UNIVERSITY OF TURIN

DOCTORAL SCHOOL

PhD in
AGRICULTURAL, FOREST AND FOOD SCIENCES

CYCLE: XXXIV

Different approaches in hazelnut and chestnut breeding: genomic and molecular analysis and *in vitro* culture

Vera Pavese

Supervisor:
Prof. Roberto Botta

Ph.D. Coordinator:
Prof. Domenico Bosco

Co-Supervisor:
Prof. Daniela Torello Marinoni

YEARS
2019; 2020; 2021

C’è un solo tipo di successo,
quello di fare della propria vita
ciò che si desidera

Contents

1. General introduction *1*
2. Hazelnut *2*
2.1 Genomic resources and hazelnut breeding 4
3. Chestnut *6*
3.1 Chestnut breeding and diseases *7*
4. New plant breeding techniques *9*
Aim and thesis structure *12*
Chapter I *15*
Abstract *15*
Introduction *17*
Materials and methods *19*
Plant materials and DNA sequencing *19*
De novo genome assembly and reference-guided scaffolding *19*
Genome annotation, integrity and completeness *20*
OrthoFinder *22*
Resistance genes analogs (RGA) *22*
Data availability *23*
Results and discussion *24*
Genome sequencing and assembly *24*
Genome annotation and Orthofinder analysis 28
Resistance genes *33*
Conclusions *38*
References *39*
Supplementary materials *45*
Chapter II *46*
Abstract *46*
Introduction *47*
Results *52*
Genes identification and structure *52*

Phylogenetic and modelling analysis......54
Transcriptional profiling in response to P. cinnamomi infection......57
Transcriptional profiling in response to C. parasitica infection......61
Discussion......64
Materials and Methods......69
Identification of chestnut S-genes orthologues......69
Phylogenetic analysis......70
Protein modeling......70
Pathogen's inoculation and samples collection......71
RNA extraction and Real-Time qPCR quantification......72
Pathogen quantification......74
Supplementary materials......75
Author Contributions......75
Funding......75
Conflicts of Interest......75
References......76
Chapter III......89
Abstract......89
Introduction......91
Materials and Methods......94
Plant material......94
Mining of pds in the C. sativa transcriptome and phylogenetic analysis......94
Vector design......95
Somatic embryo genetic transformation......95
Plant regeneration......97
Molecular analysis of kanamycin-resistant embryos......97
Statistical analysis......98
Results and discussion......99
Gene structure and phylogenetic analysis......99
CRISPR/Cas9 mediated transformation of somatic embryos......102
Somatic embryo proliferation, maturation and germination process......105

Somatic embryogenesis mutation screening 107
Conclusions 112
Data availability statement 113
Author contributions 113
Funding 114
Acknowledgments 114
Supplementary material 114
References 115
Chapter IV 123
Introduction 123
Material and Methods 124
Plant material 124
Protoplast isolation 125
GFP vector design 126
Protoplast transformation using PEG transfection system 126
Transformation efficiency 127
Regeneration 127
Results and Discussion 128
Protoplast isolation 128
Protoplast transformation 129
Regeneration 131
Conclusions 132
Final considerations and future perspectives 133
References 136
Acknowledgments 143
Appendix A 145
Appendix B 146

1. General introduction

Hazelnut and chestnut are two of the most important tree nuts in Italy, especially in the Piedmont Region (Botta *et al.,* 2019). Thanks to the valuable healthy nuts, they are commonly used in the human food diet and commercialized, raw or processed, all over the world.

Nuts are nutrient-rich foods, source of proteins, fibres, polyphenols and phytosterols, unsaturated fatty acids, minerals and vitamins; a diet enriched with the consumption of nuts can decrease cardiometabolic diseases and the level of mortality (Martini *et al.,* 2021; Micek *et al.,* 2021).

In spite of the growing demand for nuts by the confectionery industry (hazelnut and chestnut) and the fresh market (chestnut) growing, there are constrains that hinder the renewal and the new planting of orchards in many areas. In the case of hazelnut, only a small pool of cultivated varieties with excellent traits are available, and this does not provide sufficient diversity for the cultivation in the different environments. In the case of chestnut, the species is still affected by serious biotic adversities that compromise yield and, often, the plant survival.

The *Corylus avellana* species shows the sporophytic self-incompatibility, the high emission of suckers, the later production and the tendency to produce blank nuts. The *Castanea sativa* species shows high susceptibility to pathogens. For both species it is necessary to improve nut yield, quality and resistance to pests and diseases (Botta *et al.,* 2019).

For these reasons, nowadays, there is a growing interest in developing breeding programs to provide improved cultivars that increase yield and

nut quality, with a better adaptability to climate change and tolerance to pathogens and pests.

Woody species have long generation times and, in many cases, show high heterozygosity level due to self-incompatibility systems; in this condition, traditional breeding does not represent an efficient strategy to improve cultivars because it is an intensive, expensive and time-consuming process (Limera *et al.,* 2017).

For this reason, molecular biology, genome sequencing and genetic engineering, offers innovative strategies to improve plant knowledge and confer valuable genetic traits to elite genotipes in order to overcome the challenges of the XXI century: to produce more with less, overcome the risk of food reduction due to climate change, yield in a sustainable manner (Ahmar *et al.,* 2020).

2. Hazelnut

Hazelnut (*Corylus avellana* L.) is a woody species that belongs to the Fagales order, Betulaceae family, Coryloideae subfamily and *Corylus* genus which includes *Corylus avellana* L., *C. maxima* Mill.*, C. chinensis* Franch., *C. sieboldiana* Bl., *C. colurna* L. In Europe, there are two species of *Corylus*: *C. avellana* L., the European hazel, and *C. colurna* L. the Turkish hazelnut (Kasapligil, 1972). The European hazelnut *C. avellana* has a wide geographical distribution: it is widespread from the European continent to the Caucasus mountains (Boccacci and Botta, 2009), and it is absent only in peripheral regions of Europe, in Iceland and some Mediterranean islands (Palmè and Vendramin, 2002).

The average annual world production is 939,927 t (2015–2019) of in-shell hazelnuts, and the total harvested area is 1,000,231 ha (2019). Turkey is the first hazelnut producing country with 606,409 t (means 2015-2019), representing 65% of the world's production. Italy is the second producer with 116,945 t (means 2015-2019) (FAOSTAT 2021).

In Italy, cultivated germplasm accounts some 30 cultivars (Baratta *et al.,* 2016), but only 6 of them are major cultivars still planted in orchards: 'Tonda Gentile delle Langhe', 'Tonda Gentile Romana', 'Tonda di Giffoni', 'Nocchione' (syn. 'Mansa'), 'Mortarella' and 'S. Giovanni'.

The cultivar 'Tonda Gentile delle Langhe' ('TGdL'; syn 'Tonda Gentile'; 'Tonda Gentile Trilobata'), grown in Piedmont, represents an excellence for the Italian economy thanks to the valuable nuts, highly appreciated and used in the food sector. Since December 1993, the productions of 'TGdL' grown in Piedmont can be protected under the Protected Geographical Indication (PGI) designation "Nocciola Piemonte".

Corylus avellana plants are shrubs, 3-7 meters tall, characterized by a variable number of stems generated by the sucker emission.

Hazelnut is an anemophilous, monoecious plant that exhibits a sporophytic self-incompatibility system. The female flowers are grouped in inflorescences and inserted in mixed vegetative buds. The male flowers are organized in cylindrical inflorescences, the catkins, and become visible during the summer period. Anthesis takes place during the winter period from December to March, depending on climate conditions.

The fruit is an indehiscent nut surrounded by an herbaceous involucre (Botta *et al.,* 2019).

2.1 Genomic resources and hazelnut breeding

Plant Breeding is a discipline focused on obtaining new valuable cultivars, increasing resistance to pathogens and adaptability to climate change (Botta *et al.,* 2019).

Hazelnut breeding started in the 1960s in Italy, France and USA and in 1980s in China and Turkey. Breeding programs were focused on increasing yield and nut quality and resistance to diseases. These programs use the traditional technique of controlled crossing between two selected cultivars chosen for their valuable traits. In Italy the University of Torino released 4 hybrids (Daria, UNITO 101, UNITO 119 and UNITO 3L) originated by the crosses between 'Cosford' X 'Tonda Gentile delle Langhe' and one hybrid (UNITO G1) from 'Payrone' X 'Tonda Romana' (Valentini and Me, 1999).

UNITO L35 ('Tonda Gentile delle Langhe' X 'Lansing') (Valentini *et al.*, 2001a) was released for the in-shell market. In addition, four clones (UNITO-AD17, UNITO-MT4, UNITO-MT5, and UNITO-PD6) of 'Tonda Gentile delle Langhe' were selected to be used in new orchards. The University of Tuscia (Viterbo) developed 'Madonnella' and 'Romanella', two selections with high quality kernel (Tombesi *et al.,* 2017), while the University of Perugia released two cultivars originated from 'Tonda Romana' X 'Tonda di Giffoni' crosses, four selections from open pollination of 'Tonda Romana' or 'Tonda di Giffoni' (Volumnia I, II, III, IV; Tombesi *et al.,* 2017).

Since the conventional breeding of woody species is limited due to the long generation times (Limera *et al.,* 2017), in recent years, thanks to the introduction of bioinformatics techniques, capable of relating molecular

data and field observations, breeding has achieved many advances. Despite this, there is a large gap between the resources available for the other woody fruit species and the current information on hazelnut.

Hazelnut is a diploid species with eleven pairs of homologous chromosomes (2n = 2x = 22). Its genome size is estimated to be around 378 Mbp (https://www.cavellanagenomeportal.com).

The genomic resources available for hazelnut consist of two *de novo* assemblies: the cultivar 'Jefferson' genome (Rowley *et al.,* 2018) and the chromosome-scale genome of the Turkish ‘Tombul’ cultivar (Lucas *et al.,* 2021).

Rowley and co-authors (2012; 2018) developed transcriptomes starting from four different 'Jefferson' plant materials (leaves, catkins, bark and whole young seedlings) and the re-sequencing of seven European cultivars. Recently, the *C. heterophylla* Fisch. (Chen *et al.,* 2014) and *C. mandshurica* Maxim. (Ma *et al.*, 2013) transcriptome sequences were made available.

A high-density genetic map (Mehlenbacher *et al.,* 2006), was constructed using Random Amplified Polymorphic DNA (RAPD) and Simple Sequence Repeats (SSR) markers, and subsequently optimized using further SSR markers (Gürcan and Mehlenbacher, 2010; Gürcan *et al.,* 2010; Colburn *et al.,* 2017; Bhattarai and Mehlenbacher, 2017). Two other genetic maps available are those developed by Beltramo *et al.* (2016) and Ozturk *et al.* (2017): Quantitative Trait Loci (QTL) associated with phenotypic traits were identified in these maps. Torello Marinoni *et al.* (2018) constructed a genetic map based on SNP marker based using a progeny ‘TGdL’ X ‘Merveille de Bollwiller’ and revealed the QTL regions associated with the time of leaf budburst. Finally Valentini *et al.,* 2021, based on

the same genetic map, revealed the QTL regions associated with flowering time, dichogamy and nut maturity time.

3. Chestnut

The *Castanea* genus belongs to the Fagaceae family that comprises six genera: *Castanea*, *Castanopsis*, *Fagus*, *Lithocarpus*, *Nothofagus* and *Quercus*. It is widespread in the boreal hemisphere and includes 12 or 13 species (depending on the classification), among which the most economically important are the European chestnut *Castanea sativa* (Miller), the Japanese chestnut *Castanea crenata* (Siebold and Zuccarini), the Chinese chestnut *Castanea mollissima* (Blume) and the American chestnut Castanea *dentata* (Bork) (Beccaro *et al.,* 2020).

The European chestnut (*Castanea sativa* Mill.) is a multipurpose tree, appreciated worldwide for timber and nut production (Fernandes *et al.,* 2020). The European chestnut offers a wide range of secondary products and ecosystem services and is recognised wordwide for the excellent nuts quality.

Chestnut wood, thanks to the high tannin content, is highly appreciated for outdoor uses. Chestnuts have been widely used in the human diet and consumed in different ways: roasted, candied, boiled, dried, or transformed into flour (Conedera *et al.,* 2016).

The European chestnut is a vigorous 30-35 m height tree that can exceed 400 years of age. The leaves are deciduous, oblong-lanceolate with a crenate margin and with a lighter green abaxial leaf surface.

It is a species that prefers temperate climates but also tolerates intense cold winters.

Chestnut is a monoecious species with staminate and pistillate flowers arranged in male and androgynous catkins. The male flowers are disposed in a spiral along the catkin axis, while the female flowers are grouped in globose inflorescences at the base of bisexual catkins (Beccaro *et al.,* 2020). Sprouting occurs in late March, vegetative activity lasts until November, while anthesis takes place from mid-June to mid-July, depending on cultivar type and environmental conditions (Freitas *et al.,* 2021).

Fruits are nuts protected by the burr, a spiny involucre that opens at maturity.

The Italian cultivated germplasm accounts for over 300 cultivars (Barrel *et al.,* 2016), many of which endangered. The cultivar 'Marrone' is appreciated worldwide for its fine taste and is found from North to Central Italy. In Piedmont, the cultivar 'Garrone Rosso' is grown in the Cuneo province where it represents most of the production. It is appreciated for external traits and the excellent organoleptic qualities; it is mainly used for fresh consumption and in the confectionery industry.

3.1 Chestnut breeding and diseases

Chestnut breeding has focused on obtaining cultivars with higher yield and nut and timber quality (large nut size, high peeling and good wood quality, fast growth), increasing the resistance to biotic and abiotic stresses. Another breeding objective has been the selection of rootstocks compatible with *C. sativa* (Beccaro *et al.,* 2019) and tolerant to root pathogens.

One of the main chestnut coltural problem is the high susceptibility to two severe diseases that threaten its survival: i) ink disease caused by the

oomycete *Phytophthora* spp. and ii) chestnut blight caused by the fungus *Cryphonectria parasitica*.

In addition, chestnut is affected by the Asian gall wasp *Dryocosmus kuriphilus* Yasumatsu (Hymenoptera Cynipidae), an invasive insect that causes gall formation (Torello Marinoni *et al.,* 2020), and by the nut rot and canker agent *Gnomoniopsis castaneae* G. Tamietti (Lione *et al.,* 2020).

i) Ink disease is caused by soil organisms belonging to *Phytophthora* genus with includes *P. cambivora* and *P. cinnamomi.* The pathogen spread is favoured by movements of agricultural vehicles, cars, people and water.

Infection occurs through wounds localized at the plant collar or at root level (Dilzahan *et al.,* 2021) and the typical symptoms are flame-shaped necrosis at the base of the trunk with the appearance of ink-blue exudate. Other symptoms are foliage yellowing and microfilias, branch dieback, until the death of the plant and the complete loss of suckering capability (Jung *et al.,* 2018).

Among chestnut species, *C. mollissima* and *C. crenata* are tolerant to *Phytophthora* spp (Fernandes *et al.,* 2021) and have been used to obtain hybrids bearing the trait. Breeding programs carried out at INRA, crossing *C. sativa* and *C. crenata,* produced a set of hybrid selections with a higher tolerance to *Phytophthora* that are currently used as rootstocks and as direct producers, despite the lower quality of nuts (Serrazina *et al.,* 2015; Santos *et al.,* 2015).

ii) The second disease, chestnut blight, is caused by the fungus *Cryphonectria parasitica*. All the organs of the plant, except roots, are susceptible to the disease. The fungus causes tissue necrosis and the appearance of red coloured spots.

The necrotic areas, subjected to strong tensions, cause the appearance of the cankers characterized by the presence of yellow-orange fungus fruiting papules (Lione *et al.,* 2020). In addition to active cankers, more superficial and healed cankers caused by hypovirulent strains of *C. parasitica* were detected in chestnut. In this case, the chestnut plant reacts by healing the tissues and producing new healthy tissues (Muñoz-Adalia *et al.,* 2021). The hypovirulent strains are used in the biological control of virulent *C. parasitica* by inoculation close to active cankers (Milgroom and Cortesi, 2004; Rigling and Prospero, 2018). An alternative to biological control can be the remotion of active cankers, taking care to remove the infected material to prevent a subsequent infection. At the moment, there are not breeding programs aimed at obtaining resistance or increasing tolerance to canker blight in the European chestnut. On the contrary, a large breeding program has been carried out in the USA to restore the species *C. dentata* by introgressing resistance genes from *C. mollissima*.

4. *New plant breeding techniques*

In 2050, the world population is expected to increase to over 9 billions people and the harvest needed will be 60% higher than today's production. For this reason, new strategies are necessary to achieve plant resilience to climate change, higher yields and nutritional quality (Tilman *et al.,* 2011). The traditional breeding, as previously described, doesn't represent a valid strategy for the development of new improved woody fruit and nut species. For this reason, New Plant Breeding Techniques (NPBTs) can be a powerful tool to improve plant breeding in a short time and economic way (Osakabe *et al.,* 2018).

Currently, the CRISPR/Cas9 (*Clustered Regularly Interspaced Short Palindromic Repeats-Cas9*) technique is considered one of the most effective low-cost tool for plant genetic engineering among NPBTs.

The CRISPR/Cas9 technology system is based on a guide RNA sequence (gRNA) complementary designed to a target genome site. The gRNA identifies the target gene and then the Cas9 nuclease provokes a DNA double-strand break, promoting the insertion, deletion, or single nucleotide modifications through two mechanisms, non-homologous end joining (NHEJ) and homology-directed recombination (HDR).

Genetic engineering using the CRISPR/Cas9 editing system has been efficiently adopted in several plant species, such as *Arabidopsis*, tobacco, rice, wheat, maize, soybean, tomato, poplar and citrus (Osakabe *et al.,* 2018). There is a growing interest in applying CRISPR/Cas9 system on woody plants in order to generate rapidly ideal cultivars deprived of negative genetic undesired traits (Scintilla *et al.,* 2021).

In woody plants, genetic engineering is still limited due to regeneration recalcitrance and low transformation efficiency. Currently, there is a growing amount of work aimed at better understanding these biological processes and detecting genes able to promote *in vitro* regeneration.

CRISPR/Cas9 complex is usually delivered using *Agrobacterium tumefaciens* strain or through particle bombardment; the complex can be thus integrated into the plant genome with the drawback that both the Cas9 enzyme and the gRNA can remain active for a long time causing off-target events.

Since GMOs in Europe are subjected to strict government restrictions (Woo *et al.,* 2015) and the public opinion throws doubts on the safety of the GMO products (Chen *et al.,* 2019), researchers are searching for new

OGM-free strategies, without transgene integration. To avoid this event, the CRISPR/Cas9 construct can be introduced as ribonucleoprotein (RNP), a form that avoids the introgression of exogenous DNA (GMO-free).
The RNP-based system is efficient because immediately acts on the target site without requiring the activation of the transcription process, being the RNP assembled *in vitro* and delivered directly into the cell; then the ribonucleoprotein complex is rapidly degraded (Chen *et al.,* 2019) thanks to the natural cellular mechanisms of protein and RNA turnover. The regeneration from a single cell allows to maintain genetic uniformity.
The major limit of this technology is that woody species are characterized by low editing efficiencies and low regeneration rates (Corredoira *et al.,* 2019).
The development of a gene editing protocol using CRISPR/Cas9 precise editing and in addition the RNP delivery can be a useful tool to improve and accelerate chestnut breeding of elite cultivars.

Aim and thesis structure

Since the chestnut and hazelnut industries are two very important crops for Italy and for the Piedmont region, the purpose of the thesis is to increase the genetic knowledge on these species and make this knowledge available for developing strategies of targeted breeding programs. The thesis is divided into 4 chapters, the first one focused on hazelnut (Chapter I) and the following 3 chapters focused on chestnut (Chapter II, Chapter III and Chapter IV).

Chapter I "Whole-genome assembly of *Corylus avellana* cv 'Tonda Gentile delle Langhe' using linked-reads (10X Genomics)", presents the genome sequencing and assembly of the *C. avellana* cultivar 'Tonda Gentile delle Langhe', which represents an excellence for the Italian economy and is highly used in the food sector.
The availability of a chromosome scale genome sequence of this cultivar would allow to increase the genetic knowledge of the species. The identification of genes involved in several biological processes, such as sporophytic incompatibility and host-pathogen interaction mechanisms, could be useful for genome editing programs.
Moreover, since hazelnut shows an *in vitro* recalcitrant response (Contessa *et al.,* 2011), detecting genes related to the regeneration process can be of interest to increase *in vitro* plant production.

The availability of a complete genome would also be a powerful tool for SNP markers detection, to be used in cultivar identification programs in hazelnut-based food products, to prevent or reveal commercial fraud.
A paper has already been published (https://doi.org/10.1093/g3journal/jkab152).

Chapter II "Identification of susceptibility genes in *Castanea sativa* and their transcription dynamics following pathogen infection" presents the first example of identification and characterization of susceptibility genes involved in plant-pathogen interaction in chestnut. It was demonstrated in several species that the knock out of these genes reduces the ability of the pathogen to infect the host and induces plant tolerance. In this chapter the involvement of susceptibility genes is demonstrated in *Castanea sativa* and *Castanea crenata* response to *Cryphonectria parasitica* and *Phytophthora cinnamomi* infections. The selected candidate genes, involved in the infection process, can be tested in future targeted genome editing programs.
A paper has already been published
(https://doi.org/10.3390/plants10050913).

Chapter III "First Report of CRISPR/Cas9 Gene Editing in *Castanea sativa* Mill" presents the results of the first experiment of CRISPR-Cas9 genetic transformation in *C. sativa*, since no evidence of CRISPR-Cas9 genetic transformation in *C. sativa* is present in literature so far. The target chosen for this editing experiment was the *phytoene desaturase* (*pds*) gene

involved in chlorophyll biosynthesis; its mutation causes an albino phenotype. For this reason, *pds* is usually used as a visual marker to test the effectiveness of a new technique.
A paper has already been published (https://doi.org/10.3389/fpls.2021.728516).

Chapter IV "First protoplasts isolation and transformation protocol in *C. sativa* Mill." presents the first protoplasts isolation protocol and set up starting from zygotic embryos ever reported in *C. sativa*. Protoplasts were afterward transformed using the Green Fluorescence Protein (GFP) reporter gene to test the transformation success.

Chapter I

G3, 2021, 11(7), jkab152
DOI: 10.1093/g3journal/jkab152
Advance Access Publication Date: 8 May 2021
Genome Report

Whole-genome assembly of *Corylus avellana* cv 'Tonda Gentile delle Langhe' using linked-reads (10X Genomics)

Vera Pavese[1]*, Emile Cavalet Giorsa[1]*, Lorenzo Barchi[1], Alberto Acquadro[1]**, Daniela Torello Marinoni[1], Ezio Portis[1], Stuart James Lucas[2], Roberto Botta[1]

1 Dipartimento di Scienze Agrarie, Forestali e Alimentari—DISAFA, Università degli Studi di Torino, Largo Paolo Braccini 2, Grugliasco, 10095 Torino, Italy
2 Sabanci University SUNUM Nanotechnology Research and Application Centre, Istanbul, Turkey
* These authors contributed equally to this work.
** Correspondence: alberto.acquadro@unito.it

Abstract

The European hazelnut (*Corylus avellana* L.; 2n=2x=22) is a worldwide economically important tree nut that is cross-pollinated due to sporophytic incompatibility. Therefore, any individual plant is highly heterozygous. Cultivars are clonally propagated using mound layering, rooted suckers

and micropropagation. In recent years, the interest in this crop has increased, due to a growing demand related to the recognized health benefits of nut consumption. *C. avellana* cv 'Tonda Gentile delle Langhe' ('TGdL') is well-known for its high kernel quality, and the premium price paid for this cultivar is an economic benefit for producers in northern Italy. Assembly of a high-quality genome is a difficult task in many plant species because of the high level of heterozygosity. We assembled a chromosome-level genome sequence of 'TGdL' with a two-step approach. First, 10X Genomics Chromium Technology was used to create a high-quality sequence, which was then assembled into scaffolds with cv 'Tombul' genome as the reference.

Eleven pseudomolecules were obtained, corresponding to 11 chromosomes. A total of 11,046 scaffolds remained unplaced, representing 11% of the genome (46,504,161 bp). Gene prediction, performed with Maker-P software, identified 27,791 genes (AED ≤ 0.4 and 92% of BUSCO completeness), whose function was analysed with BlastP and InterProScan software. To characterise 'TGdL' specific genetic mechanisms, Orthofinder was used to detect orthologs between hazelnut and closely related species. The 'TGdL' genome sequence is expected to be a powerful tool to understand hazelnut genetics and allow the detection of markers/genes for important traits to be used in targeted breeding programs.

Keywords: genomics, NGS, 10X genomics, hazelnut

Introduction

The European hazelnut (*Corylus avellana* L.) is a woody species belonging to the *Betulaceae* family. It is an economically important tree nut whose production is mostly destined to the confectionery industry with a demand that has rapidly increased (Molnar, 2011). As a consequence, hazelnut harvested areas showed a 58% increase in 2019 (1,000,231 ha) compared to 2014 (FAOSTAT, 2019). Hazelnut is cultivated in many countries, including Turkey (65% World production), Italy (12.5%), Azerbaijan (4.6%), USA (3.9%), Chile, China, and Georgia (Botta *et al.,* 2019). In the Piedmont Region of Italy, hazelnut production is mainly based on the cultivar 'Tonda Gentile delle Langhe' (syn. 'Tonda Gentile', 'Tonda Gentile Trilobata', hereafter 'TGdL'), a small-sized and trilobate shaped kernels command a premium price due to their high quality, especially after roasting (Valentini *et al.,* 2014). In December 1993, the European Union recognized the Protected Geographical Indication (PGI) "Nocciola Piemonte" to 'TGdL' produced in the piedmont areas of northern Italy (https://eur-lex.europa.eu). 'TGdL' is considered to have a monoclonal origin and it is clonally propagated by mound layering, rooted suckers and micropropagation (Valentini *et al.,* 2014).

High-quality genome assembly in many fruit tree species is a difficult task due to their high heterozygosity and thus haploid or doubled haploid plants have been often used to accomplish this goal (Jaillo *et al.,* 2007) or genomes have been assembled with specialized algorithms (e.g. Platanus, Kajitani *et al.,* 2014). Recently, long-read sequencing technologies, such as single-molecule real-time sequencing (SMRT, Pacific Biosciences) or

nanopore sequencing (Oxford Nanopore Technologies) have been adopted to face this task. Moreover, scaffolding-like technologies such as optical mapping (Bionano Genomics, Barchi *et al.,* 2019), proximity ligation methods (Hi-C, Dovetail Genomics, Acquadro *et al.,* 2020b) and linked-reads (10X Genomics, Hulse-Kemp, 2018) are generally used as companion strategies. The latter, a low-cost approach, can also be used to improve assembly metrics and to reconstruct long-range haplotypes. The 10X Linked-Reads technique amplifies the potential of short-read sequencing to achieve a much more complete genomic analysis. Using this technology, it is possible to discriminate the two haplotypes and also to analyze regions with high repetitiveness.

European hazelnut is a diploid species with 11 chromosomes (2n=2x=22), with an estimated genome content (1C) of 0.43 pg (Pustahija *et al,* 2013). Being an outbred species (Beltramo *et al.,* 2016) hazelnut shows a high level of heterozygosity. In 2009, a draft genome of the European hazelnut cultivar 'Jefferson' was released (https://www.cavellanagenomeportal.com/), while a chromosome-scale assembly of the Turkish cv. 'Tombul' was recently reported (Lucas *et al.,* 2020). Rowley *et al.* (2012, 2018) studied the cv 'Jefferson' transcriptome (4 tissues), and re-sequenced seven European cultivars (~20x coverage). More recently, transcriptome sequences were obtained for *C. heterophylla* Fisch. (Chen *et al.,* 2014) and *C. mandshurica* Maxim. (Ma *et al.,* 2013). Mehlenbacher *et al.* (2006) constructed a genetic linkage map of *C. avellana* based on RAPD and SSR markers, while Torello Marinoni *et al.* (2018) developed SNP-based genetic maps for 'TGdL' x 'Merveille de Bollwiller' and detected QTL regions associated with time

of leaf budburst. Many studies in addition to Öztürk *et al.* (2018) have used SSR markers to study diversity sequencing/assembly.

There is a large gap between the tools available for other fruit species and the existing knowledge on hazelnut. This work aims to fill this gap, considering that hazelnut is a strategic crop for Italy. The European hazelnut genome sequencing will allow the study of the *Corylus* pan-genome, the identification of variants for traceability or the implementation of genome wide association studies.

For this reason, here we report the chromosome-scale assembly of the European hazelnut cultivar 'TGdL' established through a two-tiered approach: i) 10X Genomics sequencing/assembling and ii) scaffolding using the RaGOO pipeline (Alonge *et al.*, 2019) and the 'Tombul' genome as a guide.

Materials and methods

Plant materials and DNA sequencing

Young fresh leaves were collected from the 'TGdL' UNITO-AD17 clone. DNA extraction was performed by Novogene (Genome Sequencing Company, Hong Kong) and used to construct 10X Genomics Chromium technology (Weisenfeld *et al.,* 2017) libraries. Sequencing was then performed on an Illumina NovaSeq 6000 System.

De novo genome assembly and reference-guided scaffolding

The 'TGdL' genome was *de novo* assembled using Supernova Assembler v 2.1.1 (Weisenfeld *et al.,* 2017) software (10X Genomics) using 10X

linked-reads as input. The Supernova Assembler was run directly on raw data derived from the sequencing process without any read cleaning process. The output format chosen for the subsequent analyses was “pseudohap”. The gap-closing process was performed using GapCloser script from SOAPdenovo2 pipeline (Luo *et al.*, 2012).
The reference-guided scaffolding was performed using the RaGOO v1.1 (Alonge *et al.*, 2019) scaffolder with the ‘Tombul’ genome as reference (PRJEB31933, https://www.ebi.ac.uk/ena/), with default parameters. The gap-closing process was repeated to further decrease the rate of indeterminate bases (N). Quality assessment of the genome assemblies was obtained using the QUAST tool (http://quast.sourceforge.net/). SNP/Indels were counted and analyzed using custom bash scripts. The estimation of the genome heterozygosity level was calculated by considering the ratio between the number of SNP/Indels (called in heterozygous state) and the size of the assembled genome after removal of Ns (404.097.498 bp) as previously reported (Acquadro *et al.*, 2020a).

