure 4.1. Estimates of Δ k calculated on the basis of Evanno et al. (2005) based on k-subdivision for 319 samples.

Genotypes were grouped in two main clusters with a Q > 80%: Cluster1 consisting of 102 genotypes, Cluster 2 consisting of a set of 163 genotypes (Table 4.3), with a clear distinction between Italian (Cluster2) and Spanish genotypes (Cluster1). Additionally, there were 54 admixed cultivars (Q < 80%).

We also tested K = 3 (Figure 4.2, Table 4.3), showing that the separation between the Italian and Spain cultivars remains: Cluster 1 included the Italian cultivars, mainly from Tuscan-Emilian Apennines, and Cluster 2 identified the Spanish genotypes, with little distinction between regions.

An interesting and distinct cluster (Cluster 3) emerged from the results and included the Euro-Japanese hybrids from France ('Marrone Comballe', 'Bouche de Balzac') and Spain (‘Portugesa’, ‘Gallega’, ‘Super’, ‘TG 90025’, ‘125-10 TG’ and ‘324-29 TG’), and the species of *C. pumila*, *C. mollissima*, *C. crenata* and ‘Volos’ from Greece, which clustered with the some main Italian cultivars (‘Marrone Fiorentino’; ‘Madonna’, ‘Lucente’ and ‘Inserta’) and two from Andalucía region (‘Tomasa’ and ‘Capilla’).

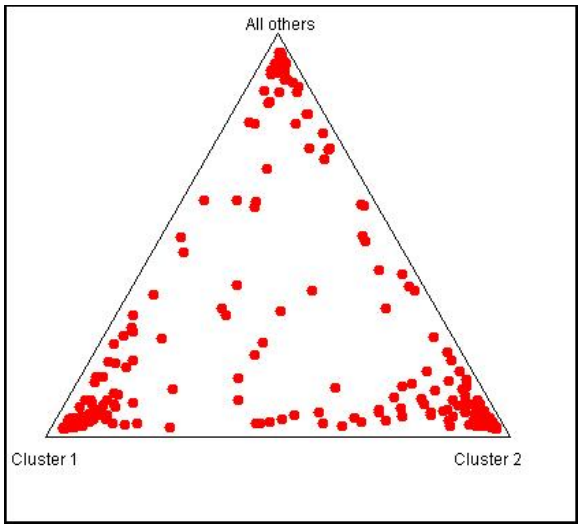

Figure 4.2: Triangle plot of 319 samples for K=3 by STRUCTURE Software.

Table 4.3: list of the different varieties collected with the origin countries, the STRUCTURE subdivision for K=2 and K=3 for 319 samples.

N° STRUCTURE	Samples	Stato	Region		Samples of Pereira-Lorenzo et al. (2017)	K=2	K=3
1	Marrone Fiorentino	Italy	Granaglione	Emilia-Romagna		Cluster 1	Cluster 3
2	Unknown 1	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
3	Unknown 2	Italy	Granaglione	Emilia-Romagna		AD	AD
4	Unknown 3	Italy	Granaglione	Emilia-Romagna		Cluster 2	AD
5	Unknown 4	Italy	Granaglione	Emilia-Romagna		AD	AD
6	Unknown 5	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
7	Unknown 6	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
8	Unknown 7	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
9	Unknown 8	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
10	Unknown 9	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
11	Unknown 10	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
12	Unknown 11	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
13	Unknown 12	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
14	Unknown 13	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
15	Unknown 14	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
16	Lisanese	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
17	Unknown 15	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
18	Unknown 16	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
19	Unknown 17	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
20	Pastanese	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
21	Unknown 18	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
22	Unknown 19	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
23	Unknown 20	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
24	Unknown 21	Italy	Granaglione	Emilia-Romagna		AD	AD
25	Legno1	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
26	Legno2	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
27	Legno3	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
28	Legno4	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
29	Legno6	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
30	Legno7	Italy	Granaglione	Emilia-Romagna		AD	AD
31	Legno9	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
32	Legno10	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
33	Legno11	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
34	Legno12	Italy	Granaglione	Emilia-Romagna		Cluster 2	AD
35	Legno13	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
36	Legno14	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
37	Legno15	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
38	Legno16	Italy	Granaglione	Emilia-Romagna		AD	AD
39	Legno20	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
40	Legno21	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
41	Legno22	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
42	Legno23	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
43	Legno24	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
44	Legno25	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
45	Legno26	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
46	Legno27	Italy	Granaglione	Emilia-Romagna		Cluster 2	AD
47	Legno28	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
48	Legno29	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
49	Legno30	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
50	Legno31	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
51	Legno32	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
52	Matildico1	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
53	Matildico2	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
54	Matildico3	Italy	Granaglione	Emilia-Romagna		AD	AD
55	Matildico4	Italy	Granaglione	Emilia-Romagna		Cluster 2	AD
56	Matildico5	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
57	Matildico6	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
58	Matildico7	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
59	Matildico8	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
60	Matildico9	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
61	Matildico10	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
62	Matildico11	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
63	Matildico12	Italy	Granaglione	Emilia-Romagna		Cluster 2	AD
64	Matildico13	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
65	Unknown 22	Italy	Zocca	Emilia-Romagna		Cluster 2	AD
66	Unknown 23	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
67	Unknown 24	Italy	Zocca	Emilia-Romagna		AD	AD
68	Ceppa	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
69	Unknown 25	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
70	Unknown 26	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
71	Unknown 27	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
72	Unknown 28	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
73	Madonna	Italy	Zocca	Emilia-Romagna		AD	Cluster 3
74	Garfagnina	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
75	Unknown 29	Italy	Monzuno	Emilia-Romagna		AD	AD
76	Unknown 30	Italy	Monzuno	Emilia-Romagna		Cluster 2	Cluster 1
77	Biancherina1	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
78	Biancherina2	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
79	Calarese	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
80	Carrarese	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
81	Unknown 31	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
82	Loiola	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
83	Mascherina	Italy	Zocca	Emilia-Romagna		AD	Cluster 3
84	Massangaia	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
85	Molana	Italy	Zocca	Emilia-Romagna		Cluster 2	AD
86	Pelosa	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
87	Pistolese	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
88	Piusella	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
89	Svizzera	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
90	Tosca	Italy	Zocca	Emilia-Romagna		Cluster 2	Cluster 1
91	Loglia	Italy	Zocca	Emilia-Romagna		Cluster 2	AD
92	Tizzano4	Italy	Monteombraro (MO)	Emilia-Romagna		Cluster 1	Cluster 3
93	Tizzano5	Italy	Monteombraro (MO)	Emilia-Romagna		Cluster 2	Cluster 1
94	Unknown 32	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
95	Unknown 33	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
96	Unknown 34	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
97	Unknown 35	Italy	Granaglione	Emilia-Romagna		AD	AD
98	Matildico 1.1	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
99	Matildico 2.1	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
100	Matildico 3.1	Italy	Granaglione	Emilia-Romagna		AD	AD
101	Matildico 5.1	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
102	Legno21	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1
103	Precoce Migoule	Italy	Granaglione	Emilia-Romagna		AD	Cluster 3
104	Martina5	Italy	Monghidoro	Emilia-Romagna		AD	AD
105	Unknown 36	Italy	Granaglione	Emilia-Romagna		Cluster 2	Cluster 1

N° STRUCTURE	Samples	Stato	Region		Samples of Pereira-Lorenzo et al. (2017)	K=2	K=3
106	Unknown 37	Italy	Badolo (BO)	Emilia-Romagna		Cluster 2	AD
107	Unknown 38	Italy	Badolo (BO)	Emilia-Romagna		AD	Cluster 3
108	Unknown 39	Italy	Badolo (BO)	Emilia-Romagna		Cluster 2	Cluster 1
109	Unknown 40	Italy	Badolo (BO)	Emilia-Romagna		Cluster 1	Cluster 3
110	Unknown 41	Italy	Badolo (BO)	Emilia-Romagna		Cluster 1	Cluster 3
111	Unknown 42	Italy	Badolo (BO)	Emilia-Romagna		Cluster 2	Cluster 1
112	Unknown 43	Italy	Badolo (BO)	Emilia-Romagna		AD	Cluster 3
113	Unknown 44	Italy	Badolo (BO)	Emilia-Romagna		AD	AD
114	Unknown 45	Italy	Badolo (BO)	Emilia-Romagna		Cluster 2	AD
115	Unknown 46	Italy	Badolo (BO)	Emilia-Romagna		Cluster 2	Cluster 3
116	Stanco2	Italy	Stanco (Grizzana Morandi)	Emilia-Romagna		Cluster 1	Cluster 3
117	Biancherina3	Italy	Cavona (RE)	Emilia-Romagna		Cluster 2	Cluster 1
118	Piusella1	Italy	Cavona (RE)	Emilia-Romagna		Cluster 2	Cluster 1
119	Rossola	Italy	Cavona (RE)	Emilia-Romagna		Cluster 2	Cluster 1
120	Mand al Broc	Italy	Carola (RE)	Emilia-Romagna		Cluster 1	AD
121	Punghenta	Italy	Carola (RE)	Emilia-Romagna		Cluster 2	Cluster 1
122	Petra	Italy	Tiolla (RE)	Emilia-Romagna		Cluster 2	Cluster 1
123	Sfronzola	Italy	Monghidoro	Emilia-Romagna		Cluster 2	Cluster 1
124	Unknown 47	Italy	Monteombraro (MO)	Emilia-Romagna		Cluster 1	Cluster 3
125	Unknown 48	Italy	Monteombraro (MO)	Emilia-Romagna		Cluster 2	AD
126	Unknown 49	Italy	Brisighella (FA)	Emilia-Romagna		Cluster 1	Cluster 3
127	Caprarola	Italy	Brisighella (FA)	Emilia-Romagna		AD	Cluster 3
128	Montemarano	Italy	Brisighella (FA)	Emilia-Romagna		AD	AD
129	Mozza	Italy	Brisighella (FA)	Emilia-Romagna		AD	AD
130	Mondistollo	Italy	Brisighella (FA)	Emilia-Romagna		Cluster 1	Cluster 3
131	Castiglion dei Pepoli	Italy	Brisighella (FA)	Emilia-Romagna		Cluster 1	Cluster 3
132	Unknown 50	Italy	Brisighella (FA)	Emilia-Romagna		AD	Cluster 3
133	CastagnaG1	Italy	Trentino Alto Adige			Cluster 1	Cluster 3
134	Marrone di Knoll	Italy	Trentino Alto Adige			Cluster 1	Cluster 3
135	Marrone di Unterganzner	Italy	Trentino Alto Adige			Cluster 1	Cluster 3
136	Bouche de Betizac	France	Piemonte			AD	Cluster 3
137	C.Pumila	America	Piemonte			AD	Cluster 3
138	Garrone Rosso	Italy	Piemonte			Cluster 1	Cluster 3
139	C.Crenata	Giappone	Piemonte			AD	Cluster 3
140	Volos	Grecia	Piemonte			AD	Cluster 3
141	C.Mollissima	Cina	Piemonte			AD	Cluster 3
142	Amarelante 1	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
143	Blanca Canarias	Spain	Canary Islands		EU DATASET	Cluster 1	Cluster 2
144	Blanca Galicia	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
145	Calva2	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
146	Calva3	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
147	Campano	Spain	Castilla-León		EU DATASET	AD	AD
148	Cedo	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
149	Chiusa Pesio2	Italy	Piemonte		EU DATASET	Cluster 1	Cluster 3
150	Courela	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
151	Garrida	Spain	Galicia		EU DATASET	Cluster 1	AD
152	Grua	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
153	Inxerta	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
154	Judia	Portugal	Trás-os-Montes		EU DATASET	Cluster 1	AD
155	Mollar1	Spain	Canary Islands		EU DATASET	Cluster 1	Cluster 2
156	Mollar2	Spain	Canary Islands		EU DATASET	Cluster 1	Cluster 2
157	Mulata2	Spain	Canary Islands		EU DATASET	Cluster 1	Cluster 2
158	Negral	Spain	Galicia, Castilla-Leon		EU DATASET	Cluster 1	Cluster 2
159	País	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
160	Pelona Andalucia	Spain	Andalucía		EU DATASET	Cluster 1	Cluster 2
161	Porteliña	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
162	Rapuca2	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
163	Serodia	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
164	Tempuriva	Italy	Campania, Piemonte		EU DATASET	AD	AD
165	Unknown1	Spain	Canary Islands		EU DATASET	Cluster 1	Cluster 2
166	Unknown5	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
167	Unknown6	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
168	Vazqueña	Spain	Andalucía		EU DATASET	Cluster 1	Cluster 2
169	Vegamesada	Spain	Asturias		EU DATASET	Cluster 1	AD
170	Verata	Spain	Extremadura		EU DATASET	Cluster 1	Cluster 2
171	Verdina	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
172	Amarela	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
173	Bermello	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
174	Campilla	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
175	Famosa	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
176	Lebre	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
177	Raigona1	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
178	Raigona2	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
179	Rapada	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
180	Unknown4	Spain	Galicia		EU DATASET	Cluster 1	AD
181	Montagne	France	Dordogne		EU DATASET	Cluster 2	Cluster 1
182	Verdale	France	Dordogne		EU DATASET	Cluster 2	Cluster 1
183	Ventura	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
184	Xabrega	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
185	Comisaria Pelona	Spain	Andalucía		EU DATASET	AD	Cluster 2
186	Laga	Spain	Andalucía		EU DATASET	Cluster 1	Cluster 2
187	Laguilla	Spain	Andalucía		EU DATASET	Cluster 1	AD
188	Planta Alajar	Spain	Andalucía		EU DATASET	AD	AD
189	Temprana Jubrique	Spain	Andalucía		EU DATASET	Cluster 1	Cluster 2
190	Miguelina	Spain	Asturias		EU DATASET	Cluster 1	Cluster 3
191	Negrera	Spain	Asturias		EU DATASET	Cluster 1	AD
192	Abada	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
193	Cerredo	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
194	Injerta Bierzo	Spain	Castilla-León		EU DATASET	Cluster 1	Cluster 2
195	Pelado	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
196	Unknown7	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
197	Verdial	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
198	Armentina	Spain	Canary Islands		EU DATASET	Cluster 1	Cluster 2
199	Boroñona	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
200	Chamberga1	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
201	Galliciana	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
202	Llanisca	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
203	Loura	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2
204	Parruquina	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
205	Pelona Asturias	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
206	Valduna1	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
207	Valduna2	Spain	Asturias		EU DATASET	Cluster 1	Cluster 2
208	Puga	Spain	Galicia		EU DATASET	Cluster 1	AD
209	Rapuca1	Spain	Asturias		EU DATASET	Cluster 1	AD
210	Unknown8	Spain	Galicia		EU DATASET	Cluster 1	Cluster 2

