DIPARTIMENTO SCIENZE DELLA VITA

DOTTORATO DI RICERCA IN SCIENZE DELLA VITA

CICLO XXXII

COORDINATORE Prof. Massimo Valoti

TITOLO DELLA TESI

VALORISING TUSCAN VARIETIES OF SWEET CHERRY (*PRUNUS AVIUM* L.) THROUGH AN INTEGRATED APPROACH: GENE EXPRESSION ANALYSIS COUPLED TO METABOLOMICS AND PROTEOMICS

SETTORE SCIENTIFICO-DISCIPLINARE: BIO/01

TUTOR: Dr. *Claudio Cantini*

Istituto per la BioEconomia (IBE CNR)

Dipartimento di Scienze BioAgroAlimentari

DOTTORANDO: Dr. *Roberto Berni*

A.A. 2019/2020

Summary

Acknowledgments

Writing a PhD thesis is certainly an exciting adventure. These were years of hard work, sometimes difficult because in Science things do not always go as initially predicted, but also constellated of pleasant moments, both personal and professional.

I would not have been able to reach this point in my scientific formation without the help of Dr. Claudio Cantini and Prof. Giampiero Cai, who took me under their guidance and mentored me during my PhD course. Thank you so much for for having given me this great opportunity and for your constant support and help.

I would like to thank also Dr. Marco Romi because he contributed substantially to this thesis and because he was among the first to propose to me to start a PhD course.

I thank Dr. Massimo Nepi and Dr. Massimo Guarnieri for providing me help with matters related to metabolite quantification (among other things) when I needed it.

I thank my laboratory Colleagues Dr. Cecilia Del Casino, Dr. Claudia Faleri, Dr. Lavinia Mareri, Dr. Chiara Piccini, Dr. Veronica Conti, Dr. Luigi Parrotta for keeping a nice and friendly atmosphere and for the fruitful scientific exchanges.

I had the chance to spend a period abroad during my PhD course. For this, I thank the Luxembourg Institute of Science and Technology for having hosted me. A special thank goes to Dr. Gea Guerriero, Dr. Jenny Renaut, Dr. Sophie Charton, Dr. Sébastien Planchon, Dr. Jean-Francois Hausman and Dr. Sylvain Legay. I discovered not only that Luxembourg is a place where excellent science is made, but also that there football is seriously played!

I thank the Region Tuscany for having supported me financially during my PhD course and for having given me the possibility of seeing how science is done abroad.

Finally, I thank my family and friends for their constant presence, comfort and support.

Without all of you, this work would not have been possible.

Abstract

A strong interest is given to sustainable agricultural practices valorising the local biodiversity of territories and promoting, at the regional level, products deriving from locally-grown non-commercial varieties. These varieties designate plants that were grown in the past, but that have not been subject to a market-driven pressure, thereby falling out of agricultural interest. Ancient varieties are considered wild, since they can thrive in soils where the human input is minimal and are sources of interesting agronomic characters, notably higher resilience to adverse environmental conditions. Italy has started conservation programs aimed at preserving the local biodiversity via *ex situ* collections; in this respect Tuscany is at the forefront, with a regional law (law 64/04) created for the recovery and preservation of local varieties of woody and herbaceous plant species. The fruits produced by ancient varieties can diversify the options offered by local markets and are equally rich sources of functional molecules with nutraceutical value. With a view to valorize non-commercial regional sweet cherry varieties of Tuscany, the present PhD thesis characterised the Tuscan repertoire of *Prunus avium* L., by using complementary approaches comprising genotyping, gene expression analysis, as well as metabolomics and proteomics. The fruits of sweet cherry are rich sources of health-promoting compounds. Among the beneficial phytochemicals are different classes of phenolic compounds produced via the phenylpropanoid pathway, an important hub of plant secondary metabolism.

The content of phenolic compounds was measured in six ancient *P. avium* varieties from Tuscany (Benedetta, Carlotta, Crognola, Maggiola, Moscatella and Morellona) sampled at the stage of commercial harvest (ca. 60 dpa, days post anthesis). The results cover three consecutive years of sampling (2016-2017-2018) to highlight the variations in functional molecule content linked to the environmental conditions encountered during the years. Higher levels of phenolic compounds were quantified in the Tuscan sweet cherries, as compared to the fruits of the commercial variety Durone. In particular, the varieties Crognola and Morellona are the top-producers of phenolics. The expression analysis reveals the presence of different expression patterns of the genes involved in the early and late steps of the phenylpropanoid pathway, both among the genotypes and the years. Finally, the soluble proteomes of the two highest producers of functional molecules, i.e. Crognola and Morellona, reveal differences in proteins related to cell wall remodelling, redox balance and stress response. In conclusion, the results presented draw attention on the nutraceutical potential of ancient *P. avium* varieties from Tuscany and stress the importance of a multidisciplinary approach encompassing gene expression and metabolomic analyses.

Introduction

Biodiversity and agrobiodiversity

The term *biodiversity* was defined at the Rio de Janeiro environmental conference in 1992 as the *variability of organisms living in terrestrial, marine, and aquatic ecosystems and biological complexes of which they are a part*. The different species inhabiting the various ecosystems, together with their genetic traits, are fundamental aspects defining the variability in biological systems/habitats[1]. Natural events, climatic changes and anthropogenic activities severely impact the ecosystems[2] by negatively affecting the biodiversity and favouring events of genetic erosion with subsequent progressive loss of richness in species. In this context, anthropogenic activities are certainly the main cause of environmental alterations, namely due to industrial pollution and deforestation.

Since the Rio de Janeiro conference, countries around the world have made efforts to preserve biodiversity through legislations and regulations safeguarding biodiversity. In this respect, Italy has defined specific rules for each region[3]. Each Italian region has created a database that contains the phenotypic and genetic characteristics of all the indigenous varieties in the territory[4]. Italy showed significant interest in its own agrobiodiversity through the creation of laws specific for each region. Fideghelli and colleagues defined the Italian regional legislative model as an example of biodiversity protection and safeguard of agricultural species that are interesting in the European context[5]. Thanks to the application of local laws, recovery activities started that spread from provincial campaigns to regions and, eventually, to the entire country. The aim of these laws is to promote the preservation of the ancient plants in the regional territory, preventing the risk of genetic erosion to which some varieties are subjected. Preserving ancient varieties can also help the national agriculture, through the use of species with agricultural or forestry interest.

The diversity of the crop/seed species used in agricultural practices is defined by the term "agrobiodiversity". Human intervention is most evident in agriculture; for years, men have selected more productive species to satisfy the industrial demands. Consequently, less productive species were gradually abandoned and confined to wild lands away from agricultural fields, while these last were destined to massive crop production[6]. The loss of agrobiodiversity has gone for years hand in hand with the loss of territorial biodiversity. The potential use and valorization of these ancient plants has fostered their introduction into the commercial landscape and shifted the purpose of most laws towards the consideration of economic aspects[4]. Non-commercial local varieties show also characters that contribute to a wide genetic reservoir and linked to a better adaptation to environmental constraints.

How Tuscany preserves the regional plant biodiversity and regional germplasm banks

Tuscany is historically characterized by a rich plant biodiversity, well exemplified by the different fruit tree species. These wild species have played a significant role in nutrition, especially during times of weather extremes, thanks to their ability to withstand exogenous stresses, thereby representing a feedstock for human nutrition[7]. For many years these local plants have been neglected because they do not fulfil the commercial standard requirements. These plants represent a cultural heritage of the region and are now considered an agricultural source of unique characters, such as specific phenotypic traits like shape, colour and biochemical composition[4].

Tuscany has shown interest towards recovery programs aimed at preserving the local seeds/crops in dedicated repositories through the law 64/04[8]. It made an effort to valorise its own regional heritage of species, with the specific aim of protecting and recovering ancient autochthonous endangered species. Since a considerable number of classified/protected species are of agronomic interest, such an initiative fosters, on a long-term perspective, sustainable agronomic techniques relying on the local agrobiodiversity and inspires innovative agricultural practices compliant with green technologies. Since 1997, the Tuscan regional biodiversity repertoire has been periodically enriched with new plant and animal species found across the territory. The recovery program of Tuscany relies on the conservation and propagation of the germplasm which is guaranteed by specific institutes, called "Regional Germplasm Banks" (Banche Regionali del Germoplasma).

The Banks dedicated to plants avoid any form of contamination, alteration, or degradation through *ex situ* conservation via seed banks and dedicated fields located in different geographical locations. Additionally, there are *in situ* conservation programs maintaining the plants in their original environment, which ensures their optimal growth.

The Regional Germplasm Banks are divided into various sections distributed throughout the region (Figure 1), which combine the information into a single database containing the descriptive data of the species. For plants, phenotypic and genetic descriptions are supported by photos (e.g. fruits and leaves of each variety) and explicatory cards of accessions.

The regional database is organized in 5 sections, as follows:

1. native animal genetic resources,
2. woody and fruit species,
3. ornamental and flowering species,
4. herbaceous species,
5. species of interest for forestry.

The inclusion of a species into the database follows a specific process, composed of two steps. The first one classifies species and varieties according to the morphological aspect. Subsequently, the classification procedure is based on plant genetic features, as determined by microsatellites (Simple Sequence Repeats, SSR, see section "Nutraceutical and genetic analysis for biodiversity conservation")[9]. This second step discriminates the varieties and collects genetic data that are subsequently made available in the regional dataset.

Nowadays there are about 877 plant and animal species registered in the Regional Germplasm Banks; notably, some of these (751) are subjected to genetic erosion risks.

The data reported in the Regional Plant Database underline the immense heritage of ornamental and fruit plants cultivated in the past and present across the whole Tuscan territory. The past uses were multiple: consumption of fruits, wood and fodder, to signal borders, provide shelter to birds, support other plants and wild animals[10].

These repositories foster the rediscovery of the seasonality of fruits and vegetables: taking these features into account would guarantee products of high quality to consumers, unlike what happens in large retail, where the use of artificial preservatives negatively affects the organoleptic qualities of end-products. The use of ancient varieties paves the way to new commercial local strategies that different Italian regions have already adopted[6] with the ultimate goal of manufacturing food products deriving from the cultivation of local varieties.

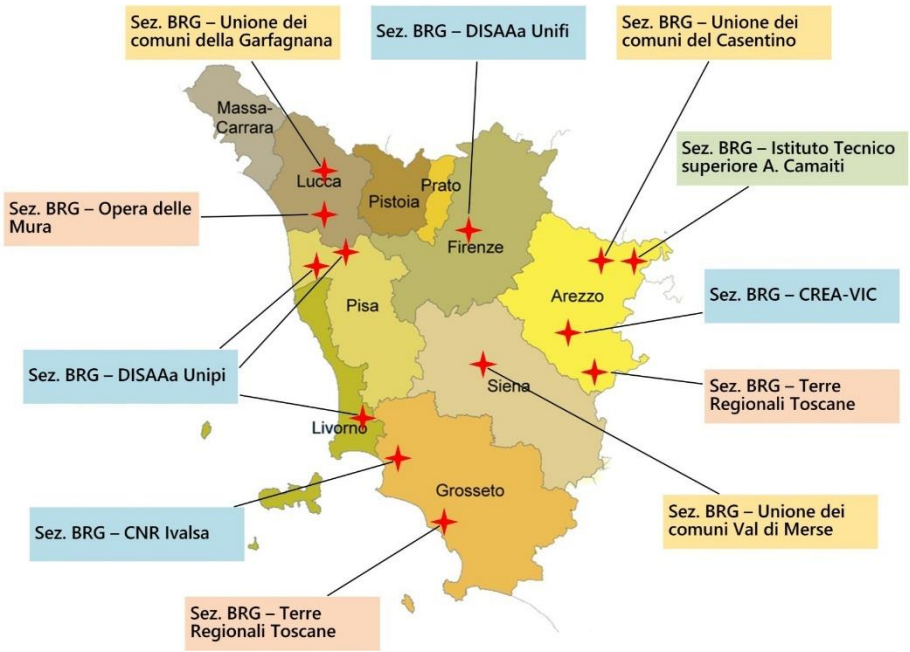


Figure 1: Locations of the Regional Germplasm Banks across Tuscany.

Valorisation of non-commercial autochthonous varieties

The loss of habitat diversification is a major threat linked to the ever-growing industrialization, climatic changes and land-use. Future forecasts highlight the dramatic scenarios of habitat loss with the increasing use of wild lands for the massive cultivation of man-selected plants[12]. Landscape and plant biodiversity are directly linked: the loss of one involves the loss of the other. The cultivation of native crops can play an important role in the preservation of both, while more traditional agricultural practices, especially if intense, can have negative consequences because of the massive use of land resources to promote large-scale products. Therefore, a well-conceived agricultural management based on the regional valorisation of autochthonous species, can preserve biodiversity and local ecosystems[13]. Native plants are also useful to maintain the local soil microbiota and composition, thereby preserving an optimal interaction between plants and microorganisms. Lange and colleagues reported the importance of plant diversity in the microbial soil composition with indirect effects also on C storage of the whole ecosystem[14].

Besides these important ecological aspects, there are other reasons inspiring researches on ancient varieties. Hereafter, two aspects will be considered: environmental impact linked to specific physiological characteristics and content of bioactives (i.e. functional molecules displaying a biological effect).

Commercial fruits fulfil the requirements of the market demand, but often require chemical treatments to guarantee maximum production. However, chemicals are well-known as cause of environmental pollution and the limitations in their use is one of the most discussed topic in agriculture[15]. The use of ancient crops could provide a solution to this problem. The ancient varieties have developed physiological mechanisms enabling them to thrive in a wild context. Some authors indeed demonstrated the lower environmental impact of ancient varieties per unit of cultivated land[16].

The local ancient fruits show unique characters not present in those destined to large distribution. Indeed, some authors have emphasized the unconventional qualitative traits of ancient varieties through sensory analyses; these fruits are different in shape and colour from the commercial ones, they display peculiar organoleptic features and high nutritional value[17]. Considerable differences were also discovered in qualitative and quantitative profiles of the bioactive compounds (e.g. polyphenols), a sought-after feature for the consumer, given the antioxidant potential of these molecules[18]. Hence, the ancient varieties have characteristics that encourage the consumption of their fruits and their introduction in the daily diet[4].

Considering the scientific value of proposing new food products with enriched functional aspects, ancient varieties are valuable resources that can help improve the daily diet via the supply of molecules with nutraceutical value[11].

Genetic analyses for biodiversity conservation and food traceability

Genetic analyses are widely used methods for the classification of plant species and are important tools for field studies. The use of genetic markers is indeed an accurate method to detect DNA polymorphic regions, i.e. discontinuous genetic variations allowing the discrimination of individuals belonging to the same species. Polymorphisms can be of several types: deletions, insertions, inversions, single base pair changes and are used as discriminative genetic markers. Among these, SSR (Single Sequence Repeats, also referred to as *microsatellites*) are widely used markers in genetic studies that may discriminate accessions through the amplification of specific genomic regions. SSR regions are now considered fundamental for many biological studies (e.g. determination of gene gain or loss, or the regulation of transcription and translation) and any alteration of these region may affect the susceptibility to diseases in plants[8,13]. The advantages of this type of technique are the applications to a wide range of organisms and the discrimination due to the high specificity, which make SSR very useful in agronomic studies.

The use of SSR for DNA fingerprinting generates a numerical value corresponding to the length of the amplified fragment and it is associated to each single variety, thereby univocally recognizing the accessions. Results are then shown with a graphical representation based on the analysis of the raw data representing the genetic distance between the studied genotypes. In literature, there are many examples of the SSR analytical power, even for ancient Tuscan fruits (such as apples, pears, peaches and olives, stored in a section of the Tuscan Regional Germplasm Bank named “Il Campino”; (http://www.provincia.siena.it/var/prov/storage/original/application/298a1f0c1d8e0f95659cfd3c5b00f880.pdf).

For example, Cantini and colleagues characterised 150 Tuscan olives (*Olea europaea* L.) belonging to the regional germplasm using 12 SSR markers. The authors proved the high potential of this method when used in combination with morphological analyses[19].

The SSR markers are also widely used in food quality analyses for applications in product traceability. The increased customers’ awareness of product quality requires the development of methods which are able to certify the food provenance and composition[20]. For instance, the use of genetic markers can detect frauds in commercial food products through PCR amplification and, therefore, can identify contaminations. Numerous cases of food frauds have been discovered mostly in transformed food products thanks to SSR. The use of fingerprinting ensures efficacy in a wide range of food products (both animal- and plant-derived)[21–23].

High-quality plant-derived food products need to be fully traceable and should also contain beneficial bioactive molecules with nutraceutical value. In this perspective, local non-commercial plants are interesting, because they display resistance to exogenous stresses and also contain high amounts of

bioactive molecules. Therefore, they fulfil the requirements needed to manufacture high quality and fully traceable food products.

Nutraceuticals and functional biodiversity

The impact of nutraceuticals on human health is one of the most important scientific discoveries of the last decades[24]. Molecules, such as polyphenols and antioxidants, are naturally present in plants and impact the human health through the diet. Thanks to their particular chemical structures, polyphenols are able to counteract reactive oxygen species (ROS) and are thus considered natural scavengers. For instance, the antioxidant property of flavonoids is due to their hydroxyl groups which donate hydrogen and an electron to radicals, thereby stabilizing them[25]. The role of these molecules is directly linked to the decreased susceptibility in developing cardiovascular diseases[26].

Red fruits like cherries typically contain high levels of anthocyanins and their dietary intake could help in the prevention of cardiovascular diseases[27]. From a nutraceutical point of view, it is fundamental to look for fruits that have a higher content of functional compounds and to study the biosynthetic pathways leading to the formation of these molecules. Ancient wild varieties represent an innovative source of genetic characters and may be considered as important bioresources for the valorization of the local economy. Genetic and functional studies of these species shed light on the potential of underused or non-commercial fruits with the aim of protecting and rediscovering the local biodiversity. The study of biodiversity is moving from the simple concept of safeguarding and preservation to the use of these territorial varieties in order to exploit their features and evolve the concept of biodiversity towards functional biodiversity.

Bioactives are products of the plant secondary metabolism and partake in different functions, namely response to environmental constraints[28]. For their chemical structures, bioactive molecules (e.g. terpenoids, phenolic compounds, alkaloids and sulphur-containing compounds) have many medical and pharmaceutical applications and are considered functional. A strong body of evidence suggests that the consumption of these fruits could boost the immune defenses and lower the risk of developing diseases[29].

Polyphenols are a class of phenolic compounds and are an example of functional molecules that have multiple roles in plants. Phenolics are typically stored in fruits and are assumed by humans through the diet. In the body, polyphenols act as natural antioxidants that scavenge ROS, which are dangerous if over-accumulated in the tissues. Ancient varieties such as cherries, onions, tomatoes, apples, peaches, represent important sources of bioactive molecules that are often in higher content than the commercial counterparts[30]. The adaptation to the original territory, over the centuries, likely forced these plants to elaborate new strategies to survive in wild environments characterized by minimal/no

human impact. This adaptation may be (partly) due to epigenetic modifications affecting the secondary metabolism and, ultimately, the production of bioactive molecules. The fruits of such locally adapted ancient varieties can be classified as *functional food* in the light of the high content of bioactives beneficially impacting human health. In this sense, the regional plant biodiversity fulfils the nutraceutical requirement of the modern diet.

Definition of functional foods

The human diet is characterized by highly energetic molecules, but it also requires non-energetic compounds that are equally useful for cell functioning and for preserving the organism's health status. These functional molecules are represented by a wide variety of plant secondary metabolites, such as terpenoids, vitamins and polyphenols with antioxidant power. Indeed today's foods are not intended to only satisfy hunger and to provide necessary nutrients for humans, but also to prevent nutrition-related diseases and improve physical and mental well-being[31]. De Felice in 2005 defined these double action (i.e nutrition and health prevention) in foods as *nutraceuticals*, combining the features of *nutrition* and *pharmaceutical*[32]. Nutraceuticals are considered compounds between food and medicine, that may be present in dietary supplements, herbal products and processed foods. In summary, nutraceutical is any substance present in a food, or part of a food, with health beneficial effects, including the prevention of diseases[33,34].

Most of these functional compounds derive from the secondary metabolism of plants and are synthesized in response to (a)biotic stresses, as well as during normal physiological processes (e.g., monolignols for lignin biosynthesis during secondary cell wall formation[35]).

Food products enriched with natural ingredients and with the ability to positively affect one or more functions of the organism and/or reduce the risk of the diseases, are considered *functional food*. In this sense, the term was first defined in Japan in the 1980s, to indicate a component generally absent or present at reduced concentrations in other foods. According to the International Life Science Institute, functional food should be taken by healthy subjects as an integral part of a proper diet[36].

Plants synthesize a huge variety of secondary metabolites, with complex chemical composition, which are produced in response to different forms of (a)biotic stresses, as well as to fulfil important physiological tasks, like attracting pollinators, establishing symbiosis, providing structural components to lignified cell walls of vascular tissues[37]. Importantly, many of the secondary metabolites produced by plants are used by pharmaceutical industries (since these bioactive compounds trigger a pharmacological or toxicological effect in humans and animals), in cosmetics, nutrition, for the manufacture of drugs, dyes, fragrances, flavors, dietary supplements. Hence, both the scientific and industrial interest around plant secondary metabolites is enormous[38].

In recent decades, the needs of consumers in the food sector have changed significantly. Indeed, as De Felice and colleagues had predicted in 1995, a progressive increase of nutraceutical product consumption was observed[32]. Eskin and colleagues described over 470 nutraceutical products available with documented health benefits in 2005 and according to De Felice's predictions, Nasri presented a prospective value of 250 billion brands by 2018[39,40]. Although nutraceuticals constitute a high commercial value, they are currently not regulated by a legislation in Europe[41]. Indeed, the European legislation did not define the functional foods as foods and/or drugs, allowing thus the trade of non-certified products[36].

An effort in this sense should be made to understand which products may be considered as nutraceuticals. Furthermore, biological and chemical approaches should be required to monitor the functional molecule contents of the newly harvested fruits up to the finished products. In this sense, multi -*omics* (e.g. genomic, proteomic and metabolomic studies supported by in *vitro* and *in vivo* tests) will help unravel the actual content of nutraceutical compounds and may ultimately lead to the establishment of a legislation. These methods could allow the introduction of new products on the market as a source of beneficial characteristics and define products worthy from an industrial perspective[42].

The importance of –*omics* for the study of plant secondary metabolism

Plants synthesize a huge variety of secondary metabolites with complex chemical composition. Their chemical structures are rich and diverse and are responsible for their bioactivities. Although called *secondary*, this type of metabolism is crucial for the survival of plants, which are sessile organisms. Indeed, many of the plants' response to exogenous cues is mediated by specific secondary metabolites. Since these compounds are synthesized by specific classes of enzymes, which, in their turn, are translated from genes, it is clear that a comprehensive understanding of plant secondary metabolism requires knowledge of the three levels of biological complexity, i.e. genes-proteins-metabolites.

The content of bioactives in each species depends on the variety/cultivar analysed, on the growth condition and environmental constraints, as well as on the specific organ investigated[43,44]. This obviously adds further complexity to the study, characterisation and determination of the genes/enzymes involved in the biosynthesis of secondary metabolites. However, the recent progresses in the field of molecular analyses via Next-Generation Sequencing (NGS) provide unprecedented analytical depth, thereby opening the way to new perspectives in the study of the plant secondary metabolome. -*Omic* approaches applied to ancient local varieties can help shed light on the mechanism of adaptation to the territory and will enable the identification of those varieties

expressing the best characters for a potential exploitation. In this paragraph, examples of untargeted approaches (-*omics*) used in systems biology for the study of secondary metabolism in non-model plants are provided. What can be learned about non-model species can therefore be transferred to native (ancient) plants. Such old varieties, when compared to traditional commercial crops, can be considered non-model organisms.

For example, a transcriptomic and proteomic shot-gun approach was adopted in the ornamental plant *Peperomia obtusifolia* L. to study the production of benzopyrans derived from orsellinic acid, a phenolic acid usually found in fungi[45]. The analyses led to the identification of both the mevalonate and methylerythritol pathways as being active in the leaves and showed that terpenoid biosynthesis was the pathway with the highest number of enzymes identified.

Proteomics was very useful in studying fruit ripening in non-climacteric species, such as sweet cherry. A study compared ripe and unripe fruits of sweet cherry Hedelfinger and its somaclonal variant HS[46]: 39 differentially abundant proteins were identified, among which members related to cell wall remodelling and secondary metabolism. In yet another study, the treatment of unripe and ripe sweet cherries with salicylic acid (SA) revealed differences in proteins related to antioxidant production, stress response, heat shock proteins (HSPs) and dehydrogenases[47]: interestingly, only unripe fruits showed increased resistance against pathogens after SA treatment.

While many studies have used complementary -*omics* on model organisms[48,49], little is known about the proteomic/transcriptomic profiling of ancient fruits. Future efforts should be devoted to the analysis, via systems biology techniques merging -*omics*, of underutilized varieties (such as ancient local varieties) to compare the content of bioactive molecules with respect to commercial ones and understand the molecular basis of such differences. These studies will favour the diversification, on a local level, of the current market of fruits and vegetables and promote regional programs aiming at the preservation of the regional agrobiodiversity.

The importance of the plant secondary metabolism: the phenylpropanoid pathway as an example of central *hub*

The phenylpropanoid pathway (PPP) is a central metabolic *hub* providing precursors for important structural molecules, such as lignin, as well as molecules responsible for pigments, aroma and flavour of fruits/flowers. These macromolecules derive from the polymerization of polyphenols[50,51]. Over 8000 phenolic molecules have been identified and classified based on their chemical structure, biological function and origin[52]. From a biochemical point of view, polyphenols are characterized by aromatic rings linked to different chemical groups and they display specific functions depending on the elements bound. An initial classification was made by Manach and co-authors, who distinguished

phenolic acids, flavonoids, stilbenes and lignans[53]. Tsao and colleagues extended this classification by further ranking polyphenols in 8 groups (i.e. 1. phenolic acids; 2. flavonoids; 3. isoflavones, neoflavonoids and chalcones; 4. flavones, flavonols, flavanones and flavanonols; 5. flavanols and proanthocyanidins; 6. anthocyanidins; 7. polyphenolic amides; 8. other polyphenols), based on their aglycone structures[54]. Thanks to these particular chemical features, polyphenols are able to scavenge ROS, via the so-called hydrogen atom transfer (HAT) mechanism, whereby the -OH groups act as donors[55].

The PPP starts with the amino acid phenylalanine (Phe), derived from the shikimate pathway, which in plants is located in the chloroplasts[56]. The pathway is divided into three main parts, each comprising three different enzymes (Figure 2). The upper part of the pathway constitutes the general PPP and is characterized by the reactions of the enzymes phenylalanine ammonia lyase (PAL), cinnamate-4-hydroxylase (C4H) and 4-coumarate coenzyme A ligase (4CL).

PAL is a tetrameric enzyme with different isoforms encoded by a multigene family in plants [57–60]. For instance, 4 different *PAL* isoforms were found in the model plant *Arabidopsis thaliana* and at least 20 genes have been described in *Lycopersicon esculentum*[61]. Such a large number of genes witnesses the importance of PAL in plant physiology. For example, in thale cress, the analysis of *pal1* and *pal2* mutants revealed a more important role of PAL1 in the biosynthesis of phenylpropanoids. Indeed, genes related to phenylpropanoid biosynthesis had altered expression in *pal1* mutants and less extractable phenolics were recovered[62]. PAL is regulated by ubiquitination: PAL1 and PAL2 interact with 3 KFB (Kelch repeat F-box) proteins (namely KFB0-1, -20, -50) which control the proteolytic turnover through the ubiquitin-26S proteasome system[63].