Genome annotation, integrity and completeness

The *de novo* assembly was masked using RepeatMasker (Smit *et al.*, 2013–2015) and the gene prediction used Maker-P (Campbell *et al.*, 2014). The prediction process was made using Augustus (Stanke *et al.*, 2006) Hidden Markov Models and SNAP (Bromberg and Rost, 2007) algorithms aided by a set of NCBI available hazelnut proteins and transcripts. All the genes detected were evaluated considering AED values and only genes with AED ≤ 0.4 were maintained. The AED values measure the concordance between the predicted gene and a transcript, mRNA-seq and protein

homology library data. In a case of perfect concordance, the score is 0, in the opposite case 1. To measure the quality and completeness of the predicted proteomes, a quantitative assessment was carried out based on evolutionary informed expectations of gene content known as Benchmarking Universal Single-Copy Orthologs (BUSCO89 v3.0.2., Embryophyta odb 10).

Gene function was attributed using BLASTP (Altschul *et al.,* 1990) comparing data with the Uniprot/Swissprot Viridiplantae database. Default parameters, except for the e-value (<1e-5) were applied. InterProScan (v. 5.33-72.0; Jones *et al.*, 2014) was also introduced for domain inspection using all the available databases (ProSitePro les-20.119, PANTHER-10.0, Coils-2.2.1, PIRSF-3.01, Hamap-201511.02, Pfam29.0, ProSitePatterns - 20.119, SUPERFAMILY-1.75, ProDom-2006.1, SMART-7.1, Gene3D-3.5.0 and TIGRFAM-15.0).

OrthoFinder

OrthoFinder software was used for the detection of putative orthologs and orthology groups. The comparisons were made among three *C. avellana* cultivars ('TGdL', 'Jefferson', 'Tombul'), *Quercus suber*, *Betula pendula* and *Carpinus fangiana*. Gene ontology (GO) term enrichment was carried out with AGRIGOv2 (http://systemsbiology.cau.edu.cn/agriGOv2/) to find a representative subset of the GO terms previously identified with the Interproscan pipeline.

Resistance genes analogs (RGA)

Candidate resistance genes were identified using RGAugury (Li *et al.*, 2016). RGA candidates were classified into four major families based on the presence of combinations of these RGA domains and motifs: NBS-encoding (subsequently divided in subgroups according to their domain architecture, namely NBS (NBS domain), CNL (CC-NBS-LRR domains), TNL (TIR-NBS-LRR), TN (TIR-NBS), CN (CC-NBS), NL (NBS-LRR), TX (TIR-unknown domain and other), TM-CC, and membrane associated RLP and RLK. MAFFT v7.450 (https://mafft.cbrc.jp/) was used for protein alignment with the following parameters: –ep 0 –reorder –maxiterate 1000 –genafpair. Genetic relationships were described by constructing a phylogenetic tree by maximum likelihood by using the IQ-TREE software (v.1.6.12, http://www.iqtree.org/); branch supports were obtained with the ultrafast bootstrap with 1000 replicates. Trees were visualized using interactive Tree of Life (iTOL v3, https://itol.embl.de/).

Data availability

Raw reads are publicly available in the NCBI sequence read archive under the bioproject: PRJNA694440. The reference assembly and annotation data are also available for downloading from https://zenodo.org/deposit/4454484/. Supplementary material is available at figshare: https://doi.org/10.25387/g3.14502048/.

Results and discussion

Genome sequencing and assembly

The chromosome-scale ‘TGdL’ genome was developed using two-tiered approach. The 10X Genomics Chromium Technology was firstly used to obtain a high-quality preliminary assembly. Reference-guided scaffolding was then implemented using the ‘Tombul’ genome as the reference. The 10X genomic library was sequenced with Illumina technology and 138.56 million raw pair-end reads were generated (52X coverage). The average read length was 138.50 bp, with 86.06% of them having Q>30. These data are comparable to the optimal standard values suggested by Supernova Assembler software manufacturer (Table 1). In details, Supernova assembled 47,216 scaffolds having a total length of 414.38 Mb and an N_{50} of 51,567 bp. The results were similar (Table 2) to the other genome assemblies at a contig level (‘Jefferson’, Rowley *et al.,* 2018; ‘Tombul’, Lucas *et al.,* 2020; Table 2). A single library (10X Genomics Chromium Technology) produces a more optimized ‘TGdL’ genome assembly compared to the ‘Jefferson’ assembly (Table 2), the latter being obtained using three different Illumina libraries, 250-bp and 350-bp Illumina paired-end (PE) libraries and a 4.5-Kb mate-pair (MP) library. Moreover, the 10X strategy proved to be a cost-effective route being the ‘TGdL’ assembly highly comparable to the ‘Tombul’ (contigs) draft made with a higher coverage (108X, short reads) and 9.3X of long Nanopore reads (Figure 1) prior to scaffolding.

The reference guided scaffolding was made using the RaGOO pipeline, which was able to optimize the ‘TGdL’ assembly (contigs) adopting the ‘Tombul’ genome (PRJEB31933), previously obtained with proximity

ligation technology (Dovetail Genomics using Chicago & HiC protocols), as reference. It produced 11 pseudomolecules (11 chromosomes) and 11,046 scaffolds belonging to chromosome 0, which represent 11% of the genome (46,504,161 bp). The resulting assembly was a complete 'TGdL' chromosome-scale genome (Table 2), whose total length (without chr0) resembled that on the 'Tombul' assembly. Following the scaffolding process, we renamed the super scaffolds based on Torello Marinoni *et al.* (2018) linkage groups (Table 3).

The rate of heterozygosity of the hazelnut genome was calculated as 0.84% and was similar for all the pseudomolecules, ranging from 0.91% (chromosome 6) to 0.82% (chromosome 7). Unplaced scaffolds showed a lower rate of heterozygosity (0.60%). These data were expected due to the allogamous behavior of hazelnut and are comparable to those of other outbred species (~1%; Acquadro *et al.,* 2017, Velasco *et al.,* 2007), and higher than the ones obtained in inbred species (~0.1%; Barchi *et al.,* 2019, Acquadro *et al.,* 2020a).

Table 1: Summary of the metrics of the 'TGdL' 10X Genomics Chromium Technology

	'TGdL' Results	**Optimal standard values**
Reads number	138.56 M	-
Reads average length	138.50 bp	140 bp
Coverage	52X	56X
% reads with Q30 quality	86.06	75-85%

Table 2: Metrics of the genome assemblies of the 'TGdL', 'Tombul' and 'Jefferson' cultivar.

	'TGdL' (contigs)	'Tombul' (contigs)	'Jefferson'	'TGdL'	'Tombul'
Scaffold number (#)	47,216	12,557	36,641	11,059	2,207
Total Length (Mb)	414.38	383.1	345.54	419.46	384.2
N_{50}	51,567	1,299	21,508	35,598,485	36,653,616
L_{50}	1,457	78,800	4,253	5	5
Largest contig (bp)	1,152,936	-	274,525	53,036,447	50,950,907
# N's/ 100 Kb	3,811.59	180	12,054.75	3,514.85	468.09
GC (%)	36.87	-	36.3	36.84	35.91

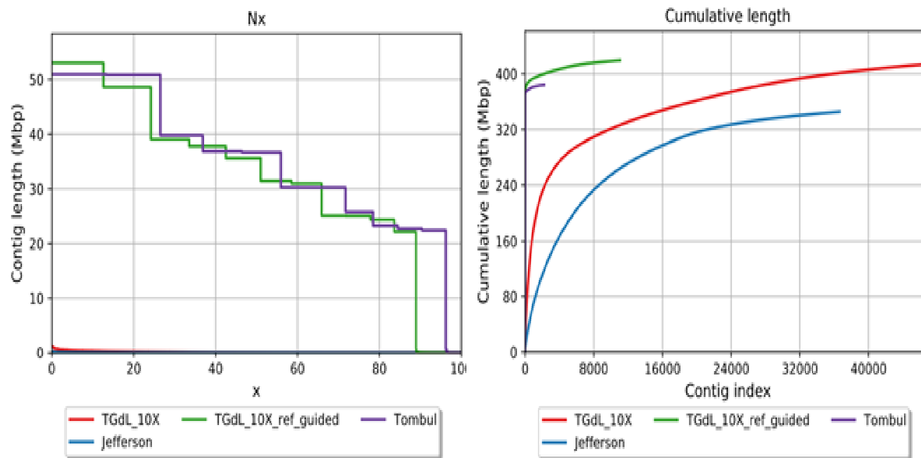

Figure 1: Contiguity statistics performed on 'TGdL' (contigs), 'TGdL' (pseudomolecules plus unplaced scaffolds), 'Tombul' (pseudomolecules) and 'Jefferson' genomes. Left picture: Nx statistics (Nx is the largest contig length, L, such that using contigs of length ≥ L accounts for at least x% of the bases of the assembly) with x varying between 1 and 100. Right picture represents the cumulative length increment of the genome through the scaffold/contig addition.

Table 3: Pseudomolecules reconstructed in 'TGdL' and nomenclature according to the genetic map by Torello Marinoni *et al.,* 2018. The nomenclature of 'Tombul' pseudomolecules is reported as in Lucas *et al* (2020). Observed SNP/indel in heterozygous state and their frequency are calculated using the size of the assembled genome, after removal of Ns (404,097,498 bp).

'TGdL' pseudomolecules	**'Tombul' Nomencla-ture**	**size (bp)**	**Ns**	**SNP/ indels**	**SNP/ indels frequency**
1	1	53,036,447	2,194,688	430,039	0.85%
2	2	48,611,531	1,656,681	390,798	0.83%
3	8	25,054,957	974,784	216,892	0.90%
4	3	39,027,746	1,377,931	322,996	0.86%
5	5	35,598,485	1,323,708	299,132	0.87%
6	10	24,339,126	955,975	211,629	0.91%
7	6	30,979,729	1,185,588	244,134	0.82%
8	11	22,131,124	858,998	182,492	0.86%
9	4	37,785,217	1,350,719	299,648	0.82%
10	9	25,093,934	1,084,840	201,790	0.84%
11	7	31,441,515	1,359,091	262,222	0.87%
unplaced scaffolds	-	45,736,493	415,803	272,272	0.60%
Whole genome	-	418,836,304	14,738,806	3,334,091	

Genome annotation and Orthofinder analysis

Globally, ~41.5% of the assembled genome were repeated, 17% of which consisted of LTR elements (Table 4). The assembled genome was then structurally annotated with the Maker-P suite and the total number of genes identified was 27,791 (AED 0.4). The proteome was validated using BUSCO; overall, more than 92% of 1,614 expected embryophyta genes

were identified in the 'TGdL' genome annotations as the complete and partial BUSCO profiles. The number of predicted genes is similar to the one predicted in 'Tombul' (27,270) and in 'Jefferson' (28,167); a similar number of genes were also identified in the close species *Carpinus fangiana* (27,384) and *Betula pendula* (28,153), while fewer genes were predicted in *Quercus suber* (25,808).

The function attributed to each predicted protein was based on the results of the BLASTP (SwissProt) and the InterProScan domain inspection. InterProScan analyses highlighted about 80% of the predicted proteins with at least one IPR domain. Among the top 20 SUPERFAMILY domains (Table 5), the most abundant in all the genomes was SSF56112 (protein Kinase-like domain), which acts on signalling and regulatory processes in the eukaryotic cell. The other most abundant Superfamilies were: SSF52540 (P-loop containing nucleoside triphosphate hydrolase), which is involved in several UniPathways, including chlorophyll or coenzyme A biosynthesis and SSF52058 (Leucine-rich repeat domain, L domain-like), which is related to resistance to pathogens.

Clustering by Orthofinder the proteomes (164,573 sequences) of the three hazelnut genomes together with the ones from *Betula pendula*, *Carpinus fangiana* and *Quercus suber*, produced a set of 21,239 gene families (plus 24,639 unassigned genes), of which 5,892 (including 59,597 genes) were shared (Figure 2). Focusing on hazelnut, the 'Jefferson' proteome showed the highest percentage of unassigned genes (41.6%), presumably due to the fragmented assembly which limited the annotation procedure. On the other hand, the 'TGdL' and 'Tombul' assemblies showed a high percentage of assigned genes to orthogroups (93.1% and 96.3%

respectively). The ‘TGdL’ proteome contained 388 genome-specific orthogroups (1,279 genes), while 732 (with 2,040 genes) were shared between ‘TGdL’ and ‘Tombul’, but not the other genomes. For the former, the analysis revealed significant gene enrichment for some GO terms (Table S1), including GO:0042908 (xenobiotic transport) as well as GO terms related to nuclease activity (GO:0016891 (endoribonuclease activity, producing 5'-phosphomonoesters), GO:0004540 (ribonuclease activity)) and transport (GO:0008559 (xenobiotic-transporting ATPase activity) and GO:0090484 (drug transporter activity)). For genes shared by ‘TGdL’ and ‘Tombul’, enriched GO terms included GO:0044092 (negative regulation of molecular function), GO:0043086 (negative regulation of catalytic activity) and GO:0050790 (regulation of catalytic activity), as well as several nuclease related terms (as GO:0004523 (RNA-DNA hybrid ribonuclease activity) GO:0004521 (endoribonuclease activity) and GO:0004540 (ribonuclease activity) (Table S2).

Table 4: Masking statistics for the 'TGdL' hazelnut genome.

Class	Superfamily	Count	Masked (bp)	Masked (%)
DNA	--	--	--	
	hAT	46,506	9,536,392	2.30%
	CACTA	66,513	10,364,758	2.50%
	PIF/Harbinger	33,649	5,713,915	1.38%
	Mutator	208,743	34,956,690	8.43%
	Tcl/Mariner	9,668	1,607,102	0.39%
	Helitron	71,851	14,190,408	3.42%
LTR	--	--	--	
	Copia	46,495	19,214,573	4.63%
	Gypsy	48,852	26,434,445	6.37%
	unknown	95,170	26,065,393	6.28%
MITE	--	--	--	
	hAT	9,724	1,295,113	0.31%
	CACTA	1,080	128,610	0.03%
	PIF/Harbinger	10,503	1,717,093	0.41%
	Mutator	60,562	6,926,393	1.67%
	Tcl/Mariner	388	33,917	0.01%
Unspecified		22,038	4,627,197	1.12%
	total interspersed	**731,742**	**162,811,999**	**39.24%**
Low_complexity		30,853	1,453,286	0.35%
Simple_repeat		224,449	7,743,022	1.87%
Total		**987,044**	**172,008,307**	**41.46%**

Table 5: Top 20 SUPERFAMILY domains in the 'TGdL' hazelnut genome

Superfamily	Description	Count
SSF56112	Protein kinase-like domain	1577
SSF52540	P-loop containing nucleoside triphosphate hydrolase	1565
SSF52058	L domain-like	935
SSF51735	NAD(P)-binding domain	484
SSF57850	RING/U-box	468
SSF48371	Armadillo-type fold	428
SSF48452	Tetratricopeptide-like helical domain	413
SSF57889	Cysteine-rich domain	409
SSF52047	RNI-like	380
SSF48264	Cytochrome P450	376
SSF53474	Alpha/Beta hydrolase fold	362
SSF53756	UDP-Glycosyltransferase/glycogen phosphorylase	362
SSF46689	Homeobox-like domain	336
SSF53335	S-adenosyl-L-methionine-dependent methyltransfer-ase	335
SSF48403	Ankyrin repeat-containing domain	324
SSF81383	F-box-like domain	322
SSF54928	RNA-binding domain	321
SSF51445	Glycoside hydrolase	313
SSF50978	WD40-repeat-containing domain	291

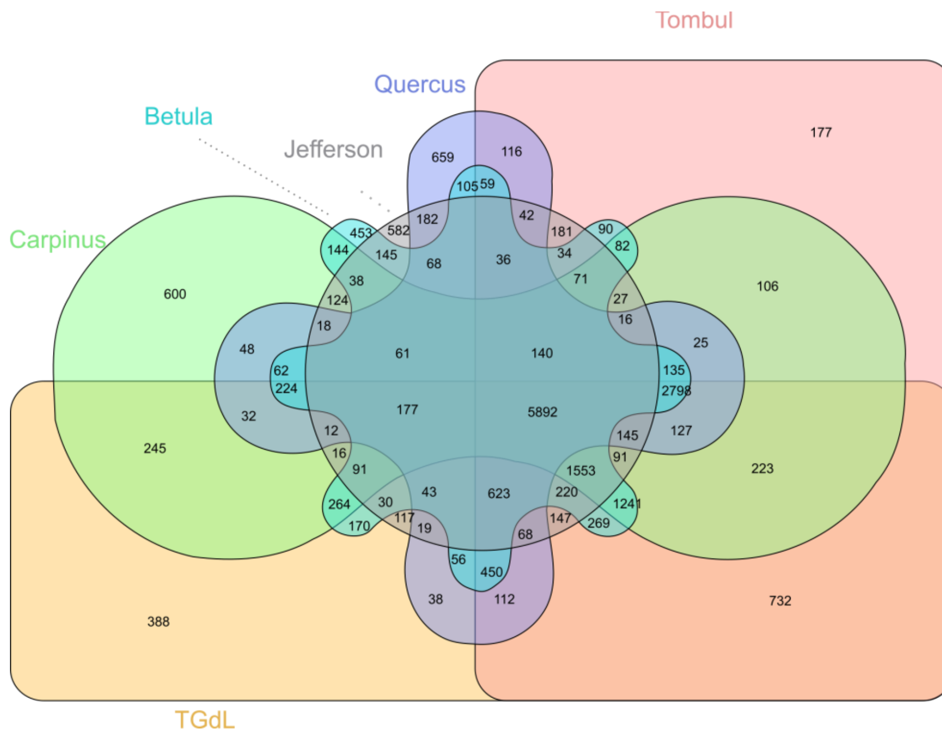

Figure 2: Orthofinder analysis performed using 'TGdL', 'Jefferson', 'Tombul', *Carpinus fangiana*, *Betula pendula* and *Quercus suber* genomes.

Resistance genes

Many plant-pathogen interactions are determined by the presence of resistance (R) genes/alleles, which enable plants to recognize pathogen effectors and subsequently activate effector-triggered immunity (ETI) (Sekhwal *et al.,* 2015), followed by a defense response often leading to cell death or a hypersensitive response (HR) (Zaidi *et al.,* 2018). Most intracellular immune receptors in plants belong to the nucleotide-binding site and leucine-rich repeat (NLR, also known as NB-LRR) superfamilies

(Eitas *et al.,* 2010; Lee *et al.,* 2015). The NLR superfamily proteins include two classes on the basis of the presence of a toll and interleukin-1 receptor domain in the N-terminus (TIR-NLR or TNL) or its absence (non-TIR-NLR or non-TNL). Some non-TNL proteins have a coiled-coil motif (CC-NLR or CNL).

The RGAugury pipeline detected between 86 and 2,017 resistance gene analogues (RGAs) among the species/genotypes analysed (Table 6). The highest percentage of RGAs compared to the total number of genes was found in *Quercus suber* (7.82%), while the lowest was detected in 'Jefferson' (0.31%, presumably as a consequence of the low quality genome annotation) and *Carpinus* (1.91). In the 'TGdL' assembly we identified a total of 810 RGAs. Furthermore, clustering of RLKs, RLPs, NBS-encoding and TM-CC genes in some chromosomes were detected (Table 7), in agreement with classical genetics and analysis from large scale sequencing data in plant genomes (Rody *et al.,* 2019). The chr. 2 was the richest in RGAs followed by 5, 3 and 4, while chr. 10 was the poorest. The majority of RLK genes was found on chrs. 2, 5, and 6, while the majority of RLP on chr. 2, 9, and 7.

The majority of RGAs were receptor like kinases (RLKs), followed by receptor like proteins (RLP), while only few RGAs contain at least one NB-ARC domain. Similarly, in other members of the order Fagales, i.e. *Juglans microcarpa* and *J. regia*, the most represented RGAs belong to RLK while few TNLs were identified (Zhu *et al.,* 2019). Comparable results have been obtained in other non-woody species, like Solanaceae species such as *Capsicum annuum*, *Solanum melongena*, *Solanum lycopersicum* and *Solanum tuberosum* (Barchi *et al.*, 2019 and Acquadro

et al., 2020a). Furthermore, Kim *et al.*, 2012 highlighted that some Asterids contain functional TNLs, whereas others do not. This resulted in the identification of only 19 and 13 full length CNLs in sunflower and lettuce respectively, but no full length TNLs. Recently, Acquadro *et al.* (2017) reported that in the *Cynara cardunculus* genome, the RGAs belong almost exclusively to the RLK/RLP families, while no TNLs and few CNLs were identified. This species-specific RGAs distribution was also observed in *Brassica oleracea, B. rapa, Arabidopsis* and *Theobroma cacao*, where the number of TNL was higher than CNL, while an opposite situation was found for *Populus trichocarpa, Vitis vinifera* and *Medicago truncatula* (Yu *et al.,* 2014).

The alignments of the amino acid sequences and subsequent IQ-TREE analyses generated phylogenetic trees for CNL-TNL, RLP and RLK RGA classes (Figure S1 3a-3c).

It has been reported that several R-genes (*Triticum aestivum Pm3*, *Arabidopsis thaliana RPP13*, *Linum usitatissimum* and *Capsicum annuum eIF4E*) seem to have evolved following a co-evolutionary relationship with pathogens and thus environment. The difference in terms of number and phylogenetic relationship represent a valuable information for conducting future in-depth studies on particular genes that are associated with their local environment (Charron *et al.,* 2008; Rose *et al.,* 2004).

Table 6: Resistance (R) genes in the 'TGdL' hazelnut genome compared with the genomes of 'Jefferson' and 'Tombul' and those of *Betula pendula*, *Carpinus fangiana* and *Quercus suber*. For each resistance gene class, the number as well as the percentage over the total number of genes is reported. Resistance genes abbreviations (from Li et al. 2016): NBS: nucleotide-binding site; CNL: CC (coiled-coil)-NBS-LRR; TNL: TIR (Toll/Interleukin-1 receptor)-NBS-LRR; CN: CC-NBS; TN: TIR-NBS; NL: NBS-LRR; TX: TIR-unknown domain; RLK: receptor like kinase; RLP: receptor like protein; TM (transmembrane)-CC.

Species/ Genotype	NBS	CNL	TNL	CN	TN	NL	TX	Others	RLP	RLK	TM-CC	Total
'TGdL'	18 (0.06%)	32 (0.12%)	1 (0%)	23 (0.08%)	1 (0%)	33 (0.12%)	6 (0.02%)	1 (0%)	93 (0.33%)	547 (1.97%)	55 (0.2%)	810 (2,91%)
'Jefferson'	2 (0.01%)	0 (0%)	0 (0%)	1 (0%)	0 (0%)	2 (0.01%)	2 (0.01%)	0 (0%)	11 (0.04%)	67 (0.24%)	1 (0%)	86 (0,31%)
'Tombul'	14 (0.05%)	43 (0.15%)	14 (0.05%)	12 (0.04%)	5 (0.02%)	28 (0.1%)	23 (0.08%)	6 (0.02%)	123 (0.44%)	673 (2.42%)	190 (0.68%)	1131 (4,07%)
Betula pendula	20 (0.07%)	10 (0.04%)	48 (0.17%)	0 (0%)	16 (0.06%)	31 (0.11%)	70 (0.25%)	5 (0.02%)	30 (0.11%)	662 (2.38%)	50 (0.18%)	942 (3,39%)
Carpinus fangiana	3 (0.01%)	23 (0.08%)	0 (0%)	4 (0.01%)	1 (0%)	30 (0.11%)	4 (0.01%)	0 (0%)	55 (0.2%)	291 (1.05%)	51 (0.18%)	462 (1,66%)
Quercus suber	47 (0.17%)	240 (0.86%)	174 (0.63%)	30 (0.11%)	16 (0.06%)	309 (1.11%)	118 (0.42%)	26 (0.09%)	286 (1.03%)	736 (2.65%)	35 (0.13%)	2017 (7,26%)

Table 7: Distribution of the resistance R genes among the 11 pseudomolecules of the 'TGdL' hazelnut genome

'TGdL' pseudomol-ecules	CN	CNL	NBS	NL	other	RLK	RLP	TM-CC	TN	TNL	TX	Total
1	-	1	-	-	-	45	10	7	-	-	-	63
2	6	9	9	18	-	96	18	2	-	-	3	161
3	2	4	2	3	-	51	9	12	1	-	1	85
4	2	3	-	2	1	53	8	6	-	1	-	76
5	1	5	-	-	-	67	4	7	-	-	2	86
6	1	-	-	4	-	58	6	-	-	-	-	69
7	-	2	1	1	-	44	9	8	-	-	-	65
8	9	4	3	1	-	33	7	3	-	-	-	60
9	-	1	-	2	-	31	11	2	-	-	-	47
10	1	3	2	2	-	20	5	1	-	-	-	34
11	-	-	-	-	-	43	5	4	-	-	-	52
unplaced scaffolds	1	-	1	-	-	6	1	3	-	-	-	12
Total	23	32	18	33	1	547	93	55	1	1	6	810

Conclusions

We performed a whole-genome assembly, using a combination of 10X Chromium linked-read technology and accurate 150 bp paired-end short-read Illumina sequencing, to generate the genome of the European hazelnut cv. 'TGdL', one of the best cultivars for processing due to its high kernel quality. A chromosome-scale assembly of 'TGdL' was built and will facilitate the detection of genomic variants, including copy number variations and large insertions/deletions. About 28.000 genes were identified and annotated with known homology. Since the European hazelnut 'TGdL' has excellent kernel quality and its genome sequences will be useful for studying important traits, predicting genes, and developing markers for use in breeding programs.

References

Acquadro, A., L. Barchi, E. Portis, M. Nourdine, C. Carli, *et al.,* 2020a Whole genome resequencing of four Italian sweet pepper landraces provides insights on sequence variation in genes of agronomic value. *Sci Rep.* 10(1):9189. doi:10.1038/s41598-020-66053-2

Acquadro, A., E. Portis, D. Valentino, L. Barchi, S. Lanteri, 2020b "Mind the Gap": Hi-C technology boosts contiguity of the globe artichoke genome in low-recombination regions. *G3* (Bethesda), 10(10):3557-3564. doi:10.1534/g3.120.401446

Acquadro, A., L. Barchi, E. Portis, G. Mangino, D. Valentino, *et al.,* 2017 Genome reconstruction in *Cynara cardunculus* taxa gains access to chromosome-scale DNA variation. *Sci Rep.* 7(1):5617. doi:10.1038/s41598-017-05085-7

Alonge, M., S. Soyk, S. Ramakrishnan, X. Wang, S. Goodwin *et al.,* 2019 RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol.* 20:224. doi:10.1186/s13059-019-1829-6

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, D.J. Lipman, 1990 Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410. doi:10.1016/S0022-2836(05)80360-2

Barchi, L., M. Pietrella, L. Venturini, A. Minio, L. Toppino *et al.,* 2019 A chromosome-anchored eggplant genome sequence reveals key events in Solanaceae evolution. *Sci Rep.* 9(1):11769. doi: 10.1038/s41598-019-47985-w

Beltramo, C., N. Valentini, E. Portis, D. Torello Marinoni, P. Boccacci *et al.,* 2016 Genetic mapping and QTL analysis in European hazelnut

(*Corylus avellana* L.). *Mol Breeding* 36:27. doi:10.1007/s11032-016-0450-6

Botta, R., T.J. Molnar, V. Erdogan, N. Valentini, D. Torello Marinoni, *et al.,* 2019 Hazelnut (*Corylus spp.*) Breeding: 157-219. Chapter 6 in: Advances in Plant Breeding Strategies, Vol 4: Nut and Beverage Crops. Al-Khayri JM, Jain SM, Johnson DV (eds). Springer Nature 607 p. ISBN 978-3-030-23111-8. doi:10.1007/978-3-030-23112-5_6

Bromberg, Y., and B. Rost, 2007 SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res.* 35: 3823–3835. doi:10.1093/nar/gkm238

Campbell, M. S., M. Law, C. Holt, J. C. Stein, G. D. Moghe *et al.,* 2014 MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiol.* 164: 513–524. doi:10.1104/pp.113.230144

Charron, C., M. Nicolaï, J.L. Gallois, C. Robaglia, B. Moury, A. *et al.* 2008 Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. Plant J. 54(1):56-68. doi:10.1111/j.1365-313X.2008.03407.x.