N° STRUCTURE	Samples	Stato	Region	Samples of Pereira-Lorenzo et al. (2017)	K=2	K=3
211	Arial	Spain	Galicia	EU DATASET	Cluster 1	Cluster 2
212	Monfortina	Spain	Galicia	EU DATASET	Cluster 1	AD
213	Parede	Spain	Asturias, Castilla-León, Galicia	EU DATASET	Cluster 1	Cluster 2
214	Picona	Spain	Galicia	EU DATASET	Cluster 1	Cluster 2
215	Sergude	Spain	Galicia	EU DATASET	Cluster 1	Cluster 2
216	Comisaria2	Spain	Andalucía	EU DATASET	AD	AD
216	Temprana1	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 2
218	Injerta Roja	Spain	Extremadura	EU DATASET	AD	AD
219	Unknown2	Spain	Galicia	EU DATASET	AD	AD
220	Antigua	Spain	Extremadura	EU DATASET	Cluster 1	Cluster 2
221	Calvotera	Spain	Extremadura	EU DATASET	Cluster 1	Cluster 2
222	Comisaria1	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 2
223	De Pablo	Spain	Extremadura	EU DATASET	Cluster 1	Cluster 2
224	Majadas	Spain	Castilla-León	EU DATASET	Cluster 1	Cluster 2
225	Cachero	Spain	Canary Islands	EU DATASET	Cluster 1	Cluster 2
226	Calambres	Spain	Castilla-León	EU DATASET	Cluster 1	Cluster 2
227	Injerta Gorda	Spain	Extremadura	EU DATASET	Cluster 1	Cluster 2
228	Injerta Guadalupe	Spain	Extremadura	EU DATASET	Cluster 1	Cluster 2
229	Injerta Tío Sabino	Spain	Extremadura	EU DATASET	Cluster 1	Cluster 2
230	Longal	Spain	Galicia, Extremadura, Portugal	EU DATASET	Cluster 1	Cluster 2
231	Martahiña	Portugal	Trás-os-Montes	EU DATASET	AD	AD
232	Mollar3	Spain	Canary Islands	EU DATASET	Cluster 1	Cluster 2
233	Mulata1	Spain	Canary Islands	EU DATASET	Cluster 1	Cluster 2
234	Pelona Avila	Spain	Castilla-León	EU DATASET	Cluster 1	Cluster 2
235	Sietepernadas	Spain	Canary Islands	EU DATASET	Cluster 1	Cluster 2
236	Tardía Clara	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 2
237	Unknown3	Spain	Galicia	EU DATASET	Cluster 1	Cluster 2
238	Pilonga2	Spain	Andalucía	EU DATASET	AD	AD
239	Pilonga	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 2
240	Tardía Oscura	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 2
241	Capilla1	Spain	Andalucía	EU DATASET	Cluster 1	AD
242	Colorada	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 2
243	Marrone Roccadaspide	Italy	Campania	EU DATASET	Cluster 1	Cluster 2
244	Camberoune	France	Dordogne	EU DATASET	Cluster 1	AD
245	Luguesa	Spain	Galicia	EU DATASET	Cluster 1	Cluster 3
246	Peluda Tardía2	Spain	Andalucía	EU DATASET	Cluster 1	AD
247	Pertinaccio	Italy	Toscana, Emilia Romagna, Piemonte	EU DATASET	Cluster 1	AD
248	Temprana Genalguacil	Spain	Andalucía	EU DATASET	Cluster 1	AD
249	Tomasa	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 3
250	Capilla2	Spain	Andalucía	EU DATASET	Cluster 1	Cluster 3
251	Bracalla	Italy	Piemonte	EU DATASET	Cluster 1	Cluster 3
252	Chiusa Pesio1	Italy	Piemonte	EU DATASET	AD	Cluster 3
253	Marrone Comballe	France	South-East of France	EU DATASET	AD	Cluster 3
254	Inserta	Italy	Calabria	EU DATASET	Cluster 1	Cluster 3
255	Lucente1	Italy	Campania	EU DATASET	AD	AD
256	Lucente2	Italy	Campania	EU DATASET	AD	Cluster 3
257	Mamma	Italy	Calabria	EU DATASET	AD	AD
258	Marrone di Melfi	Italy	Basilicata, Campania	EU DATASET	Cluster 1	Cluster 2
259	Mercogliana	Italy	Campania	EU DATASET	Cluster 2	AD
260	Riggiola1	Italy	Calabria	EU DATASET	Cluster 1	AD
261	Riggiola2	Italy	Calabria	EU DATASET	AD	Cluster 3
262	Galega	Spain	Asturias	EU DATASET	Cluster 1	Cluster 2
263	Portuguesa	Spain	Andalucía		AD	Cluster 3
264	CA-15	France	INRA- Bordeaux		AD	Cluster 3
265	Gallega	Spain	Andalucía		AD	Cluster 3
266	Super	Spain	Extremadura		AD	Cluster 3
267	TG 90025	Spain	Galicia		AD	Cluster 3
268	125-10 TG	Spain	Galicia		AD	Cluster 3
269	324-29 TG	Spain	Galicia		AD	Cluster 3
270	Paderne4	Spain	Galicia		Cluster 1	Cluster 2
271	Doney 1 Sanabria	Spain	Castilla-León		Cluster 1	Cluster 2
272	Villasumil 1	Spain	Castilla-León		Cluster 1	Cluster 2
273	Entrambosrios 1	Spain	Galicia		Cluster 1	Cluster 2
274	Necas 3	Spain	Galicia		Cluster 1	Cluster 2
275	San Roman De Sanabria 1	Spain	Castilla-León		Cluster 1	Cluster 2
276	Baamonde 1 No Capilla	Spain	Galicia		Cluster 1	Cluster 2
277	Cerdedelo2 Das Viñas 1	Spain	Galicia		Cluster 1	Cluster 2
278	Beira Valente 3	Portugal			Cluster 1	Cluster 2
279	Chaguazoso Cementerio 2	Spain	Galicia		Cluster 1	Cluster 2
280	Vega Selorio Villaviciosa	Spain	Asturias		Cluster 1	Cluster 2
281	Pollayo 5	Spain	Cantabria		Cluster 1	Cluster 2
282	Nacerona 1 Ocejo	Spain	Cantabria		Cluster 1	Cluster 2
283	Pesaguero 5	Spain	Cantabria		Cluster 1	Cluster 2
284	Alcobilla12	Spain	Castilla-León		Cluster 1	Cluster 2
285	Cerdedelo B1	Spain	Galicia		Cluster 1	Cluster 2
286	Beira Valente 2	Portugal			Cluster 1	Cluster 2
287	Santa Eufemia (Baños Molgás) 1	Spain	Galicia		Cluster 1	Cluster 2
288	Alcobilla 6	Spain	Castilla-León		Cluster 1	Cluster 2
289	Las Caldas, Caces	Spain	Asturias		Cluster 1	Cluster 2
290	Peixeroos 2	Spain	Galicia		Cluster 1	Cluster 2
291	Antes De Villasumil 1	Spain	Castilla-León		Cluster 1	Cluster 2
292	Peixeroos 1 Raiz	Spain	Galicia		Cluster 1	Cluster 2
293	Trogais 1	Spain	Galicia		Cluster 1	Cluster 2
294	Vime De Sanabria 1	Spain	Castilla-León		Cluster 1	Cluster 2
295	Mendoia-2	Spain	Galicia		Cluster 1	Cluster 2
296	Verduenga 6 (Raiz)	Spain	Castilla-León		Cluster 1	AD
297	Alcobilla 8	Spain	Castilla-León		Cluster 1	Cluster 2
298	Ribeira 3	Spain	Galicia		Cluster 1	Cluster 2
299	Alcobilla1	Spain	Castilla-León		Cluster 1	Cluster 2
300	La Pesanca Riofabar 2	Spain	Asturias		Cluster 1	Cluster 2
301	Villaorille, Souto Quirós 1	Spain	Asturias		Cluster 1	Cluster 2
302	Verduenga 3 (Raiz)	Spain	Castilla-León		Cluster 1	Cluster 2
303	Pesaguero 1	Spain	Cantabria		Cluster 1	Cluster 2
304	Pesaguero 8	Spain	Cantabria		Cluster 1	Cluster 2
305	Verduenga 1	Spain	Castilla-León		Cluster 1	AD
306	Alcobilla10	Spain	Castilla-León		AD	AD
307	Chaguazoso Fonte 1	Spain	Galicia		Cluster 1	Cluster 2
308	Armeiriz2	Spain	Galicia		Cluster 1	Cluster 2
309	Cerdedelo B3	Spain	Galicia		Cluster 1	Cluster 2
310	Ribeira 1	Spain	Galicia		Cluster 1	Cluster 2
311	Villanueva Iglesia	Spain	Asturias		Cluster 1	Cluster 3
312	Villaorille, Souto Quirós 2	Spain	Asturias		Cluster 1	Cluster 2
313	Denia De Onís	Spain	Asturias		Cluster 1	Cluster 2
314	Pollayo 1	Spain	Cantabria		AD	AD
315	Necas 2	Spain	Galicia		Cluster 1	Cluster 2
316	Amola 1	Spain	Galicia		Cluster 1	Cluster 2
317	Nespereira4	Spain	Galicia		Cluster 1	Cluster 2
318	Carrelao 12	Spain	Galicia		Cluster 1	Cluster 2
319	Garrone	Italy	Piemonte		Cluster 1	AD

4.4.3 Population genetic structure of the European cultivars

In order to check if the hybrid samples could have affected the STRUCTURE analysis of the population, 13 hybrid genotypes were removed, and another analysis was performed under the same experimental conditions, for a total of 306 samples.

The results confirmed the clear separation between the two main clusters (C1, Spanish cluster with 167 samples, C2 Italian cluster with 104 and 35 admixed samples).

We also considered the separation for K=3 and K=4 with a ΔK=27.34 and 25.72, respectively (Figure 4.3, Table 4.4).

The separation between the Italian and Spanish varieties was also kept in further subdivisions: the Italian cluster was represented by Cluster 1 for the subdivision K=3 with 89 samples and by the Cluster 4 for K=4 with 86 samples, with important varieties from the Tuscan-Emilian Apennines, such as 'Pastanese', 'Ceppa', 'Pistolese', 'Piusela' and 'Lisanese'. This cluster also included the two France cultivars 'Montagne' and 'Verdale'.

An important group of varieties was Cluster 2 (C2) for K=3 and most of the Cluster 1 (C1) for K=4 , with different relevant varieties inside: 'Marrone Fiorentino', 'Madonna', 'Chiusa Pesio', 'Bracalla' from the North of Italy (Emilia-Romagna, Trentino Alto Adige and Piedmont) and 'Lucente', 'Inserta', 'Marron di Melfi', 'Riggiola' from the South of Italy (Calabria and Campania), 'Tomasa', 'Capilla' and 'Temprana' from South Spain (Andalucia) but also 'Luguesa' and 'Miguelina' from Galicia and Asturias. In addition, this sub-cluster included the variety 'Marrone di Comballe' from France.

Cluster 2 was represented for K=3 and K=4 by the cultivars from Canary island, Extremadura, Andalucía, Castilla-Leon and Galicia (Central Spain) and with the most important cultivar on the Iberian Peninsula, 'Longal' from Spain and Portugal, and 'Martahiña' from Portugal (Table 4.4).

The Cluster 3 included the main varieties from northern Spain, Galicia ('Famosa', 'Inxerta'), Castilla-Leon ('Negral'), Asturias ('Parede', 'Rapuca' and 'Chamberga') and Cantabria; and some accessions from Extremadura, with the main cultivar 'Verata', and from Canary Islands ('Mollar', 'Mulata' and 'Armentina') (Table 4.4). With 97 samples in Cluster 3 for K=3 and 79 for K=4, the subdivision within this subgroup was well defined by both STRUCTURE results.

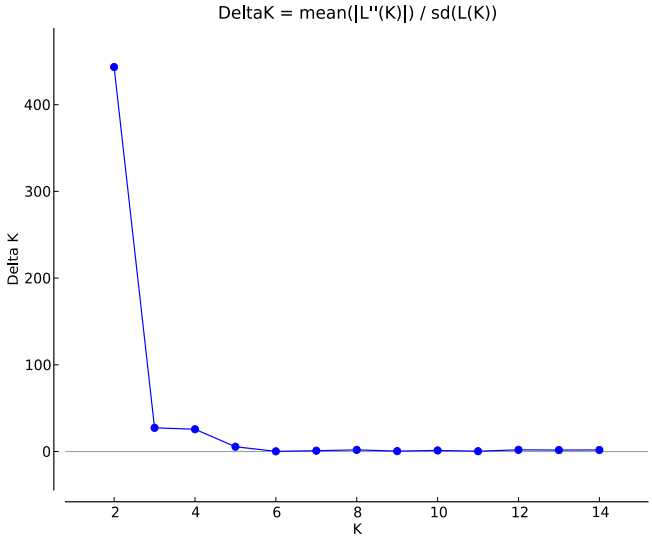

Figure 4.3. Estimates of Δ k calculated as described by Evanno et al. (2005) based on k-subdivision for 306 samples.

Table 4.4: list of the different varieties collected with the origin countries, the STRUCTURE subdivision for K=2, K=3 and K=4 for 306 samples (no hybrids included) and cultivated/wild subdivision.

N° STRUCTURE	Samples	COUNTRY	K=2	K=3	K=4	Division EU DATASET for K=2	Cultivated/Wild
1	Marrone Fiorentino	Italy	Cluster 1	Cluster 2	Cluster 1		C
2	Unknown 1	Italy	Cluster 2	Cluster 1	Cluster 4		W
3	Unknown 2	Italy	AD.2	AD.1	AD.3		W
4	Unknown 3	Italy	Cluster 2	AD.2	AD.1		W
5	Unknown 4	Italy	Cluster 2	AD.1	AD.4		W
6	Unknown 5	Italy	Cluster 2	Cluster 1	Cluster 4		W
7	Unknown 6	Italy	Cluster 2	Cluster 1	Cluster 4		C
8	Unknown 7	Italy	Cluster 2	Cluster 1	Cluster 4		C
9	Unknown 8	Italy	Cluster 2	Cluster 1	Cluster 4		W
10	Unknown 9	Italy	Cluster 2	Cluster 1	Cluster 4		W
11	Unknown 10	Italy	Cluster 2	Cluster 1	Cluster 4		W
12	Unknown 11	Italy	Cluster 2	Cluster 1	Cluster 4		W
13	Unknown 12	Italy	Cluster 2	Cluster 1	Cluster 4		C
14	Unknown 13	Italy	Cluster 2	Cluster 1	Cluster 4		C
15	Unknown 14	Italy	Cluster 2	Cluster 1	Cluster 4		W
16	Lisanese	Italy	Cluster 2	Cluster 1	Cluster 4		C
17	Unknown 15	Italy	Cluster 2	Cluster 1	Cluster 4		W
18	Unknown 16	Italy	Cluster 2	Cluster 1	Cluster 4		W
19	Unknown 17	Italy	Cluster 2	Cluster 1	Cluster 4		C
20	Pastanese	Italy	Cluster 2	Cluster 1	Cluster 4		C
21	Unknown 18	Italy	Cluster 2	Cluster 1	Cluster 4		C
22	Unknown 19	Italy	Cluster 2	Cluster 1	Cluster 4		W
23	Unknown 20	Italy	Cluster 2	Cluster 1	Cluster 4		W
24	Unknown 21	Italy	AD.1	AD.1	AD.4		W
25	Legno1	Italy	Cluster 2	Cluster 1	Cluster 4		W
26	Legno2	Italy	Cluster 2	Cluster 1	Cluster 4		W
27	Legno3	Italy	Cluster 2	Cluster 1	Cluster 4		C
28	Legno4	Italy	Cluster 2	Cluster 1	Cluster 4		C
29	Legno6	Italy	Cluster 2	Cluster 1	Cluster 4		C
30	Legno7	Italy	AD.2	AD.3	AD.3		W
31	Legno9	Italy	Cluster 2	Cluster 1	Cluster 4		W
32	Legno10	Italy	Cluster 2	Cluster 1	Cluster 4		C
33	Legno11	Italy	Cluster 2	Cluster 1	Cluster 4		W
34	Legno12	Italy	Cluster 2	Cluster 1	AD.4		W
35	Legno13	Italy	Cluster 2	Cluster 1	Cluster 4		W
36	Legno14	Italy	Cluster 2	Cluster 1	Cluster 4		W
37	Legno15	Italy	Cluster 2	Cluster 1	Cluster 4		W
38	Legno16	Italy	AD.1	AD.1	AD.4		W
39	Legno20	Italy	Cluster 2	Cluster 1	Cluster 4		C
40	Legno21	Italy	Cluster 2	Cluster 1	Cluster 4		W
41	Legno22	Italy	Cluster 2	Cluster 1	Cluster 4		W
42	Legno23	Italy	Cluster 2	Cluster 1	AD.4		W
43	Legno24	Italy	Cluster 2	Cluster 1	Cluster 4		C
44	Legno25	Italy	Cluster 2	Cluster 1	Cluster 4		C
45	Legno26	Italy	Cluster 2	Cluster 1	Cluster 4		C
46	Legno27	Italy	Cluster 2	AD.1	AD.1		C
47	Legno28	Italy	Cluster 2	Cluster 1	Cluster 4		W
48	Legno29	Italy	Cluster 2	Cluster 1	Cluster 4		W
49	Legno30	Italy	Cluster 2	Cluster 1	Cluster 4		W
50	Legno31	Italy	Cluster 2	Cluster 1	Cluster 4		C
51	Legno32	Italy	Cluster 2	Cluster 1	AD.4		C
52	Matildico1	Italy	Cluster 2	Cluster 1	Cluster 4		W
53	Matildico2	Italy	Cluster 2	Cluster 1	Cluster 4		W
54	Matildico3	Italy	AD.2	AD.1	AD.4		W
55	Matildico4	Italy	Cluster 2	AD.1	AD.4		W
56	Matildico5	Italy	Cluster 2	Cluster 1	Cluster 4		C
57	Matildico6	Italy	Cluster 2	Cluster 1	Cluster 4		C
58	Matildico7	Italy	Cluster 2	Cluster 1	Cluster 4		C
59	Matildico8	Italy	Cluster 2	Cluster 1	Cluster 4		C
60	Matildico9	Italy	Cluster 2	AD.1	AD.4		W
61	Matildico10	Italy	Cluster 2	Cluster 1	Cluster 4		C
62	Matildico11	Italy	Cluster 2	Cluster 1	Cluster 4		C
63	Matildico12	Italy	Cluster 2	AD.1	AD.4		W
64	Matildico13	Italy	Cluster 2	Cluster 1	Cluster 4		W
65	Unknown 22	Italy	Cluster 2	AD.1	AD.4		C
66	Unknown 23	Italy	Cluster 2	Cluster 1	Cluster 4		C
67	Unknown 24	Italy	AD.2	AD.2	AD.1		W
68	Ceppa	Italy	Cluster 2	Cluster 1	Cluster 4		C
69	Unknown 25	Italy	Cluster 2	Cluster 1	Cluster 4		W
70	Unknown 26	Italy	Cluster 2	Cluster 1	Cluster 4		W
71	Unknown 27	Italy	Cluster 2	Cluster 1	Cluster 4		W
72	Unknown 28	Italy	Cluster 2	Cluster 1	Cluster 4		C
73	Madonna	Italy	AD.1	Cluster 2	Cluster 1		W
74	Garfagnina	Italy	Cluster 2	Cluster 1	Cluster 4		W
75	Unknown 29	Italy	AD.2	AD.2	AD.1		W
76	Unknown 30	Italy	Cluster 2	Cluster 1	Cluster 4		W
77	Biancherina1	Italy	Cluster 2	Cluster 1	Cluster 4		C
78	Biancherina2	Italy	Cluster 2	Cluster 1	Cluster 4		C
79	Calarese	Italy	Cluster 2	Cluster 1	Cluster 4		C
80	Carrarese	Italy	Cluster 2	Cluster 1	Cluster 4		C
81	Unknown 31	Italy	Cluster 2	Cluster 1	Cluster 4		C
82	Loiola	Italy	Cluster 2	Cluster 1	Cluster 4		C
83	Mascherina	Italy	AD.2	AD.2	Cluster 1		W
84	Massangaia	Italy	Cluster 2	Cluster 1	Cluster 4		C
85	Molana	Italy	Cluster 2	AD.2	AD.4		C
86	Pelosa	Italy	Cluster 2	Cluster 1	Cluster 4		C
87	Pistolese	Italy	Cluster 2	Cluster 1	Cluster 4		C
88	Piusella	Italy	Cluster 2	Cluster 1	Cluster 4		C
89	Svizzera	Italy	Cluster 2	Cluster 1	AD.4		C
90	Tosca	Italy	Cluster 2	Cluster 1	Cluster 4		C
91	Loglia	Italy	AD.2	AD.2	AD.4		C
92	Tizzano4	Italy	Cluster 1	Cluster 2	Cluster 1		C
93	Tizzano5	Italy	Cluster 2	Cluster 1	Cluster 4		W
94	Unknown 32	Italy	Cluster 2	Cluster 1	Cluster 4		C
95	Unknown 33	Italy	Cluster 2	Cluster 1	Cluster 4		C
96	Unknown 34	Italy	Cluster 2	Cluster 1	Cluster 4		W
97	Unknown 35	Italy	AD.2	AD.1	AD.4		W
98	Matildico 1.1	Italy	Cluster 2	Cluster 1	Cluster 4		W
99	Matildico 2.1	Italy	Cluster 2	Cluster 1	Cluster 4		W
100	Matildico 3.1	Italy	AD.2	AD.2	AD.1		W
101	Matildico 5.1	Italy	Cluster 2	Cluster 1	Cluster 4		W
102	Legno21	Italy	Cluster 2	Cluster 1	Cluster 4		W