4CL is the last enzyme intervening in the three shared steps of the general phenylpropanoid pathway and as such it channels activated thioesters of hydroxycinnamic acids to the synthesis of different phenylpropanoids[64]. Several *4CL* genes are present in plants, they respond to environmental cues and phytohormones and are classified in type I and II: the first ones are related to lignin biosynthesis, while the latter are related to the synthesis of phenylpropanoids other than lignin[64]. In thale cress, different *4CL* isoforms are implicated in the shunting of the metabolic flux: *AtCL1* is co-expressed with lignin biosynthetic genes, while *AtCL3* with flavonoid biosynthetic ones[64].

The central steps of PPP lead to the synthesis of flavonoids through the reactions of chalcone synthase, chalcone isomerase and flavanone 3-hydroxylase (*CHS*, *CHI*, *F3H*).

CHS is considered the gatekeeper of flavonoid biosynthesis[65]: it belongs to the plant polyketide synthase superfamily and it catalyses a condensation reaction of one activated thioester (e.g. *p*-coumaroyl-CoA) to three molecules of malonyl-CoA. Studies demonstrate that *CHS* is under control of the circadian rhythm and proteolytic regulation[65]. More specifically, *CHS* is negatively regulated

by a KFB protein mediating degradation and whose expression is developmentally controlled in *Arabidopsis*[66].

CHI acts as an isomerase, by transforming chalcone in flavanone. A recent study has revealed that CHI evolved from a non-catalytic ancestor related to fatty acid binding proteins and a non-catalytic protein, CHIL (chalcone isomerase-like)[67]. Small changes in the positioning of the substrates, as well as modifications in the positioning and flexibility of the catalytic residues with respect to the substrate, favoured the emergence of catalysis in CHI[67].

F3H catalyses the last step of the central portion of the pathway leading to the formation of dihydroflavonols. These are substrates of the last part of PPP catalysed by dihydroflavonol 4-reductase (DFR). DFR reduces dihydroflavonols to leucoanthocyanidins, which are substrates of the next enzymatic reactions leading to the synthesis of anthocyanidins[68].

Anthocyanidin synthase (ANS) belongs to the 2–oxoglutarate dependent dioxygenases family (OGD) and catalyses the conversion of colourless leucoanthocyanidins to coloured anthocyanidins. Mutations in *ANS* cause yellow pigmentation, as observed in yellow raspberry fruits, where a 5 bp insertion in the coding region was associated with the observed phenotype[69].

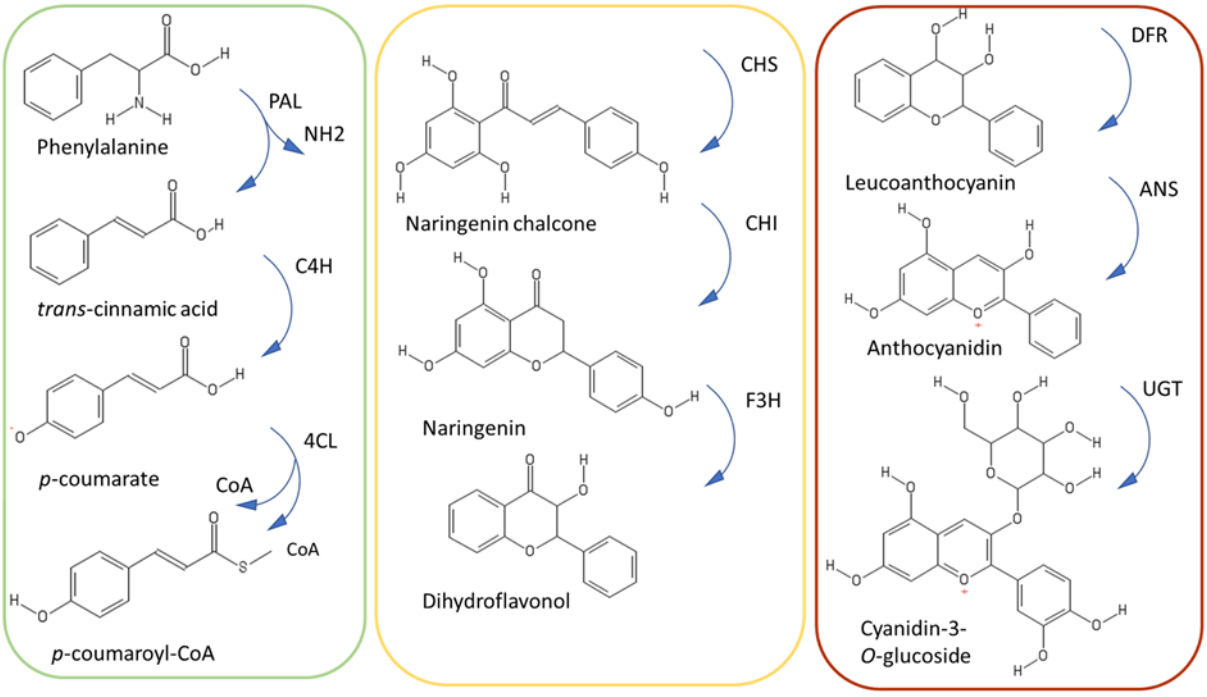


Figure 2: Simplified scheme of PPP, with the upper, central and late steps in green, yellow and brown, respectively. PAL: phenylalanine ammonia lyase, C4H: cinnamate 4-hydroxylase, 4CL: 4-coumarate CoA ligase, CHS: chalcone synthase, CHI: chalcone isomerase, F3H: flavanone 3-hydroxylase, DFR: dihydroflavonol reductase, ANS: anthocyanidin synthase, UGT: UDP-glycosyltransferase.

The transcriptional regulation of the PPP pathway

In order to fully understand the complexity of the plant secondary metabolism, it is here necessary to mention the importance of the regulatory factors driving gene expression. Transcription factors (TFs) are proteins with the ability to modulate, either alone or in combination, gene expression, by binding to specific sequences in the promoters of target genes. TFs are able to positively or negatively regulate the transcription of their targets[70]. TFs are characterized by modular structures in which specific regions are responsible for binding to the DNA using different domains (e.g helix loop helix, two cysteine-two histidine zinc finger, multi-cysteine zinc fingers) and activation regions triggering transcription[71].

Three different factors were found as regulators of PPP: MYB, bHLH and WD40 that actively interact with each other driving the pathway towards phenolic synthesis[72]. Baudry and colleagues described the role of the MYB/bHLH/WD complex in flavonoid biosynthesis during different plant developmental programs, such as trichome and root-hair organogenesis, as well as mucilage biosynthesis in the seed coat[73]. Basically, the interaction between MYB and bHLH proteins is the main factor triggering the activation of PPP (in particular flavonoid formation) and WDR proteins act by influencing the heterodimer (MYB/bHLH) activity[73,74].

Further yeast experiments strongly evidenced the direct interaction of TT1 (TRANSPARENT TESTA1, a WD40 family member) simultaneously with TT2 and TT8 (respectively a MYB and bHLH protein)[75]. In plants, WD40 proteins are well described as inducers of the anthocyanin pathway, through the activation of the BANYULS (BAN) gene. The interaction between MYB/bHLH and WD40 results in an a ternary complex fundamental for BAN activation, where WD40 partakes as chief component favouring the stabilization of the complex[75,76].

Xu and co-authors described the interaction of R2R3-MYB TFs and the R/B-like bHLH as the main event regulating gene expression. MYB directly interacts with bHLH through a conserved motif (D/E)LX2(R-K)X3LX6LX3R); this motif binds the N-terminal MYB-interacting region (MIR), which contains an arginine residue. Subsequently, WD40 is able to form the ternary protein complexes named MBW (MYB-bHLH-WD40)[77] (Figure 3). Xu and colleagues also suggested that MYB plays a regulatory role and WD40 is crucial for the MBW activity[77,78]. Studies focused on plant evolution found a high conservation of genes coding for the MBW complex, both in angiosperms and gymnosperms[79,80].

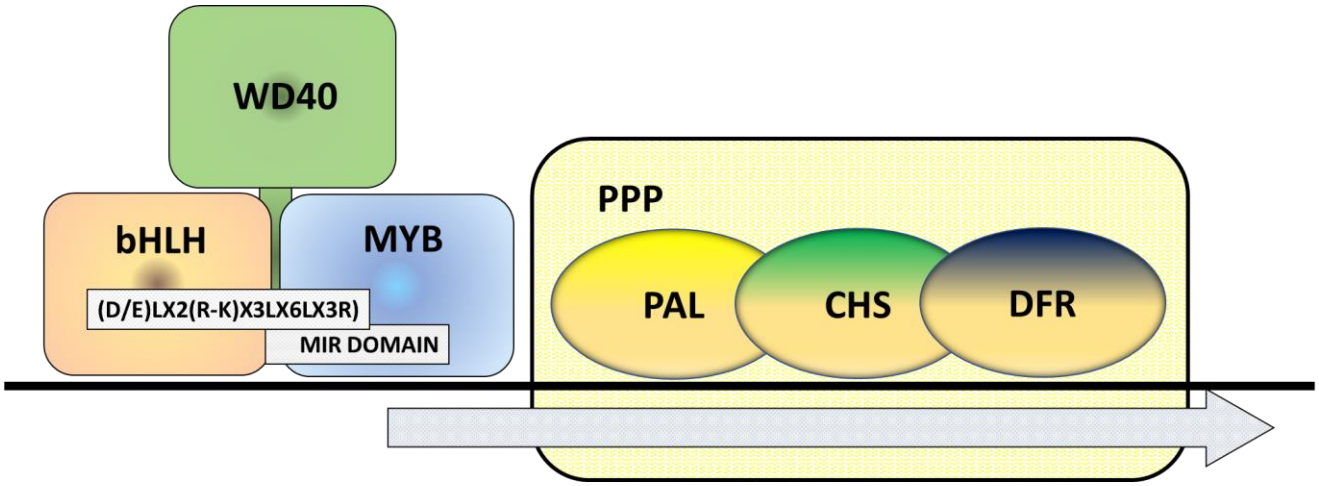


Figure 3: Graphical representation of the MBW complex involved in flavonoid biosynthesis. The heterodimer bHLH/MYB is stabilized by WD40 and then the ternary complex formed is able to stimulate the PPP.

The concept of plant *metabolons*

The rich variety of plant secondary metabolites is possible thanks to redundant families of enzymes decorating molecules' backbones and showing regioselectivity, more than substrate selectivity. Based on these features, enzymes can cluster together in certain cell areas to form molecular complexes that are responsible of substrate channelling and therefore favour the formation of specific secondary metabolites. The presence of a supercomplex formed by the association of metabolic enzymes acting sequentially was introduced by Srere in 1987[81]. The author named this catalytic supercomplex *metabolon*. It ensures fast substrate conversions, immediate availability of intermediates which do not diffuse away and are therefore not diluted, confinement of toxic or labile intermediates in a circumscribed microenvironment with rapid conversion to more stable and less toxic compounds[82]. The formation of such a complex, which can be considered a metabolic unit, has also a role in controlling the effects of inhibitors which have more difficulties in reaching the active sites of enzymes[82].

Metabolons show also a plasticity in terms of enzyme composition: they typically involve soluble enzymes synthesizing intermediates required in different pathways and, therefore, their protein composition may change according to the specific metabolic needs of the cell. In this case, metabolons are transient structures which fine-tune the metabolic needs of the cell in a specific situation. Such a plasticity, where downstream enzymes (glycosyltransferases, methyltransferases) can associate to different metabolons containing upstream core enzymes, guarantees the synthesis of the rich palette of secondary metabolites using a minimum number of enzymes.

Additionally, the metabolon is also characterized by a high level of plasticity thanks to the presence of scaffold proteins which act as Lego bricks, enabling modularity in composition.

Metabolons play a fundamental part in many biological processes. Rakus and colleagues demonstrated the importance of this physiological system in glycolysis and gluconeogenesis with *in vivo* experiments, while You and co-authors highlighted the essential role played by scaffold proteins in the speed of the reactions[83,84].

A strong body of evidence is available in the literature on the formation of metabolons in the PPP[85]. In quaking aspen different PAL isoforms were identified and assigned to different functions, together with the downstream 4CL. For example, PAL1 was found to be partially soluble and membrane-bound, while PAL2 was exclusively cytosolic. However, upon overexpression of C4H, both isoforms were membrane-bound, suggesting a potential role of C4H as nucleation site[86].

Studies in tobacco showed the whole nucleation process. Initially, C4H interacts directly with the cytosolic enzymes PAL1 and PAL2, then the formed heterodimer (C4H-PAL1/2) binds the endoplasmic reticulum (ER)[78]. C4H is a cytochrome P450, like isoflavone synthase (IFS), which is involved in isoflavonoid biosynthesis in soybean by interacting with C4H and other soluble enzymes of the PPP[86].

Interesting is the cellular location of the *phenylpropanoid metabolon*. Indeed, scientific evidence localizes the complex close to the ER surface[86]. This evidence was supported by ER proteins that were described as anchor elements, acting as structural portions without catalytic activities[86]. Recently, membrane steroid-binding proteins were identified as scaffolding proteins organizing monolignol P450 complexes regulating lignin biosynthesis (Figure 4)[87].

Nowadays the metabolon is considered as a highway for plant cells, able to form a restricted reactive area where enzymes and substrates are close to each other, increasing the speed of the reactions, limiting energy dispersion and possible toxic intermediates. Data regarding the mechanisms triggering the formation of the metabolon are still scarce. Future studies may consider other biological factors (e.g. TFs or signaling pathways[88]) in order to deepen the knowledge on this fascinating process. Understanding more about the mechanisms underlying the formation of the metabolon in plants can indeed inspire strategies in biotechnology aimed at boosting the production of specific secondary metabolites of industrial interest.

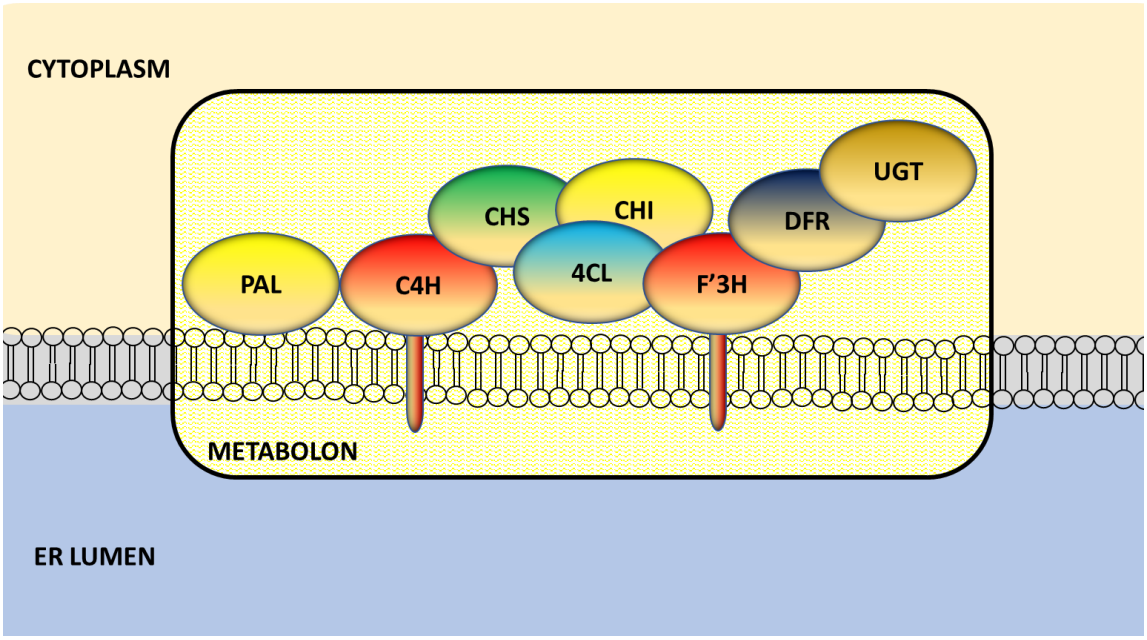


Figure 4: Model of metabolon for flavonoid biosynthesis. PAL: phenylalanine ammonia lyase, C4H: cinnamate 4-hydroxylase, CHS: chalcone synthase, CHI: chalcone isomerase, 4CL: 4-coumarate CoA ligase F3H: flavanone 3-hydroxylase, ANS: anthocyanidin synthase, UGT: UDP-glycosyltransferase.

Prunus avium L. fruits as models for nutraceutical studies

Italy is an important producer of sweet cherry (*Prunus avium* L.) fruits, with a stable production of around 110000–120000 tons over ca. 30000 ha of orchards[89]. Therefore, this fruit-tree plays a prominent role in the agricultural and economic landscape of Italy. Understanding more about the physiology and bioactive contents of non-commercial sweet cherry varieties of Italian collections can inspire exploitation programs valorising these local fruits at the regional level.

Tuscany is known for the high quality of its food products, exported worldwide (wine, oil, cheese, meat) and for specific geographic areas within its territory that have obtained the Protected Geographical Indication (IGP) label. Such an example is Lari, where a specific variety of sweet cherry is cultivated (https://www.rossorubino.tv/en/ciliegia-di-lari-primo-passo-verso-ligp/).

P. avium is a fruit-tree belonging to the Rosaceae family (genus *Prunus* to which many edible fruits belong, notably almond and peach) that produce stone fruits (they are drupes with a woody endocarp, Figure 5) with a characteristic aroma and taste.

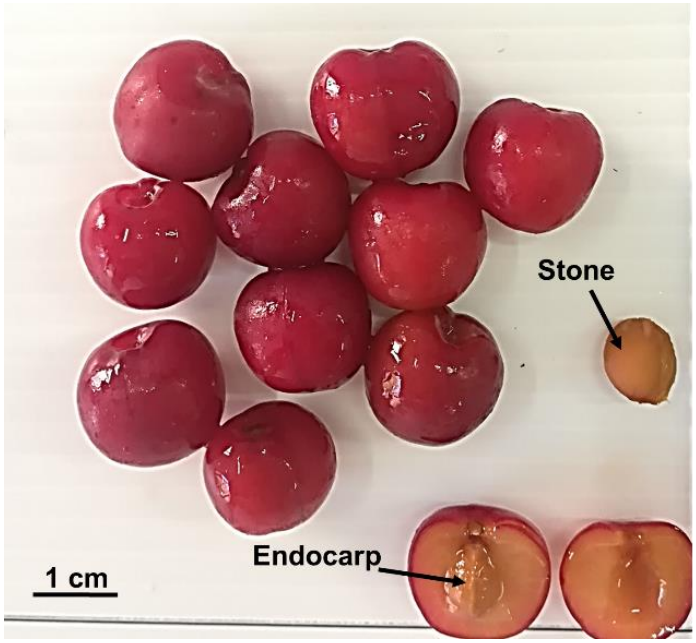


Figure 5: Image showing sweet cherry fruits (var. Crognola) with details of the arrangement of tissues (the arrows indicate the woody endocarp and the stone).

This fruit-tree is native to many regions worldwide, among which Europe, with a preference for temperate climates, like the Mediterranean area. It has a diploid genome of 16 chromosomes (2*n*=16). Like other members of the Rosaceae family, also sweet cherry contains toxic cyanogenic glycosides[90,91], which are present in low concentrations in the stone (0.8%)[91].

The fruits of *P. avium* are rich sources of health-promoting compounds[37,92]. They have a moderate content of simple sugars (and therefore a low glycemic index), as well as organic acids. They are cholesterol-free, low in calories with a high content of water. These drupes are also rich sources of vitamins (notably vitamin C) and minerals (K, P, Ca, Mg). Among the beneficial phytochemicals composing the rich palette of bioactives, there are polyphenols and triterpenes[30,92]. The latter compose the cuticle of sweet cherry fruits and, more specifically, they are found almost exclusively associated with the intracuticular waxes[93,94].

The availability of a sequenced genome greatly helps molecular and functional studies on this economically- important fruit-tree[95]. It has an estimated genome size of ca. 225-330 Mb, with 43349 complete and partial genes coding for proteins. The advent of NGS has fostered RNA-Seq studies on sweet cherry where the transcriptomic signature was studied at different developmental stages[96], or in different fruit tissues[97]. Functional studies are limited by the recalcitrance to regeneration via organogenesis; however, a recent study described a successful protocol for transient and stable transformation starting from leaf explants[98].

The technique of fruit agroinjection is also very useful to assess gene functions in fruits. This protocol is well established for strawberry[99] and tomato[100] and has provided valuable insights into the function of genes involved in anthocyanin biosynthesis [101]. Nobody has tried this protocol on sweet cherry fruits.

In this work an integrated approach was adopted to analyse the bioactive molecules present in 6 ancient varieties of Tuscan sweet cherry and to identify the molecular determinants responsible for the high levels measured. This work is part of a bigger effort undertaken jointly by the CNR in Follonica and the University of Siena and aimed at preserving the collection of Tuscan herbaceous and woody plants, as well as at valorising their use in local products (see http://www.valdimersegreen.com/en/basiq/).
The ultimate goal of the present work is to stimulate future studies on non-commercial varieties and promote their local valorisation via the manufacture of traceable and high-quality food products.

Materials and methods

Sample collection

Sampling was carried out in 3 consecutive years, i.e. 2016, 2017 and 2018, from approximately 18-year-old cherry trees (ancient local varieties of *P. avium* on P-HL-B rootstocks). The trees were grown in the experimental field of the CNR IBE in Follonica (GR, Italy). The experimental field is located at the following coordinates: 42°55′59″N, 10°45′57″E. From a total of 42 regional cherry varieties present in the regional germplasm banks, a sub-set of 6 was selected for this study. The reasons for selecting these varieties are due to their presence in the collection field of Follonica and to their productivity, giving an appropriate number of fruits.

For each variety, the corresponding trees were present in different numbers, depending on how many could be recovered across Tuscany. The total number of trees for each variety was: 'Benedetta' (2), 'Maggiola' (8), 'Morellona' (5), 'Crognola' (3), 'Carlotta' (4), 'Moscatella' (5). The variety Benedetta was sampled only in 2016, since no fruits were obtained in 2017 and 2018. A total of 20 cherry fruits were taken from each tree to have enough biological replicates (4 independent biological replicates), each consisting of a pool of at least 5 fruits. Samples were collected on May the 16th in 2016, May 19th in 2017 and May 18th in 2018 from trees exposed to natural variations of temperature and solar radiation.

The sample collection was carried out by harvesting fruits located at different places on the tree canopies, in order to minimize the bias due to variations in solar exposition. Harvesting took place at an average temperature of 17–20°C (min 12°C-max 22°C in 2016, min 13°C-max 26°C in 2017, min 8°C-max 24°C in 2018) and at an average of 71–74% of humidity for all the years of study. The fruit samples used for the analyses were harvested in the morning between 9:00–10.00 am and were positioned at around 1.70–1.90 m from the soil. Sweet cherry fruits were picked ca. 60 dpa (days post-anthesis). Sweet cherry fruits were picked ca. 60 dpa (days post-anthesis). This period was chosen following the information reported in the regional germplasm bank database (http://germoplasma.regione.toscana.it/index.php?option=com_content&view=article&id=5&Itemid=110). Additionally, colour and fruit size were also taken into consideration as ripening indicators (Figure 6). Each fruit was harvested from the tree by detaching it with the stem, which was subsequently rapidly removed. After removal of the stem, the fruits were immediately plunged in liquid nitrogen, brought to the laboratory and stored at −80°C in Ziplock bags until RNA extraction. The commercial variety Durone was purchased at a local grocery shop in Siena. The phenotypic characteristics of each variety are shown in Figure 6.

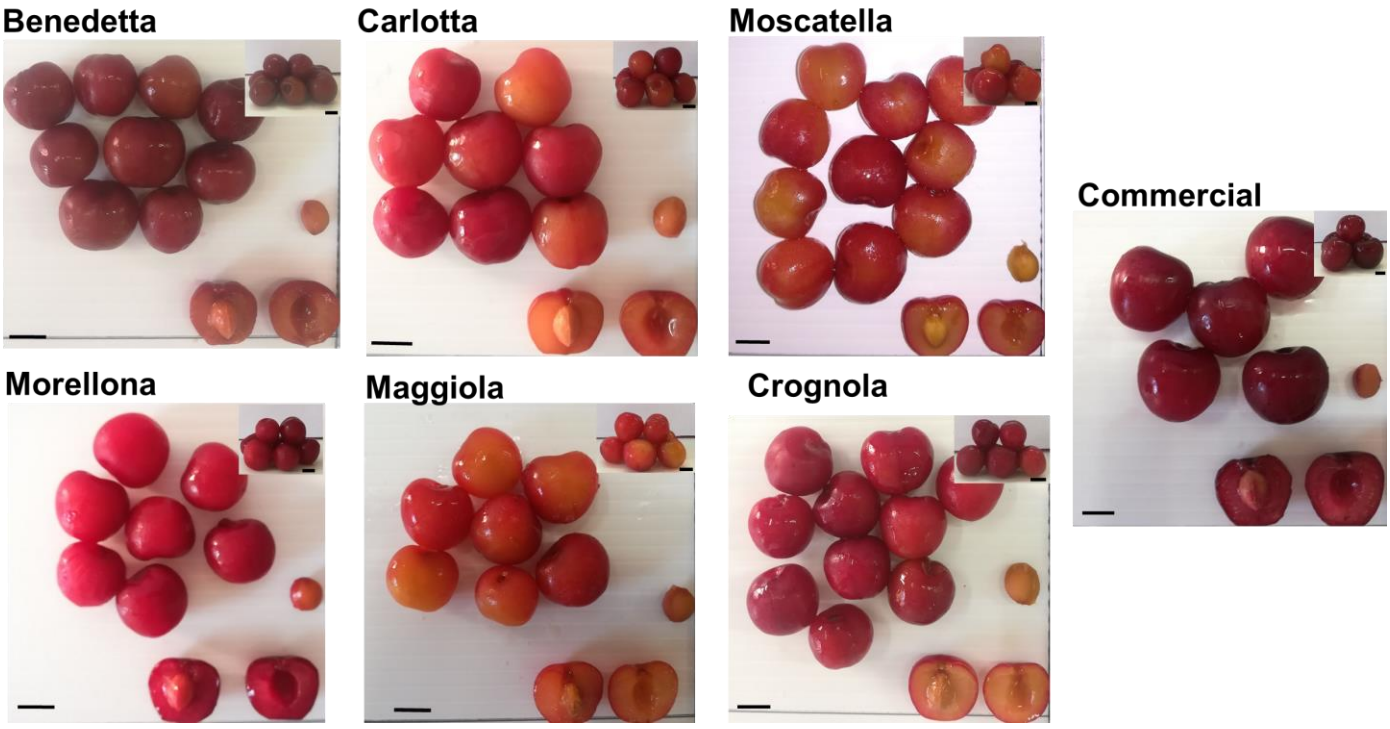


Figure 6: Phenotypic aspects of the 6 ancient sweet cherry fruits used and the commercial variety Durone, obtained from a local market.

DNA extraction

DNA was extracted exclusively from the fruit (pericarp, mesocarp). Samples of six cherry varieties were stored at -80°C and were reduced to a fine powder with liquid nitrogen. DNA was extracted with the quiagen DNeasy Plant Mini Kit, following the manufacturer's instructions; each sample extraction was repeated three times. After extraction, the concentration values were measured with the Nanodrop; the measurements were repeated three times for each DNA extracted and the data were collected. The data showed that the kit used guaranteed a good concentration of DNA and a good degree of purity.

SSR markers

Fourteen primer pairs were used, UDP98411, UDP98412, UDAp420, BPPCT039, AMPA101, UDAp-414, UDAp-415, UDP96008, BPPCT034, BPPCT040, EMPaJ15, EMPaS11, EMPaS12 and UCD-CH17, taken from the literature[102–106]. The primers were previously tested on species belonging to the same family, i.e. *P. avium*, *P. cerasus* (sour cherry) and *P. persica* (peach). SSR markers were specific for genomic regions with a high coefficient of polymorphism and different length of the amplicons. Different fluorophores were chosen to discriminate the amplicons. All the primer details are reported in Table 1.