Chen, X., J. Zhang, Q. Liu, W. Guo, T. Zhao, *et al.,* 2014 Transcriptome sequencing and identification of cold tolerance genes in hardy *Corylus* species (*C. heterophylla* Fisch) floral buds. *PLoS One*. e108604. doi:10.1371/journal.pone.0108604

Eitas, T. K., and J.L. Dangl, 2010 NB-LRR proteins: Pairs, pieces, perception, partners, and pathways. *Current Opinion in Plant Biology* 13:472–477

FAOSTAT, 2019 Food and Agriculture Organization of the United Nations http://www.fao.org/faostat/ (2019), Accessed 20th Apr 2020

Hulse-Kemp, A.M., S. Maheshwari, K. Stoffel, T.A. Hill, D. Jaffe *et al. 2018* Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Hortic Res.* 5, 4. doi:10.1038/s41438-017-0011-0

Jaillon, O., and J.M. Aury 2007 The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*. 449: 463–467

Jones, P., D. Binns, H.-Y. Chang, M. Fraser, W. Li *et al.,* 2014 InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240

Kajitani, R., K. Toshimoto, H. Noguchi, A. Toyoda, Y. Ogura, *et al.,* 2014 Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* 24(8):1384-95. doi:10.1101/gr.170720.113

Kim, J., C.J. Lim, B.W. Lee, J.P. Choi, S.K. Oh *et al.,* 2012 A genome-wide comparison of NB-LRR type of resistance gene analogs (RGA) in the plant kingdom. *Mol Cells* 33(4):385-92. doi: 10.1007/s10059-012-0003-8

Lee, H.A., and S.I. Yeom 2015 Plant NB-LRR proteins: tightly regulated sensors in a complex manner. *Brief. Funct. Genomics* 14:233–242

Li, P., X. Quan, G. Jia, J. Xiao, S. Cloutier *et al.* 2016 RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. *BMC Genomics* 17(852). doi:10.1186/s12864-016-3197-x

Lucas S.J., K. Kahraman, B. Avşar, R.J.A. Buggs, I. Bilge 2020 A chromosome-scale genome assembly of European Hazel (*Corylus avellana* L.) reveals targets for crop improvement. *Plant J.* doi:10.1111/tpj.15099

Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang *et al.,* 2012 SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1(1):18. doi:10.1186/2047-217X-1-18

Ma, H., Z. Lu, B. Liu, Q. Qiu, J. Liu, 2013 Transcriptome analyses of a Chinese hazelnut species *Corylus mandshurica*. *BMC Plant Biol.* 13:152. doi:10.1186/1471-2229-13-152

Mehlenbacher, S. A., R. N. Brown, E. R. Nouhra, T. Gökirmak, N. V. Bassil *et al.,* 2006 A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. Genome 49: 122–133. doi.org/10.1139/g05-091

Molnar, T. J., 2011 *Corylus*. In: Kole C. (eds) Wild Crop Relatives: Genomic and Breeding Resources. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-21250-5_2

Öztürk, S.C., M. Göktay, J. Allmer, S. Doğanlar, A. Frary, 2018 Development of simple sequence repeat markers in hazelnut (*Corylus avellana L.*) by next-generation sequencing and discrimination of turkish hazelnut cultivars. *Plant Mol Biol Report* 36:800-811. doi: 10.1007/s11105-018-1120-0

Pustahija, F., S.C. Brown., F. Bogunic, N. Basic, E. Muratovic*, et al.,* 2013 Small genomes dominate in plants growing on serpentine soils in West Balkans, an exhaustive study of 8 habitats covering 308 taxa. Plant and Soil 373:427–453.

Rody, H.V.S., R.G.H. Bombardelli, S. Creste, L.E.A. Camargo, M.A. Van Sluys *et al.,* 2019 Genome survey of resistance gene analogs in sugarcane: genomic features and differential expression of the innate immune system from a smut-resistant genotype. *BMC Genomics* 20(1):809. doi:10.1186/s12864-019-6207

Rose, L. E., P. D.Bittner-Eddy, C. H. Langley, E.B. Holub, R.W. Michelmore *et al.,* 2004 The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in *Arabidopsis thaliana*. *Genetics*, *166(*3), 1517–1527.doi:10.1534/genetics.166.3.1517

Rowley, E.R., S. E. Fox, D. W. Bryant, C. M. Sullivan, S. A. Givan *et al.,* 2012 Assembly and characterization of the European hazelnut (*Corylus avellana* L.) 'Jefferson' transcriptome. *Crop Science* 52: 2679–2686. doi:10.2135/cropsci2012.02.0065

Rowley, E.R., R. Vanburen, D.W. Bryant, H. D. Priest, S.A Mehlenbacher, *et al.,* 2018. A Draft genome and High-Density Genetic Map of European hazelnut (*Corylus avellana* L.). *bioArxiv* 1–25. doi:10.1101/469015

Sekhwal, M.K., P. Li, I. Lam, X. Wang, S. Cloutier *et al.,* 2015 Disease Resistance Gene Analogs (RGAs) in *Plants. Int. J. Mol. Sci.* 16:19248-19290.

Smit, A., R. Hubley, and P. Green, 2013–2015 RepeatMasker Open-4.0. http://www.repeatmasker.org/faq.html.

Stanke, M., O. Keller, I. Gunduz, A. Hayes, S. Waack *et al.,* 2006 AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34: W435–W439. doi:10.1093/nar/gkl200

Torello Marinoni, D., N. Valentini, E. Portis, A. Acquadro, C. Beltramo *et al.,* 2018 High density SNP mapping and QTL analysis for time of leaf

budburst in *Corylus avellana* L. *PLoS One* 13(4): e0195408. doi:10.1371/journal.pone.0195408

Valentini, N., F. Calizzano, P. Boccacci, R. Botta, 2014 Investigation on clonal variants within the hazelnut (*Corylus avellana* L.) cultivar 'Tonda Gentile delle Langhe'. *Scientia Horticulturae* 165: 303-310. doi:10.1016/j.scienta.2013.11.027

Velasco, R., A. Zharkikh, M. Troggio, D.A. Cartwright, A. Cestaro *et al.,* 2007 A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS One*. 2(12):e1326. doi:10.1371/journal.pone.0001326

Weisenfeld, N.I., P. Kumar, P. Shah, D. M. Church, D. B. Jaffe, 2017 Direct determination of diploid genome sequences. *Genome Research* 27: 757-767. doi:10.1101/gr.214874.116

Yu, J., S. Tehrim, F. Zhang, C. Tong, J. Huang, *et al.* 2014 Genome-wide comparative analysis of NBS-encoding genes between *Brassica* species and *Arabidopsis thaliana*. *BMC Genomics* 15(1):3. doi:10.1186/1471-2164-15-3.

Zaidi, S. S., M. S. Mukhtar, S. Mansoor 2018 Genome Editing: Targeting susceptibility genes for plant disease resistance. *Trends in Biotechnology* 36: 898–906

Zhu, T., L. Wang, F.M. You, J.C. Rodriguez, K.R. Deal, L. *et al.,* 2019 Sequencing a *Juglans regia* × *J. microcarpa* hybrid yields high-quality genome assemblies of parental species. *Hortic Res*. 6:55. doi:10.1038/s41438-019-0139-1

Supplementary materials

Supplementary materials (Table S1, S2 and Figure S1) are available at the following link: https://doi.org/10.25387/g3.14502048/.

Chapter II

Identification of susceptibility genes in *Castanea sativa* and their transcription dynamics following pathogen infection

Vera Pavese[1], Andrea Moglia[1]*, Paolo Gonthier[1], Daniela Torello Marinoni[1], Emile Cavalet-Giorsa[1], Roberto Botta[1]

1. Dipartimento di Scienze Agrarie, Forestali e Alimentari-DISAFA, Università degli Studi di Torino, Largo Paolo Braccini 2, Grugliasco, 10095 Torino, Italy

*Correspondence: **andrea.moglia@unito.it**

Abstract

Castanea sativa is one of the main multipurpose tree species valued for its timber and nuts. This species is susceptible to two major diseases, ink disease and chestnut blight, caused by *Phytophthora* spp. and *Cryphonectria parasitica*, respectively. The loss-of-function mutations of genes required for the onset of pathogenesis, referred to as plant susceptibility (S) genes, are one mechanism of plant resistance against pathogens. On the basis of sequence homology, functional domain identification, and phylogenetic analyses, we report for the first time on

the identification of S-genes (*mlo1, dmr6*, *dnd1*, and *pmr4*) in the *Castanea* genus. The expression dynamics of S-genes were assessed in *C. sativa* and *C. crenata* plants inoculated with *P. cinnamomi* and *C. parasitica*. Our results highlighted the upregulation of *pmr4* and *dmr6* in response to pathogen infection. *Pmr4* was strongly expressed at early infection phases of both pathogens in *C. sativa*, whereas in *C. crenata,* no significant upregulation was observed. The infection of *P. cinnamomi* led to a higher increase in the transcript level of *dmr6* in *C. sativa* compared to *C. crenata*-infected samples. For a better understanding of plant responses, the transcript levels of defence genes *gluB* and *chi3* were also analysed.

Keywords: chestnut, susceptibility genes, *Phytophthora cinnamomi*, *Cryphonectria parasitica*

Introduction

The *Castanea* genus belongs to the Fagaceae family and includes four major species of commercial and ecosystemic interest: *Castanea sativa* Mill. (European chestnut), *Castanea crenata* Sieb. et Zucc. (Japanese chestnut), *Castanea mollissima* Bl. (Chinese chestnut), and *Castanea dentata* Borkh (American chestnut). *C. sativa* is a woody species common in all Mediterranean countries and Asia Minor. It has been widely used since ancient times, not only for the consumption of its edible nuts, but also for wood and the products of its ecosystem, such as mushrooms and honey. It is a forest tree, relevant for landscape ecology and biodiversity of mountain and rural areas [1].

Over the last century, the number of chestnut trees decreased in growing areas in Europe due to the depopulation of mountains, climate change, and the spread of two severe diseases: ink disease and chestnut blight [2,3]. Ink disease is caused by the Oomycete *Phytophthora cinnamomi* and *Phytophthora cambivora*. Both species are pathogenic to *C. sativa*, although *P. cinnamomi* generally displays greater virulence than *P. cambivora* [4,5]. Among *Castanea* species, only *C. crenata* exhibits high tolerance to *P. cinnamomi* [6]. The disease, which affects both young and old trees, leads to subcortical necrosis of the root system and the basal part of the stem; this is followed by the appearance of wasting symptoms in the foliage until the total desiccation and death of the plant occur [7,8,9,10]. These pathogens spread mainly through the movement of soil harboring inoculum and the dissemination of asexual flagellated spores (i.e., zoospores) that can actively travel short distances or passively travel long distances in flowing water [10,11]. The use of resistant rootstocks represents one possible solution to protect against these pathogens, although, at present, only tolerant selections obtained from hybridization between *C. sativa* and *C. crenata* are available [12].

Chestnut blight stands among the most destructive fungal tree diseases ever [10,13]. The causal agent, *Cryphonectria parasitica*, infects trees through dead plant tissue and wounds, including those caused by pruning, graft, and hail [13,14]. The symptoms involve bark cankers that can develop on suckers, young branches, and adult branches and trunks [15]. Chestnut blight was one of the causes of the abandonment of chestnut orchards in Europe until the end of the 1970s, when the natural spread of the hypovirulent form of the fungus favored a slow but progressive

recovery of chestnut orchards and coppices. However, the fungus still represents a relevant problem in many areas of Europe. It is very harmful to young grafted trees in particular, hampering the establishment of new orchards in many areas [10,13].

C. dentata forests in Eastern North America were wiped out by *C. parasitica* in the early 20th century [16]. Extensive studies and breeding activities have been carried out to restore the American chestnut species introgressing resistance genes of *C. mollissima* [17,18]. More recently, researchers discovered that the onset of the disease is associated with the release of oxalic acid by the pathogen during infection. Blight-resistant *C. dentata* trees were obtained by transferring a wheat gene that encodes oxalate oxidase [19].

Recently, a new interest and sensitivity towards the preservation of the local landscape generated a growing interest in silviculture and chestnut trees [20]. Moreover, the market demand for chestnuts in European countries has been strong in the last two decades and has often been supplied by importations. This has been due in part to the gall wasp (*Dryocosmus kuriphilus* Yasumatsu) infestation, which only recently has been controlled effectively [1,21], and to the general difficulty of developing a modern chestnut industry based on quality cultivars of *C. sativa* that are more tolerant to pathogens. The elucidation of the genetic mechanism behind host–pathogen interaction could thus be useful for the development of novel breeding strategies aimed at achieving resistance or higher tolerance to these pathogens.

Plants take advantage of different defense mechanisms during pathogen attack, and pathogens trigger counter-defense mechanisms. Plants carry

pattern recognition receptors (PRRs) able to perceive pathogen-associated molecular patterns (PAMPs); this perception leads to intracellular signal transduction culminating in PAMP-triggered immunity (PTI). PTI is characterized by the production of reactive oxygen species (ROS), the secretion of antimicrobial compounds, and hydrolytic enzymes targeting the pathogen cell wall (chitinase and glucanase) and local cell wall fortifications (through callose deposition) [22].

To suppress PTI, pathogens developed effector molecules able to facilitate pathogen infection by manipulating the host response to support compatibility. Plant resistance (R) genes can detect effectors and trigger effector-triggered immunity (ETI) [23]. The recognition between R genes and effectors causes a cascade of responses involving jasmonic acid (JA) and salicylic acid (SA), culminating in a hypersensitive response (HR) [24].

Most pathogens require the cooperation of the host to establish a compatible interaction. Plant genes supporting compatibility and facilitating infection are called susceptibility (S) genes. S-genes can be divided into three main classes: a) genes required for the early pathogen infection step (basic compatibility); b) genes encoding negative regulators of plant immunity; c) genes necessary for pathogen proliferation (sustained compatibility) [22].

The mutation or loss of an S-gene can thus limit the ability of the pathogen to infect the host and the spread of the disease. The resistance mediated by the S-gene mutation can be pathogen-specific or broad-spectrum. In the former case, the pathway can be implicated in the penetration phase; in the latter, one of the target genes can be involved in constitutive defense

responses [22]. Resistance due to the loss of S-genes is generally recessive, differing from the generally dominant resistance mediated by R genes.
Among the S-genes, *Mildew resistance locus O* (*mlo1*), *Powdery mildew resistance 4* (*pmr4*), *Downy Mildew Resistance 6* (*dmr6*), and *Defense no death* (*dnd1*) have been characterized in many plant species. The *Mlo* gene family, encoding seven transmembrane domain proteins, has been characterized in many plant species [25]. Some *mlo* homologs act as PM susceptibility factors, as their loss of function results in a distinguished type of resistance known as *mlo* resistance. Originally discovered in barley (*Hordeum vulgare* L.), *mlo* resistance was later shown to occur in several monocot and eudicot species, namely Arabidopsis, tomato (*Solanum lycopersicum* L.), pea (*Pisum sativum* L.), pepper (*Capsicum annum* L.), tobacco (*Nicotiana tabacum* L.) and wheat (*Triticum aestivum* L.) [26,27] plants. The callose synthase encoded by *pmr4* is responsible for the production of callose in response to biotic and abiotic stresses. In tomato and potato plants the knockout and silencing of *pmr4* led to *Oidium neolycopersici* and *Phytophthora infestans* tolerance [28,29]. *Dmr6* is involved in the conversion of salicylic acid (SA) to 2,3-dihydroxybenzoic acid (2,3-DHBA) and negatively regulates defense gene expression [30]. Its silencing caused resistance to hemi-biotrophic *Phythophthora capsici*, *Hyaloperonospora arabidopsidis,* and *Pseudomonas syringae* [31]. Mutants of *dnd1*, encoding for a cyclic nucleotide-gated cation channel, showed *P. infestans* resistance [29].
At the moment, studies on S-genes in woody plant species have been carried out only for *mlo* genes in rubber trees [32], poplar trees [33], apple trees, and grapevines [34]. In our work, we report on the S-genes

identification and characterization in *C. sativa* on the basis of sequence homology, functional domain detection and phylogenetic relationships. In addition, the expression dynamics of S-genes were assessed in *C. sativa* and *C. crenata* plants inoculated with the two pathogens, *P. cinnamomi* and *C. parasitica*, belonging to different kingdoms. Using the same plant material, the transcription levels of key genes involved in pathogen resistance, *chi3* (*acidic 26 kDa endochitinase*) and *gluB* (*glucan endo-1,3-beta- glucosidase),* were also determined (S1 File). Our analysis revealed the strong activation of *pmr4* and *dmr6 genes* in response to infection by both *P. cinnamomi* and *C. parasitica.*

Results

Genes identification and structure

Chestnut susceptible (S) genes were identified in the *C. mollissima* v1.1 reference genome using available coding sequences of gene orthologs as a query (S2 File). Based on the blastn survey, four loci with high similarity were identified and attributed to different subclasses of S-genes due to the presence of specific domains: *mlo1*, *dmr6*, *dnd1*, *and pmr4* (Fig. 1). The coding sequence length of *mlo1* is 1425 bp (composed of 13 exons); the protein size is 474 amino acids (aa) (Table 1). A single Mlo domain (PF03094) is present within the protein sequence. The *Dmr6* gene, whose coding sequence is 1128 bp, contains four exons and is translated into a 375 aa protein (Table 1). Two specific domains are characterized: 2OG-FeII_Oxy and DIOX_N (PF03171; PF14226). *Dnd1* is 1407 bp in length, codes for 468 aa proteins, and is composed of six exons. Two structural domains, cNMP_binding and Ion_trans, were highlighted (PF00027;

PF00520) (Fig 1; Table 1). The *Pmr4* gene is characterized by one single 5346 bp exon. The protein size is 1781 aa, and the structural domains are FKS1 dom1 and Glucan_synthase (PF14288; PF02364) (Table 1).

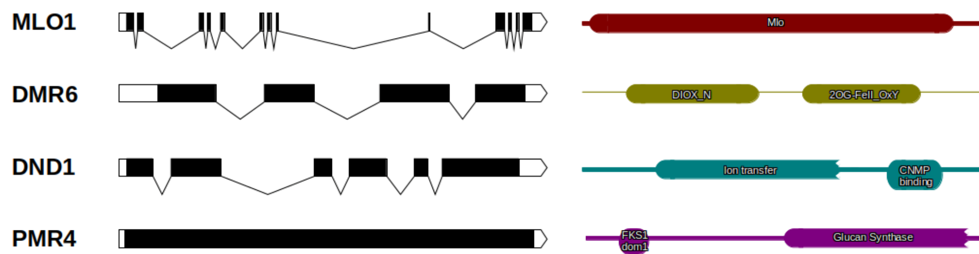

Figure 1. Chestnut S-genes and their protein structures. The graphical representations of gene exon/intron structures were generated using the http://wormweb.org/exonintron/ tool and are shown in the left panel. The exons are indicated with black boxes, whereas introns are shown with lines. In the right panel, the protein structural domains are displayed

Table 1. S-genes detected in the *C. mollissima* v1.1 genome and protein domain annotations.

Gene Name	Scaffold	ORF length (bp)	N° Exons	Size (aa)	Domains	PFAM DOMAINS
MLO1	Scaffold 00101	1425	13	474	Mlo	PF03094
DMR6	Scaffold 02358	1128	4	375	2OG-FeII_Oxy; DIOX_N	PF03171;PF14226
DND1	Scaffold 00410	1407	6	468	cNMP_bind-ing; Ion_trans	PF00027;PF00520
PMR4	Scaffold 00300	5346	1	1781	FKS1_dom1; Glucan_syn-thase	PF14288;PF02364

Phylogenetic and modelling analysis

Available full-length NCBI S-gene coding sequence orthologues (S3 File) were used for phylogenetic tree construction. The resulting unrooted maximum-likelihood trees are shown in Fig. 2, and all the phylogenetic trees are available separately in the S4 File.

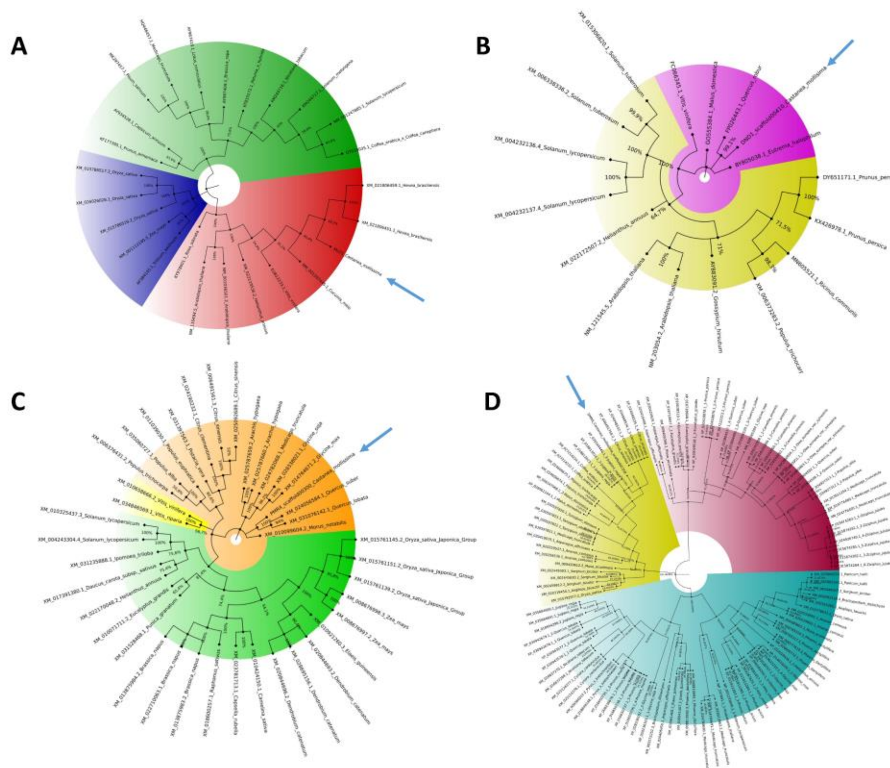

Figure 2. Phylogenetic analysis of the S-genes. The 4 phylogenetic trees of mlo1 (A), dnd1 (B), pmr4 (C), and dmr6 (D) were constructed using MEGAX software by aligning chestnut S-gene coding sequences with NCBI S-gene ortholog coding sequences (available in file S3). The colors indicate the main clades detected, and the arrows underline the location of *C. mollissima*. To visualize details, all the phylogenetic trees are available in file S4.

The *mlo1* tree was grouped into three clades (blue, green, red) with 100% bootstrap value. The monocot proteins formed a separate clade (blue) with respect to those of the dicotyledonous species. *CmMlo1* is located in the red clade with *Vitis vinifera* and *Hevea brasiliensis* orthologs (Fig.

2A). The 17 *dnd1* coding sequences were divided into two subclades; no monocotyledon genes were included. *CmDnd1* is in the violet subclade, phylogenetically close to *Quercus robur* ortholog (Fig. 2B). For the construction of *pmr4* and *dmr6* trees, a greater number of coding sequences were available: 40 and 115, respectively. The *Pmr4* phylogenetic tree showed the division in three main clades, with monocots in the green clade intermixed with dicots. *CmPmr4* is located in the orange clade and clusters together with *Quercus lobata* and *Quercus* spp. orthologs (Fig. 2C). The *Dmr6* tree is divided into three clades intermixed with monocot/dicot proteins. *CmDmr6* is in the yellow clade, phylogenetically close to the *Juglans regia* ortholog gene (99% bootstrap value) (Fig. 2D).

By comparing the 3D protein structure model of *C. mollissima DMR6* with *Arabidopsis thaliana DMR6* via Modeller software, a high degree of structural conservation was observed (Fig. 3). DMR6 is a putative oxygenase involved in the conversion of salicylic acid (SA) to 2,3-dihydroxybenzoic acid (2,3-DHBA), and its catalytic activity is probably necessary to suppress plant immunity. The catalytic triad that binds the iron atom (grey sphere) is made by two histidines (H212 and H269), and an aspartic acid residue (D214). Zeilmaker et al. [31] demonstrated, by removing the histidines, that these residues were fundamental for the catalytic activity of DMR6, as well as for its role in immune suppression.

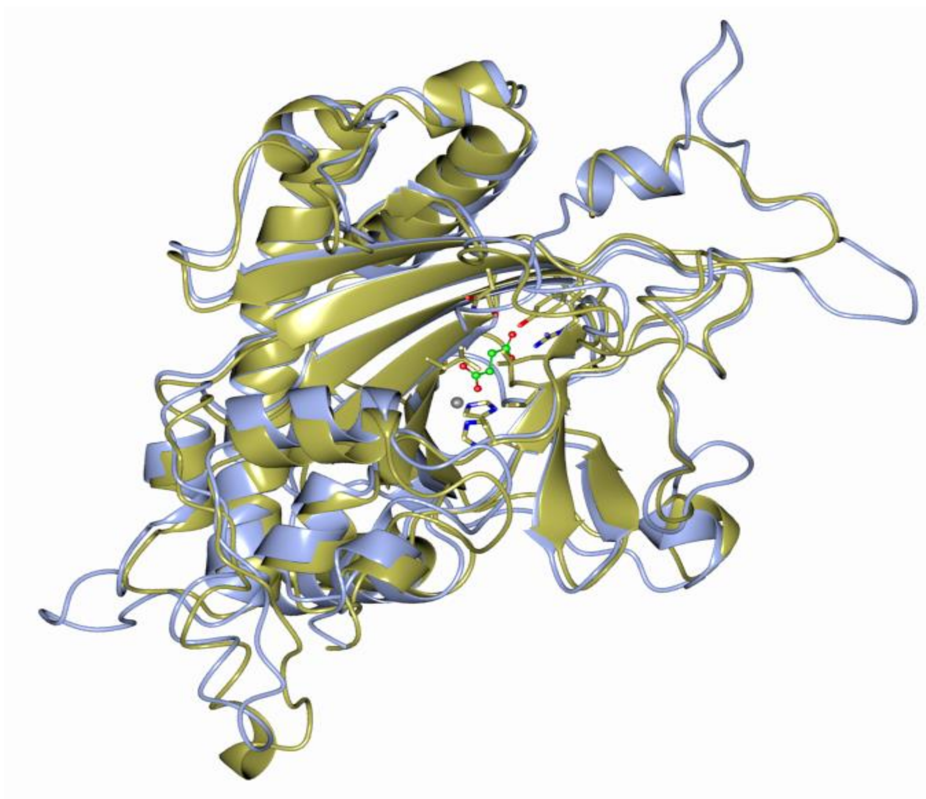

Figure 3. The DMR6 3D protein model created using Modeller software and visualized using Ccp4mg software. The *C. mollissima* (Cm) DMR6 (blue) protein and A. thaliana (At) DMR6 (yellow) protein are shown.

Transcriptional profiling in response to P. cinnamomi infection

S-genes are genes related to plant–pathogen interaction and are supposed to be activated during the early stages of infection before symptoms emerge. The susceptible species *C. sativa* was used as a reference to define the onset of symptoms due to *P. cinnamomi* inoculation on the stem. Five days after inoculation, lesions [35] were observed, followed by total leaf

desiccation. Based on the evidence from the preliminary inoculation tests, *C. sativa* and *C. crenata* species inoculated with *P. cinnamomi* and samples were collected at 0, 3, 6, 12, 24, 48, and 72 hours post-inoculation (hpi). The wound area left a lesion of 0.5 cm^2. *C. sativa* plants inoculated with *P. cinnamomi* showed an enlargement of the lesion of 0.6 cm^2 at 24, 48, and 72 hpi compared to the initial area of the lesion at time 0 (control). No visible enlargement of lesions was recorded in the case of *C. crenata*. Quantitative PCR from infected stem tissues was applied to quantify *P. cinnamomi* and the assay confirmed a higher amount of the pathogen in plant tissues at 72 hpi (Fig. 4). *C. crenata* showed a lower abundance of the pathogen compared to *C. sativa* at all tested experimental time points. S-gene expression was analysed using the same time points (Fig. 5A). *Mlo1* was mainly expressed at early infection phases, peaking at 6 and 3 hpi in *C. sativa* and *C. crenata,* respectively. *Dnd1* showed an analogous trend in both plant species, and its transcription was strongest at 6 hpi. Regarding *pmr4*, a differential transcript regulation in response to *P. cinnamomi* infection was highlighted in *C. sativa* and *C. crenata*. In *C. sativa, pmr4* was strongly expressed at 3 and 6 hpi, with transcript level around 6-fold higher compared to 0 hpi. On the contrary, the expression of p*mr4* remained very limited in *C. crenata* infected tissues, and a significant downregulation at 48 and 72 hpi was observed. Significant increases in the transcript level of *dmr6* were observed only at 12 hpi (for *C. crenata*) and at 3 and 12 hpi (for *C. sativa*). For a better understanding of the plant response against *P. cinnamomi,* genes coding for pathogenesis-related (PR) proteins, *chi3* and *gluB*, were analysed (Fig. 5B). *Chi3* peaked at 6 hpi both in *C. sativa* and in *C. crenata,* but the increment was higher in the

former species (~3 fold vs ~2 fold). *GluB* increased with the progression of infection in *C. sativa*, whereas a higher upregulation was observed during the early infection phases (at 3 and 6 hpi) in *C. crenata*.

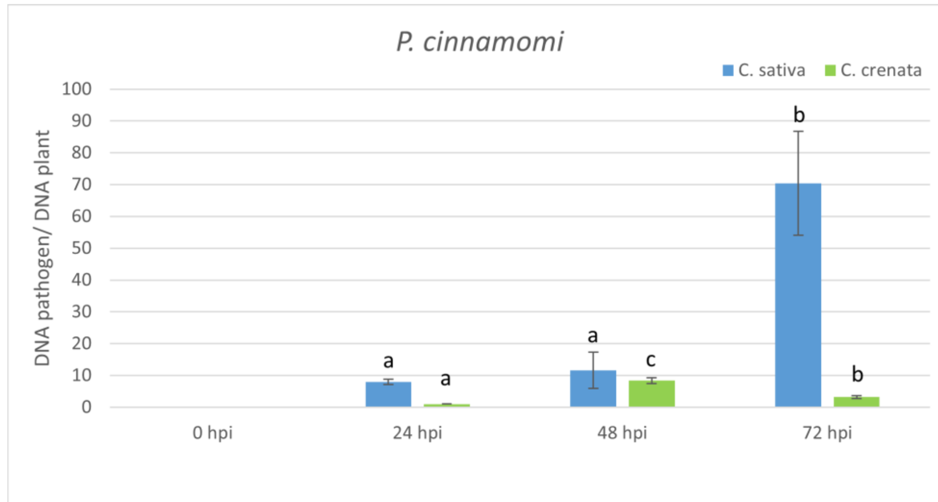

Figure 4. qRT-PCR pathogen DNA quantification after *P. cinnamomi* inoculation. Data were quantified using the $2^{-\Delta\Delta Ct}$ method based on the Ct values of pathogen genes (ypt and mf1) and actin-7 as a housekeeping gene. Data are the means of three biological replicates ± SE. *C. sativa* data are normalized with *C. sativa* 0 hpi control; *C. crenata* data are normalized with *C. crenata* 0 hpi control. Different letters associated with the set of means indicate a significant difference based on Tukey's HSD test ($P \leq 0.05$).