N° STRUCTURE	Samples	COUNTRY	K=2	K=3	K=4	Division EU DATASET for K=2	Cultivated/Wild
103	Martina5	Italy	AD.2	AD.2	AD.1		W
104	Unknown 36	Italy	Cluster 2	Cluster 1	Cluster 4		C
105	Unknown 37	Italy	Cluster 2	AD.1	AD.4		W
106	Unknown 38	Italy	AD.2	Cluster 2	Cluster 1		C
107	Unknown 39	Italy	Cluster 2	AD.1	Cluster 4		W
108	Unknown 40	Italy	Cluster 1	Cluster 2	Cluster 1		C
109	Unknown 41	Italy	Cluster 1	Cluster 2	Cluster 1		W
110	Unknown 42	Italy	Cluster 2	Cluster 1	AD.4		W
111	Unknown 43	Italy	AD.2	Cluster 2	Cluster 1		W
112	Unknown 44	Italy	AD.2	AD.2	AD.4		W
113	Unknown 45	Italy	Cluster 2	AD.1	AD.1		C
114	Unknown 46	Italy	Cluster 2	Cluster 2	Cluster 1		W
115	Stanco2	Italy	Cluster 1	Cluster 2	Cluster 1		C
116	Biancherina3	Italy	Cluster 2	Cluster 1	Cluster 4		W
117	Piusella1	Italy	Cluster 2	Cluster 1	Cluster 4		C
118	Rossola	Italy	Cluster 2	Cluster 1	Cluster 4		C
119	Mand al Broc	Italy	Cluster 1	Cluster 2	AD.1		C
120	Punghenta	Italy	Cluster 2	Cluster 1	Cluster 4		W
121	Petra	Italy	Cluster 2	Cluster 1	Cluster 4		W
122	Sfronzola	Italy	Cluster 2	AD.1	AD.4		W
123	Unknown 47	Italy	Cluster 1	Cluster 2	Cluster 1		C
124	Unknown 48	Italy	Cluster 2	AD.1	AD.4		W
125	Unknown 49	Italy	Cluster 1	Cluster 2	Cluster 1		C
126	Caprarola	Italy	Cluster 1	Cluster 2	AD.1		W
127	Montemarano	Italy	AD.1	Cluster 2	AD.2		C
128	Mozza	Italy	AD.2	AD.2	AD.4		W
129	Mondistollo	Italy	Cluster 1	Cluster 2	Cluster 1		C
130	Castiglion dei Pepoli	Italy	Cluster 1	Cluster 2	Cluster 1		C
131	Unknown 50	Italy	Cluster 1	Cluster 2	Cluster 1		C
132	CastagnaG1	Italy	Cluster 1	Cluster 2	Cluster 1		C
133	Marrone di Knoll	Italy	Cluster 1	Cluster 2	Cluster 1		C
134	Marrone di Unterganzner	Italy	Cluster 1	Cluster 2	Cluster 1		C
135	Amarelante 1	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
136	Blanca Canarias	Spain	Cluster 1	Cluster 3	AD.2	RPP1	C
137	Blanca Galicia	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
138	Calva2	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
139	Calva3	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
140	Campano	Spain	AD.1	AD.3	AD.3	RPP2	C
141	Cedo	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
142	Chiusa Pesio2	Italy	Cluster 1	Cluster 2	Cluster 1	Admixed	C
143	Courela	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
144	Garrida	Spain	Cluster 1	AD.3	AD.3	Admixed	C
145	Grua	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
146	Inxerta	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
147	Judia	Portugal	Cluster 1	Cluster 2	AD.2	Admixed	C
148	Mollar1	Spain	Cluster 1	Cluster 3	AD.2	RPP1	C
149	Mollar2	Spain	Cluster 1	AD.3	AD.2	RPP1	C
150	Mulata2	Spain	Cluster 1	Cluster 3	Cluster 2	RPP1	C
151	Negral	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
152	País	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
153	Pelona Andalucia	Spain	Cluster 1	AD.3	Cluster 2	Admixed	C
154	Porteliña	Spain	Cluster 1	Cluster 3	AD.3	RPP1	C
155	Rapuca2	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
156	Serodia	Spain	Cluster 1	Cluster 3	AD.3	RPP1	C
157	Tempuriva	Italy	AD.1	AD.3	AD.3	RPP2	C
158	Unknown1	Spain	Cluster 1	Cluster 3	AD.3	RPP1	C
159	Unknown5	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
160	Unknown6	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
161	Vazqueña	Spain	Cluster 1	Cluster 2	Cluster 2	Admixed	C
162	Vegamesada	Spain	Cluster 1	Cluster 3	AD.3	Admixed	C
163	Verata	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
164	Verdina	Spain	Cluster 1	Cluster 3	AD.3	RPP1	C
165	Amarela	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
166	Bermello	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
167	Campilla	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
168	Famosa	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
169	Lebre	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
170	Raigona1	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
171	Raigona2	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
172	Rapada	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
173	Unknown4	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
174	Montagne	France	Cluster 2	Cluster 1	Cluster 4	RPP2	C
175	Verdale	France	Cluster 2	Cluster 1	Cluster 4	RPP2	C
176	Ventura	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
177	Xabrega	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
178	Comisaria Pelona	Spain	AD.1	AD.2	AD.2	Admixed	C
179	Laga	Spain	Cluster 1	AD.2	Cluster 2	RPP1	C
180	Laguilla	Spain	Cluster 1	Cluster 2	AD.2	Admixed	C
181	Planta Alajar	Spain	AD.1	Cluster 2	AD.2	Admixed	C
182	Temprana Jubrique	Spain	Cluster 1	AD.2	AD.2	Admixed	C
183	Miguelina	Spain	Cluster 1	Cluster 2	Cluster 1	Admixed	C
184	Negrera	Spain	Cluster 1	Cluster 3	Cluster 3	Admixed	C
185	Abada	Spain	Cluster 1	Cluster 3	AD.3	RPP1	C
186	Cerredo	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
187	Injerta Bierzo	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
188	Pelado	Spain	Cluster 1	AD.3	AD.3	RPP1	C
189	Unknown7	Spain	Cluster 1	Cluster 3	AD.3	RPP1	C
190	Verdial	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
191	Armentina	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
192	Boroñona	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
193	Chamberga1	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
194	Galliciana	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
195	Llanisca	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
196	Loura	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
197	Parruquina	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
198	Pelona Asturias	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
199	Valduna1	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
200	Valduna2	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
201	Puga	Spain	Cluster 1	Cluster 3	Cluster 3	Admixed	C
202	Rapuca1	Spain	Cluster 1	Cluster 3	Cluster 3	Admixed	C
203	Unknown8	Spain	Cluster 1	Cluster 3	Cluster 3	Admixed	C
204	Arial	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C

N° STRUCTURE	Samples	COUNTRY	K=2	K=3	K=4	Division EU DATASET for K=2	Cultivated/Wild
205	Monfortina	Spain	Cluster 1	AD.3	AD.3	RPP1	C
206	Parede	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
207	Picona	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
208	Sergude	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
209	Comisaria2	Spain	AD.1	AD.1	AD.2	Admixed	C
210	Temprana1	Spain	Cluster 1	Cluster 2	Cluster 2	Admixed	C
211	Injerta Roja	Spain	AD.1	Cluster 2	Cluster 2	Admixed	C
212	Unknown2	Spain	AD.1	AD.2	Cluster 2	Admixed	C
213	Antigua	Spain	Cluster 1	AD.3	Cluster 2	RPP1	C
214	Calvotera	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
215	Comisaria1	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
216	De Pablo	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
217	Majadas	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
218	Cachero	Spain	Cluster 1	Cluster 3	Cluster 2	RPP1	C
219	Calambres	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
220	Injerta Gorda	Spain	Cluster 1	AD.2	Cluster 2	RPP1	C
221	Injerta Guadalupe	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
222	Injerta Tío Sabino	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
223	Longal	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
224	Martahiña	Portugal	AD.1	Cluster 2	Cluster 2	RPP1	C
225	Mollar3	Spain	Cluster 1	Cluster 3	Cluster 2	RPP1	C
226	Mulata1	Spain	Cluster 1	Cluster 3	Cluster 2	RPP1	C
227	Pelona Avila	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
228	Sietepernadas	Spain	Cluster 1	AD.3	Cluster 2	RPP1	C
229	Tardía Clara	Spain	Cluster 1	AD.2	Cluster 2	RPP1	C
230	Unknown3	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
231	Pilonga2	Spain	AD.2	AD.2	AD.2	RPP1	C
232	Pilonga	Spain	Cluster 1	AD.3	AD.2	RPP2	C
233	Tardía Oscura	Spain	Cluster 1	Cluster 2	Cluster 2	RPP1	C
234	Capilla1	Spain	Cluster 1	AD.3	AD.3	Admixed	C
235	Colorada	Spain	Cluster 1	Cluster 2	AD.2	Admixed	C
236	Marrone Roccadaspide	Italy	Cluster 1	AD.3	AD.2	Admixed	C
237	Camberoune	France	Cluster 1	AD.2	AD.1	RPP2	C
238	Luguesa	Spain	Cluster 1	Cluster 2	Cluster 1	RPP2	C
239	Peluda Tardía2	Spain	Cluster 1	Cluster 2	AD.1	RPP2	C
240	Pertinaccio	Italy	Cluster 1	Cluster 3	AD.3	RPP2	C
241	Temprana Genalguacil	Spain	Cluster 1	Cluster 2	AD.2	RPP2	C
242	Tomasa	Spain	Cluster 1	Cluster 2	Cluster 1	RPP2	C
243	Capilla2	Spain	Cluster 1	Cluster 2	Cluster 1	RPP2	C
244	Bracalla	Italy	Cluster 1	Cluster 2	Cluster 1	RPP2	C
245	Chiusa Pesio1	Italy	AD.1	Cluster 2	Cluster 1	RPP2	C
246	Marrone Comballe	France	AD.1	Cluster 2	Cluster 1	RPP2	C
247	Inserta	Italy	Cluster 1	Cluster 2	Cluster 1	RPP2	C
248	Lucente1	Italy	AD.2	AD.2	AD.4	RPP2	C
249	Lucente2	Italy	AD.2	Cluster 2	Cluster 1	RPP2	C
250	Mamma	Italy	AD.2	Cluster 2	AD.1	RPP2	C
251	Marrone di Melfi	Italy	Cluster 1	Cluster 2	AD.2	RPP2	C
252	Mercogliana	Italy	Cluster 2	AD.1	AD.4	RPP2	C
253	Riggiola1	Italy	Cluster 1	Cluster 2	Cluster 1	RPP2	C
254	Riggiola2	Italy	AD.2	Cluster 2	Cluster 1	RPP2	C
255	Galega	Spain	Cluster 1	Cluster 3	Cluster 3	RPP1	C
256	Paderne4	Spain	Cluster 1	AD.3	AD.3		W
257	Doney 1 Sanabria	Spain	Cluster 1	Cluster 3	Cluster 3		W
258	Villasumil 1	Spain	Cluster 1	Cluster 3	Cluster 3		W
259	Entrambosrios 1	Spain	Cluster 1	Cluster 3	Cluster 3		C
260	Necas 3	Spain	Cluster 1	Cluster 3	Cluster 3		C
261	San Roman De Sanabria 1	Spain	Cluster 1	Cluster 3	Cluster 3		W
262	Baamonde 1 No Capilla	Spain	Cluster 1	Cluster 3	Cluster 3		W
263	Cerdedelo2 Das Viñas 1	Spain	Cluster 1	AD.3	AD.3		C
264	Beira Valente 3	Portugal	Cluster 1	Cluster 3	Cluster 3		W
265	Chaguazoso Cementerio 2	Spain	Cluster 1	Cluster 2	Cluster 2		W
266	Vega Selorio Villaviciosa	Spain	Cluster 1	Cluster 3	Cluster 3		W
267	Pollayo 5	Spain	Cluster 1	Cluster 3	Cluster 3		W
268	Nacerona 1 Ocejo	Spain	Cluster 1	Cluster 3	Cluster 3		W
269	Pesaguero 5	Spain	Cluster 1	Cluster 3	Cluster 3		W
270	Alcobilla12	Spain	Cluster 1	Cluster 3	Cluster 3		W
271	Cerdedelo B1	Spain	Cluster 1	Cluster 3	Cluster 3		W
272	Beira Valente 2	Portugal	Cluster 1	Cluster 2	Cluster 2		C
273	Santa Eufemia (Baños Molgás)	Spain	Cluster 1	Cluster 3	Cluster 3		W
274	Alcobilla 6	Spain	Cluster 1	Cluster 3	Cluster 3		W
275	Las Caldas, Caces	Spain	Cluster 1	Cluster 3	Cluster 3		W
276	Peixeroos 2	Spain	Cluster 1	Cluster 3	Cluster 3		W
277	Antes De Villasumil 1	Spain	Cluster 1	Cluster 3	Cluster 3		W
278	Peixeroos 1 Raiz	Spain	Cluster 1	Cluster 3	Cluster 3		W
279	Trogais 1	Spain	Cluster 1	AD.2	Cluster 2		W
280	Vime De Sanabria 1	Spain	Cluster 1	AD.3	AD.3		W
281	Mendoia-2	Spain	Cluster 1	Cluster 3	Cluster 3		W
282	Verduenga 6 (Raiz)	Spain	Cluster 1	AD.3	AD.3		W
283	Alcobilla 8	Spain	Cluster 1	Cluster 3	Cluster 3		W
284	Ribeira 3	Spain	Cluster 1	AD.2	AD.2		C
285	Alcobilla1	Spain	Cluster 1	Cluster 3	Cluster 3		W
286	La Pesanca Riofabar 2	Spain	Cluster 1	Cluster 3	Cluster 3		W
287	Villaorille, Souto Quirós 1	Spain	Cluster 1	Cluster 3	Cluster 3		W
288	Verduenga 3 (Raiz)	Spain	Cluster 1	Cluster 3	Cluster 3		W
289	Pesaguero 1	Spain	Cluster 1	Cluster 3	Cluster 3		W
290	Pesaguero 8	Spain	Cluster 1	Cluster 3	AD.3		W
291	Verduenga 1	Spain	Cluster 1	Cluster 3	AD.3		C
292	Alcobilla10	Spain	AD.2	AD.3	AD.4		W
293	Chaguazoso Fonte 1	Spain	Cluster 1	AD.3	Cluster 2		W
294	Armeiriz2	Spain	Cluster 1	Cluster 3	Cluster 3		W
295	Cerdedelo B3	Spain	Cluster 1	Cluster 3	Cluster 3		C
296	Ribeira 1	Spain	Cluster 1	Cluster 3	Cluster 3		W
297	Villanueva Iglesia	Spain	Cluster 1	AD.2	AD.1		W
298	Villaorille, Souto Quirós 2	Spain	Cluster 1	Cluster 3	Cluster 3		W
299	Denia De Onís	Spain	Cluster 1	Cluster 3	Cluster 3		W
300	Pollayo 1	Spain	AD.1	Cluster 3	Cluster 3		W
301	Necas 2	Spain	Cluster 1	Cluster 2	Cluster 2		W
302	Amola 1	Spain	Cluster 1	AD.3	Cluster 2		W
303	Nespereira4	Spain	Cluster 1	Cluster 3	AD.3		C
304	Carrelao 12	Spain	Cluster 1	Cluster 3	AD.3		C
305	Garoone Rosso	Italy	Cluster 1	AD.2	Cluster 1		C
306	Garrone	Italy	Cluster 1	AD.2	AD.1		C