Primer name	5' labelling	Sequence	Size range	T_m	Allele N°	Reference
UDP98411 Fwd	Dragonfly Orange	AAGCCATCCACTCAGCACTC	150–230	57	2	Testolin 2000,
UDP98411 Rev		CCAAAAACCAAAACCAAAGG				Cipriani et al. 1999
UDP98412 Fwd	FAM	AGGGAAAGTTTCTGCTGCAC	80–150	57	2	Testolin 2000,
UDP98412 Rev		GCTGAAGACGACGATGATGA				Cipriani et al. 1999
UDAp420 Fwd	FAM	TTCCTTGCTTCCCTTCATTG	175	56	2	Messina et al. 2004
UDAp420 Rev		CCCAGAACTTGATTCTGACC				
BPPCT039 Fwd	Dragonfly Orange	ATTACGTACCCTAAAGCTTCTGC	154	55	3	Dirlewanger et al. 2002
BPPCT039 Rev		GATGTCATGAAGATTGGAGAGG				
AMPA101 Fwd	FAM	CAGTTTGATTTGTGTGCCTCTC	190/210	56	3	Hagen et al. 2004
AMPA101 Rev		GATCCACCCTTTGCATAAAATC				
UDAp-414 Fwd	HEX	CAAGCACAAGCGAACAAAAT	174	56	2	Messina et al. 2004
UDAp-414 Rev		GGTGGTTTCTTATCCGATG				
UDAp-415 Fwd	Dragonfly Orange	AACTGATGAGAAGGGGCTTG	156	56	2	Messina et al. 2004
UDAp-415 Rev		ACTCCCGACATTTGTGCTTC				
UDP96008 Fwd	Dragonfly Orange	TTGTACACACCCTCAGCCTG	148-152	60	2	Testolin 2000,
UDP96008 Rev		TGCTGAGGTTCAGGTGAGTG				Cipriani et al. 1999
BPPCT034 Fwd	Dragonfly Orange	CTACCTGAAATAAGCAGAGCCAT	221-255	57	2	Dirlewanger et al. 2002
BPPCT034 Rev		CAATGGAGAATGGGGTGC				
BPPCT040 Fwd	Dragonfly Orange	ATGAGGACGTGTCTGAATGG	135-148	57	2	Dirlewanger et al. 2002
BPPCT040 Rev		AGCCAAACCCCTCTTATACG				
EMPaJ15 Fwd	FAM	TTTTGGTCAATCTGCTGCTG	216-222	60	2	Xuan et al 2009
EMPaJ15 Rev		CTCTCATCTTCCCCCTCCTC				
EMPaS11 Fwd	HEX	ACCACMGAGGAACTTGGG	59-103	60	2	Xuan et al 2009
EMPaS11 Rev		CTGCCTGGAAGAGCAATAAC				
EMPaS12 Fwd	FAM	TGTGCTAATGCCAAAAATACC	137-143	60	2	Xuan et al 2009
EMPaS12 Rev		ACATGCATTTCAACCCACTC				
UCD-CH17 Fwd	HEX	TGGACTTCACTCATTTCAGAGA	188-212	60	2	Xuan et al 2009
UCD-CH17 Rev		ACTGCAGAGAATTTCCACAACCA				

Table 1: Sequences of the 14 primer pairs used for genotyping and relative details, notably fluorophore, size range, number of alleles, melting temperature and references.

PCR conditions

PCR conditions consisted of an initial cycle of 30 sec at 98ºC, followed by 35 cycles of 10 sec at 98ºC, 1 min at 60ºC, and 30 sec at 72ºC, with a final extension at 72ºC for 2 min. PCR products were visualized on agarose gel and then diluted in double distilled water (1:50 v/v) for sequencing. PCR was performed in a final volume of 20 µl containing 2 µl of genomic DNA at 2ng/µl, 10 ul of Mastermix 2X and 5 µl of primers forward and reverse 5µM. All the amplicons were verified by sequencing on an Applied Biosystems 3500 Genetic Analyser using the BigDye Terminator v3.1 Cycle Sequencing and the BigDye XTerminator Purification kits, according to the manufacturer's instructions. The cherry genetic profiles obtained by sequencing were used to generate a phylogenetic tree via the unweighted pair-group method for arithmetic averages (UPGMA) with the NTSYSpc software (Exeter Software, USA).

Extraction of phenolics

For the quantification of total antioxidants, polyphenols and flavonoids the same extraction procedure described by Henríquez and colleagues[107] was used. Whole fruits were ground to a fine powder using liquid nitrogen. 1.5 mL of 70% (v/v) acetone were added to 500 mg of fruit powder and homogenized using an Ultra-Turrax® T-25 basic (IKA®-Werke GmbH & Co., IKA, Staufen, Germany). The mixture was then sonicated for 20 min with an Elma Transsonic T 460/H and homogenized once again to achieve total lysis of the plant material. The final mixture was centrifuged for 5 min at 12000 rpm (Microcentrifuge 5415D, Eppendorf®, Hamburg, Germany) and finally filtered through a 0.45 µm membrane to eliminate the sample impurities.

For anthocyanins, a different solvent was used to perform extraction, as reported by Hosu and colleagues[108]. Anthocyanins were extracted with an acidified ethanol mixture (0.1% HCl 1:1 v/v) adjusted at pH 2.8, to maintain the anthocyanins stability in a liquid solution[109]. Five hundred mg of frozen fruits powder were dissolved in 1.5 ml of extraction solvent, then the previously described steps for sample preparation were followed.

Spectrophotometric analyses

To evaluate the antioxidant capacity, the FRAP reagent was freshly prepared and it consisted of 2040 µL of sodium acetate buffer (300 mM pH 3.6) mixed with 200 µL of 10 mM TPTZ (2,4,6-tripyridyl-s-triazine) and 200 µL of 20 mM ferric chloride. In the last step, 20 µL of sample extract were added to the solution and the solution was incubated for 1 h at 37°C[110]. FRAP values were obtained by comparing the absorbance change at 593 nm in a spectrophotometer (Shimadzu UV Visible

Recording Spectrophotometer UV 160, Shimadzu, Kyoto, Japan) with a previously prepared ferric chloride standard curve. The values were expressed as mmol of ferrous chloride (Fe^{2+}) equivalents per 100 g fresh weight (FW).

The phenolic content was evaluated with the Folin-Ciocalteu method which is used as a standardized method for the analysis of food products. The method uses a mixture of phosphomolybdic acid and phosphotungstic acid allowing electron transfer in an alkaline medium[111]. For the quantifications, 500 µL of a sample extract were added to 3.0 mL of distilled water and 250 µL of Folin-Ciocalteu reagent (Sigma Chemical, St. Louis, MI, USA). Then, 750 µL of saturated sodium carbonate and 950 µL of distilled water were added. The mixture was incubated for 30 min at 37°C; at the end, the absorbance was recorded at 765 nm. The spectrophotometric results were compared to a pre-made gallic acid standard (Sigma Chemical, St. Louis, MI, USA) curve. The total phenolic content was expressed as milligrams (mg) of gallic acid equivalents (GAE) per 100 g FW.

The aluminium chloride method allows the determination of the total flavonoid content, separating their contribution from that of polyphenols. The principle is that aluminium chloride can form complexes with flavonoids in an acidic medium[112]. Briefly, 500 µL solutions of each fruit extract were mixed with 1.5 mL of methanol, 100 µL of 10% aluminium chloride, 100 µL of 1 M potassium acetate and 2.8 mL of distilled water. After incubating at room temperature for 30 min, the absorbance of the solutions was measured at 415 nm. The total flavonoid content was calculated in relation to a quercetin standard (Sigma Chemical, St. Louis, MI, USA) from a calibration curve and the values were expressed as mg of quercetin equivalents (QeE) per 100 g of fruit.

Total anthocyanins were determined according to the pH differential method[113]. One ml of sample solution was dissolved in 3 ml of 0.025 M potassium chloride pH = 1.0 and 1 ml of the same sample was mixed with 3 ml of 0.4 M sodium acetate pH = 4.5. Mixtures were incubated at room temperature and then measured at different absorbance values (510 nm and 700 nm).

The following mathematical formulas were applied[109,113].

$$Asp = (A_{510} - A_{700})\ pH\ 1.0 - (A_{510} - A_{700})\ pH\ 4.5$$

$$\text{Total anthocyanins (TA)} = (Asp \times M \times DF \times 1000) / (\varepsilon \times \lambda \times m)$$

Asp = spectrophotometrically measured absorption values M = molecular weight, DF = dilution factor, ε = molar absorptivity coefficient, λ = cuvette optical path-length (1 cm), m = weight of the sample (g).

The total anthocyanin content was expressed as mg of cyanidin-3-glucoside equivalents (CyE) per 100 g of FW.

HPLC assays

Approximately 500 g of fruit powder were dissolved in 1 mL of acidified methanol containing 1% HCl (v/v). The standard procedure used for acid hydrolysis of the flavonoid components was followed, as described for quercetin, myricetin, and kaempferol in tomatoes and processed products[114]. The mixture was incubated at 90°C for 2 h and continuously stirred. The sample was cooled to room temperature and sonicated for 3 min to remove oxygen before injection. The final extract was filtered through a 0.45 μm membrane to remove impurities.

Anthocyanins were prepared for HPLC analysis following the method described by Nyman and Kumpulainen[115]. 500 mg of fruit powder were dissolved in acidified methanol (0.1 % HCL 1:1 v/v) pH 2.8 and heated at 90°C in a water bath for 50 min, avoiding light exposure. The sample was cooled to room temperature and sonicated for 3 min, then the extract was filtered through a 0.45 μm membrane, before the HPLC injection.

The HPLC method was adapted from the previous work of Kumar and colleagues[116]. The analysis was carried out with an RP-C18 column (SUPELCO Kromasil 100A-5u-C18 4.6 mm × 250 mm) at a flow rate 1 mL/min and the absorbance set at 280 nm in a run time of 21 min. The mobile phase consisted of two solutions, A) H2O and B) acetonitrile with 0.02% trifluoroacetic acid, and the following gradient: 0–5 min A (20%)–B (80%); 5–8 min A (60%)–B (40%); 8–12 min A (50%)–B (50%); 12–17 min A (60%)–B (40%); 17–21 min A (80%)–B (20%). Quantification was performed using an external standard calibration curve consisting of six points at the increasing concentrations of 0.5, 2, 5, 12, 25, and 50 μg per mL for the standards caffeic acid, ferulic acid, chlorogenic acid, quercetin and naringenin (Sigma Chemical, St. Louis, MI, USA).

Anthocyanins were detected following the method previously reported by Luczkiewicz and Cisowski[117]. The analysis was carried out with an RP-C18 column (SUPELCO Kromasil 100A-5u-C18 4.6 mm × 250 mm), at a flow rate 1 mL/min and the absorbance set at 530 nm, in a run time of 65 min.

The mobile phase consisted of two solutions, A) H_2O and B) acetonitrile with 0.02% trifluoroacetic acid and the following gradient: 0–5 min A (20%)–B (80%); 5–8 min A (60%)–B (40%); 8–12 min A (50%)–B (50%); 12–17 min A (60%)–B (40%); 17–21 min A (80%)–B (20%).

Results were compared with previously prepared standard curves (at the increasing concentrations of 0.5, 2, 5, 12, 25, 50 μg per mL) of cyanidin-3-glucoside standard reagents and reported as μg of molecules per g of FW.

Sample preparation for untargeted metabolomics

Whole fruits comprising the flesh and the skin were ground using a mortar and pestle in liquid nitrogen. Approximately 500 mg of frozen powders were lyophilised with an Acid-Resistant CentriVap Vacuum Concentrator (LABCONCO, Kansas City, MO). Lyophilised samples were then weighed (14.5-15.5 mg) and stored at -80°C before the analysis.

Before resuspension, the samples were put at room temperature for a few minutes to avoid condensation. Two µL of chloramphenicol ET 379c (5mg/mL), used as internal standard, were put directly on the powders and then the samples were resuspended in 998 µL of the extraction solvent (MeOH 80% v/v)[118]. The samples were vortexed until complete solubilisation and then put in a thermomixer at 1400 rpm for 4 hours at room temperature. The solutions were vortexed again and centrifuged for 30 min at 20000g at 4°C. 750 µL of the supernatants were collected and completely evaporated using the CentriVap Vacuum Concentrator. Finally, the samples were solubilised in 1 mL of solution (MeOH 85% v/v and 0.1% of FA) and filtered through 0.22 µm filters[118].

UPLC-UV-MS/MS and untargeted metabolomics

Sample separation was performed using a mass spectrometer, TripleTOF 6600-1 (AB Sciex Instruments) with a DuoSpray Ion Source operating in negative and positive ion mode in a hybrid quadrupole-TOF detector. The ESI parameters were set as follows: source temperature of 650°C, ion spray voltage of −4.5 kV. Precursor charge state selection was set at 1. For information-dependent acquisition in high sensitivity mode, survey scans were acquired in 175 ms and the 10 most abundant product ion scans were collected if exceeding a threshold of 100 counts/s.

A sweeping collision energy setting of 15 eV was applied to all precursor ions for collision-induced dissociation. The declustering potential was set at −60 eV. Dynamic exclusion was set for 8 s after two occurrences and then the precursor was refreshed of the exclusion list. The full high resolution (HR)-MS spectra between 190 and 800 mass-to-charge ratio (m/z) were recorded[118].

Five ul of the sample were injected in the UPLC eLambda 800 nm system and analysed in a run time of 60 min at 10 points/sec of sampling rate. The separation was performed on a reverse-phase Acquity UPLC BEH C18 column (2.1 × 100 mm, 1.7-µm particle size; Waters).

The solvents used were A) water + 0.1% FA and B) ACN + 0.1% FA and the column was maintained at 50°C for all the run time. The gradient was applied as follows: 0-16 min 99% of A, 16-35 min 95% of A, 35-45 min 60% of A, 45-50 min 100% of B, 50-54 min 99% of A and 54-60 min 99% of A at 0.5 ml/min flow rate.

Chromatograms were acquired through the Acquity UPLC software and the peaks were identified in PeakView (v 1.2 0.3; AB Sciex) combined with Metlin[118].

RNA extraction and reverse transcription

Sweet cherry fruits were kept frozen in liquid nitrogen and quickly cut into pieces (comprising the exocarp and ca. 5 mm of the mesocarp tissue) using a sterile liquid nitrogen-cooled scalpel. The tissue pieces were collected in a pre-sterilized frozen mortar filled with liquid nitrogen and immediately ground to a fine powder using a pestle. This procedure was necessary to ensure extra care, as previous tests showed that the RNA of sweet cherry fruits is extremely sensitive to degradation caused by even minimal tissue thawing[30,119,120].

RNAs were extracted using the CTAB-based protocol described by Guerriero and colleagues[121]. The CTAB reagent was freshly prepared (2% (% w/v) CTAB, 2.5% (% w/v) polyvinylpyrrolidone (PVP-40), 2 M NaCl, 100 mM Tris-HCl pH 8, 25 mM EDTA and 20 µL β-ME/mL buffer added just prior to use). Five hundred µL of the buffer were added to 100–150 mg of frozen tissue powder and an incubation with gentle shaking was carried out at 60°C for 10 min. Subsequently, 500 µL of chloroforrm/isopropanol (24:1) were added and the tubes were then vortexed and centrifuged as described by the method[120]. The aqueous phase was withdrawn and a precipitation was done with cold isopropanol (2/3 of the volume) for 1 h at −20°C. The RNA was then bound to the silica membranes of the RNeasy columns (Qiagen) and washed/eluted following the manufacturer's instructions.

The purity and quantity of the extracted RNAs were measured at the NanoDrop ND-1000 (Thermo Scientific, Villebon-sur-Yvette, France). In case of low A260/230 nm ratios, a further precipitation/wash step with ammonium acetate/ethanol was performed[121].

The RNA integrity values (RINs) were determined using a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA); for all the samples the RINs were >7.

The extracted RNAs were reverse-transcribed to cDNA with the ProtoScript II reverse transcriptase (New England Biolabs, Leiden, The Netherlands) and random primers, according to the manufacturer's instructions. Samples were reverse transcribed following the ProtoScript II method. Cycling parameters were set according to the manufacturer's instructions. The obtained cDNAs were diluted to 2 ng/µL.

Bioinformatic analyses and primer design

Sweet cherry genes of interest were obtained by blasting the thale cress protein sequences in NCBI, as well as by querying the Genome Database for Rosaceae (GDR; available at https://www.rosaceae.org/) and the cherry database (available at http://cherry.kazusa.or.jp).

The multiple alignments were performed in CLUSTAL-Ω (http://www.ebi.ac.uk/Tools/msa/clustalo/) to identify conserved residues and motifs.

The primers were designed with Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and checked with OligoAnalyzer 3.1 (http://eu.idtdna.com/calc/analyzer). Ten genes were selected as reference for the relative quantification analyses.

All the primers and their relative features are reported in Table 2.

Name	Primer sequence (5' →3')	Tm (°C)	Efficiency (%)	Amplicon size (bp)
Reference genes				
PavPOLYU Fwd	CCCTTGCGGATTACAACATC	80.7	106.29	87
PavPOLYU Rev	GGGTCTTCACGAAAATCTGC			
PavUBQE Fwd	TGAAGGAGCTGAAGGACCTC	84.1	97.00	104
PavUBQE Rev	AGCAGGACCCATTATTGTCG			
PavACT7 Fwd	CCATGTATGTTGCCATCCAG	83.8	97.3	149
PavACT7 Rev	AAGGTCCAGACGAAGAATGG			
PavSERTHR Fwd	AACTCAATCCGCAGGCTATC	82.7	99.43	149
PavSERTHR Rev	TATGGAATGAAGACCCCCAA			
PavPP2A Fwd	CTTTCCCCATCTTTGGACAC	81.2	104.77	149
PavPP2A Rev	CCACAAGAGATCGCACATTG			
PavAP4 Fwd	CCATGATCTCGCAGTTCTTC	83.4	92.73	137
PavAP4 Rev	CTCTTGCCCATCTTCTTTCC			
PavEF2 Fwd	CATTATTGCTGGTGCTGGTG	80.2	105.9	114
PavEF2 Rev	TCACGGAAAGACACAACAGG			
PavTIP41 Fwd	GAAATGGTGTTTGGGGACAG	81.0	98.2	118
PavTIP41 Rev	ACTTCAACTGGTGGCAAAGC			
PavETIF4E Fwd	GGCAAAGCCTCGATACAATG	78.2	89.98	80
PavETIF4E Rev	TTGGTTATGGAGAGCGAAGAC			
PavGAPDH Fwd	TATCAAAGCCACAGCCACTG	82.9	82.27	118
PavGAPDH Rev	TGCTATTCGGAGAACCAACC			
Target genes phenylpropanoid				
PavPAL2 Fwd	CTGCGAGGGAAAGATTATCG	83.9	92.04	114
PavPAL2 Rev	AGTGGAATGGAATGCAGCAC			
PvPAL4 Fwd	AGCCTCTTCCTTTCCCATTC	79.1	95.88	155
PvPAL4 Rev	AATGCCAAACTTGACGAACC			
Pav4CH Fwd	TCCGCATTTTTCCTCTGC	84.8	88.32	111
Pav4CH Rev	ATGATGGCGATGAAGAGACC			
Pav4CL2 Fwd	GTTGCGATGCCGTATTCTTC	85.7	100.45	105
Pav4CL2 Red	TTCTCCCCATCCACTTGTTG			
Pav4CL5 Fwd	TGATGGTGAGGAAGGAAAGG	84.2	109.59	147
Pav4CL5 Rev	TCAGATTCTTGTGCGACGAC			
PavCHS2 Fwd	GTACCAACAAGGCTGTTTTGC	88.1	93.28	146
PavCHS2 Rev	TGTCAAGGTGGGTATCACTGG			
PavCHI3 Fwd	ATAGATTGGCAGCCGATGAC	80.9	91.46	143
PavCHI3 Rev	AATCTCAGCAGTGGCAGAAG			
PavCHI Fwd	TTTCCACCGTCAGTCAAACC	85.7	96.42	102

PavCHI Rev	TCACGAAGTTCCCCTGAATC			
PavF3H Fwd	GAAGATTGTGGAGGCTTGTG	78.8	87.1	72
PavF3H Rev	ATGAGCTTGGCATCAACTCC			
PavDFR Fwd	GCCCATTTCTCATGTCATCC	81.9	86.6	116
PavDFR Rev	TCGTCCAAGTGAACGAACTG			
PavANS Fwd	ATGGGCAGTTTTCTGTGAGC	83.27	88.78	100
PavANS Rev	GTTCTTGGTGGGAAGATTGG			
PavUDP Fwd	ACAACTTGGGCACCTCAAAC	81.26	93.55	80
PavUDP Rev	AGTGAACTCCAACCGCAATG			
PavPPO Fwd	ATGCGAGCCTTACCAGATG	86.07	110.51	104
PavPPO Rev	AGATCCGAATACCCGACTTG			
Target genes cell wall				
PavBGLUC Fwd	AGAATGGCATGGACGAGTTC	79.64	92.17	97
PavBGLUC Rev	TAACAGAGGTGGCGATAGCAG			
PavXTH31 Fwd	TCTCTGGTTTGACCCAACAC	78.84	98.12	95
PavXTH31 Rev	TTCTTACGGGCACATCATCC			
PavXTH6 Fwd	GTGCGTGATGAGCTAGACTTTG	81.92	96.09	103
PavXTH6 Rev	TTTGCTCCCTGTTACCCTTC			
PavAX1 Fwd	CTTCCTCAACCCGAAAACTG	82.32	91.02	100
PavAX1 Rev	AGCCTCGTTCATGTCAATCC			
Target genes stress response				
PavABF1 Fwd	CCTGCTTGGTTGGATCAATG	80.71	91.66	83
PavABF1 Rev	TCCTTGGAATGGTCTGAAGG			
Pav70HS Fwd	TGATGAGGTGGAAAGGATGG	82.05	92.87	93
Pav70HS Rev	TCAGCCTGGTTCTTTGTGTC			
Pav70HS2 Fwd	TGGGAGGAGAGGATTTTGAC	78.97	95.27	91
Pav70HS2 Rev	CTTGGGTTTCCGATGATGTC			
PavUGFU Fwd	TCGTCTCTCGTTATGTCAGTGG	81.38	96.85	104
PavUGFU Rev	AGGTGCCAAGCCATCATAAG			
PavSRC2 Fwd	TGACGTTAAGGTTACGATTAGGG	85.67	98.85	108
PavSRC2 Rev	GGCGGAGCAGGATAATCTC			
PavLTCP Fwd	TGTATTTACGGGACGGTGTG	80.45	94.21	108
PavLTCP Rev	CCACCCCATGAATTTCTCAC			
PavGLUST Fwd	TATTGGGAACAACCCTGGAG	81.65	96.09	108
PavGLUST Rev	GCACCAGAAGTTGAAGTACCAG			
PavHS17 Fwd	AGGACGACAATGTGCTTCTG	83.66	96.02	107
PavHS17 Rev	TAAACTTGCCGACTCTCCTCTC			
PavGLY2 Fwd	CGAGTCAAAGAAAGACCCAGAG	84.2	92.14	104
PavGLY2 Rev	GATGCTCATGGAAGGCAAAC			
Pav12SCRB Fwd	ATTCCCCAACACATCAGAGG	80.85	103.05	110
Pav12SCRB Rev	CAGAGTTTGAGCCACCATTG			
Pav20CHA Fwd	TCAAGGTTGCTGAGGTTGAG	82.32	90.81	82
Pav20CHA Rev	GTGCCAATCGAAGGTTTCTC			
PavTHIRED Fwd	GAGCAACCCGAAAATCAGAG	84.2	85.51	107
PavTHIRED Rev	CCCCAGTCACCAAATTCTTC			

PavDHAR2 Fwd	CTCAGCGACAAACCCAAATG	82.59	104.4	97
PavDHAR2 Rev	TCACGTCAGAATCAGCCAAC			
PavBAS1 Fwd	GTTTGCCCCACAGAAATCAC	82.32	90.81	106
PavBAS1 Rev	CAAGGTGCGAAAACACACTG			

Table 2: Details of all the primers used for gene expression analysis. Details relative to the sequences of the reference and target genes, together with primer efficiency %, melting temperature and amplicon sizes are provided.

qPCR parameters

The qPCR reactions were set up and run as described previously by Trott and co-authors[30,119,120,122]. The cDNA (2 µL) was dispensed in 384-well plates which were prepared with an automated liquid handling robot (epMotion 5073, Eppendorf, Hamburg, Germany). Five µL of a SYBR Green (Takyon™ No Rox SYBR® MasterMix dTTP Blue), 0.4 µL of 2.5 µM primer forward and 0.4 µL of 2.5 µM primer reverse and 2.2 µL of ultrapure H_2O were distributed in each well. The qPCR instrument used was a ViiA 7 Real-Time PCR System (Thermofisher scientific) and the cycling conditions were as follows: an initial 95°C for 10 minutes, followed by 40 cycles of 95°C for 10 sec, 60°C for 1min, 95°C for 15 sec[122].

A melt curve analysis was performed at the end of the amplification cycles, in order to assess the specificity of the primers. The primer efficiencies were determined using a 5-fold dilution series of 6 points (10–2–0.4–0.08–0.016–0.0032 ng/µL).

The expression values were calculated with qBasePLUS (version 2.5, Biogazelle, Ghent, Belgium) by using the reference genes indicated by geNormPLUS. Ten reference genes were tested for stability; *PavACT7* and *PavTIF4E* were identified as sufficient for data normalization when tested together with *PavPP2A*, *PavPOLYU*, *PavSERTHR*, *PavAP4*, *PavGAPDH*, *PavTIP41*, *PavETIF4E* and *PavUBQE*. The data were analyzed by a one-way ANOVA with a Tukey's post-hoc test performed with IBM SPSS Statistics v19 (IBM SPSS, Chicago, IL, USA) to determine the statistically significant differences among groups.

TCA/Phenol-SDS protein extraction

Whole cherry fruits (except the stones) were ground in liquid nitrogen with a mortar and a pestle. One g of fine powder was transferred to 2 ml tubes and resuspended in 1 ml of cold acetone with 10% of trichloroacetic acid (TCA) and 0.07% of dithiothreitol (DTT)[123].

The samples were vortexed and left at -20°C for 60 min, then centrifuged 5 min at 10000 g.

The pellets thus obtained were separated from supernatants and dried at room temperature overnight. The dried pellets were dissolved in 0.8 ml of phenol (Tris-buffer, pH 8.0) and 0.8 ml of SDS buffer [30% (w/v) sucrose, 2% (v/v) SDS, 0.1 M Tris-HCl pH 8.0, 5% (v/v) 2-mercaptoethanol]. The mixtures were vortexed and centrifuged for 3 min at 10000 g. Three hundred µl of the upper phase were transferred in a new 2 ml tube and diluted in 5 volumes of cold ammonium acetate ($NH_4CH_3CO_2$) in methanol. The resuspended samples were centrifuged 2 times to wash the pellet, removing the supernatants each time. Finally, the samples were washed with 80% (v/v) acetone 2 times and the pellets were dried. The dried pellets were resuspended in a buffer of urea 7M, thiourea 2M, Tris 30 mM and CHAPS 4% (w/v). The extracted proteins were quantified using the Bradford method[124].

2D-DIGE

A volume of sample equivalent to 50 µg of proteins was transferred to a new 500 µl tube. The biological replicates of each sample were split and marked half with CyDye 3 fluorochrome and half with CyDye 5. The CyDye 2 fluorochrome was added to the internal standard, which is a mixture of all samples in equal amount. The labelling was done by the addition of 400 pmol of dye, followed by a 30 min incubation on ice in the dark. Then, 1 µl of lysine 10 mM was added to stop the reaction. Samples were mixed and centrifuged again and left 10 min on ice, avoiding the light, before electrophoresis.