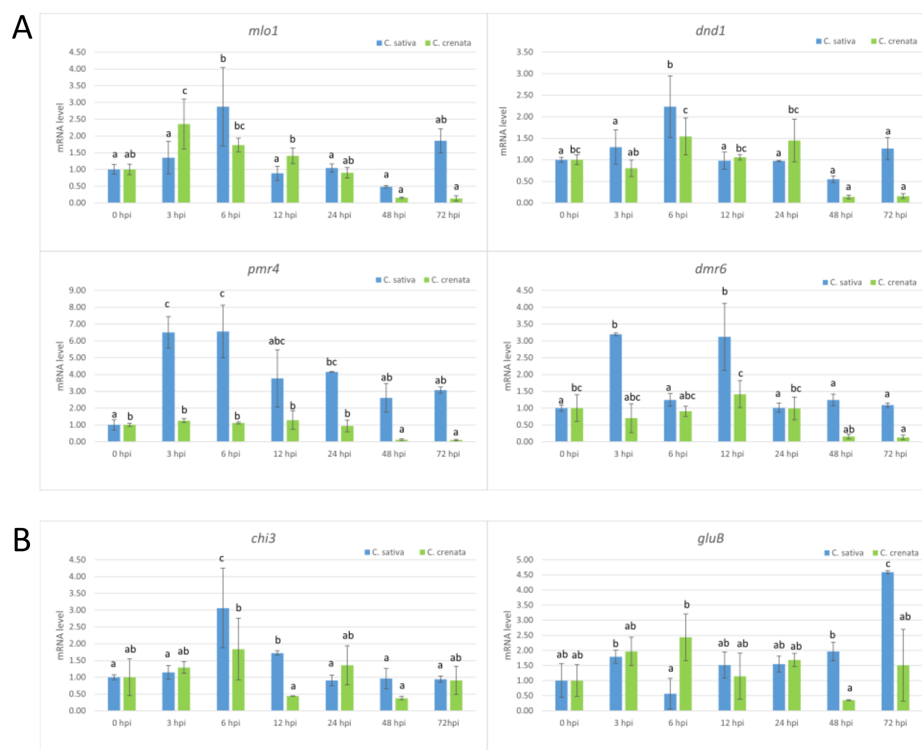

Figure 5. qRT-PCR-based transcription profiling after *P. cinnamomi* inoculation. A) The S-gene transcription profiles in *C. sativa* (blue) and *C. crenata* (green) chestnut species. B) The expression analysis of genes coding for several pathogenesis-related (PR) proteins in *C. sativa* (blue) and *C. crenata* (green) species. In all analyses, Cm7-actin was used as a housekeeping gene. Data are the means of three biological replicates ± SE. *C. sativa* data are normalized with *C. sativa* 0 hpi control; *C. crenata* data are normalized with *C. crenata* 0 hpi control. Different letters associated with the set of means indicate a significant difference based on Tukey's HSD test ($P \leq 0.05$).

Transcriptional profiling in response to C. parasitica infection

As described for *P. cinnamomi*, a preliminary stem inoculation assay on *C. sativa* plants was done with *C. parasitica*. Seven days after inoculation, a necrotic lesion around the inoculation point and orange fruiting bodies were observed in *C. sativa*. Based on the results of the preliminary inoculation test, *C. sativa* and *C. crenata* plants were inoculated and sampled at 0, 12, 24, 48, 72, 96, and 120 hpi. The size of the wound area was 0.5 cm^2. *C. sativa* plants inoculated with *C. parasitica* showed enlargements of the lesion of 0.5 cm^2 (at 48 and 120 hpi) and of 1 cm^2 (at 96 hpi) compared to the initial wound area at 0 hpi. No visible enlargement of the inoculation lesions was recorded in *C. crenata* at 72 and 120 hpi, and a limited enlargement of 0.5 cm^2 was observed at 96 hpi.

In *C. sativa,* qPCR analysis showed an increase in the abundance of pathogen inoculum with time elapsing from infection, peaking at 72 hpi. In *C. crenata*, no statistical differences were observed among the different experimental times (from 72 to 120 hpi) (Fig. 6). The transcript levels of S-genes were analysed using the same time points (Fig. 7A). *Mlo1* was mainly expressed at 24 hpi in *C. sativa*. No significant upregulation of *mlo1* was observed in *C. crenata*-infected tissues. An upregulation of *dnd1* was detected at 24 hpi in *C. sativa* infected tissues. In *C. sativa, pmr4* was strongly expressed at the early infection phases (12 and 24 hpi), with transcript levels around 3-fold higher as compared with the level at 0 hpi. On the contrary, *pmr4* was downregulated at all experimental times in inoculated *C. crenata* plants. A significant upregulation of *dmr6* was observed only at late infection phases, 48 hpi in *C. crenata* and 24 hpi in

C. sativa. The transcription of *chi3* was strongest at 96 hpi in *C. sativa* and at 48 hpi in *C. crenata*. The transcript level of *gluB* increased until reaching 96 hpi in *C. sativa* and 24 hpi in *C. crenata* (Fig. 7B).

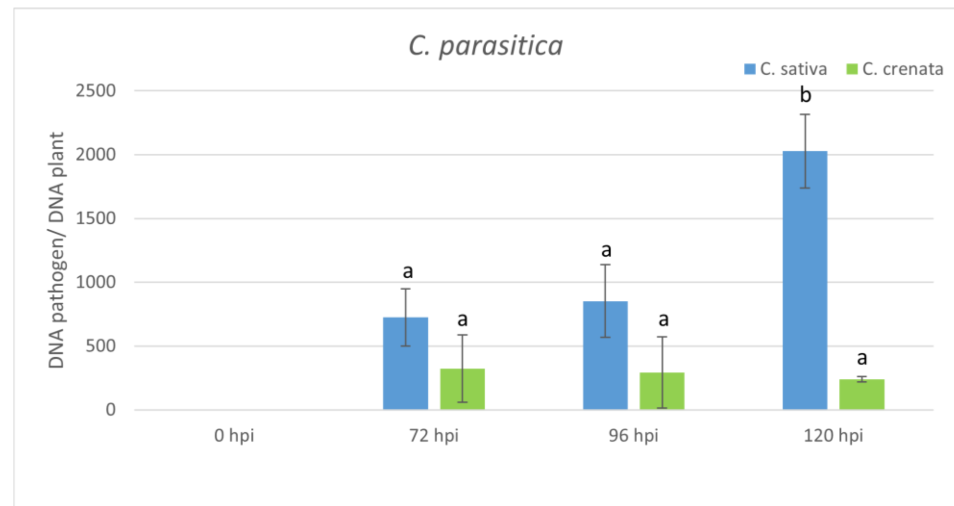

Figure 6. qRT-PCR pathogen DNA quantification after *C. parasitica* inoculation. Data were quantified using the $2^{-\Delta\Delta Ct}$ method based on the Ct values of fungal genes (ypt and mf1) with actin-7 as a housekeeping gene. Data are the means of three biological replicates ± SE. *C. sativa* data are normalized with the *C. sativa* 0 hpi control; *C. crenata* data are normalized with *C. crenata* 0 hpi control. Different letters associated with the set of means indicate a significant difference based on Tukey's HSD test ($P \leq 0.05$).

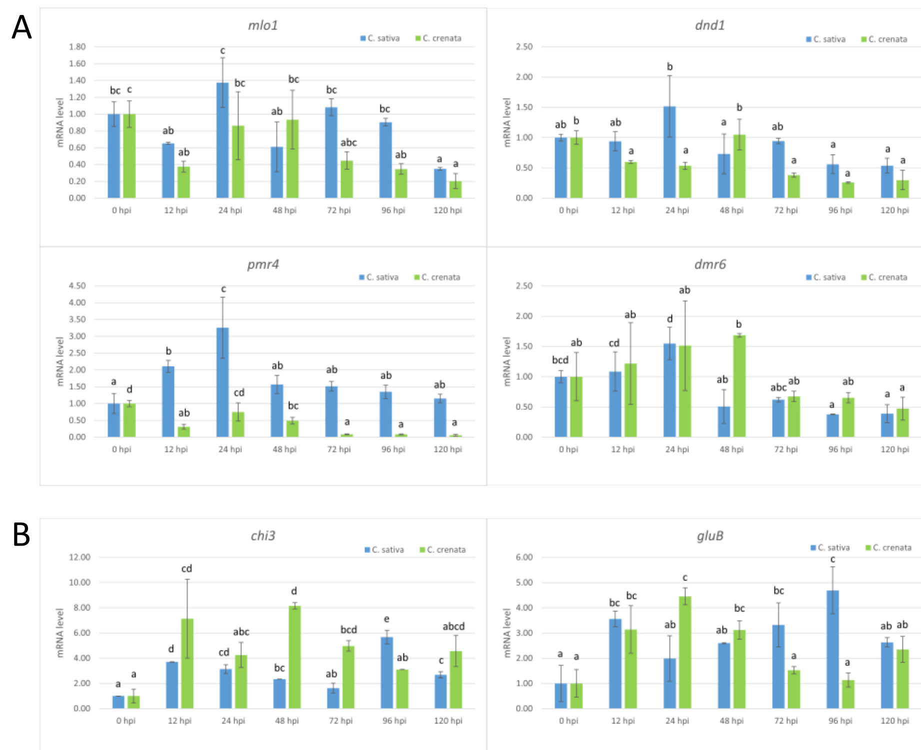

Figure 7. qRT-PCR-based transcription profiling after *C. parasitica* inoculation. A) The S-gene transcription profile in *C. sativa* (blue) and *C. crenata* (green) chestnut species. B) The expression analysis of genes coding for several pathogenesis-related (PR) proteins of *C. sativa* (blue) and *C. crenata* (green) species. In all the analyses, Cm7-actin was used as the housekeeping gene. The data are the means of three biological replicates ± SE. *C. sativa* data are normalized with *C. sativa* 0 hpi control; *C. crenata* data are normalized with *C.*

crenata 0 hpi control. Different letters associated with the set of means indicate a significant difference based on Tukey's HSD test ($P \leq 0.05$).

Discussion

C. sativa is a European woody tree species commonly used across the globe in the food and timber industries. This chestnut species is susceptible to the two major pathogens, *P. cinnamomi* and *C. parasitica* [10,36]. In contrast, the Asian chestnut species *C. crenata* and *C. mollissima* exhibit higher tolerance to *P. cinnamomi* and *C. parasitica* [6,39,42,50]. Achieving tolerance or resistance to pathogens is the major aim of rootstock breeding. Blight-resistant trees were obtained through backcross breeding of introgression genes from Asian species into American chestnut trees. [37]. However, this approach, although successful in developing blight-resistant American chestnut selections has been slowed by a lack of genetic tools. In Europe, ink disease tolerant hybrids were obtained through interspecific crosses between *C. sativa* and *C. crenata*, although the nut quality produced by these hybrids is below current market standards [38,39].

It has long been recognized that a deep understanding of a pathogen's biology, host–pathogen interactions, and the resistance mechanisms are fundamental to improving breeding programs. Genomic and transcriptomic analyses have provided the first genetic insights into mechanisms underlying susceptible and resistant chestnut species responses to *P. cinnamomi* and *C. parasitica* [37,38,40,41,42]. Santos et

al. [40] reported the upregulation of a set of candidate genes (e.g., *Cast_Gnk2-like* and *Calcium-dependent protein kinase*) after *P. cinnamomi* infection, which may trigger HR-like cell death in *C. crenata* cells. A significant number of genes involved in the defence against chestnut blight were identified. [37].

Traditionally, the introduction of resistance gene analogues into plants was the most promising approach to facilitate the acquisition of resistance. However, it did not prove to be durable enough because the widespread use of R genes caused the selection of pathogens capable of overcoming them [24]. Susceptibility (S) genes can be interesting candidates to be used in target breeding programs [22,23,24]. Furthermore, on the basis of previous studies, it was highlighted that the disabling of susceptibility genes may facilitate durable resistance since the pathogen needs to gain a new function to replace the lost host factor it was exploiting [43].

In woody species, the investigation of S-genes has been performed only for MLO genes in rubber trees [32], poplar trees [33], apple trees, and grapevines [34,44]. Due to the absence of a *C. sativa* genome, highly similar S-genes were selected using the *C. mollissima* v 1.1 genome. Based on the blastn survey, four loci with high similarity were identified in the *C. mollissima* genome and attributed to different subclasses of S-genes [31,45,46,47] due to the presence of specific domains: *mlo1, dmr6*, *dnd1 and pmr4* (Fig. 1, Table 1). As previously observed [31], in the phylogenetic trees, monocot proteins formed a separate clade with respect to those of dicotyledonous species, supporting the hypothesis that an independent evolution occurred for these genes (Fig. 2). Quantitative PCR analysis has been carried out to identify the differential expression of

candidate S-genes in response to *P. cinnamomi* and *C. parasitica* in the stems of a susceptible species, *C. sativa*, and of a tolerant one, *C. crenata*. Lesion analysis and DNA quantification of the pathogen (Fig. 4; Fig. 6) confirmed the higher tolerance level of *C. crenata* in response to both *P. cinnamomi* and *C. parasitica* infection. Our qPCR results highlighted the main upregulation of *pmr4* and *dmr6* in response to infection by both *P. cinnamomi* and *C. parasitica*. As expected, a greater increase in the transcription of these susceptibility genes was observed in the susceptible species *C. sativa*. Remarkably, *pmr4* was strongly expressed at early infection phases of both pathogens in *C. sativa*; in the tolerant *C. crenata*, significant upregulation was observed (Fig. 5; Fig. 7). *Pmr4* codifies for a callose synthase, which is necessary to create a physical barrier to avoid pathogen penetration and is also implicated in plant-triggered immunity suppression. *Pmr4* is thus not only involved in the synthesis of callose, but it also acts as a negative regulator of the salicylic acid pathway [28].
Huibers *et al.* [48] demonstrated that resistance due to the silencing of *Pmr4* is associated with salicylic acid (SA) accumulation rather than with the callose deposition absence. Salicylic acid signalling plays a key role protecting against biotrophic pathogens through the establishment of a hypersensitive response (HR). Saiz-Fernandez *et al.* [49] revealed the increment of SA levels in *P. cinnamomi* inoculated stems, indicating that *P. cinnamomi* activates a defence response similar to that triggered by biotrophic pathogens. Inoculation with both virulent and hypovirulent strains of *C. parasitica* led to SA accumulation in European chestnut plantlets that were grown *in vitro* [50]. Transcriptome analyses carried out

in both *C. dentata* and *C. mollissima* highlighted activation of salicylic-acid-related genes in canker tissues [37].

In chestnut trees, callose deposition around *P. cinnamomi* hyphae was detected early in the infection process; however, it does not seem to play a key role in the associated interactions since the pathogen can reach the vascular cylinder in both susceptible and resistant plant genotypes [51]. This result was validated by transcriptomes analyses of *C. sativa* and *C. crenata,* in which no overexpression of *Callose synthases* after *P. cinnamomi* infection was observed [38].

Based on our results and the literature, we can hypothesize that callose accumulation due to the *pmr4* upregulation in inoculated *C. sativa* lines may not be responsible for controlling *P. cinnamomi* colonization. We suggest that the upregulation of *pmr4* could lead to a negative regulation of the SA pathway that in turn provokes the susceptibility of *C. sativa* to both *P. cinnamomi* and *C. parasitica*. A clear link with SA pathway has emerged even with the other chestnut gene candidate *dmr6* (downy mildew resistance 6). The mutation of *Arabidopsis dmr6* gene, associated with salicylic acid (SA) homeostasis [31], results in the generation of plants that are resistant to bacteria and oomycetes. *Dmr6* is involved in the conversion of salicylic acid (SA) to 2,3-dihydroxybenzoic acid (2,3-DHBA) and negatively regulates the expression of defence genes (PR-1, PR-2, and PR-5) [30].

The expression trend of the *Dmr6* gene in response to *P. cinnamomi* infection turned out to be similar to the profile of *pmr4*. Indeed, *dmr6* was strongly expressed at early infection phases of *P. cinnamomi* in *C. sativa*; in *C. crenata* no significant upregulation was detected (Fig. 5). No

upregulation of *dmr6* in response to *C. parasitica* was highlighted in both plant species (Fig. 7). We can thus hypothesize that *dmr6* upregulation observed in *C. sativa* could negatively regulate defence gene expression, leading to susceptibility to *P. cinnamomi*.

Plants produce a variety of hydrolytic defence enzymes against pathogens, including chitinases, proteases, and also glucanases [52]. The genes coding for several pathogenesis-related (PR) proteins, *Acidic 26 kDa endochitinase* gene (*chi3*) and *Glucan endo-1,3-beta-glucosidase B* gene (*gluB*), were selected in our analysis because they are considered as responsive to SA-dependent signalling [53,54]. *Chi3* and *gluB* are enzymes that cause the lysis of hyphae of various pathogens, resulting in growth inhibition [55,56,57].

In both *C. sativa* and *C. crenata* plants inoculated with *C. parasitica*, *chi3* and *gluB* were significantly upregulated. The transcription of *chi3* was higher in *C. crenata* than in *C. sativa*, presumably as a consequence of the improved defence mechanism against *C. parasitica*. Our results are in agreement with Shain et al. [58], who demonstrated the involvement of *b-1,3-glucanase* and *chitinase* in chestnut species affected by *C. parasitica*. Studies on the role of *chitinase* in blight infection mostly involved *C. sativa* as a model system [50,59]. In both *C. dentata* and *C. mollissima*, transcripts of several compounds expressing *chitinase* accumulated more in canker tissues than healthy stem tissues [37]. In order to obtain chestnut plants with potentially increased resistance/tolerance to chestnut blight, the endogenous Ch3gene encoding a chitinase-like protein was over-expressed in the European chestnut through Agrobacterium-mediated transformation [60].

The emergent CRISPR/Cas9 technology is expected to play a key role in future crop breeding as it allows highly efficient gene editing and generates genetic changes indistinguishable from those arising spontaneously in nature or through conventional breeding [61]. Several examples of edited plants resistant to fungal pathogens have been described [62,63]. For example, genome editing was successfully applied to knock out *mlo* S-genes, leading to Powdery mildew (PM) resistance [44,64,65,66]. *Pmr4* and *dmr6* loss-of-function through CRISPR/Cas reduced the susceptibility to PM in tomato plants [28,67]. In our laboratory we are setting up a CRISPR/Cas9 transformation protocol in *Castanea sativa*. Our future goal will be to perform the functional characterization using the CRISPR/Cas9 approach of the two candidate genes (*dmr6* and *pmr4*), while checking if the two genes may also play a role in the interaction between *C. sativa* and the emergent nut rot and canker agent *Gnomoniopsis castaneae* [68].

Materials and Methods

Identification of chestnut S-genes orthologues

S-gene sequences (S2 File), available in the NCBI database (https://www.ncbi.nlm.nih.gov/, accessed on 31 March 2021), were used as a query in the BLAST+ program (blastn task) against *Castanea mollissima* v1.1 reference genome (https://www.hardwoodgenomics.org/Genome-assembly/1962958/, accessed on 31 March 2021) to find chestnut S-gene orthologs. Hits were filtered using the e-value cut-off of 1e^{-5}.

The domain structures of chestnut S-genes were predicted using Pfam (pfam.xfam.org/, accessed on 31 March 2021) and Uniprot (https://www.uniprot.org/, accessed on 31 March 2021) databases. The graphical gene structure with exons/introns representation was realized using the script accessible at http://wormweb.org/exonintron/ (accessed on 31 March 2021).

Phylogenetic analysis

The alignment of chestnut S-gene coding sequences (*mlo*, *pmr4, dmr6* and *dnd1)* and of known related S-genes in other plant species were generated via multiple sequence alignment using the ClustalW algorithm (http://www.clustal.org/ , accessed on 31 March 2021). All the sequences used for tree construction are accessible in file S3. MEGAX software (https://www.megasoftware.net/ , accessed on 31 March 2021) was used for phylogenetic tree construction, applying maximum likelihood algorithms. To obtain a confidence level for each branch, bootstrap analysis was performed with 1000 iterations. All the phylogenetic trees are available in file S4.

Protein modeling

The Modeller (https://salilab.org/modeller/, accessed on 31 March 2021) software was applied to generate 3-D protein structure models. The Modeller software generates the 3D structure of a given target protein sequence based on its alignment with a known protein structure (templates) [69].

The alignment file of *Arabidopsis thaliana* (Q9FL0) and *C. mollissima* protein sequences was obtained using the Emboss Needle online tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle/, accessed on 31 March 2021). The 3D model was developed with an automodel class using the 3D *A. thaliana* model and the alignment file. Ccp4mg software (www.ccp4.ac.uk/, accessed on 31 March 2021) was used for protein 3D structure visualization, which is useful for studies of catalytic and regulatory domain conservation/divergence.

Pathogen's inoculation and samples collection

P. cinnamomi (ID_C4) and *C. parasitica* (ID_5183 L2d) isolates used in the experiment were originally isolated from symptomatic *C. sativa* trees in Piedmont and Aosta Valley, northwestern Italy, respectively, and preserved in the plant pathogen culture collection at DISAFA (University of Turin). Isolates were subcultured in Potato Dextrose Agar (PDA) before inoculations. The inoculation trial was carried out on *Castanea sativa* and *Castanea crenata* plants (1-year-old) grafted on *C. sativa* and *C. sativa* x *C. crenata* rootstocks, respectively. Plants were grown in pots under greenhouse conditions. The identity of the plant material was preliminary checked through marker analysis using 10 SSR loci from Marinoni et al. [70] Plants were 80-100 cm high and 0.9-1.5 cm in diameter at the collar.

Plants were inoculated 20 cm above the collar by placing a colonized plug of PDA (0.5 cm diameter) in a slit, superficially cleaned with 70% ethanol, obtained by excising a small portion of the bark with a sterile scalpel according to the methods described by Zampieri et al. [71]. After the

inoculation process, the inoculation point was wrapped with parafilm to prevent tissue dryness and external contamination [71,72,73]. As negative controls (0 hpi), plants were wounded in the same way but inoculated with a sterile PDA plug. Plants were incubated in greenhouse conditions at 28 ± 2 °C with a 16-h photoperiod.

S-genes are genes related to plant-pathogen interaction and are expressed during the first step of inoculation, before symptoms manifest. The time points used in our analysis were selected using the susceptible *C. sativa* as a reference on the basis of the onset of evident symptoms, i.e., bark necrosis and leaf dryness/browning, on *C. sativa* reference plants inoculated with each pathogen. The selected time points were 5 days after inoculation for *P. cinnamomi* and 7 days after inoculation for *C. parasitica*. Three biological replicates for seven experimental time points, including 0 hpi (control), were tested both for *Castanea sativa* and *Castanea crenata* and for the two pathogens (84 plants in total).

Plant material was harvested at 0, 3, 6, 12, 24, 48, 72 hours after *P. cinnamomi* inoculation. For the *C. parasitica* experiment, material was collected at 0, 12, 24, 48, 72, 96, and 120 hours after inoculation. For each time point, two disks of the stem were cut 0.5 cm above and below the wound. Bark was then removed to reduce the polyphenol contamination of RNA, and samples were frozen in liquid nitrogen to preserve RNA integrity. All samples were stored at -80 °C before RNA extraction.

RNA extraction and Real-Time qPCR quantification

RNA was extracted from both inoculated and control (0 hpi) samples. A total of 0.1 g of frozen tissue was manually ground into a fine powder and

liquid nitrogen was added. RNA was extracted using Spectrum Plant Total RNA Kit (Sigma-Aldrich) following the manufacturer's protocol. Extracted RNA was treated with DNase I (Thermo Fisher Scientific) following the manufacturer's instructions.

RNA was quantified by a NanoDrop spectrophotometer (Thermo Scientific, Hudson, NH, USA). cDNA was synthesized from 2 µg RNA using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, USA). Primer sequences for candidate S-genes were designed using the Primer3 online tool (https://primer3.ut.ee/, accessed on 31 March 2021) and are available in file S5. All primers were in silico tested through Primer-BLAST program (https://www.ncbi.nlm.nih.gov/tools/primer-blast/, accessed on 31 March 2021)

Chi3 and *gluB* genes, coding for several pathogenesis-related (PR) proteins, were also analysed in order to observe their role in the defence response of chestnut trees [74]. Transcript abundance was quantified in three biological replicates by the StepOnePlus Real-Time PCR System (Applied Biosystems). Real-Time qPCR was performed using the Power SYBR® Green Master Mix added with Bovine Serum Albumin (BSA) to reduce the action of PCR inhibitors. The amplification protocol included an initial denaturation step at 95 °C for 5min, followed by 40 cycles of 95 °C/5s and 60 °C/1min. Data were quantified using the $2^{-\Delta\Delta Ct}$ method based on Ct values of candidate genes and actin (as a housekeeping gene) [38]. IBM SPSS statistical software was used to carry out a one-way analysis of variance test (ANOVA). Each value represented the mean of three biological replicates, which were compared using Tukey's HSD Test ($P \leq 0.05$).

Pathogen quantification

Samples inoculated with *P. cinnamomi* (24, 48, 72 hours) and *C. parasitica* (72, 96, 120 hours) were used for DNA extraction and pathogen quantification. Plants were debarked and the necrotic area in the cambium layer was measured using ImageJ v. 1.8.0 software.

The DNA extraction was performed using an E.Z.N.A.® Stool DNA Kit following the manufacturer's protocol. Standard curves were prepared for the quantification of DNAs by qPCR using primers designed as follows: *7-actin* for chestnut DNA, the *ypt* gene for *P. cinnamomi* DNA, and the *mf1* species-specific gene for *C. parasitica* DNA [75] (S5 file). All the inoculated and control (0 hpi) samples were analysed through real-time qPCR both with pathogens genes (*ypt* and *mf1*) and with *7-actin*. The results, normalized by standard curves, were used for the calculation of the ratio of DNA fungus/plant DNA. Real-Time qPCR was performed using the experimental conditions previously described: initial denaturation step at 95 °C for 5min, followed by 40 cycles at 95 °C/5s and 60 °C/1min. Data were quantified through the $2^{-\Delta\Delta Ct}$ method based on Ct values of pathogen genes and *actin-7* as a housekeeping gene [38]

IBM SPSS statistical software was applied to perform a one-way analysis of variance test (ANOVA). Each value represented the mean of three biological replicates compared using Tukey's HSD Test ($P \leq 0.05$).

Supplementary materials

The following are available online at https://www.mdpi.com/article/10.3390/plants10050913/s1. S1 File: experiments pipeline, S2 File: S-gene coding sequences available in the NCBI database (https://www.ncbi.nlm.nih.gov/, accessed on 31 March 2021) used for S-gene detection in the *C. mollissima* genome, S3 File: S-gene coding sequences used for tree construction, S4 File: phylogenetic trees, S5 File: primer sequences

Author Contributions

Conceptualization, R.B., A.M., P.G., and D.T.M.; methodology, A.M. and V.P.; software, A.M., V.P., and E.C.G; validation, A.M. and V.P.; formal analysis, A.M. and V.P; investigation, A.M. and V.P; resources R.B. and P.G.; data curation, A.M. and V.P.; writing—original draft preparation, A.M and V.P.; writing—review and editing, R.B, P.G., D.T.M., A.M., and V.P.; visualization, A.M. and V.P.; supervision, A.M., R.B., P.G., and D.T.M.; project administration, R.B.; funding acquisition, R.B. All authors have read and agreed to the published version of the manuscript.

Funding

The research was funded by Fondazione CRT. This research was also partially funded by the "MONGEFITOFOR" Project Interreg Cooperation Program Va ITA-CH 2014/2020.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Torello Marinoni, D.; Nishio, S.; Valentini, N.; Shirasawa, K.; Acquadro, A.; Portis, E.; Alma, A.; Akkak, A.; Pavese, V.; Cavalet-Giorsa, E.; Botta, R. Development of High-Density Genetic Linkage Maps and Identification of Loci for Chestnut Gall Wasp Resistance in *Castanea* spp. Plants **2020**, 9, 1048.
2. Conedera, M.; Krebs, P.; Tinner, W.; Pradella, M.; Torriani, D. The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. *Veg. Hist. Archaeobot.* **2004a**, *13*, 161-179, **doi: 10.1007/s00334-004-0038-7.**
3. Müllerová, J.; Hédl, R.; Szabó, P. Coppice abandonment and its implications for species diversity in forest vegetation. *For. Ecol. Manag.* **2015**, *343*, 88-100.
4. Zentmyer, G.A. *Phytophthora cinnamomi* and The Diseases it Causes. *The American Phytopathological Society.* St. Paul, MN, USA. **1980**, Monograph. *10*, 96.
5. Cristinzio, G.; Grassi, G. Valutazione di resistenza a *Phytophthora cambivora* e *Phytophthora cinnamomi* in cultivar di *Castanea sativa. Monti e Boschi* **1993**, *1*, 54-58.
6. Serrazina, S.; Machado, H.; Costa, R.; Duque, P.; Malhó, R. Expression of *Castanea crenata* Allene Oxide Synthase in *Arabidopsis* Improves the Defense to *Phytophthora cinnamomi.* bioRxiv; **2020**, doi: 10.1101/2020.10.09.333781.