4.4.4 Genetic differentiation

A Principal Coordinate Analysis (PCoA) for 306 samples with 14 SSR markers was performed on the genetic distance matrix using the program GenAlEx, to confirm the division resulting from STRUCTURE analysis for K=2, K=3 and K=4 (Figure 4.4). The PCoA showed the large diversity existing between *C. sativa* genotypes.

Figure 4.4 shows, in particular, that the Italian cluster (in green) was differentiated from the main Spanish cultivars (in red) with the admixture sample that connected the clusters (in grey). The two main groups presented a meeting point with the admixture samples from the STRUCTURE analysis.

The Spanish cluster for K=3 showed a differentiation between the northern Spain (in blue) and the southern (in red). In this condition the northern Italian cluster appeared to be very well separated (in green; Figure 4.4).

In K=4 it was possible to recognize cultivars of both South Spain and Italy from central-northern Spain (represented by C1 and C2).

The results of the STRUCTURE analysis were used for the calculation of the fixation coefficient (FST) and the Hierarchical analysis of molecular variance (AMOVA), calculated with GenAlEx.

The genetic differentiation between the two main clusters was FST = 0.077, $P < 0.001$ (Table 4.5, panel A), suggesting a genetic structure for the chestnut at European level, confirmed also by the AMOVA results (8%). Similar AMOVA results were found for K=3 and K=4, with a 6% and 7% of variance component among the populations, respectively (Table 4.5, panel B and C).

The largest differentiation between pairs of groups was found between the northern Italian cluster (C4) with mainly samples from the Tuscan-Emilian Apennines, and the North-Central Spain cluster (C3) for K=4 (FST=0.133, $P<0.001$), as showed in Table 4.5 panel C.

A high FST value was observed also between the C1 with cultivars from Italy, southern Spain and France and C4 represented by the Italian cluster with FST=0.113 $P<0.001$; similarly, between C1 and C2, which included central Spain's varieties (FST=0.112, $P<0.001$).

With the dendrogram, also, we distinguished the cultivars between the Wild samples of the Italian accessions using the Jaccard dissimilarity coefficients (Dixon 2003; Oksanen et al., 2017).

The AMOVA analysis revealed no substantial differences between wild samples and cultivated varieties, in relation with the STRUCTURE division in two main clusters (Table 4.5, panel D and E). The variance components among populations were 1% and 5% respectively, confirmed also by the FST index (0.0012 with $P<0.001$). Instead, 99% and 95% of the variance were found within populations.

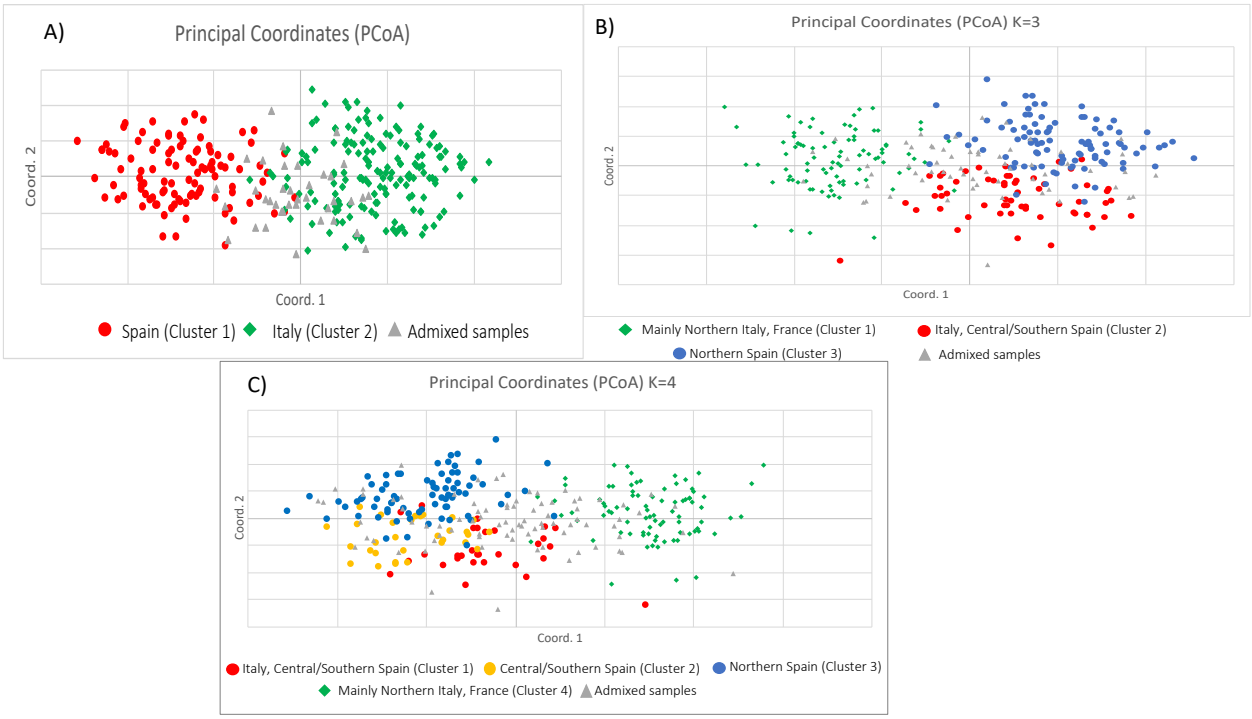

Figure 4.4: Principal Coordinate Analysis (PCoA) based on polymorphism at 14 SSR loci for 306 unique genotypes. Accession color reflects the consistent assignment using Bayesian analysis to the sub-groups defined in Fig. 3.

Table 4.5: Analysis of molecular variance (AMOVA) and FST value based on the 14 SSR loci of 306 chestnut accessions corresponding to: A),B),C) the groups K=2, K=3 and K=4 as defined by STRUCTURE analysis; D) the total population (n= 306) divided in wild and the cultivated samples; E) Wild between Cultivated samples in K=2. All estimate was highly significant - P<0,001.

A) Structure Cluster K=2, No hybrids included (306 samples)

Pairwise estimate of Fst value

	Q value	Country	N° samples	C1	C2	Admixture
C1	> 0,8	Spain	164	-	0,077	0,019
C2	> 0,8	Italy	104		-	0,032
Admixture	< 0,8		38			-

Summary AMOVA Table

Source	df	SS	MS	Est. Var.	%
Among Pops	2	119,673	59,837	0,308	8%
Within Pops	609	3053,864	5,015	5,015	92%
P-value	0,077				

B) Structure Cluster K=3, No hybrids included (306 samples)

Pairwise estimate of Fst value

	Q value	Country	N° samples	Cluster 1	Cluster 2	Cluster 3	Admixture
Cluster 1	>0,8	Northern Italy, France	89	-			
Cluster 2	>0,8	Mainly Italy, central and southern Spain	59	0,087	-		
Cluster 3	>0,8	Northern Spain	97	0,097	0,058	-	
Admixture	<0,8		61	0,041	0,020	0,025	-

Summary AMOVA

Source	df	SS	MS	Est. Var.	%
Among Pops	3	158,264	52,755	0,317	6%
Within Pops	608	3015,273	4,959	4,959	94%
P-value	0,060				

C) Structure Cluster K=4, No hybrids included (306 samples)

Pairwise estimate of Fst value

	Q	Country	N° samples	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Admixture
Cluster 1	>0,8	Italy, central and southern Spain	30	-				
Cluster 2	>0,8	Central and southern Spain	31	0,112	-			
Cluster 3	>0,8	Northern Spain	76	0,094	0,075	-		
Cluster 4	>0,8	Mainly Northern Italy, France	85	0,113	0,133	0,100	-	
Admixture	<0,8		84	0,048	0,050	0,025	0,042	-

Summary AMOVA

Source	df	SS	MS	Est. Var.	%
Among Pops	4	192,643	48,161	0,369	7%
Within Pops	607	2980,895	4,911	4,911	93%

P-value 0,070

D) 306 samples divided in Cultivated and Wild

Pairwise estimate of Fst value

	N° samples	Cultivated	Wild
Cultivated	198	-	0,012
Wild	108		-

Summary AMOVA Table

Source	df	SS	MS	Est. Var.	%
Among Pops	1	22,811	22,811	0,063	1%
Within Pops	610	3154,328	5,171	5,171	99%

P-value 0,012

E) 306 samples divided in Cultivated and Wild in C1/C2

Pairwise estimate of Fst value

	N° samples	Country	Q value	C2.Cultivated	C2.Wild	C1.Cultivated	C1.Wild	Admixture
C2.Cultivated	51	Italy	> 0,8	-	0,009	0,087	0,096	0,049
C2.Wild	53	Italy	> 0,8		-	0,068	0,078	0,025
C1.Cultivated	126	Spain	> 0,8			-	0,003	0,013
C1.Wild	39	Spain	> 0,8				-	0,023
Admixture	37		< 0,8					-

Summary AMOVA Table

Source	df	SS	MS	Est. Var.	%
Among Pops	4	135,076	33,769	0,253	5%
Within Pops	607	3042,063	5,012	5,012	95%

P-value 0,048

4.5 Discussion

Our results showed a high polymorphism in *Castanea* spp. as in previous studies (Pereira-Lorenzo et al., 2011; Martin et al., 2017). Moreover, Urrestarazu et al. (2015) demonstrated that a number of 12 – 15 microsatellite markers is sufficient for accurate fingerprinting purposes.

The high discriminating power of the microsatellites used was confirmed also by Pereira-Lorenzo et al. (2011) with 13.25 average alleles per locus for 10 SSRs used and by Pereira-Lorenzo et al. (2017) with 8.92 using 24 SSRs.

In particular, the loci CsCAT3 (PIC=0.879) and EMCs15 (PIC=0.604) appeared to be the most and least informative loci, as reported also by Pereira-Lorenzo et al. (2010, 2011) and by Martín et al. (2012).

The EMCs markers are trinucleotide SSRs (EMCs) and mutate with a lower rate than dinucleotide SSRs (CsCAT), thus resulting with less polymorphic**.**

A clear separation therefore emerged initially between the Spanish and the Italian varieties and then among the samples belonging to the north and south of Spain.

This finding was explained before due to the genetic differentiation between northern and central Iberian Peninsula can be explained by genetic adaptations to climatic conditions, mainly temperature and precipitation gradient (Pereira-Lorenzo et al., 2010) and the instant domestication process by grafting (Pereira-Lorenzo et al., 2019) followed by centuries of diversification process by hybridisation (Pereira-Lorenzo et al., 2011) from main cultivars.

In Italy during the Middle Age, a key player for the chestnut cultivation was Matilde di Canossa, who increased the area of chestnut cultivation with the help of the Tuscan Benedictine monks, especially in the area of the Tuscan-Emilian Apennines. She also introduced a variety of chestnut, named upon her ('Matildico' or 'Pastanese'), an ancient local chestnut cultivar known in the past for the production of high-quality flour.

Moreover, from the analysis of STRUCTURE it was found that some varieties of southern Spain (Andalucia) shared a higher number of alleles with the varieties of southern Italy (Calabria and Campania regions), but also with the main important northern chestnut cultivars (as 'Marrone Fiorentino'), maybe due to the historical relationships among Spanish kingdom and southern Italy (Figure 4.5A and 4.5B).

The genetic differentiation between the two main clusters and admixed samples was low (FST=0.019 and 0.032 $P<0.001$, respectively), which could support the early hybridization between the Italian and the Spanish group. Admixed samples can indicate an earlier hybridization between the two clusters, as suggested by Pereira-Lorenzo et al. (2012), taking into account the oldest giant tree in Andalucía with Italian genetic background, it should have occurred before the XV century (Pereira-Lorenzo et al., 2019).

Our results are in line with Pereira-Lorenzo et al. (2019), supporting the hypothesis of an earlier introduction of chestnuts cultivation from Italy to Spain, in particular in the Andalucía and the Extremadura regions, with contacts also in Castilla-León and Galicia. This can be noted with the cultivar 'Luguesa' that was included in C2 and C1 for K=3 and K=4 STRUCTURE subdivisions with the main southern Italian varieties.

In the history of chestnut cultivation, the reduction of diversity produced by grafting may have been compensated by the use of seedlings as reported by Auge and Brandl (1997), Forneck (2005) and Pereira-Lorenzo (2010): a seedling plant of a local cultivar has been selected for the superior traits of the nuts and used for the multiplication.

Pereira-Lorenzo et al. (2010) showed that the distribution of seedlings from the main cultivar groups of the north and central Iberian Peninsula was used to create new orchards in South of Spain, in particular in Andalucía and in the Canary Islands, as the cultivar 'Longal', 'Reborda' and 'Dieguina'.

Our results confirmed the relationship between 'Longal' and different main varieties from South Spain, such as 'Laga', 'Temprana' and 'Pelona' and in Extremadura with 'Injerta'. This was in line also with previous studies (Pereira-Lorenzo et al., 2006 and Costa et al., 2008), identifying 'Longal' as a cultivar used for genetic contribution to create new cultivars in different regions of Spain, and explains the reason of the huge number of admixture samples between the sub-clusters under exam.

Our results highlighted that the genetic of chestnut trees is characterized by a complex structure and genetic diversity. Hybridization could therefore have played an important role in the diversification process as previously suggested by Pereira-Lorenzo et al. (2011). It also explains the great diversity found in a small geographic area such as the Tuscan-Emilian Apennines (central-northern Italy), and in Galicia region (northern Spain).

The AMOVA analysis showed a low FST value among the wild and the cultivated chestnuts (FST= 0.0012 with $P<0.001$).

This result underline that *C. sativa* is an outcrossing species and for this reason the gene flow between wild and cultivated samples is low. In addition, it is important to consider the changes in usage over time and in agricultural practices of the territory and the "instant domestication" (Pereira-Lorenzo et al., 2019).

Our results confirm previous studies on autochthone chestnut populations (Martín et al., 2012; Mattioni et al., 2013) and local cultivated varieties (Gobbin et al., 2007; Martín et al., 2009; Pereira-Lorenzo et al., 2010).

Genetic differentiation of Spanish cultivars from other gene pools has also been found in other tree crops (Pereira-Lorenzo et al., 2017) as pear (Miranda et al., 2010; Dos Santos et al., 2011; Urrestarazu et al., 2015), peach (Aranzana et al., 2010; Fonti Forcada et al., 2013), apricot (Bourguiba et al., 2012) or grapevine (Díaz-Losada et al., 2012; Emanuelli et al., 2013).