The samples were combined as follows: 1 Cy3-labelled, 1 Cy5 labelled and 1 internal standard. They were then loaded on strips (pH 3−10 non-linear, 24cm) for the first dimension, by using the passive rehydration method. Nine µl of ampholytes and 2.7 µl of destreak reagent were added to 450 µl of sample in buffer solution; then, the mixture was centrifuged briefly and loaded in the rehydration support on the strip. The rehydration required between 12 and 18 h.

The strips were transferred to the isoelectric focusing (IEF) system (Ettan IPGphor 3) with paper wicks (wetted with deionized water), put at the ends of the strips. The strips were covered with dry strip cover fluid and the IEF was started. A gradual increase of the voltage was used to reach a total of ca. 90000 V h within 25 h, through the 5 steps planned as follow: 0-3 h 100 V, 3-7 h 1000 V, 7-14 h 1000 V, 14-20 h 10000 V and 20-25 h 10000 V.

The second dimension gels (25x20 cm) were prepared and put in a horizontal electrophoresis chamber (HPE FlatTOP Tower). This machine allows four simultaneous 2D gel runs. The strips were equilibrated with two different buffers (1: urea and DTT, 2: urea and iodoacetamide IAA) and placed directly on the second dimension in the dedicated wells.

The 2D gels were scanned using a laser scanner (Typhoon FLA 9500) and consequently analysed with the software SameSpots (http://totallab.com/home/samespots/).

Spot picking and mass spectrometry

One of the 2D gels was used for spot picking using the Ettan spot picker. The picked spots were previously selected using the SameSpots program and selecting filters of *p*-value <0.01 and the max fold change >2 (a total of 166 proteins was obtained). The gel spots were put in a new 96 well-plate and trypsinised using an automated instrument for trypsinisation. The dried samples were resuspended in 0.7 µl of the α-cyano-4-hydroxycinnamate solution and loaded in a MALDI plate.
A MALDI mass spectrum was acquired using the SCIEX 5800 TOF/TOF (Sciex). The 10 most abundant peaks, excluding known contaminants, were automatically selected and fragmented. MS and MS/MS were submitted to an in-house MASCOT server (version 2.6.1; Matrix Science, www.matrixscience.com) for database-dependent identifications against the NCBI non-redundant protein sequence database (NCBInr) limited to the taxonomy *P. avium* (taxID4229; 10 July 2019; 35758sequences). The parameters were as follows: peptide mass tolerance 100 ppm, fragment mass tolerance 0.5 Da, cysteine carbamidomethylation as fixed modification (alkylation was performed during the equilibration step between IEF and second dimension) and methionine or tryptophan oxidation, double oxidation of tryptophan and tryptophan to kynurenine as variable modifications. Kynurenine, resulting from tryptophan oxidation, is an artefact often observed during automatic digestion in the laboratory were the analysis was performed (Luxembourg Institute of Science and Technology-LIST). Up to two miscleavages were allowed.

Results

The Results section is divided according to the different techniques used. Genotyping is presented first, to provide results concerning the genotype distance of the studied varieties. Subsequently, the coarser spectrophotometric analyses are described, followed by the targeted quantification of phenolic compounds carried out with HPLC on fruits collected in 2016-2017-2018. It should be noted that the variety Benedetta does not appear in 2017 and 2018 because the trees did not produce any fruits. The results obtained with untargeted metabolomics are then shown and complemented by qPCR analyses on selected PPP-related genes. Finally, the soluble proteomes of the 2 highest bioactive-producing varieties (Morellona and Crognola) are shown.

Genotyping results

All the primers, selected by literature search, matched with polymorphic genomic regions. Therefore, it was possible to discriminate the ancient varieties. A similarity tree was generated using commercial varieties originating from France, Turkey Luxembourg and the Tuscan sweet cherries (Figure 7). The varieties from other countries were here considered to better discriminate the phylogenetic relatedness of the Tuscan fruits, thus enriching the tree and highlighting the genetic differences in this small group of samples.

However, these varieties will not appear in the other analyses.

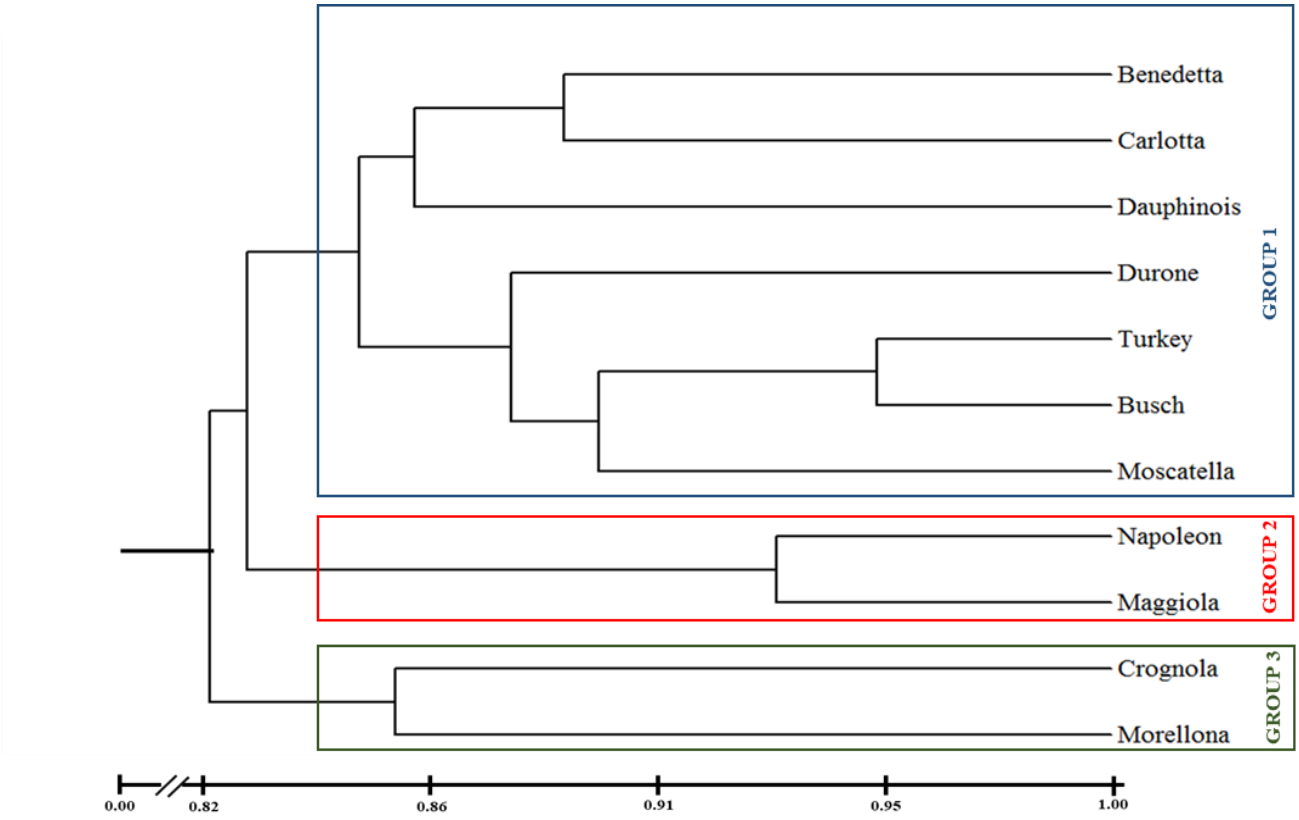


Figure 7: Maximum likelihood phylogenetic tree (bootstraps: 100) of the *P. avium*. Boostrap values ranging from 0.82 to 1 are displayed in the black bar below the tree.

The phylogenetic tree shows a good separation of the ancient varieties from the commercial ones studied. Figure 7 shows 3 main genetic clusters belonging to 3 different branches of the tree. The biggest cluster includes 3 ancient varieties (Benedetta, Carlotta, Moscatella) and 4 commercial ones (Durone from Italy, Dauphinois from France, commercial from Turkey and commercial from Luxembourg referred to as Busch). The second branch comprises the ancient variety Maggiola and the commercial French variety Napoleon, which display a high genetic similarity. The last branch includes the varieties Crognola and Morellona, which show higher genetic distance as compared to the other Tuscan cherries studied.

The phylogenetic results were then complemented with analyses on the metabolites, genes and proteins, in order to assess whether the detected genetic differences were reflected in varying metabolite levels, as well as PPP-related gene expression profiles and protein abundances.

Spectrophotometric and targeted metabolite quantification

In Table 3 the results relative to the spectrophotometric analyses are reported for the 3 years of study. The objective of these analyses was to confirm the stability in the trend of the measured parameters over different years of harvest. For comparative purposes, only the Italian commercial variety Durone is reported.

The results are shown in Table 3. With respect to the general trend of the total antioxidants, the variety Crognola ranks first, followed by Morellona, while the commercial fruits are the lowest. The same ranking is observed for the polyphenols and flavonoids, while anthocyanins are highest in Morellona in 2016 and 2018. It should be noted that this variety is red fleshed (Figure 6), differently from all the others. The commercial fruits always rank lowest, with the exception of anthocyanins: the reason may be due to the commercial selection based on characters that are appealing for the consumer, namely red skin and flesh. It is interesting to note that, in 2017, Carlotta was the lowest producer of antioxidants, polyphenols and anthocyanins. Despite showing a yellow pulp (Figure 6), Crognola is the highest producer of bioactive molecules: one can suppose that the functional molecules are concentrated in the red skin.

Variety name	Antioxidants (mmol Fe^{2+}/100 g FW) ± SD	Polyphenols (mg GaE/100 g FW) ± SD	Flavonoids (mg QeE/100 g) ± SD	Anthocyanins (mg CyE/100 g FW) ± SD
2016				
Carlotta	1.24 ± 0.04^{d}	115.23 ± 5.41^{d}	46.37 ± 1.40^{d}	39.57 ± 0.76^{c}
Benedetta	1.52 ± 0.03^{c}	255.21 ± 10.56^{b}	65.71 ± 2.89^{c}	27.51 ± 1.35^{f}
Morellona	2.15 ± 0.06^{b}	276.29 ± 8.73^{b}	77.91 ± 2.33^{b}	55.37 ± 2.15^{b}

Maggiola	2.09±0.09^{b}	267.64±7.12^{b}	47.51±1.87^{d}	30.32±0.87^{ef}
Moscatella	1.42±0.05^{c}	150.66±7.86^{c}	44.38±2.21^{d}	35.14±2.28^{cd}
Crognola	3.06±0.11$^{^{\circ}}$	403.40±6.41^{a}	88.26±3.05^{a}	65.12±2.43^{a}
Durone (Commercial)	1.08±0.05^{e}	105.56±3.65^{d}	31.94±1.51^{e}	33.51±2.39^{de}
2017				
Carlotta	1.49±0.06^{a}	200.49±15.82^{b}	47.21±8.42^{b}	30.8±1.84^{a}
Morellona	2.51±0.03^{b}	300.87±13.14^{c}	63.01±3.71^{b}	59.41±2.53^{b}
Maggiola	1.92±0.12^{c}	276.71±7.57^{d}	46.41±2.51^{a}	33.17±1.79^{a}
Moscatella	1.51±0.05^{a}	215.15±11.91^{a}	44.82±4.82^{a}	27.04±2.40^{a}
Crognola	3.17±0.01^{d}	368.18±5.72^{e}	81.62±3.28^{c}	67.75±2.71^{c}
Durone (Commercial)	1.13±0.04^{e}	146.05±2.71^{f}	32.21±2.79^{d}	34.65±2.32^{a}
2018				
Carlotta	1.73±0.01^{c}	159.18±0.41^{d}	49.33±0.71^{b}	34.46±0.81^{c}
Morellona	2.22±0.02^{b}	313.11±3.55^{b}	97.86±1.56^{a}	65.27±1.03^{a}
Maggiola	1.43±0.02^{d}	137.70±2.01^{e}	40.06±1.23^{c}	32.41±0.91^{c}
Moscatella	1.34±0.03^{e}	175.99±2.45^{c}	47.61±1.01^{b}	35.15±1.84^{c}
Crognola	3.07±0.01$^{^{\circ}}$	387.11±1.29^{a}	102.05±2.42^{a}	58.46±1.11^{b}
Durone (Commercial)	1.22±0.02^{f}	126.58±0.97^{f}	33.20±2.67^{d}	34.24±0.98^{c}

Table 3: Values (± standard deviation, number of independent biological replicates = 3) obtained from the sweet cherry varieties sampled in 2016-2017-2018. The results show the total content of antioxidants expressed as mmol Fe^{2+} per 100 g FW, polyphenols as mg of GAE (gallic acid equivalents) per 100 g of FW, flavonoids as mg of QeE (quercetin equivalents) per 100 g of FW and anthocyanins as CyE (cyanidin-3-*O*-glucoside equivalents) per 100 g of FW. Different letters indicate statistically significant differences (p-value <0.05) among groups at the one-way ANOVA with Tukey's post-hoc test.

The spectrophotometric assays gave a general overview of the different bioactive classes present in the sweet cherries and their relative contents. It was then necessary to quantify key phenolics, in order to enrich and confirm the spectrophotometric assays.

The HPLC analysis was performed on 5 phenolic components belonging to different sub-groups and frequently investigated in nutraceutical studies: chlorogenic acid, *p*-coumaric acid, catechins, rutin and cyanidin-3-*O*-glucoside (Table 4).

Variety name	Chlorogenic acid (µg/g FW) ± SD	*p*-coumaric acid (µg/g FW) ± SD	Catechins (µg/g FW) ± SD	Rutin (µg/g FW) ± SD	Cyanidin-3-*O*-glucoside (µg/g FW) ± SD
2016					
Benedetta	81.27±0.45^{f}	16.67±0.35^{d}	186.76±1.41^{a}	44.32±1.19^{b}	33.85±1.53^{d}
Carlotta	111.81±0.60^{d}	29.47±0.19^{c}	27.90±1.92^{f}	23.43±1.31^{de}	48.97±1.84^{c}
Morellona	224.98±0.40^{b}	34.03±0.25^{a}	82.32±2.33^{c}	40.67±1.41^{c}	84.87±1.10^{b}
Maggiola	74.32±0.22^{e}	18.12±0.25^{e}	65.76±1.93^{d}	31.12±1.82^{cd}	37.76±1.22^{d}
Moscatella	139.15±0.11^{c}	14.18±0.31^{g}	30.78±2.28^{e}	29.07±0.94^{df}	32.22±1.64^{d}

Crognola	287.76±0.56[a]	21.11±0.32[b]	123.45±1.53[b]	75.98±2.25[a]	149.54±0.79[a]
Durone (Commercial)	68.29±0.45[g]	19.05±0.10[f]	11.99±0.67[g]	22.54±2.15[f]	40.43±1.74[d]
2017					
Carlotta	121.91 ± 0.62[a]	19.41 ± 0.22[a]	25.93 ± 1.98[a]	30.83 ± 1.31[a]	51.78 ± 1.83[a]
Morellona	324.59 ± 0.44[c]	30.05 ± 0.32[c]	72.54 ± 2.34[c]	38.54 ± 1.42[c]	74.82 ± 1.15[c]
Maggiola	94.32 ± 0.25[d]	12.43 ± 0.22[d]	45.99 ± 1.95[d]	34.46 ± 1.83[c]	35.28 ± 1.35[b]
Moscatella	239.25 ± 0.13[e]	10.13 ± 0.37[e]	34.86 ± 2.79[e]	27.93 ± 1.63[a]	33.54 ± 1.66[b]
Crognola	387.73 ± 0.25[f]	28.11 ± 0.38[f]	163.51 ± 1.47[f]	95.33 ± 2.51[d]	151.2 ± 1.22[d]
Durone (Commercial)	78.25 ± 0.47[b]	10.85 ± 0.17[e]	15.85 ± 1.15[g]	25.65 ± 2.15[a]	36.27 ± 1.77[b]
2018					
Carlotta	179.42±0.83[c]	29.02±0.57[e]	122.83±4.96[c]	32.54±0.63[d]	59.50±1.02[c]
Morellona	276.38±0.98[b]	51.17±1.11[d]	172.55±1.10[b]	39.05±0.43[c]	80.62±0.77[b]
Maggiola	99.23±0.57[d]	76.50±0.88[c]	46.74±0.99[e]	29.11±0.72[e]	31.26±1.47[e]
Moscatella	117.03±0.86[e]	87.61±0.74[b]	53.20±0.25[d]	43.01±0.21[b]	32.51±1.05[e]
Crognola	312.67±1.11[a]	123.49±0.64[a]	219.44±2.49[a]	99.78±0.55[a]	149.77±1.29[a]
Durone (Commercial)	148.33±0.43[e]	74.07±0.87[c]	21.71±1.31[f]	27.85±0.53[e]	40.32±0.83[d]

Table 4: HPLC quantifications of specific bioactive molecules in the sweet cherry varieties sampled in 2016-2017-2018 expressed as μg per g of FW. The values are indicated with the relative standard deviation (number of independent biological replicates = 3); different letters indicate statistically significant differences (*p*-value <0.05) among groups for each year at the one-way ANOVA with Tukey's post-hoc test.

Despite the strong red colour, the most abundant compound quantified in cherries is chlorogenic acid, showing high values in all of the years investigated. Values ranging from a minimum of 68.29 μg per g of FW in Durone to a maximum of 380.73 μg per g of FW in Crognola were measured in the fruits collected in 2016, meaning that the chlorogenic acid represents over 40% of the total phenolic compounds detected. Crognola and Morellona results always as the highest producing varieties, followed by Moscatella in 2016-2017 and Carlotta in 2018; the variety Durone shows the lowest values in all the years studied.

The second phenolic acid detected is *p*-coumaric acid, showing a concentration between a minimum of 10.85 μg per g of FW in Durone in 2017 and a maximum of 123.49 μg per g of of FW in Crognola in 2018. Coumaric acid is less abundant than chlorogenic acid among the varieties and it shows a variable trend of concentration during the years. Indeed, in 2016 Morellona and Carlotta are the main representative varieties, differently from 2017 in which Morellona and Crognola rank first and in 2018 where Crognola and Moscatella are the highest producing ones. The Durone variety is the lowest in 2016 and 2017 together with Moscatella; however, in 2018 the value is considerably higher and Carlotta ranks last.

With respect to flavonoids, the most abundant compound is catechin with values of 219.44 µg per g of FW in Crognola in 2018 and a minimum of 11.99 µg per g of FW in Durone in 2017. Because of the lack of fruit production in 2017 and 2018, results relative only to 2016 are shown for Benedetta, which shows high contents of catechins (Table 3).

In addition to Benedetta, Morellona and Crognola also show high contents of catechins for all the years, followed by Maggiola in 2017 and Carlotta in 2018. Rutin could also be detected, a flavonoid present in plants together with quercetin. In local cherries, rutin is present in concentrations ranging between 99.78 µg per g of FW in Crognola in 2018 and 22.54 µg per g of FW in Durone in 2016. The highest producer remains Crognola over the three years.

Benedetta, Morellona and Moscatella show the second highest values in 2016, 2017 and 2018, respectively. As previously reported for chlorogenic acid and catechin compounds, the Durone variety shows the lowest concentrations.

Concerning anthocyanidins, we identified the cyanidin-3-*O*-glucoside molecule in a maximum concentration of 149.77 µg per g of FW in Crognola and a minimum of 31.26 µg per g of FW in Maggiola in 2018. These compounds are synthesized by the enzymes belonging to the last stage of the PPP (Figure 2) and are responsible for the red colour in cherries.

Despite a flesh devoid of red colour, the variety Crognola has the highest concentration of cyanidin-3-*O*-glucoside in all of the years investigated. Morellona ranks second, with high values of anthocyanins maintained in all of the years studied. Unlike the major part of the detected compounds, Durone often contains higher levels of cyanidin-3-*O*-glucoside, as compared to the ancient varieties Maggiola and Moscatella.

Untargeted metabolomics analyses

The cherry varieties were then studied using untargeted metabolomic techniques. It was thus possible to confirm the compounds previously detected with HPLC and identify further molecules. Besides the previously shown compounds (Table 4), 15 different metabolites were identified in positive mode and 14 in negative mode. The attention was therefore concentrated on these newly identified compounds.

A hierarchical clustering of the heatmap was generated, using the maximum fold change values which express the different relative abundance of the compounds (Figure 8).

Most of the molecules detected belong to the flavonol and proanthocyanidin classes, which are present in different forms (Tables 5-6).

As reported in Tables 5-6, flavonols and proanthocyanidins are present as A-type or B-type classes. In general A and B-type compounds are classified based on their chemical bonds. Compounds with

double bonds, such as C4→C8 and an additional ether bond between O7→C2, are classified as A-type, while a single bond between C4→C6 or C4→C8 characterize the B-type[125] (Figure 8). Proanthocyanidins are products of the flavonoid pathway and require the same metabolic intermediates but, differently from other compounds, such as anthocyanins, proanthocyanidins are produced from catechins and epicatechins[126].

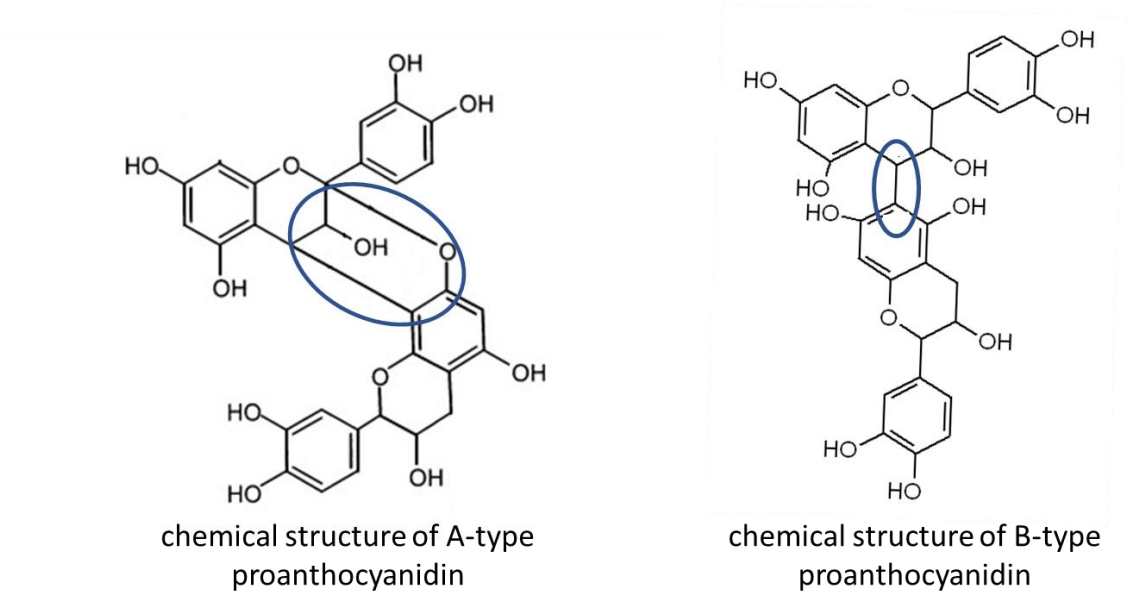


Figure 8: Chemical structures of A- and B-type proanthocyanidins.

As Figure 9A shows, the upper cluster of the heatmap includes epicatechin, i.e. the precursor of the proanthocyanidin pathway[126], as well as flavonols and proanthocyanidins, which are the intermediates and final compounds. In the negative mode (Figure 9B), the pattern is instead slightly different, because of the clustering of coumaroyl quinic acid together with epicatechin.

The heatmap hierarchical clustering reveals different abundances of the identified flavonoids, with Crognola and Morellona as the highest producing varieties and the commercial one as the lowest. These results confirm the ranking obtained previously with HPLC (Table 4).

Coumaroyl quinic acid belongs to the polyphenol group of the chlorogenic acids (CGAs). Figure 9 shows that Crognola and Morellona have the highest concentration of this compound, both in positive and negative mode. Moreover, the variety Maggiola also contains high amounts of coumaroyl quinic acid, while Benedetta has the lowest quantity.

Trihydroxyl flavanone, also called naringenin, is present in high amounts in the Tuscan non-commercial cherry varieties, both as aglycone and conjugated (trihydroxyl flavanone hexosyl) (Tables 5-6). Naringenin is present at higher levels in the varieties Morellona and Crognola. Durone and Carlotta show instead the lowest concentrations (Tables 5-6).

Cinchonain is also present in the ancient sweet cherries. It is a molecule that belongs to the flavonolignan group. Flavonolignans are related to lignans and neolignans because of their similar biosynthetic pathways and their chemical structure, deriving from the polymerization of two phenylpropanoid units[127]. Based on the results shown in Tables 5-6, cinchonain shows high concentration in all the ancient varieties, notably in Crognola and Morellona. Furthermore, the commercial variety is characterised by the lowest content of cinchonain.

Putative identification	$t_{r\ (min)}$	Formula	Theoretical *m/z*	Observed *m/z*	Adduct interpretation of observed *m/z*	Mass error (ppm)	main MS2 fragments
Flavanol hexoside	8.79	C21H24O11	453.13914	453.1380593	$[M+H]^+$	-2.38	139.0381
A-type flavanol dimer	13.87	C30H24O12	577.13405	577.1323654	$[M+H]^+$	-2.92	245.0434
Coumaroyl quinic acid	16.00	C16H18O8	339.10744	339.1066404	$[M+H]^+$	-2.36	147.0433 - 119.0481 - 91.0534
(epi)afzelechin–(epi)catechin	17.65	C30H26O11	563.15479	563.1536	$[M+H]^+$	-2.11	107.0480 - 147.0431 - 287.0544
A-type procyanidin trimer	19.02	C45H36O18	865.19744	865.1971339	$[M+H]^+$	-0.35	245.0441 - 287.0544 - 163.0375
B-type flavanol hexamer	19.02	C45H38O18	867.21309	867.2131657	$[M+H]^+$	0.09	245.0422 - 127.0379 - 163.0382
B-type procyanidin tetramer	19.64	C60H50O24	1155.27648	1155.277155	$[M+H]^+$	0.58	245.0440 - 247.0593 - 163.0375
A-type procyanidin trimer	19.64	C45H36O18	865.19744	865.1967749	$[M+H]^+$	-0.77	287.0539 - 247.0587 - 135.0428
A-type flavanol dimer	19.95	C30H24O12	577.13405	577.1325196	$[M+H]^+$	-2.65	135.0441 - 123.0416 - 163.0376
B-type flavanol pentamer	19.98	C45H38O18	867.21309	867.211714	$[M+H]^+$	-1.59	127.0374 - 163.0394 - 135.0409
Trihydroxyflavanone	20.79	C15H12O5	273.07575	273.0752237	[M+H]+	-1.93	153.0167
B-type proanthocyanidin dimer	21.52	C30H26O12	579.1497	579.1490729	$[M+H]^+$	-1.08	123.0429 -127.0378 - 163.0331
A-type dimer proanthocyanidin	21.52	C30H24O12	577.13405	577.1319435	$[M+H]^+$	-3.65	123.0431 - 245.0442 - 135.0430
Cinchonain	24.52	C24H20O9	453.11801	453.1169285	$[M+H]^+$	-2.39	191.0333 - 163.0362
Trihydroxyflavanone	25.74	C21H22O10	435.12857	435.126972	$[M+H]^+$	-3.67	169.0122 - 273.0757

Table 5: List of differentially abundant compounds in sweet cherries. Data were obtained by UPLC–TTOF in positive ESI mode, with MS/MS. The relative abundance of each peak is based on the selected ion current and is presented as the average normalised value of two technical replicates from four biological replicates. tr, retention time.