7. Maurel, M.; Robin, C.; Simonneau, T.; Loustau, D.; Dreyer, E.; Desprez-Loustau M.-L. Stomatal conductance and root-to-shoot signalling in chestnut saplings exposed to *Phytophthora cinnamomi* or partial soil drying. *Functional Plant Biol* **2004**, *31*, 41-51, doi: 10.1071/FP03133.
8. Vannini, A.; Vettraino, A. M. Ink disease in chestnuts: impact on the European chestnut. *Forest Snow and Landscape Research* **2002**, *76*, 345-350.
9. Robin, C.; Smith, I.; Hansen, E.M. *Phythophthora cinnamomi*. *Forest Phytophthoras* **2012**, *2*,1 doi: 10.5399/osu/fp.2.1.3041.
10. Gonthier, P.; Robin, C. Diseases. Chapter 11. In *The Chestnut Handbook* 2020: Crop and Forest Management, Beccaro, G.; Alma, A.; Bounous, G.; Gomes-Laranjo, J. Eds; CRC Press, Taylor & Francis Group, Boca Raton, FL, USA. pp. 297-316. ISBN: 9781138334021.
11. Hayden, K.J.; Garbelotto, M.; Dodd, R.; Wright, J.W. Scaling up from greenhouse resistance to fitness in the field for a host of an emerging forest disease. *Evol App*. **2013**, *6*, 970-82, doi: 10.1111/eva.12080.
12. Bounous, G. Il castagno: risorsa multifunzionale in Italia e nel mondo. *Edagricole* **2014**, 312.
13. Rigling, D; Prospero, S. *Cryphonectria parasitica*, the causal agent of chestnut blight: invasion history, population biology and disease control. *Mol Plant Pathol.* **2018**,*1*, 7-20, doi: 10.1111/mpp.12542.

14. Lione, G.; Giordano, L.; Turina, M.; Gonthier, P. Hail-Induced Infections of the Chestnut Blight Pathogen *Cryphonectria parasitica* Depend on Wound Size and May Lead to Severe Diebacks. *Phytopathology* **2020**, *110*(*7*),1280-1293, doi: 10.1094/PHYTO-01-20-0006-R.
15. Trapiello, E.; González-Varela, G.; González, A.J. Chestnut Blight Control by Agrochemicals in *Castanea sativa* under Managed Conditions. *J Plant Dis Prot* **2015**, *122*, 120-124, doi: **10.1007/BF03356540.**
16. Westbrook, J. W.; Zhang, Q.; Mandel, M. K.; Jenkins, E. V.; Barth, L. E.; Jenkins, J. W.; Grimwood, J.; Schmutz, J.; Holliday, J. A. Genomic selection analyses reveal tradeoff between blight tolerance and genome inheritance from American chestnut (*Castenea dentata*) in [*C. dentata* x *C. mollissima*] x *C. dentata* backcross populations. BioRxiv. **2019**, doi: **10.1101/690693.**
17. Kubisiak, T.L.; Hebard, F.V.; Nelson, C.D.; Zhang, J.; Bernatzky, R.; Huang, H.; Anagnostakis, S.L.; Doudrick, R.L. Molecular Mapping of Resistance to Blight in an Interspecific Cross in the genus *Castanea*. *Phytopathology* **1997**, *87*, 751-759.
18. Kubisiak, T.L.; Nelson, C.D.; Staton, M.E.; Zhebentyayeva, T.; Smith, C.; Olukolu, B.A.; Fang, G.-C.; Hebard, F.V.; Anagnostakis, S.; Wheeler, N.; *et al.* A transcriptome-based genetic map of Chinese chestnut (*Castanea mollissima*) and identification of regions of segmental homology with peach (*Prunus persica*). *Tree Genet. Genomes* **2012**, *9*, 557-571.

19. Zhang, B.; Oakes, A.D.; Newhouse, A.E.; Baier, K.M.; Maynard, C.A.; Powell, W.A. A threshold level of oxalate oxidase transgene expression reduces *Cryphonectria parasitica*-induced necrosis in a transgenic American chestnut (*Castanea dentata*) leaf bioassay. *Transgenic Res.* **2013**, *22(5)*, 973-82, doi: 10.1007/s11248-013-9708-5.
20. Marcolin, E.; Manetti, M. C.; Pelleri, F.; Conedera, M.; Pezzatti, G. B.; Lingua, E.; Pividori, M. Seed regeneration of sweet chestnut (*Castanea sativa* Miller) under different coppicing approaches. *Forest Ecology and Management* **2020**, *472*, 118273, doi: **10.1016/j.foreco.2020.118273.**
21. Aebi, A.; Schönrogge, K.; Melika, G.; Alma, A.; Bosio, G.; Quacchia, A.; Picciau, L.; Abe, Y.; Moriya, S.; Yara, K.; Seljak, G.; Stone, G.N. Parasitoid recruitment to the globally invasive chestnut gall wasp *Dryocosmus kuriphilus*. In: Ozaki K, Yukwa J, Ohgushi T, Price PW Eds Ecology and evolution of galling arthropods and their associates. *Springer-Verlag*, Tokyo **2006**, 103-121.
22. Van Schie, C.C.; Takken, F.L. Susceptibility genes 101: how to be a good host. Annu. Rev. *Phytopathol.* **2014**, *52*, 551-581
23. *Zaidi, S.S.*; *Mukhtar, M.S.; Mansoor, S.* Genome editing: targeting susceptibility genes for plant disease resistance. *Trends Biotechnol.* **2018,** *36*, 898-906.
24. Pavan, S.; Jacobsen, E.; Visser, R.G.; Bai, Y. Loss of susceptibility as a novel breeding strategy for durable and broad-spectrum resistance. *Mol. Breed.* **2010**, *25*, 1-12, doi: 10.1007/s11032-009-9323-6

25. Devoto, A.; Hartmann, H.A.; Piffanelli, P.; Elliott, C.; Simmons, C.; Taramino, G.; Goh, C.S.; Cohen, F.E.; Emerson, B.C.; Schulze-Lefert, P.; Panstruga, R. Molecular phylogeny and evolution of the plant-specific seven-transmembrane MLO family. *J Mol Evol.* **2003**, 56(1),77-88, doi: 10.1007/s00239-002-2382-5.
26. Büschges, R.; Hollricher, K.; Panstruga, R.; Simons, G.; Wolter, M.; Frijters, A.; van Daelen, R.; van der Lee, T.; Diergaarde, P.; Groenendijk, J. *et al.* The barley Mlo gene: a novel control element of plant pathogen resistance. *Cell.* **1997,** *88*, 695-705.
27. Appiano, M.; Pavan, S.; Catalano, D.; Zheng, Z.; Bracuto, V.; Lotti, C.; Visser, R.G.; Ricciardi, L.; Bai, Y. Identification of candidate MLO powdery mildew susceptibility genes in cultivated *Solanaceae* and functional characterization of tobacco NtMLO1. *Transgenic Res.* **2015**, *24(5),* 847-58, doi: 10.1007/s11248-015-9878-4.
28. Santillán Martínez, M.I., Bracuto, V., Koseoglou, E., Appiano, M., Jacobsen, E., Visser, R.G.F., Wolters, A.A., Bai, Y. CRISPR/Cas9-targeted mutagenesis of the tomato susceptibility gene *PMR4* for resistance against *powdery mildew*. *BMC Plant Biol* **2020**, *20*, 284, doi: **10.1186/s12870-020-02497-y.**
29. Sun, K.; A-MA, W.; Loonen, A.E.; Huibers, R.P.; van der Vlugt, R.; Goverse, A.; Jacobsen, E.; Visser, R.G.; Bai, Y. Down-regulation of *Arabidopsis* DND1 orthologs in potato and tomato leads to broad-spectrum resistance to late blight and powdery mildew. *Transgenic Res.* **2016**, *25(2),* 123-38.
30. Van Damme, M.; Huibers, R. P.; Elberse, J.; Van Den Ackerveken, G. *Arabidopsis* DMR6 encodes a putative 2OG-Fe(II) oxygenase

that is defense-associated but required for susceptibility to *downy mildew*. *Plant J.* **2008**, *54*, 785-793, doi: 10.1111/j.1365-313X.2008.03427.x.

31. Zeilmaker, T; Ludwig, N.R.; Elberse, J.; Seidl, M.F.; Berke, L.; Van Doorn, A.; Schuurink, R.C.; Snel, B; Van den Ackerveken, G. '*Downy mildew resistant 6* and DMR6-like oxygenase 1 are partially redundant but distinct suppressors of immunity in Arabidopsis', *Plant Journal.* **2015**, *81(2),* 210-222, doi: 10.1111/tpj.12719.
32. Liyanage, K.K.; Khan, S.; Herath, V.; Brooks, S.; Mortimer, P.E.; Nadir, S.; Hyde, K.D.; Xu, J. Genome Wide Identification of the MLO Gene Family Associated with Powdery Mildew Resistance in Rubber Trees (*Hevea brasiliensis*). *Tropical Plant Biol.* **2020**, *13*, 331-342, doi: 10.1007/s12042-020-09262-3.
33. Filiz, E.; Vatansever, R. Genome-wide identification of mildew resistance locus O (*MLO*) genes in tree model poplar (*Populus trichocarpa*): powdery mildew management in woody plants. *Eur J Plant Pathol.* **2018**, *152*, 95-109, doi: **10.1007/s10658-018-1454-3.**
34. Pessina, S.; Lenzi, L.; Perazzolli, M.; Campa, M.; Dalla Costa, L.; Urso, S.; Valè, G.; Salamini, F.; Velasco, R.; Malnoy, M. Knockdown of *MLO* genes reduces susceptibility to powdery mildew in grapevine. *Hortic Res* **2016**, *3*, 16016, doi: 10.1038/hortres.2016.16.
35. Jung, T.; Pérez-Sierra, A.; Durán, A.; Jung, M.H.; Balci, Y.; Scanu, B. Canker and decline diseases caused by soil- and airborne Phytophthora species in forests and woodlands. *Pers. Mol. Phylogeny Evol. Fungi* **2018**, *40*, 182-220.

36. Pereira-Lorenzo, S.; Ballester, A.; Corredoira, E.; Bounous, G.; Botta, R.; Beccaro, G. L.; Kubisiak, T. L.; Conedera, M.; Krebs, P.; Yamamoto, T.; *et al.*Chestnut. In: M. L. Badenes and D. H. Byrne (eds.). *Fruit Breeding*. Springer Science + Business Media, New York, pp 729-769.
37. Barakat, A.; Staton, M.; Cheng, C.H.; Park, J.; Yassin, N.B.M.; Ficklin, S.; Yeh, C.C.; Hebard, F.; Baier, K.; Powell, W. *et al.,* Chestnut resistance to the blight disease: insights from transcriptome analysis. *BMC Plant Biol.* **2012**, *12,* 38.
38. Serrazina, S.; Santos, C.; Machado, H.; Pesquita, C.; Vicentini, R.; Pais, M.S.; Sebastiana, M.; Costa, R. *Castanea* root transcriptome in response to P*hytophthora cinnamomi* challenge. *Tree Genet. Genomes* **2015**, 11, 19.
39. Santos, C.; Machado, H.; Correia, I; Gomes, F.; Gomes-Laranjo, J.; Costa, R. Phenotyping *Castanea* hybrids for *Phytophthora cinnamomi* resistance. *Plant Pathology* **2015**, *64*, 901-910, doi: 10.1111/ ppa.12313.
40. Santos, C.; Duarte, S.; Tedesco, S.; Fevereiro, P.; Costa, R.L. Expression profiling of *Castanea* genes during resistant and susceptible interactions with the oomycete pathogen *Phytophthora cinnamomi* reveal possible mechanisms of immunity. *Frontiers in Plant Science* **2017a**, *8*, 515, doi: 10.3389/ fpls.2017.00515.
41. Santos, C.; Nelson, C.D.; Zhebentyayeva, T.; Machado, H; Gomes-Laranjo, J.; Costa, R.L. First interspecific genetic linkage map for Castanea sativa x Castanea crenata revealed QTLs for resistance

to Phytophthora cinnamomi. *PLoS ONE* **2017b**, *12*, 9. e0184381, doi: 10.1371/journal.pone.0184381.

42. Barakat, A.; DiLoreto, D.S.; Zhang, Y.; Smith, C.; Baier, K.; Powell, W.A.; Wheeler, N.; Sederoff, R.; Carlson, J.E. Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection. *BMC Plant Biol.* **2009,** *9*, 51, doi: **10.1186/1471-2229-9-51.**
43. van Esse, H.P.; Reuber, T.L.; van der Does, D. Genetic modification to improve disease resistance in crops. *New Phytol.* **2020**, *225(1),* 70-86, doi: 10.1111/nph.15967.
44. Wan, D.Y.; Guo, Y.; Cheng, Y.; Hu, Y.; Xiao, S.; Wang, Y.; Wen, Y.Q. CRISPR/Cas9-mediated mutagenesis of *VvMLO3* results in enhanced resistance to *powdery mildew* in grapevine (*Vitis vinifera*). *Hortic. Res.* **2020**, *7*, 116, doi: **10.1038/s41438-020-0339-8.**
45. Clough, S.J.; Fengler, K.A.; Yu, I.C.; Lippok, B.; Smith, R.K.; Bent, A.F. The Arabidopsis dnd1 "defense, no death" gene encodes a mutated cyclic nucleotide-gated ion channel. *Proc. Natl. Acad. Sci.* USA **2000**, *97*, 9323-9328.
46. Nishimura, M.T.; Stein, M.; Hou, B.H.; Vogel, J.P.; Edwards, H.; Somerville, S.C. Loss of a callose synthase results in salicylic acid-dependent disease resistance. *Science* **2003,** *301*, 969-972.
47. Zhang, Y.; Zhao, L.; Zhao, J.; Li, Y.; Wang, j.; Guo, R.; Gan, S; Liu, C.J.; Zhang, K. S5H/DMR6 encodes a salicylic acid 5-hydroxylase that fine-tunes salicylic acid homeostasis. *Plant Physiol.* **2017**, *175*, 1082-1093.

48. Huibers, R.P.; Loonen, A.E.; Gao, D.; Van den Ackerveken, G.; Visser, R.G.; Bai, Y. Powdery mildew resistance in tomato by impairment of SlPMR4 and SlDMR1. *PLoS One*. **2013**, *8 (6)* e67467, doi: 10.1371/journal.pone.0067467.
49. Saiz-Fernández, I.; Milenković, I.; Berka, M.; Černý, M.; Tomšovský, M.; Brzobohatý, B.; Kerchev, P. Integrated Proteomic and Metabolomic Profiling of *Phytophthora cinnamomi* Attack on Sweet Chestnut (*Castanea sativa*) Reveals Distinct Molecular Reprogramming Proximal to the Infection Site and Away from It. Int. J. Mol. Sci. **2020**, 21, 8525, doi: **10.3390/ijms21228525.**
50. Lovat, C.-A.; Donnelly, D. J. Mechanisms and metabolomics of the host-pathogen interactions between Chestnut (*Castanea* species) and Chestnut blight (*Cryphonectria parasitica*). *For. Pathol.* **2019**, 49:e12562, doi: **10.1111/efp.12562.**
51. Fernandes, P.; Machado,H.; Silva, M.D.C.; Costa, R.L. A Histopathological Study Reveals New Insights Into Responses of Chestnut (*Castanea* spp.) to Root Infection by *Phytophthora cinnamomi*. *Phytopathology* **2021**, doi: 10.1094/PHYTO-04-20-0115-R.
52. Veluthakkal, R.; Dasgupta, M.G. Pathogenesis-related genes and proteins in forest tree species. *Trees* **2010**, *24*, 993-1006, doi: **10.1007/s00468-010-0489-7.**
53. Puthoff, D. P.; Holzer, F. M.; Perring, T. M.; Walling, L. L. Tomato Pathogenesis-related Protein Genes are Expressed in Response to *Trialeurodes vaporariorum* and *Bemisia tabaci* Biotype B Feeding. *J Chem Ecol.* **2010**, *36*, 1271-1285, doi: **10.1007/s10886-010-9868-1.**

54. Molinari, S.; Fanelli, E.; Leonetti, P. Expression of tomato salicylic acid (SA)-responsive pathogenesis-related genes in Mi-1-mediated and SA-induced resistance to root-knot nematodes. *Mol Plant Pathol.* **2014**, *15(3),* 255-64, doi: 10.1111/mpp.12085.
55. Broglie, K.; Chet, I.; Holliday, M.; Cressman, R.; Biddle, P.; Knowlton, S.; Mauvais, C.J.; Broglie, R. Transgenic plants with enhanced resistance to the fungal pathogen, *Rhizoctonia solani*. *Science* **1991**, *254*, 1194-1197.
56. Mauch, F.; Staehelin, L.A. Functional Implications of the Subcellular Localization of Ethylene-Induced Chitinase and [beta]-1,3-Glucanase in Bean Leaves. *Plant Cell.* **1989**,*1(4),* 447-457, doi: 10.1105/tpc.1.4.447.
57. Schlumbaum, A.; Mauch, F.; Vögeli, U.; Boller, T. Plant chitinases are potent inhibitors of fungal growth. *Nature* **1986,** *324*, 365-367, doi: **10.1038/324365a0.**
58. Shain, L.; Miller, J.B.; Spalding, R.J. Responses of American and Chinese Chestnut to *Cryphonectria parasitica* and Ethylene. In "*Proceedings of the International Chestnut Conference".* Double M.L.; MaCDonald W.L. eds. pp 97-101.
59. Vannini, A.; Caruso, C.; Leonardi, L.; Rugini, E., Chiarot, E.; Caporale, C., Buonocore, V. Antifungal properties of chitinases from *Castanea sativa* against hypovirulent and virulent strains of the chestnut blight fungus *Cryphonectria parasitica*. *Physiol Mol, Plant Pathol.* **1999**, *55*, 29-35.
60. Corredoira, E.; San José, M.C.; Vieitez, A.M.; Allona, I.; Aragoncillo, C.; Ballester, A. *Agrobacterium*-mediated transformation of

European chestnut somatic embryos with a *Castanea sativa* (Mill.) endochitinase gene. *New Forests* **2016**, *47*, 669-684, doi: **10.1007/s11056-016-9537-5.**

61. Zhu, H.; Li, C; Gao, C. Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat Rev Mol Cell Biol* **2020**, 21, 661-677, doi: **10.1038/s41580-020-00288-9.**
62. Schenke, D.; Cai, D. Applications of CRISPR/Cas to Improve Crop Disease Resistance: Beyond Inactivation of Susceptibility Factors. *iScience* **2020**, *23(9),* 101478, doi: 10.1016/j.isci.2020.101478.
63. Tyagi, S.; Kesiraju, K.; Saakre, M.; Rathinam, M.; Raman, V.; Pattanayak, D.; Sreevathsa, R. Genome editing for resistance to insect pests: an emerging tool for crop improvement. *ACS Omega* **2020**, *5*, 20674-20683, doi: 10.1021/acsomega.0c01435.
64. Wang, Y.; Cheng, X.; Shan, Q.; Zhang, Y.; Liu, J.; Gao, C.; Qiu, J.L. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol.* **2014**, *32(9)*, 947-51, doi: 10.1038/nbt.2969.
65. Malnoy, M.; Viola, R.; Jung, M.H.; et al. DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. *Front Plant Sci*. **2016**, *7*, 1904, doi:10.3389/fpls.2016.01904.
66. Nekrasov, V.; Wang, C.; Win, J.; Lanz, C.; Weigel, D.; Kamoun, S. Rapid generation of a transgene-free *powdery mildew* resistant tomato by genome deletion. *Sci Rep*. **2017**, *7(1),* 482, doi:10.1038/s41598-017-00578-x.

67. de Toledo Thomazella, D. P.; Brail, Q.; Dahlbeck, D.; Staskawicz, B. J. CRISPR-Cas9 mediated mutagenesis of a DMR6 ortholog in tomato confers broad-spectrum disease resistance. bioRxiv. **2016**, 1-23, doi: 10.1101/064824.
68. Lione, G.; Danti, R.; Fernandez-Conradi, P.; Ferreira-Cardoso, J.V.; Lefort, F.; Marques, G.; Meyer, J.B., Prospero, S.; Radòcz, L.; Robin, C. *et al.* The emerging pathogen of chestnut *Gnomoniopsis castaneae*: the challenge posed by a versatile fungus. *European Journal of Plant Pathology* **2019,** *153*, 671-685.
69. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* **2016,** *54*, 5.6.1-5.6.37, doi: 10.1002/cpbi.3.
70. Marinoni, D.; Akkak, A., Bounous, G.; Edwards, K. J; Botta, R. Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Molecular Breeding* **2003**, *11*, 127-136, doi: **10.1023/A:1022456013692.**
71. Zampieri, E.; Giordano, L.; Lione, G.; Vizzini, A.; Sillo, F.; Balestrini, R.; Gonthier, P. A nonnative and a native fungal plant pathogen similarly stimulate ectomycorrhizal development but are perceived differently by a fungal symbiont. *New Phytol.* **2017,** *213*, 1836-1849.
72. Robin, C.; Morel, O.; Vettraino, A.M.; Perlerou, C.; Diamondis, S.; Vannini, A. Genetic variation in susceptibility to Phytophthora cambivora in European chestnut (*Castanea sativa*). *Forest Ecology and Management* **2006,** *226*, 199-207.

73. Dinis, L.-T.; Peixoto, F.; Zhang, C.; Martins, L.; Costa, R.; Gomes-Laranjo, J. Physiological and biochemical changes in resistant and sensitive chestnut (*Castanea*) plantlets after inoculation with *Phytophthora cinnamomi*. *Physiological and Molecular Plant Pathology* **2011**, *75*, 146-56.
74. Wu, C.T.; Leubner-Metzger, G.; Meins, F. Class I beta- 1,3-glucanase and chitinase are expressed in the micropylar endosperm of tomato seeds prior to radicle emergence. *Plant Physiol.* **2001**,*126*,1299-1313.
75. Popov, A.P.; Tsvetkov, I.L.; Belov, A.A.; Konichev, A. S.; Ivanushkina, N. E.; Kochkina, G. A.; Ozerskaya, S. M. Molecular genetic identification of the phytopathogenic fungus *Cryphonectria parasitica*. *Microbiology* **2010**, *79*, 223-228, doi: 10.1134/S0026261710020141.

Chapter III

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 August 2021
doi: 10.3389/fpls.2021.728516

First report of CRISPR/Cas9 gene editing in *Castanea sativa* Mill.

Vera Pavese [1], Andrea Moglia [1*], Elena Corredoira[2], Mª Teresa Martínez[2], Daniela Torello Marinoni[1], Roberto Botta[1]

[1.] Dipartimento di Scienze Agrarie, Forestali e Alimentari (DISAFA), Università degli Studi di Torino, Largo Paolo Braccini 2, Grugliasco, 10095 Torino, Italy.
[2.] Instituto de Investigaciones Agrobiológicas de Galicia (IIAG)-CSIC, Avd Vigo s/n, 15705 Santiago de Compostela, Spain.

*Correspondance:
Andrea Moglia
andrea.moglia@unito.it

Abstract

CRISPR/Cas9 has emerged as the most important tool for genome engineering due to its simplicity, design flexibility, and high efficiency. This technology makes it possible to induce point mutations in one or some target sequences simultaneously, as well as to introduce new genetic variants by homology-directed recombination. However, this approach

remains largely unexplored in forest species. Here, we report the first example on CRISPR/Cas9-mediated gene editing in *Castanea* genus. As a proof of concept, we targeted the gene encoding *phytoene desaturase* (*pds*), whose mutation disrupts chlorophyll biosynthesis allowing for visual assessment of knockout efficiency. Globular and early torpedo-stage somatic embryos of *C. sativa* (European chestnut) were co-cultured for 5 days with a CRISPR/Cas9 construct targeting two conserved gene regions of *pds* and subsequently cultured on selection medium with kanamycin. After eight weeks of subculture on selection medium, 4 kanamycin resistant embryogenetic lines were isolated. Genotyping of these lines through target Sanger sequencing of amplicons revealed the successful gene editing. Cotyledonary somatic embryos were maturated on maltose 3% and cold-stored at 4ºC for 2 months. Subsequently, embryos were subjected to the germination process to produce albino plants. This work opens the way to the use of the CRISPR/Cas9 system in European chestnut for biotechnological applications.

Keywords: European chestnut, *Agrobacterium*-mediated transformation, *phytoene desaturase,* somatic embryos, targeted mutagenesis, gene knockout.

Introduction

European chestnut or sweet chestnut (*Castanea sativa* Mill.) is a worldwide widely distributed tree species with an important economic role in Spain and Italy. It is highly appreciated for both timber and fruit production (Conedera *et al.,* 2004; Conedera *et al.,* 2016). In addition to its main productive role, European chestnut may also play a key role in wildlife and landscape conservation, rural tourism, recreation and protection from erosion (Merkle *et al.,* 2020). However, European chestnut populations are seriously threatened for two severe diseases: ink disease mainly caused by the oomycete *Phytophthora cinnamomi* and chestnut blight provoked by the fungus *Cryphonectria parasitica*. During the first third of the 20th, the crossings between European chestnut and Asian tolerant species (*C. crenata* and *C. mollissima*) were the only valid option to deal with the ink disease. Although the hybrids obtained showed tolerance to *P. cinnamomi*, their nut quality was lower when compared to the European chestnut cultivars. The spread of hypovirulence has reduced the impact of canker blight in Europe but the disease is still a threat in orchards and young plantings, since extensive breeding of *C. sativa* for resistance to *C. parasitica* has not been carried out yet.

Biotechnological methods, such as genetic transformation, can represent an interesting alternative to traditional breeding of chestnut and could contribute to overcome the disease issue, a major limiting factor for the expansion of modern orchard planting. A prerequisite for transformation is the availability of an adequate *in vitro* plant regeneration procedure. Somatic embryos are considered the best explant to be used as a target in

the transformation systems (Corredoira *et al.,* 2019), due to the higher transformation rate than other regeneration methods and to the reduced number of escapes and chimeric plants. In recent years, several procedures have been reported for the establishment of European chestnut somatic embryos lines from zygotic embryos, and from shoot apices and leaves isolated from axillary shoot cultures (Corredoira *et al.,* 2006; Merkle *et al.,* 2020). Using these embryogenic systems, efficient protocols of genetic transformation were set up; chestnuts were transformed introducing both marker genes (Corredoira *et al.,* 2004) and genes coding for pathogenesis related proteins such as thaumatin-like proteins and chitinases, in order to confer tolerance against ink and blight diseases, respectively (Corredoira *et al.,* 2012; 2016).

Targeted genome editing allows the introduction of precise modifications directly into a commercial cultivar, offering a viable alternative to traditional breeding methods. The 2020 Nobel Prize CRISPR/Cas9 technology (www.nobelprize.org/prizes/chemistry/2020/) has shown high efficiency in the knock out, insertion and correction of genes, and has sparked great enthusiasm in the scientific community being expected to play a key role in future efforts to improve crop traits. This technology makes it possible to induce point mutations in one or some target sequences simultaneously, as well as to introduce new genetic variants by homology directed recombination (HDR) or to target and modify the transcription. In addition, CRISPR/Cas9 is inexpensive, simple and highly flexible allowing the rapid target plant genome editing (Walawage *et al.,* 2019). The development of new genome editing technologies in plant breeding fostered a growing interest for *in vitro* culture and regeneration

protocols, which represent a major bottleneck in the application of these techniques in many plant species of agricultural and industrial interest.

So far, reports of genome editing in tree species are still limited (Bewg *et al.,* 2018). Nowadays, gene editing has been only reported in a few fruit tree species such as coffee (Breitler *et al.,* 2018), apple (Nishitani *et al.,* 2016), grape (Wang *et al.,* 2018), cocoa (Fister *et al.,* 2018) and walnut (Walawage *et al.,* 2019), while in forest species, gene editing has been only achieved in poplar (Fan *et al.,* 2015).

Phytoene desaturase (*pds*) gene plays a central role in chlorophyll biosynthesis and is considered a popular marker for CRISPR/Cas9 editing since its knockout leads to an albino phenotype (Qin *et al.,* 2007). For this reason, *pds* gene is used as endogenous reporter gene for proof-of-concept gene editing in plants (Odipio *et al.,* 2017; Shan *et al.,* 2018; Bernard *et al.,* 2019; Charrier *et al.,* 2019; Ma *et al.,* 2019; Wilson *et al.,* 2019).

In this study we reported, for the first time, the establishment of a CRISPR/Cas9-based transformation protocol in *Castanea sativa* using somatic embryos. Our results demonstrated that genome editing through CRISPR/Cas9 can be efficiently applied for chestnut genome modification.

Materials and Methods

Plant material

The embryogenic line CI-3 initiated from zygotic embryos of *Castanea sativa* Mill. (Corredoira *et al.,* 2006) was used for gene editing experiments. This embryogenic line was maintained by secondary embryogenesis with sequential subcultures at 5/6 weeks intervals onto a multiplication medium consisting of MS (Murashige and Skoog, 1962) mineral salts (half-strength macronutrients; ½ MS) and vitamins, 3 mM glutamine, 0.1 mg/L benzyladenine (BA), 0.1 mg/L 1-naphthaleneacetic acid (NAA), 3% sucrose (w/v), and 0.7% Sigma-agar (w/v). The pH was adjusted to 5.6-5.7, and the medium was autoclaved at 115ºC for 20 min. Cultures were incubated under a 16-h photoperiod (provided by cool-white fluorescent lamps at a photon flux density of 50-60 µmol m^{-2} s^{-1}) and 25ºC light/20ºC dark temperatures (standard conditions).

Mining of pds in the C. sativa transcriptome and phylogenetic analysis

C. sativa pds (*Cspds*) gene sequence was kindly provided by Dr. Susana Serrazina, from the *C. sativa* transcriptome database (Serrazina *et al.,* 2015). *Cspds* gene was analysed using Prosite (https://prosite.expasy.org/**,** accessed number 01-01-2021) to annotate functional domains.