In this study, by adding 138 Northern Italian genotypes to the EU database, we were able to differentiate a similar genetic differentiation, although of lower intensity, between northern and central-southern Italy as it was previously demonstrated between the northern and central-southern Iberian Peninsula (Pereira-Lorenzo et al., 2010, 2019). In addition, northern Italy produced also a high introgression (30% of the genotypes, Figure 4.5A) in central-southern Spain but lower than central-southern Italy (63%). Moreover, introgression from both genetics origins in Italy was also noticed in northern Spain (16%), with 5% from northern and 11% from central-southern Italy.

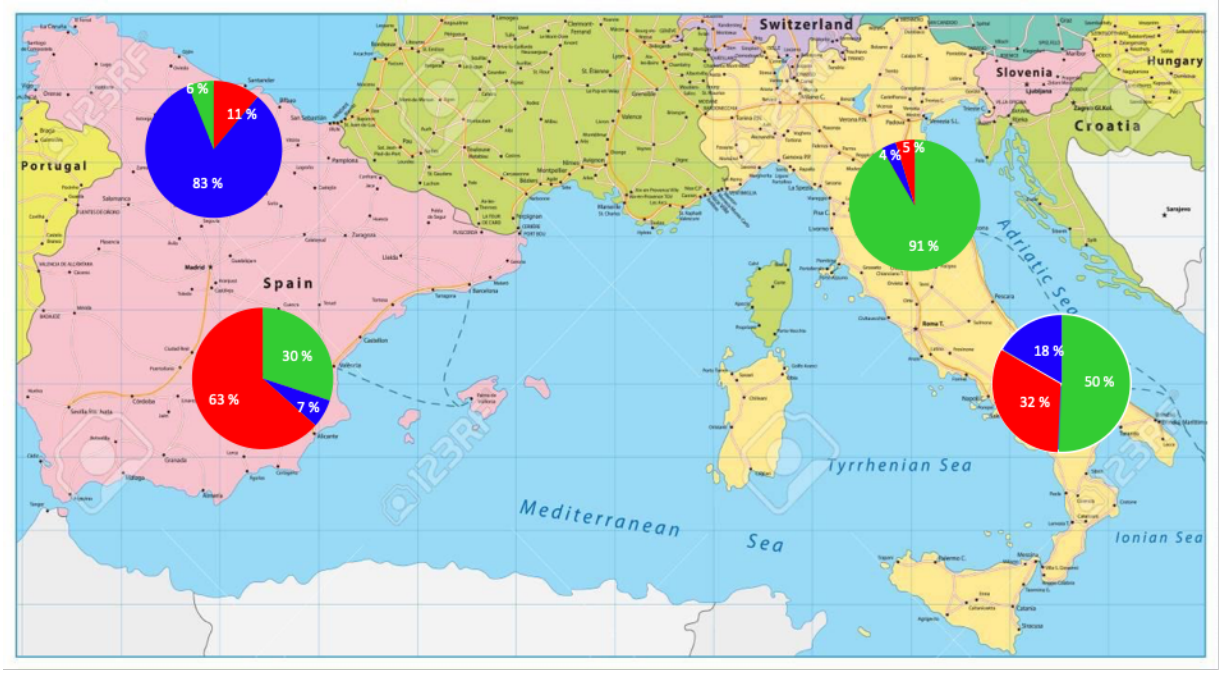

Figure 4.5A: Gene pool distribution of chestnut cultivars for K=3 between the north and the south, both of Italy and Spain (Green – Mainly Northern Italian and France (Cluster 1); Red – Italy, Central and Southern Spain (Cluster 2); Blue –Northern of Spain (Cluster 3).

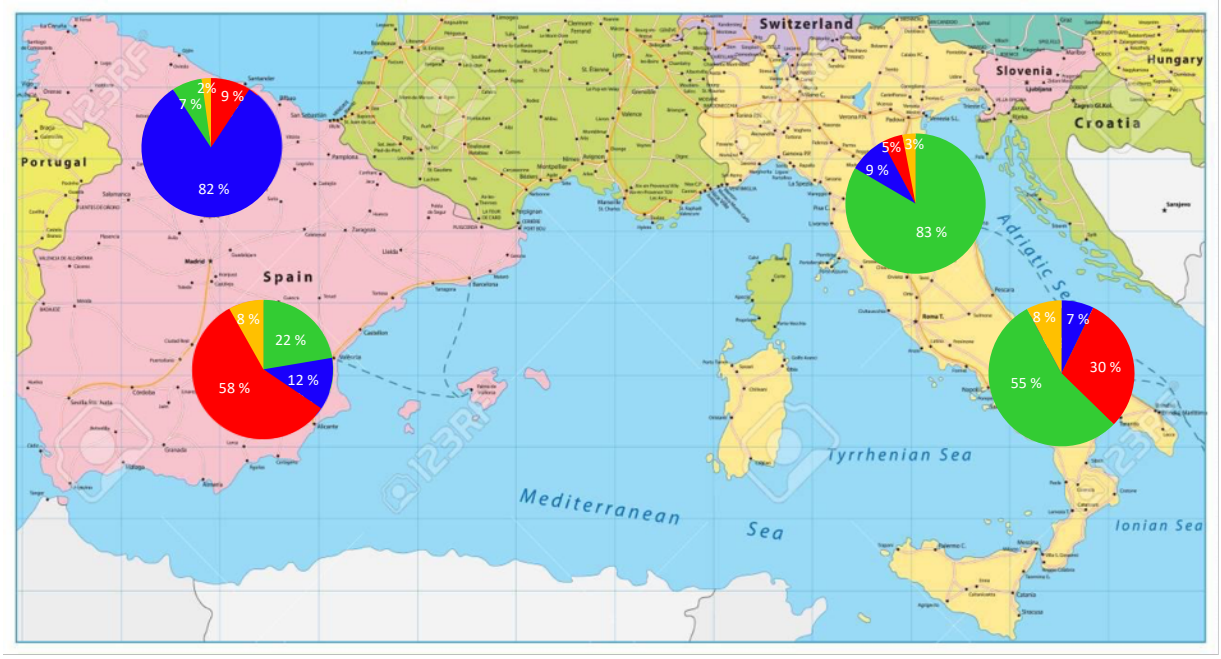

Figure 4.5B: Gene pool distribution of chestnut cultivars for K=4 between the north and the south, both of

Italy and Spain (Red –Italy, Central and Southern Spain (Cluster 1); Yellow – Central and Southern Spain (Cluster 2; Blue –Northern Spain (Cluster 3) and Green – Mainly Northern Italian and France (Cluster 4)).

4.6 Conclusions

This study aimed to be a systematic genetic structure analysis of grafted chestnuts, traditional varieties and wild populations in four different countries (Spain, Italy, France and Portugal).

In particular our specific aims were a) to contributed to the expansion of the European Genetic Dataset based on the reference SSR for the identification of the main chestnut cultivars; b) to study the genetic structure in order to define the historical connections that occurred in the past; c) to define the fixation index between wild chestnuts and cultivated local varieties.

The analysis performed on 319 unique genotypes indeed contributes to expand the utility of the Chestnut European Genetic Dataset, identifying 138 genotypes from northern Italy that helped to understand the genetic structure of European chestnut in southern Europe with a genetic differentiation between northern and central-southern Italy as it was found before between northern and southern Spain. Moreover, the important introgression from both genetic origins of Italy in central-southern Spain, and at much lower level in the northern Iberian Peninsula.

The STRUCTURE analysis highlighted that the chestnut trees are characterized by a complex structure and considerable genetic diversity, confirming results of previous studies with a greater number of accessions (Gobbin et al., 2007; Martín et al., 2009; Martín et al., 2012; Mattioni et al., 2013; Pereira-Lorenzo et al., 2010 and 2019). The study revealed the existence of two genetically and, to a large extent, geographically distinct groups of chestnut populations corresponding to a Spanish and an Italian cluster.

The variation found between the groups and within individual groups, reflects a combination of historical migration / selection processes, represented by the high number of admixtures, as for the chestnut populations of South Spain and South of Italy. Factors of adaptation to different environments lead to a wide genetic variation in a limited population structure such as the Tuscan-Emilian Apennines and the Galicia region. These data are important to guarantee identification and future preservation of specific cultivars.

Finally, this study showed a low genetic structure between wild chestnuts and cultivated chestnuts from different European countries (Spain, Portugal, France and Italy), that emphasizes the outcross nature of chestnut tree.

In conclusion, the results of this study could contribute to better understand the human role in the evolution of this species, to expand its genetic knowledge and open up the possibility of making new orchards in Europe. This would be the starting point for future selection programs useful for the revival of chestnut as a fresh product or for the production of flours.

4.7 Bibliography

Aranzana, M.J.; Abbasi, E.K.; Howard, W.; Arus, P. (2010). Genetic variation, population structure and linkage disequilibrium in peach commercial varieties. BMC Genet. 11, 69.

Auge, H.; Brandl, R. (1997). Seedling recruitment in the invasive clonal shrub, *Mahonia aquifolium* Pursh (Nutt.). Oecologia (Berl.), 110(2): 205–211. doi: 10.1007/s004420050151.

Bouffartigue, C.; Debille, S.; Fabreguette, O.; Ramos Cabrer, A.; Pereira-Lorenzo, S.; Flutre, T.; Harvengt, L. (2019). Genetic diversity and structure of sweet chestnut (*Castanea sativa* Mill.) in France: At the intersection between Spain and Italy
bioRxiv 792259; doi: https://doi.org/10.1101/792259

Bouffartigue, C.; Debille, S.; Fabreguettes, O.; Cabrer, A. R.; Pereira-Lorenzo, S.; Flutre, T.; & Harvengt, L. (2020). High admixture between forest and cultivated chestnut (*Castanea sativa* Mill.) in France. bioRxiv, 792259.

Bourguiba, H.; Khadari, B.; Krichen, L.; Trifi-Farah, N.; Mamouni, A.; Trabelsi, S.; ... & Audergon, J. M. (2012). Genetic diversity analysis of Mediterranean apricot geographic groups. Acta Horticulturae, (966), 269-274.

Buck, E.J.; Hadonou, M.; James, C.J.; Blakesley, D.; Russell, K. (2003) Isolation and characterization of polymorphic microsatellites in European chestnut (*Castanea sativa Mill*.) Mol Ecol Notes 3:239–241. https://doi.org/10.1046/j.1471-8286.2003.00410.x

Chiocchini, F.; Mattioni, C.; Pollegioni, P.; Lusini, I.; Martín, M. A.; Cherubini, M.; ... & Villani, F. (2016). Mapping the genetic diversity of *Castanea sativa*: exploiting spatial analysis for biogeography and conservation studies. Journal of Geographic Information System, 8(02), 248.

Conedera, M. (1996). Die Kastanie: der Brotbaum. Vergangenheit, Gegenwart und Zukunft der ‘‘Waldfrucht par excellence’’. Bu¨ndnerwald, 49: 28–46.

Conedera, M.; Krebs, P.; Tinner, W.; Pradella, M.; & Torriani, D. (2004). The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. Vegetation History and Archaeobotany, 13(3), 161–179. https://doi.org/10.1007/s00334-004-0038-7

Díaz-Losada, E.; Tato Salgado, A.; Ramos-Cabrer, A. M.; Díaz-Hernández, B.; & Pereira-Lorenzo, S. (2012). Genetic and geographical structure in grapevines from northwestern Spain. Annals of applied biology, 161(1), 24-35.

Dos Santos, A. R. F.; Ramos-Cabrer, A. M.; Díaz-Hernández, M. B.; & Pereira-Lorenzo, S. (2011). Genetic variability and diversification process in local pear cultivars from northwestern Spain using microsatellites. Tree genetics & genomes, 7(5), 1041-1056.

Earl, D.A.; Vonholdt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Emanuelli, F.; Lorenzi, S.; Grzeskowiak, L.; Catalano, V.; Stefanini, M.; Troggio, M.; ... & Grando, M. S. (2013). Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC plant biology, 13(1), 39.

Evanno, G.; Regnaut, S.; & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, 14(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fernández Cruz, J.; Fernández López, J. (2016) Genetic structure of wild sweet chestnut (*Castanea sativa* Mill.) populations in northwest of Spain and their differences with other European stands. Conserv Genet 17:949–967. https://doi.org/10.1007/s10592-016-0835-4

Fideghelli, C. (2017). Atlante dei fruttiferi autoctoni italiani. Crea - Centro di frutticoltura, Roma.

Fineschi, S.; Malvolti, M.E.; Morgante, M.; Vendramin, G.G. (1994). Allozyme variation within and among cultivated varieties of sweet chestnut (*Castanea sativa*). Can. J. For. Res. 24(6): 11601165. doi:10.1139/x94-153.

Forneck, A. (2005). Plant breeding: clonality — a concept for stability and variability during vegetative propagation. Prog. Bot. 66: 164–183. doi: 10.1007/3-540-27043-4_8

Fonti Forcada, C.; Gogorcena, Y.; Moreno, M.A. (2013). Fruit sugar profile and antioxidants of peach and nectarine cultivars on almond × peach hybrid rootstocks. Scientia Horticulturae 164: 563-572.

Gobbin, D.; Hohl, L.; Conza, L.; Jermini, M.; Gessler, C.; Conedera, M. (2007) Microsatellite-based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. Genome 50:1089–1103. doi: 10.1139/G07-086

Galderisi, U.; Cipollaro, M.; Di Bernardo, G.; De Masi, L.; Galano, G.; & Cascino, A. (1998). Molecular typing of Italian sweet chestnut cultivars by random amplified polymorphic DNA analysis. The Journal of Horticultural Science and Biotechnology, 73(2), 259-263.

Hummer, K.E.; Dempewolf, H.; Bramel, P.; Markham, R.; Stover, E. (2015) Status of global strategies for horticultural fruit crops. Acta horticulturae

Kalinowski, S. T.; Taper, M. L.; & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology, 16(5), 1099–1106. https://doi.org/10.1111/j.1365294X.2007.03089.x

Kampfer, S.; Lexer, C.; Glossl, J.; & Steinkellner, H. (1998). Characterization of (GA)(n) microsatellite loci from Quercus robur. Hereditas, 129(2), 183–186. https://doi.org/10.1111/j.1601-5223.1998.00183.x

Krebs, P.; Conedera, M.; Pradella, M.; Torriani, D.; Felber, M.; Tinner W. (2004) Quaternary refugia of the sweet chestnut (*Castanea sativa* Miller): an extended palynological approach. Vegetation History and Archaeobotany, 13, 145–160.

Lusini, I.; Velichkov, I.; Pollegioni, P.; Chiocchini, F.; Hinkov, G.; Zlatanov, T.; ... & Mattioni, C. (2014). Estimating the genetic diversity and spatial structure of Bulgarian *Castanea sativa* populations by SSRs: implications for conservation. Conservation genetics, 15(2), 283-293.

Maguire, T. L.; Collins, G. G.; & Sedgley, M. (1994). A modified CTAB DNA extraction procedure for plants belonging to the family proteaceae. Plant Molecular Biology Reporter, 12(2), 106–109. https://doi.org/10.1007/BF02668371

Marinoni, D.; Akkak, A.; Bounous, G.; Edwards, K.J.; and Botta, R. (2003). Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). Mol. Breed. 11 (2), 127–136 https://doi.org/10.1023/ A:1022456013692.

Marshall, T.C.; Slate, J.; Kruuk, L.E.B.; Pemberton, J.M. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655

Martín, M.A.; Alvarez, J.B.; Mattioni, C.; Cherubini, M.; Villani, F.; Martín, L.M. (2009) Identification and characterization of traditional chestnut varieties of southern Spain using morphological and simple sequence repeat (SSRs) markers. Ann Appl Biol 154:389–398. https://doi.org/10.1111/j.1744-7348.2008.00309.x

Martín, M.A.; Mattioni, C.; Molina, J.R.; Alvarez, J.A.; Cherubini, M.; Herrera, M.A.; Villani, F.; Martín, L.M. (2012) Landscape genetics structure of chestnut (*Castanea sativa* Mill.) in Spain. Tree Genet Genomes 8:127–136. https://doi.org/10.1007/s11295-011-0427-x

Martín, M. A.; Monedero, E.; & Martín, L. M. (2017). Genetic monitoring of traditional chestnut orchards reveals a complex genetic structure. Annals of Forest Science, 74(1). https://doi.org/10.1007/s13595-016-0610-1

Martín, M. A.; Mattioni, C.; Cherubini, M.; Villani, F.; & Martín, L. M. (2017). A comparative study of European chestnut varieties in relation to adaptive markers. Agroforestry systems, 91(1), 97-109.