Putative identification	$t_{r\ (min)}$	Formula	Theoretical *m/z*	Observed *m/z*	Adduct interpretation of observed *m/z*	Mass error (ppm)	main MS2 fragments
A-type flavanol dimer	13.85	C30H24O12	575.1195	575.1180972	$[M-H]^-$	-2.439080709	125.0249 - 163.0017 - 255.0300
B-type flavanol	13.93	C45H38O18	865.19854	865.197616	$[M-H]^-$	-1.068018575	125.0231 - 161.0243 - 407.0770

Coumaroyl quinic acid	15.99	C16H18O8	337.09289	337.0928345	$[M-H]^-$	-0.164571234	173.0462 - 93.0341 - 119.0494
(epi)afzelechin–(epi)catechin	17.63	C30H26O11	561.14024	561.1387	$[M-H]^-$	-2.744411985	289.0733 - 245.0821 - 203.0747
A-type procyanidin trimer	19.01	C45H36O18	863.18289	863.1819872	$[M-H]^-$	-1.045934021	285.0370 - 125.0265 - 161.0259
B-type flavanol hexamer	19.01	C45H38O18Na	887.18048	887.1794619	M+Na-2H	-1.147536622	125.0205 - 161.0141 - 243.0284
B-type procyanidin tetramer	19.63	C60H50O24	1153.26193	1153.257	$[M-H]^-$	-4.27483113	125.0279 - 243.0303 - 161.0254
A-type procyanidin trimer	19.63	C45H36O18	863.18289	863.1816647	$[M-H]^-$	-1.419463094	125.0230 - 161.0276 - 243.0317
B-type flavanol pentamer	19.93	C75H62O30	1441.32532	1441.3232	$[M-H]^-$	-1.470868492	125.0219 - 243.0327 - 287.0709
trihydroxyflavanone	20.78	C15H12O5	271.0612	271.0605	[M-H]-	-2.582442637	153.0167
B-type proanthocyanidin dimer	21.5	C30H26O12	577.13515	577.1345859	$[M-H]^-$	-0.977462236	125.0237 - 289.0737 - 161.0258
A-type dimer proanthocyanidin	21.5	C30H24O12	575.1195	575.1181919	$[M-H]^-$	-2.274410482	125.0224 - 161.0292 - 177.0172
Cinchonain	24.57	C33H26O11	597.14024	597.1593279	$[M-H]^-$	31.96554834	341.0686 - 189.0171 - 217.0126
Trihydroxyflavanone	25.72	C21H22O10	433.11402	433.1128338	$[M-H]^-$	-2.738880348	271.0612 - 243.0652

Table 6: List of differentially abundant compounds in sweet cherries. Data were obtained by UPLC–TTOF in negative ESI mode, with MS/MS. The relative abundance of each peak is based on the selected ion current and is presented as the average normalised value of two technical replicates from four biological replicates. tr, retention time.

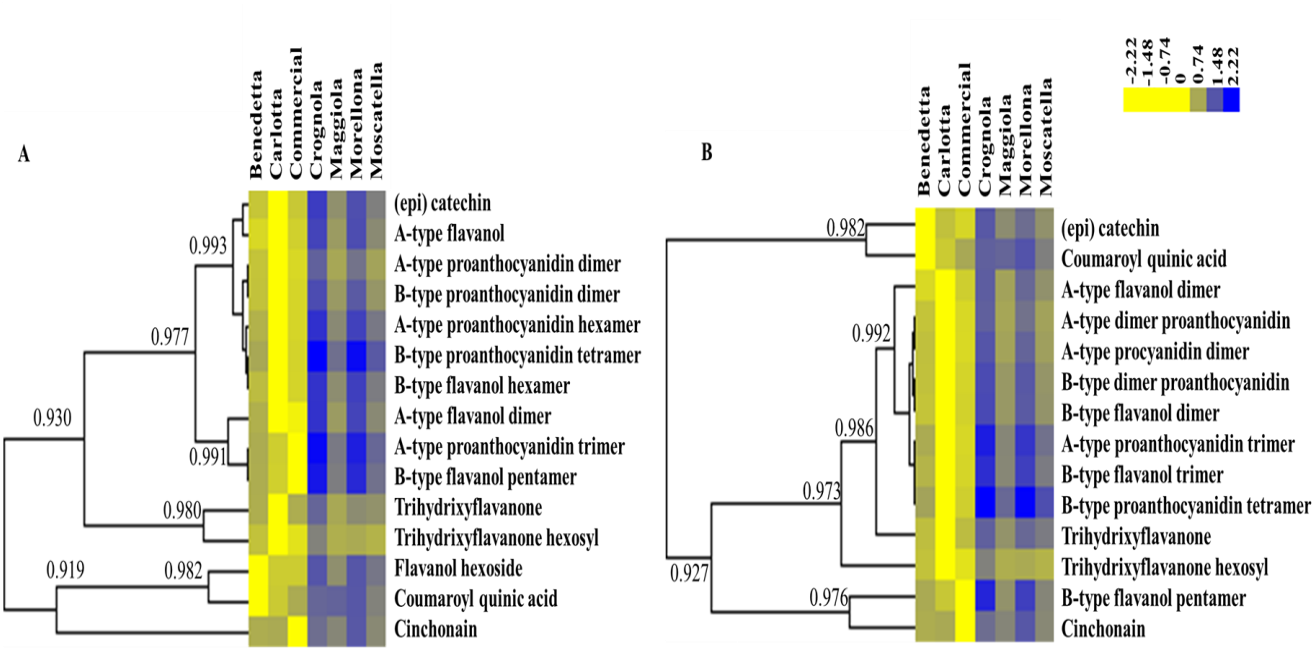


Figure 9: Heat map hierarchical clustering of the compounds identified in positive (A) and negative (B) modes using the maximum fold change (average of the values obtained in the 3 years). Numbers indicate the Pearson correlation coefficients. The colour bar indicates the abundances.

Targeted gene expression analysis

The gene expression analysis was carried out on the three years of harvest 2016, 2017 and 2018 and on genes intervening in the PPP. More specifically, the genes are *PAL*, *C4H*, *4CL*, *CHS*, *CHI*, *F3H*, *DFR*, *ANS* and a *UGT* (responsible for the glycosylation of anthocyanin aglycones). The NCBI protein accessions are given in Table 2.

The genes involved in the PPP are notoriously multigenic[128,129]; for *PAL*, *4CL* and *CHI*, two isoforms were analysed because of the roles that these genes have as gatekeepers (*PAL*) and members of the general steps (*4CL*), respectively, as well as their implication in branch points (*CHI*).

By studying the isoforms, it is possible to speculate about their potential role in the provision of precursors needed for the synthesis of aromatic macromolecules.

The PCA (Principal Component Analysis) was carried out on the 3 years to get an idea of the separation of each biological replicate and, above all, of each variety (Figure 10).

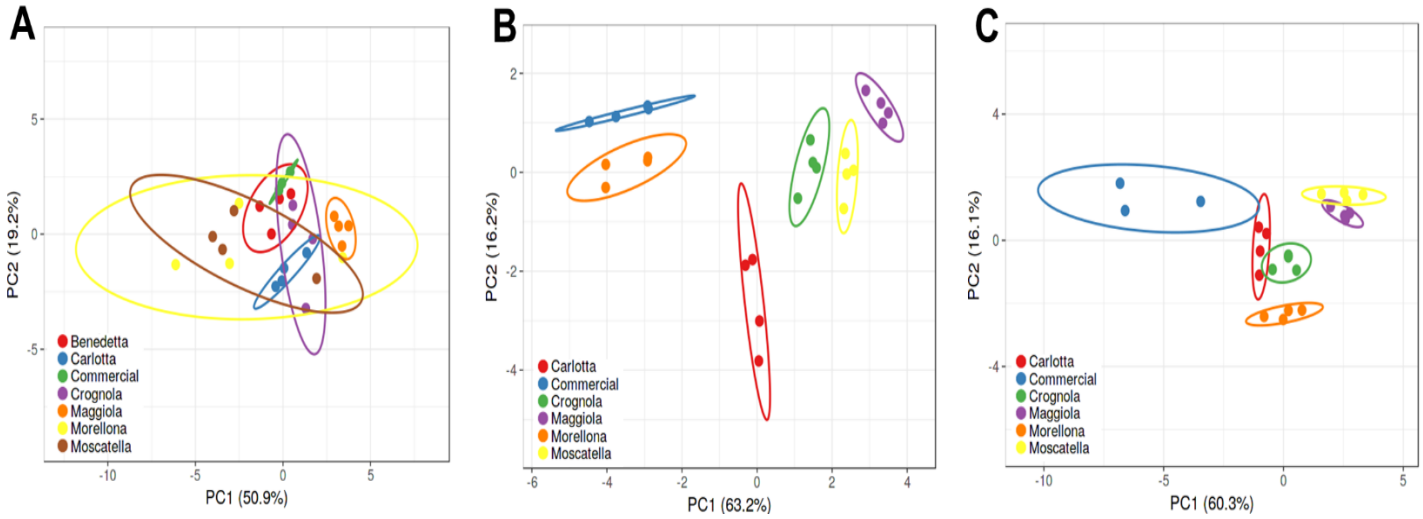


Figure 10: Principal Component Analysis (PCA) of the gene expression values in the 2016 (A), 2017 (B), 2018 (C). Circles indicate the 4 biological replicates for each variety (represented using different colours, according to the legends on the left-hand side).

The PCAs show that 2017 is characterized by a nice and clear separation of the varieties, while in 2018, Morellona and Crognola, together with Moscatella and Maggiola are less separated (Figure 10). The year 2016 is characterized by a less distinct and more overlapping distribution of the varieties.

Subsequently, the hierarchical clustering of the gene expression was carried out to unveil eventual correlations of expression patterns among the genes studied (Figure 11).

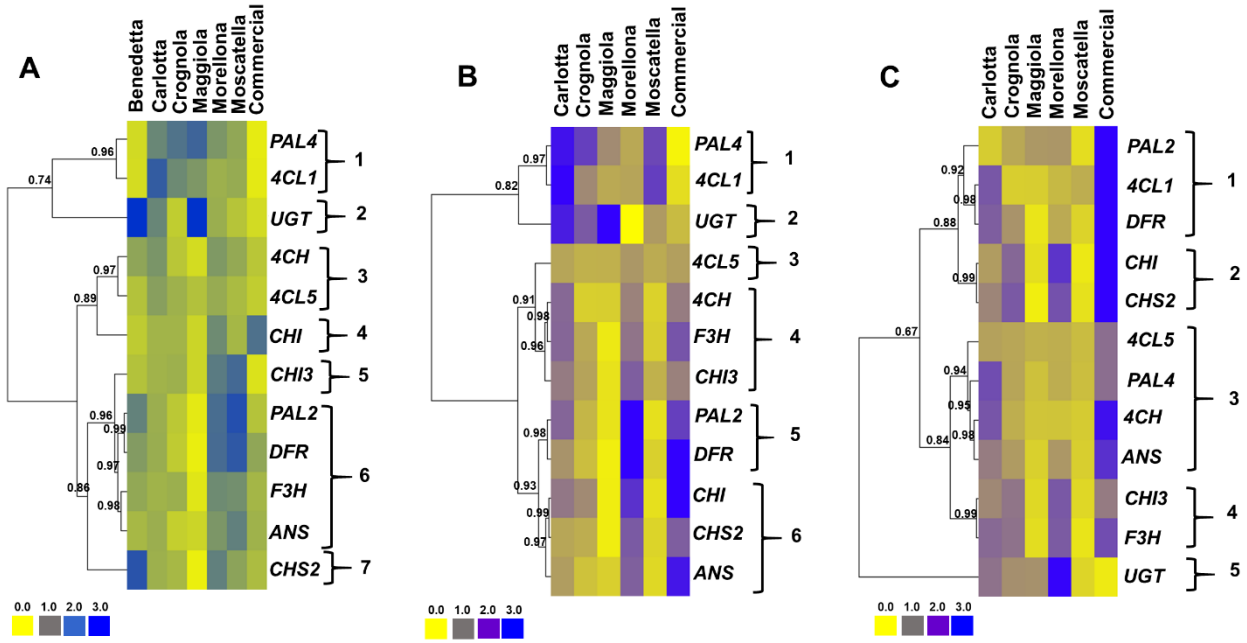


Figure 11: Heat map hierarchical clustering of the PPP-related gene expression data across the 3 years of study. Numbers indicate the Pearson correlation coefficients. The colour bar indicates the expression intensities.

In 2016 it is possible to distinguish 7 major patterns, by setting 0.96 as threshold value for the Pearson correlation. The first one groups *PAL4* and *4CL1*, the second is given by *UGT*, the third is composed of *C4H* and *4CL5*, the fourth and fifth by *CHI* and *CHI3*, the sixth comprises *PAL2*, *DFR*, *F3H* and *ANS* and the last *CHS* (Figure 11A).

In 2016, Benedetta is present together with a commercial Italian variety Durone, included for comparative purposes. Indeed, although the details of growth conditions and post-harvest storage are not known and were certainly different from those of the ancient fruits, its inclusion can provide reference data enriching those obtained with the targeted metabolite approach.

Besides *CHI*, the commercial fruits display lower expressions as compared to the ancient fruits and this is particularly evident for the genes partaking in the general phase of the PPP (Figure 11A). It is interesting to note that the two *PAL* isoforms cluster with different genes: *PAL4* groups with *4CL1*, while *PAL2* with *DFR*.

Carlotta, Crognola and Maggiola showed overall higher expression of the gatekeeper gene *PAL4*, while *CHI3*, *PAL2*, *DFR*, *F3H*, *ANS* and *CHS* were highly expressed in the varieties Morellona and Moscatella.

Interestingly, Morellona is the ancient variety showing the strongest red pulp; therefore, the high content of anthocyanins shown above (Table 3) is explained by the higher expression of the genes involved in the central and late stages of the PPP. The variety Benedetta shows high expression of *UGT* and *CHS2*.

In 2017 six major clusters can be recognized by setting as threshold value for the Pearson correlation a value of 0.93 (Figure 11B): the first two are the same as those of 2016, *PAL4*/*4CL1* and *UGT*, the third comprises only *4CL5*, the fourth *C4H*, *F3H* and *CHI3*, the fifth cluster comprises *PAL2* and *DFR*, the last *CHI*, *CHS* and *ANS*. It is possible to observe again the clustering of *PAL4* with *4CL1* and of *PAL2* with *DFR*, as previously seen for the fruits sampled in 2016. The commercial variety shows instead differences, notably higher expression of the genes involved in the central and late stages of the phenylpropanoid pathway. Morellona confirms the high expression of genes involved in flavonoids/anthocyanin biosynthesis. Maggiola displays a high expression of *UGT* as previously observed in 2016 (Figure 11A).

The year 2018 differs from the previous 2 years in terms of gene expression (Figure 11C). By setting a threshold value of 0.88 for the correlation coefficient, five major expression clusters are observed: the first groups *PAL2*, *4CL1* and *DFR*, the second *CHI* and *CHS*, the third comprises *4CL5*, *PAL4*, *C4H* and *ANS*, the fourth cluster groups *CHI3* and *F3H* and the last one is represented by *UGT*. *PAL2* and *DFR* are in the same cluster, however in the year 2018 *4CL1* is also present, differently from the previous years, where it groups with *PAL4*. *PAL2* shows lower expression in Morellona with respect to 2016 and 2017, while the commercial fruits show much higher expression of all the genes, with the exception of *UGT*.

Analysis of the soluble proteome

The spectrophotometric and metabolite quantification (both targeted and untargeted) revealed that the varieties Morellona and Crognola are the highest producers of bioactive compounds. Therefore, they are interesting for an economic local exploitation as resources for the manufacture of high-quality food products. Part of the PhD studies was therefore devoted to high-throughput analyses of these two varieties. An analysis of the soluble proteome was carried out, in order to highlight differences in the abundance of soluble enzymes, such as those intervening in the PPP (CHS, CHR[130]), that could explain the metabolite and gene expression differences observed in the three years.

The two-dimensional difference gel electrophoresis (2D-DIGE) experiments showed 166 differentially abundant spots corresponding each to a single protein. The spots were selected according to the previously described SameSpot settings (max fold change>2 and *p*-value <0.01).

The spot identification was done through peptide sequence searches in the MASCOT database. The search was conducted against the NCBI non-redundant protein database of Viridiplantae and of *P. avium*. From an initial number of 166 proteins, after the removal of hits corresponding to the same protein, 52 proteins were retrieved with a good e-value and similarity sequence score, according to BLASTp.

Spot number	Accession number	Protein Name	Category	*p*-value (year)	*p*-value (variety)
753	XP_021804554.1	Thioredoxin reductase NTRB	REDOX	0.12515	0.00297
996	XP_021801592.1	Superoxide dismutase	REDOX	0.00965	0.14682
632	XP_021832997.1	Protein SRC2	REDOX	0.27232	0.00004
1004	XP_021816866.1	2-Cys peroxiredoxin BAS1	REDOX	0.7117	0.00026
988	XP_021823852.1	Glutathione S-transferase F11	REDOX	0.66672	0.00001
976	XP_021813812.1	Glutathione S-transferase DHAR2	REDOX	0.00007	0.32618
439	XP_021830782.1	Polyphenol oxidase	REDOX	0.00005	0.80611
299	XP_021815855.1	Stromal 70 kDa heat shock-related	ENVIRON	0.17303	0.00017
948	XP_021820221.1	Low-temperature-induced cysteine proteinase	ENVIRON	0.81184	0.00003
845	XP_021820409.1	Glycine-rich protein 2	ENVIRON	0.0056	0
1163	XP_021824432.1	17.1 kDa class II heat shock protein	ENVIRON	0.00695	0.00163
305	XP_021813758.1	Heat shock cognate 70 kDa protein 2	ENVIRON	0.08196	0.00561
985	XP_021805019.1	20 kDa chaperonin	ENVIRON	0.0128	0.00019
792	XP_021812280.1	Xyloglucan endotransglucosylase 31	WALL	0.4535	0.00001
477	XP_021823827.1	UTP-glucose-1-phosphate uridylyltransferase	WALL	0.40482	0.00437
822	XP_021802094.1	Xyloglucan endotransglucosylase 6	WALL	0.01025	0.00207
347	AAA91166.1	Beta-glucosidase	WALL	0.00037	0.56739
221	XP_021814417.1	Alpha-xylosidase 1	WALL	0.08851	0.00005
201	XP_021823469.1	Acid beta-fructofuranosidase 1	WALL	0.08196	0.00561
511	XP_021814422.1	Ubiquitin receptor RAD23c	PRIMARY	0.00018	0.91013
1012	XP_021825182.1	Proteasome subunit beta type-6	PRIMARY	0.0371	0.00475
936	XP_021828441.1	Proteasome subunit alpha type-6	PRIMARY	0.00762	0.03074
46	XP_021826137.1	Patellin-3-like	PRIMARY	0.26719	0.00001

1009	XP_021804963.1	Uncharacterized protein LOC110749212	PRIMARY	0.43022	0.00502
910	XP_021819098.1	Soluble inorganic pyrophosphatase 4	PRIMARY	0.04654	0.00001
814	XP_021820843.1	Short-chain dehydrogenase TIC 32	PRIMARY	0.14667	0.00616
378	XP_021806563.1	Protein disulfide-isomerase	PRIMARY	0.84562	0
725	XP_021806982.1	Voltage-gated potassium channel subunit beta	PRIMARY	0.00007	0
1023	XP_021808495.1	NAD(P)H dehydrogenase (quinone) FQR1	PRIMARY	0.00036	0
674	XP_021804100.1	Plastoglobulin-1	PRIMARY	0.05678	0.00082
802	XP_021807453.1	Mitochondrial outer membrane protein porin of 36 kDa	PRIMARY	0.43668	0.00712
709	XP_021804616.1	Malate dehydrogenase	PRIMARY	0.0004	0.00121
258	XP_021832122.1	Phosphoenolpyruvate carboxykinase	PRIMARY	0.00066	0.00002
1293	XP_021813779.1	2-oxoglutarate dehydrogenase	PRIMARY	0.00439	0.00327
1453	XP_021823850.1	Isocitrate dehydrogenase	PRIMARY	0.06206	0.00213
887	XP_021828712.1	GTP-binding nuclear protein Ran-3	PRIMARY	0.58551	0.00403
411	XP_021825040.1	Glucose-6-phosphate 1-dehydrogenase	PRIMARY	0.15166	0.00721
883	XP_021831593.1	Gamma carbonic anhydrase 1	PRIMARY	0.39206	0
624	XP_021810138.1	Fructose-bisphosphate aldolase 1	PRIMARY	0.01983	0.00042
209	XP_021818418.1	Eukaryotic translation initiation factor 3 subunit B	PRIMARY	0.01917	0.00332
1024	XP_021804938.1	Auxin-binding protein ABP19a	PRIMARY	0.00007	0.00033
66	XP_021824244.1	ATP synthase subunit beta	PRIMARY	0.14432	0.00849
772	XP_021825752.1	Annexin-like protein RJ4	PRIMARY	0.00198	0
709	XP_021804616.1	Glutathione *S*-transferase	PRIMARY	0.24001	0.00043
161	XP_021822858.1	Aconitate hydratase	PRIMARY	0.00664	0.41726
505	XP_021817392.1	6-phosphogluconate dehydrogenase. decarboxylating 3	PRIMARY	0.16838	0.00277
1238	XP_021831271.1	Phosphatidylglycerol/phosphatidylinositol transfer protein	OTHER	0.15061	0.00045
957	XP_021826507.1	Ferritin-4	OTHER	0.03802	0.00003
944	XP_021820122.1	Ferritin-3	OTHER	0.00021	0.00003
188	XP_021828563.1	Elongation factor 2	OTHER	0.06334	0.00455
761	XP_021821654.1	Elongation factor 1	OTHER	0.45155	0.00002
251	XP_021834633.1	5-methyltetrahydropteroyltriglutamate	OTHER	0.02626	0.000000

Table 7: Details of the spot numbers, accession numbers, annotations and *p*-values (the *p*-values <0.01 are highlighted in light green) obtained with the ANOVA (reporting differences among years and between varieties).

The hierarchical clustering of the heat maps shows two major patterns which correspond to the 2 genotypes, by choosing a Pearson correlation coefficient >0.94: one group is represented by proteins that are more abundant in Morellona and one by those that are higher in Crognola (Figure 12). It is possible to observe variability in protein abundance for Crognola in the harvest year 2016: if this result is compared to the PCA described previously for the genes (Figure 10), one confirms the greater heterogeneity of the biological material in 2016.

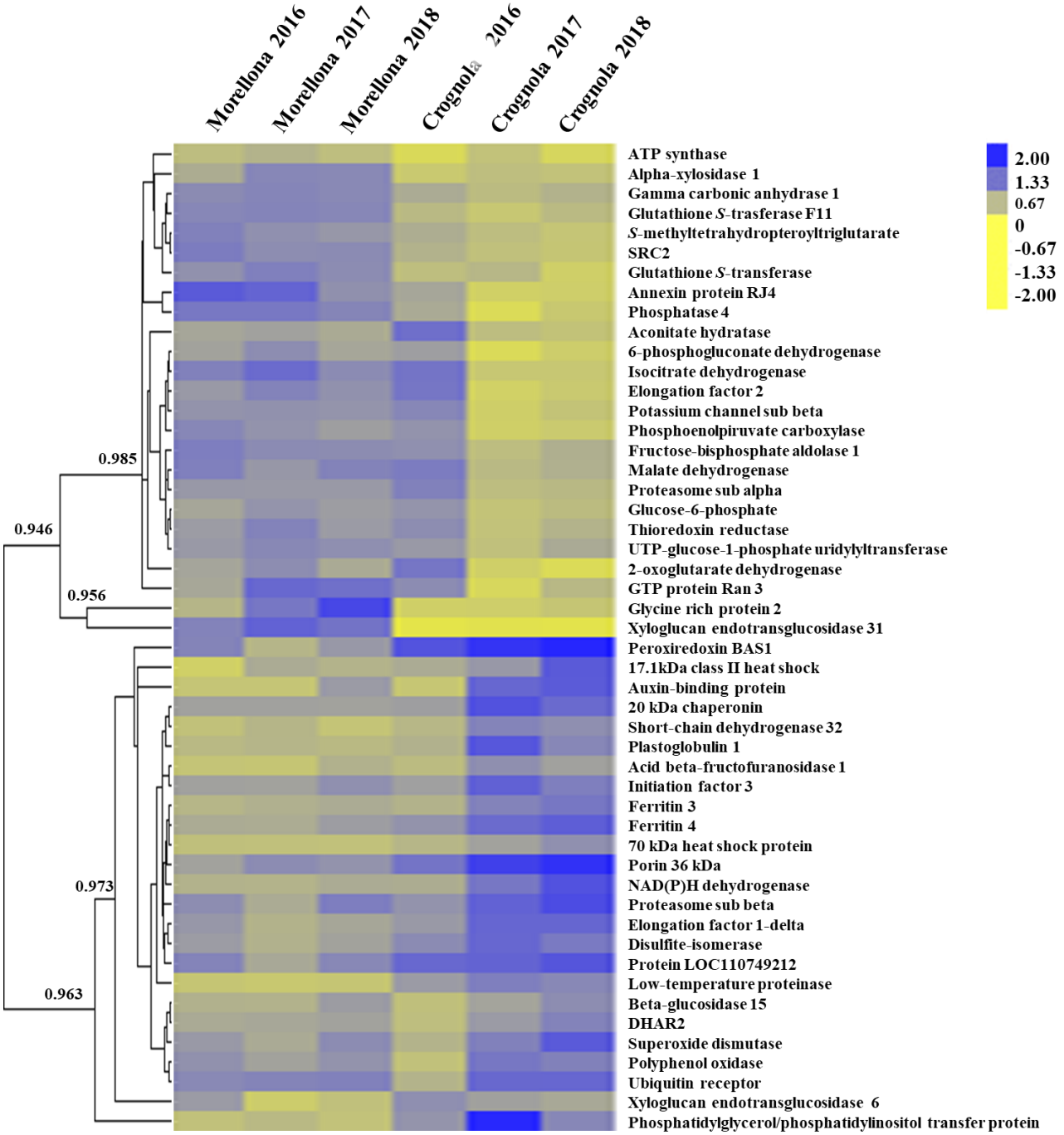


Figure 12: Heat map hierarchical clustering of 52 proteins obtained by the significant variations detected between the two varieties and among the three years of study. The scale bar indicates the abundances.

The results are hereafter divided by protein category.

Metabolism of carbohydrates and proteasome-related pathway

As Table 7 shows, more than the 50% of the differences between the two varieties is represented by proteins related to primary metabolism. Although these proteome fluctuations may be explained by the different genotypes, plant primary metabolism, which provides the precursors needed for the synthesis of secondary metabolites, is significantly influenced by environmental conditions [such as the (a)biotic constraints encountered in the field during the different years studied].

We identified 3 different proteins related to carbohydrate metabolism (Table 2), glucose-6-phosphate dehydrogenase, 6-phosphogluconate dehydrogenase and fructose-bisphosphate aldolase. All of these proteins show significant differences between the two cherry varieties and therefore appear to be genotype-dependent. Glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase are involved in the oxidative pentose phosphate pathway and their role, despite linked to glucose oxidation, is anabolic, rather than catabolic.

We detected 14 differentially abundant proteins directly related to cell energy production. Most of these differences are between the two varieties, suggesting a potential genotype effect. Otherwise few proteins, such as phosphoenolpyruvate carboxykinase, malate dehydrogenase, NAD(P)H dehydrogenase and 2-oxoglutarate dehydrogenase, show variations both between the varieties and harvest years.

In order to cope with the organism's demands and to maintain the normal functions, cells require a continuous turn-over of proteins, operated, among other actors, by the proteasome system. Ubiquitination and the proteasome activity are responsible for the degradation of proteins, a mechanism allowing stressed plants to fuel the mitochondrial metabolism[131]. In our dataset, we identified the ubiquitin receptor RAD23c, the proteasome subunit alpha type-6 and the proteasome subunit beta type-6.

Proteins related to cell redox balance

Concerning the proteins related to stress response, 7 proteins linked to different cellular redox mechanisms were identified.