The alignment of the *Cspds* gene sequence together with 44 other plant *pds* coding sequences, available on NCBI database (accessed number 01-01-2021), was performed via multiple sequence alignment using the ClustalW algorithm (http://www.clustal.org/, accessed number 01-01-2021).

MEGAX software (https://www.megasoftware.net/, accessed number 01-01-2021) was adopted to generate the phylogenetic tree, by applying maximum likelihood algorithm. The individual branch statistical significance was assessed by bootstrap analysis with 1,000 iterations.

Vector design

The two gRNAs were designed on the *Cspds* sequence using CRISPR-P 2.0 (http://crispr.hzau.edu.cn/CRISPR2/, accessed number 01-01-2021) (Supplementary File 1). Putative off-target sites were identified with the CRISPOR software (http://crispor.tefor.net/crispor.py/), using the *Castanea mollissima* genome as reference. The gRNAs were assembled into a CRISPR/Cas9 vector carrying the h*Cas9* and the *nptII* gene for kanamycin (kan) resistance, using the GoldenBraid (GB) assembly system and following GB software-directed procedures (https://gbcloning.upv.es/, accessed number 01-01-2021). CaMV 35S promoter and AtU6-26 RNA PolIII promoter were used to drive the hCas9 and gRNA expression, respectively.
The final vector named nptII_Cas9_*pds*_gRNA1_gRNA2 (Supplementary File 2) was transferred into *Agrobacterium tumefaciens* strain EHA105 (Hood *et al.,* 1993) by using the freeze/thaw method (Xu and Li, 2008).

Somatic embryo genetic transformation

Cultures of EHA105-nptII_Cas9_*pds*_gRNA1_gRNA2 for transformation experiments were prepared according to Corredoira et al. (2015). A single colony was inoculated in 2 mL of LB medium (Sambroock *et al.,* 1989) supplemented with 50 mg/L kan and the

culture was grown overnight at 28°C with shaking (200 rpm) in darkness. One mL of this bacterial suspension was added into 600 mL of LB liquid medium added with 50 mg/L kan and was grown at 28°C at 100 rpm in darkness until the attainment of an OD_{600}= 0.6. Bacteria were recovered by centrifugation (6500 rpm for 10 min at 10°C), and the pellet was resuspended in 200 mL of MS liquid medium supplemented with 5% sucrose (w/v) (infection medium).

For transformation, clumps of 2-3 somatic embryos at globular or early torpedo stage were employed as target explants. Somatic embryos were pre-cultured on free growth plant regulators multiplication medium for one day before the transformation trial. The explants were immersed during 30 min in the infection medium and transferred to multiplication medium for co-cultivation in dark at 25°C. After 5 days of co-cultivation, explants were washed for 30 min with sterilized water with 500 mg/L cefotaxime and subsequently somatic embryo groups were transferred to multiplication medium supplemented with carbenicillin (300 mg/L), cefotaxime (200 mg/L) and kanamycin (150 mg/L) (selection medium) and cultured under standard conditions. In this experiment, 60 explants (6 Petri dishes with 10 explants per dish) were transformed. In addition, 10 clumps of somatic embryos transformed with the nptII / Cas9 vector (gRNA-free control) and 20 groups of un-infected embryos were cultured to be used as negative and positive control, respectively (Figure 1).

After 8 weeks on selection medium, kanamycin-resistant embryos were detected and transferred to fresh selection medium to establish the different mutated lines. These transformed lines were routinely maintained by secondary embryogenesis with sequential subcultures at 6-week

intervals on selection medium according to Corredoira *et al.* (2015). The multiplication capacity of two transformed/mutated lines was evaluated by recording the number of somatic embryos produced by explant and compared with those of the control embryogenic line.

Plant regeneration

Plant regeneration was performed according to Corredoira *et al.* (2008). Cotyledonary somatic embryos (>5 mm) were isolated and cultured on maturation medium consisting of ½ MS medium supplemented with 3% maltose (w/v) and 0.8% Sigma-agar (w/v). After 4 weeks of culture on maturation medium at standard conditions, somatic embryos were transferred to empty Petri dishes and stored at 4°C. After 2 months, somatic embryos were cultured on germination medium consisting of ½MS supplemented with 0.1 mg/L BA, 0.1 mg/L indole-3-butyric acid (IBA), glutamine (200 mg/L) and 0.7% Sigma-agar (w/v). After 8 weeks at standard conditions, the percentage of regenerated plants/ shoot development was evaluated.

Molecular analysis of kanamycin-resistant embryos

Genomic DNA was extracted from 0.1 g of white kanamycin-resistant embryos (obtained through transformation with nptII_*Cas9_pds*_gRNA1_gRNA2) and from nptII_*Cas9* control using the Qiagen DNeasy Plant Pro Kit. The transgene integration was investigated through h*Cas9* gene amplification (in three technical replicates) using SYBR Green PCR Master Mix (Applied Biosystems) and the following

PCR program: 95°C/10 min, followed by 40 cycles of 95°C/15 s and 60°C/1 min cycle. All primer sequences are available in Supplementary File 3.

Mutation frequencies at the target sites were evaluated through PCR amplification using primers designed on gRNAs flanking regions (Supplementary File 3). DNA was amplified using KAPA HIFI Taq and the following PCR program: 95°C/3 min, followed by 30 cycles of 98°C/20 sec, 60°C/20 sec, 72°C/45 sec and 72°C/3 min. The PCR products were purified using DNA/RNA Clean Up E.Z.N.A.® kit. Samples were sequenced using the Sanger method and the chromatograms obtained were analyzed using the TIDE online software (https://tide.deskgen.com; accessed number 01-01-2021).

Statistical analysis

The influence of mutation on somatic embryo proliferation (Table 1) was evaluated by one-way factorial analysis of variance (ANOVA I) applying SPSS for Windows (version 26.0, Chicago, USA).

Results and discussion

Gene structure and phylogenetic analysis

The *pds* gene sequence is 1498 bp long and the protein size is 472 amino acids (aa). A single amino_oxidase domain (PF01593) is present within the protein sequence (Figure 2). Available full-length NCBI *pds* coding sequence orthologues (Supplementary File 4) were used for phylogenetic tree construction. The resulting unrooted maximum-likelihood tree is shown in Figure 2.

Underlined in yellow a sub-clade containing *Cspds* together with gene sequences closely related from *Quercus* spp. and *Juglans regia* (100% bootstrap value) is reported. Both are nut species belonging to the *Fagales* order (Bernard *et al.,* 2018). *Castanea* and *Quercus* genus belong to the same *Fagacea*e family while *Juglans* to *Juglandaceae* family. This result suggests the common origin of chestnut, oak and walnut *pds* gene. The phylogenetic relationship of gene sequences from *Quercus*, *Juglans* and *Castanea* was also confirmed in Pavese *et al.* (2021).

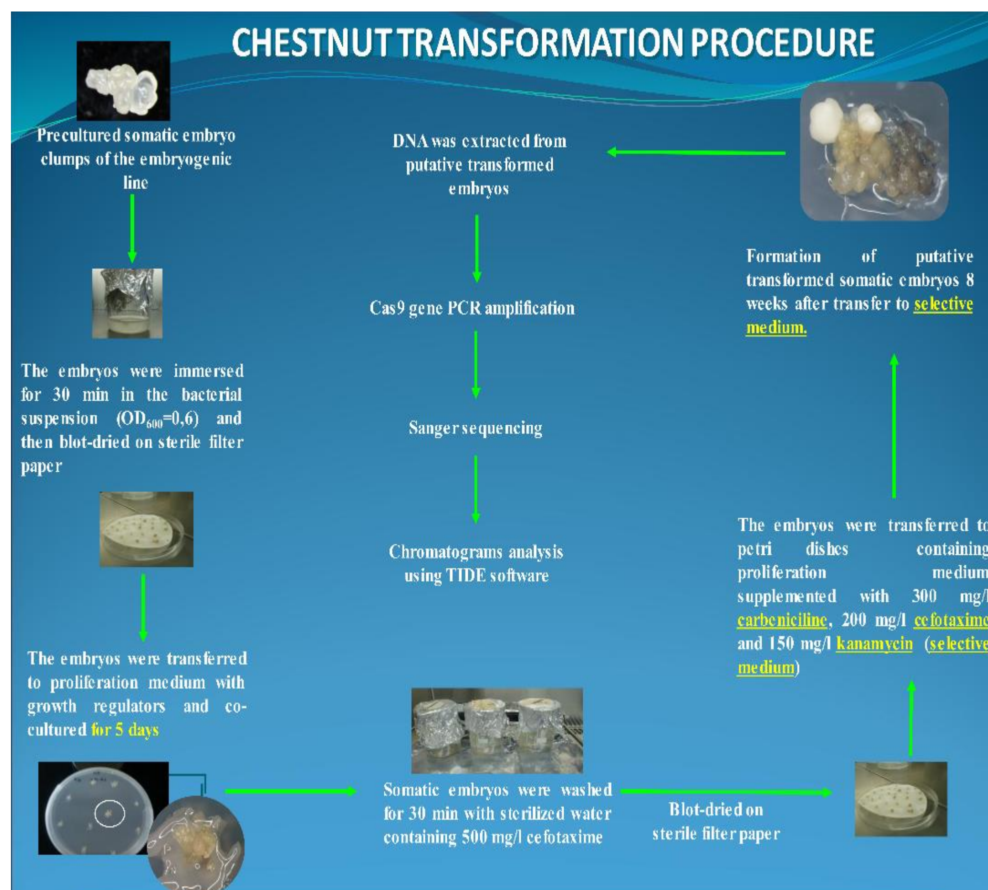

Figure 1. Flow chart of the chestnut somatic embryo transformation protocol using EHA105-nptII_Cas9_*pds*_gRNA1_gRNA2 construct. The putative transformed lines were analysed at DNA level through the amplification of the *Cas9* gene and Sanger sequencing.

Table 1. Secondary embryogenesis ability of non-transgenic line (CI-3-wt) and two mutated lines (W1, W2).

Embryogenic Line	**Embryo-stage**		
	Globular-torpedo **nº/explant**	**Cotyledonary** **nº/explant**	**Total nº of somatic embryos/explant**
CI-3-wt	7.0 ± 0.9	8.5 ± 0.9	15.5 ± 0.8
W1	6.2 ± 1.2	4.8 ± 0.8	11.0 ± 0.8
W2	7.5 ± 0.7	5.3 ± 0.7	12.8 ± 1.0
ANOVA I	Ns	$p \leq 0.05$	$p \leq 0.05$

The data show the total number of somatic embryos per explant and the number of embryos per developmental stage, recorded after 6 weeks of culture in selection medium. Values are means ± SE of 6 Petri dishes with 8 explants each. ANOVA I significances are shown for each parameter. ns, not significant.

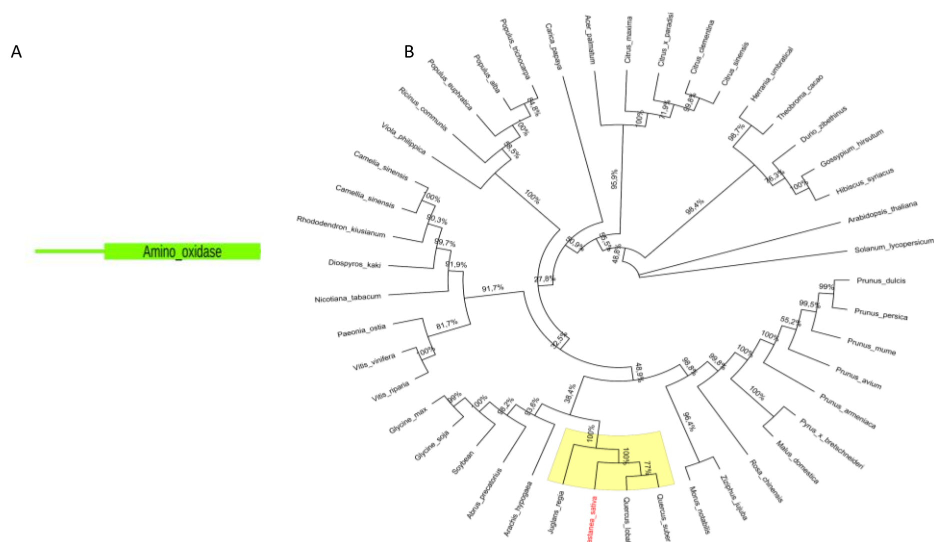

Figure 2. A) Structural domains of *Cspds*. The amino_oxidase domain (PF01593) ranges from 71 to 467 amino acids. B) Phylogenetic analysis of *pds* gene. The phylogenetic tree was constructed using MEGAX software by aligning chestnut *pds* coding sequences with NCBI S-gene ortholog coding sequences. The yellow colour indicates the C. *sativa* clade. The *C. sativa* sequence is highlighted in red colour.

CRISPR/Cas9 mediated transformation of somatic embryos

The main problems in the application of genome editing technologies in woody species include low transformation efficiency, recalcitrance to transformation and difficulties in plant regeneration. Moreover, the predominantly outcrossing nature of trees and highly heterozygous genomes represent additional challenges. The limited number of whole genome sequences available hamper the design of effective sgRNA and the reduction of off-target effects.

To determine whether the CRISPR/Cas9 system may be suitable for gene editing in *C. sativa*, we used as a target the *pds* gene. As this gene is related to chlorophyll biosynthesis (Walawage *et al.,* 2019; Wang *et al.,* 2020), an albino phenotype of somatic embryos is expected, which would allow easier discrimination of mutated embryos. This visual marker has also been applied in defining editing methods in other woody species such as cassava (Odiplo *et al.,* 2017), coffee (Breitler *et al.,* 2018) and walnut (Walawage *et al.,* 2019). Besides proof-of-concept studies, the CRISPR/Cas9 system has been applied to obtain disease resistant fruit trees (Bewg *et al.,* 2018).

Two unique gRNAs were designed, directed at the coding sequences of *pds* gene. One gRNA targeted the Amino_oxidase domain, while the other was chosen to be as close as possible to the 5' end of the coding sequence, in order to ensure that mutations would affect the protein translation. No loci were found in the *Castanea* genome which could be considered as a likely source of off-target gene editing. Each gRNA was put under the control of the Arabidopsis U6-26 promoter. The transcription efficiency of sgRNA is pivotal for an efficient CRISPR/Cas9 genome editing. Both endogenous and exogenous and U6 promoters have been successfully deployed for controlling sgRNA transcription in plants. The AtU6 promoter was used in poplar (Fan *et al.,* 2015), apple (Nishitani *et al.,* 2016) and grape (Wang *et al.,* 2018).

GoldenBraid (GB) cloning system, suited for gene editing experiments, was previously used (Maioli *et al.,* 2020; Vazquez-Vilar *et al.,* 2016). We adopted a dual sgRNA construct in order to increase the genome editing efficiency either by increasing the possibility that at least one gRNA will

be active for mutagenesis or deleting large gene fragments, in case double strand break are simultaneous (Supplementary File 5) (Pauwels *et al.,* 2018; Xie *et al.,* 2015).

Somatic embryos at the globular or torpedo stage were transformed with EHA105- nptII_Cas9_*pds*_gRNA1_gRNA2, by applying the previously defined protocol (Corredoira *et al.,* 2012; 2016). Since many hardwood species are recalcitrant to *in vitro* regeneration by organogenesis, somatic embryogenesis (SE) is considered one of the best methods of producing modified plants by genetic engineering (Peña and Séguin, 2001; Corredoira *et al.,* 2019). When somatic embryos are used as target explants an important factor to achieve a successful genetic transformation is the election of the most suitable embryo developmental stage. It is known that in somatic embryos at early developmental stage many cells are undergoing active cell division, and *Agrobacterium* infection is therefore more feasible (Corredoira *et al.,* 2007). In the present study, using globular and or torpedo stage embryos 6.6% of kanamycin-resistant explants were obtained after 5 days of co-cultivation with *Agrobacterium* and 8 weeks on selection medium. This rate is similar to those reported in previous studies, in which the same embryogenic lines were transformed with a thaumatin-like gene (Corredoira *et al.,* 2012) or a chitinase gene (Corredoira *et al.,* 2016). In addition, visual evidence of altered *pds* function was given in all kanamycin-resistant explants by producing albino somatic embryos (Figure 3A, B). By contrast, white somatic embryos were not detected in the nptII/Cas9 control (Figure 3C).

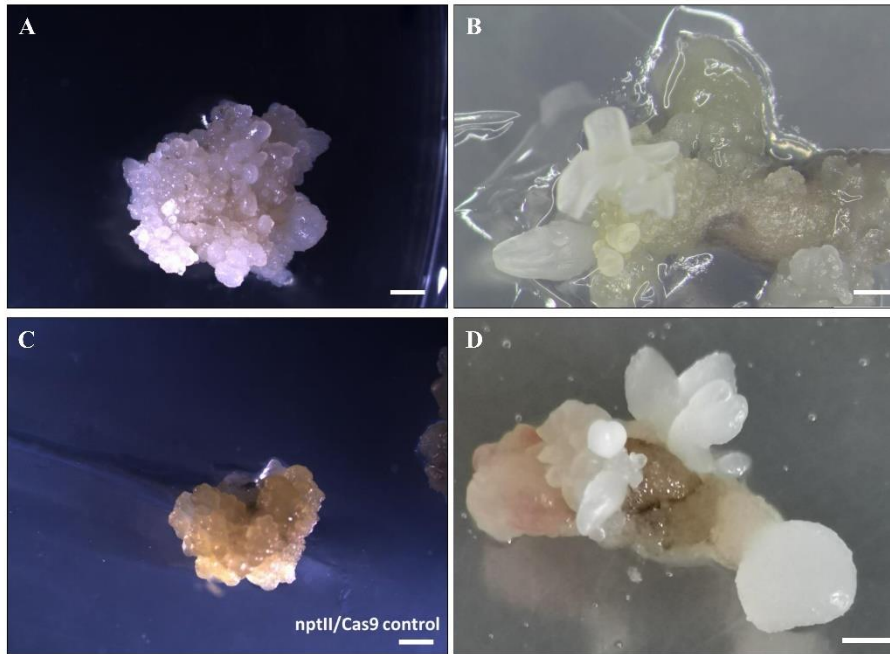

Figure 3. Gene editing of somatic embryos of European chestnut. A, B) White transgenic somatic embryos at different developmental stages formed after 12 weeks on selection medium. C) nptII/Cas9 control. D) Secondary embryos of W1 line generated following 6 weeks of culture on selection medium. Bar: 1mm.

Somatic embryo proliferation, maturation and germination process

One somatic embryo was isolated from each kanamycin-resistant explant and independently cultured to establish four different embryogenetic lines: W1, W2, W3 and W4. These mutated lines were successfully multiplied by secondary embryogenesis on selection medium (Figure 3D) in order to produce enough material for both maturation and germination steps and for molecular mutation screening. To ascertain whether the multiplication

ability is affected by the transformation, the number of secondary somatic embryos of lines W1 and W2 relative to the non-transgenic line was determined after 6 weeks of culture on selective medium (Table 1). The mutated lines produced significantly fewer somatic embryos (11.0-12.8) than the non-transgenic line (15.5). Regarding the number of somatic embryos in relation to the developmental stage, we did not find any differences in the number of globular-torpedo developmental stage; however, the number of cotyledonary embryos was significantly higher in the non-transgenic line ($p \leq 0.05$), which limited the subsequent maturation step (Table 1).

Mutated cotyledonary embryos were allowed to mature on maltose medium followed by 2 months cold storage period. Only 50% embryos of W1 line, 30 % of W2 line, 13 % of W3 line and 6 % of W4 line survived to cold period. Surviving explants were transferred to germination medium. Only shoot-development was observed (13% of W1 line, 3% of W2 line, 7% of W3 and 2% of W4 line). The low plant regeneration rate of somatic embryos is a common problem in SE systems in hardwoods (Corredoira *et al.,* 2019). This problem has also been reported in European chestnut, especially in transgenic somatic embryos in which simultaneous development of shoot and root is occasionally observed (Corredoira *et al.,* 2004; 2012; 2015). Moreover, it is known that the mutation of *pds* gene negatively affects to plant conversion; Breitler *et al.* (2018) pointed out that the loss of function of *pds* gene causes a reduction of photosynthetic pigments that provoke the low germination rates. Similarly, Walawage *et al.* (2019) found in walnut that *pds* gene is essential for proper plant growth. As expected, chestnut transformed somatic embryos showed shoots displaying the typical albino phenotype, but a stunted phenotype

with shorter internodes and small leaves was also observed (Figure 4A, B). Similar phenotypic aberrations were described in rice, apple, sweet orange, and poplar (Odipio *et al.,* 2017). In *Arabidopsis* the alteration of *pds* gene function causes dwarfism and albino phenotypes due to impaired chlorophyll, gibberellin and carotenoid biosynthesis (Breitler *et al.,* 2018).

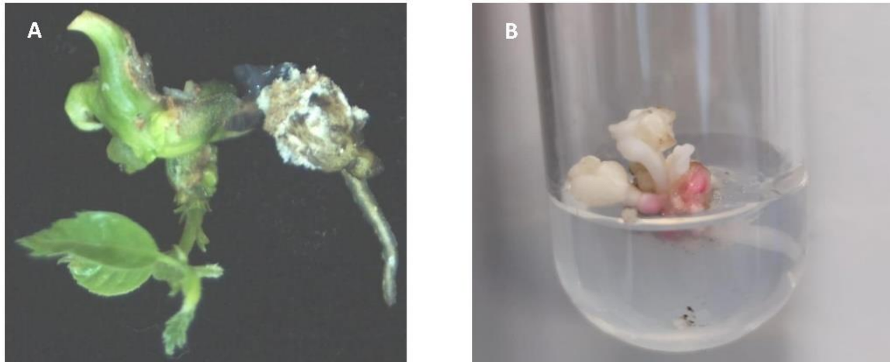

Figure 4. (A) Green plant regenerated from an untransformed somatic embryo. **(B)** Albino shoot originated from a germinated somatic embryo of transgenic line W1, showing a stunted morphology with shorter internodes and small leaves.

Somatic embryogenesis mutation screening

The qPCR analysis using Cas9 gene-specific primers revealed genomic integration of the construct in W1-W4 embryo lines (Supplementary File 6). In order to detect *pds* gene editing efficiency and the types of mutations, the Sanger sequencing was used in association with TIDE software (Table 2). The average gene editing efficiency was 61% for gRNA1 and 56% for gRNA2. Although the editing efficiency was different between the two gRNAs, it was not possible to attribute this

dissimilarity to the different GC content of target sequences, or to the sgRNA secondary structure, or to the promoters that direct Cas9 and gRNAs expression (Ma *et al.,* 2015). This reinforces the importance of assuring efficient knock-out by employing different gRNAs.

Molecular data demonstrated a higher editing efficiency in W1 and W2 lines than in W3 and W4 lines for both gRNAs tested. W1 showed the greatest editing efficiency, i.e. 94% for gRNA1 and 80% for gRNA2 (Figure 5, 6, 7, 8). In the case of gRNA1, three lines were heterozygous and one chimeric; in the case of gRNA2, two lines were chimeric and two heterozygous mutants. Our results are in contrast with what was observed in other woody plants (grape, pear, apple, or poplar) showing high level of homologous and biallelic mutants in their T_0 generation (Dai *et al.,* 2021). The obtaining of chimeric genotypes by transgenic protocols or genome editing is a very challenging problem (Ding *et al.,* 2020). Self-pollination is usually used in most model and crop plants to obtain homozygous individuals from heterozygous transgenic or gene edited plants. However, it is difficult to obtain homozygous mutants by this method in woody trees that have long vegetative period and are often self-incompatible. A recent study has reported that a second round of shoot regeneration as well as a further selection with kanamycin can produce homozygous mutant shoots at a high frequency in poplar (Ding *et al.,* 2020).

The most common mutations in our T_0 chestnut plantlets were represented by a single nucleotide insertion followed by deletions of 1, 2, 4 and 10 nucleotides. Previous observations showed that small indels are the predominant mutations introduced in plants by gene editing, with 1 bp insertions, especially +T and +A, predominant in most cases (Bortesi *et*

al., 2016). However, considerable variations in type and size of mutations are reported in literature (Jacobs *et al.,* 2015; Xu *et al.,* 2015), highlighting a possible influence of target site sequences and/or of their genomic contexts.

Table 2. Genotyping of targeted gene mutations induced by CRISPR/Cas9 in the T_0 generations in the four edited lines. Editing efficiency, allelic asset (heterozygous (HE) and chimeric (CH)) and allele frequency are indicated.

Sample	**-10**	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**Efficiency**	**R^2**	**Zygosity**
W1_gRNA1				2.8	5.4	2.8	85.9	94.30%	0.97	CH
W2_gRNA1						19.1	76.5	77.70%	0.97	HE
W3_gRNA1						61.2	35.3	35.40%	0.97	HE
W4_gRNA1						59.2	36.8	36.80%	0.96	HE
W1_gRNA2				21.8	37.9		18.4	80.40%	0.8	CH
W2_gRNA2	10.9	15.8		8.3	30.7		14.1	79.80%	0.8	CH
W3_gRNA2						54.2	26.8	27.60%	0.82	HE
W4_gRNA2					35.4	48.7		36.10%	0.85	HE

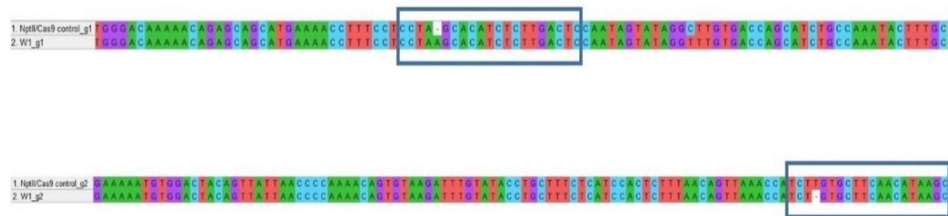

Figure 5. Example of genotyping of targeted gene mutations induced by CRISPR/Cas9 in the W1 line. The blue box underlines the gRNAs sequences. As compared to nptII/Cas9 control, the W1 line showed a nucleotide insertion at gRNA1 level and a nucleotide deletion at gRNA2 level.

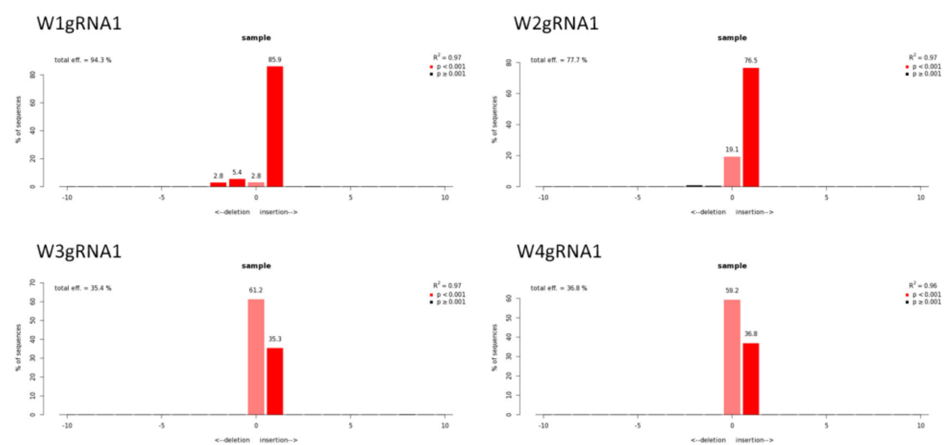

Figure 6. Genotyping of targeted gene mutations induced by CRISPR/Cas9 on gRNA1 in the transformed lines (W1-W4). Editing efficiency and mutagenesis frequencies are reported.

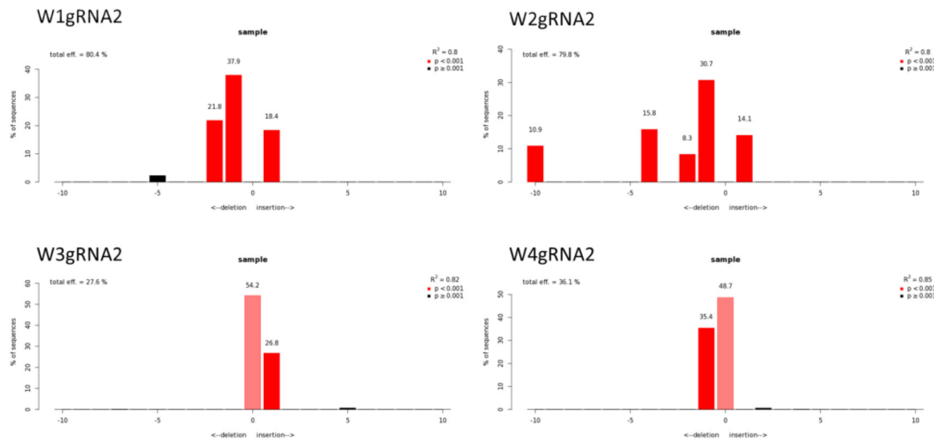

Figure 7. Genotyping of targeted gene mutations induced by CRISPR/Cas9 on gRNA2 in the transformed lines (W1-W4). Editing efficiency and mutagenesis frequencies are reported.