Mattioni, C.; Cherubini, M.; Micheli, E.; Villani, F.; Bucci, G. (2008) "Role of domestication in shaping *Castanea sativa* genetic variation in Europe". Tree Genetics & Genomes 4(3):563574.

Mattioni, C.; Martin, A.M.; Pollegioni, P.; Cherubini, M.; & Villani, F. (2013). Microsatellite markers reveal a strong geographical structure in European populations of *Castanea sativa* (Fagaceae): Evidence for multiple glacial refugia. American Journal of Botany, 100(5), 951–961. https://doi.org/10.3732/ajb.1200194

Mattioni, C.; Martin, M. A.; Chiocchini, F.; Cherubini, M.; Gaudet, M.; Pollegioni, P.; … Villani, F. (2017). Landscape genetics structure of European sweet chestnut (*Castanea sativa* Mill): indications for conservation priorities. Tree Genetics and Genomes, 13(2). https://doi.org/10.1007/s11295-017-1123-2

Mellano, M.G.; Torello-Marinoni, D.; Boccacci, P. (2018) Ex situ conservation and characterization of the geneticdiversity of Castanea spp. Acta Horticulturae 1–6. doi: 10.17660/ActaHortic.2018.1220.1

Miranda, C.; Urrestarazu, J.; Santesteban, L.G.; Royo, J.B.; Urbina, V. (2010) Genetic diversity and structure in a collection of ancient Spanish pear cultivars assessed by microsatellite markers. J Amer SocHortic Sci 135:428–437

Neri, L.; Dimitri, G.; & Sacchetti, G. (2010). Chemical composition and antioxidant activity of cured chestnuts from three sweet chestnut (*Castanea sativa* Mill.) ecotypes from Italy. Journal of Food Composition and Analysis, 23(1), 23–29. https://doi.org/10.1016/j.jfca.2009.03.002

Peakall, R.; Smouse, P.E. (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Pereira-Lorenzo, S.; Costa, R.; Ramos-Cabrer, A.M.; Ribeiro, C.; Serra da Silva, C.; Manzano, G.; Barreneche, T. (2010) Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. Tree Genet Genomes 6:701–715. https://doi.org/10.1007/s11295-010-0285-y

Pereira-Lorenzo, S.; Lourenço Costa, R.M.; Ramos-Cabrer, A.M.; Ciordia-Ara, M.; Marques Ribeiro, C.A.; Borges, O.; Barreneche, T. (2011) Chestnut cultivar diversification process in the Iberian Peninsula, Canary Islands, and Azores. Genome 54:301–315. https://doi.org/10.1139/g10-122

Pereira-Lorenzo, S.; Ramos-Cabrer, A. M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, M.B.; … Martín, Á. (2017). Database of European chestnut cultivars and definition of a core collection using simple sequence repeats. Tree Genetics and Genomes, 13(5). https://doi.org/10.1007/s11295-017-1197-x

Pereira-Lorenzo, S.; Ramos-Cabrer, A. M.; Barreneche, T.; Mattioni, C.; Villani, F.; Díaz-Hernández, B.; Martín, A. (2019). Instant domestication process of European chestnut cultivars. Annals of Applied Biology, 174(1), 74–85. https://doi.org/10.1111/aab.12474

Pereira-Lorenzo, S.; Bischofberger, Y.; Conedera, M.; Piattini, P.; Crovadore, J.; Chablais, R.; Rudow, A.; Hatt, S.; Ramos-Cabrer, A.M.; Barreneche, T.; Lefort F. (2020) Reservoir of the European chestnut diversity in Switzerland. Biodivers Conserv (2020). https://doi.org/10.1007/s10531-020-01970-2

Porras-Hurtado, L.; Ruiz, Y.; Santos, C.; Phillips, C.; Carracedo, A.; Lareu, M. (2013) An overview of STRUCTURE: applications, parameter settings, and supporting software. Front Genet 4:98. https://doi.org/10. 3389/fgene.2013.00098

Pritchard, J.K.; Stephens, M.; Donnelly, P. (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pitte, J.R. (1986). Terres de castanide. Hommes et paysages du chataignier de l'Antiquite´a nos jours. Librairie Artheme Fayard.

Quintana, J.; Contreras, A.; Merino, I.; Vinuesa, A.; Orozco, G.; Ovalle, F.; & Gomez, L. (2015). Genetic characterization of chestnut (*Castanea sativa* Mill.) orchards and traditional nut varieties in El Bierzo, a glacial refuge and major cultivation site in northwestern Spain. Tree Genetics and Genomes, 11(1). https://doi.org/10.1007/s11295-014-0826-x

Silvanini, A.; Marinoni, D. T.; Beccaro, G. L.; & Ganino, T. (2011). La caratterizzazione varietale del germoplasma di Castanea sativa Mill. Italus Hortus, 18, 47-61.

Sint, D.; Raso, L.; & Traugott, M. (2012). Advances in multiplex PCR: Balancing primer efficiencies and improving detection success. Methods in Ecology and Evolution, 3(5), 898–905. https://doi.org/10.1111/j.2041-210X.2012.00215.x

Torello Marinoni, D.; Akkak, A.; Beltramo, C.; Guaraldo, P.; Boccacci, P.; Bounous, G.; … Botta, R. (2013). Genetic and morphological characterization of chestnut (*Castanea sativa* Mill.) germplasm in Piedmont (north-western Italy). Tree Genetics and Genomes, 9(4), 1017–1030. https://doi.org/10.1007/s11295-013-0613-0

Urrestarazu, J.; Royo, J. B.; Santesteban, L. G.; & Miranda, C. (2015). Evaluating the influence of the microsatellite marker set on the genetic structure inferred in *Pyrus communis* L. PLoS ONE, 10(9), 1–23. https://doi.org/10.1371/journal.pone.0138417

CHAPTER 5: A CASE OF GERMOPLASM VALORIZATION – 'ROSA ROMANA' apple

5.1 Abstract

A molecular characterization on the ancient variety 'Rosa Romana' was carried out to improve biodiversity knowledge and preserve these trees from extinction risks.

In this work, 47 accessions were collected throughout an investigation in the Emilia-Romagna region (Italy) and particularly in the mountain area of the Bologna province (19 collection sites). The analysis at molecular level by using 15 SSR (microsatellites) identified two main genotype groups and ascertained their relationship with several phenotypic traits. This two clusters contained most of the collected accessions, while remaining genotypes differ clearly, according with the phenotypic diversity on the behavior of the trees or fruits.

This study also revealed the highest quality traits of 'Rosa Romana' apple grown in the Apennines mountain around Bologna (in a range between 400 and 1000 m.s.l.) if compared to each main clone produced at the lowland corresponding to the Bologna University Agricultural Experimental Station (30 m.s.l). Therefore, the apple quality as color, appearance, taste (flesh firmness and texture, sweetness, acidity, aroma, polyphenol soreness), were improved in the higher altitudes. In conclusion, the results of this environmental and genetic investigation on the residual cultivation of 'Rosa Romana' apple provided a genomic validation of its best identified clones (correspondent to the main two clusters), which now can be recovered and promoted as new planting, with a own brand 'Rosa Romana' produced in the Apennines mountain of north Italy.

Keywords: *Malus X domestica* Borkh., accessions, SSR, molecular characterization, qualitative parameters, Cluster analysis.

5.2 Introduction

Apple (*Malus x domestica* Borkh.) is the main fruit crop of temperate regions of the world such as Europe, the west area of Turkestan and the south-east and Centre of Asia (Velasco et al., 2010), in terms of production levels. It occupies a central position as nutritional value and also in culture, art and folklore (Janick, 2005; Cornille et al., 2014). Much of the genetic diversity of the old cultivated apples is currently maintained in germoplasm repositories and amateur collections (Alessandri et al., 2016).

The 'modern' apple was domesticated in Central Asia from *Malus sieversii* (Velasco et al., 2010; Cornille et al., 2012; Volk et al., 2005 and 2013) and was brought to Europe through human migrations between 6,000 and 3,000 years ago (Janick, 2005; Ross-Ibarra et al., 2007; Cornille et al., 2012). Humans have been exploiting, selecting, and transporting apples for centuries, and several thousand apple cultivars have been historically documented (Ross-Ibarra et al., 2007; Cornille et al., 2014). Over time, many of the 'old' varieties of Italian apple trees however have been marginalized and now are present only in small local area. In some cases, only single specimens of trees have survived, a memory of a glorious past, while unfortunately some genotypes have been forever lost.

The first historical quotation of a Rosa apple in Emilia-Romagna region dates back to the 16th century, by the famous naturalist Aldrovandi and a first pictorial representation was released at the end of the 17th century by Bartolomeo Bimbi, a famous painter of the Medici's court, who painted more than one hundred apple varieties and reported their correct names (Fideghelli, 2016).

'Rosa Romana' was grown in the Reno Valley for its high fruit quality (flavor, taste, texture), high storability, easy harvesting, short juvenile phases, synchronicity in blooming and fruit ripening (Gregori et al., 2013). The Reno valley represents the propagation point of the 'Rosa Romana' variety. Probably because this valley was a passage area during the Roman age since it allowed the connection between the regions of Emilia-Romagna and Tuscany.

In 1929, this variety represented the 25% of the apple production in the Bologna area. However, this apple has almost disappeared in the time frame of thirty years (Sansavini et al., 2018).

The 'Rosa Romana' fruit descriptor evidences a flattened shape, a short peduncle, a yellow ground color with bright red on 20-30% of the skin (only in the mountain areas). The fruit has a thick and slightly waxy skin when the apple is ripe. Normally the peduncle cavity is covered by russeting (Figure 1). Flesh is firm,

juicy, fine, non-crispy and non-astringent. The taste highlights a well-balanced equilibrium of sweetness and acidity with a slightly bitter aftertaste. Storability without refrigeration is excellent (even till 4 months), but a controlled atmosphere it can be suggested for much longer storage. Fruits are susceptible to physiological disorders such as bitter pit, especially in young, too vigorous, over-nourished trees. The picking time is late autumn as well as the ripening time (Fideghelli et al., 2017; Sansavini et al., 2018).

The international literature on commercial, nutritional and genetic information relating to the 'Rosa Romana' variety is scanty despite its cultivation and use in the Reno Valley dates back to ancient times (at least since the Roman age).

Sansavini et al. (2018) showed like this variety currently consumed and promoted in the market by local farmers pointing out its health and gustatory qualities together with its strong link with the Reno Valley territory and history – heritage which deserves proper protection and interest.

Farneti et al. (2015) evidenced that ancient apple varieties as 'Rosa Romana' have a higher level of phenols compared to commercial apple cultivars. In particular, the organic acids and the phenolics compounds were significantly influenced and dependent by human selection. Bignami et al. (2001) carried out the only reported work on the variability of qualitative traits of the 'Rosa Romana' genotype. The analysis of nutrients and polyphenols showed the high quality of this apple.

The local germplasm of apple varieties represents a good source for breeding programs so as to guarantee the availability of a wide genetic variability (Bignami et al., 2001). To preserve this genotype, in particular, it is necessary to identify and classify the possible variables that can be differentiated over the long cultivation time.

Other two accessions grown in this area are: 'Rosa Romana Gentile' and 'Rosa Nostrana'. 'Rosa Romana Gentile' differs from 'Rosa Romana' for its low russeting, the smaller extension of the red fruit skin overcolor (Figure 5.1c), the greater greasiness and for its earlier ripening while 'Rosa Nostrana' differs from the other apple Roses for its conical fruit shape (Figure 5.1b), its high greasiness after storage and for the not excellent sensorial traits, susceptible to scald.

It is important to be not confuse this apple genotypes with the other Rose apple varieties which present distinctive characteristics such as different fruits and lenticelllar shape and coloration.

As showed by Figure 1, ‘Rosa d'Osta’ and ‘Rosa Mantovana’ differ mainly in the rounder shape compared to the flat shape of ‘Rosa Romana’ (Figure 5.1 d, e). In addition, ‘Rosa d’Osta’ is characterized by a scarce over-color and absence of rust. ‘Rosa d’Oliveto’ has a longer stalk and a more uniform red color diffused at lenticellar level compared to 'Rosa Romana' (Figure 5.1f). Lastly, ‘Rosa Marchigiana’ presents a shorter stalk and more evident lenticels on the skin compared with ‘Rosa Romana’ fruit (Figure 5.1g).

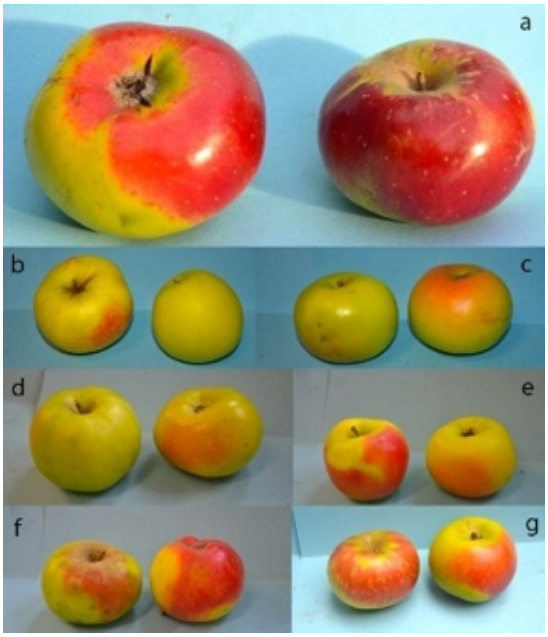

Figure 5.1 – Seven ‘Rosa’ varieties which differ for several fruit traits but having the same root name ‘Rosa’: a) ‘Rosa Romana’ apple (accession n°17, cluster 1); b) ‘Rosa Nostrana’ (accession n°3); c) ‘Rosa Romana Gentile’ (accession n°2); d) ‘Rosa d'Osta’ (accession n°15); e) ‘Rosa Mantovana’; f) ‘Rosa d'Oliveto’ (accession n°14); g) ‘Rosa Marchigiana’ (accession n°9).

Molecular markers [Simple Sequence Repeat (SSR)] are fundamental for verifying the correct propagation in the nurseries, the true-to-type correspondence and for reducing redundancies in collections. In particular, microsatellites are considered the most suitable and useful markers for exploring the genetic diversity because they are i) abundant and well distributed in the genome; ii) codominant and multi-allelic; and iii) analyzed by multiplexed PCR (Polymerase Chain Reaction) assays (Hayden et al., 2008; Patocchi et al., 2009; Urrestarazu et al., 2016; Larsen et al., 2017; Testolin et al., 2019; Baric et al., 2020).

The aims of this work are: i) the phenotyping of the selected clones for the fruit quality traits and the relative comparison with fruits from the lowland; ii) the identification of the genetic variability present among the 'Rosa Romana', 'Rosa Romana Gentile' and 'Rosa Nostrana' accessions sampled in Reno valley (hill around 400-600 m.s.l. and mountain area around 600-1000 m.s.l.).

The identification of historical trees and best reference plants for propagation are fundamental steps for the development of nursery activities. This will also promote and support the exploitation and protection of such ancient Italian apple cultivars. An increased interest in local products and ancient flavors is expected to follow.

5.3 Materials and methods

5.3.1 Plant material

The fruit and leaf samples were collected by historical trees in 20 different locations of the Reno Valley and in two sites of the Bologna plain (for a total of 47 accessions, Table 5.1; Figure 5.2). The sampled trees from the mountain area are grafted on non-characterized apple seedling and maintained *in situ* collections by guardian farmers and are grown following the organic farming guidelines.

The sample list includes 3 Rosa accessions from the Marche region and 9 Rosa accessions from the apple collection of the University of Bologna (Table 5.1). Fruit samples for quality analyses were collected in two consecutive harvesting years (2018-2019).