In Table 7, significant differences for ROS scavenging enzymes are reported, notably for SOD superoxide dismutase (SOD), glutathione-*S*-transferase (GST), dehydroascorbate reductase (DHAR2) and GST F11. DHAR belongs to a subclass of enzymes in the wide group of the GST enzymes, able to catalyse the reduction of dehydroascorbate to ascorbate, with the concomitant oxidation of reduced GSH to glutathione disulfide, as reported by Dixon and Edwards[132]. GST F11 belongs to another GST group carrying out different reactions from the main GST class. Indeed, GST

F11 lacks a serine at the active site residue and, therefore, does not catalyse classical GST reactions. Instead, GST F11 seems to have a role in glucosinolate metabolism[132,133].

The different abundance of these two proteins is mostly related to the year of harvest, a finding suggesting an influence of the environmental conditions. Conversely, GST F11 is highly abundant in the variety Morellona, showing a pattern tightly related to the genotype.

Although the well-known enzymes SOD and GST carry out a central role in scavenging ROS, other enzymes are implicated in this process, with both direct and undirect effects. In the protein dataset obtained, other proteins related to ROS scavenging were indeed identified and these are thioredoxin (TRX), peroxiredoxin (PRX), polyphenoloxidase (PPO) and the product of the *SRC2* gene (*SOYBEAN GENE REGULATED BY COLD 2*).

In plants, the presence of TRX and PRX (BAS1) is linked with a control of ROS. These two proteins act together by using NADPH as a source of reducing power to contribute to the plant antioxidant systems, mostly in the chloroplasts[134,135]. PRX is commonly detected in plant proteomes and is responsible for the signalling related to ROS[136].

Two additional proteins related to redox reactions and the production of ROS are here presented, i.e. PPO and SRC2. In the dataset, PPO shows a significant difference in abundance among the three years of study, with the most noteworthy differences in Crognola (Figure 12).

Although, as explained in the Discussion section, SRC2 is related to temperature stress, it is more appropriate to show it here, as its role is to activate the NADPH oxidase-mediated ROS production.

The abundance of SRC2 is significantly higher in Morellona than Crognola, which suggests a genotype effect, since the cherries were always harvested at the same moment.

Proteins responding to external cues

Proteins related to the response to external cues (linked to the environment) were also identified as differentially abundant in the 2 varieties: more specifically, a stromal 70 kDa heat shock-related protein, a low-temperature-induced cysteine proteinase, a glycine-rich protein 2 (GRP2), a 17.1 kDa class II and a 70 kDa heat shock protein (HSP) and a 20 kDa chaperonin were identified (Table 5).

With the exception of GRP2, the proteins related to the response to external cues are expressed at higher levels in the variety Crognola.

Domain analysis of sweet cherry GRP2 with Motif Scan (https://myhits.isb-sib.ch/cgi-bin/motif_scan) revealed the presence of a glycine-rich, as well as of a WDS domain (e-value = 2.9e-45). ABA/WDS domain refers to abscisic acid and water deficit stress and is found in the dual transcription factor/chaperone protein ASR1 (ABA, stress and ripening) which is induced in tomato upon drought and is expressed during ripening[137].

Crognola showed higher abundance of HSP70, 20 and 17.1 (members of small HSPs; Table 7) which are involved in the tolerance to abiotic stresses[138,139], as well as of a low-temperature-induced cysteine proteinase showing the presence of granulin domains (e-values = 1.2e-10, 2.1e-06).

Proteins related to the cell wall

The proteomic analysis also revealed 5 proteins related to the cell wall: xyloglucan endotransglucosylase/hydrolase 31 (XTH31), xyloglucan endotransglucosylase/hydrolase 6 (XTH6), alpha-xylosidase 1, a β-glucosidase and a UTP-glucose-1-phosphate uridylyltransferase.

XTHs display both xyloglucan *endo*-transglucosylase (XET, cutting and rejoining xyloglucan chains) and xyloglucan *endo*-hydrolase (XEH, hydrolysis of xyloglucan) activities[140–143]. The majority of XTHs enzyme kinetics data showed the predominant presence of XET activity; a bioinformatic analysis coupled to structural data and enzymology predicted AtXTH31 and 32 as potential hydrolases belonging to clade III-A[144]. Subsequent studies showed that AtXTH31 accounts for the majority of XET activity in thale cress roots and has a pivotal role under Al stress[145]. A phylogenetic analysis of thale cress, tomato and nasturtium XTH showed that the cherry XTH31 clusters in group III-A, together with AtXTH31, the paralog AtXTH32 and nasturtium TmNXG1 (a predominant xyloglucan hydrolase,[144] Figure 13).

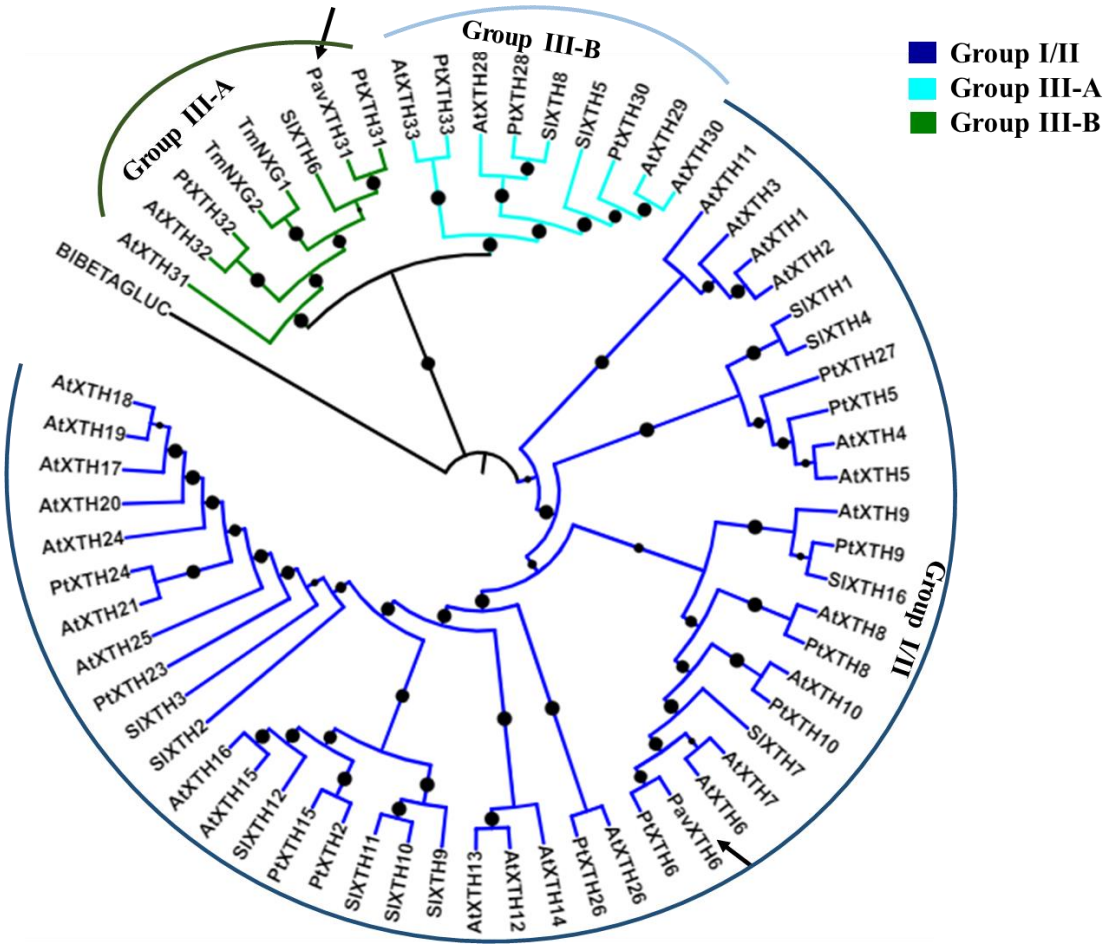


Figure 13: Maximum-likelihood phylogenetic tree of poplar, nasturtium, thale cress and sweet cherry XTHs. The tree is rooted with a bacterial lichenase (CAA40547). Bootstraps = 100. Circles indicate bootstrap values ≥80%. The bigger the circles, the higher the bootstrap values. The arrows indicate the sweet cherry XTHs detected with proteomics.

The other XTH detected in the soluble proteomes of the two Tuscan varieties is XTH6, which clusters together with AtXTH6 (Figure 13). The abundance of XTH6 increased in thale cress shoots and roots under heat stress in response to cytokinin[146], while the transcript was downregulated upon drought stress in 6 different accessions of *A. thaliana*[147]. It is therefore reasonable to assume that in the Tuscan sweet cherry Crognola, the higher abundance of XTH6 is linked to environmental cues to which the variety reacted. This finding is corroborated by the previously described higher abundance of stress-related proteins in Crognola.

A β-glucosidase (BGLU) with sequence similarity to the apoplast-localized *A. thaliana* BGLU15 was also identified as more abundant in Crognola, with respect to Morellona. Interestingly, despite the apoplastic localization, this protein is not related to cell wall processes, but to the hydrolysis of flavonol 3-*O*-β-glucoside-7-*O*-α-rhamnoside (a flavonol bisglycoside acting as antioxidant and

reducing ROS damage), which occurs in thale cress during recovery from synergistic abiotic stresses (i.e. N deficiency, low temperature, high light intensity, UV light)[148,149].

Gene expression analysis on targets identified with proteomics

The expression of genes belonging to the categories REDOX, ENVIRON and WALL was measured to find a correlation with the abundances studied with proteomics. The results will be discussed in relation to both the year of harvest and to the genotype, as previously done for proteomics.
Nineteen qPCR primers were designed on the targets reported in Table 5. The primers are found in Table 2 and the relative genes coding for the proteins of interest are summarised in Table 8.

Target proteins	Category
Polyphenol oxidase	REDOX
SRC2	REDOX
Glutathione *S*-transferase F11	REDOX
Superoxide dismutase	REDOX
Thioredoxin reductase NTRB	REDOX
Glutathione *S*-transferase DHAR2	REDOX
2-Cys peroxiredoxin BAS1	REDOX
Stromal 70 kDa heat shock-related protein	ENVIRON
Heat shock cognate 70 kDa protein 2	ENVIRON
Low-temperature-induced cysteine proteinase	ENVIRON
17.1 kDa class II heat shock protein	ENVIRON
Glycine-rich protein 2	ENVIRON
20 kDa chaperonin	ENVIRON
Beta-glucosidase	WALL
Xyloglucan endotransglucosylase 31	WALL
Xyloglucan endotransglucosylase 6	WALL
Alpha-xylosidase 1	WALL
Acid beta-fructofuranosidase 1	WALL
UTP-glucose-1-phosphate uridylyltransferase	WALL

Table 8: Target proteins on which qPCR primers were designed. Their corresponding categories are also indicated.

Out of 19 genes studied, 7 (*PavBAS1*, *PavGLUST*, *PavHSP17*, *PavPPO*, *PavSRC2*, *PavXTH31* and *PavXTH6*) display statistically significant differences in expression between the two varieties and/or among the three years studied. Their gene expression pattern is shown in Figure 14.
The gene encoding the 2-Cys peroxiredoxin BAS1 shows no difference among the three years in Crognola, while in Morellona sampled in 2017, it shows the lowest expression. Generally, *BAS1*

shows lower expression in Morellona, thereby confirming the results obtained with proteomics (Figure 12 and Table 7).

Glutathione *S*-transferase F11 shows the highest amount in Morellona, as it is visible in Figure 14. Despite the low expression in Morellona in 2018, the relative gene *GST F11* is highly expressed in this variety in the years 2016 and 2017, following the trend of proteomics.

The gene *HSP17* shows variations among the years of harvest, mostly in the variety Crognola. According to Table 7, the protein HSP17 shows statistically significant variations across the years of study, confirming the gene expression results. Moreover, the high protein abundance in Crognola in 2018 matches the high expression in Figure 14.

The gene *PPO* shows variations in the 3 years in Morellona, especially in 2016 and 2018. In accordance with these data, the relative protein shows statistically significant variations across the years.

Figure 14 shows a clear trend for the gene *SRC2* that is expressed at higher levels in the variety Morellona. Interestingly, the relative protein is also more abundant in Morellona (Figure 12).

The gene coding for XTH31 is expressed at higher levels in Morellona (Figure 14), as previously seen also for the protein abundances (Figure 12). Differently from what observed with proteomics, the gene encoding XTH6 does not show a statistically significant difference between the two varieties. In 2018 *XTH6* is induced in Morellona (Figure 14).

It is known that gene expression changes are not always accompanied by a similar trend in the corresponding proteins[150]. Indeed, only 7 out of 19 targets show statistically significant differences. Of these 7 genes, 3 follow the same trend as the corresponding proteins.

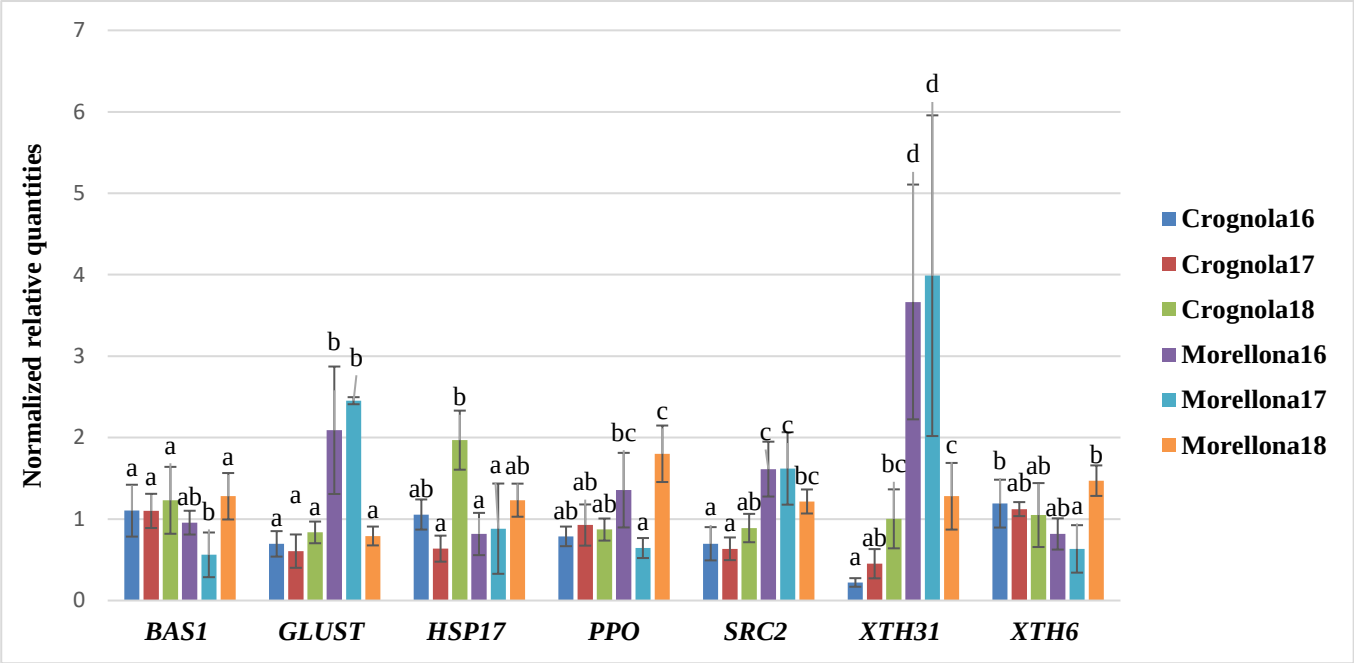


Figure 14: Relative expression of some genes coding for differentially abundant proteins in the two Tuscan varieties. Error bars refer to the standard deviation (number of independent biological replicates = 4). Different letters indicate statistically significant differences among the groups at the one-way ANOVA with Tukey's post-hoc test.

Discussion

The research done during the PhD course allowed, for the first time, a thorough characterisation of the Tuscan repertoire of sweet cherries. Several interesting features were thus revealed. This section of the thesis will start with the discussion of the data relative to the functional molecules.

From a nutraceutical point of view, the Tuscan cherry varieties show a wide range of health-beneficial compounds, such as chlorogenic acids (CGAs), flavonoids and anthocyanins with antioxidant potential.

Plants use these molecules to counteract the excess of ROS inside their cells.

Most of the compounds detected belong to the CGA class. Interestingly, CGAs (i.e. chlorogenic acid, *p*-coumaric acid and coumaroyl quinic acid) have been quantified and studied in beverages, such as coffee, because of their health beneficial role[151]. The presence of CGAs (Table 3 and 4) confirm the high amount of phenolic acids in the ancient sweet cherry varieties (especially in Morellona and Crognola) and, therefore, their nutraceutical value.

The high health-promoting activity of the Tuscan cherries is confirmed by the data relative to the content of flavonoids and anthocyanins. Naringenin, rutin and catechins are the most important flavonoids belonging to the flavanone and flavan-3-ol classes, which are predominantly found in some edible fruits, such as those of the species *Citrus* and in tomatoes[152]. Wang and colleagues studied the effects of flavonoids (naringenin, in particular) in rat livers and kidneys to counteract the oxidative stress caused by lead. The authors proved a significative protective effect, thereby proposing this compound as useful in the treatment of lead toxicity[153]. Nemes and co-authors detected the presence of naringenin in polyphenolic extracts of sour cherries[154].

The spectrophotometric assay (Table 3), as well as the targeted (Table 4) and untargeted approaches (Figure 9) revealed that the Tuscan cherries are also rich in anthocyanins and proanthocyanidins. These molecules have a wide range of health-beneficial properties, such as antioxidant, anticancer and immunostimulatory[155]. Studies have shown that proanthocyanidins improve vision, flexibility of joints, blood circulation by strengthening capillaries, arteries and veins[156].

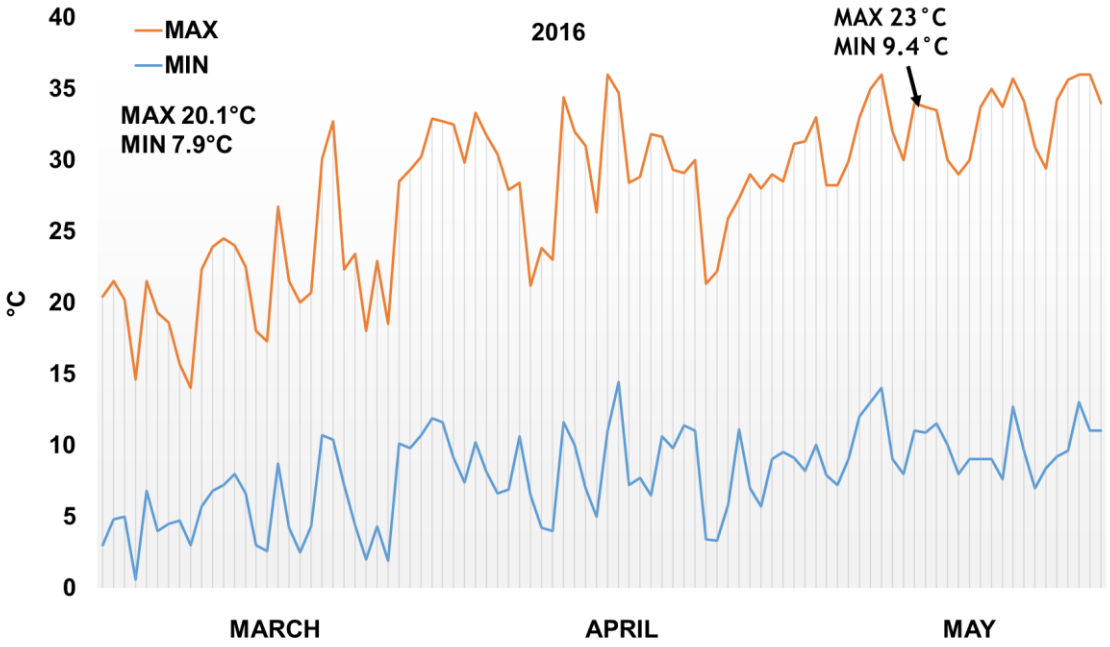
It is worth noting that untargeted metabolomics revealed the presence of a flavonolignan, cinchonain. Cinchonain is known to be a strong antioxidant compound able to modulate a variety of cell-signalling pathways, resulting in the reduction of pro-inflammatory responses[157]. Furthermore, Olszewska and colleagues showed that its antiradical properties (towards DPPH) are up to four times stronger than those of (+)-catechin—one of the most effective antioxidants both *in vitro* and *in vivo*[158]. In addition to its antioxidant potential, *in vitro* experiments showed the insulinotropic effects of cinchonain and its possible use for type 2 diabetes[159].

Although further studies including commercial fruits grown in the same conditions as the local varieties are needed, the results obtained with the complementary approaches here used suggest that the Tuscan cherries are often richer in functional molecules. These results are in agreement with previous work by Iacopini and colleagues[18] and focused on other Tuscan fruits, notably apples. Exploiting this potential via a local cultivation would certainly contribute to foster the regional economy and promote more sustainable agricultural practices, since these varieties need minimal human input. However, it is worth discussing here another potential way to valorise these varieties: besides the *in vitro* propagation of the ancient varieties, the establishment of cell suspension cultures, for example from the fruits of the highest producers, namely Crognola and Morellona, would be an interesting way to scale up the production of the health-promoting compounds. Green extraction procedures avoiding the use of toxic solvents can then be applied in downstream processing to recover the active ingredients produced by the plant cell cultures. Such an approach is particularly promising if one considers that the nutraceuticals market is witnessing an increase, according to the latest estimates. Indeed, it is predicted to record a revenue of USD 671.30 billion by 2024, registering a Compound Annual Growth Rate (CAGR) of 7.5% during the forecast period 2019–2024 (source: https://www.mordorintelligence.com/industry-reports/global-nutraceuticals-market-industry). This trend reflects the key role of nutraceuticals in the consumer's daily diet and is the result of the increasing prevalence of chronic diseases on one hand and of the preventive healthcare measures consciously taken by people on the other.

The production of metabolites is the result of the expression of genes. The data obtained on the PPP-related genes indicate that the environmental conditions of the 3 years play a prominent role in controlling the expression of the genes. In the literature, it is known that PPP-related genes respond to exogenous stresses[160]. For example, *Colobanthus quitensis* grown in the field showed an induction of genes involved in the PPP as compared to chamber-grown plants[138]. One of the promptest effects of plant cells exposed to abiotic stresses is the production of reactive oxygen species (ROS), i.e. superoxide ($O^{\bullet-}_2$) and hydroxyl radicals ($^{\bullet}OH$), as well as non-radicals, such as hydrogen peroxide (H_2O_2) and singlet oxygen ($^1O_2$)[161,162]. A dual cellular system based on both enzymatic and non-enzymatic players scavenge ROS produced in plants[163]. The first mechanism relies on the enzymes superoxide dismutase (SOD), ascorbate/guaiacol peroxidase (APX/GPX), glutathione-*S*-transferases (GST) and catalase (CAT). The second mechanism is the non-enzymatic response regulating the production of polyamines, as well as low-molecular weight compounds. These are reduced glutathione (GSH), vitamins like ascorbic acid and α-tocopherol, osmolytes like proline and secondary metabolites deriving from PPP, notably flavonoids, phenolics, carotenoids[164].

Plant secondary metabolites are typically produced in response to changes in the environment or pathogen/herbivore attack and are currently in the spotlight of biotechnology because of their applications in different industrial sectors (i.e. healthcare, cosmetics, to name a few[37]). The chemical structures of plant phenolics are able to counteract ROS and reactive nitrogen species (RNS) by acting as natural scavengers[38] and via the previously-mentioned hydrogen atom transfer (HAT) mechanism, involving -OH groups that act as donors[55]. In strawberry, an effect of the environment was reported on the content of flavonoids and the expression of the related genes in genotypes growing in different geographical locations[165]. In this thesis genotypes growing in the same location were studied; however, the results show differences among the 3 years investigated.

To understand the environmental causes that could be (at least in part) responsible for the differential gene expression in the Tuscan varieties, the daily temperatures were recorded with the LaMMA meteorological station placed in the experimental field. As can be seen in Figure 15, 2018 was the hottest year, if the minimum mean temperatures are considered. Moreover, on the day of sampling in May 2018 (reported in Figure 15), the temperatures were ca. 3°C higher.



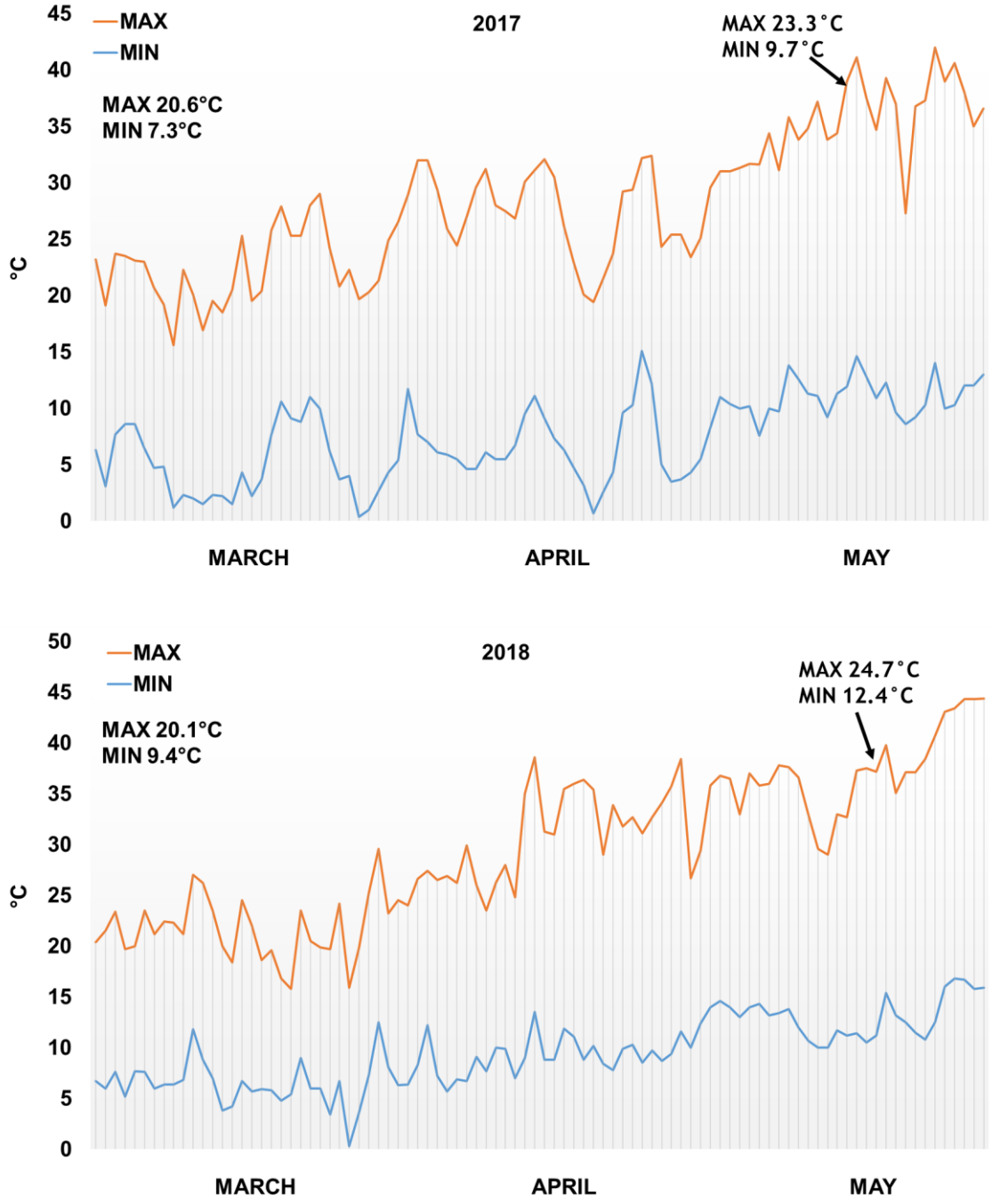


Figure 15: Temperature profiles reported by the LaMMA station placed in the experimental field and relative to the tree years of harvesting (i.e.2016, 2017, 2018). The average of maximum and minimum temperatures are shown in the graphs. The arrows show the day of harvest.