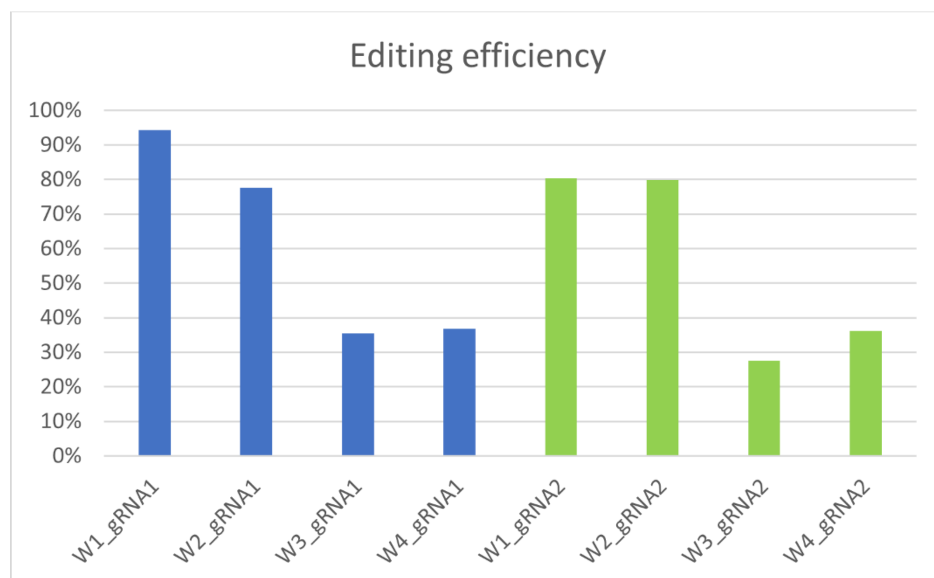

Figure 8. Genotyping of targeted gene mutations induced by CRISPR/Cas9 on gRNA1 and gRNA2 in the transformed lines (W1-W4).

Conclusions

In conclusion, we report for the first time the application of CRISPR/Cas9 technology in *Castanea* genus. Our system, based on the use of somatic embryos and two guide RNAs directed simultaneously at *pds* locus demonstrated to be specific for the target gene. Non-pigmented "albino" shoots obtained from *in vitro* cultures were associated with the successful editing of this gene. Since the antibiotics reduced the percentage and efficiency of regeneration, it will be interesting to optimize the transformation protocol trying to minimize the effect of these substances on the regeneration process.

The presence of an effective gene editing method will facilitate the chestnut breeding improvement, acting on genes responsible for pathogen resistance/susceptibility, such as the two candidate S genes (*pmr4* and *dmr6*) potentially involved in *C. sativa* susceptibility to *C. parasitica* and *P. cinnamomi* (Pavese *et al.,* 2021).
However, a further step of the research may consider the development of genome-editing tools that do not require integration of the CRISPR/Cas9 cassette. An efficient DNA-free genome editing system was developed using *in vitro* assembled Cas9/sgRNA ribonucleoproteins (RNPs), that are delivered in plant protoplasts using polyethylene glycol-mediated transfection (Woo *et al.,* 2015; Malnoy *et al.,* 2016).

Data availability statement

The original contributions presented in the study are included in the article/**SupplementaryMaterials**, further inquiries can be directed to the corresponding author.

Author contributions

AM, DT, and RB: conceptualization. VP, AM, and EC: data curation and writing—original draft preparation. VP, AM, EC, and TM: investigation. AM, DT, RB, EC, and TM: supervision and writing—review and editing preparation. All authors contributed to the article and approved the submitted version.

Funding

This study was financially supported by Fondazione CRT and by Ministerio de Ciencia e Innovación (Spain) through the project PID2020-112627RB-C33.

Acknowledgments

We thank Dr. Susana Serrazina for having provided the *pds* sequence.

Supplementary material

The Supplementary Material for this article can be found online at:

https://www.frontiersin.org/articles/10.3389/fpls.2021.728516/full#supplementary-material

Supplementary File 1 | Selected gRNA sequences predicted from *C. sativa pds* (Correspond to Data sheet 1).

Supplementary File 2 | CRISPR/Cas9 vector nptII_Cas9_pds_gRNA1_gRNA2 assembled using GoldenBraid (GB) cloning system (Correspond to image 1).

Supplementary File 3 | Primer sequences (Correspond to Data sheet 2).

Supplementary File 4 | *Pds* coding sequences used for tree construction (Correspond to Data sheet 3).

Supplementary File 5 | Focus on nptII_Cas9_pds_gRNA1_gRNA2 vector at gRNA level (Correspond to image 2).

Supplementary File 6 | PCR analysis of Cas9 integration in the different embryogenetic lines (Correspond to image 3).

References

Bernard, A., Barreneche, T., Lheureux, F., Dirlewanger, E. (2018). Analysis of genetic diversity and structure in a worldwide walnut (*Juglans regia* L.) germplasm using SSR markers. *PLoS One*.13(11), e0208021. doi:10.1371/journal.pone.0208021.

Bernard, G., Gagneul, D., Alves Dos Santos, H., Etienne, A., Hilbert, J. L., Rambaud, C. (2019). Efficient genome editing using CRISPR/Cas9 technology in chicory. *Int. J. Mol. Sci.* 20, E1155. doi: 10.3390/ijms20051155.

Bewg, W. P., Ci, D., Tsai, C.-J. (2018). Genome Editing in Trees: From Multiple Repair Pathways to Long-Term Stability. *Front. Plant Sci.* 9, 1732. doi: 10.3389/fpls.2018.01732.

Bortesi, L., Zhu, C., Zischewski, J., Perez, L., Bassié, L., Nadi, R., et al. (2016). Patterns of CRISPR/Cas9 activity in plants, animals and microbes. *Plant Biotechnol. J.* 14, 2203-2216. doi: 10.1111/pbi.12634.

Breitler, J. C., Déchamp, E., Campa, C., Rodrigues, L. A. Z., Guyot, R., Marraccini, P., et al. (2018). CRISPR/Cas9-mediated efficient targeted mutagenesis has the potential to accelerate the domestication of *Coffea canephora*. *Plant Cell Tissue Organ Cult.* 134, 383-394. doi: 10.1007/s11240-018-1429-2.

Charrier, A., Vergne, E., Dousset, N., Richer, A., Petiteau, A., Chevreau, E. (2019). Efficient targeted mutagenesis in apple and first time edition of pear using the CRISPR-Cas9 system. *Front. Plant Sci.* 10, 40. doi: 10.3389/fpls.2019.00040.

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. *Veget. Hist. Archaeobot.* 13.

Conedera, M., Tinner, W., Krebs, P., de Rigo, D., Caudullo, G. (2016). *Castanea sativa* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, pp. e0125e0+

Corredoira, E., Ballester, A., Vieitez, A.M. (2006). Somatic embryogenesis in chestnut. In: Mujib S, Samaj J (eds) Somatic embryogenesis. Plant cell monogr. (2). *Springer* 177–199.

Corredoira, E., Merkle, S.A., Martínez, M.T., Toribio, M., Canhoto, J.M., Correia, S.I., Ballester, A., Vieitez, A.M. (2019). Non zygotic embryogenesis in hardwood species. *CRC Crit. Rev. Plant Sci.* 38, 29-97.

Corredoira, E., Montenegro, D., San José, M.C., Vieitez, A.M., Ballester, A. (2004). *Agrobacterium*-mediated transformation of European chestnut embryogenic cultures. *Plant Cell Rep.* 23, 311-318.

Corredoira, E., San-José, M.C., Vieitez, A.M. *et al.* (2007). Improving genetic transformation of European chestnut and cryopreservation of transgenic lines. *Plant Cell Tiss Organ Cult.* 91, 281. doi:10.1007/s11240-007-9295-3.

Corredoira, E., San José, M.C., Vieitez, A.M. *et al.* (2016). *Agrobacterium*-mediated transformation of European chestnut somatic embryos with a *Castanea sativa* (Mill.) endochitinase gene. *New Forests* 47, 669-684. doi: 10.1007/s11056-016-9537-5

Corredoira, E., Valladares, S., Allona, I., Aragoncillo, C., Vieitez, A.M., Ballester, A. (2012). Genetic transformation of European chestnut somatic embryos with a native thaumatin-like protein (*CsTL1*) gene isolated from *Castanea sativa* seeds. *Tree Physiol.* 32, 1389-1402.

Corredoira, E., Valladares, S., Vieitez, A.M. *et al.* (2008). Improved germination of somatic embryos and plant recovery of European chestnut. *In Vitro Cell.Dev.Biol.-Plant* 44, 307. doi:10.1007/s11627-008-9105-6.

Corredoira, E., Valladares, S., Vieitez, A.M., Ballester, A. (2015). Chestnut, European (*Castanea sativa*). In: Wang K (ed) Methods in molecular biology *Agrobacterium* protocols. *Springer* 163–176.

Dai, X., Yang, X., Wang, C., Fan, Y., Xin, S., Hua, Y., Wang, K., Huang, H. (2021). CRISPR/Cas9-mediated genome editing in *Hevea brasiliensis*. *Ind. crop. prod.* 164, 113418.

Ding, L., Chen, Y., Ma, Y., Wang, H., Wei, J. (2020). Effective reduction in chimeric mutants of poplar trees produced by CRISPR/Cas9 through a second round of shoot regeneration. *Plant Biotechnol Rep.* 14, 549-558. doi:10.1007/s11816-020-00629-2.

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* 5, 12217. doi: 10.1038/srep12217.

Fister, A.S., Landherr, L., Maximova, S.N., Guiltinan, M.J. (2018). Transient Expression of CRISPR/Cas9 Machinery Targeting *TcNPR3* Enhances Defense Response in *Theobroma cacao*. *Front. Plant Sci.* 9, 268. doi: 10.3389/fpls.2018.00268.

Hood, E.E., Gelvin, S.B., Melchers, L.S., Hoekema, A. (1993). New Agrobacterium helper plasmids for gene transfer to plants. *Transgenic Res.* 2, 208-218. doi: 10.1007/BF01977351.

Jacobs, T. B., Lafayette, P. R., Schmitz, R. J., and Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol.* 15, 16. doi: 10.1186/s12896-015-0131-2.

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., et al. (2015). A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* 8, 1274-1284.

Ma, C., Zhu, C., Zheng, M., Liu, M., Zhang, D., Liu, B., et al. (2019). CRISPR/Cas9-mediated multiple gene editing in *Brassica oleracea* var. capitata using the endogenous tRNA-processing system. *Hortic. Res.* 6, 20. doi: 10.1038/s41438-018-0107-1

Maioli, A., Gianoglio, S., Moglia, A., Acquadro, A., Valentino, D., Milani, A.M., Prohens, J. et al.(2020) Simultaneous CRISPR/Cas9 Editing of Three PPO Genes Reduces Fruit Flesh Browning in Solanum melongena L. Front. Plant Sci. 11, 607161. doi: 10.3389/fpls.2020.607161.

Malnoy, M., Viola, R., Jung, M. H., Koo, O. J., Kim, S., Kim, J. S., Velasco, R., & Nagamangala Kanchiswamy, C. (2016). DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. Front. Plant Sci. 7, 1904. https://doi.org/10.3389/fpls.2016.01904

Merkle, S.A., Vieitez, F.J., Corredoira, E., Carlson, J.E. (2020). *Castanea* spp. Chesnut. In: Biotechnology of Fruit and Nut Crops. Litz RE, Pliego Alfaro F, Hormaza JI (Eds). *CABI Publishing* 206-237.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *tiPhysiol Plant* 15, 473-497.

Naidoo, S., Slippers, B., Plett, J.M., Coles, D., Oates, C.N. (2019). The Road to Resistance in Forest Trees. *Front. Plant Sci.* 10, 273. doi: 10.3389/fpls.2019.00273.

Nishitani, C., Hirai, N., Komori, S., Wada, M., Okada, K., Osakabe, K., et al. (2016). Efficient genome editing in apple using a CRISPR/Cas9 system. *Sci. Rep.* 6, 31481. doi: 10.1038/srep31481.

Odipio, J., Alicai, T., Ingelbrecht, I., Nusinow, D. A., Bart, R., and Taylor, N. J. (2017). Efficient CRISPR/Cas9 genome editing of phytoene desaturase in *Cassava*. *Front. Plant Sci.* 8, 1780. doi: 10.3389/fpls.2017.01780.

Pan, C., Ye, L., Qin, L., Liu, X., He, Y., Wang, J., et al. (2016). CRISPR/Cas9- mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. *Sci. Rep.* 6, 24765. doi: 10.1038/srep24765.

Pauwels, L., De Clercq, R., Goossens, J., Iñigo, S., Williams, C., Ron, M., Britt A. and Goossens, A. (2018). A Dual sgRNA Approach for Functional Genomics in *Arabidopsis thaliana*. *G3 Genes Genom. Genet.* 8, 2603-2615.

Pavese, V., Moglia, A., Gonthier, P., Torello Marinoni, D., Cavalet-Giorsa, E., Botta, R. (2021). Identification of Susceptibility Genes in *Castanea sativa* and their transcription dynamics following pathogen infection. Plants 10, 913. doi:10.3390/plants10050913.

Peña, L., Séguin, A. (2011). Recent advances in the genetic transformation of trees. *Trends Biotechnol.* 19, 500-506.

Qin, G., Gu, H., Ma, L., Peng, Y., Deng, X.W., Chen, Z., Qu, L.J. (2007) Disruption of *phytoene desaturase* gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res.* 17, 471-82. doi: 10.1038/cr.2007.40.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd edn. *Cold Spring Harbor Laboratory Press,* Cold Spring Harbor.

Santillán Martínez, M. I., Bracuto, V., Koseoglou, E., Appiano, M., Jacobsen, E., Visser, R. G. F., Wolters, A-M. A., Bai, Y. (2020). CRISPR/Cas9-targeted mutagenesis of the tomato susceptibility gene PMR4 for resistance against powdery mildew. *BMC Plant Biol.* 20(1), 284. doi:10.1186/s12870-020-02497-y.

Serrazina, S., Santos, C., Machado, H., Pesquita, C., Vicentini, R., Pais, M.S., Sebastiana, M., Costa, R. (2015). *Castanea* root transcriptome in response to *Phytophthora cinnamomi* challenge. *Tree Genet. Genomes* 11, 6.

Shan, S., Mavrodiev, E. V., Li, R., Zhang, Z., Hauser, B. A., Soltis, P. S., et al. (2018). Application of CRISPR/Cas9 to *Tragopogon* (Asteraceae), an evolutionary model for the study of polyploidy. *Mol. Ecol. Resour.* 18, 1427-1443. doi: 10.1111/1755-0998.12935.

Vazquez-Vilar, M., Bernabe-Orts, J.M., Fernandez-Del-Carmen, A., Ziarsolo, P., Blanca, J., and Granell, A. (2016). A modular toolbox for gRNA-Cas9 genome engineering in plants based on the golden braid standard. *Plant Methods* 12, 10-21. doi: 10.1186/s13007-016-0101-2.

Walawage, S.L., Zaini, P.A., Mubarik, M.S., Martinelli, F., Balan, B., Caruso, T., Leslie, C.A., Dandekar, A.M. (2019). Deploying Genome

Editing Tools for Dissecting the Biology of Nut Trees. *Front. Sustain. Food Syst.* 3, 100. doi: 10.3389/fsufs.2019.00100.

Wang, Z., Wang, S., Li, D., Zhang, Q., Li, L., Zhong, C., et al. (2018). Optimized paired-sgRNA/Cas9 cloning and expression cassette triggers high efficiency multiplex genome editing in kiwifruit. *Plant Biotechnol. J.* 16, 1424-1433. doi: 10.1111/pbi.12884.

Wang, J., Wu, H., Chen, Y., Yin, T. (2020). Efficient CRISPR/Cas9-Mediated gene editing in an interspecific hybrid *poplar* with a Highly Heterozygous Genome. *Front Plant Sci.* 11, 996. doi:10.3389/fpls.2020.00996.

Wilson, F. M., Harrison, K., Armitage, A. D., Simkin, A. J., and Harrison, R. J. (2019). CRISPR/Cas9-mediated mutagenesis of *phytoene desaturase* in diploid and octoploid strawberry. *Plant Methods* 15, 45. doi: 10.1186/s13007-019-0428-6.

Woo J.W., Kim J., Kwon S.I., Corvalán C., Cho S.W., Kim H., Kim S.-G., Kim S.-T., Choe S., Kim J.-S. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* 33, 1162-1164.

Xie, K., Minkenberg, B., Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. USA* 112, 3570-3575.

Xu, R., Li, Q.Q. (2008). Protocol: Streamline cloning of genes into binary vectors in *Agrobacterium* via the Gateway®TOPO vector system. *Plant Methods* 4, 4.

Xu, R.-F., Li, H., Qin, R.-Y., Li, J., Qiu, C.-H., Yang, Y.-C., et al. (2015). Generation of inheritable and "transgene clean" targeted genome-modified

rice in later generations using the CRISPR/Cas9 system. *Sci. Rep.* 5, 11491. doi: 10.1038/srep11491.

Chapter IV

First protoplasts isolation and transformation protocol in *C. sativa* Mill.

Introduction

CRISPR-Cas9 machinery can be delivered to the cell both as nucleic acid form and as RNPs complex. DNA can be integrated into the host genome while RNP complex can be delivered directly to the cell, causing a stable mutation and then being rapidly degraded without leaving exogenous DNA traces (Woo *et al.,* 2015).
Since in Chapter III, an efficient CRISPR/Cas9 protocol in chestnut has been firstly developed, ongoing research is focused on developing DNA-free transformation gene editing methods in chestnut.
In this case, the components of the CRISPR/Cas9 system are *in vitro* synthesized, pre-assembled and then delivered to the plant cell protoplasts as RNP (Bernard *et al.,* 2019).
A protoplast is a plant cell without a wall, removed through enzymatic or mechanical processes. The wall removal releases spherical osmotically fragile protoplasts, characterized by only a cell membrane able to isolate the cytoplasm from the external environment.
Since the cell wall is no longer present, protoplasts are a useful tool for genetic transformation due to their permeability to exogenous DNA

molecules. Furthermore, if the plant is regenerated starting from a single edited cell, its genetic homogeneity is guaranteed (Scintilla *et al.,* 2021).
The major problem related to the woody plant genetic transformation is the difficulty to develop an efficient regeneration protocol due to the release of phenolic compounds in the culture medium (Bertini *et al.,* 2019; Osakabe *et al.,* 2020).
Protoplasts can be directly transformed through the electroporation and PEG-mediated transformation techniques; the particle bombardment and *Agrobacterium*-mediated transformations can be also adopted but remain mainly used on walled plant cells (Scintilla *et al.,* 2021).
Since there is no evidence of published protocols for protoplasts isolation and transformation in *C. sativa*, here we present the first results of a protocol for obtaining transformed protoplasts of chestnut, using the Green Fluorescent Protein (GFP) as a marker gene to test the transformation capacity. Once the effectiveness of the transformation is validated the future goal will be the protoplasts transformation using RNPs.

Material and Methods

Plant material

The isolation and the genetic transformation of *C. sativa* protoplasts were performed following Osakabe *et al.* (2018), making appropriate changes to better adapt the method to chestnut plants.
The explants used for the protoplast's isolation were *C. sativa* embryogenic calli. Embryogenic cultures were induced following Corredoira *et al.,* 2005 from immature seeds grown on MS medium supplemented with 2,4-D (0.5 mg / L) and BAP (1 mg / L). The seeds were

incubated for two months in dark conditions and then transferred to MS medium containing a reduced concentration of BAP (0.1 mg / L) and kept in the growth chamber with a 16/8-hour photoperiod. After 3 months the embryogenic cultures were obtained (Corredoira *et al.,* 2005) and used as starting materials for protoplasts isolation.

Protoplast isolation

C. sativa embryogenic calli (0.1 g) were used as starting materials.
Calli were dissected into small clumps and immersed in cell-wall digestion enzyme solution (0.5% Macerozyme R-10 and 1% Cellulase R-10). These enzymes degrade the cell wall to release the protoplasts. Explants were subjected to a 4-hour digestion on a rotary shaker.
After digestion, the protoplasts were filtered using a nylon mesh (100 µM) to remove cell wall residues and WS solution (Washing Solution, Osakabe *et al.,* 2018) was added to maintain the osmolarity.
Protoplasts were centrifuged at 50 g for 5' and supernatant was discarded. The protoplast pellet was slowly resuspended in WS solution, transferred to 5 ml of 21% (wt/vol) sucrose solution and then centrifuged at 50 g for 5'.
The ring of viable protoplasts was detected in the interface layer; it was aspirated using a Pasteur pipette and resuspended in WS solution. Protoplasts were centrifuged at 50 g for 5' and the pellet was resuspended in 200 µL of MMG solution (Osakabe *et al.,* 2018). Protoplast's viability was tested using Fluorescein diacetate (FDA; Sigma, St. Louis, MO, USA) test (green cells/ total cells × 100%).

Viable protoplasts were counted using a hemacytometer and diluted to obtain a concentration of 2×10^5 in 200 µL and then were stored at 4°C overnight.

GFP vector design

The plasmid pAVA393 (Ochatt *et al.,* 2005), containing the ampicillin resistance gene and the gene coding for GFP, under the control of the 35SCaMV promoter and the Nos terminator was used for the first transformation process.
The plasmid was cloned in *Escherichia coli* DH5α competent cells. Positive colonies were grown in LB liquid medium and then Miniprep (l'E.Z.N.A.® Plasmid Mini Kit I, Omega Bio-tek) was used to achieve high quality plasmid DNA concentration, tested using Qubit® 2.0 fluorimeter (ThermoFisher Scientific).

Protoplast transformation using PEG transfection system

One aliquot of 200 µL of concentrated protoplast was used for pAVA393 transformation experiment. Protoplasts were slowly mixed with 5 µg of pAVA393 plasmid and then 200 µL of PEG 4000 was added. The solution was slowly mixed and incubated for 10' at room temperature. Two WS washing rinses were performed, and the pellet was resuspended in WS solution and then maintained overnight in dark conditions.
After overnight incubation, the solution appeared divided into two layers. The lower layer was picked up and the viable protoplasts were counted using a hemacytometer and diluted to a final 1×10^6 protoplasts/ml.

Transformation efficiency

The protoplasts transformation efficiency for pAVA393 plasmid was evaluated using a fluorescence microscope (Nikon Eclipse Ti2), 40x objective, 5 days after the inoculation process. The excitation was produced by a LED fluorescent source ($\lambda = 470$ nm) and the GFP emission was collected at 516 nm.

Regeneration

Three regeneration media (C1, C2, C3) were tested (Table 1). The protoplasts, after being incorporated in the regeneration media, were incubated in dark conditions at 24°C. Protoplast growth was monitored weekly using a Leica-Wild Heerbrugg M8 stereoscope.

Table 1: Composition of the 3 regeneration media tested.

Medium	**C1**	**C2**	**C3**
NAA	1 mg/L	1 mg/L	\
BAP	0.5 mg/L	0.5 mg/L	0.2 mg/L
2,4-D	\	\	2 mg/L
Sucrose	30 g/L	5 g/L	5 g/L
Mannitol	\	30 g/L	30 g/L
Glucose	50 g/L	50 g/L	50 g/L
MS	\	\	4.4 g/L
Nitsch Medium	\	2.2 g/L	\
Agar	6 g/L	6 g/L	6 g/L

Results and Discussion

Genome editing may represent the future of breeding in woody species which present a high genome complexity and a long juvenile phase. Thanks to genome editing techniques it is possible to perform target mutations in short time. Despite this, since part of the construct may be integrated into the host genome, the edited plants in many countries, including EU, are considered GMO and subjected to strict rules.

For this reason, researchers are trying to develop new strategies to circumvent DNA integration, like the CRISPR/Cas9 delivery using RNPs (Poovaiah *et al.,* 2021), a complex consisting of the recombinant Cas9 nuclease and the gRNA transcribed *in vitro*.

Protoplast isolation

Protoplasts were isolated from *C. sativa* embryogenic calli using an enzymatic digestion mixture which involved the use of Cellulase R-10 (1%) and Macerozyme R-10 (0.5%). The final yield of the protoplasts, evaluated by counting the cells with a hemacytometer, was 4,500,000 protoplasts/mL (Figure 1). The number of isolated protoplasts is comparable to data by Bertini *et al.* (2019) and Malnoy *et al.* (2016). The protoplasts observed under the microscope were vital (91%) and intact, with a spherical shape and a 20-70 µm diameter.

It is necessary to evaluate several parameters during the protoplast isolation: the starting plant material, the enzymes concentration and the incubation time in the enzymatic mixture (Shen *et al.,* 2014). Embryogenic calli are an excellent starting material, as previously confirmed by Bertini

et al. 2019 and Malnoy *et al.* 2016. Being a very friable matrix, they are easily disaggregated into small pieces, unlike the leaf tissue which has a higher level of cellulose and lignin that reduce the digestion capacity (Kuzminsky *et al.,* 2016; Brandt *et al.,* 2020).

The enzymatic solution recipe and the incubation times adopted for chestnut embryogenic calli were also effective, in fact, the protoplasts obtained showed a perfect spherical shape and no aggregates of undigested cells were detected.

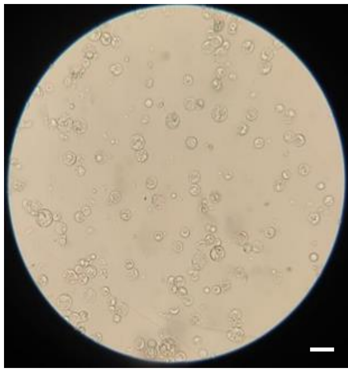

Figure 1: Protoplasts isolated from *C. sativa* embryogenic calli observed under an optical microscope (Bresser TFM 201/301). Scale bar = 100 µm

Protoplast transformation

Protoplasts were transformed with GFP marker, to test the protoplast transfection efficiency. The protoplast transformation capacity allows the future protoplast transformation using RNPs.

The plasmid DNA penetrates directly into the protoplast cell by direct absorption, thanks to the PEG action, which makes the cell membranes permeable to DNA. PEG-mediated transformation technique, thanks to its simplicity and low cost application, has been used in several plant species (Shen *et al.,* 2014; Yoo *et al.,* 2007; Ohnuma *et al.,* 2008).

The plasmid pAVA393, containing the gene coding for the GFP synthesis, was selected as the candidate vector. This fluorescent protein is an excellent marker to test the transformation efficiency for the first time in a new species. Several species have been transformed using GFP marker, including *Elaeis guineensis* (Masani *et al.,* 2014), *Brassica oleracea* (Sun *et al.,* 2019) and *Cucumis sativus* (Huang *et al.,* 2013).

The GFP transformation efficiency was evaluated 5 days after the transfection process using the fluorescence microscope. The results revealed a good protoplasts conservation, which are intact and spherical even after 5 days from the transfection event; 51% of the protoplasts showed GFP expression, highlighting the protoplasts transformation capacity (Figure 2).

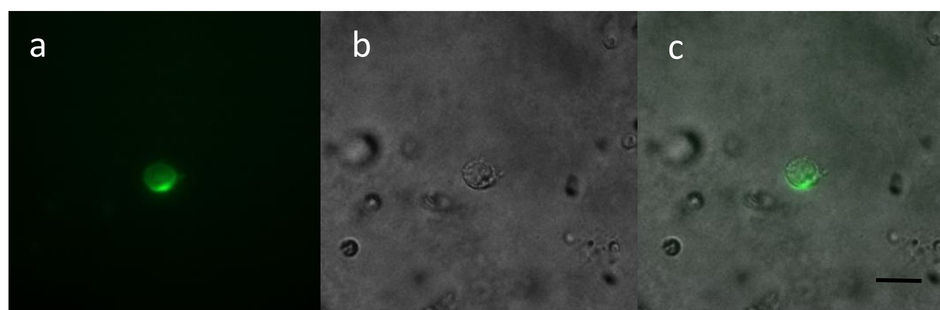

Figure 2: protoplast transformed with the pAVA393 plasmid observed with (a) blue light, (b) white light and (c) fusion of the two images, 5 days after transfection. Scale bar = 100 µm

Regeneration

The transformed protoplasts with pAVA393 plasmid were incubated in three different regeneration media.

The first cell divisions occurred after ten days and the microcolonies formation was observed after 30 days on the C2 medium (Figure 3). Figure 4 shows the embryogenic callus obtained three (a) and four (b) months on C2 medium. The embryogenic callus shows a white aspect with a size that in one month increased from 1 mm to about 3 mm.

.

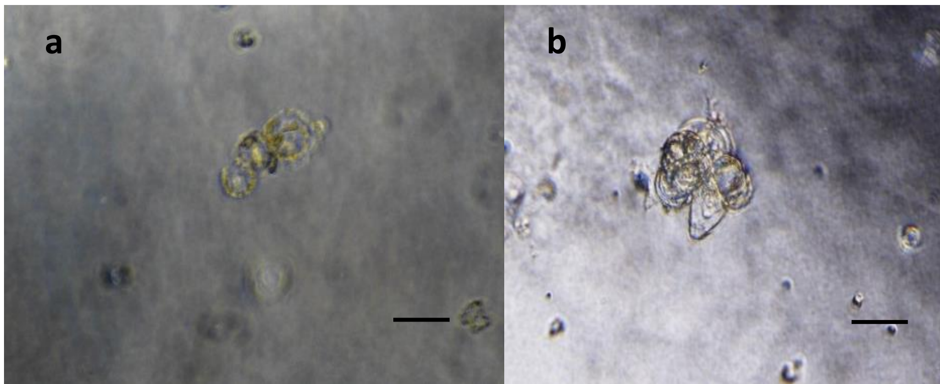

Figure 3: (a) Cellular divisions observed after 10 days; (b) Microcolony observed 30 days after insertion in C2 medium. Scale bar = 200 µm

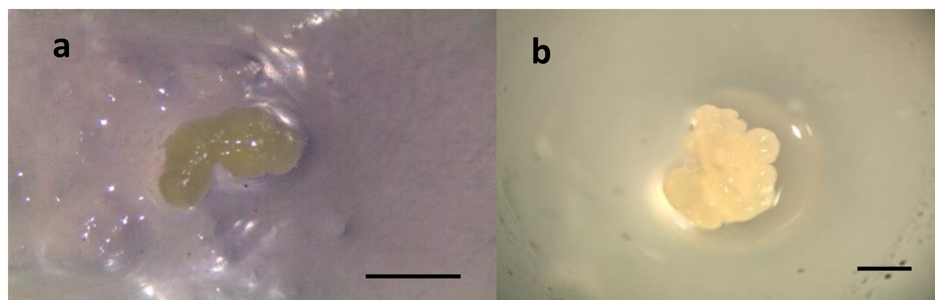

Figure 4: Embryogenic callus of *C. sativa* developed from protoplasts after three months of culture in C2 medium (a) and after four months (b). Observations made under the stereo microscope (Leica-Wild Heerbrugg M8). Scale bar = 1 mm

Conclusions

In conclusion, here we report the first protocol for *C. sativa* protoplasts isolation and transformation with GFP gene, starting from *C. sativa* embryogenic lines.
This protocol needs further improvements to increase the transformation efficiency. In addition, since the starting material for the experiments were embryos originated from seed that do not genetically match the cultivar of origin, the future work will consist in the development of a protoplast's isolation protocol from *in vitro* somatic explants such as leaves and young stems. Since it was possible to transform chestnut protoplasts, the future main goal will be to the protoplast transformation using RNP.