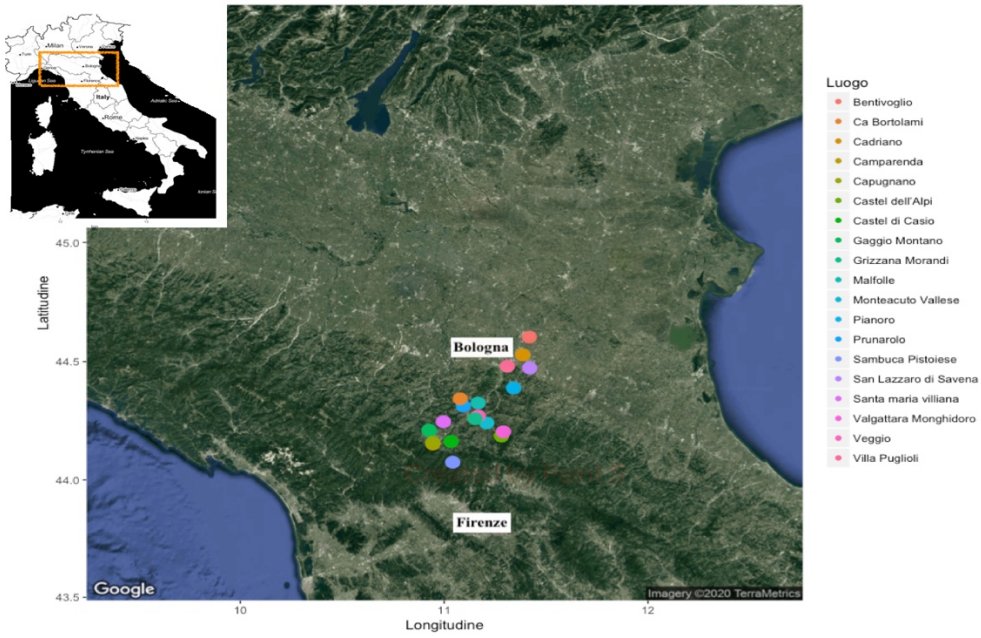

Figure 5.2 - Map of the sites where the samples were collected, elaborated with the software R, package (ggmaps), source from Google, Emilia-Romagna region, Italy.

Table 5.1 - List of analyzed accessions and their sampling sites. BO: Bologna; PT: Pistoia.

	Accessions	Sampling Area	Altitude (m.s.l.)
#1	Rosa Romana	Santa Maria Villiana	643
#2	Rosa Romana Gentile	Santa Maria Villiana	643
#3	Rosa Nostrana	Santa Maria Villiana	643
#4	Rosa 1	Gaggio Montano (BO)	944
#5	Rosa 2	Gaggio Montano (BO)	944
#6	Rosa 3	Gaggio Montano (BO)	944
#7	Rosa Romana	Castel di Casio (BO)	533
#8	Rosa Romana Gentile	Castel di Casio (BO)	533
#9	Rosa Marchigiana R101	Marche	600
#10	Rosa Marchigiana R108	Marche	600
#11	Rosa Marchigiana R60	Marche	76
#12	Rosa Nostrana o Locale	Bentivoglio (BO) - Villa Smeraldi	19
#13	Rosa Romana	Bentivoglio (BO) - Villa Smeraldi	19
#14	Rosa D'Oliveto	Bentivoglio (BO) - Villa Smeraldi	19
#15	Rosa Osta	Cadriano (BO) - UNIBO	32

#16	Rosata Russolina	Cadriano (BO) - UNIBO	32
#17	Rosa Romana	Castal dell'Alpi (BO)	737
#18	Rosa Romana	Monteacuto (BO)	915
#19	Rosa Romana	Malfolle (BO)	500
#20	Musabo Rossa	Castal dell'Alpi (BO)	737
#21	Rugginosa	Castal dell'Alpi (BO)	737
#22	Rosa Romana	Bologna	40
#23	Rosa R. Gentile	Bologna	40
#24	Rosa 1	Bologna - Villa Puglioli	270
#25	Rosa 2	Bologna - Villa Puglioli	270
#26	Rosa 3	Bologna - Villa Puglioli	270
#27	Rosa Romana	Ecchia- Prunarolo (BO)	193
#28	Rosa Romana	Ca Bortolami (BO)	334
#29	Rosa Romana	Ca Bortolami (BO)	334
#30	Rosa Romana	Ca Bortolami (BO)	334
#31	Rosa Romana	Grizzana Morandi (BO)	547
#32	Rosa Romana	Grizzana Morandi (BO)	547
#33	Rosa Romana	Veggio (BO)	550
#34	Rosa Romana	Veggio (BO)	550
#35	Rosa Romana	Pianoro (BO)	200
#36	Rosa Romana	Sambuca Pistoiese (PT)	504
#37	Rosa Romana	Capugnano (BO)	820
#38	Rosa Romana	Camparenda (BO)	800
#39	Rosa Romana	Valgattara (BO)	700
#40	Rosa Romana	Camparenda (BO)	815
#41	Rosa Romana (strain 24)	Cadriano (BO) – UNIBO	32
#42	Rosa Romana Gentile (strain 43)	Cadriano (BO) – UNIBO	32
#43	Rosa Romana (strain A23)	Cadriano (BO) – UNIBO	32
#44	Rosa Mn_(Tn)	Cadriano (BO) – UNIBO	32
#45	Mela Rosa (Pd)	Cadriano (BO) – UNIBO	32
#46	Rosa D'oliveto	Cadriano (BO) – UNIBO	32
#47	Mela Rosa (Tn)	Cadriano (BO) - UNIBO	32

5.3.2 Apple phenotyping: qualitative parameters

After harvesting, fruits were immediately stored at cold room at 0°C with high humidity for about one month and then kept out in shelf-life for three days to ripen the fruit (Gorny and Kader, 1997). Fruit weight (g), percentage of overcolour, russeting (%), bitterness (%), soluble solid (%) and organic acid content (malic acid g/L) have been evaluated on pools of 10 fruits (Gregori et al., 2013). Firmness was measured by a penetrometer (11 mm diameter probe) on apple surfaces from opposite sides of each fruit (Kg/cm2). A pool of ten apples was analysed for each sampled tree. Soluble Solids Content (SSC) was determined by a digital refractometer (Atago) on filtrated apple juice obtained by homogenizing two slices taken from each of the 10 fruits. Titratable acidity (TA) was detected by automatic titrator (Crison). Twenty millilitres of juice diluted with additional twenty millilitres of distilled water were titrated to pH 8,1 with 0,25N NaOH. Trees and fruits were evaluated with pomological descriptors in field after fruit harvesting, according to Gregori et al. (2013). Percentage of fruit skin overcolour was empirically classified. Bitterness was estimated by a sensory panel test by ranking the evaluations in classes from 1 to 9 on an empirical scale (1, absence; 9, maximum intensity). The data were processed (i) by unpaired t-test to compare means between fruits collected in mountains (the Reno Valley) vs reference those of the plains (Bologna); ii) by variance analysis (ANOVA) according to Fisher's Least Significant Difference (LSD) test at P = 0,05 to compare the single samples of different mountain areas with 4 number of replicates per sample in each of the two harvesting years.

5.3.3 DNA extraction, SSR genotyping and allele characterization

For each accession, genomic DNA was extracted from 50 mg of young freeze-dried leaves following the standard CTAB protocol (Maguire et al., 1994). Genomic DNA was quantified by Nanodrop^{TM} ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and diluted to 10 ng/µl. Samples were analyzed with 15 SSR markers according to Liang et al. (2015).

The molecular data were compared and aligned with the SSR profiles of the references conserved in the collection of the Department of Agricultural Sciences and Technologies of the University of Bologna (DISTAL) in Cadriano: 'Rosa Romana' (strain 24), 'Rosa Romana Gentile' (strain 43), 'Rosa Romana' (strain A23), 'Rosa d'Osta', 'Rosata Russolina', 'Rosa Mantovana [Trento (TN)]', 'Mela Rosa [Padova (PD)]', 'Rosa d'Oliveto' and 'Mela Rosa (TN)' to better understand the variability present among the accessions collected.

The number of alleles per locus (k), the observed and the expected heterozygosities (Ho and He) and polymorphism information content (PIC) of the SSRs were estimated using the CERVUS Software Version 3.0.3 (Marshall et al., 1998; Kalinowski et al., 2007). A PIC value greater than 0.7 was considered to be highly polymorphic and informative for a certain locus. Subsequently, the dendrogram tree was calculated by using the NTSYSpc 2.0 software with the coefficient of DICE (Dice, 1945) and the software R (Project for Statistical Computing). The cluster analysis and the construction of the dendrogram related to genetic distances were obtained by the UPGMA method (Unweighted Pair-Grop Method).

5.4 Results

5.4.1 Pomological and qualitative parameters

Pomological observation and the analyses of several fruit quality parameters (fruit weight, percentage of overcolor, russeting, bitterness, soluble solid and organic acid content) was carried out on a pool of 10 representative fruits for the Rose genotypes (‘Rosa Romana’, ‘Rosa d’Osta’, ‘Rosa Romana Gentile’, ‘Rosa Nostrana’,’Rosa d’Oliveto’, ‘Rosa Marchigiana’ e ‘Rosa Mantovana’).

The statistical analysis was initially elaborated by comparing the pools of individual trees harvested of ‘Rosa Romana’ in different locations of the Reno Valley with each other and those in the Bologna plains (Experimental farm of the University of Bologna, Cadriano and Villa Smeraldi).

First of all, the fruits from accessions belonging to the ‘Rosa Romana’ were not statistically distinguishable for all the analyzed traits. In fact, all the qualitative parameters analyzed did not show significant differences among the samples which presented the phenotypic characteristics typical of the variety (Figure 5.1a).

Differences were observed by comparing samples of ‘Rosa Romana’ collected in the Reno Valley to those harvested in the plain (Table 5.2).

In particular, the apples of the plains had a greater fruit weight but with a reduced fruit overcolor (Table 5.2). As reported in Table 5.2, the ‘Rosa Romana’ plain samples also presented a lower russeting in the peduncular region of the fruit. In addition, the juice of the apples of the plains had 1° Brix less than those of the mountains while the acidity was about a half. These data evidence that the fruit quality traits are enhanced in areas at medium and high altitude (400-800m). This observation was also confirmed by the analysis of the variance (ANOVA).

Table 5.2 - Pomological profile of the 'Rosa Romana' apples from Cluster 1 and Cluster 2;Data are divided separating Mountain and Plain samples, collected in locations with different altitude [Gaggio Montano (944 m a.s.l.), Castel dell'Alpi (737 m a.s.l.) and S. Maria Villiana (643 m a.s.l.)] and in Plain (Bologna) [Cadriano (32 m a.s.l.), S. M. Bentivoglio (19 m a.s.l.)].

Area	**Mountain (Reno Valley)**				**Plain (Bologna)**		
N° of Samples	4. Rosa 1	17. Rosa Romana	1.Rosa Romana	**Means**	41.Rosa Romana	13.Rosa Romana	**Means**
Medium weight (g)	124b	167a	153a	150 B	230a	198b	214 A
Over-colored (%)	35a	35a	23b	24 A	3b	7a	5 B
Russeting (%)	20a	26a	11b	14 B	24a	26a	25 A
Bitterness (index 1-9)*	8,7a	7b	7,7a	7,8 A	6a	5,6a	5,8 B
Brix (%)	16,8a	16b	15,8b	15,1 A	14,1a	13,9a	14 B
Firmenss (kg)	10,4a	8,39c	9,04b	9,27 A	6,79a	6,91a	6,85 B
Acidity (g/l malic acid)	4,9 b	6,5a	6,9a	7,0 A	3,4a	3,6a	3,5 B

Analysis of variance (ANOVA) according to Fisher's LSD. The medium with different letters are significantly different (P≤0.05); *Index of empirical scale (1, absence; 9, maximum intensity)

Finally, data collected on fruits of other Rosa accessions (such as 'Rosa Nostrana') showed difference respect the 'Rosa Romana' accessions (Table 5.3). In particular, 'Rosa Nostrana' is differing from the other Roses for a conical shape of the fruit, less percentage of russeting of the skin apple (1%) and less pulp firmness at harvest (4,24 Kg/cm^2), a high greasiness after storage and for the organoleptic characteristics (Figure 5.1, Table 5.3).

Table 5.3 - Pomological profile of the other apple varieties which differ for several fruit traits from the 'Rosa Romana' phenotype but having partially the same root name 'Rosa'.

Accessions	Rosa Nostrana	Rosa Romana Gentile	Rosa d'Osta	Rosa Mantovana	Rosa d'Oliveto	Rosa Marchigiana
Place	S. Maria Villiana (BO)	S. Maria Villiana (BO)	Cadriano (BO)	Cadriano (BO)	Bentivoglio (BO)	Macerata (MC)
Altitude (m a.s.l)	643	643	32	32	19	600
Fruit weight (g)	186a	148c	162b	82e	139d	135d
Overcolor (%)	22b	6c	1d	32a	30a	38a
Russeting (%)	1a	2a	3a	3a	2a	2a
Brix (%)	14,5a	13,9b	14b	13c	14,6a	14,1ab
Firmness (Kg)	4,24d	9,29a	7,30b	5,99c	7,56b	7,49b
Acidity (g/l malic acid)	5,3b	5,4b	3,8d	2,9e	6,4a	3,9c

Analysis of variance (ANOVA) according to Fisher's LSD. The medium with different letters are significantly different ($P \leq 0.05$)

5.4.2 SSR and Cluster analysis

The 47 samples collected were amplified with 15 pairs of primers already used by scientific community for their good discriminating ability (Liang et al., 2015). An average of 9 alleles per locus are observed for a total of 126 alleles. For each analyzed locus the observed and expected heterozygosity was calculated with CERVUS Software as showed in Table 5.4. Ho ranged from 0.333 for CN444542 to 0.917 CH01A09, CH03G07 and GD12; He ranged from 0,631 for CH02C09 to 0,861 for CH04C07 (Table 5.4). The highest PIC values of 0.850 and 0.803 were observed for the markers CH01H01 and GD12, respectively. Values greater than 0.7, were also observed for all the other SSRs used in the present research. More in detail, SSR loci CH01F02 and CH04C07 were able to distinguish 12 alleles (Table 5.4), thus showing their high discrimination power as reported by Liebhard et al. (2002) and by Cavanna et al. (2008) for apple and pear accessions.

Table 5.4 - Genetic variability parameters: number of alleles per locus (k); observed heterozygosity (HObs); Expected heterozygosis (HExp) and the PIC index.

Locus	K	Hobs	HExp	PIC
CH01A09	8	0.917	0.824	0.784
CH02C09	7	0.583	0.631	0.589
CH03G07	5	0.917	0.687	0.611
CHVf1	7	0.583	0.585	0.544
GD12	10	0.917	0.842	0.803
CH01F2	12	0.833	0.838	0.799
CH02D08	9	0.875	0.820	0.781
CH04C07	12	0.833	0.861	0.828
CH01F03	9	0.875	0.810	0.767
CH01H01	10	0.708	0.883	0.850
CH01H10	8	0.833	0.834	0.794
CH01H02	9	0.542	0.773	0.729
Hi05E07	8	0.875	0.769	0.723
CH05C06	6	0.667	0.809	0.761
CN444542	6	0.333	0.714	0.652

UPGMA cluster analysis, based on DICE genetic distance, evidenced the presence of two main groups of 'Rosa Romana' (namely C1 and C2) that share a high number of alleles, confirming a high degree of similarity between the analyzed samples but also the allele differences (Figure 5.3).

The first cluster includes 12 accessions (#1, #5, #6, #7, #19, #27, #28, #30, #31, #35, #38, #40) with 100 % of similarity with the reference ‘Rosa Romana (strain 24)’ (#41) and ‘Rosa Romana Gentile (strain 43)’(#42) of the University of Bologna and other 3 samples (#22, #23, #39) with very low allele variations (Figure 2). The second cluster could be divided in two subgroups: the former is represented by 5 accessions (#4, #8, #13, #17, #18) that were identical to the reference of the University of Bologna 'Rosa Romana (strain A23)'(#43) and the latter one including samples collected in the area around the Grizzana Morandi site (‘Rosa Romana’ #32, #33 and #34; Figure 5.3). Other samples of ‘Rosa Romana’ not included in these two clusters (#2, #24, #25, #26, #29, #36, #37) should be consider as misnomer (Figure 5.3).

Between those groups of accessions, it should also be noted that there are three accessions of ‘Rosa Marchigiana’ that are very similar but not identical to each other and differ in 8 alleles from the ‘Rosa Romana’ samples. In addition, the two representative samples of the ‘Rosa Nostrana’ accession (#3 and #12) are clearly separated from the ‘Rosa Romana’ clusters and they are distinguishable each other for a number of allele polymorphisms. It is important to underline that the present results could not uniquely identify the ‘Rosa Romana Gentile’ accession, as the samples labelled with this name were all different. In particular, ‘Rosa Romana Gentile (strain 43)’(#42) from Bologna was found to belong to Cluster 1 (with a few polymorphic alleles) while the ‘Rosa Romana Gentile’ (sample #8) was found in Cluster 2. A third sample (number #2) clearly deviates from the main clusters and it represents another misnomer.The dendrogram also included other Rosa accessions clearly separated from the ‘Rosa Romana’ clusters. ‘Rosa Rosata Russolina’ (#16), ‘Rosa Mantovana (TN)’(#47) and ‘Mela Rosa (TN)’(#44) presented identical allelic profile and they can be considered as synonyms (Figure 5.3).