Despite the less discriminative PCA pattern in 2016, the hierarchical clustering of the heat map revealed a similar clustering of the genes in 2016 and 2017 (Figure 11A and B). The year 2018 shows differences instead. Furthermore, the contents of flavonoids, namely catechins and rutin, are higher in 2018 compared to the other years of the survey, a finding confirming the different trend observed in 2018.

The hierarchical clustering of the heat maps revealed interesting co-expression patterns of the 2 *PAL* isoforms investigated, *PAL2* and *PAL4*, in the Tuscan fruits (Figure 11). *PAL2* clusters with *DFR*, while *PAL4* clusters with *4CL1*. In thale cress four different *PAL* isoforms have been described[166] and a redundant role for PAL1 and PAL2 was shown in flavonoid formation[167]. The sweet cherry PAL2 may be involved in the PPP branch shunting precursors towards the synthesis of flavonoids. It will be interesting to carry out functional studies using for example transient expression in tobacco (*Nicotiana benthamiana*) leaves, to check whether the overexpression of sweet cherry *PAL2* and *DFR* leads to an increase in flavonoid synthesis. It will also be highly interesting to study whether PAL2 and DFR interact in sweet cherry, either directly or indirectly, to form a complex.

The gene expression analysis also revealed a higher expression of the genes involved in the central and late stages of the PPP in Morellona (Figure 11). It should be noted that this ancient variety is the only one to have a red pulp; therefore, the high content of anthocyanins (Table 3) can be explained by the higher gene expression.

The entry point into the production of catechins, which were shown to be abundant in Benedetta (Table 4), is the enzyme CHS and the corresponding gene here studied, *CHS2*, which is highly expressed in this variety (Figure 11), may thus represent a regulatory check-point controlling the shunt of phenolic precursors needed for the synthesis of catechins.

The thesis provides a molecular snapshot of the Tuscan fruits relative to one sampling point. In the future, it will be interesting to address the gene expression changes on fruits sampled at different developmental stages. Particularly relevant in this respect will be the analysis of the two highest producers of bioactives, i.e. Crognola and Morellona, which are the most promising for an eventual exploitation, e.g. the manufacture of local products of high nutraceutical value.

The targeted and untargeted approaches, together with the gene expression analysis, revealed differences among the Tuscan varieties and the years, which are explained by the genotypes and the environmental conditions.

To shed light on the phylogenetic relatedness of the Tuscan cherries, an SSR screening was performed, by including varieties from Luxembourg, France and Turkey. The results showed that the varieties Crognola and Morellona cluster together in a separate branch (Figure 7), a result showing that, also genotypically, they are more distant from all the other Tuscan varieties studied. It is natural, therefore, to associate this phylogenetic distance to the higher production of bioactives of these two regional varieties. Genetic polymorphisms can indeed account for metabolomics differences among a plant population, as recently demonstrated in rice[168].

Since metabolites are the product of enzyme activities, a proteomic analysis was performed on the 2 highest producers of phenolics, Morellona and Crognola. Few differences were found in PPP-related

proteins between the two varieties, as the main distinguishing traits are related to primary metabolism, maintenance of redox balance, stress response and cell wall proteins (Figure 12).

Primary metabolism ensures basal cell functions, like the provision of carbohydrates, which are fundamental for plants and are directly related to the photosynthetic activity. In stress conditions, the photosynthetic performance is compromised and the storage carbohydrates may be converted into free sugars and used as energy reserves. Free sugars are also described by Arbona and colleagues as osmolytes and, in certain cases, ROS scavengers[169]. Indeed, it is known that glucose-6-phosphate dehydrogenase generates reducing power in the form of NAD(P)H and that it has a role both to sustain nitrogen assimilation[170] and to counteract stress conditions[171].

Carbohydrate metabolism is tightly linked to cell energy production; these systems act together with chloroplast and mitochondria activities to maintain optimal energy levels inside plant cells, especially under stress conditions, when defence responses are activated[172].

In this respect, the proteins phosphoenolpyruvate carboxykinase, malate dehydrogenase, NAD(P)H dehydrogenase and 2-oxoglutarate dehydrogenase, detected as differentially abundant in the Tuscan varieties, play a fundamental role in the response to external stimuli, as reported in the literature for drought and salt stress[173].

Within plant cells, during normal growth and development, there is a dynamic balance between oxidants and antioxidants: for example, the ROS hydrogen peroxide (H_2O_2) is important in coordinating elongation, however its abundance has to be controlled to avoid oxidative damages to important macromolecules[38]. One of the most dramatic consequences caused by oxidative stress is premature cell senescence[174], that, in special organs such as fruits, triggers a loss of texture, flavour and a decrease in the abundance of health beneficial molecules.

Proteomics highlighted differences in ROS scavenging enzymes, i.e. SOD, GST, DHAR2 and GST F11 (Figure 12). However, although the high abundance of proteins related both to the enzymatic and non-enzymatic scavenging mechanisms is indicative of an efficient antioxidant system in the ancient Tuscan sweet cherries, it is difficult to explain their different abundance in the two varieties without further functional studies.

The higher abundance of GST F11, involved in the glucosinolate pathway[175], is interesting in light of the health-related effects of isothiocyanates, molecules obtained through the action of myrosinase on glucosinolates (and typical of vegetables, such as members of the Brassicaceae) and displaying anticancerogenic activity[176]. It remains to be verified whether Morellona produces more glucosinolates than Crognola, as the differences observed may also be due to biotic stress, since glucosinolates are typically synthesized in response to herbivore attack[177].

PPO is the only detected protein directly related to PPP: this enzyme catalyses the polymerization of quinones formed through the oxidation of phenols[178] to produce brown pigments. Its role was linked to the defence against biotic stresses, such as insect attacks, as quinones bind to leaf proteins thereby inhibiting digestion by herbivores[178]. Other authors also reported an increase in abundance of this class of proteins in response to altered environmental conditions[179].

SRC2 is recognized as a cold stress marker. Indeed, scientific works reported an increase of this protein in plant tissues exposed to low temperature[180]. Despite the main role in cold stress response, Kawarazaki and co-authors proposed a novel implication of SRC2 in ROS scavenging: it acts together with the physiological Ca^{2+} release[181], in cold stress conditions, to activate the NADPH oxidase RbohF[180]. Indeed SRC2 can bind of the N-terminal region of RbohF and in this way it regulates ROS release[180] (Figure 16).

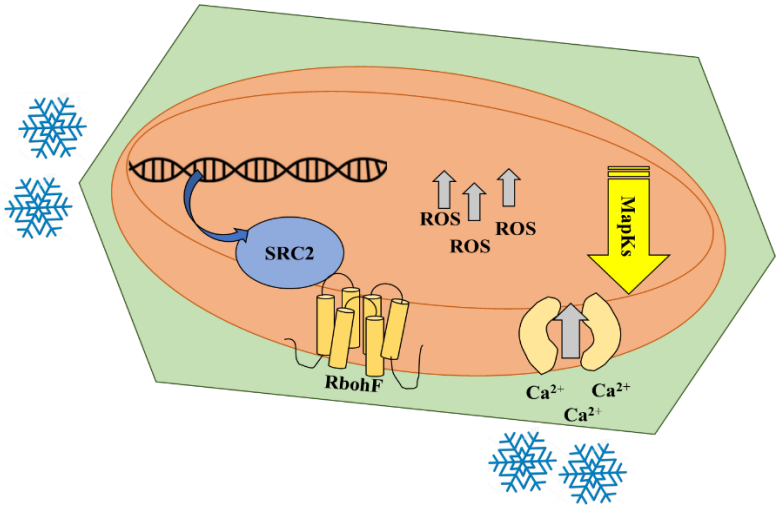


Figure 16: Mechanism of action of SRC2. The cold-induced SRC2 protein binds the N-terminus of RbohF (NADPH oxidase) inducing its activation and the consequent released of ROS. This may either activate cold-responsive genes or MAPK signalling[180].

In general, Crognola shows a higher abundance of proteins responding to external cues, a finding suggesting a higher tolerance to environmental stress. The protein GRP2 is instead more abundant in Morellona: this could be linked to abiotic stress event(s) (e.g. drought) and/or to ripening. The glycine-rich WDS-containing protein from *Tetragonia tetragonoides* TtASR increased ROS scavenging and tolerance to salt and drought, when overexpressed in thale cress[182].

Among the proteins with multiple functions, it is worth discussing the low-temperature-induced cysteine proteinase. This enzyme belongs to the C1 peptidases and are involved in stress response, as well as in fruit ripening, senescence, accumulation and mobilization of storage proteins[183]. The higher abundance of the protease in Crognola may be related to the post-transcriptional activities required during the fruit maturation stage[184], as well as to immunity. Indeed, a BLASTp against thale cress

using the MEROPS database[185] revealed high sequence similarity (e-value = 0, identities% = 67.5%, positives% = 82.4%) with RD21A papain-like cysteine protease which was previously shown to be involved in leaf senescence[186], as well as immunity[187].

From the point of view of cell wall related proteins, the main differences between the two varieties are detected in XTHs, alpha-xylosidase 1, a β-glucosidase and UTP-glucose-1-phosphate uridylyltransferase. Xyloglucan endotransglucosylase/hydrolase (XTH) proteins have a crucial role in the remodelling of plant cell walls, by acting on xyloglucans, the predominant hemicellulose in primary cell walls of plants. Xyloglucans bridge cellulose microfibrils, thereby contributing to the mechanical properties of the cell walls and to morphogenesis[144]. It was demonstrated that thale cress XTH from group III-A are endohydrolases involved in tissue expansion and are dispensable for normal growth[188]. The higher abundance of XTH31 in Morellona with respect to Crognola may explain the differences in sizes of the fruits produced by the 2 varieties: as can be seen in Figure 16, Morellona shows indeed a statistically significant higher height and diameter. Alpha-xylosidases are involved in xyloglucan remodelling; the sweet cherry XYL1 identified via proteomics is orthologous to thale cress XYL1, an enzyme acting on the non-reducing end of xyloglucan oligosaccharides[189]. The higher abundance in Morellona is indicative of a higher xyloglucan remodelling at maturity. A previous study showed that *xyl1*/*axy3* mutants displayed reduced silique length and altered xyloglucan structure, where the hemicellulose is less tightly bound to other cell wall components[190]. Therefore, the BGLU protein seems to be linked to stress-related pathways, which in Crognola are more represented than in Morellona. Differences in the abundance of a UTP-glucose-1-phosphate uridylyltransferase were also detected between the two Tuscan varieties. A BLASTp analysis against thale cress revealed sequence similarity with UGP2, one of the two genes contributing to sucrose and cell wall biosynthesis[191]. The higher expression in Morellona may indicate an involvement in cell wall-related processes and in accommodating the request of nucleotide sugars during fruit maturation. The bigger size of Morellona cherries as compared to Crognola (Figure 17) may indeed require a higher provision of precursors for cell wall biosynthesis.

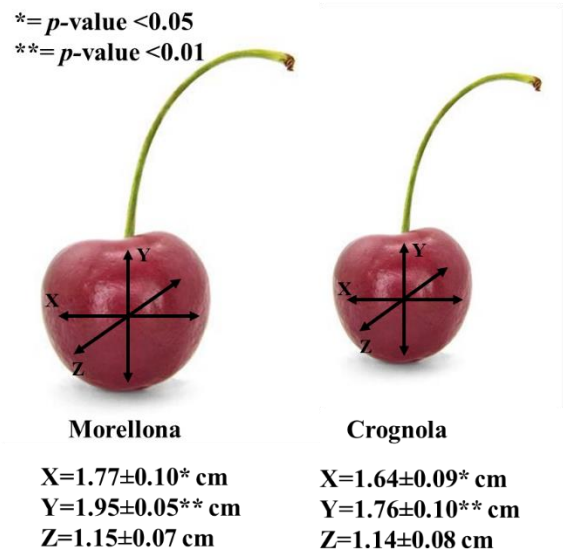


Figure 17: Details of the size of the two ancient sweet cherry varieties. The measures were performed on 20 fruits per variety. The asterisks highlight the significant differences in size.

Key results, conclusions and future perspectives

Innovation in the agricultural field is one of the most important topics in the scientific community, especially to foster and meet the requirements of a bioeconomy where *sustainability* is a keyword. One of the major societal challenges that Europe 2020 has to address is ensuring safe and healthy food while maintaining a minimal environmental impact (https://ec.europa.eu/programmes/horizon2020/en/h2020-section/food-security-sustainable-agriculture-and-forestry-marine-maritime-and-inland-water). Finding new and possibly sustainable agronomic techniques to improve crop productivity and resistance to climate factors is fundamental to meet the ever-growing commercial demand. Furthermore, cultivation strategies are being developed in the field of green technology. From this perspective, scientific interest in the use of compounds from natural sources (such as biostimulants), rather than synthetic ones has grown, as witnessed by the recent European Fertilising Products Regulation approved by both the Parliament and the Council of Europe (https://www.cencenelec.eu/News/Publications/Publications/WorkProgramme-2018_UK_acces.pdf).

Plants are renewable resources providing raw material (like lignocellulosic biomass) and phytochemicals (notably secondary metabolites) for different industrial applications, namely in the textile, construction, pharmaceutical, nutraceutical and cosmetic sectors.

In this scenario, non-commercial local varieties (designated as heirloom or ancient) play an important role in the diversification of the current market options and to give an economic boost at a regional level, via the manufacture of fully traceable products.

This thesis aimed at investigating, for the first time, a sub-set of sweet cherries belonging to the Tuscan plant repertoire, by adopting complementary molecular techniques. Hereafter is a summary of the key findings obtained in the thesis. It is important to highlight them to introduce the conclusions and future perspectives.

A genetic characterization of the six Tuscan cherries (Benedetta, Maggiola, Moscatella, Morellona, Crognola, Carlotta) was performed by using molecular markers (Figure 7). The results showed a good separation of the ancient varieties and commercial ones originating from Turkey, Luxembourg, Italy and France. The genotyping performed also revealed that two varieties, Crognola and Morellona, have the highest genetic distance, since they separate earlier from the cluster comprising all the other varieties investigated.

These molecular results were subsequently complemented by targeted and untargeted metabolite analyses on cherries sampled in three consecutive years (2016-2017-2018): the rationale behind this experimental choice was the need to assess whether the genotypic distance would be reflected in a different content in bioactive molecules. At first, a spectrophotometric quantification and HPLC analysis was carried out, by quantifying the total content of polyphenols/flavonoids/anthocyanins, as well as specific phenolic compounds known to be particularly abundant in sweet cherry (i.e. rutin, catechins, *p*-coumaric acid, cyanidin 3-*O*-glucoside,[192]). The results showed different abundances in the Tuscan fruits (Tables 3-4) which were particularly high in the varieties Crognola and Morellona. Interestingly, despite some differences across the years (likely due to environmental conditions), the ranking of the varieties in terms of bioactive content was constant. For comparative purposes, an Italian widespread commercial variety purchased at a local store was also analysed (variety Durone). Despite the different agronomic practices adopted for the Durone variety, both in pre-harvest and post-harvest treatments, it was decided to include them to have a comparison value from a commercial counterpart.

The targeted quantification of phenolics was enriched by an untargeted analysis which was chosen to confirm the results and to identify new compounds which could not be identified by HPLC. The data obtained confirmed the ranking of Crognola and Morellona as the highest producers of bioactives (Tables 5-6 and Figure 9) and provided information on other classes of phenolics, namely proanthocyanidins, flavonols, coumaroyl quinic acid and the flavolignan cinchonain.

Since it is known that the genes intervening in the PPP are differentially expressed in sweet cherry varieties[193,194], a targeted gene profiling was performed on the Tuscan fruits. The year 2016 was less

discriminative in terms of varieties, while a better separation was observed in 2017 and 2018 (Figure 10). The results showed also a high variability in the expression profiles (Figures 11). The hierarchical clustering of the data revealed very similar expression profiles for *PAL2* and *DFR*, regardless of the year, a finding suggesting a role for the PAL2 isoform in the shunting of precursors for flavonoid formation.

Crognola and Morellona were selected for an analysis of the soluble proteome: these varieties were chosen because they were the most distant in terms of genotype and the best producers of bioactives. The proteomic analysis on the mature cherries provided important information on aspects relevant for the post-harvest handling of the fruits. Overall, Morellona fruits are bigger than Crognola's and this is reflected in the higher abundance of proteins involved in the remodelling of xyloglucan (XTH31 and XYL1), a hemicellulose of the primary cell wall (Figure 12). The soluble proteome of Crognola is characterized by proteins responding to external cues, namely HSPs, as well as chaperonins, a finding suggesting a possible higher response to external cues.

From the results obtained it emerges that the Tuscan sweet cherries are interesting from a nutraceutical point of view. In particular, the varieties Crognola and Morellona are the most valuable in terms of bioactive content. At the sampling time, Morellona is characterized by processes that are indicative of cell wall remodelling: fruit maturation is accompanied by changes in the texture, due to the activities of enzymes acting on both cellulosic and non-cellulosic polysaccharides[195–197]. Crognola fruits, despite sampled at the same time-point as Morellona, do not show proteins related to cell wall remodelling and these results open the way to future experiments assessing the post-harvest stability. It may be that the 2 varieties show a distinct post-harvest behaviour which makes one of the 2 more suitable for a commercial exploitation.

In conclusion, the results obtained stress the importance of preserving and studying the regional repertoire of plants, because they can promote local agriculture and economy through their valorisation. Ancient non-commercial varieties represent a bridge between the past, in terms of popular culture and the future, given their possible use in innovative agricultural development and sustainable technologies. Future efforts should be devoted to the analysis, via systems biology approaches merging -*omics*, of underutilized varieties, such as the ancient local ones, to compare their content of bioactive molecules with respect to that found in commercial ones and to understand the molecular basis of such differences.

-*Omics* can be merged with epigenetics and phenotyping to unveil the molecular mechanisms underlying the expression of specific characters in ancient varieties.

Such studies favour the diversification of the current market of fruit and vegetables and promote programs aimed at the preservation of the regional agrobiodiversity.

References

1. DeLong, D. C. Defining biodiversity. *Wildl. Soc. Bull. 1973-2006* **24**, 738–749 (1996).
2. Likens, G. E. Some consequences of long-term human impacts on ecosystems. *Rev. Chil. Hist. Nat.* **64**, 597–614 (1991).
3. Bertacchini, E. Regional legislation in Italy for the protection of local varieties. *J. Agric. Environ. Int. Dev. JAEID* **103**, 51–63 (2009).
4. Berni, R. *et al.* Agrobiotechnology Goes Wild: Ancient Local Varieties as Sources of Bioactives. *Int. J. Mol. Sci.* **19**, 2248 (2018).
5. Fideghelli, C. & Engel, P. L'attività di raccolta, caratterizzazione, valorizzazione e conservazione della biodiversità vegetale di interesse agricolo in Italia con particolare riguardo alle risorse genetiche frutticole. *Italus Hortus* **18**, 33–45 (2011).
6. Tait, J. & Morris, D. Sustainable development of agricultural systems: competing objectives and critical limits. *Futures* **32**, 247–260 (2000).
7. Berni, R. *et al.* Functional Molecules in Locally-Adapted Crops: The Case Study of Tomatoes, Onions, and Sweet Cherry Fruits From Tuscany in Italy. *Front. Plant Sci.* **9**, (2019).
8. Timpanaro, G., Giudice, V. L. & Foti, V. T. The interest of consumers to safeguard plant biodiversity: The case of vegetable products. *Calitatea* **16**, 237 (2015).
9. Federico, M. *et al.* Ancient Pomoideae (*Malus domestica* Borkh. and *Pyrus communis* L.) cultivars in "Appenino Toscano"(Tuscany, Italy): molecular (SSR) and morphological characterization. *Caryologia* **61**, 320–331 (2008).
10. ASCANI, E. Piante autoctone. *Storia E Coltiv. Delle Piante Frutto Nel Territ. Senese Siena* (2007).
11. Berni, R. *et al.* Agrobiotechnology Goes Wild: Ancient Local Varieties as Sources of Bioactives. *Int. J. Mol. Sci.* **19**, 2248 (2018).
12. Pereira, H. M. *et al.* Scenarios for global biodiversity in the 21st century. *Science* **330**, 1496–1501 (2010).
13. Tscharntke, T., Klein, A. M., Kruess, A., Steffan-Dewenter, I. & Thies, C. Landscape perspectives on agricultural intensification and biodiversity–ecosystem service management. *Ecol. Lett.* **8**, 857–874 (2005).
14. Lange, M. *et al.* Plant diversity increases soil microbial activity and soil carbon storage. *Nat. Commun.* **6**, 6707 (2015).
15. Grzywacz, D., Stevenson, P. C., Mushobozi, W. L., Belmain, S. & Wilson, K. The use of indigenous ecological resources for pest control in Africa. *Food Secur.* **6**, 71–86 (2014).

16. Cerutti, A. K., Bruun, S., Donno, D., Beccaro, G. L. & Bounous, G. Environmental sustainability of traditional foods: the case of ancient apple cultivars in Northern Italy assessed by multifunctional LCA. *J. Clean. Prod.* **52**, 245–252 (2013).
17. Legay, S. *et al.* Differential Lipid Composition and Gene Expression in the Semi-Russeted "Cox Orange Pippin" Apple Variety. *Front. Plant Sci.* **8**, (2017).
18. Iacopini, P., Camangi, F., Stefani, A. & Sebastiani, L. Antiradical potential of ancient Italian apple varieties of *Malus* x *domestica* Borkh. in a peroxynitrite-induced oxidative process. *J. Food Compos. Anal.* **23**, 518–524 (2010).
19. Cantini, C., Cimato, A., Autino, A., Redi, A. & Cresti, M. Assessment of the Tuscan olive germplasm by microsatellite markers reveals genetic identities and different discrimination capacity among and within cultivars. *J. Am. Soc. Hortic. Sci.* **133**, 598–604 (2008).
20. Scarano, D. & Rao, R. DNA markers for food products authentication. *Diversity* **6**, 579–596 (2014).
21. Barbuto, M. *et al.* DNA barcoding reveals fraudulent substitutions in shark seafood products: the Italian case of "palombo"(*Mustelus* spp.). *Food Res. Int.* **43**, 376–381 (2010).
22. Rodriguez-Mateos, A. *et al.* Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: an update. *Arch. Toxicol.* **88**, 1803–1853 (2014).
23. Scarano, D., Rao, R., Masi, P. & Corrado, G. SSR fingerprint reveals mislabeling in commercial processed tomato products. *Food Control* **51**, 397–401 (2015).
24. Das, L., Bhaumik, E., Raychaudhuri, U. & Chakraborty, R. Role of nutraceuticals in human health. *J. Food Sci. Technol.* **49**, 173–183 (2012).
25. Kabera, J. N., Semana, E., Mussa, A. R. & He, X. Plant secondary metabolites: biosynthesis, classification, function and pharmacological properties. *J Pharm Pharmacol* **2**, 377–392 (2014).
26. Wallace, T. C. Anthocyanins in cardiovascular disease. *Adv. Nutr.* **2**, 1–7 (2011).
27. He, J. & Giusti, M. M. Anthocyanins: natural colorants with health-promoting properties. *Annu. Rev. Food Sci. Technol.* **1**, 163–187 (2010).
28. Wink, M. Introduction: biochemistry, physiology and ecological functions of secondary metabolites. *Annu. Plant Rev. Online* 1–19 (2018).
29. Seca, A. M. L. & Pinto, D. C. G. A. Plant Secondary Metabolites as Anticancer Agents: Successes in Clinical Trials and Therapeutic Application. *Int. J. Mol. Sci.* **19**, 263 (2018).
30. Berni, R. *et al.* Tuscan Varieties of Sweet Cherry Are Rich Sources of Ursolic and Oleanolic Acid: Protein Modeling Coupled to Targeted Gene Expression and Metabolite Analyses. *Mol. Basel Switz.* **24**, (2019).

31. Takachi, R. *et al.* Fruit and vegetable intake and risk of total cancer and cardiovascular disease: Japan Public Health Center-Based Prospective Study. *Am. J. Epidemiol.* **167**, 59–70 (2007).
32. DeFelice, S. L. The nutraceutical revolution: its impact on food industry R&D. *Trends Food Sci. Technol.* **6**, 59–61 (1995).
33. Pandey, K. B. & Rizvi, S. I. Plant polyphenols as dietary antioxidants in human health and disease. *Oxid. Med. Cell. Longev.* **2**, 270–278 (2009).
34. Rajasekaran, A., Sivagnanam, G. & Xavier, R. Nutraceuticals as therapeutic agents: A Review. *Res J Pharm Tech* **1**, 328–340 (2008).
35. Berni, R., Romi, M., Parrotta, L., Cai, G. & Cantini, C. Ancient Tomato (*Solanum lycopersicum* L.) Varieties of Tuscany Have High Contents of Bioactive Compounds. *Horticulturae* **4**, 51 (2018).
36. Siro, I., Kapolna, E., Kapolna, B. & Lugasi, A. Functional food. Product development, marketing and consumer acceptance—A review. *Appetite* **51**, 456–467 (2008).
37. Guerriero, G. *et al.* Production of plant secondary metabolites: Examples, tips and suggestions for biotechnologists. *Genes* **9**, 309 (2018).
38. Berni, R. *et al.* Reactive Oxygen Species and heavy metal stress in plants: impact on the cell wall and secondary metabolism. *Environ. Exp. Bot.* (2018).
39. Eskin, M. & Tamir, S. *Dictionary of nutraceuticals and functional foods*. (Crc Press, 2005).
40. Nasri, H., Baradaran, A., Shirzad, H. & Rafieian-Kopaei, M. New concepts in nutraceuticals as alternative for pharmaceuticals. *Int. J. Prev. Med.* **5**, 1487 (2014).
41. Coppens, P., da Silva, M. F. & Pettman, S. European regulations on nutraceuticals, dietary supplements and functional foods: A framework based on safety. *Toxicology* **221**, 59–74 (2006).
42. A current look at nutraceuticals – Key concepts and future prospects. Trends. Food. Sci. Tech. ***62***, 68-78 (2017).
43. Moore, B. D., Andrew, R. L., Külheim, C. & Foley, W. J. Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytol.* **201**, 733–750 (2014).
44. Kooke, R. & Keurentjes, J. J. B. Multi-dimensional regulation of metabolic networks shaping plant development and performance. *J. Exp. Bot.* **63**, 3353–3365 (2011).
45. Batista, A. N. L. *et al.* The Combined Use of Proteomics and Transcriptomics Reveals a Complex Secondary Metabolite Network in Peperomia obtusifolia. *J. Nat. Prod.* **80**, 1275–1286 (2017).
46. Prinsi, B., Negri, A. S., Espen, L. & Piagnani, M. C. Proteomic Comparison of Fruit Ripening between 'Hedelfinger' Sweet Cherry (*Prunus avium* L.) and Its Somaclonal Variant 'HS'. *J. Agric. Food Chem.* **64**, 4171–4181 (2016).