Final considerations and future perspectives

In this thesis, hazelnut and chestnut biotechnological studies have been carried out to increase the knowledge on genetic resources and innovative tools available for hazelnut and chestnut plant breeding.

In the Chapter I, we report the first whole-genome sequencing and assembly of *Corylus avellana* cultivar ‘Tonda Gentile delle Langhe’. The 10X Genomics Chromium Technology was firstly applied to obtain an high-quality sequence of the genome; then, the scaffolding process was generated using the cultivar 'Tombul' genome as a reference guide. The genome was assembled at chromosome-scale level with only 11% of sequences corresponding to unplaced scaffolds. Gene prediction and gene function detected 27,791 genes with an AED ≤ 0.4 and 92% of BUSCO completeness. This high-quality genome represents a valid resource for further hazelnut genetic studies to understand better the unclear genetic processes, such as self-incompatibility and recalcitrance to regeneration. Detecting genes involved in these processes will be useful for target gene-editing programs.

The Chapter II represents the first example of a study on S-genes (*mlo1, dmr6*, *dnd1*, and *pmr4*) in *C. sativa* and *C. crenata* following the infection by *P. cinnamomi* and *C. parasitica*.

Among S-genes, only the *mlo* gene was previously studied in woody plants, including rubber, poplar, apple, and grapevine (Liyanage *et al.,* 2020; Filiz *et al.,* 2018; Pessina *et al.,* 2016). No previous evidence of S-genes studies is found in the *Castanea* genus.

This work highlights the involvement of both p*mr4* and *dmr6* genes in the infection process; *pmr4* was observed to be over-expressed in the infections of both pathogens but only in *C. sativa* plants, which are highly susceptible. *Dmr6* was principally expressed, during *P. cinnamomi* infection, in *C. sativa* plants.

Since it has been observed that the S-genes silencing allows for greater disease tolerance (Santillán Martínez *et al.,* 2020; Sun *et al.*, 2016), *pmr4* and *dmr6* are two new possible candidates to be used in target gene editing programs, to increase chestnut tolerance to diseases.

The discovery of genes to be used in transformation programs, yet, is not useful without an efficient transformation protocol.

For this reason, Chapter III and Chapter IV are focused on optimizing the first CRISPR/Cas9 transformation protocol (III) and protoplast isolation and transformation (IV) in the *Castanea* genus.

In the Chapter III, the *pds* gene was selected for the first editing experiment in *C. sativa*. The silencing of *pds* causes an albino phenotype (Pan *et al.,* 2016; Qin *et al.,* 2007) and is commonly induced to validate transformation efficiency of a protocol. It allows to visually detect the explants successfully edited, which indeed show a white phenotype. The success of the *C. sativa* somatic embryo transformation opens the way to further genetic transformations using target genes involved in interesting biological processes or related to pathogen infection. Future perspectives will be the application of CRISPR/Cas9 on the new candidate S-genes, studied in Chapter II.

Chapter IV is the logical continuation of the research of Chapter III. Protoplasts are an excellent target for obtaining GMO-free plants.

Chestnut is a recalcitrant plant with a high level of polyphenols content; obtaining a good transformation protocol is very important goal for the future of chestnut breeding.

The ongoing research is further developing the results obtained transforming the protoplasts with RNP targeted on *pds* and on the susceptibility genes.

References *(General introducion, Chapter IV, Final conclusions and future perspectives)*

Ahmar, S., Gill, R.A., Jung, K.-H., Faheem, A., Qasim, M.U., Mubeen, M., Zhou, W. (2020) Conventional and Molecular Techniques from Simple Breeding to Speed Breeding in Crop Plants: Recent Advances and Future Outlook. *Int. J. Mol. Sci.,* 21*:* 2590.

Barrel, A., Beccaro, G.L., Bellini, E., Botta, R., Bonous, G. *et al.,* (2016) Capitolo: Il castagno. In: Atlante dei fruttiferi autoctoni nazionali. Vol. II: 557

Baratta, B., Botta, R., Cristofori, V., Farinelli, D., Nicolosi, E., Palombi, M.A., Petriccione, M., Piccirillo, P., Todde, M.G., Ughini, V., Valentini, N. (2016) Capitolo: Il nocciolo. In: Atlante dei fruttiferi autoctoni nazionali. Vol. III: 1261-1288

Beccaro, G., Alma, A., Bounous, G., Gomes-Laranjo, J. (2019) The Chestnut Handbook: Crop & Forest Management; CRC Press: Boca Raton, FL, USA.

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*: 1062.

Beltramo, C., Valentini, N., Portis, E., Torello Marinoni, D., Boccacci, P., Sandoval Prando, M. A., Botta, R. (2016) Genetic mapping and QTL analysis in European hazelnut (*Corylus avellana* L.). *Molecular Breeding,* 36.

Bernard, G., Gagneul, D., Alves Dos Santos, H., Etienne, A., Hilbert, J.L, Rambaud, C. (2019) Efficient Genome Editing Using CRISPR/Cas9 Technology in Chicory. *Int J Mol Sci.,* 20(5):1155.

Bertini, E., Tornielli, G. B., Pezzotti, M., Zenoni, S. (2019) Regeneration of plants from embryogenic callus-derived protoplasts of Garganega and Sangiovese grapevine (*Vitis vinifera* L.) cultivars. *Plant Cell, Tissue and Organ Culture* (PCTOC), 138: 239–246.

Bhattarai, G., Mehlenbacher, S. A. (2017) In silico development & characterization of trinucleotide simple sequence repeat markers in hazelnut (*Corylus avellana* L.). *PLOS ONE.,* 12. e0178061.

Boccacci, P., Botta, R. (2009) Investigating the origin of hazelnut (*Corylus avellana* L.) cultivars using chloroplast microsatellites. *Genetic Resources and Crop Evolution*, 56:851-859.

Botta, R., Molnar, T.J., Erdogan, V., Valentini, N., Torello Marinoni, D., Mehlenbacher, S.A. (2019) Hazelnut (*Corylus* spp.) Breeding. In: Al-Khayri J., Jain S., Johnson D. (eds) Advances in Plant Breeding Strategies: Nut and Beverage Crops. Springer, Cham.

Brandt, K. M., Gunn, H., Moretti, N., Zemetra, R. S. (2020) A Streamlined Protocol for Wheat (*Triticum aestivum*) Protoplast Isolation and Transformation With CRISPR-Cas Ribonucleoprotein Complexes. *Plant Science,* 11: 769.

Chen, X., Zhang, J., Liu, Q., Guo, W., Zhao, T., et al. (2014) Transcriptome sequencing and identification of cold tolerance genes in hardy *Corylus* species (*C. heterophylla* Fisch) floral buds. *PLoS One,* 9: e108604.

Chen, G., Abdeen, A.A., Wang, Y. *et al.* (2019) A biodegradable nanocapsule delivers a Cas9 ribonucleoprotein complex for in vivo genome editing. *Nat. Nanotechnol.,* 14: 974–980.

Colburn, B., Mehlenbacher, S. A., Sathuvalli, V. (2016) Development and mapping of microsatellite markers from transcriptome sequences of European hazelnut (*Corylus avellana* L.) and use for germplasm characterization. *Molecular Breeding,* 37: 16.

Conedera, M., Tinner, W., Krebs, P., de Rigo, D., and Caudullo, G. (2016) "*Castanea sativa* in Europe: distribution, habitat, usage and threats," in European Atlas of Forest Tree Species, eds J. San-Miguel-Ayanz, D. de Rigo, G. Caudullo, T. Houston Durrant, and A. Mauri (Luxembourg: Publications Office of the EU), e0125e0.

Contessa, C., Valentini, N., Caviglione, M., Botta, R. (2011b) Propagation of *Corylus avellana* L. by means of semi-hardwood cutting: rooting and bud retention in four Italian cultivars. *Eur J Hortic Sci.,* 76(5):170–175.

Corredoira, E., Montenegro, D., San José, M. C., Vieitez, A. M., and Ballester, A. (2004) Agrobacterium-mediated transformation of European chestnut embryogenic cultures. *Plant Cell Rep.,* 23: 311–318.

Corredoira, E., Ballester, A., Vieitez, F. J., Vieitez, A. M. (2005) Somatic Embryogenesis in Chestnut. *Plant Cell Monograph.,* vol. 2.

Corredoira, E., Valladares, S., Allona, I., Aragoncillo, C., Vieitez, A. M., and Ballester, A. (2012) Genetic transformation of European chestnut somatic embryos with a native thaumatin-like protein (CsTL1) gene isolated from *Castanea sativa* seeds. *Tree Physiol.,* 32: 1389–1402.

Corredoira, E., San José, M. C., Vieitez, A. M., Allona, I., Aragoncillo, C., and Ballester, A. (2016) Agrobacterium-mediated transformation of European

chestnut somatic embryos with a *Castanea sativa* (Mill.) endochitinase gene. *New Forests,* 47: 669–684.

Corredoira, E., Merkle, S. A., Martínez, M. T., Toribio, M., Canhoto, J. M., Correia, S. I., et al. (2019) Non zygotic embryogenesis in hardwood species. CRC Crit. Rev. *Plant Sci.,* 38: 29–97.

Dilzahan, H. A., Okamura, A., Calubaquib, M. A., Organo, N. D., Kobayashi, M., & Cruz, A. F. (2021) Function and Characterization of Fungal Communities in Chestnut Soils (*Castanea crenata*) of Kansai Region, Japan. *Asian Journal of Soil Science and Plant Nutrition*, *7*(4): 63-75.

FAOSTAT 2021: http://www.fao.org/faostat/en/#data/QC, accessed 20 November 2021.

Fernandes, P., Tedesco, S., Vieira da Silva, I., Santos, C., Machado, H., Lourenço Costa, R. (2020) A New Clonal Propagation Protocol Develops Quality Root Systems in Chestnut. *Forests*, 11*:* 826.

Filiz, E., Vatansever, R. (2018) Genome-wide identification of mildew resistance locus O (MLO) genes in tree model poplar (*Populus trichocarpa*): Powdery mildew management in woody plants. *Eur. J. Plant Pathol.*, 152: 95–109.

Freitas, T.R., Santos, J.A., Silva, A.P., Fraga, H. (2021) Influence of Climate Change on Chestnut Trees: A Review. *Plants*, 10*:* 1463.

Gürcan, K., Mehlenbacher, S. A., Botta, R., Boccacci, P. (2010) Development, characterization, segregation, and mapping of microsatellite markers for European hazelnut (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. *Tree Genetics Genomes,* 6: 513-531.

Gürcan, K., Mehlenbacher, S. A. (2010) Development of microsatellite loci for European hazelnut (*Corylus avellana* L.) from ISSR fragments. *Molecular Breeding*, 26: 551-559.

Huang, H., Wanga, Z., Chenga, J., Xhaoa, W., Li, X., Wanga, H., Zhanga, Z., Sui, X. (2013) An efficient cucumber (*Cucumis sativus* L.) protoplast isolation and transient expression system. *Scientia Horticulturae*, 105: 206-212.

Jung, T., Pérez-Sierra, A., Durán, A., Horta Jung, M., Balci, Y., Scanu, B. (2018) Canker and decline diseases caused by soil- and airborne *Phytophthora* species in forests and woodlands. *Persoonia*, 40:182-220.

Kasapligil, B. (1972). A bibiography on Corylus (Betulaceae) with annotations. *Annual Report of the Northern Nut Growers Association*, 63: 107-162.

Kuzminsky, E., Meschini, R., Terzoli, S., Scarascia-Mugnozza, G. (2016) Isolation of Mesophyll Protoplasts from Mediterranean Woody Plants for the Study of DNA Integrity under Abiotic Stress. *Frontiers in Plant Science*, 7(75).

Limera, C., Sabbadini, S., Sweet, J.B., Mezzetti, B. (2017) New biotechnological tools for the genetic improvement of major woody fruit species. *Front. Plant Sci.*, 8: 1418.

Lione, G., Giordano, L., Turina, M., Gonthier, P. (2020) Hail-nduced Infections of the Chestnut Blight Pathogen *Cryphonectria parasitica* Depend on Wound Size and May Lead to Severe Diebacks. *Phytopathology*, 110: 1280–1293

Liyanage, K.K., Khan, S., Herath, V., Brooks, S., Mortimer, P.E., Nadir, S., Hyde, K.D., Xu, J. (2020) Genome Wide Identification of the MLO Gene Family Associated with Powdery Mildew Resistance in Rubber Trees (Hevea brasiliensis). *Trop. Plant Biol.,* 13: 331–342.

Lucas, S.J., Kahraman, K., Avşar, B., Buggs, R.J.A., Bilge, I. (2021) A chromosome-scale genome assembly of European hazel (*Corylus avellana* L.) reveals targets for crop improvement. *Plant J.,* 105(5):1413-1430.

Ma, H., Lu, Z., Liu, B., Qiu, Q., Liu, J. (2013) Transcriptome analyses of a Chinese hazelnut species *Corylus mandshurica*. *BMC Plant Biol.,* 13: 152.

Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., Velasco, R. and Nagamangala Kanchiswamy, C. (2016) DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. *Front. Plant Sci.,* 7:1904.

Martini, D., Godos, J., Marventano, S., Tieri, M., Ghelfi, F., Titta, L., Lafranconi, A., Trigueiro, H., Gambera, A., Alonzo, E., et al. (2021) Nut and legume consumption and human health: An umbrella review of observational studies. *Int. J. Food Sci. Nutr.*, 1–8.

Masani, M. Y. A., Noll, G. A., Parveez, G. K. A., Sambanthamurthi, R., Prufer, D. (2014) Efficient Transformation of Oil Palm Protoplasts by PEGMediated Transfection and DNA Microinjection. *Plos One,* 9(5).

Mehlenbacher, S.A., Brown, R.N., Nouhra, E.R., Go¨ kirmak, T., Bassil, N.V., et al. (2006) A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. *Genome,* 49:122–133.

Merkle, S. A., Vieitez, F. J., Corredoira, E., and Carlson, J. E. (2020). "Castanea spp. Chesnut," in Biotechnology of Fruit and Nut Crops, eds R. E. Litz, F. Pliego Alfaro, and J. I. Hormaza (Pondicherry, CABI Publishing), 206–237.

Micek, A., Godos, J., Cernigliaro, A., Cincione, R. I., Buscemi, S., Libra, M., Galvano, F., Grosso, G. (2021) Total Nut, Tree Nut, and Peanut Consumption and Metabolic Status in Southern Italian Adults. *International journal of environmental research and public health*, 18(4): 1847.

Milgroom, M. G., and Paolo Cortesi (2004) Biological control of chestnut blight with hypovirulence: a critical analysis. *Annu. Rev. Phytopathol.,* 42: 311-338.

Muñoz-Adalia, E.J., Meijer, A., Abel, J., Colinas, C., Aletà, N., Guàrdia, M. (2021) Evaluation of Chestnut Susceptibility to *Cryphonectria parasitica*: Screening under Controlled Conditions. *Agriculture*, 11*:* 1158.

Ochatt, S., Lionnenton, E., Delaitre, C., Kahane, R. (2005) One team, PCMV, and one approach, in vitro biotechnology, for one aim, the breeding of quality plants with a wide array of species. *Crops: Growth, Quality and Biotechnology,* 1038-1067.

Ohnuma, M., Yokoyama, T., Inouye, T., Sekine, Y., Tanaka, K. (2008) Polyethylene glycol (PEG)-mediated transient gene expression in a red alga, *Cyanidioschyzon merolae* 10D. *Plant Cell Physiol.,* 49(1):117-20.

Osakabe, Y., Liang, Z., Ren, C. *et al.* (2018) CRISPR–Cas9-mediated genome editing in apple and grapevine. *Nat Protoc* 13, 2844–2863.

Öztürk, S., Balik, H., Balık, S., Kızılcı, G., Duyar, Ö., Doganlar, S., Frary, A. (2017) Molecular genetic diversity of the Turkish national hazelnut collection and selection of a core set. *Tree Genetics & Genomes*, 13.

Palmè, A. E., Vendramin, G. G. (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel*, Corylus avellana. Molecular Ecology,* 11: 1769-1779.

Pan, C., Ye, L., Qin, L. *et al.* (2016) CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. *Sci Rep* 6**:** 24765.

Pessina, S., Lenzi, L., Perazzolli, M., Campa, M., Dalla Costa, L., Urso, S., Valè, G., Salamini, F., Velasco, R., Malnoy, M. (2016) Knockdown of MLO genes reduces susceptibility to powdery mildew in grapevine. *Hortic. Res.,* 3: 16016.

Poovaiah, C., Phillips, L., Geddes, B., Reeves, C., Sorieul, M., Thorlby, G. (2021) Genome editing with CRISPR/Cas9 in *Pinus radiata* (D. Don). *BMC Plant Biol.,* 21(1):363.

Qin, G., Gu, H., Ma, L. *et al.* (2007) Disruption of *phytoene desaturase* gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res* 17, 471–482.

Rigling, D., and Prospero, S. (2018) *Cryphonectria parasitica*, the causal agent of chestnut blight: invasion history, population biology and disease control. *Molecular Plant Pathology,* 19.1: 7-20.

Rowley, E., Fox, S., Bryant, D., Sullivan, C., Priest, H., Givan, S., Mehlenbacher, S., Mockler, T. (2012) Assembly and Characterization of the European Hazelnut 'Jefferson' Transcriptome. *Crop Science,* 52: 2679.

Rowley, E. R., Vanburen, R., Bryant, E.R., Priest, H.D., Mehlenbacher, S.A., and Mockler, T.C. (2018) A Draft Genome and High-Density Genetic Map of European Hazelnut (*Corylus avellana* L.). bioRxiv.

Santillán Martínez, M.I., Bracuto, V., Koseoglou, E., Appiano, M., Jacobsen, E., Visser, R.G.F., Wolters, A.A., Bai, Y. (2020) CRISPR/Cas9- targeted mutagenesis of the tomato susceptibility gene PMR4 for resistance against powdery mildew. *BMC Plant Biol.,* 20: 284

Santos, C., Machado, H., Correia, I., Gomes, F., Gomes-Laranjo, J., Costa, R. (2015) Phenotyping *Castanea* hybrids for *Phytophthora cinnamomi* resistance. *Plant Pathol.,* 64: 901–910.

Scintilla, S., Salvagnin, U., Giacomelli, L., Zeilmaker, T., Malnoy, M. A., van der Voort, J. R., Moser, C. (2021) Regeneration of Plants from DNA-free Edited Grapevine Protoplasts. bioRxiv; 2021. DOI: 10.1101/2021.07.16.452503

Serrazina, S., Santos, C., Machado, H., Pesquita, C., Vicentini, R., Pais, M.S., Sebastiana, M., Costa, R. (2015) Castanea root transcriptome in response to *Phytophthora cinnamomi* challenge. *Tree Genet. Genomes,* 11: 19.

Shen, J., Fu, J., Ma, J., Wang, X., Gao, C., Zhuang C, Wan J, Jiang L. (2014) Isolation, culture, and transient transformation of plant protoplasts. *Curr Protoc Cell Biol.,* 63: 2.8.1-17.

Sun, K., Wolters, A.M.A., Loonen, A.E., Huibers, R.P., van der Vlugt, R., Goverse, A., Jacobsen, E., Visser, R.G., Bai, Y. (2016) Downregulation of Arabidopsis DND1 orthologs in potato and tomato leads to broad-spectrum resistance to late blight and powdery mildew. *Transgenic Res.,* 25: 123–138.

Sun, B., Yuan, Q., Zheng, H., Hiang, S., Jiang, M., Wang, M., Chen, Q., Li M., Zhang, Y., Luo, Y., Gong, R., Zhang; F., Tang, H. (2019) An Efficient and Economical Protocol for Isolating, Purifying and PEG-Mediated Transient Gene Expression of Chinese Kale Hypocotyl Protoplasts. *Plants,* 8: 385.

Tilman, D., Balzer, C., Hill, J., Befort, B.L. (2011) Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U S A*, 108(50):20260-4.

Tombesi, S., Botta, R., Valentini, N. et al (2017) Piattaforma varietale e orientamenti produttivi per i nuovi impianti. *Frutticol.,*81 (1/2):23–29.

Torello Marinoni, D., Valentini, N., Portis, E., Acquadro, A., Beltramo, C., Mehlenbacher, S. A., Mockler, T. C., Rowley, E. R., Botta, R. (2018) High density SNP mapping and QTL analysis for time of leaf budburst in *Corylus avellana* L. *PLoS One*, 13(4):e0195408.

Torello Marinoni, D., Nishio, S., Valentini, N., Shirasawa, K., Acquadro, A., Portis, E., Alma, A., Akkak, A., Pavese, V., Cavalet-Giorsa, E., Botta, R. (2020) Development of High-Density Genetic Linkage Maps and Identification of Loci for Chestnut Gall Wasp Resistance in *Castanea* spp. *Plants*, *9*: 1048.

Valentini, N., Me, G. (1999) Nuove selezioni di nocciolo per usi industriali. *Frutticolt.,* 61(11):36–38.

Valentini, N., Me, G., Vallania, R., Zeppa, G. (2001a) New hazelnuts selections for direct consumption. *Acta Hort.*, 556:103–108.

Valentini, N., Portis, E., Botta, R., Acquadro, A., Pavese, V., Cavalet Giorsa, E. and Torello Marinoni, D. (2021) Mapping the Genetic Regions Responsible for Key Phenology-Related Traits in the European Hazelnut. *Front. Plant Sci.,* 12:749394.

Woo, J.W., Kim, J., Kwon, S.I., Corvalán, C., Cho, S.W., Kim, H., Kim, S.G., Kim, S.T., Choe, S., Kim, J.S. (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol.,* 33(11):1162-4.

Yoo, S., Cho, Y., Sheen, J. (2007) *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols,* 2(7).

Acknowledgments

Vorrei ringraziare tutti coloro che mi hanno sostenuta in questi 3 anni fatti di gioie, soddisfazioni e duro lavoro.
Un ringraziamento particolare al Prof. Roberto Botta che mi ha sempre incoraggiata a non mollare e a credere in me stessa e alla Prof.ssa Daniela Torello Marinoni che è sempre stata una spalla a cui appoggiarmi durante i momenti difficili. Grazie per avermi permesso di intraprendere questo percorso e per avermi aiutata a crescere professionalmente e personalmente.
Un ringraziamento speciale al Prof. Andrea Moglia perché senza il suo supporto non ce l'avrei mai fatta! È stato una guida molto importante e con il suo aiuto mi ha permesso di superare le difficoltà e raggiungere gli obiettivi.
E grazie ai prof. Alberto Acquadro e Lorenzo Barchi per aver cercato di far entrare qualche piccolo briciolo di bioinformatica nel mio cervello, sempre con pazienza e piena disponibilità. Grazie a tutto il gruppo di genetica agraria per avermi sempre accolta come se fossi parte del team. Grazie al Prof. Gonthier e a tutto il gruppo di patologia forestale.
Un ringraziamento speciale va ad Anna, la mia compagna di follie scientifiche. Grazie per essere mia amica e per avermi ascoltata e supportata sempre.
Grazie alla mia compagna di viaggio Paola, questi tre anni in laboratorio sono stati stupendi grazie a te! Durante i momenti difficili e anche di gioia, ho sempre potuto contare su di te, spero di poter condividere altrettanti momenti felici con te. Grazie a Nadia per avermi guidata con spirito critico a fare le scelte migliori per me. Grazie a tutto il gruppo di colture arboree per avermi accolta.
Grazie a te Eric, senza di te non avrei mai fatto il dottorato. Mi ricordo ancora quando ti dicevo: "io il dottorato?? Ma va, non ne sono in grado, è troppo difficile per me" e tu ostinato mi hai sempre sostenuta e mi hai aiutata a credere in me stessa. So che non sono stati anni facili, pieni di insicurezze e momenti stressanti ma tu hai saputo sempre stare al mio fianco in ogni caso. Spero di poter

condividere la vita con te e gioire insieme di tutti i traguardi futuri, io te e la nostra 'picla Tac'.
Grazie alla mia mamma, al mio pippo, nonnino e nonnina. Voi siete le mie gioie. Grazie per avermi sostenuta e incoraggiata, siete sempre stati il mio supporto più grande. Senza di voi non ce l'avrei mai fatta! Mi ricordo ancora quel giorno in cui nonnino mi hai detto: "ma quindi tu fai il CRISPR? quello si che è bello! Sono proprio un nonno fortunato".
Grazie alle amiche di sempre. Grazie a te Vale che ancora oggi dopo 12 anni sei sempre qui con me, sono felice di poter essere al tuo fianco oggi e raggiungere con te questi bellissimi traguardi. Grazie a te Ade per essermi stata accanto in tutti questi anni e per avermi "obbligata" a fare palestra. Hai salvato la mia salute fisica e mentale!! Siete amiche importanti per me e spero di rendervi orgogliose oggi. Grazie alla più bella scoperta che la palestra potesse portarmi, la mia Nene. Grazie a Cri per essere la mia 'Mamacita' del cuore, sempre premurosa e attenta. Grazie a Chiara per tutto l'affetto che mi doni ogni giorno. Grazie a Chiara Ar. per essere al mio fianco e per condividere con me l'ossessione dei panettoni e a Marta. Senza le tue battute e i tuoi tick tock le giornate sarebbero state grigie.
Grazie ad Elena, Teresa, Rita e Fernanda. Da voi ho imparato tantissimo e ve ne sono molto grata. Mi avete accolta con gioia e avete reso il mio periodo all'estero unico ed indimenticabile. Grazie alle mie amiche Andreia, Patricia, Belen e Vanesa che mi hanno accompagnata in questo viaggio straordinario.
Grazie a tutti coloro che ci sono e ci saranno sempre.
Questo è un momento molto bello e spero che tutti voi siate felici di condividerlo con me.

Appendix A

Reference for Chapter I

Pavese V., Cavalet-Giorsa E., Barchi L., Acquadro A., Torello Marinoni D., Portis E., Lucas S.J., Botta R. (2021) Whole-genome assembly of Corylus avellana cv "Tonda Gentile delle Langhe" using linked-reads (10X Genom ics). G3 Genes|Genomes|Genetics, Volume 11, 7. https://doi.org/10.1093/g3journal/jkab152

Reference for Chapter II

Pavese V., Moglia A., Gonthier P., Torello Marinoni D., Cavalet-Giorsa E., Botta R. (2021) Identification of susceptibility genes in Castanea sativa and their transcription dynamics following pathogen infection. Plants, 10, 913. https://doi.org/10.3390/plants10050913

Reference for Chapter III

Pavese V., Moglia A., Corredoira E., Martínez M.T., Torello Marinoni D., Botta R. (2021) First report of CRISPR/Cas9 Gene Editing in Castanea sativa Mill. Front. Plant Sci. 12, 728516. https://doi.org/10.3389/fpls.2021.728516

Appendix B

Pubblications

-Valentini N., Portis E., Botta R., Acquadro A., Pavese V., Cavalet Giorsa E., Torello Marinoni D. (2021) Mapping the Genetic Regions Responsible for Key Phenology-Related Traits in the European Hazelnut. *Front. Plant Sci.* 12: 749394. doi: 10.3389/fpls.2021.749394

-Torello Marinoni D., Nishio S., Valentini N., Shirasawa K., Acquadro A., Portis E., Alma A., Akkak A., Pavese V., Cavalet-Giorsa E., Botta R. (2020) Development of high-density genetic linkage maps and identification of loci for chestnut gallwasp resistance in *Castanea* spp. Plants 9, 1048. https://doi.org/10.3390/plants9081048

-Lamberti C., Nebbia S., Balestrini R., Marengo E., Manfredi M., Pavese V., Cirrincione S., Giuffrida M.G., Cavallarin L., Acquadro A., Abbà S. (2020) Identification of a caleosin associated with hazelnut (*Corylus avellana* L.) oil bodies. Plant Biology, 22, 3, 404-409. https://doi.org/10.1111/plb.13096.

-Torello Marinoni D., Nishio S., Valentini N., Acquadro A., Portis E., Akkak A., Ruffa P., Pavese V., Botta R. (2019) Costruzione di una mappa di linkage ad alta densita' per il breeding in castagno. ACTA Italus Hortus 25, Pergine Valsugana (TN), 11-14 giugno 2019

-Poster presso XIII Giornate Scientifiche SOI (22-23 giugno 2021) Pavese V., Moglia A., Gonthier P., Torello Marinoni D., Cavalet-Giorsa E., Botta R. Identification of susceptibility genes in *Castanea sativa* and their transcription dynamics following pathogen Infection

-Poster presso XVII Congreso Hispano-Luso de Biología de Plantas y XXIV Reunión de la Sociedad Española de Biología de Plantas (7-9 luglio 2021) Pavese V., Moglia A., Corredoira E., Martínez M.T., Torello Marinoni D., Botta R. First report of CRISPR/Cas9 Gene Editing in *Castanea sativa* Mill.