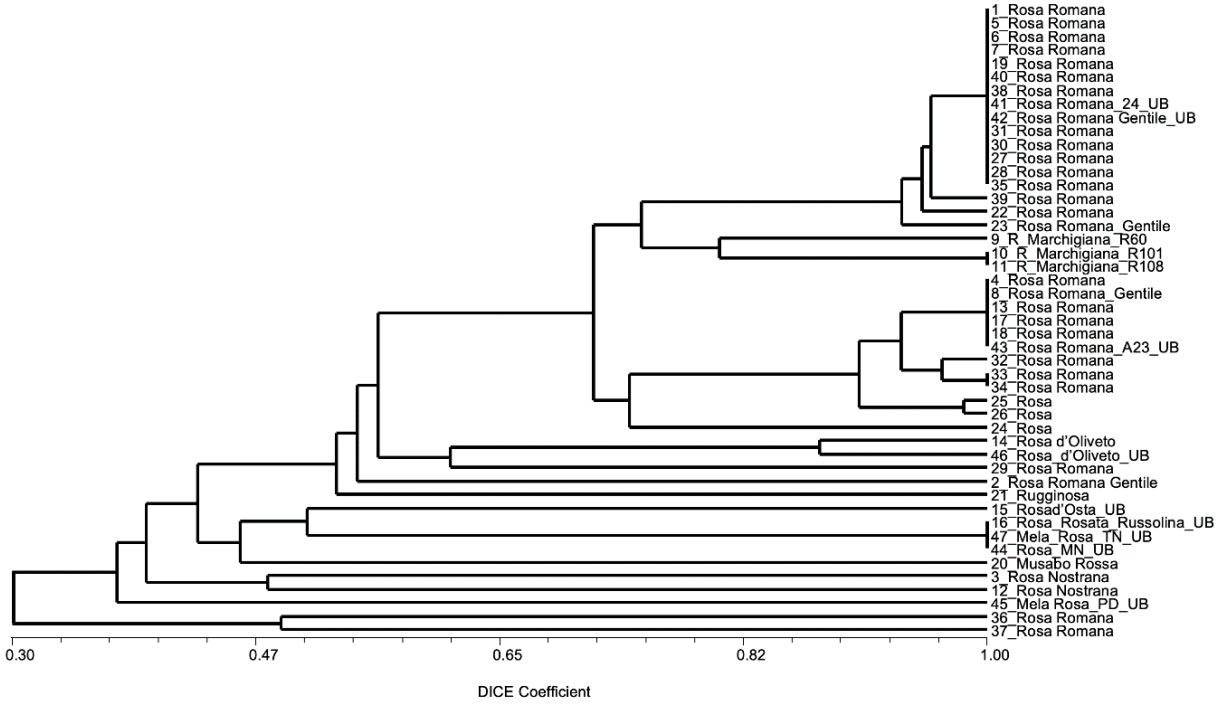

Figure 5.3 - UPGMA tree indicating the relationships among the 47 varieties, built using the NTSYS program.

5.5 Discussion

This study represents the first step of the re-evaluation process for the old Italian variety ‘Rosa Romana’. The area between 400 and 1000 m.l.s within the Reno Valley has been recognized as the traditional area for cultivation of the ‘Rosa Romana’ apple which is particularly widespread therein since the Roman times.

Unfortunately, the cultivation of this apple was abandoned for a long time and replaced by cultivation of conventional apple cultivars. ‘Rosa Romana’ is currently considered of great interest for promoting apple cultivation in the Apennines area in analogy with the model studies carried out on ‘Annurca’, the ancient apple variety of Naples (Melchiade et al., 2007; Iannaccone et al., 2007).

The recovery of surviving trees present in this territory is the first step for the conservation and valorization of such an old variety of apple germplasm (Bignami et al., 2001; Sansavini et al., 2018).

The analysis of the different fruit quality parameters for samples of ‘Rosa Romana’ showed no statistically significant difference with reference to the analyzed traits. The main morphological difference can be found between fruit samples collected in the plains compared to those collected in areas of medium and high altitude (400-800m). Due to the higher altitude the fruits of the latter area present better-quality features, such as less russeting and a more over-colored expression.

At molecular level, the high discrimination power of the 15 SSR used suggests a good differentiation of Rosa apple accessions. The average number of alleles per locus was similar to the values reported by Liang et al., 2015 (in which the CH03G07 locus also resulted less polymorphic).

With the Cluster analysis most of the 47 accessions classified as ‘Rosa Romana’ were divided into two main clusters that share a high number of alleles. In both groups at least a reference accession from the apple germplasm collection of the University of Bologna (‘Rosa Romana (strain 24)’(#41) and ‘Rosa Romana Gentile (strain 43)’(#42) for the first cluster and ‘Rosa Romana (strain A23)’(#43) for the second one) was included. In both clusters only a few accessions showed a limited number of polymorphic alleles, indicating probably the presence of mutations accumulated during the ages. The first cluster included most of the oldest trees which are phenotypically correspondent to the ‘Rosa Romana’ descriptions. The second group, on the other hand, showed a few differences at the phenotypic level, especially in the fruits of the accessions #33 and #34 of ‘Rosa Romana’.

Furthermore, the ‘Rosa Romana Gentile’ and ‘Rosa Nostrana’ accessions presented different genetic profiles that created difficulties in defining the correct genotype.

Finally, data collected on fruits of other Rosa accessions (‘Rosa Mantovana’, ‘Rosa d’Osta’, ‘Rosa d’Oliveto’ and ‘Rosa Marchigiana’) confirmed the difference with the ‘Rosa Romana’ cultivar in relation to molecular data and the pomological descriptions.

These resulted evidenced that ‘Rosa Romana’ is an ancient genotype, propagated in the area of the Tuscan-Emilian Apennines since hundreds of years being well adapted to the different pedoclimatic environments that characterize this area. The adaptation of this genotype to specific agroclimatic conditions has created allele diversity within the samples collected.

The conservation of this variety implies the discrimination of the different accessions with very similar phenotype that are present in the original cultivation area (Sansavini et al., 2018). A certain degree of genetic heterogeneity is acceptable for old varieties (Sansavini et al., 2018). Molecular analysis with microsatellites demonstrated to be the most efficient approach for variety fingerprinting, for recognizing incorrectly labeled material (homonymy and synonymy) and, consequently, for preserving the original ‘Rosa Romana’ genotype.

The identification of the most adequate reference plants is a key step for setting up the correct propagation of this old variety by nurseries and for defining a business plan for its re-evaluation and promotion for a new market niche. In fact, the organoleptic characteristics of the ‘Rosa Romana’ fruits are exalted in the Apennines environmental conditions. If it will be possible to adopt organic cultivation techniques and control the production costs, in all likelihood ‘Rosa Romana’ can represent a new opportunity of income for the farmers of the mountain areas.

5.6 Conclusions

In this study we assessed the phenotype and molecular diversity of ‘Rosa Romana’ accessions collected in residual cultivation of this ancient variety. The area of these old trees was located in the middle and upper Reno Valley in the Tuscan-Emilian Apennines. The sampled accessions (47) also included several varieties (7) locally known with a denomination of Rosa, specified with other secondary names.

SSR results evidenced the presence of two main groups of ‘Rosa Romana’ accessions corresponding to genomic Cluster1 and Cluster2. A number of 12 accessions placed in cluster 1 showed the allelic profile of the oldest trees and of two of the references collected in the apple germplasm collection of the University of Bologna while other 5 are identical to the third reference used for the analyses. All these accessions produce fruits that, as listed in the descriptors, are attributable to the variety ‘Rosa Romana’. This seems to be the effect of mutations that could be probably accumulated during the centuries and that produced some allelic difference between the two clusters. Clusters 1 and 2 represent two clones of ‘Rosa Romana’ and, as consequence, this should be taken into account for a proper identification of reference plants for setting up the nursery propagation activity and supporting the protected variety name for the market.

Despite the observed mutations, ‘Rosa Romana’ appears as relatively stable apple, especially for the fruiting traits: it is not easy to recognize and separate the fruit of the two clusters. ‘Rosa Romana’ is a very pleasant, well recognizable and good tasting apple which get a great improvement by interaction between its genotype and the environmental mountain conditions which exist in agricultural area from 400 to 800 m of altitude in Apennines. Finally, these two clones of ‘Rosa Romana’ are so deeply rooted in the history of the Bologna mountain area that their recovery should be recommended, for a better welfare and the environment and landscape attractiveness.

5.7 Bibliography

Alessandri S., Tartarini S., Buscaroli C., Dondini L. (2016) Biodiversità nel germoplasma del melo emiliano-romagnolo e marchigiano. Riv. di Frutticoltura (N. 11-November).

Baric, S., Storti, A., Hofer, M., Guerra, W., & Dalla Via, J. (2020) Molecular genetic identification of apple cultivars based on microsatellite DNA analysis. I. The database of 600 validated profiles. Erwerbs-Obstbau, 62:117–154.

Scossa, A., Vagnoni, G., Bignami, C. (2001). Evaluation of Old Italian Apple Cultivars by Means of Sensory Analysis. In International Symposium on Sustainable Use of Plant Biodiversity to Promote New Opportunities for Horticultural Production 598, November, 85-90.

Cavanna M, Torello Marinoni D., Bounous G, Botta R. (2008) Genetic diversity in ancient apple germoplasm from northwest Italy, The Journal of Horticultural Science and Biotechnology, 83:5, 549-554, DOI: 10.1080/14620316.2008.11512421

Cornille A., Giraud T., Smulders M.J.M., Roldan-Ruiz I., Gladieux P. (2014) The domestication and evolutionary ecology of apples. Trends in Genetics, February, 30, 2:57-65

Cornille A., Gladieux P., Smulders M.J.M., Roldánruiz I., Laurens F., Le Cam B., Nersesyan A., Clavel J., Olonova M., Feugey L., Gabrielyan I., Zhang X., Tenaillon M.I., Giraud T. (2012) New insight into the history of domesticated apple: Secondary contribution of the european wild apple to the genome of cultivated varieties. PLoS Genetics, 8, 5, e1002703

Dice L. R. (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Fideghelli C. (2017) Atlante dei fruttiferi Italiani, 3 Volumi, III. 1732 pp.

Gorny J.R., Kader A.A. (1997) Low Oxygen and Elevated Carbon Dioxide Atmospheres Inhibit Ethylene Biosynthesis in Preclimacteric and Climacteric Apple Fruit. Journal of the American Society for Horticultural Science, 542-546

Gregori R., Guerra W., Berra L., Bassi G., Sansavini S. (2013) Panel test sensoriale e comparazione di varietà in diversi contesti ambientali. Riv. Frutticoltura 11, 44-52.

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

Janick J. (2005) The origin of fruits, fruit growing, and fruit breeding. Plant Breeding Rev. 25:255-320.

Iannaccone M., Palumbo D., Ventimiglia I., Patocchi A., Spigno P., & Capparelli R. (2007) Use of molecular markers and flow cytometry to preserve ancient Annurca apple germplasm. Biotechnology letters, 29(2): 279-284.

Larsen B., Toldam-Andersen T.B., Pedersen C., Ørgaard M. (2017) Unravelling genetic diversity and cultivar parentage in the Danish apple gene bank collection. Tree Genet Genomes 13:14

Liang W., Dondini L., De Franceschi P., Paris R., Sansavini S., Tartarini S. (2015) Genetic diversity, population structure and construction of a core collection of apple cultivars from Italian germplasm. Plant Mol Biol Rep, 33:458-473

Liebhard R., Gianfranceschi L., Koller B., Ryder CD., Tarchini R., Van de Weg E., Gessler C. (2002) Development and characterisation of 140 new microsatellites in apple (*Malus* × *domestica* Borkh.). Mol Breeding 10:217–241

Maguire T. L., Collins G. C., Sedgley M. (1994) A modified CTAB DNA extraction procedure for plants belonging to the family Proteaceae. Plant Mol Biol Rep, 12:106-109.

Marshall T.C., Slate J., Kruuk L., Pemberton J.M. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology, 7, 639-655.

Melchiade, D., Foroni, I., Corrado, G., Santangelo, I., & Rao, R. (2007) Authentication of the 'Annurca' apple in agro-food chain by amplification of microsatellite loci

. Food Biotechnology, 21(1), 33-43.

Patocchi, A., Fernández-Fernández, F., Evans, K. et al. Development and test of 21 multiplex PCRs composed of SSRs spanning most of the apple genome (2009) Tree Genetics & Genomes 5, 211–223. https://doi.org/10.1007/s11295-008-0176-7

Ross-Ibarra J., Morrell P.L., Gaut B.S. (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. PNAS, 15, 104, 1, 8641-864.

Sansavini S., Alessandri S., Buscaroli C., Gregori R., Dondini L. (2018) Riscoperta e valorizzazione della mela Rosa Romana. Riv. di Frutticoltura, 8:60-64.

Testolin R., Foria S., Baccichet I., Messina R., Danuso F., Losa A., Scarbolo E., Stocco M., Cipriani G. (2019) Genotyping apple *(Malus ×domestica* Borkh.) heirloom germplasm collected and maintained by the Regional Administration of Friuli-Venezia Giulia (Italy). Sci Hortic 252:229–237

Velasco R., Zharkikh A., Affourtit J., Dhingra A., Cestaro A., Kalyanaraman A et al. (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nature Genetics, 42, 10:833-841

Volk G.M., Henk A.D., Richards C.M., Forsline P.L., Chao C.T. (2013) *Malus sieversii*: A diverse Central Asian apple species in the USDA-ARS National Plant Germoplasm System. HortScience, 48(12):1440-1444.

Volk G.M., Richards C.M., Reilley A.A., Henk A.D., Forsline P.L., Aldwinckle H.S. (2005) Ex situ conservation of vegetative propagated species: development of a seed-based core collection for *Malus sieversii*. J. Amer. Soc. Hort. Sci. 130(2): 203-210

Urrestarazu J., Denancé C., Ravon E., Guyader A., Guisnel R., Feugey L., Poncet C., Lateur M., Houben P., Ordidge M., Fernandez-Fernandez F., Evans K.M., Paprstein F., Sedlak J., Nybom H., Garkava Gustavsson L., Miranda C., Gassmann J., Kellerhals M., Suprun I., Pikunova A.V., Krasova N.G., Torutaeva E., Dondini L., Tartarini S., Laurens F., Durel C.E. (2016) Analysis of the genetic diversity and structure across a wide range of germplasm reveals prominent gene flow in apple at the European level. BMC Plant Biol, 16:130

CHAPTER 6: CONCLUSIONS

The biodiversity of the different species, which have played an important role in human history, is in danger of being lost. This research has focused on the characterization of germplasm of two different species: European chestnut (*Castanea sativa*) and the apple ancient genotype 'Rosa Romana'.

The results indicate that the continuous verification of varietal correspondence and the characterization of cultivars, traditionally performed with morphological and biochemical observations, now accompanied by molecular ones, play an increasingly important role in the identification and conservation of germplasm.

The genetic characterization for the recovery of the germplasm is motivated also by the interest for the valorization of local productions in order to make possible their traceability and for the individuation of denomination errors in the collection fields in which they are preserved.

In particular, this research analyzed the genetic diversity of European chestnut cultivars to evaluate their genetic diversity with the aim of enriching collection fields in the Emilia-Romagna region through identified unique varieties. The results confirm that the Italian chestnut germplasm is an important source of genetic biodiversity and contributes to the preservation and enhancement of the entire chestnut genetic heritage. The results of this study could contribute to better understand the human role in the evolution of chestnuts, to expand its genetic knowledge and open up the possibility of making new orchards in Europe. This would be the starting point for future selection programs useful for the revival of chestnut as a fresh product or for the production of flours.

The identification of synonymous accessions, therefore, emphasized the importance of verifying germplasm collections with powerful tools such as microsatellites (SSR). The latter are considered the best markers to explore genetic diversity as they are multi-allelic well distributed in the genome.

The results of this research also show that the revival of local varieties, such as apple 'Rosa Romana' and different local chestnut cultivars that characterize the area of the Tuscan-Emilian Apennines, provide for the preliminary acquisition of information on local germplasm and the cataloguing of accessions through the use of both morphological descriptors and molecular markers.

The protection and maintenance of wild and ancient varieties, such as ‘Rosa Romana’ apple, is of fundamental importance so as to allow an increase in the allelic variability present within the collections of

germplasm and to safeguard the different genotypes in singular environments, thus increasing the diffusion areas of the species.

In conclusion, it is hoped that in the near future we can proceed to further integration of data from collections from different Italian and European regions in order to rationalize, preserve and enhance the entire genetic heritage of these two species.

Moreover, in this way it would be possible to recover the largest number of allelic variants for genes that control the most important agronomic characteristics for the cultivation and for the adaptation of the species to the different environments also in consideration of the climate change observed in recent decades.