47. Chan, Z. *et al.* Functions of defense-related proteins and dehydrogenases in resistance response induced by salicylic acid in sweet cherry fruits at different maturity stages. *Proteomics* **8**, 4791–4807 (2008).
48. Sant'Ana, D. V. P. & Lefsrud, M. Tomato proteomics: Tomato as a model for crop proteomics. *Sci. Hortic.* **239**, 224–233 (2018).
49. Rödiger, A., Agne, B., Baerenfaller, K. & Baginsky, S. Arabidopsis proteomics: a simple and standardizable workflow for quantitative proteome characterization. *Methods Mol. Biol. Clifton NJ* **1072**, 275–288 (2014).
50. Douglas, C. J. Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends Plant Sci.* **1**, 171–178 (1996).
51. Pietta, P., Minoggio, M. & Bramati, L. Plant polyphenols: Structure, occurrence and bioactivity. in *Studies in Natural Products Chemistry* vol. 28 257–312 (Elsevier, 2003).
52. Manach, C., Scalbert, A., Morand, C., Rémésy, C. & Jiménez, L. Polyphenols: food sources and bioavailability. *Am. J. Clin. Nutr.* **79**, 727–747 (2004).
53. Manach, C., Scalbert, A., Morand, C., Rémésy, C. & Jiménez, L. Polyphenols: food sources and bioavailability. *Am. J. Clin. Nutr.* **79**, 727–747 (2004).
54. Tsao, R. Chemistry and biochemistry of dietary polyphenols. *Nutrients* **2**, 1231–1246 (2010).
55. Di Meo, F. *et al.* Free Radical Scavenging by Natural Polyphenols: Atom versus Electron Transfer. *J. Phys. Chem. A* **117**, 2082–2092 (2013).
56. Santos-Sánchez, N. F., Salas-Coronado, R., Hernández-Carlos, B. & Villanueva-Cañongo, C. Shikimic Acid Pathway in Biosynthesis of Phenolic Compounds. *Plant Physiol. Asp. Phenolic Compd.* (2019) doi:10.5772/intechopen.83815.
57. Reichert, A. I., He, X.-Z. & Dixon, R. A. Phenylalanine ammonia-lyase (PAL) from tobacco (Nicotiana tabacum): characterization of the four tobacco PAL genes and active heterotetrameric enzymes. *Biochem. J.* **424**, 233–242 (2009).
58. Huang, J. *et al.* Functional Analysis of the Arabidopsis PAL Gene Family in Plant Growth, Development, and Response to Environmental Stress. *Plant Physiol.* **153**, 1526–1538 (2010).
59. Cochrane, F. C., Davin, L. B. & Lewis, N. G. The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry* **65**, 1557–1564 (2004).
60. Hamberger, B. *et al.* Genome-wide analyses of phenylpropanoid-related genes in Populus trichocarpa, Arabidopsis thaliana, and *Oryza sativa*: the Populus lignin toolbox and conservation and diversification of angiosperm gene families. *Can. J. Bot.* **85**, 1182–1201 (2007).

61. Chang, A., Lim, M.-H., Lee, S.-W., Robb, E. J. & Nazar, R. N. Tomato phenylalanine ammonia-lyase gene family, highly redundant but strongly underutilized. *J. Biol. Chem.* **283**, 33591–33601 (2008).
62. Rohde, A. *et al.* Molecular Phenotyping of the pal1 and pal2 Mutants of *Arabidopsis thaliana* Reveals Far-Reaching Consequences on Phenylpropanoid, Amino Acid, and Carbohydrate Metabolism. *Plant Cell* **16**, 2749–2771 (2004).
63. Zhang, X., Gou, M. & Liu, C.-J. Arabidopsis Kelch repeat F-box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell* **25**, 4994–5010 (2013).
64. Lavhale, S. G., Kalunke, R. M. & Giri, A. P. Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants. *Planta* **248**, 1063–1078 (2018).
65. Dao, T. T. H., Linthorst, H. J. M. & Verpoorte, R. Chalcone synthase and its functions in plant resistance. *Phytochem. Rev.* **10**, 397–412 (2011).
66. Zhang, X., Abrahan, C., Colquhoun, T. A. & Liu, C.-J. A Proteolytic Regulator Controlling Chalcone Synthase Stability and Flavonoid Biosynthesis in Arabidopsis. *Plant Cell* **29**, 1157–1174 (2017).
67. Kaltenbach, M. *et al.* Evolution of chalcone isomerase from a noncatalytic ancestor. *Nat. Chem. Biol.* **14**, 548 (2018).
68. Wilmouth, R. C. *et al.* Structure and Mechanism of Anthocyanidin Synthase from *Arabidopsis thaliana*. *Structure* **10**, 93–103 (2002).
69. Rafique, M. Z. *et al.* Nonsense Mutation Inside Anthocyanidin Synthase Gene Controls Pigmentation in Yellow Raspberry (*Rubus idaeus* L.). *Front. Plant Sci.* **7**, (2016).
70. Hobert, O. Gene regulation by transcription factors and microRNAs. *Science* **319**, 1785–1786 (2008).
71. Latchman, D. S. Transcription factors: an overview. *Int. J. Biochem. Cell Biol.* **29**, 1305–1312 (1997).
72. Schaart, J. G. *et al.* Identification and characterization of MYB-b HLH-WD 40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* x *ananassa*) fruits. *New Phytol.* **197**, 454–467 (2013).
73. Baudry, A., Caboche, M. & Lepiniec, L. TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in Arabidopsis thaliana. *Plant J.* **46**, 768–779 (2006).
74. Lepiniec, L. *et al.* Genetics and biochemistry of seed flavonoids. *Annu Rev Plant Biol* **57**, 405–430 (2006).

75. Baudry, A. *et al.* TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* **39**, 366–380 (2004).
76. Payne, C. T., Zhang, F. & Lloyd, A. M. GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics* **156**, 1349–1362 (2000).
77. Xu, W., Dubos, C. & Lepiniec, L. Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* **20**, 176–185 (2015).
78. Pattanaik, S., Xie, C. H. & Yuan, L. The interaction domains of the plant Myc-like bHLH transcription factors can regulate the transactivation strength. *Planta* **227**, 707–715 (2008).
79. Feller, A., Machemer, K., Braun, E. L. & Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* **66**, 94–116 (2011).
80. Bedon, F. *et al.* Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid-and flavonoid-oriented responses. *J. Exp. Bot.* **61**, 3847–3864 (2010).
81. Srere, P. A. Complexes of Sequential Metabolic Enzymes. *Annu. Rev. Biochem.* **56**, 89–124 (1987).
82. Jørgensen, K. *et al.* Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant Biol.* **8**, 280–291 (2005).
83. Rakus, D., Pasek, M., Krotkiewski, H. & Dzugaj, A. Interaction between muscle aldolase and muscle fructose 1, 6-bisphosphatase results in the substrate channeling. *Biochemistry* **43**, 14948–14957 (2004).
84. You, C. & Zhang, Y.-H. P. Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. *ACS Synth. Biol.* **2**, 102–110 (2012).
85. Dastmalchi, M., Bernards, M. A. & Dhaubhadel, S. Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. *Plant J.* **85**, 689–706 (2016).
86. Achnine, L., Blancaflor, E. B., Rasmussen, S. & Dixon, R. A. Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* **16**, 3098–3109 (2004).
87. Gou, M., Ran, X., Martin, D. W. & Liu, C.-J. The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nat. Plants* **4**, 299–310 (2018).
88. Wang, B. & Zhao, Q. Membrane-bound metabolons. *Nat. Plants* **4**, 245–246 (2018).
89. Quero-García, J., Iezzoni, A., Pulawska, J. & Lang, G. A. *Cherries: Botany, Production and Uses*. (CABI, 2017).

90. Del Cueto, J. *et al.* Cyanogenic Glucosides and Derivatives in Almond and Sweet Cherry Flower Buds from Dormancy to Flowering. *Front. Plant Sci.* **8**, (2017).
91. Püssa, T. *Principles of Food Toxicology*. (CRC Press, 2007).
92. Berni, R. *et al.* Functional molecules in locally-adapted crops: the case study of tomatoes, onions and sweet cherry fruits from Tuscany in Italy. *Front. Plant Sci.* **9**, 1983 (2018).
93. Szakiel, A., Pączkowski, C., Pensec, F. & Bertsch, C. Fruit cuticular waxes as a source of biologically active triterpenoids. *Phytochem. Rev. Proc. Phytochem. Soc. Eur.* **11**, 263–284 (2012).
94. Buschhaus, C. & Jetter, R. Composition differences between epicuticular and intracuticular wax substructures: how do plants seal their epidermal surfaces? *J. Exp. Bot.* **62**, 841–853 (2011).
95. Shirasawa, K. *et al.* The genome sequence of sweet cherry (*Prunus avium*) for use in genomics-assisted breeding. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* **24**, 499–508 (2017).
96. Alkio, M., Jonas, U., Declercq, M., Van Nocker, S. & Knoche, M. Transcriptional dynamics of the developing sweet cherry (*Prunus avium* L.) fruit: sequencing, annotation and expression profiling of exocarp-associated genes. *Hortic. Res.* **1**, 11 (2014).
97. Maldonado, J., Dhingra, A., Carrasco, B., Meisel, L. & Silva, H. Transcriptome datasets from leaves and fruits of the sweet cherry cultivars 'Bing', 'Lapins' and 'Rainier'. *Data Brief* **23**, 103696 (2019).
98. Zong, X. *et al.* Adventitious shoot regeneration and Agrobacterium tumefaciens-mediated transformation of leaf explants of sweet cherry (*Prunus avium* L.). *J. Hortic. Sci. Biotechnol.* **94**, 229–236 (2019).
99. Carvalho, R. F., Carvalho, S. D., O'Grady, K. & Folta, K. M. Agroinfiltration of Strawberry Fruit — A Powerful Transient Expression System for Gene Validation. *Curr. Plant Biol.* **6**, 19–37 (2016).
100. Orzaez, D., Mirabel, S., Wieland, W. H. & Granell, A. Agroinjection of Tomato Fruits. A Tool for Rapid Functional Analysis of Transgenes Directly in Fruit. *Plant Physiol.* **140**, 3–11 (2006).
101. Ring, L. *et al.* Metabolic Interaction between Anthocyanin and Lignin Biosynthesis Is Associated with Peroxidase FaPRX27 in Strawberry Fruit. *Plant Physiol.* **163**, 43–60 (2013).
102. Messina, R., Lain, O., Marrazzo, M. T., Cipriani, G. & Testolin, R. New set of microsatellite loci isolated in apricot. *Mol. Ecol. Notes* **4**, 432–434 (2004).
103. Cipriani, G. *et al.* AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterisation and cross-species amplification in Prunus. *Theor. Appl. Genet.* **99**, 65–72 (1999).

104. Development and characterization of polymorphic microsatellites from *Prunus avium* 'Napoleon' - Clarke - 2003 - Molecular Ecology Notes - Wiley Online Library.

105. Dirlewanger, E. *et al.* Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theor. Appl. Genet.* **105**, 127–138 (2002).

106. Testolin, R. *et al.* Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* **43**, 512–520 (2000).

107. Henríquez, C. *et al.* Determination of antioxidant capacity, total phenolic content and mineral composition of different fruit tissue of five apple cultivars grown in Chile. *Chil. J. Agric. Res.* **70**, 523–536 (2010).

108. Hosu, A., Cristea, V.-M. & Cimpoiu, C. Analysis of total phenolic, flavonoids, anthocyanins and tannins content in Romanian red wines: Prediction of antioxidant activities and classification of wines using artificial neural networks. *Food Chem.* **150**, 113–118 (2014).

109. Oancea, S. & Draghici, O. pH and thermal stability of anthocyanin-based optimised extracts of romanian red onion cultivars. *Czech J. Food Sci.* **31**, 283–291 (2013).

110. Benzie, I. F. & Strain, J. J. The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. *Anal. Biochem.* **239**, 70–76 (1996).

111. Singleton, V. L. & Rossi, J. A. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Vitic.* **16**, 144–158 (1965).

112. Ebrahimzadeh, M. A., Pourmorad, F. & Hafezi, S. Antioxidant activities of Iranian corn silk. *Turk. J. Biol.* **32**, 43–49 (2008).

113. Tonutare, T., Moor, U. & Szajdak, L. Straw1 berry anthocyanin determination by ph differential spectroscopic method – how to get true results? *Acta Sci. Pol.***13(3)**, 35-47 (2014).

114. Tokuşoğlu, Ö., Ünal, M. K. & Yıldırım, Z. HPLC-UV and GC-MS characterization of the flavonol aglycons quercetin, kaempferol, and myricetin in tomato pastes and other tomato-based products. *Acta Chromatogr* **13**, 196–207 (2003).

115. Nyman, N. A. & Kumpulainen, J. T. Determination of anthocyanidins in berries and red wine by high-performance liquid chromatography. *J. Agric. Food Chem.* **49**, 4183–4187 (2001).

116. Kumar, N., Bhandari, P., Singh, B., Gupta, A. P. & Kaul, V. K. Reversed phase-HPLC for rapid determination of polyphenols in flowers of rose species. *J. Sep. Sci.* **31**, 262–267 (2008).

117. \Luczkiewicz, M. & Cisowski, W. The RP-HPLC analysis of anthocyanins. *Chromatographia* **48**, 360–364 (1998).

118. Porsche, F. M. *et al.* Antifungal Activity of Saponins from the Fruit Pericarp of *Sapindus mukorossi* against *Venturia inaequali*s and *Botrytis cinerea*. *Plant Dis.* **102**, 991–1000 (2017).

119. Berni, R. *et al.* Identification of the laccase-like multicopper oxidase gene family of sweet cherry (*Prunus avium* L.) and expression analysis in six ancient Tuscan varieties. *Sci. Rep.* **9**, 3557 (2019).

120. Berni, R., Cai, G., Xu, X., Hausman, J.-F. & Guerriero, G. Identification of Jasmonic Acid Biosynthetic Genes in Sweet Cherry and Expression Analysis in Four Ancient Varieties from Tuscany. *Int. J. Mol. Sci.* **20**, 3569 (2019).

121. Guerriero, G., Mangeot-Peter, L., Hausman, J.-F. & Legay, S. Extraction of High Quality RNA from Cannabis sativa Bast Fibres: A Vademecum for Molecular Biologists. *Fibers* **4**, 23 (2016).

122. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).

123. Wu, X., Xiong, E., Wang, W., Scali, M. & Cresti, M. Universal sample preparation method integrating trichloroacetic acid/acetone precipitation with phenol extraction for crop proteomic analysis. *Nat. Protoc.* **9**, 362–374 (2014).

124. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).

125. Occurrence and biological significance of proanthocyanidins in the American diet - ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0031942205001408.

126. Rauf, A. *et al.* Proanthocyanidins: A comprehensive review. *Biomed. Pharmacother.* **116**, 108999 (2019).

127. Csupor, D., Csorba, A. & Hohmann, J. Recent advances in the analysis of flavonolignans of Silybum marianum. *J. Pharm. Biomed. Anal.* **130**, 301–317 (2016).

128. Vogt, T. Phenylpropanoid Biosynthesis. *Mol. Plant* **3**, 2–20 (2010).

129. Tohge, T., Watanabe, M., Hoefgen, R. & Fernie, A. R. The evolution of phenylpropanoid metabolism in the green lineage. *Crit. Rev. Biochem. Mol. Biol.* **48**, 123–152 (2013).

130. Biała, W. & Jasiński, M. The Phenylpropanoid Case – It Is Transport That Matters. *Front. Plant Sci.* **9**, (2018).

131. Araújo, W. L., Tohge, T., Ishizaki, K., Leaver, C. J. & Fernie, A. R. Protein degradation – an alternative respiratory substrate for stressed plants. *Trends Plant Sci.* **16**, 489–498 (2011).

132. Dixon, D. P. & Edwards, R. Glutathione Transferases. *Arab. Book Am. Soc. Plant Biol.* **8**, (2010).

133. Hirai, M. Y. A robust omics-based approach for the identification of glucosinolate biosynthetic genes. *Phytochem. Rev.* **8**, 15–23 (2009).

134. Asada, K. Production and Scavenging of Reactive Oxygen Species in Chloroplasts and Their Functions. *Plant Physiol.* **141**, 391–396 (2006).

135. Nikkanen, L. & Rintamäki, E. Thioredoxin-dependent regulatory networks in chloroplasts under fluctuating light conditions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **369**, 20130224 (2014).

136. Dietz, K.-J. *et al.* The function of peroxiredoxins in plant organelle redox metabolism. *J. Exp. Bot.* **57**, 1697–1709 (2006).

137. Maskin, L. *et al.* Differential expression of the members of the Asr gene family in tomato (*Lycopersicon esculentum*). *Plant Sci.* **161**, 739–746 (2001).

138. Cho, S. M. *et al.* Comparative transcriptome analysis of field- and chamber-grown samples of *Colobanthus quitensis* (Kunth) Bartl, an Antarctic flowering plant. *Sci. Rep.* **8**, 1–14 (2018).

139. Sun, W., Van Montagu, M. & Verbruggen, N. Small heat shock proteins and stress tolerance in plants. *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* **1577**, 1–9 (2002).

140. Nishitani, K. & Tominaga, R. Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. *J. Biol. Chem.* **267**, 21058–21064 (1992).

141. Fry, S. C. *et al.* Xyloglucan endotransglycosylase, a new wall-loosening enzyme activity from plants. *Biochem. J.* **282**, 821–828 (1992).

142. Thompson, J. E. & Fry, S. C. Restructuring of wall-bound xyloglucan by transglycosylation in living plant cells. *Plant J. Cell Mol. Biol.* **26**, 23–34 (2001).

143. Zhu, X. F. *et al.* XTH31, encoding an in vitro XEH/XET-active enzyme, regulates aluminum sensitivity by modulating in vivo XET action, cell wall xyloglucan content, and aluminum binding capacity in *Arabidopsis. Plant Cell* **24**, 4731–4747 (2012).

144. Structural Evidence for the Evolution of Xyloglucanase Activity from Xyloglucan Endo-Transglycosylases: Biological Implications for Cell Wall Metabolism | Plant Cell. http://www.plantcell.org/content/19/6/1947.

145. Yang, J. L. *et al.* Cell Wall Hemicellulose Contributes Significantly to Aluminum Adsorption and Root Growth in *Arabidopsis*. *Plant Physiol.* **155**, 1885–1892 (2011).

146. Stimulation of ipt overexpression as a tool to elucidate the role of cytokinins in high temperature responses of *Arabidopsis thaliana. J. Exp. Bot*. ***67*(9)**, 2861-2873 (2016).

147. Clauw, P. *et al.* Leaf Responses to Mild Drought Stress in Natural Variants of *Arabidopsis*. *Plant Physiol.* **167**, 800–816 (2015).

148. Roepke, J. *et al.* An Apoplastic β-Glucosidase is Essential for the Degradation of Flavonol 3-O-β-Glucoside-7-O-α-Rhamnosides in Arabidopsis. *Plant Cell Physiol.* **58**, 1030–1047 (2017).

149. *Arabidopsis thaliana* β-glucosidase BGLU15 attacks flavonol 3-O-β-glucoside-7-O-α-rhamnosides. *Phytochemistry*, ***109,*** 14-24 (2015).

150. Stare, T., Stare, K., Weckwerth, W., Wienkoop, S. & Gruden, K. Comparison between Proteome and Transcriptome Response in Potato (*Solanum tuberosum* L.) Leaves Following Potato Virus Y (PVY) Infection. *Proteomes* **5**, (2017).

151. Farah, A. & de Paula Lima, J. Consumption of Chlorogenic Acids through Coffee and Health Implications. *Beverages* **5**, 11 (2019).

152. Mbaveng, A. T., Zhao, Q. & Kuete, V. 20 - Harmful and Protective Effects of Phenolic Compounds from African Medicinal Plants. in *Toxicological Survey of African Medicinal Plants*. (pp. 577-609). Elsevier.

153. Wang, J. *et al.* Protective Effect of Naringenin Against Lead-Induced Oxidative Stress in Rats. *Biol. Trace Elem. Res.* **146**, 354–359 (2012).

154. Nemes, A. *et al.* Determination of Flavonoid and Proanthocyanidin Profile of Hungarian Sour Cherry. *Mol. Basel Switz.* **23**, (2018).

155. Analysis of A-Type and B-Type Highly Polymeric Proanthocyanidins and Their Biological Activities as Nutraceuticals. *Journal of Chemistry*, (*2013)*.

156. Testai, L. & Calderone, V. Nutraceutical Value of Citrus Flavanones and Their Implications in Cardiovascular Disease. *Nutrients* **9**, 502 (2017).

157. Bijak, M., Synowiec, E., Sitarek, P., Sliwiński, T. & Saluk-Bijak, J. Evaluation of the Cytotoxicity and Genotoxicity of Flavonolignans in Different Cellular Models. *Nutrients* **9**, (2017).

158. The Effect of Standardised Flower Extracts of *Sorbus aucuparia* L. on Proinflammatory Enzymes, Multiple Oxidants, and Oxidative/Nitrative Damage of Human Plasma Components In Vitro. *Oxidative medicine and cellular longevity*, (2019).

159. Qa'dan, F., Verspohl, Eugen. J., Nahrstedt, A., Petereit, F. & Matalka, K. Z. Cinchonain Ib isolated from *Eriobotrya japonica* induces insulin secretion in vitro and in vivo. *J. Ethnopharmacol.* **124**, 224–227 (2009).

160. Sharma, A. *et al.* Response of Phenylpropanoid Pathway and the Role of Polyphenols in Plants under Abiotic Stress. *Molecules* **24**, 2452 (2019).

161. Smeets, K. *et al.* Induction of oxidative stress and antioxidative mechanisms in *Phaseolus vulgaris* after Cd application. *Plant Physiol. Biochem.* **43**, 437–444 (2005).

162. Tamás, L., Mistrík, I. & Zelinová, V. Heavy metal-induced reactive oxygen species and cell death in barley root tip. *Environ. Exp. Bot.* **140**, 34–40 (2017).

163. Gill, S. S. & Tuteja, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem. PPB* **48**, 909–930 (2010).

164. Akram, N. A., Shafiq, F. & Ashraf, M. Ascorbic Acid-A Potential Oxidant Scavenger and Its Role in Plant Development and Abiotic Stress Tolerance. *Front. Plant Sci.* **8**, (2017).

165. Carbone, F. *et al.* Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits*. *Plant Cell Environ.* **32**, 1117–1131 (2009).

166. Cochrane, F. C., Davin, L. B. & Lewis, N. G. The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry* **65**, 1557–1564 (2004).

167. Huang, J. *et al.* Functional analysis of the Arabidopsis PAL gene family in plant growth, development, and response to environmental stress. *Plant Physiol.* **153**, 1526–1538 (2010).

168. Matsuda, F. *et al.* Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. *Plant J.* **81**, 13–23 (2015).

169. Arbona, V., Manzi, M., Ollas, C. D. & Gómez-Cadenas, A. Metabolomics as a Tool to Investigate Abiotic Stress Tolerance in Plants. *Int. J. Mol. Sci.* **14**, 4885–4911 (2013).

170. Esposito, S. *et al.* Glutamate synthase activities and protein changes in relation to nitrogen nutrition in barley: the dependence on different plastidic glucose-6P dehydrogenase isoforms. *J. Exp. Bot.* **56**, 55–64 (2005).

171. Cardi, M. *et al.* The effects of salt stress cause a diversion of basal metabolism in barley roots: possible different roles for glucose-6-phosphate dehydrogenase isoforms. *Plant Physiol. Biochem. PPB* **86**, 44–54 (2015).

172. ROS and redox signalling in the response of plants to abiotic stress. *Plant, Cell & Environment*, ***35*(2)**, 259-270 (2012).

173. Golldack, D., Li, C., Mohan, H. & Probst, N. Tolerance to drought and salt stress in plants: Unraveling the signaling networks. *Front. Plant Sci.* **5**, (2014).

174. Tian, S., Qin, G. & Li, B. Reactive oxygen species involved in regulating fruit senescence and fungal pathogenicity. *Plant Mol. Biol.* **82**, 593–602 (2013).

175. Bell, L. The Biosynthesis of Glucosinolates: Insights, Inconsistencies, and Unknowns. *Annual Plant Reviews online*, 1-31 (2019).

176. Soundararajan, P. & Kim, J. S. Anti-Carcinogenic Glucosinolates in Cruciferous Vegetables and Their Antagonistic Effects on Prevention of Cancers. *Molecules* **23**, (2018).

177. Herbivore induction of the glucosinolate–myrosinase defense system: major trends, biochemical bases and ecological significance. *Phytochemistry Reviews*, ***8*(1)**, 149-170 (2009).

178. War, A. R. *et al.* Mechanisms of plant defense against insect herbivores. *Plant Signal. Behav.* **7**, 1306–1320 (2012).

179. Boeckx, T., Winters, A. L., Webb, K. J. & Kingston-Smith, A. H. Polyphenol oxidase in leaves: is there any significance to the chloroplastic localization? *J. Exp. Bot.* **66**, 3571–3579 (2015).

180. Kawarazaki, T. *et al.* A low temperature-inducible protein AtSRC2 enhances the ROS-producing activity of NADPH oxidase AtRbohF. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1833**, 2775–2780 (2013).

181. Knight, M. R., Campbell, A. K., Smith, S. M. & Trewavas, A. J. Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium. *Nature* **352**, 524–526 (1991).

182. Ye, Y. *et al.* The functional identification of glycine-rich TtASR from *Tetragonia tetragonoides* (Pall.) Kuntze involving in plant abiotic stress tolerance. *Plant Physiol. Biochem.* **143**, 212–223 (2019).

183. Multifunctional role of plant cysteine proteinases. *ACTA BIOCHIMICA POLONICA-ENGLISH EDITION*, 609-624 (2004).

184. Post-transcriptional regulation of fruit ripening and disease resistance in tomato by the vacuolar protease SlVPE3. *Genome biology*, ***18*(1)**, 47 (2017).

185. Rawlings, N. D. *et al.* The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632 (2018).

186. Yamada, K., Matsushima, R., Nishimura, M. & Hara-Nishimura, I. A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing Arabidopsis leaves. *Plant Physiol.* **127**, 1626–1634 (2001).

187. Shindo, T., Misas-Villamil, J. C., Hörger, A. C., Song, J. & van der Hoorn, R. A. L. A Role in Immunity for Arabidopsis Cysteine Protease RD21, the Ortholog of the Tomato Immune Protease C14. *PLoS ONE* **7**, (2012).

188. Kaewthai, N. *et al.* Group III-A XTH Genes of Arabidopsis Encode Predominant Xyloglucan Endohydrolases That Are Dispensable for Normal Growth. *Plant Physiol.* **161**, 440–454 (2013).

189. Sampedro, J., Sieiro, C., Revilla, G., González-Villa, T. & Zarra, I. Cloning and Expression Pattern of a Gene Encoding an α-Xylosidase Active against Xyloglucan Oligosaccharides from Arabidopsis. *Plant Physiol.* **126**, 910–920 (2001).

190. Günl, M. & Pauly, M. AXY3 encodes a α-xylosidase that impacts the structure and accessibility of the hemicellulose xyloglucan in *Arabidopsis* plant cell walls. *Planta* **233**, 707–719 (2011).

191. Meng, M. *et al.* UDP-glucose pyrophosphorylase is not rate limiting, but is essential in Arabidopsis. *Plant Cell Physiol.* **50**, 998–1011 (2009).

192. Commisso, M. *et al.* Multi-approach metabolomics analysis and artificial simplified phytocomplexes reveal cultivar-dependent synergy between polyphenols and ascorbic acid in fruits of the sweet cherry (*Prunus avium* L.). *PLoS ONE* **12**, (2017).

193. Liu, Y. *et al.* Expression Analysis of Anthocyanin Biosynthetic Genes in Different Colored Sweet Cherries (*Prunus avium* L.) During Fruit Development. *J. Plant Growth Regul.* **32**, 901–907 (2013).

194. Wei, H. *et al.* Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in the red and yellow fruits of sweet cherry (*Prunus avium* L.). *PloS One* **10**, (2015).

195. Dheilly, E. *et al.* Cell wall dynamics during apple development and storage involves hemicellulose modifications and related expressed genes. *BMC Plant Biol.* **16**, (2016).

196. Posé, S. *et al.* A nanostructural view of the cell wall disassembly process during fruit ripening and postharvest storage by atomic force microscopy. *Trends Food Sci. Technol.* **87**, 47–58 (2019).

197. Ng, J. K. T. *et al.* Solid-state 13C NMR study of the mobility of polysaccharides in the cell walls of two apple cultivars of different firmness. *Carbohydr. Res.* **386**, 1–6 (2014).